



Wydział Nauk o Żywności i Żywieniu

Katedra Biotechnologii i Mikrobiologii Żywności

PAULINA KORPYS-WOŹNIAK

Intensyfikacja produkcji heterologicznych białek w komórkach niekonwencjonalnych drożdży *Yarrowia lipolytica* poprzez zastosowanie inżynierii molekularnych mechanizmów dojrzewania i sekrecji polipeptydów

Intensification of heterologous protein production in nonconventional yeast *Yarrowia lipolytica* through the engineering of molecular mechanisms of polypeptide maturation and secretion

PRACA DOKTORSKA

Dziedzina nauk ścisłych i przyrodniczych

Dyscyplina: Nauki Biologiczne

Promotor: dr hab. Ewelina Celińska, Prof. UPP

Poznań 2023

Pragnę szczególnie podziękować

Pani Promotor Prof. UPP dr hab. Ewelinie Celińskiej

za możliwość realizacji nieniejszej pracy, poświęcony czas, wsparcie merytoryczne, przekazaną wiedzę, miłą atmosferę, okazaną życzliwość i wyrozumiałość,

Panu Prof. UPP dr hab. Wojciechowi Białasowi

za pomoc w przeprowadzeniu analiz statystycznych,

Panu dr hab inż. Piotrowi Kubiakowi

za fachowe wskazówki dotyczące prowadzenia ciągłych hodowli bioreaktorowych, okazaną pomoc w przeprowadzeniu analiz statystycznych, a także miłą atmosferę,

Pani dr inż. Monice Kubiak-Szymenderze

za pomoc w prowadzeniu eksperymentów

**Zespołowi Chromatograficznemu Katedry Biotechnologii i Mikrobiologii
Żywności**

za przeprowadzenie analiz HPLC

Firmie Genomed

za wykonanie sekwencjonowania RNA

Rodzinie

za wsparcie i wyrozumiałość

Pracę dedykuję Mężowi i Synowi Kajetanowi

Wykaz artykułów wchodzących w skład rozprawy doktorskiej

P.1. Korpys-Woźniak P., Kubiak P., Białas W., Celińska E. (2020) Impact of overproduced heterologous proteins characteristics on physiological response in *Yarrowia lipolytica* steady-state maintained continuous cultures. *Applied Microbiology and Biotechnology*. Springer. 104:9785-9800. <https://doi.org/10.1007/s00253-020-10937-w>.

2-letni Impact Factor: 5.56, 5-letni Impact Factor: 5.365, **100** punktów MEiN (2020)

P.2. Korpys-Woźniak P., Celińska E. (2021) Global transcriptome profiling reveals genes responding to overproduction of a small secretory, a high cysteine- and a high glycosylation-bearing protein in *Yarrowia lipolytica*. *Biotechnology Reports*. Elsevier. 31: e00646. <https://doi.org/10.1016/j.btre.2021.e00646>.

2-letni Impact Factor: 4.982, 5-letni Impact Factor: 5.557, **100** punktów MEiN (2021)

P.3. Korpys-Woźniak P., Kubiak P., Celińska E. (2021) Secretory helpers for enhanced production of heterologous proteins in *Yarrowia lipolytica*. *Biotechnology Reports*. Elsevier. 32: e00669. <https://doi.org/10.1016/j.btre.2021.e00669>.

2-letni Impact Factor: 4.982, 5-letni Impact Factor: 5.557, **100** punktów MEiN (2021)

P.4. Korpys-Woźniak P., Celińska E. (2023). Molecular background of HAC1-driven improvement in the secretion of recombinant protein in *Yarrowia lipolytica* based on comparative transcriptomics. *Biotechnology Reports*. Elsevier. 38: e00801. <https://doi.org/10.1016/j.btre.2023.e00801>.

2-letni Impact Factor: 4.982, 5-letni Impact Factor: 5.557, **100** punktów MEiN (2023)

Źródło finansowania badań

Projekt *Diamantowy Grant* VII nr DI 2017 000947 finansowany przez Ministerstwo Nauki i Szkolnictwa Wyższego (Obecnie Ministerstwo Edukacji i Nauki), Kierownik projektu.

Spis treści

Wykaz skrótów stosowanych w pracy	9
Streszczenie	12
Streszczenie w języku angielskim	14
1. Wstęp	16
1.1. Produkcja rekombinowanych białek w systemach drożdżowych	16
1.2. Charakterystyka molekularnych podstaw syntezy i sekrecji polipeptydów w komórkach drożdżowych.....	17
1.3. Strategie eliminacji „wąskich gardeł” szlaku translacyjno-sekrecyjnego	21
1.4. Hodowle ciągłe w stanie ustalonym – solidne narzędzie w analizach omicznych....	25
2. Hipotezy i cele badań.....	27
3. Materiały i metody	28
3.1. Podłoża hodowlane.....	28
3.2. Konstrukcja rekombinowanych szczepów	29
3.2.1. Przygotowanie konstrukcji do nadekspresji genów kodujących r(s)-Prots - pochodne szczepów Po1g oraz Po1h (P.1, P.2)	29
3.2.2. Przygotowanie konstrukcji do ko-nadekspresji genów SCYFP i SHs –.....	33
3.2.3. pochodne szczepu Po1f (P.3, P.4)	33
3.2.4. Transformacja bakterii <i>E. coli</i> (JM109)	35
3.2.5. Transformacja drożdży <i>Y. lipolytica</i> - pochodne Po1g, Po1h, i Po1f	36
3.3. Metody hodowlane	37
3.3.1. Hodowle okresowe – screening szczepów <i>Y. lipolytica</i> (P.1, P.3).....	37
3.3.2. Hodowle okresowe – ocena wpływu obniżonej temperatury na działanie poszczególnych SHs (P.3).....	37
3.3.3. Hodowle ciągłe w warunkach chemostatu – globalna analiza transkryptomu (P.1, P.2, P.4)	38
3.4. Metody analityczne.....	38
3.4.1. Pomiar gęstości biomasy drożdżowej	38
3.4.2. Pomiar żywotności komórek (P.1)	39
3.4.3. Morfologia komórek (P.2).....	39
3.4.4. Pomiar stężenia substratu i metabolitów – wysokosprawna chromatografia cieczowa (HPLC).....	39
3.4.5. Analiza aktywności enzymatycznej – microSIT i microDNS (P.1, P.2)	40

3.4.6.	Pomiar intensywności fluorescencji.....	40
3.4.7.	Izolacja białek i ocena aktywności proteolitycznej (P.4).....	41
3.4.7.1.	Ekstrakcja białka całkowitego i oznaczenie stężenia metodą BCA.....	41
3.4.7.2.	Oznaczenie aktywności proteolitycznej.....	42
3.5.	Oznaczenie poziomu ekspresji genów.....	42
3.5.1.	Izolacja RNA.....	42
3.5.2.	Odwrotna transkrypcja oraz PCR w czasie rzeczywistym (qPCR; <i>quantitative Polymerase Chain Reaction</i>) (P.1, P.3)	42
3.5.3.	Sekwencjonowanie RNA - RNAseq (P.2, P.4)	43
3.6.	Globalna analiza transkryptomu (P.2, P.4).....	43
3.6.1.	Przetwarzanie i filtrowanie surowych danych pochodzących z RNAseq	43
3.6.2.	Analiza porównawcza - wykresy Venn'a, heatmapy, test nadprezentacji	44
3.6.3.	Identyfikacja miejsc splicingowych w regionie genu <i>HAC1 (B12716g)</i> (P.4) ..	45
3.6.4.	Narzędzia bioinformatyczne i bazy danych	45
3.7.	Obliczenia.....	46
3.7.1.	Stężenie białek reporterowych (P.1)	46
3.7.2.	Uproszczony bilans masowy dystrybucji C i N między produktami metabolizmu <i>Y. lipolytica</i> (P.1)	46
4.	Wyniki.....	47
4.1.	Charakterystyka szczepów <i>Y. lipolytica</i> z nadekspresją r(s)-Prots w warunkach chemostatu	47
4.1.1.	Klonowanie r(s)-Prots i selekcja szczepów <i>Y. lipolytica</i> (P.1).	47
4.1.2.	Wpływ nadprodukcji r(s)-Prots na przebieg ciągłych hodowli drożdży <i>Y. lipolytica</i> (P.1).....	47
4.1.3.	Analiza ekspresji heterologicznych genów podczas wzrostu drożdży <i>Y. lipolytica</i> w warunkach chemostatu (P.1).....	50
4.2.	Globalna analiza transkryptomu rekombinowanych szczepów <i>Y. lipolytica</i> (pochodne Po1h) – sekwencjonowanie RNA.....	51
4.2.1.	Podstawowa charakterystyka szczepów – obserwacje ogólne.....	51
4.2.2.	Zmiany w transkryptomie rekombinowanych szczepów <i>Y. lipolytica</i> w warunkach nadprodukcji r(s)-Prots – analiza nieukierunkowana.....	52
4.2.3.	Podobieństwa i różnice w transkryptomie indukowane nadprodukcją r(s)-Prots	59
4.2.4.	Ukierunkowana analiza transkryptomu pod kątem specyficznej odpowiedzi szlaku wydzielniczego na nadprodukcję r(s)-Prots.....	61

4.3. Hodowle okresowe rekombinowanych szczepów <i>Y. lipolytica</i> z ko-nadekspresją rs-YFP i SH (P.3).....	68
4.3.1. Klonowanie homologicznych SHs i selekcja szczepów	68
4.3.2. Kinetyka wzrostu, syntezy i sekrecji rs-YFP	68
4.3.3. Szczegółowa analiza wpływu nadekspresji SHs na zdolność wydzielniczą komórek <i>Y. lipolytica</i>	72
4.3.4. Wpływ temperatury na działanie SHs	74
4.3.5. Analiza poziomu ekspresji heterologicznego genu <i>SCYFP</i> w warunkach nadekspresji SHs	75
4.4. Profil transkrypcyjny rekombinowanego szczepu <i>Y. lipolytica</i> z ko-nadekspresją <i>HAC1</i> i <i>SCYFP</i> – transkryptomyczna analiza porównawcza (P.4)	78
4.4.1. Synteza i sekrecja YFP w szczepach <i>Y. lipolytica</i> utrzymywanych w stanie ustalonym (P.2: <i>inYFP</i> , <i>scYFP</i> , P.4: <i>HAC1_scYFP</i>)	78
4.4.2. Silnie deregulowane DEGs w warunkach nadekspresji <i>HAC1</i> i <i>rs-YFP</i>	79
4.4.3. Procesy biologiczne o szczególnym znaczeniu dla produkcji rs-Protos	84
4.4.3.1. Regulacja ekspresji białek	84
4.4.3.2. Degradacja białek	85
4.4.3.3. Transport pęcherzykowy	85
4.4.4. Cykl komórkowy	86
5. Wykaz analiz.....	88
6. Dyskusja.....	90
7. Podsumowanie i kluczowe wnioski	106
8. Literatura.....	110
Wykaz Tabel	132
Wykaz Rycin.....	132

Wykaz skrótów stosowanych w pracy

(sp_)exYFP – (specyficzna) zewnątrzkomórkowa fluorescencja białka YFP (*(specific) extracellular fluorescence of YFP*) wyrażona w [FU OD600nm⁻¹]

(sp_)inYFP – (specyficzna) wewnątrzkomórkowa fluorescencja białka YFP (*(specific) intracellular fluorescence of YFP*) wyrażona w [FU OD600nm⁻¹]

AA - aminokwasy

AAU – jednostka aktywności α -amylazy (*α -amylase Activity Unit*)

AG – aparat Golgiego (*Golgi Apparatus*)

AU – jednostka aktywności enzymatycznej

CA – kwas cytrynowy (*citric acid*)

Cys – cysteina (*Cysteine*)

D - szybkość rozcieńczania [h⁻¹] (*Dilution Rate*)

DCW – sucha masa komórkowa (*Dry Cellular Weight*)

DEG(s) – gen (geny) różnicowo ekspresjonowane (*Differentially Expressed Gene(s)*)

ER - reticulum endoplazmatyczne (*Endoplasmic Reticulum*)

ERAD – zjawisko degradacji nieprawidłowo sfałdowanych białek związane z ER (*ER-associated degradation*)

ERY – erytrytol (*erythritol*)

FC - krotności zmiany ilości transkryptu próby badanej względem próby kontrolnej (*Fold Change*)

FL – fluorescencja (*Fluorescence*)

FU – jednostka fluorescencji (*Fluorescence Unit*)

GAU – jednostka aktywności glukoamylazy (*Glucoamylase Activity Unit*)

GLY – glicerol (*glycerol*)

HSS – szczepy wydajnie wydzielające białka (*High Secretory Strains*)

in/exYFP – bezwzględna wartość fluorescencji wewnątrz-/zewnątrzkomórkowej

in/scYFP – wewnątrzkomórkowe/sekrecyjne białko żółtej fluorescencji

log2_Estimated_FoldChange - logarytm dwójkowy z krotności zmiany ilości transkryptu próby badanej względem próby kontrolnej

MAN – mannitol (*mannitol*)

NAC – kompleks wiążący powstający polipeptyd (*Nascent-polypeptide-Associated Complex*)

OD – gęstość optyczna (*Optical Density*)

PM – błona komórkowa (plazmatyczna) (*Plasma membrane*)

qPCR – reakcja ilościowego PCR w czasie rzeczywistym (*quantitative PCR*)

r(s)-Prot(s) – rekombinowane (sekrecyjne) białko (białka) (*recombinant secretory Protein(s)*)

RE – enzym restrykcyjny (*Restriction Enzyme*)

RFU – jednostka względnej fluorescencji (*Relative Fluorescence Unit*)

RQ – względna analiza ilościowa (*Relative Quantification*)

sc – sekrecyjny (*secretory*)

SH(s) – białko (białka) wspomagające sekrecję (*Secretory Helper(s)*)

SH(s) – gen (geny) kodujący białko wspomagające sekrecję

SNARE - (*Soluble N-ethylmaleimide-sensitive factor attachment protein receptor*)

SoA – α -amylaza pochodząca z wołka ryżowego *Sitophilus oryzae*

SoA/TIG/inYFP/scYFP – białko heterologiczne

SOA/TLG/INYFP/SCYFP – gen kodujący określony rodzaj białka heterologicznego

SoA/TIG/inYFP/scYFP – szczep produkujący określony rodzaj białka heterologicznego

SP – peptyd sygnałny (*Signal Peptide*)

TF – czynnik transkrypcyjny (*Transcription Factor*)

TIG – glukoamylaza pochodząca z grzyba *Thermomyces lanuginosus*

TU – Jednostka transkrypcyjna (*Transcription Unit*)

UPR – reakcja na nieprawidłowo sfałdowane białka (*Unfolded Protein Response*)

VPS – wakuolarnie białka sortujące (*Vacuolar protein sorting*)

YFP – białko żółtej fluorescencji (*Yellow Fluorescence Protein*)

α -KG – kwas α – ketoglutarynowy (*α -ketoglutaric acid*)

W pracy zastosowano skróconą wersję numerów genów bez przedrostka Yali0_

Streszczenie

Produkcja rekombinowanych białek (r-Protów) to jeden z kluczowych obszarów współczesnej biotechnologii. Wśród szeregu mikrobiologicznych producentów r-Protów, gatunek drożdży *Yarrowia lipolytica* zyskuje coraz większą popularność przez wzgląd na wiele cech korzystnych w tym zakresie. Od lat prowadzone są prace z zakresu inżynierii genetycznej nad ulepszaniem szczepów *Y. lipolytica* w celu intensyfikacji syntezy i sekrecji r-Protów. Zdobywanie nowej, szczegółowej wiedzy na temat funkcjonowania szlaku translacyjno-sekrecyjnego pozwoli na udoskonalenie opracowywanych strategii modyfikacji. Szlak translacyjno-sekrecyjny to wielofunkcyjna, zintegrowana, wielo-organelkowa sieć powiązań strukturalno-funkcjonalnych. Stąd, w obliczu wysokiego poziomu nadprodukcji r-Protów, dysfunkcja występująca w jednym z punktów, wpływa na funkcjonowanie szerokiego spektrum zdarzeń molekularnych w komórce. Nadprodukcja sekrecyjnych r-Protów (rs-Protów) bardzo często prowadzi do przeciążenia mechanizmów komórkowych i indukuje silny stres, co w konsekwencji prowadzi do utraty docelowego białka. Globalne poznanie elementów, jak również zrozumienie procesów biologicznych zaangażowanych w fałdowanie, dojrzewanie i sekrecję polipeptydów, i ich powiązań, jest możliwe dzięki rozwojowi analiz omicznych. Biorąc pod uwagę, że światowy rynek r(s)-Protów stanowi jedną z kluczowych gałęzi współczesnego przemysłu biotechnologicznego, badania nad biologią szlaku translacyjno-sekrecyjnego są istotne z punktu widzenia nauk podstawowych, jak i aplikacyjnych.

W niniejszej rozprawie doktorskiej podjęto próbę wyjaśnienia mechanizmów działania szlaku translacyjno-sekrecyjnego w komórkach *Y. lipolytica*. Celem zachowania inercji badanego układu względem czynników zewnętrznych, badane szczepy nadprodukujące r(s)-Protów utrzymywano w stanie ustalonym. Analizie poddano szczepy nadprodukujące r(s)-Protów o odmiennych właściwościach biochemicznych, aby wskazać procesy specyficzne dla danego procesu biologicznego (jak glikozylacja, fałdowanie oksydacyjne itd.). Tak przygotowany układ eksperymentalny następnie poddano globalnej analizie transkryptomicznej. Szczegółowa analiza danych omicznych, dostarczyła wglądu w procesy biologiczne mające miejsce w komórce, w tym m.in. inicjacja stresu oksydacyjnego i niesfałdowanego białka (UPR; *Unfolded Protein Response*), jak również identyfikację elementów zaangażowanych w proces glikozylacji, fałdowania oraz translokacji r(s)-Protów. Szczególnie interesujących obserwacji dokonano w obszarze funkcjonowania niekonwencjonalnego mechanizmu wydzielania białek, zaangażowania regulatorów transkrypcji, inicjacji lub inhibicji

wakuolarnej proteolizy, czy też zatrzymania progresji cyklu komórkowego. Na podstawie danych omicznych, wskazano geny, które następnie wykorzystano jako czynniki wspomagające (*helper factors*) syntezę i sekrecję rs-Prots w warunkach ich ko-nadekspresji z białkiem reporterowym. Wyniki tego eksperymentu wyraźnie wskazały, że ko-nadekspresja kluczowych genów zaangażowanych w syntezę rs-Prots znacząco zwiększyła efektywność syntezy docelowego białka przez *Y. lipolytica* w zależności od przyjętej temperatury hodowli. Ponadto, dokonano szczegółowej analizy mechanizmów molekularnych indukowanych w warunkach ko-nadekspresji kluczowego czynnika transkrypcyjnego zaangażowanego w inicjację odpowiedzi UPR – Hac1. Wskazano, że ko-ekspresja natywnego *HAC1* usprawnia proces wydzielania rs-Prots i znacząco wpływa na deregulację genów zaangażowanych w biogenezę rybosomów, procesy zachodzące w jądrze komórkowym i mitochondriach, zatrzymanie cyklu komórkowego, a także modulację proteolizy i metabolizmu RNA.

Uzyskane w niniejszej pracy wyniki przyczyniły się do lepszego poznania biologii syntezy i sekrecji r(s)-Prots w komórkach *Y. lipolytica*. Zdobyta wiedza umożliwiła opracowanie efektywnej strategii modyfikacji genetycznej tychże drożdży celem intensyfikacji produkcji rs-Prots w komórkach tego gospodarza.

Streszczenie w języku angielskim

The production of recombinant proteins (r-Protos) is one of the key areas of modern biotechnology. Among a number of microbial producers of r-Protos, the *Yarrowia lipolytica* yeast species is gaining high interest due to many relevant and beneficial features. For years, work has been carried out in the field of genetic engineering on the improvement of *Y. lipolytica* strains in order to intensify the synthesis and secretion of r-Prot. Acquiring new, detailed knowledge on the functioning of the translational-secretory pathway will allow for the refinement of the modification strategies being developed. The translational-secretory pathway is a multifunctional, integrated, multi-organelle network of structural and functional interaction. Hence, under high levels of r-Protos overproduction, dysfunction at one point affects the functioning of a wide spectrum of molecular events in the cell. Overproduction of secretory-Protos (rs-Protos) frequently leads to an overload of cellular mechanisms and induces severe stress, which in turn leads to the loss of the target protein. Global knowledge of the elements, as well as an understanding of the biological processes involved in the folding, maturation and secretion of polypeptides and their interrelationships, is possible thanks to the development of omics analysis. Considering that the global r(s)-Protos market is one of the key branches of the modern biotechnology industry, research on the biology of the translational-secretory pathway is important from the point of view of basic and applied sciences.

This work aimed at describing the elements and mechanisms of the translational-secretory pathway operation in *Y. lipolytica*. To endow the experimental set-up with an inert character, the tested strains overproducing r(s)-Protos were maintained in a steady state. Strains overproducing r(s)-Protos with different biochemical properties were analyzed to identify elements and mechanisms specific to a given biological process (such as glycosylation, oxidative folding, etc.). The experimental set-up prepared in this way was then subjected to global transcriptomic analysis. Detailed analysis of the omics data provided insight into the biological processes taking place in the cell, including initiation of oxidative stress and unfolded protein response (UPR), as well as identification of elements involved in the process of glycosylation, folding and translocation of r(s)-Protos. Particularly interesting observations were made on the functioning of the unconventional protein secretion mechanism, the involvement of transcription regulators, the initiation or inhibition of vacuolar proteolysis, or the arrest of cell cycle progression. Based on the omics data, some of the identified genes were used as helper factors for the synthesis and secretion of rs-Protos.

The results of this experiment clearly showed that co-overexpression of the helper genes significantly increased the efficiency of target protein synthesis by *Y. lipolytica*, which was dependent on the adopted culture temperature. In addition, a detailed analysis of the molecular mechanisms induced under co-overexpression of the key transcription factor involved in the initiation of the UPR response - Hac1, was performed. It was shown that the co-expression of native *HAC1* improves the secretion of rs-ProtS and significantly affects the deregulation of genes involved in ribosome biogenesis, processes occurring in the cell nucleus and mitochondria, cell cycle arrest, as well as modulation of proteolysis and RNA metabolism.

The results obtained in this work contributed to a better understanding of the biology of r(s)-ProtS synthesis and secretion in *Y. lipolytica* cells. The acquired knowledge enabled the development of an effective strategy of genetic modification of these yeasts in order to intensify the production of rs-ProtS in the cells of this host.

1. Wstęp

1.1. Produkcja rekombinowanych białek w systemach drożdżowych

Produkcję rekombinowanych białek (r-Protów) w systemach mikrobiologicznych zainicjowano na początku lat 80 wraz z rozwojem narzędzi inżynierii genetycznej. Takie podejście stanowiło atrakcyjną alternatywę dla ekstrakcji białek ze źródeł naturalnych. Obecnie światowy rynek sekrecyjnych białek rekombinowanych (rs-Protów) stanowi jedną z kluczowych gałęzi przemysłu, przynoszącą wielomiliardowe dochody (USD) każdego roku [1]. Wśród znanych mikrobiologicznych systemów ekspresyjnych powszechnie wykorzystywanych do produkcji r(s)-Protów do celów medycznych, diagnostycznych i przemysłowych wymienia się: dobrze scharakteryzowane drożdże *Saccharomyces cerevisiae*, metylotroficzne drożdże: *Pichia pastoris* (*Komagataella phaffii*) i *Hansenula polymorpha*, drożdże rozszczepkowe *Schizosaccharomyces pombe*, a także drożdże niekonwencjonalne jak: *Kluyveromyces lactis*, i *Yarrowia lipolytica* [1,2].

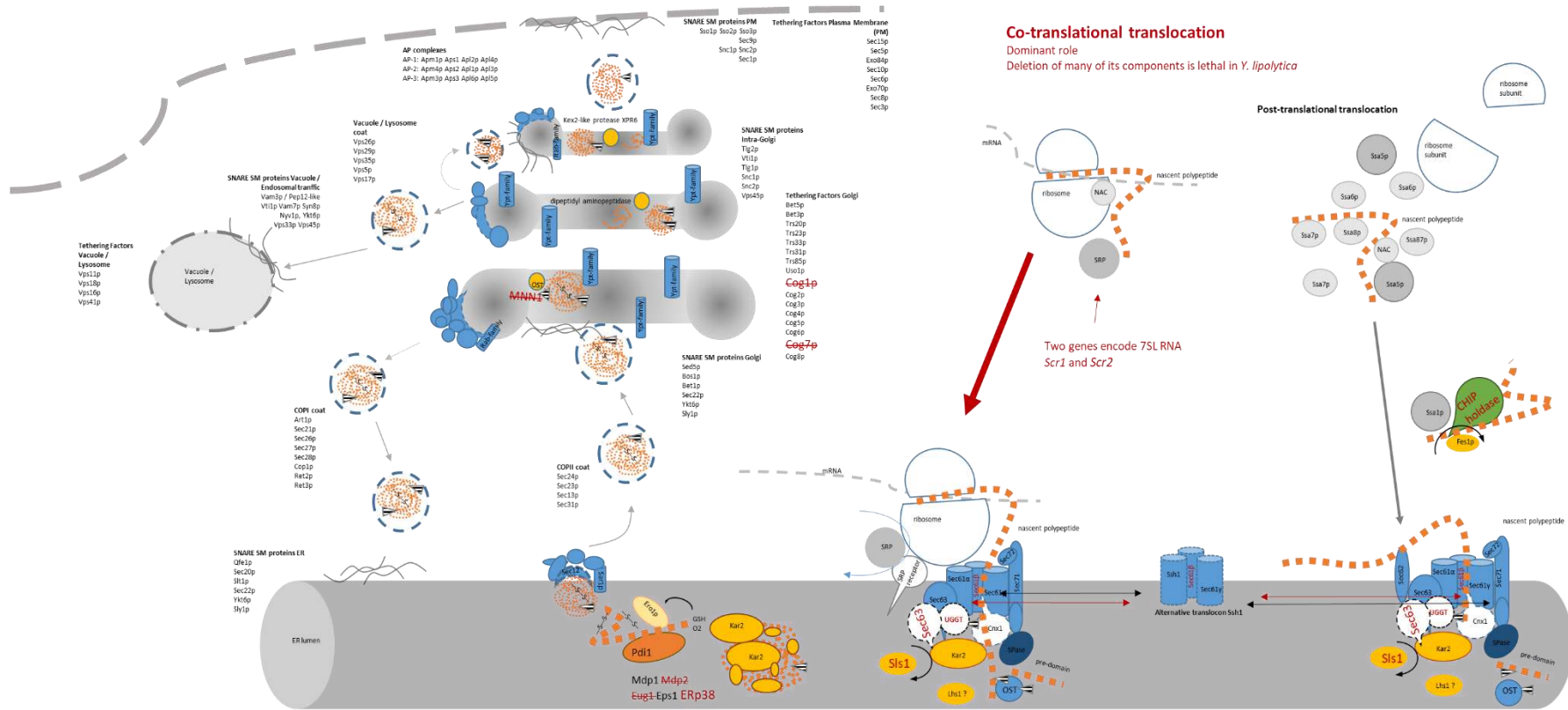
Gatunek drożdży niekonwencjonalnych *Y. lipolytica* obecnie stanowi uznaną platformę do ekspresji heterologicznych genów w celu produkcji r-Protów. Do tej pory w systemie ekspresyjnym *Y. lipolytica* nadprodukowano ponad 150 r-Protów z co najmniej 80 gatunków [3,4]. Korzystne właściwości fizjologiczne, łatwość manipulacji genetycznych, wysoka wydajność produkcji, bezpieczeństwo stosowania (status GRAS), a także wyróżniające cechy szlaku wydzielniczego (np. kotranslacyjny system translokacji) [5–7], predysponują ten gatunek do produkcji rs-Protów. Jak wykazano, w trakcie hodowli bioreaktorowych, drożdże *Y. lipolytica* przyrastają do wysokich gęstości biomasy, efektywnie wydzielają rs-Protów, co przekłada się na wysokie produktywności rs-Protów odnotowywane dla tych bioprocessów [8–10]. Inne badania wykazały, że drożdże *Y. lipolytica* charakteryzuje także wysoki potencjał do syntezy i sekrecji rs-Protów o masie molekularnej przekraczającej 120 [kDa] [11]. Na podstawie przeprowadzonych analiz porównawczych, wykazano, że wydzielanie aktywnej formy lipazy I z *Thermomyces lanuginosus* było 10-krotnie wyższe, a celulazy II z *Humicola insolens* nawet 44-krotnie wyższe w komórkach *Y. lipolytica* w porównaniu z *S. cerevisiae* [12]. Niedawne badania porównawcze jednoznacznie wskazały na przewagę *Y. lipolytica* w odniesieniu do syntezy i sekrecji heterologicznej lipazy CalB w porównaniu do klasycznego systemu bazującego na indukowanej metanolem ekspresji w komórkach *P. pastoris* [8]. Jednak, pomimo natywnych predyspozycji

Y. lipolytica do wydajnej produkcji r(s)-Prots, jak również szerokiej dostępności narzędzi inżynierii genetycznej [13,14], a także pomyślnie zastosowanych strategii zwiększających produkcję r-Prots [3,15–17], nadal pozostaje przestrzeń do ulepszania tych właściwości. Szczególnie, wiedza dotycząca działania maszynerii translacyjno-sekrecyjnej w trakcie produkcji rs-Prots, a także kierunki intensyfikacji ich syntezy i sekrecji, wciąż wymagają uzupełnienia.

1.2. Charakterystyka molekularnych podstaw syntezy i sekrecji polipeptydów w komórkach drożdżowych

Dogłębne poznanie mechanizmów działania szlaku wydzielniczego u organizmów służących jako platformy ekspresyjne jest istotne zarówno dla badań podstawowych, jak i aplikacyjnych. Wiedza na temat mechanizmów związanych z fałdowaniem, dojrzewaniem i sekrecją polipeptydów w komórkach *Y. lipolytica* było możliwe dzięki badaniom prowadzonym z wykorzystaniem klasycznych podejść z zakresu biologii molekularnej [6,18–24], jak również dzięki dynamicznemu rozwojowi analiz omicznych, takich jak: genomika, transkryptomika czy proteomika. Techniki te umożliwiły także identyfikację specyficznych genów przyczyniających się do zwiększonej/zmniejszonej zdolności wydzielniczej komórek [25].

W trakcie transportu polipeptydu poprzez szlak sekrecyjny, szereg mechanizmów molekularnych zostaje uruchomionych w sposób skoordynowany, aby umożliwić wydzielenie białka i jego pełne dojrzewanie. Procesy zachodzące w szlaku sekrecyjny obejmują wiele złożonych etapów, jak translokacja, fałdowanie, modyfikacja potranslacyjna oraz transport pęcherzykowy między organellami i błoną plazmatyczną (PM; *Plasma membrane*), w których zaangażowanych jest wiele genów/białek, a także liczne organelle komórkowe (**Rycina 1**).



Rycina 1. Szlak wydzielniczy w *Yarrowia lipolytica*. Szczegółowe informacje o roli poszczególnych elementów aparatu wydzielniczego przedstawionych na rycinie znajdują się w tekście. Wszystkie białka przedstawione na rysunku zostały zidentyfikowane w *Y. lipolytica* za pomocą podejścia proteomicznego lub genomiki porównawczej [26–28]. Kolorem ciemnoczerwonym zaznaczono białka odróżniające *Y. lipolytica* od drożdży modelowych (przekreślenia – brak genu kodującego daną aktywność; duża czcionka – unikalne występowanie genu kodującego daną aktywność w *Y. lipolytica* lub jego dominująca rola, np. Sls1 NEF). Nieznana funkcja Lhs1 została zaznaczona cudzysłowem. Dominującą rolę szlaku translokacji kotranslacyjnej wskazuje grubsza czerwona strzałka w porównaniu z cienką szarą translokacją oznaczającą translokację potranslacyjną. Powstający polipeptyd jest zaznaczony na pomarańczowo. Trójkąty w linii – N-glikany. Białka opiekuńcze, czynniki wiążące, SNARE, białka SM zostały wymienione jako zidentyfikowane przez analizę proteomiczną w [27]. Rycina z Celińska i Nicaud [5].

Typowo, białka sekrecyjne (w tym polipeptydy związane z błoną komórkową i organellami) rozpoczynają swoją „podróż” w przestrzeni wewnątrzkomórkowej szlaku wydzielniczego poprzez translację ko- lub potranslacyjną przez system translokonu Sec61 do światła retikulum endoplazmatycznego (ER; *Endoplasmic Reticulum*), gdzie ulegają fałdowaniu przy udziale licznych białek rezydujących w ER [29,30]. Kierowanie do ER odbywa się na podstawie rozpoznania tzw. peptydu sygnałowego (SP; *Signal Peptide*), najczęściej (choć nie zawsze) ulokowanego na N-końcu polipeptydu [31,32]. W momencie ekspozycji SP z rybosomu podczas translacji, jest on rozpoznawany przez kompleks NAC (*Nascent-polypeptide-Associated Complex*). Rolą NAC jest zapewnienie odpowiedniego środowiska chroniącego krótkie, aminowe końce powstającego łańcucha polipeptydowego. Mechanizm translacji potranslacyjnej (SRP-niezależnej; *Signal Recognition Particle*) polega na wiązaniu powstającego polipeptydu, uwolnionego z dysocjującego rybosomu, przez białka opiekuńcze/chaperonowe (Ssas) obecne w cytozolu (u *Y. lipolytica*: Ssa5 - Ssa8) [33]. Główną rolą tych białek jest ochrona polipeptydu w niesfałdowanej konformacji, udział w fałdowaniu, transporcie oraz degradacji nieprawidłowo sfałdowanych białek [33]. W trakcie translacji kotranslacyjnej (SRP-zależnej) nowo syntetyzowane białka są rozpoznawane przez NAC, a następnie cząsteczkę SRP, która działa jako adapter pomiędzy maszyną translacyjną w cytoplazmie a maszyną translokacyjną w błonie ER. SRP utrzymuje stan przejściowego blokowania procesu translacji, aby spowolnić elongację powstającego polipeptydu i kieruje kompleks „rybosom-polipeptyd-SRP” do receptora SRP w błonie ER, gdzie rybosom oddziałuje z elementami kanału translokacji – kompleksu Sec61/Ssh1 [5]. Preferencję dla poszczególnych mechanizmów translacji u drożdży *S. cerevisiae* przypisuje się hydrofobowości N-końcowego SP, która wpływa na wiązanie z SRP [34]. W przypadku *Y. lipolytica* skład aminokwasowy (AA; *Amino Acid*) oraz konformacja SP wydaje się być ważniejszą, choć nie jedyną, determinantą interakcji z SRP [31,35]. Ponadto inne szczegółowe badania, wykazały, że niekonwencjonalne drożdże *Y. lipolytica* wykazują „preferencję” w stosunku do kotranslacyjnego systemu translokacji (75% w porównaniu z 30% u *S. cerevisiae*) [36].

W trakcie etapu translacji, w świetle ER następuje wiązanie powstającego polipeptydu przez białko opiekuńcze (BiP; Kar2) i foldazy rezydujące w ER. Białko Kar2 stanowi jeden z kluczowych elementów wspomagających transport polipeptydu do światła ER. Foldazy obecne w ER wspomagają fałdowanie polipeptydu do struktury natywnej, w tym tworzenie mostków disiarczkowych (-S-S-), a także przeprowadzają modyfikacje potranslacyjne. Jako „wąskie gardło” procesów zachodzących w świetle ER wskazano m.in.

tworzenie wiązań -S-S- [37,38]. W procesie tworzenia wiązań -S-S- uczestniczy białkowa izomeraza disiarczkowa Pdi, oksydaza izomerazy Pdi - Ero1, oraz syntetaza glutationu Gsh1 [38]. Jak wiadomo, reakcja ta opiera się na stochastycznym utlenianiu-redukcji łańcuchów bocznych cysteiny, które zużywa znaczne ilości czynników utleniających (O₂) i redukujących (glutationu). Precyzyjne zrównoważenie procesów fałdowania i formowania -S-S- jest kluczowe dla zachowania homeostazy komórki. W przeciwnym razie, komórka jest narażona na ciężki stres oksydacyjny powodowany nieskoordynowaną działalnością Pdi1p/Ero1p i formowaniem reaktywnych form tlenu [38].

Przechodzące do światła ER białka zawierające miejsca glikozylacji są wiązane przez foldazę Cne1 (kalneksynę); białko chaperonowe zaangażowane w fałdowanie, obróbkę N-glikanów i kontrolę jakości przeprowadzonej glikozylacji. Zgodnie z danymi literaturowymi, cykl kalneksyny funkcjonuje w komórkach *Y. lipolytica*, natomiast nie został zidentyfikowany w drożdżach *S. cerevisiae* [26,28]. Boisrame i wsp. wskazali na 45% podobieństwo kalneksyny *Y. lipolytica* z ludzkim odpowiednikiem [20].

Prawidłowe fałdowanie białka w ER jest kluczowe i decyduje o tym, czy dany polipeptyd przechodzi do kolejnych etapów szlaku sekrecyjnego (transport do aparatu Golgiego (AG; *Golgi Apparatus*) przy udziale pęcherzyków COPII), czy staje się celem degradacji związanej z ER (ERAD; *ER-associated degradation*). Funkcjonalny ERAD chroni komórkę gospodarza przed stresem związanym z UPR (*Unfolded Protein Response*) [39–41]. Szlak UPR jest regulowany przez czynnik transkrypcyjny Hac1, którego funkcjonowanie jest, co do zasady, filogenetycznie zakonserwowane wśród *Eukarya* [42–46]. Pewna zmienność dotyczy struktury transkryptu *HAC1*, jego ekspresji i wzoru splicingu, a także sekwencji motywu wiążącego DNA oraz regulowanego przez Hac1 sub-transkryptomu [44–47]. Czynnikiem Hac1 pośredniczy w deregulacji ponad 100 różnych genów, zaangażowanych w procesy tworzenia, fałdowania i sekrecji białek (np. *KAR2*, *PDI*, *ERO1*, *MDP1*, *CNE1* i inne), ale także w syntezę lipidów czy ekspansję błon [44]. Błędnie sfałdowane lub zmodyfikowane polipeptydy eksportowane są z ER, ubikwitynowane, a następnie kierowane do degradacji przez proteasom. Pomyślnie zakończony etap fałdowania i kontroli jakości ER jest sygnałem do transportu białek prekursorowych do AG za pośrednictwem pęcherzyków COPII, które składają się głównie z Sar1, Sec16, Sec23/Sec24 i Sec13/31. W AG powstający polipeptyd podlega dalszym modyfikacjom i oczekuje na przejście do innych miejsc docelowych w komórce, w tym PM. Wśród modyfikacji post-translacyjnych następujących w AG funkcjonuje kolejna seria glikozylacji. Co ważne w tym kontekście, *Y. lipolytica* jest gatunkiem drożdży (wraz z *P. pastoris*), który charakteryzuje się brakiem α -1,3-

mannozylotransferazy Mnn1 (**Rycina 1**), czyli często wskazywanego czynnika prowadzącego do nadmiernej mannozylacji wydzielanych heterologicznych glikoprotein. Brak Mnn1 stanowi cenny atut do wykorzystania tych platform drożdżowych do produkcji białek terapeutycznych [48,49].

Dystalne etapy sekrecji polipeptydów na drodze transportu pęcherzykowego w komórkach *Y. lipolytica* są najmniej poznane. Badania proteomiczne przeprowadzone przez Swennen i Beckerich [27] pozwoliły wskazać szereg czynników wiążących (*Tethering factors*) białek kompleksu SNARE (*Soluble N-ethylmaleimide-sensitive factor attachment protein receptor*) i SM funkcjonujących w tych komórkach (**Rycina 1**).

W konsekwencji silnej nadekspresji heterologicznego genu kodującego r(s)-Prot dochodzi do przeciążenia systemu translacyjno-sekrecyjnego komórki, wywołując stres ER, inicjujący zjawisko UPR. Biologiczny sens UPR polega na przywróceniu homeostazy komórkowej poprzez usunięcie wąskich gardeł szlaku synteza-łańdowanie-wydzielanie. Nagromadzenie błędnie sfałdowanych polipeptydów powoduje zwiększone wiązanie z białkiem Kar2 i w konsekwencji dimeryzację sensora UPR Ire1, co z kolei aktywuje splicing mRNA czynnika transkrypcyjnego (TF; *Transcription Factor*) Hac1. W zrównoważonych warunkach Ire1 jest wiązany przez Kar2, w którym pośredniczy Sls1 i nie jest dostępny do przetwarzania transkryptu *HAC1* [19,21]. W komórkach *Y. lipolytica* nadekspresja *SLS1* wzmocniła interakcję Kar2-Ire1, uniemożliwiając przetwarzanie *HAC1*, podczas gdy genotyp Δ *sls1* wykazywał konstytutywny UPR na wysokim poziomie [21], tym samym wskazując na istotną rolę Sls1 w UPR.

1.3. Strategie eliminacji „wąskich gardeł” szlaku translacyjno-sekrecyjnego

W zależności od stanu fizjologicznego komórki oraz właściwości biochemicznych syntetyzowanego polipeptydu, stopień „przeciążenia” w/w etapów szlaku translacyjno-sekrecyjnego może być różny, stąd i różny będzie charakter powstających „wąskich gardeł”, a w konsekwencji - miejsce blokowania tranzytu polipeptydu [5]. W procesach tych, poza zapotrzebowaniem na znaczną ilość elementów budulcowych i energii, wymagana jest także określona pula wyspecjalizowanych białek i błon komórkowych. Bardzo często ich dostępność warunkuje płynność systemu translacyjno-sekrecyjnego [50,51]. Niektóre doniesienia literaturowe wskazują właśnie procesy łańdowania i sekrecji polipeptydów jako „wąskie gardła” całego procesu produkcji r-Protów [52–54]. Szczególnie podatne etapy obejmują transport między ER a AG, translokacji do ER, oraz proces łańdowania [55].

W zależności od specyficznych właściwości produkowanego białka (struktura pierwszorzędowa / drugorzędowa, wielkość, stopień oligomeryzacji, glikozylacja) odmienne zjawiska mogą ograniczać ogólną zdolność wydzielniczą komórki. Co ważne, intensywna synteza rs-Protów ma bezpośredni wpływ na metabolizm komórki gospodarza, nierzadko negatywnie wpływając na parametry biologiczne, takie jak tempo wzrostu czy tempo zużycia substratu [56,57]. Dlatego staranne dostosowanie poziomu produkcji danego rs-Prot do możliwości komórki gospodarza jest kluczowe.

Pomimo dobrego poznania mechanizmów molekularnych szlaku syntezy i sekrecji polipeptydów w komórkach drożdży, zaprojektowanie ukierunkowanej strategii modyfikacji tego szlaku stanowi duże wyzwanie. Ponadto uprzednio opracowane strategie nierzadko nie znajdują zastosowania w przypadku nadprodukcji rs-Protów o odmiennych właściwościach biochemicznych [56]. Wśród strategii modyfikacji genetycznych stosowanych celem usprawnienia syntezy rs-Protów wymienia się: i) inżynierię sekwencji DNA kodującej określone rs-Prot, tu np. optymalizację użycia kodonów, ii) optymalizację konstrukcji genowych, iii) dobór warunków hodowlanych, jak również iv) inżynierię molekularnych mechanizmów syntezy i sekrecji polipeptydów [28,58–60]. W każdej z zaproponowanej strategii wskazano szereg rozwiązań, które w różnym stopniu wpływają na poziom nadekspresji heterologicznego genu i sekrecji docelowego białka. Niektóre z podanych strategii zastosowano w komórkach *Y. lipolytica*, m.in. : i) optymalizację użycia kodonów, sprzężanie z białkami fuzyjnymi, czy peptydami sygnałnymi [31,61,62]; ii) zwiększanie liczby kopii heterologicznego genu, inżynierię sekwencji regulatorowych [17,62–65]; iii) stosowanie induktorów / inhibitorów procesów fizjologicznych, zabiegi z zakresu inżynierii bioprosesowej [10,66,67]; iv) nadekspresję białek chaperonowych, czynników transkrypcyjnych, delecję proteaz, bezpośrednią ingerencję w mechanizmy fałdowania i dojrzewania polipeptydów, transportu białek do miejsc docelowych itd. [68–73]. O ile strategie i) i ii) dotyczą głównie manipulacji na poziomie ekspresji heterologicznego genu, o tyle zastosowanie ostatniej strategii (iv) pozwala zwiększyć zdolność komórki w zakresie prawidłowego dojrzewania polipeptydów i przepustowość szlaku sekrecyjnego. Obecne strategie inżynierii szlaku translacyjno-sekrecyjnego koncentrują się głównie w czterech obszarach: i) inżynierii systemu fałdowania i kontroli jakości białek [69]; ii) inżynierii wewnątrzkomórkowego transportu pęcherzykowego [68]; iii) minimalizacji degradacji proteolitycznej [73–75] oraz iv) inżynierii modyfikacji potranslacyjnej [76].

Według ostatnich badań, zastosowanie pięciu nowych peptydów sygnałnych znacząco poprawiło zdolność wydzielniczą komórek drożdży *Y. lipolytica* w zakresie produkcji dwóch amylolitycznych r-Prots: α -amylazy oraz glukoamylazy [31]. Oprócz doboru sekwencji SP, alternatywną strategią stosowaną w celu zwiększenia wydajności etapu translokacji uznaje się nadekspresję składników translokonu i stabilizację/nadekspresję cytozolowych białek opiekuńczych [69,77]. Najnowsze doniesienia literaturowe wskazują na zastosowanie strategii „push and pull” (znanej w inżynierii metabolicznej) celem zwiększenia przepływu różnych r-Prots na etapie translokacji poprzez nadekspresję wybranych czynników cyklu Hsp70 w cytozolu u *P. pastoris* [78].

W zakresie manipulacji środowiskiem światła ER celem zwiększenia zdolności wydzielniczej rs-Prots zastosowano m.in. nadekspresję białek chaperonowych lub enzymów ER, w tym genów: *SIL1*, *LHS1*, *JEM1*, *SCJ1*, *KAR2*, *ERO1*, *PDII* w *S. cerevisiae* [79,80], *P. pastoris* [69,81] i *K. lactis* [82]. Podjęto wiele prób poprawy wydzielniczej produkcji białek opartych na nadekspresji izomerazy *PDII*, samej lub razem z oksydazą *ERO1*. Zastosowanie tego podejścia przyczyniło się do uzyskania wyższych ilości różnych rs-Prots produkowanych przez kilka gatunków drożdży, w tym *S. cerevisiae*, *P. pastoris* i *K. lactis* [48]. Co ciekawe – korzystny efekt nadekspresji Pdi1/Ero1 był obserwowany także w przypadku nadekspresji r-Prot nie zawierającego mostków disiarczkowych [50]. Inna strategia polegała na nadekspresji Pdi1 połączonej z jednoczesną nadekspresją białka opiekuńczego ER - Kar2 (prace przeglądowe opisujące te strategie: [52,83–85]).

W celu poprawy wydzielania rs-Prot u *S. cerevisiae* podjęto próby zatrzymania zjawiska ERAD poprzez delecję genów jak: *HTM1*, *YOS9*, *HRD1*, *HRD3* lub *UBC7*, zaangażowanych w ten proces na różnym etapie [86]. W efekcie tego zabiegu, wzrost wydzielania rs-Prots był, jeśli w ogóle, na znikomym poziomie w porównaniu np. z nadekspresją czynnika uczestniczącego w fałdowaniu [86]. Podobne efekty zaobserwowano u *P. pastoris* przy delecji genów *HRD3*, *DER1*, *RPN4* lub *DOA1* [87]. Powyższe wyniki wskazują, że wydzielenie r-Prot dalej z ER do AG nie wydaje się być głównym ograniczeniem zdolności wydzielniczej komórek producentów drożdżowych [57].

Główny czynnik transkrypcyjny regulujący ERAD i dalszy UPR, Hac1, jest jednym z kluczowych celów modyfikacji genetycznych ukierunkowanych na poprawę produkcji rs-Prots w systemach drożdżowych [47,50,54]. Aktywacja UPR przez nadekspresję indukowanej wersji Hac1ⁱ zwiększyła wydzielanie kilku natywnych i rekombinowanych białek

w *S. cerevisiae*, *P. pastoris*, i *A. niger* [44,70,88–90]. Co więcej, ponad 4-krotną poprawę w sekrecji r-Prot osiągnięto u *P. pastoris* poprzez ko-nadekspresję *HAC1* z innym TF *MSN4* oraz białkami chaperonowymi rezydującymi w ER [72]. Jak pokazują uprzednio przeprowadzone badania, w zależności od właściwości produkowanego r(s)-Prot ko-ekspresja *HAC1* przyczyniła się do zwiększenia lub obniżenia heterologicznej syntezy rs-Prot, lub nie wywarła na nią istotnego wpływu [48,90]. Pomimo tego, że nadekspresja *Hac1ⁱ* zwiększa wydzielanie rs-Protów, wykazano także, że przy zbyt silnej nadekspresji wywołuje dalszy stres w komórkach gospodarzy, co zostało zaobserwowane w [44,89].

Intensyfikację kolejnego etapu szlaku sekrecyjnego, wyjścia białek z ER i tworzenia pęcherzyków COPII, otrzymano m.in. poprzez nadekspresję genu *SEC16*, który zwiększył wydzielanie trzech modelowych rs-Protów w *S. cerevisiae*, zmniejszając w ten sposób stres ER [91]. Co więcej, delecja genu *GOS1*, kodującego białko zaangażowane w transport wsteczny AG, zwiększyła zdolność sekrecyjną komórek *S. cerevisiae* [92] i umożliwiła osiągnięcie zrównoważonego przepływu białek w AG.

Celem zwiększenia szybkości transportu rs-Protów w obrębie etapów dystalnych podjęto wiele prób obejmujących m.in. nadekspresję białek SNARE i Sec1/Munc18 (SM), które odgrywają kluczową rolę w fuzji błon komórkowych [28,84]. Nadekspresja genów *SSO1* i *SSO2*, zaangażowanych w transport białek z AG do PM, umożliwiła wzrost potencjału sekrecyjnego drożdży *P. pastoris* czy *S. cerevisiae* w zakresie produkcji rs-Protów [69,93]. Co więcej, w komórkach *S. cerevisiae*, nadekspresja genu *SECI*, bezpośrednio związanego z transportem pęcherzykowym pomiędzy AG a PM, powodowała zwiększoną sekrecję kilku białek (α -amylazy i insuliny) [68]. Z kolei nadekspresja genu *SLY1* (transport z ER do AG) wykazała pewną specyficzność względem rodzaju r-Prot [68]. Zatem, można wnioskować, że efekty nadekspresji genów związanych z szlakiem ER-AG-SNARE i białek SM wydają się być specyficzne dla danego r-Prot.

Pomimo, iż domyślny szlak transportu komórkowego rs-Protów prowadzi w kierunku PM i dalej do przestrzeni zewnątrzkomórkowej, zdarza się, że polipeptydy są błędnie sortowane do wakuoli [73,94]. Dzieje się tak, gdy prawidłowo sfałdowane peptydy zostają błędnie rozpoznane przez receptor Vps10 (*Vacuolar sorting factor 10*), który, jako element systemu kontroli jakości, natywnie kieruje nieprawidłowo sfałdowane polipeptydy do degradacji. Co ciekawe, delecja genu *VPS10* zapobiegła gromadzeniu się białka fluorescencyjnego w wakuoli drożdży *S. cerevisiae* [95]. Ponadto, usunięcie Vps10,

zwiększyło wydzielanie innych rs-Prots w komórkach: *H. polymorpha* [96], *S. cerevisiae* [97] i *S. pombe* [98]. Wykazano także, że zastosowanie szczepów z delecją genów kodujących wakuolarną proteazę (np. *Aprb1*, *Δpep4*, *Δyps1*) przyczynia się do poprawy wydajności produkcji rs-Prots w *P. pastoris* [71] i *S. cerevisiae* [99], prowadząc do około 10-krotnej poprawy wydzielania ludzkiego interferonu- β [99]. Podobnie Jonson i wsp. udowodnili, że delecja genu metalo-endoproteazy *CYMI* w *S. cerevisiae* nie tylko zmniejszyła proteolizę wewnątrzkomórkową, ale także zwiększyła wydzielanie kilku rs-Prots (m.in. hormonu wzrostu) [100]. Z kolei w komórkach *Y. lipolytica* wyeliminowano zewnątrzkomórkową proteazę (*ΔAXP*, *ΔAEP*), co przyczyniło się do zwiększenia stabilności wydzielanych białek. Ponadto ostatnie doniesienia wykazały, że mutacja w genie kodującym białko Mtc6, zaangażowane w regulację autofagii, osłabia ten proces i poprawia produkcję rs-Prots w *Kluyveromyces marxianus* [101]. Stąd ograniczenie procesu autofagii zaproponowano jako kolejną strategię poprawy wydajności produkcji rs-Prots w drożdżach.

Po wydzieleniu z PM, rs-Prots dyfundują przez ścianę komórkową, która może niekiedy stanowić dla nich barierę [102]. Wpływ białek ściany komórkowej na wydajność syntezy rs-Prots zaobserwowano dla systemów drożdżowych. Li i wsp. [103] wykazali, że delecja genu *CWP2* (mannoproteiny ściany komórkowej) usprawnia proces syntezy-sekrecji cellobiohydrolazy o 85,9% w komórkach *S. cerevisiae*. Ponadto na podstawie podwyższonego poziomu ekspresji genów zaangażowanych w biogenezę rybosomów, stwierdzono, że wprowadzona modyfikacja *Δcwp2* wpłynęła nie tylko na poziom ekspresji genów uczestniczących w biosyntezie ściany komórkowej, ale także na poziom translacji białek. Korzystny wpływ nadekspresji trzech genów kodujących białka ściany komórkowej (*CCW12*, *CWP2* i *SEDI*) zaobserwowano w trakcie produkcji jednołańcuchowych przeciwciał w komórkach *S. cerevisiae* [104].

1.4. Hodowle ciągłe w stanie ustalonym – solidne narzędzie w analizach omicznych

Podczas nadprodukcji rs-Prots wiele zmian w fizjologii komórki producenta zostaje wyzwolonych [105–108]. Ekspozycja na nieoptymalne/zmienne warunki wzrostu może dodatkowo zaburzyć procesy biologiczne. Warunki środowiskowe mogą modulować wydajność zarówno transkrypcji i translacji genu kodującego rs-Prot, jak i fałdowania polipeptydu [66,67,109]. Wobec tego, usprawnienie szlaku sekrecyjnego wymaga dogłębnej wiedzy na temat biologii komórki. Kinetyka wytwarzania rs-Prots i odpowiedź gospodarza na takie wyzwanie może być dokładnie zbadana w komórkach utrzymywanych w tzw. „stanie

ustalonym” [110–112]. Technologie ciągłej hodowli komórkowej obejmujące m.in. procesy chemostatyczne opracowano kilkadziesiąt lat temu, i są nadal stosowane celem osiągnięcia w/w stanu [113,114]. W chemostacie komórki utrzymywane są w ustalonej fazie wzrostu poprzez ciągle dodawanie składników odżywczych i jednocześnie usuwanie/odbieranie hodowli (płynu pohodowlanego) z naczynia bioreaktora z określoną zadaną szybkością rozcieńczenia (D ; [h^{-1}]). W stanie ustalonym stężenie wewnątrz- i zewnątrzkomórkowych cząsteczek wytwarzanych lub metabolizowanych przez hodowane komórki pozostaje stałe przez dłuższy czas (tzw. *residence time*) [115]. Zaobserwowano, że hodowla osiąga stan ustalony, dopasowując tempo wzrostu do stopnia rozcieńczenia w szerokim zakresie wartości współczynnika rozcieńczenia.

Jedną z kluczowych zalet takiego sposobu hodowli jest otrzymanie w miarę jednorodnej i zsynchronizowanej w określonej fazie wzrostu populacji komórek. Takie podejście eliminuje zróżnicowanie w obrębie szeregu zjawisk molekularnych związanych ze wzrostem w danej fazie wzrostu, odpowiedzią na czynniki stresowe zmieniające się w czasie hodowli lub innymi rodzajami heterogeniczności występujących w typowej hodowli okresowej. Brak kontroli tych parametrów może prowadzić do błędnej interpretacji zmian fizjologicznych zachodzących w komórce. Starannie kontrolowane i zdefiniowane warunki fizjologiczne, jakie można uzyskać w hodowlach chemostatycznych, umożliwiają otrzymywanie reprezentatywnych próbek biologicznych do analizy za pomocą technik omicznych. Połączenie hodowli chemostatycznych w stanie ustalonym i technik genomiki funkcjonalnej zapewnia solidną podstawę eksperymentalną do analizy wielu procesów biologicznych w mikroorganizmach [116]. Hodowle chemostatyczne są szeroko stosowane zarówno dla badań stosowanych, jak i podstawowych z wykorzystaniem różnych organizmów, jak np. *Escherichia coli*, *P. pastoris*, *S. cerevisiae* i *Aspergillus niger* [111,117–119]. Jak dotąd, większość badań z udziałem drożdży *Y. lipolytica* utrzymywanych w chemostacie koncentrowała się na wytwarzaniu metabolitów o niskiej masie cząsteczkowej, jak: kwas cytrynowy, erytrytol i lipidy [120–122].

2. Hipotezy i cele badań

Hipotezy

H_0 : Zjawiska molekularne wywołane w komórkach *Y. lipolytica* nadprodukcujących rs-Prot zależą od właściwości biochemicznych danego rs-Prot,

H_0 : Geny deregulowane pod wpływem nadprodukcji rs-Prot mogą być stosowane jako geny pomocnicze (*SHs*; *Secretion Helpers*), korzystnie wpływające na syntezę i sekrecję rs-Prot w drożdżach *Y. lipolytica*,

H_0 : Działanie poszczególnych genów pomocniczych *SH* jest modulowane przez warunki środowiskowe zastosowane w trakcie hodowli produkcyjnej, na przykład temperatury.

Cele badań

1. Zwiększenie możliwości komórek drożdży *Y. lipolytica* w zakresie syntezy i prawidłowego dojrzewania polipeptydów, jak również przepustowości szlaku sekrecyjnego poprzez:
 - Wskazanie genów bezpośrednio zaangażowanych w syntezę, dojrzewanie i sekrecję r-Protów na podstawie profilu ich ekspresji,
 - Modyfikację zidentyfikowanych mechanizmów badanych zjawisk celem usprawnienia produkcji rs-Protów w badanym systemie ekspresyjnym.
2. Poznanie molekularnych podstaw promującego wpływu działania wybranego *SH* na podstawie globalnej analizy transkryptomu zmodyfikowanych komórek *Y. lipolytica* w warunkach jego ko-nadekspresji wraz z rs-Prot, na przykładzie czynnika transkrypcyjnego Hac1.

Założono, że wyjaśnienie mechanizmów działania szlaku translacyjno-sekrecyjnego u *Y. lipolytica* w warunkach nadprodukcji r(s)-Protów o odmiennych właściwościach biochemicznych, pozwoli zrozumieć zjawiska zachodzące w komórce w różnych kompartmentach i procesach. Wykorzystanie zdobytej wiedzy przypuszczalnie umożliwi zaprojektowanie strategii modyfikacji genetycznych, które korzystnie wpłyną na poprawę syntezy i sekrecji rs-Protów w tym systemie ekspresyjnym.

3. Materiały i metody

3.1. Podłoża hodowlane

Hodowle komórek bakteryjnych *E. coli* transformowanych plazmidami prowadzono w temp. 37 °C w podłożu LB o następującym składzie [g L⁻¹]: ekstrakt drożdżowy, 5 (BTL, Łódź, Polska); baktopenon, 10 (BTL); NaCl, 5 (POCH, Gliwice, Polska); z dodatkiem odpowiedniego antybiotyku (ampicyliny [100 mg L⁻¹] lub kanamycyny 40 [μg L⁻¹]) oraz agaru (w podłożach stałych), 15 (Biomaxima, Lublin, Polska).

Transformowane szczepy drożdży *Y. lipolytica* inkubowano w temp. 28 °C w minimalnym podłożu YNB o następującym składzie [g L⁻¹]: YNB, 1.7 (Sigma-Aldrich, Merck KgaA, St.Louis, USA); (NH₄)₂SO₄, 5 (POCH); agar (w podłożach stałych), 15 (Biomaxima). Ponadto, rekombinowane szczepy *Y. lipolytica* hodowano w podłożu YPD [g L⁻¹]: ekstrakt drożdżowy, 10 (BTL); baktopenon, 20 (BTL); glukoza, 20 (POCH), a także YPS [g L⁻¹]: ekstrakt drożdżowy, 10 (BTL); baktopenon, 20 (BTL); glukoza, 20 (POCH); skrobia, 10 (Sigma-Aldrich) z dodatkiem agaru, 15 (Biomaxima) w podłożach stałych.

Celem przygotowania prekultur sterylne podłoża zaszczerpiono pojedynczą kolonią drożdży uprzednio ożywioną ze stoków glicerolowych (-80 °C) na stałym podłożu YPD (skład j.w.). Prekultury drożdży *Y. lipolytica* służące do zaszczerpienia hodowli bioreaktorowych były przygotowywane w kolbach Erlenmeyera w 30 ml podłoża YPG₂₀ o następującym składzie [g L⁻¹]: ekstrakt drożdżowy, 10 (BTL); baktopenon, 20 (BTL); glicerol, 20 (POCH). Natomiast, prekultury dla hodowli okresowych były przygotowywane w szklanych probówkach w podłożach: YPG₂₀, YNB oraz YPD (skład j.w.). Hodowle prekultury prowadzono w inkubatorze z wytrząsaniem (BIOSAN, ES-20, Riga, Latvia) przy 250 rpm, w temp. 28 °C, przez 22 h.

W celu wyboru reprezentatywnego szczepu *Y. lipolytica* spośród transformantów, dla wszystkich badanych wariantów prowadzono selekcję subklonów w 5 ml podłoża YPG₂₀ lub YNB (skład j.w.) w kolbach wytrząsanych przy 250 rpm, w temp. 28-30 °C, do 72 h.

Hodowle drożdży *Y. lipolytica* z nadekspresją *SHs* prowadzono w 30 ml podłoża YPD (skład j.w.) w kolbach wytrząsanych przy 220 rpm, w temp. 25/30 °C, przez 96 h.

Hodowle bioreaktorowe drożdży *Y. lipolytica* prowadzono w 300 ml podłoża YPG₁₀₀ o następującym składzie [g L⁻¹]: (ekstrakt drożdżowy, 10 (BTL); baktopenon, 20 (BTL); glicerol, 100 (POCH) w temp. 28 °C.

3.2. Konstrukcja rekombinowanych szczepów

3.2.1. Przygotowanie konstrukcji do nadekspresji genów kodujących r(s)-Prots - pochodne szczepów Po1g oraz Po1h (P.1, P.2)

Do nadekspresji wytypowano trzy r(s)-Prots służące jako białka reporterowe, w tym: dwa średniej wielkości białka enzymatyczne: α -amylazę z wołka ryżowego *Sitophilus oryzae* (SoA, [56 kDa]) [123] i glukoamylazę z termofilnego grzyba *Thermomyces lanuginosus* (TIG, [65 kDa]) [124], a także małe białko emitujące fluorescencję YFP (*Yellow Fluorescence Protein*, YFP, [27 kDa]). Biochemiczne właściwości białek reporterowych przedstawiono w **Tabeli 1**. Wszystkie heterologiczne geny zostały zaopatrzone w odpowiednie SPs w celu wydajnej sekrecji polipeptydów do przestrzeni zewnątrzkomórkowej (scSoA/scTIG/scYFP). Dodatkowo, gen kodujący YFP wbudowano w konstrukcję kasety bez SP, tym samym tworząc wariant wewnątrzkomórkowej produkcji białka fluorescencyjnego (inYFP).

Tabela 1. Właściwości biochemiczne r(s)-Prots stosowanych w tej pracy.

r-Prot	Masa molekularna [Da]	Liczba AA	Teoretyczny pI	Miejsca N-glikozylacji	Miejsca O-glikozylacji	Liczba Cys	Całkowita liczba kombinacji S-S
YFP	26 991.54	239	5.58	-	1	2	2
SoA	52 140.40	474	4.88	1	8	11	35 696
TIG	65 154.96	600	4.90	3	18	8	764

W badaniach nad odpowiedzią komórki *Y. lipolytica* na nadprodukcję rekombinowanych białek w ciągłych hodowlach bioreaktorowych wykorzystano wektory oraz szczepy przedstawione w **Tabeli 2** oraz **Tabeli 3**. Były to szczepy skonstruowane w podłożu genetycznym szczepu Po1g (*MatA*, *leu2-270*, *ura3-302::URA3*, *xpr2-332*, *axp-2*) oraz szczepu Po1h (genotyp: *MatA*, *ura3-302*, *xpr2-322*, *axp1-2*). Szczepy te wywodzą się z referencyjnego szczepu dzikiego W29 (CLIB89; ATCC20460).

Tabela 2. Szczepy bakterii zastosowane w tej pracy

Nazwa szczepu	Numer w kolekcji	Konstrukcja	Rodzaj klonowania	Pochodzenie
<i>E.coli</i> JM109		Genotyp: <i>F'</i> (<i>traD36</i> , <i>proAB</i> +, <i>lacIq</i> , Δ (<i>lacZ</i>)M15), <i>endA1</i> , <i>recA1</i> , <i>hsdR17</i> (<i>rk</i> -, <i>mk</i> +), <i>mcrA</i> , <i>supE44</i> , λ - <i>gyrA96</i> , <i>relA1</i> , Δ (<i>lacproAB</i>), <i>thi-1</i>		Sigma-Aldrich
JM109_URA3	GGE167	<i>pSB1A3-zetaUP-URA3ex-zetaDOWN</i>	Klonowanie modułarne Golden Gate	Ta praca
JM109_scSoA	GGE162	<i>pSB1A3-zetaUP-URA3ex-4UASpTEF-scSoA-tLip3-zetaDOWN</i>		
JM109_scTIG	GGE76	<i>pSB1A3-zetaUP-URA3ex-4UASpTEF-scTIG-tLip3-zetaDOWN</i>		
JM109_scYFP	GGE166	<i>pSB1A3-zetaUP-URA3ex-4UASpTEF-scYFP-tLip3-zetaDOWN</i>		
JM109_inYFP	GGE164	<i>pSB1A3-zetaUP-URA3ex-4UASpTEF-inYFP-tLip3-zetaDOWN</i>		
JM109_scSoA	4.29	<i>pYLSC-SoA</i>	Klonowanie molekularne z zastosowaniem wektora pYLSC/ pYLEX	
JM109_scTIG	2.1.2.10	<i>pYLSC-TIG</i>		
JM109_scYFP	ECB31	<i>pYLSC-YFP</i>		
JM109_inYFP	ECB05	<i>pYLEX-YFP</i>		
JM109_scYFP	ECB71	<i>JMP62-pTEF-scYFP-URA3ex</i>	Klonowanie molekularne z zastosowaniem wektora JMP62	
JM109_PDI	ECB70	<i>JMP62-pTEF-PDI-LEU2ex</i>		
JM109_SSA8	ECB72	<i>JMP62-pTEF-SSA8-LEU2ex</i>		
JM109_SEC1	ECB73	<i>JMP62-pTEF-SEC1-LEU2ex</i>		
JM109_SSA5	ECB74	<i>JMP62-pTEF-SSA5-LEU2ex</i>		
JM109_HAC1	ECB75	<i>JMP62-pTEF-HAC1-LEU2ex</i>		
JM109_SSO1	ECB78	<i>JMP62-pTEF-SSO1-LEU2ex</i>		
JM109_SLS1	ECB79	<i>JMP62-pTEF-SLS1-LEU2ex</i>		
JM109_YET3	ECB80	<i>JMP62-pTEF-YET3-LEU2ex</i>		
JM109_USO1	ECB82	<i>JMP62-pTEF-USO1-LEU2ex</i>		
JM109_RPL3	ECB84	<i>JMP62-pTEF-RPL3-LEU2ex</i>		
JM109_CWP11	ECB85	<i>JMP62-pTEF-CWP11-LEU2ex</i>		
JM109_CNE1*	ECB89	<i>JMP62-pTEF-CNE1*-LEU2ex</i>		

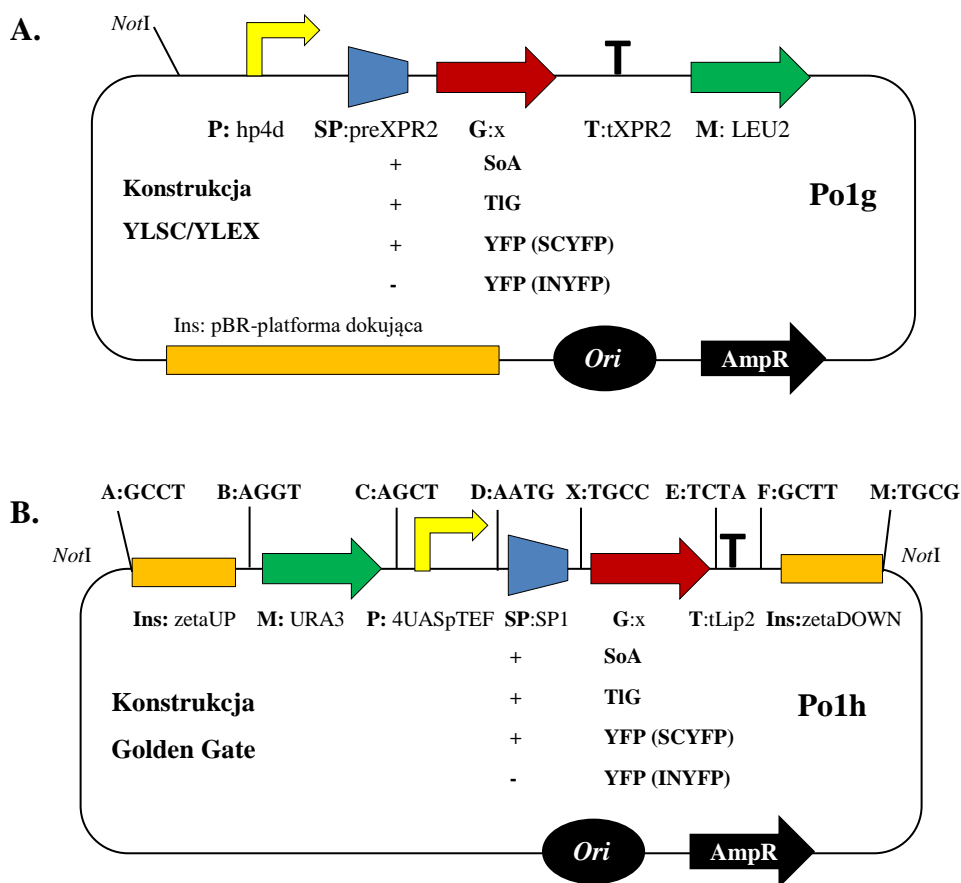
Dla pochodnych szczepu Polg, sekwencje heterologicznych genów wprowadzono przy użyciu komercyjnego systemu do klonowania YLEX™ kit (Yeastern Biotech Co. Ltd), wyposażonego w platformę dokującą pBR322 oraz marker selekcyjny *LEU2*. Ekspresja heterologicznych genów była regulowana za pomocą silnego, semikonstytutywnego promotora hp4d [125] (**Rycina 2A**). Sekwencje genów kodujących rs-Prot (scSoA, scTIG,

scYFP) wklonowano w wektor sekrecyjny pYLSC z wbudowaną sekwencją SP preXPR2 (EC: 3.4.21.62), natomiast gen kodujący wewnątrzkomórkowe białko YFP (inYFP) wklonowano w wektor ekspresyjny pYLEX, przy użyciu enzymów restrykcyjnych (RE; *Restriction Enzyme*) *SfiI* oraz *KpnI*. Szczep *Y. lipolytica* Po1g (*ura*⁺, *leu*⁻) transformowany wyłącznie kasetą markera *LEU2* zastosowano jako kontrolę negatywną do oceny tła fluorescencji (Po1g_Leu2+).

Dla pochodnych szczepu Po1h, kasetę zawierającą pojedynczą jednostkę transkrypcyjną TU (*Transcription Unit*) skonstruowano przy wykorzystaniu techniki klonowania modularnego Golden Gate zgodnie z uprzednio opracowanym protokołem [13]. Reakcje przeprowadzono stosując ligazę T4 i endonukleazę restrykcyjną *BsaI* (NEB, Ipswich, MA, USA). Każda TU składała się z sekwencji heterologicznego genu (*SoA*, *TIG*, *YFP*) pod kontrolą silnego, semi-konstytutywnego promotora 4UAS-pTEF [62] oraz terminatora tLip2 [126]. Geny kodujące białka sekrecyjne zostały opatrzone sekwencją peptydu sygnałowego SP1 (EC number: 3.2.1.58) natywną dla *Y. lipolytica* egzo-1,3-beta-glukanazy (*B03564g*) [31]. Dodatkowo, każda konstrukcja zawierała marker selekcyjny *URA3* oraz miejsca integracji zetaUP oraz zetaDOWN (**Rycina 2B**). Szczep rodzicielski *Y. lipolytica* Po1h (*ura*⁻, *leu*⁺) transformowany wyłącznie kasetą markera *URA3* zastosowano jako kontrolę negatywną do oceny tła fluorescencji (Po1h_Ura3+).

Tabela 3. Szczepy drożdży zastosowane w tej pracy

Nazwa szczepu	Numer w kolekcji	Konstrukcja	Pochodzenie
<i>Y. lipolytica</i> Po1g		Genotyp: <i>Mata</i> , <i>leu2-270</i> , <i>ura3-302::URA3</i> , <i>xpr2-332</i> , <i>axp-2</i> Fenotyp: <i>Leu</i> ⁻ , Δ <i>AEP</i> , Δ <i>AXP</i> , <i>Suc</i> ⁺ , <i>pBR</i> platforma	Yeaster Biotech Co.
Po1g_Leu2	ECY32	<i>pYLSC</i>	Ta praca
Po1g_scSoA	1.1.2.2	<i>pYLSC_SoA</i>	[123]
Po1g_scTIG	2.2.1.2	<i>pYLSC_TIG</i>	Ta praca
Po1g_scYFP	ECY205	<i>pYLSC_YFP</i>	
Po1g_inYFP	ECY18	<i>pYLEX_YFP</i>	
<i>Y. lipolytica</i> Po1h		Genotyp: <i>Mata</i> , <i>ura3-302</i> , <i>xpr2-322</i> , <i>axp1-2</i> , <i>leu2-270::LEU2</i> Fenotyp: <i>Ura</i> ⁻ , Δ <i>AEP</i> , Δ <i>AXP</i> , <i>Suc</i> ⁺	[16]
Po1h_Ura3	GGY231	<i>zetaUP-URA3ex-zetaDOWN</i>	Ta praca
Po1h_scSoA	GGY237	<i>zetaUP-URA3ex-4UASpTEF-scSoA-tLip3-zetaDOWN</i>	
Po1h_scTIG	GGY228	<i>zetaUP-URA3ex-4UASpTEF-scTIG-tLip3-zetaDOWN</i>	
Po1h_scYFP	GGY251	<i>zetaUP-URA3ex-4UASpTEF-scYFP-tLip3-zetaDOWN</i>	
Po1h_inYFP	GGY245	<i>zetaUP-URA3ex-4UASpTEF-inYFP-tLip3-zetaDOWN</i>	
<i>Y. lipolytica</i> Po1f		Genotyp: <i>Mata</i> , <i>leu2-270</i> , <i>ura3-302</i> , <i>xpr2-322</i> , <i>axp-2</i> Fenotyp: <i>Ura</i> ⁻ , <i>Leu</i> ⁻ , Δ <i>AEP</i> , Δ <i>AXP</i>	ATCC MYA2613
Po1f_scYFP	ECY220.2	<i>Zeta-lox-pTEF-LEU2ex</i> , <i>Zeta-lox-pTEF-scYFP-URA3ex</i>	Ta praca
Po1f_scYFP_HAC1	ECY212.3	<i>Zeta-lox-pTEF-HAC1-LEU2ex</i> , <i>Zeta-lox-pTEF-scYFP-URA3ex</i>	
Po1f_scYFP_SSO1	ECY214.4	<i>Zeta-lox-pTEF-SSO1-LEU2ex</i> , <i>Zeta-lox-pTEF-scYFP-URA3ex</i>	
Po1f_scYFP_SSA5	ECY215.4	<i>Zeta-lox-pTEF-SSA5-LEU2ex</i> , <i>Zeta-lox-pTEF-scYFP-URA3ex</i>	
Po1f_scYFP_SSA8	ECY226.3	<i>Zeta-lox-pTEF-SSA8-LEU2ex</i> , <i>Zeta-lox-pTEF-scYFP-URA3ex</i>	
Po1f_scYFP_SEC1	ECY218.3	<i>Zeta-lox-pTEF-SEC1-LEU2ex</i> , <i>Zeta-lox-pTEF-scYFP-URA3ex</i>	
Po1f_scYFP_CWP11	ECY219.1	<i>Zeta-lox-pTEF-CWP11-LEU2ex</i> , <i>Zeta-lox-pTEF-scYFP-URA3ex</i>	
Po1f_scYFP_YET3	ECY221.3	<i>Zeta-lox-pTEF-YET3-LEU2ex</i> , <i>Zeta-lox-pTEF-scYFP-URA3ex</i>	
Po1f_scYFP_PDI	ECY222.2	<i>Zeta-lox-pTEF-PDI-LEU2ex</i> , <i>Zeta-lox-pTEF-scYFP-URA3ex</i>	
Po1f_scYFP_USO1	ECY223.3	<i>Zeta-lox-pTEF-USO1-LEU2ex</i> , <i>Zeta-lox-pTEF-scYFP-URA3ex</i>	
Po1f_scYFP_SLS1	ECY224.1	<i>Zeta-lox-pTEF-SLS1-LEU2ex</i> , <i>Zeta-lox-pTEF-scYFP-URA3ex</i>	
Po1f_scYFP_RPL3	ECY224.1	<i>Zeta-lox-pTEF-RPL3-LEU2ex</i> , <i>Zeta-lox-pTEF-scYFP-URA3ex</i>	
Po1f_scYFP_CNE1*	ECY227.2	<i>Zeta-lox-pTEF-CNE1*-LEU2ex</i> , <i>Zeta-lox-pTEF-scYFP-URA3ex</i>	

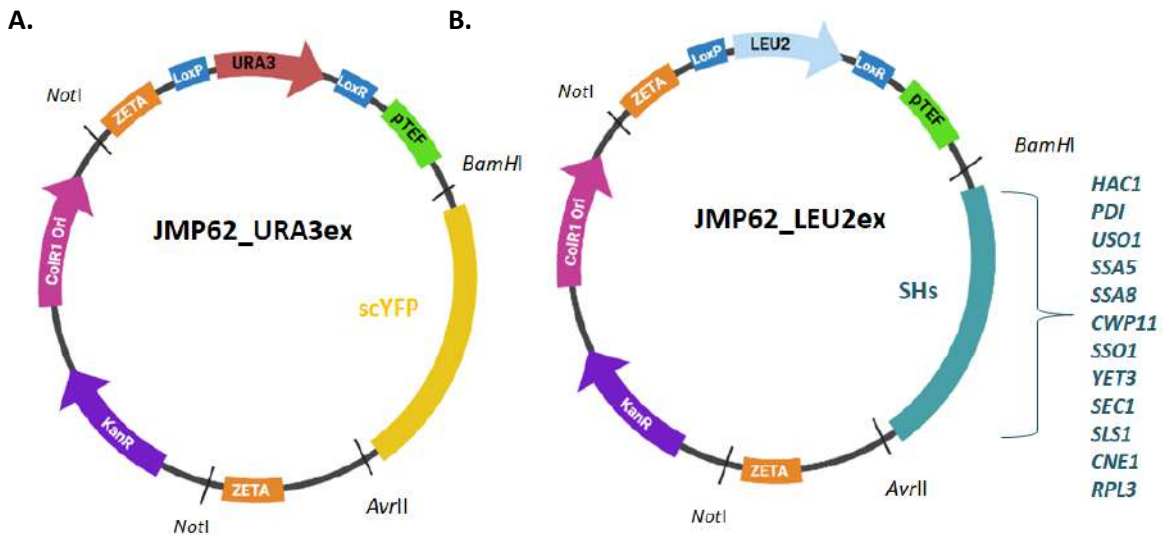


Rycina 2. Schematyczne przedstawienie strategii klonowania kasy do nadekspresji zastosowanej w tej pracy. A. Komercyjny system klonowania YLEX. B. Klonowanie modularne Golden Gate. Skróty: M – Marker aukсотroficzny (LEU2, URA3), P – promotor drożdżowy (php4d/p4UASpTEF), SP - peptyd sygnałny wydajnie kierujący r-Protos do przestrzeni zewnątrzkomórkowej (preXPR2/SP1), G – gen kodujący r-Prot (SoA, TIG, YFP), T– terminator drożdżowy (tXPR2/tLip2), Ins: zetaUP/zetaDOWN – sekwencje kierujące zeta umożliwiające losową insercję kasy do genomu gospodarza (rekombinacja niehomologiczna) lub pBR-platforma dokująca (rekombinacja homologiczna), A-J – litery oznaczające poszczególne 4 nt sekwencje flankujące, AmpR – oporność na ampicylinę, Ori – bakteryjne miejsce replikacji, *NotI* – miejsce rozpoznawane przez endonukleazę restrykcyjną *NotI* (miejsce trawienia przed transformacją drożdży).

Sekwencję DNA kodującą białko reporterowe YFP zaopatrzono w sekwencję peptydu sygnałnego SP1 celem efektywnego wydzielania polipeptydu poza komórkę. Geny do nadekspresji *SHs* namnożono na matrycy genomowego DNA dzikiego szczepu drożdży *Y. lipolytica* W29 korzystając z metody PCR przy użyciu polimerazy Phire DNA (Thermo Fisher Scientific, Waltham, USA) oraz zaprojektowanych starterów (mat uzup. Tabela S2

w **P.3**). Wyjątek stanowił gen *CNE1*, którego syntetyczną sekwencję zakupiono z firmy SYNGEN (Wrocław, Polska) ze względu na obecność licznych miejsc restrykcyjnych rozpoznawanych przez *Bam*HI, kluczowych dla klonowania w wektorach z serii JMP62. Sekwencja nukleotydowa została zoptymalizowana pod kątem użycia kodonu przez *Y. lipolytica*, natomiast sekwencja aminokwasowa kodowanego przez nie białka pozostała identyczna z kodowaną przez natywny gen.

Wszystkie zaprojektowane startery zawierały miejsca restrykcyjne *Bam*HI na końcu 5' oraz *Avr*II na końcu 3' umożliwiające wklonowanie zamplifikowanych fragmentów DNA do wektorów JMP62 (JMP62-**URA3ex**-pTEF lub JMP62-**LEU2ex**-pTEF) [127]. Wektor JMP62-**URA3ex**-pTEF wykorzystano do wklonowania sekwencji kodującej **scYFP**, natomiast wektor JMP62-**LEU2ex**-pTEF dla genów kodujących **SHs (Rycina 3)**.



Rycina 3. Schematyczne przedstawienie strategii klonowania do wektora JMP62 zastosowanej w tej pracy. Kolorami oznaczono elementy wektorów: region zeta (pomarańczowy), *LoxP* i *LoxR* (ciemnoniebieski), konstytutywny promotor *pTEF* (zielony), heterologiczne geny: *SCYFP* (A. żółty) lub *SH* (B. niebieski) wyposażone we fragmenty *loxP*, markery auksotroficzne *URA3ex* (A. czerwony) lub *LEU2ex* (B. jasnoniebieski). Bakteryjne miejsca replikacji (różowe) i gen oporności na kanamycynę *KanR* (fioletowe), zawarte w części „bakteryjnej” wektorów, odrzucone przed transformacją komórek drożdży przez trawienie restrykcyjne *NotI*.

Szczep *Y. lipolytica* Po1f (*Mata*, *leu2-270*, *ura3-302*, *xpr2-322*, *xpr-2*) zastosowano jako szczep rodzicielski do kotransformacji wektorem JMP62-URA3ex-**SCYFP** i JMP62-LEU2ex-**SH**. Natomiast, szczep *Y. lipolytica* Po1h (*ura⁻*, *leu⁺*) transformowany wyłącznie kasetą markera *URA3* zastosowano jako kontrolę negatywną do oceny tła fluorescencji (Po1h_Ura3+). Pusty wektor JMP62-LEU2ex-pTEF oraz wektor JMP62-URA3ex-**SCYFP** kotransformowano do szczepu macierzystego Po1f i użyto jako kontroli pozytywnej do oceny fluorescencji (jak opisano w sekcji 3.4.6.) oraz jako zewnętrzny kalibrator do względnej ilościowej oceny poziomu ekspresji genów (opisano w sekcji 3.5.2.).

3.2.4. Transformacja bakterii *E. coli* (JM109)

Elementy składowe kaset do nadekspresji klonowano najpierw w wektorze pCR Blunt II TOPO vector (Thermo Fisher Scientific, Waltham, MA USA) lub wektora JMP62 przy zastosowaniu ligazy DNA T4 i odpowiedniego buforu T4 DNA ligase buffer (NEB). Uzyskane konstrukcje wprowadzano do komórek kompetentnych bakterii *E. coli* JM109 zgodnie z metodą opisaną przez Russel i Sambrook [128], które następnie wysiewano na stałe podłoże LB z antybiotykiem i inkubowano w temperaturze 37 °C przez noc. Wyrosłe kolonie weryfikowano pod kątem obecności określonych fragmentów DNA metodą colony PCR przy

zastosowaniu starterów (mat. uzup. Tabela S3 w **P.1** i Tabela S2 w **P.3**) i polimerazy Phire DNA (Thermo Fisher Scientific) lub polimerazy RUN (A&A Biotechnology, Gdynia, Polska). Pozytywne klony przenoszono do płynnego podłoża LB z antybiotykiem i hodowano z wytrząsaniem w 37 °C przez noc, a następnie izolowano plazmidowy DNA przy użyciu komercyjnego zestawu Plasmid Mini Kit (A&A Biotechnology), według instrukcji producenta. Poprawność sklonowanych fragmentów DNA sprawdzano poprzez sekwencjonowanie w firmie zewnętrznej Genomed S.A. (Warszawa, Polska). Listę szczepów bakterii *E. coli* zawierających skonstruowane plazmidy przedstawiono w **Tabeli 2**.

3.2.5. Transformacja drożdży *Y. lipolytica* - pochodne Po1g, Po1h, i Po1f

Transformację komórek *Y. lipolytica* przeprowadzono zgodnie z klasyczną metodą wykorzystującą octan litu opisaną przez Barth i Gaillardin [129]. Transformacja obejmowała dwa główne etapy: i) przygotowanie komórek kompetentnych oraz ii) wprowadzenie przygotowanych konstrukcji kaset do nadekspresji do genomu drożdży. W skrócie, w pierwszym etapie, szczepy drożdży wysiewano na stałym podłożu YPD i inkubowano w temp. 28 °C od 16 do 24 h. Wyrosłą biomasę pobierano eż, zawieszano w 1 ml roztworu TE (1x) i całość wirowano. Po usunięciu supernatantu, komórki zawieszono w 600 µl roztworu octanu litu (0.1M LiAc pH 6.0) i umieszczono w łaźni wodnej na 1 h w temp. 28 °C. Po inkubacji, mieszaninę wirowano, supernatant usunięto, a komórki ponownie zawieszono w 60 µl roztworu octanu litu. W drugim etapie, 40 µl komórek kompetentnych połączono z mieszaniną transformacyjną zawierającą: 3-5 µl jednoniciowego DNA nośnikowego oraz 2-5 µl oczyszczonej kasety DNA do nadekspresji (linearyzowanej uprzednio poprzez RE *NotI*) i inkubowano przez 15 min w 28 °C w łaźni wodnej, bez mieszania. Następnie, do mieszaniny dodano 350 µl 40% glikolu polietylenowego 4000 (PEG 4000). Całość wymieszano przez ostrożne pipetowanie i ponownie inkubowano w 28 °C w łaźni wodnej przez 1 h. Po inkubacji wykonano szok cieplny umieszczając zawiesinę w 39 °C przez 10 min, po czym dodano do zawiesiny 600 µl roztworu LiAc. Tak przygotowaną zawiesinę komórek rozprowadzono za pomocą głaszczki na minimalnym stałym podłożu selekcyjnym YNB i inkubowano w temp. 28 °C przez 48 h. Uzyskane transformanty weryfikowano metodą colony PCR przy zastosowaniu starterów (mat. uzup. Tabela S3 w **P.1** i Tabela S2 w **P.3**) i polimerazy Phire DNA (Thermo Fisher Scientific) celem potwierdzenia obecności kaset nadekspresyjnych w genomie. W celu namnożenia biomasy, pozytywne transformanty przeniesiono za pomocą sterylnej eży do płynnej pożywki YPD/YNB i hodowano w 28 °C przez 24 h z wytrząsaniem

(250 rpm), po czym komórki zabezpieczono w postaci 15% stoków glicerolowych i zdeponowano w temp. – 80 °C w kolekcji Katedry. Wszystkie szczepy drożdży *Y. lipolytica* skonstruowane w tej pracy przedstawiono w **Tabeli 3**.

3.3. Metody hodowlane

3.3.1. Hodowle okresowe – screening szczepów *Y. lipolytica* (P.1, P.3)

Spośród wszystkich pozytywnych transformantów drożdży *Y. lipolytica* wytypowano trzy/cztery subklony każdego badanego wariantu, które następnie hodowano okresowo w kolbach o objętości całkowitej 50 ml i objętości roboczej 5 ml w podłożach YPG₂₀/ YNB w inkubatorze z wytrząsaniem (BIOSAN), przy 250 rpm, w temp. 28 °C / 30 °C do 72 h. Hodowle prowadzono w dwóch powtórzeniach biologicznych. Próbkę z hodowli okresowo pobierano, wirowano, a następnie odseparowaną biomasę i supernatant przechowywano w temp. – 20 °C do dalszych analiz. Następnie, rozmrożoną biomasę przemywano w sterylnym roztworze 0.85 % NaCl (POCH, Polska), ponownie zawieszano w w/w roztworze soli i odpowiednio rozcieńczano. Stężenie biomasy w próbkach oznaczano poprzez pomiar absorbancji przy długości fali 600 nm (OD₆₀₀) przy użyciu spektrofotometru (Tecan Infinite M200) (opisano w sekcji **3.4.1.**). Z kolei, próbki supernatantu analizowano odpowiednio do właściwości produkowanego r(s)-Prot (opisano w sekcji **3.4.5.** oraz **3.4.6.**).

3.3.2. Hodowle okresowe – ocena wpływu obniżonej temperatury na działanie poszczególnych SHs (P.3)

Hodowle okresowe prowadzono w kolbach wytrząsanych (BIOSAN) o objętości całkowitej 250 ml i objętości roboczej 30 ml na podłożu YPD₂₀. W skrócie, 4 ml pożywki YPD₂₀ zaszczerpiono pojedynczą kolonią reprezentatywnego szczepu (z płytki z pożywką stałą YPD) i hodowano w temperaturze 30 °C z wytrząsaniem (250 rpm). Po upływie 22 h, pobrano 3 ml prekultury, które następnie przenoszono do 30 ml pożywki YPD₂₀ i inkubowano w dwóch wariantach : i) w temp. 30 °C przez 96 h oraz ii) w 30 °C przez 24 h i w 25 °C przez kolejne 72 h. Próbkę o objętości 1 ml pobierano, wirowano a następnie analizowano w sposób opisany w **3.4.1.** i **3.4.6.** Dodatkowo, 24 h po obniżeniu temp. hodowli do 25 °C oraz w 48 h hodowli w 30 °C, próbki pobierano, izolowano RNA (**3.5.1.**) a następnie badano ekspresję genów kodujących SHs oraz r(s)-Prot w reakcji qPCR zgodnie z protokołem opisanym w punkcie **3.5.2.** Każdy wariant hodowli przeprowadzono w dwóch powtórzeniach biologicznych oraz dwóch powtórzeniach technicznych.

3.3.3. Hodowle ciągłe w warunkach chemostatu – globalna analiza transkryptomu (P.1, P.2, P.4)

Hodowle ciągłe szczepów drożdży *Y. lipolytica* prowadzono w podłożu YPG₁₀₀ w bioreaktorze mieszadłowym Multifors 2 (Infors, Bottmingen-Basel, Szwajcaria) o objętości całkowitej 0.5 L i objętości roboczej 0.3 L. W skrócie, wybrane szczepy drożdży, zabezpieczone uprzednio w stokach glicerynowych, hodowano przez noc na stałym podłożu YPD w temp. 30 °C. Następnie, prekulturę do hodowli bioreaktorowych przygotowywano w kolbach Erlenmeyera zawierających 30 ml podłoża YPG₂₀ zaszczipionych pojedynczą kolonią drożdży i hodowano w temp. 28 °C przez 22 h na wytrząsarce orbitalnej z szybkością obrotów 250 rpm (BIOSAN). Hodowle bioreaktorowe zaszczipiano prekulturą danego szczepu drożdży *Y. lipolytica*. Warunki hodowli automatycznie kontrolowano w następujących punktach: temperatura 28 °C, pO₂ utrzymywane na poziomie 20 %, stosując metodę kontroli kaskadowej z mieszaniem od 100 do 1200 obr./min i stałe napowietrzanie na poziomie 2.0 vvm. pH utrzymywano na poziomie 5.5 poprzez automatyczne dozowanie 20 % NaOH oraz 10 % H₂SO₄. Pienienie ograniczano poprzez dodanie środka przeciwpiennego AntiFoam 204 (Sigma Aldrich). W pierwszej kolejności hodowle bioreaktorowe prowadzono w trybie okresowym, do momentu całkowitego zużycia substratu - glicerolu, co charakteryzowało się wzrostem parametru pO₂. Następnie rozpoczęto hodowle ciągłe w chemostacie przy szybkości rozcieńczania równej 0.20 h⁻¹, wyznaczonej eksperymentalnie. Stan ustalony osiągnięto po co najmniej 6 pełnych wymianach podłoża, definiowanego jako stałe stężenie biomasy, utylizowanego substratu oraz produkowanych metabolitów. Próbkę pobierano okresowo, wirowano przy 12,045 x g przez 6 min, a następnie wykonywano pomiar fluorescencji (3.4.6.). Inne analizy, jak pomiar biomasy (3.4.1.), HPLC (3.4.4.), ocena aktywności amylolitycznej (3.4.5.) wykonywano na próbkach mrożonych w temp. – 20 °C. Próbkę do analiz transkryptomu pobierano po osiągnięciu stanu ustalonego i przechowywano w temp. – 80 °C do czasu izolacji RNA (3.5.1.). Hodowle prowadzono w dwóch powtórzeniach biologicznych.

3.4. Metody analityczne

3.4.1. Pomiar gęstości biomasy drożdżowej

Wzrost szczepów *Y. lipolytica* w trakcie hodowli monitorowano poprzez pomiar gęstości optycznej przy długości fali równej 600 nm (OD 600nm). Próbkę biomasy drożdży dwukrotnie przepłukiwano, zawieszano i rozcieńczano w sterylnym roztworze soli

fizjologicznej (0.85% NaCl). Odczyty przeprowadzono w płaskodennej mikropłytkce (Coring, Sigma-Aldrich) w spektrofotometrze Tecan Infinite M200 (Tecan Group Ltd., Mannedorf, Szwajcaria) w objętości 200 μ l. Pomiar dla każdego powtórzenia biologicznego wykonano w co najmniej dwóch powtórzeniach technicznych.

W trakcie ciągłych hodowli bioreaktorowych stężenie biomasy określano metodą grawimetryczną. Rozmrożoną biomasę dwukrotnie przemywano w sterylnej wodzie destylowanej a następnie suszono 2-3 dni w laboratoryjnej suszarce w temperaturze 105 °C, aż do momentu uzyskania stabilnych odczytów. Stężenie suchej masy wyrażano w gramach suchej masy na litr (DCW; [gDCW L⁻¹]). Wszystkie pomiary wykonano w dwóch powtórzeniach technicznych.

3.4.2. Pomiar żywotności komórek (P.1)

Żywotność szczepów *Y. lipolytica* w trakcie ciągłych hodowli bioreaktorowych monitorowano poprzez wykonanie serii posiewów powierzchniowych. Próbki biomasy rozcieńczano w szeregu rozcieńczeń dziesiętnych (do 10⁻¹⁰) w sterylnej soli fizjologicznej w mikropłytkce MTP (Coring, Sigma- Aldrich), a następnie wysiewano na stałe podłoże YPD przy użyciu replikatora (Sigma-Aldrich). Płytki inkubowano przez 24 h w temperaturze 30°C. Wyniki wyrażono, jako powierzchniowe jednostki tworzące kolonie [scfu] odzwierciedlające liczbę zliczonych kolonii przy najwyższym rozcieńczeniu dziesiętnym, gdzie zaobserwowano wzrost kolonii.

3.4.3. Morfologia komórek (P.2)

Morfologię wybarwionych fioletem krystalicznym komórek *Y. lipolytica* utrzymywanych w hodowlach bioreaktorowych w stanie ustalonym obserwowano w powiększeniu 1000x z użyciem mikroskopu świetlnego (Primo Star, ZEISS).

3.4.4. Pomiary stężenia substratu i metabolitów – wysokosprawna chromatografia cieczowa (HPLC)

Stężenie utylizowanego glicerolu (GLY; [g L⁻¹]) i produkowanych metabolitów w tym erytrytolu (ERY; [g L⁻¹]), mannitolu (MAN; [g L⁻¹]), kwasu cytrynowego (CA; [g L⁻¹]) oraz kwasu α -ketoglutazarowego (α -KG; [g L⁻¹]) mierzono w supernatantach pobranych próbek przy użyciu techniki wysokosprawnej chromatografii cieczowej (HPLC). Analizę

przeprowadzono przy użyciu chromatografu Agilent Technologies 1200 series (Agilent Technologies, Santa Clara, CA, USA) wyposażonego w detektor (G1362A) i kolumnę Rezex ROA-Organic Acid H⁺ (Phenomenex, Torrance, Kalifornia, USA) przy zastosowaniu następujących warunków: 0.005 N H₂SO₄ (faza ruchoma), temp. 40 °C, natężenie przepływu eluentu: 0.6 [ml min⁻¹]. Przed przystąpieniem do analizy, rozmrożone na lodzie próbki supernatantu odpowiednio rozcieńczano i klarowano przy użyciu filtrów strzykawkowych z membraną 0.45 µm (Millipore; Merck-Millipore, Burlington, MA, USA).

3.4.5. Analiza aktywności enzymatycznej – microSIT i microDNS (P.1, P.2)

Zewnątrzkomórkową aktywność amylolityczną obu białek reporterowych: α-amylazy (SoA) oraz glukoamylazy (TIG) analizowano przy użyciu kolorymetrycznych testów w mikroskali. Aktywność SoA mierzono w próbkach supernatantu przy zastosowaniu testu microSIT zgodnie z uprzednio opracowanym protokołem [130]. Jednostkę aktywności [AAU] zdefiniowano, jako taką ilość enzymu, która powoduje obniżenie wartości wybarwienia skrobi przez jod, równoważnej 1 mg skrobi w 1 ml, w ciągu 1 min przy pH 5.0 i 40 °C, w przyjętych warunkach doświadczalnych. Natomiast, aktywność TIG mierzono w próbkach supernatantu przy zastosowaniu testu microDNS (3,5-kwasu dinitrosalicylowego) zgodnie z protokołem [131]. Jednostkę aktywności [GAU] zdefiniowano, jako ilość enzymu, która uwalnia 1 µg cukrów redukujących w 1 ml, w ciągu 1 min przy pH 5.0 i 40 °C, w przyjętych warunkach doświadczalnych. W obu testach enzymatycznych analizowano również supernatanty pochodzące z hodowli szczepów referencyjnych Po1h_Ura3+/Po1g_Leu2+ w celu oceny tła reakcji. Wszystkie odczyty znormalizowano wg próby ślepej z wodą destylowaną. Odczyty wykonano w płaskodennych mikroplótkach (Coring, Sigma Aldrich) przy użyciu spektrofotometru Tecan Infinite M200 przy długości fali 580 nm (microSIT) i 540 nm (microDNS). Każdą próbkę analizowano w co najmniej dwóch powtórzeniach technicznych.

3.4.6. Pomiar intensywności fluorescencji

Pomiary fluorescencji (FL) w próbkach (biomasy i supernatantu), pobieranych w trakcie hodowli szczepów *inYFP* oraz *scYFP*, wykonywano w płaskodennych mikroplótkach MTP (Coring; Sigma-Aldrich) w automatycznym czytniku Tecan Infinite M200 przy długości fali (wzbudzenia/emisji) 495/527 nm. Zewnątrzkomórkową FL (exYFP) analizowano w 200 µl supernatantu, natomiast wewnątrzkomórkową FL (inYFP) w 200 µl

przemytej biomasy, ponownie zawieszanej i odpowiednio rozcieńczonej w roztworze soli fizjologicznej lub wody (0.85 NaCl/ddH₂O). Każdy pomiar FL normalizowano względem biomasy szczepu kontroli negatywnej (Po1h_Ura3+/Po1g_Leu2+) dla odczytów inYFP lub świeżej pożywki dla exYFP. Wyniki FL wyrażano jako: i) względną jednostkę fluorescencji [RFU] (*Relative Fluorescence Unit*) zdefiniowaną jako średnią wartość FL próbki minus wartość FL tła (świeża pożywka/biomasa kontroli negatywnej), ii) inYFP – surowy odczyt FL dla próbek biomasy w [FU], iii) sp_inYFP – specyficzna fluorescencja biomasy w [FU OD600nm⁻¹], iv) exYFP – surowy odczyt FL supernatantu w [FU], v) sp_exYFP – specyficzna fluorescencja supernatantu w [FU OD600nm⁻¹]. Celem porównania szczepów *scYFP+* niektóre dane zostały dodatkowo wyrażone jako wartości zmiany krotności (FC, *Fold Change*) obliczone jako ułamek dwóch określonych wartości FL wyrażonych w tych samych jednostkach. Wszystkie pomiary wykonano w dwóch powtórzeniach biologicznych, analizowanych w co najmniej dwóch powtórzeniach technicznych. Dodatkowo, wyniki losowo weryfikowano poprzez obserwacje pod mikroskopem fluorescencyjnym (ZEISS AxioVert, AxioCam 350 color; filterset: 09; ZEISS).

3.4.7. Izolacja białek i ocena aktywności proteolitycznej (P.4)

3.4.7.1. Ekstrakcja białka całkowitego i oznaczenie stężenia metodą BCA

Ekstrakcję białka całkowitego przeprowadzono zgodnie z wcześniej opisanym protokołem [66]. W skrócie, biomasę drożdży *Y. lipolytica* z 23 h hodowli zawieszono w lodowatym buforze do rozbijania komórek (*breaking buffer*). Komórki rozbijano przy użyciu szklanych kulek (Sigma-Aldrich, St. Louis, USA) w powtarzanych cyklach mieszania (5x) przy uderzeniach 30 s⁻¹ przez 30 s w MixerMill MM400 (Retsch GmbH, Haan, Niemcy) i inkubowano na lodzie przez 1 min. Ekstrakty wirowano (24,652 ×g, 4 °C, 10 min; Eppendorf 5430 R; Eppendorf, Hamburg, Niemcy), odrzucano osad, a supernatant zachowywano do dalszych analiz. Ekstrakty białkowe zawieszono w buforze do wymiany (*exchange buffer*), a następnie przeprowadzono ultrafiltrację przez membranę celulozową o punkcie odcięcia 3 kDa (Amicon Ultra; Merck, Darmstadt, Niemcy).

Całkowite stężenie białka w otrzymanych ekstraktach określono przy użyciu komercyjnego zestawu do oznaczania białek BCA (kwas bicynchoniowy) (Merck) na płycie mikrotitracyjnej. Krzywą wzorcową sporządzono przy użyciu roztworu wzorcowego BSA. Reakcję prowadzono zgodnie z protokołem producenta. Odczyt absorbancji przeprowadzono w przezroczystych płaskodennych 96-dółkowych mikropłytkach (Corning; Sigma-Aldrich)

przy użyciu czytnika Multiskan SkyHigh Plate Reader - UV/Vis Spectrophotometer przy długości fali 562 nm (ThermoFisher Scientific, Waltham, USA). Test przeprowadzono w dwóch niezależnych powtórzeniach biologicznych i w trzech powtórzeniach technicznych.

3.4.7.2. Oznaczenie aktywności proteolitycznej

Aktywność proteolityczną w całkowitych ekstraktach białkowych badano przy użyciu dostępnego komercyjnie zestawu Pierce™ Fluorescent Protease Assay Kit (ThermoFisher Scientific), wykorzystującego fluorescencyjnie znakowaną kazeinę jako substrat. Ekstrakty przygotowane jak w punkcie 3.4.7.1 poddano reakcji na płytkach do mikromiareczkowania. Krzywą wzorcową sporządzono stosując standaryzowany roztwór trypsyny. Odczyt fluorescencji przeprowadzono na białych nieprzezroczystych płytkach, zgodnie z zaleceniami producenta zestawu, przy użyciu automatycznego czytnika płytek Tecan Infinite M200 przy ustawieniach długości fali (wzbudzenie/emisja) 485/538 nm. Test przeprowadzono w dwóch niezależnych seriach i w trzech powtórzeniach technicznych.

3.5. Oznaczenie poziomu ekspresji genów

3.5.1. Izolacja RNA

Całkowity RNA do analizy ekspresji genów ekstrahowano z próbek biomasy pobieranych w trakcie hodowli. Biomase otrzymano z 1 ml hodowli okresowej (24 h po obniżeniu/podwyższeniu temperatury hodowli o 5 °C) oraz z 2 ml hodowli ciągłej w momencie osiągnięcia tzw. stanu ustalonego. Odseparowaną biomase poddano mechanicznej homogenizacji przy pomocy homogenizatora Mixer Mill MM400 (Retsch, Niemcy) i szklanych kulek. Izolację RNA przeprowadzono przy użyciu zestawu Bead-Beat Total RNA Mini Kit (A&A Biotechnology) zgodnie z instrukcją producenta. Tak otrzymany materiał totalnego RNA weryfikowano ilościowo (stężenie, czystość) przy użyciu spektrofotometru (NanoDrop, Thermo Fisher Scientific) oraz jakościowo poprzez elektroforezę w żelu agarozowym [128].

3.5.2. Odwrotna transkrypcja oraz PCR w czasie rzeczywistym (qPCR; *quantitative Polymerase Chain Reaction*) (P.1, P.3)

Syntezę komplementarnego cDNA na matrycy całkowitego RNA prowadzono przy użyciu transkryptazy SuperScript III Reverse Transcriptase oraz startera oligo(dT) zgodnie z instrukcją producenta (Thermo Fisher Scientific). Następnie, otrzymane próbki cDNA

wykorzystano jako matryce w reakcji PCR w czasie rzeczywistym (qPCR) prowadzonej w termocyklerze Applied Biosystems 7500 (Applied Biosystems, Foster City, USA) zgodnie z wyznaczonym profilem termicznym (95 °C 3 min, (95 °C 15 sek., 60 °C 30 sek., 72 °C 30 sek.) x 40, 72 °C 1 min, krzywa topnienia : 94 °C 15 sek., 60 °C 60 sek., 95 °C 30 sek., 60 °C 15 sek.). Reakcje qPCR nastawiano z wykorzystaniem zestawu RT HS-PCR Mix SYBR®B (A&A Biotechnology) zgodnie z instrukcją producenta, oraz starterów zaprojektowanych w programie Primer Expert Software (Applied Biosystems) (mat. uzup. Tabela S3 w **P.1** i Tabela S2 w **P.3**). Poziom ekspresji genów normalizowano względem poziomu ekspresji genu kodującego aktynę (*ACT1/D08272g*) lub białko translokonu (*SEC62/B17512g*) użytych, jako wewnętrzne kalibratory. Natomiast, preparaty cDNA szczepów kontrolnych (Po1h_Ura3+ /Po1g_Leu2+/ Control_scYFP) wykorzystano, jako kalibratory zewnętrzne, którym przypisano poziom ekspresji równy 1.0. Poziom ekspresji badanych genów analizowano w dwóch powtórzeniach biologicznych oraz w trzech powtórzeniach technicznych. Analizę otrzymanych danych przeprowadzono zgodnie z metodą delta-delta Ct ($\Delta\Delta Ct$) [132].

3.5.3. Sekwencjonowanie RNA - RNAseq (P.2, P.4)

Wyizolowany materiał całkowitego RNA pochodzący z drożdży *Y. lipolytica* hodowanych w warunkach chemostatu wykorzystano jako matrycę w reakcji sekwencjonowania nowej generacji (NGS). Przygotowanie biblioteki oraz sekwencjonowanie wykonano w zewnętrznej firmie Genomed S.A. (Warszawa, Polska) przy pomocy sekwenatora HiSeq 4000 oraz NovaSeq 6000 (Illumina) w trybie paired-end. Surowe wyniki otrzymane z sekwencjonowania RNA (RNAseq) zdeponowano w bazie NCBI SRA pod numerem akcesyjnym PRJNA701856 (**P.2**) oraz PRJNA869113 (**P.4**).

3.6. Globalna analiza transkryptomu (P.2, P.4)

3.6.1. Przetwarzanie i filtrowanie surowych danych pochodzących z RNAseq

Adaptory NGS usunięto za pomocą programu Cutadapt [133]. Początkowe filtrowanie wykonano przy użyciu parametru jakości q25 i m15 minimalnej długości odczytów, po czym przeprowadzono ponowne filtrowanie w celu odrzucenia odczytów < 20 bp. Kontrolę jakości przeprowadzono za pomocą programu FastQC [36]. Dane po kontroli jakości dopasowano i zmapowano do genomu referencyjnego *Y. lipolytica* CLIB122 (identyfikator GenBank: GCF_000002525.2) przy użyciu oprogramowania Hisat2 [134]. Liczbę par odczytów

zmapowanych do pojedynczego genu, zliczono za pomocą programu HTseq [135]. Adnotacje terminów Gene Ontology (GO) zostały przypisane przy użyciu referencyjnego szablonu genomu (GCF_000002525.2). Analizę porównawczą ekspresji genów przeprowadzono przy użyciu pakietu DESeq2 [136]. Aby porównać ekspresję różnych genów w próbkach, nieprzetworzone liczby odczytów przekształcono w znormalizowane liczby RNAseq. Alternatywnie, względny poziom ekspresji genów wyrażano jako krotność zmiany (FC) w odniesieniu do prototroficznego szczepu kontrolnego (Po1h_Ura3+). Wszystkie geny o zróżnicowanej ekspresji (DEG; *Differentially Expressed Gene*) spełniały wymaganie $FDR \leq 0.05$ (skorygowana wartość p) testowane za pomocą testu Walda w DESeq2. Tylko DEG spełniające wszystkie te kryteria uznano za reagujące na wprowadzoną zmienną (nadprodukcję określonego białka heterologicznego) i poddano dalszym analizom.

3.6.2. Analiza porównawcza - wykresy Venn'a, heatmapy, test nadprezentacji

Internetowe narzędzie bioinformatyczne Morpheus (<https://software.broadinstitute.org/morpheus>) zastosowano w celu przygotowania wykresów typu „heatmap” przedstawiających porównanie poziomu ekspresji konkretnego genu w badanych szczepach *Y. lipolytica*. Dane zaprezentowane na heatmapach wyrażono jako FC obliczony dla: i) względnej wartości ekspresji poszczególnych DEGs, lub ii) wartości sp_exYFP .

Podobieństwa i różnice w podzbiorach DEGs określano za pomocą internetowego narzędzia InteractiVenn [137]; natomiast grafikę przygotowano przy użyciu programu Inkscape (<https://inkscape.org/release/inkscape-1.0.2/>) zgodnie z instrukcją InteractiVenn. Przed analizą diagramu Venna, wartości DEG przewidywane przez DESeq2 filtrowano pod kątem wartości krotności zmiany \log_2 ilości transkryptu dla próby badanej względem próby kontrolnej ($\log_2_Estimated_FC$) > 2.0 (**P.2**) lub 1.0 (**P.4**); tylko takie DEGs uwzględniono w tej analizie.

Przypisania nazw, funkcji molekularnych, oraz procesów biologicznych dla poszczególnych genów dokonano ręcznie poprzez zestawienie informacji wygenerowanych z kilku dostępnych baz danych (opisane w 3.6.4). Wszystkie zidentyfikowane zestawy DEGs, wspólne lub specyficzne dla danego szczepu, poddano dalszemu testowaniu poprzez przeprowadzenie testu nadprezentacji statystycznej lub wzbogacenia przy użyciu internetowego narzędzia i bazy danych PANTHER v14.0 [138]. Nadprezentację określonych procesów biologicznych, funkcji molekularnych i lokalizacji komórkowych

w podzbiorze DEG testowano przy użyciu kompletnych zestawów danych adnotacji GO, testu Fishera i wskaźnika FDR (*False Discovery Rate*) dla testów wielokrotnych. Wyniki uznano za istotne statystycznie przy wartości $p < 0.05$.

3.6.3. Identyfikacja miejsc splicingowych w regionie genu *HAC1* (*B12716g*) (P.4)

Identyfikację miejsc splicingowych w rejonie genu *HAC1* oraz oznaczenie ilościowe stosunku transkryptu *HAC1* bez intronu do transkryptu *HAC1* z intronem wykonano w firmie zewnętrznej Genomed S.A. (Warszawa, Polska). Z uwagi na fakt, iż obecność intronu nie została opisana w zastosowanym genomie referencyjnym, istnienie i lokalizacja intronu w locus *B12716g* musiały zostać zidentyfikowane *de novo*. Detekcję i wizualizację potencjalnych miejsc splicingowych przeprowadzono przy użyciu pakietu SGSeq [139] w języku programowania R. Mapowanie eksonów i intronów w genie *B12716g* przeprowadzono w programie IGV [140]. Pokrycie w obrębie fragmentów genów (intron i egzony) wykorzystano jako wskaźnik rzeczywistych zdarzeń splicingowych. Pokrycie obliczono we fragmentach: intron (NC_006068.1:1702623-1702651), egzony (NC_006068.1:1702612-1702622, NC_006068.1:1702652-1702662). Średnią wartość pokrycia obliczono przy użyciu narzędzia SamTools [141].

3.6.4. Narzędzia bioinformatyczne i bazy danych

W celu poszukiwania nazw i funkcji na temat poszczególnych DEGs sekwencje nukleotydowe oraz białkowe analizowano przy użyciu dostępnych baz danych, w tym: NCBI (<https://www.ncbi.nlm.nih.gov/>), GRYC (<http://gryc.inra.fr/>), SGD (<https://www.yeastgenome.org/>), Panther (<http://www.pantherdb.org/>) czy UniProt (<https://www.uniprot.org/>). Dodatkowo, poszukiwanie homologów przeprowadzono za pomocą narzędzia blastp z ustawieniami domyślnymi. Wybrano najbardziej odpowiednie dopasowania, aby ostatecznie opisać dany gen.

Analizę sekwencji wykonywano przy użyciu narzędzia Benchling (<https://www.benchling.com/>). Do porównania sekwencji zastosowano programy BioEdit [142] i MultAlign [143].

Podstawowe właściwości biochemiczne analizowanych r-Prots określono za pomocą narzędzia ProtParam Tool z serwisu Expasy (<https://web.expasy.org/protparam/>). Przewidywanie miejsc N- i O- glikozylacji oraz kombinacji tworzenia wiązań disiarczkowych

przeprowadzono przy użyciu narzędzi NetNGlyc i NetOGlyc (<http://www.cbs.dtu.dk/services>) i NPSA (<https://npsa-prabi.ibcp.fr>).

Za pomocą narzędzi graficznych programu Microsoft Excel zaprezentowano wykresy liniowe ilustrujące kinetykę wzrostu i produkcji r(s)-Prots, metabolitów w hodowlach drożdży *Y. lipolytica* oraz przedstawiono wykresy słupkowe ilustrujące wyniki dotyczące ekspresji wybranych genów.

3.7. Obliczenia

3.7.1. Stężenie białek reporterowych (P.1)

Bezwzględne stężenie białek enzymatycznych obliczono na podstawie ich specyficznych wartości aktywności, równych 80 [GAU mg⁻¹] dla TIG [144] i 478 [AAU mg⁻¹] dla SoA [145]. Natomiast, bezwzględne stężenie reportera fluorescencyjnego eYFP obliczono na podstawie krzywej kalibracyjnej dla eGFP, dostarczonej przez (BioTek Instruments, Inc., Winooski, VT, USA). W celu ponownego obliczenia wykorzystano następujące dane do oszacowania względnej fluorescencji eYFP względem eGFP: $\epsilon_{\text{eYFP}} = 83\,400 \text{ [cm}^{-1} \text{ M}^{-1}]$, $\Phi_{\text{f eYFP}} = 0.61$, $\epsilon_{\text{eGFP}} = 55\,000 \text{ [cm}^{-1} \text{ M}^{-1}]$, $\Phi_{\text{f eGFP}} = 0.60$ [146].

3.7.2. Uproszczony bilans masowy dystrybucji C i N między produktami metabolizmu *Y. lipolytica* (P.1)

Rozkład C i N obliczono na podstawie masy cząsteczkowej i składu elementarnego głównych produktów syntetyzowanych przez szczepy *Y. lipolytica* w hodowlach bioreaktorowych w warunkach chemostatu. Do obliczenia zawartości C i N w docelowym białku zastosowano następujące wzory i wartości mas cząsteczkowych: TIG: C₂₉₁₀H₄₃₆₅N₇₈₁O₉₀₉S₁₀, i 65,154.96 [Da]; SoA: C₂₂₉₉H₃₄₃₃N₆₂₉O₇₃₀S₁₈, i 52,140.40 [Da]; eYFP: C₁₂₁₅H₁₈₆₃N₃₁₇O₃₆₄S₈, i 26,991.54 [Da]. Powyższe dane pobrano z serwera ExPASy na podstawie wyników analizy sekwencji białek reporterowych z wykorzystaniem narzędzia ProtParam [147]. Skład elementarny biomasy *Y. lipolytica* został określony wcześniej [148]. Sumę C i N zawartą we wszystkich produktach uwzględnionych w tej analizie zdefiniowano jako 100 %, a frakcję każdego konkretnego produktu obliczono na podstawie wyznaczonego stężenia DCW i metabolitów, a także ilości docelowego białka.

4. Wyniki

4.1. Charakterystyka szczepów *Y. lipolytica* z nadekspresją r(s)-Prots w warunkach chemostatu

4.1.1. Klonowanie r(s)-Prots i selekcja szczepów *Y. lipolytica* (P.1).

W badaniach dotyczących produkcji r(s)-Prots (scSoA, scTIG, scYFP i inYFP) w ciągłych hodowlach bioreaktorowych wykorzystano szczepy *Y. lipolytica* skonstruowane w podłożu genetycznym Po1g oraz Po1h przedstawione w **Tabeli 3**. Strategię klonowania r(s)-Prots przy zastosowaniu komercyjnego systemu klonowania YLEX (pochodne Po1g) lub modularnego klonowania Golden Gate (pochodne Po1h) zaprezentowano na **Rycinie 2A.B**. Pozytywne subklony otrzymane po transformacji szczepów rodzicielskich hodowano w kolbach z wytrząsaniem (**3.3.1**) w celu oceny zmienności międzyklonalnej i identyfikacji potencjalnego zaburzenia wzrostu spowodowanego wprowadzoną modyfikacją genetyczną. Jak pokazano w mat. uzup. Fig. S1 w **P.1.**, w większości analizowanych przypadków nie zaobserwowano zmienności we wzroście między subklonami a szczepami referencyjnymi (Po1h_Ura3/Po1g_LEU2). Do dalszych badań wytypowano subklon szczepu rekombinowanego reprezentujący określony wariant (rodzaj produkowanego r(s)-Prot) wykazujący wzrost identyczny/zbliżony do odpowiedniego szczepu kontrolnego.

4.1.2. Wpływ nadprodukcji r(s)-Prots na przebieg ciągłych hodowli drożdży *Y. lipolytica* (P.1)

Aby uzyskać wgląd w fizjologiczną i transkrypcyjną odpowiedź rekombinowanych komórek *Y. lipolytica* nadprodukujących r(s)-Prots przeprowadzono serię hodowli ciągłych w warunkach chemostatu. W pierwszej kolejności eksperymentalnie wyznaczono odpowiedni parametr szybkości rozcieńczania [D ; h^{-1}] wykonując hodowle ciągłe kontrolnego szczepu Po1h_Ura3 przy trzech wartościach D : 0.06, 0.12 i 0.20 [h^{-1}]. Jako główne wyznaczniki stanu stacjonarnego wykorzystano stężenie biomasy DCW [$gDCW L^{-1}$] oraz stężenie utylizowanego substratu GLY [$g L^{-1}$]. Wyniki tego eksperymentu zaprezentowano w mat. uzup. Fig. S2 w **P.1**. W trakcie hodowli przy D równym 0.20 [h^{-1}] osiągnięto stabilne odczyty stężenia GLY: 88.71 do 91.05 [$g L^{-1}$] i DCW: 8.15 do 10.35 [$gDCW L^{-1}$], a stężenie rozpuszczonego tlenu pO_2 [%] wyniosło blisko 20 [%] bez wyraźnych zmian w szybkości mieszania przy zastosowaniu kaskady, tym samym umożliwiając osiągnięcie stanu ustalonego i jednorodności populacji komórek pod względem fazy wzrostu. Stąd, wartość D równą 0.20

[h⁻¹] wykorzystano w ciągłych hodowlach bioreaktorowych w chemostacie wybranych reprezentatywnych subklonów *Y. lipolytica* oraz szczepów kontrolnych.

Kinetykę przedstawiającą akumulację biomasy, utylizację GLY oraz produkcję metabolitów (erytrytolu (ERY), mannitolu (MAN), kwasu cytrynowego (CA), kwasu α -ketoglutazarowego (a-KG)) w trakcie hodowli bioreaktorowych zaprezentowano w mat. uzup. Fig. S3. w **P.1**.

Z zaprezentowanych danych (Tabela 1; Fig.2; Fig. S3 w **P.1**) wynika, iż po osiągnięciu stanu ustalonego stężenie GLY utrzymywało się na stałym poziomie ok. 90 [g L⁻¹], tym samym zapewniając odpowiednią podaż źródła węgla do intensywnej produkcji r(s)-Prots (Fig. 2a. w **P.1**). Podczas gdy różnice w utylizacji GLY pomiędzy hodowlami były niewielkie, istotnie statystycznie wyższe stężenie GLY zaobserwowano dla pochodnych Po1g i Po1h produkujących scTIG oraz szczepu Po1g produkującego inYFP ($p < 0.05$), którym towarzyszyły 2.82-, 2.71- i 1.96-krotnie niższe wskaźniki zużycia tego substratu (Q_{GLY} [g L⁻¹ h⁻¹]) w porównaniu do szczepu kontrolnego (Tabela 1 w **P.1**). Dodatkowo, pochodne szczepu Po1g produkujące scTIG, inYFP oraz scSoA charakteryzowały się 3.09-, 2.00- i 1.70-krotnie niższą właściwą szybkością utylizacji GLY (q_{GLY} [g g_{DCW}⁻¹ h⁻¹]) względem szczepu kontrolnego ($p < 0.05$). Zgodnie z uzyskanymi danymi, pochodne Po1g produkujące r(s)-Prots wyróżniały się niższym zużyciem GLY ($Y_{S/X} < 1.53$ [g g_{DCW}⁻¹]) oraz wskaźnikiem właściwej szybkości utylizacji GLY ($q_{GLY} < 0.20 \pm 0.03$ [g g_{DCW}⁻¹ h⁻¹]) w porównaniu do pochodnych Po1h (z wyjątkiem szczepu Po1h_scYFP). Ponadto, wyższe wartości stężenia biomasy [g_{DCW} L⁻¹] w stanie ustalonym otrzymano dla pochodnych szczepu Po1g syntetyzujących rs-Prots (scSoA i scYFP) (Tabela 1, Fig. 2d w **P.1**; $p < 0.05$). Istotnie niższe wartości stężenia biomasy oraz produktywności biomasy (Q_{DCW} [g L⁻¹ h⁻¹]) zaobserwowano dla szczepów Po1h_scTIG (1.93-, 2.18-krotnie niższe) i Po1h_inYFP (1.55- i 1.97-krotnie) w odniesieniu do kontroli (Tabela 2 w **P.1**; $p < 0.05$). W kwestii produkcji polioli (ERY i MAN) wytwarzanych przez szczepy *Y. lipolytica*, stężenia bezwzględne pozostawały na niskim poziomie (< 1.30 [g L⁻¹]; Fig. 2b,c w **P.1**). Natomiast, niektóre szczepy istotnie wyróżniały się pod względem produkcji polioli, w tym szczep Po1g_scSoA z najwyższym stężeniem wyprodukowanego ERY, oraz Po1h_scTIG, Po1h_inYFP i Po1g_LEU2 z najniższym stężeniem ERY. Jak zaobserwowano, stężenie ERY były skorelowane z uzyskanymi wartościami DCW, co odpowiada specyficznej szybkości produkcji ERY (najwyższa dla Po1h_scTIG (0.033 ± 0.002 [g g_{DCW}⁻¹ h⁻¹]), a najniższa dla Po1g_LEU2 (0.015 ± 0.003 [g g_{DCW}⁻¹ h⁻¹])); pozostałe szczepy sklasyfikowano w tej samej grupie jednorodnej

w teście wielokrotnych porównań Tukey'a przy $p < 0.05$. Z kolei, synteza mannitolu (MAN) była istotnie różna dla dwóch rekombinowanych szczepów produkujących scTIG. Najwyższe stężenie MAN uzyskano dla szczepu Po1h_scTIG (0.99 ± 0.04 [g L⁻¹]), a najniższe dla Po1g_scTIG (0.09 ± 0.08 [g L⁻¹]). Ponadto, szczep Po1g_scTIG wykazał istotnie różny stosunek powstawania ERY do MAN sięgający 12-krotnie wyższą produkcję pierwszego polioliu ($p < 0.05$).

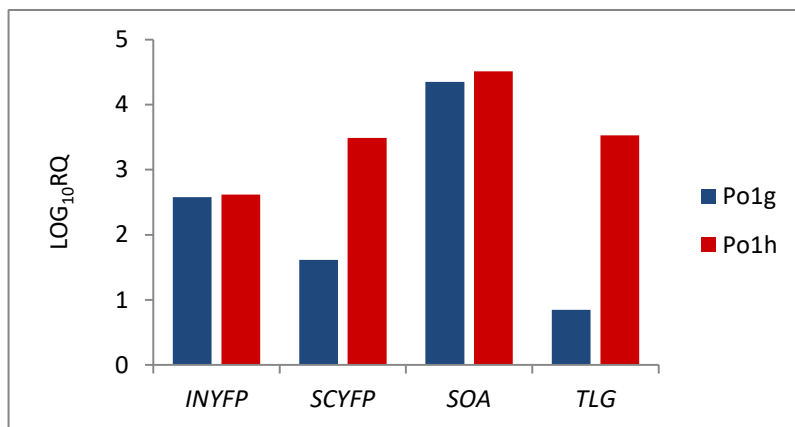
Produkcję białek reporterowych podczas stanu ustalonego wyrażono jako całkowitą jednostkę aktywności enzymatycznej AU (*Activity Unit*) lub fluorescencji FU (*Fluorescence Unit*) na objętość hodowli [AAU/GAU/kRFU L⁻¹] (Fig. 2e – h; Tabela 1 w **P.1**). Pomiar aktywności enzymatycznej scTIG w supernatancie były 3.52-krotnie wyższe [GAU L⁻¹] dla pochodnych szczepów Po1h niż dla pochodnych Po1g, natomiast SoA [AAU L⁻¹] były porównywalne, z niewielką przewagą dla pochodnych Po1h (Fig.2.e.f w **P.1**). Nie wykryto AAU lub GAU w próbkach biomasy (wykonane dla losowo wybranych próbek). Co ciekawe, **bardzo wysoki poziom fluorescencji (FL) wewnątrzkomórkowej zaobserwowano dla szczepu Po1h_scYFP, wytwarzającego YFP wyposażony w funkcjonalny SP1, który był 5-krotnie wyższy niż dla Po1h_inYFP i 10-krotnie większy niż dla Po1g_inYFP.** Niemniej jednak **zewnętrzkomórkowa FL dla tego szczepu była również najwyższa** spośród analizowanych szczepów w tym badaniu; dając ponad 2.5-krotnie wyższe wartości RFU w porównaniu do szczepu Po1g_scYFP. W przypadku tego ostatniego szczepu również wewnątrzkomórkowej FL była stosunkowo niska (Fig.2.g.h w **P.1**). Biorąc pod uwagę pomiary specyficznej AU oraz FU [(kRFU/AAU/GAU) gDCW⁻¹], pochodne Po1h wykazywały przewagę nad Po1g (1.51-krotnie [AAU gDCW⁻¹], 2.62- i 4.59-krotnie [RFU gDCW⁻¹] dla inYFP i scYFP, oraz 6.69-krotnie [GAU gDCW⁻¹] wyższe) (Tabela 1 w **P.1**). Podobnych obserwacji dokonano na podstawie parametrów produktywności objętościowej [(kRFU/AAU/GAU) L⁻¹ h⁻¹] i produktywności specyficznej [(kRFU/AAU/GAU) gDCW⁻¹ h⁻¹], z wyjątkiem produktywności objętościowej scSoA, która była porównywalna dla pochodnych szczepów Po1g i Po1h (Tabela 1 w **P.1**).

W celu bezpośredniego porównania produkcji różnych r(s)-Prots w tym badaniu, otrzymane wartości „aktywności” enzymatycznej AU oraz fluorescencyjnej FU znormalizowano wyrażając je jako ilość wytworzonego białka po stosownych przeliczeniach (opisano w **3.7.1**). Wartości stężenia białek reporterowych wyrażono w gramach oraz molach (Tabela 2 w **P.2**). Szczep Po1h_scYFP produkował istotnie więcej reportera w porównaniu do innych szczepów *Y. lipolytica* produkujących białko sekrecyjne ($p < 0.05$). W oparciu

o ilości białka wyrażone w [pmol] lub [pmol h⁻¹] szczególnie wyraźne różnice stwierdzono pomiędzy wytworzonymi ilościami małych białek (inYFP i scYFP) a większych białek sekrecyjnych (scSoA i scTIG) ($p < 0.05$). Niezależnie od przyjętej jednostki ([$\mu\text{g L}^{-1}$], [pmol], [$\mu\text{g gDCW}^{-1}$], [pmol g⁻¹ h⁻¹] itd.), szczep Po1h_inYFP produkował istotnie większą ilość białka reporterowego, w porównaniu do Po1g_inYFP, co potwierdza także % zutylicowanego węgla (C) i azotu (N) kierowanego do określonych produktów syntetyzowanych przez *Y. lipolytica* (Fig.3 w **P.1**). Z kolei, szczepy Po1h_scSoA i Po1g_scSoA, sklasyfikowano jako relatywnie mniej wydajnych producentów rs-Protów. Należy jednak pamiętać, że zastosowana metoda przeliczeń (aktywność jako wyznacznik ilości, a dalej – aktywność specyficzna jako przelicznik) jest obciążona niepewnością. W każdym parametrze i dla wszystkich badanych białek reporterowych, **poходne Po1h okazały się być lepszymi gospodarzami dla produkcji r(s)-Protów w porównaniu z Po1g** ($p < 0.05$).

4.1.3. Analiza ekspresji heterologicznych genów podczas wzrostu drożdży *Y. lipolytica* w warunkach chemostatu (P.1)

Aby dokładniej zrozumieć, czy różnice w ilości wyprodukowanych białek reporterowych wynikają z różnic w wydajności transkrypcji, analizowano poziom ekspresji genów kodujących r(s)-Protów w stanie ustalonym w reakcji qPCR w czasie rzeczywistym (**Rycina 4**). Jak zaobserwowano, **poziom ekspresji genu SOA był najwyższy**, niezależnie od zastosowanego szczepu gospodarza ($> 4.0 \log_{10}\text{RQ}$). **Nie zaobserwowano istotnego wpływu zastosowanego promotora** (4UAS-pTEF dla pochodnych szczepów Po1h lub hp4d dla pochodnych szczepów Po1g) ($p < 0.05$), chociaż ekspresja odpowiedniego genu była za każdym razem nieco wyższa dla pochodnych szczepów Po1h (4UAS-pTEF). Znacznie wyższą ekspresję z promotora 4UAS-pTEF zaobserwowano dla genów *TIG* i *SCYFP*, podczas gdy porównywalny poziom ekspresji z obu promotorów zaobserwowano dla *SOA* i *INYFP*, co odpowiada wartościom bezwzględny (Fig.2. e–h w **P.1**) i obliczonym ilościom białek (Tabela 2 w **P.1**).



Rycina 4. Poziom ekspresji genów kodujących rs-Prots analizowany metodą qPCR. Próbkę do analizy ekspresji genów pobrano po osiągnięciu stanu ustalonego przez komórki *Y. lipolytica* w ciągłych hodowlach w warunkach chemostatu. Oś X: geny reporterowe (*INYFP*, *SCYFP*, *SOA*, *TLG*); oś Y: LOG₁₀RQ (zlogarytmowana wartość względnego poziomu ekspresji). Względny poziom ekspresji (RQ) określany metodą $\Delta\Delta Ct$.

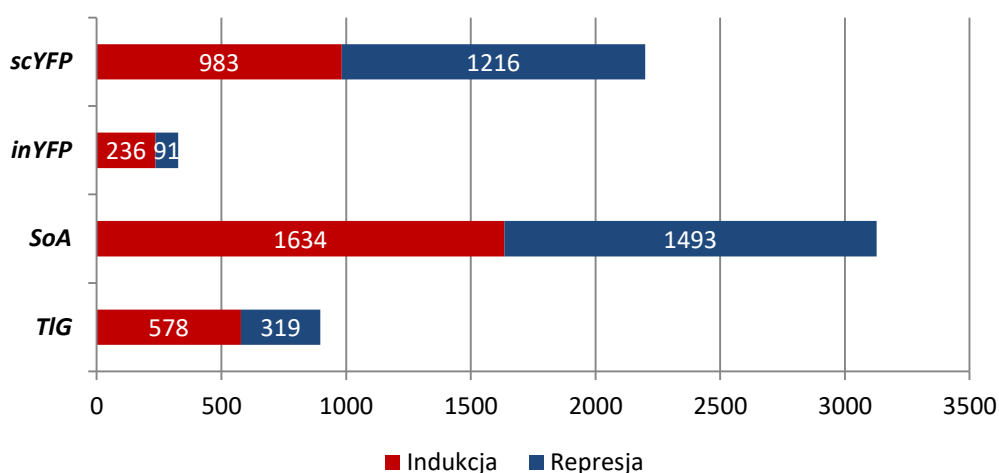
4.2. Globalna analiza transkryptomu rekombinowanych szczepów *Y. lipolytica* (pochodne Po1h) – sekwencjonowanie RNA

4.2.1. Podstawowa charakterystyka szczepów – obserwacje ogólne

W celu kompletnej charakterystyki rekombinowanych szczepów *Y. lipolytica* (pochodne Po1h) nadprodukujących biochemicznie różne r(s)-Prots przeprowadzono wysokoprzepustowe sekwencjonowanie RNA (RNAseq). Próbkę do analizy pobierano w trakcie hodowli ciągłych w momencie osiągnięcia stanu ustalonego (mat. uzup. Fig. S1 w **P.2**), z których następnie izolowano totalny RNA i sekwencjonowano. Otrzymane dane wykorzystano do globalnej analizy transkryptomu. Porównawczą analizę poziomu ekspresji genów pomiędzy badanymi szczepami *Y. lipolytica* przeprowadzono w dwóch „trybach”: 1) nieukierunkowanym – prowadzącym do pozyskania nowej wiedzy na temat innych procesów komórkowych, istotnych dla produkcji r(s)-Prots oraz 2) ukierunkowanym na analizę genów zaangażowanych w szlak translacyjno-sekrecyjny.

W pierwszej kolejności, zaobserwowano znaczną przewagę liczby genów różnicowo ekspresjonowanych (DEGs) w szczepie syntetyzującym sekrecyjną formę białka scYFP w porównaniu do jego wewnątrzkomórkowego odpowiednika (inYFP). Nadprodukcja inYFP spowodowała indukcję/represję sumarycznie 327 DEGs względem szczepu kontrolnego, podczas gdy nadprodukcja scYFP aż 2199 DEGs, czyli prawie 7-krotnie więcej (**Rycina 5**). Porównanie to jest szczególnie uzasadnione, jako że białko o tych samych właściwościach biochemicznych jest nadprodukowane w obu szczepach. Najwyższą całkowitą liczbę DEGs

zaobserwowano dla *SoA* (3127 DEGs). Ponadto szczepy produkujące sekrecyjne r-ProtS charakteryzowały się zwiększonym poziomem deregulacji genów w porównaniu do *inYFP*.



Rycina 5. Całkowita liczba genów różnicowo ekspresjonowanych (DEGs) określona dla każdego szczepu rekombinowanego w porównaniu ze szczepem kontrolnym. Oś X: liczba DEGs: czerwone słupki – geny o zwiększonej ekspresji vs. kontroli; niebieskie słupki – geny o obniżonej ekspresji vs. kontroli. Oś Y: rekombinowane szczepy wytwarzające różne białka reporterowe.

Na podstawie wstępnej charakterystyki szczepów obejmującej: ilościową analizę wyprodukowanych r(s)-ProtS (Tabela 2 w **P.2**), analizę poziomu ekspresji heterologicznych genów (Fig. 1 w **P.2**), a także całkowitą liczbę DEGs (**Rycina 5**) przedstawiono eksperymentalny zestaw czterech rekombinowanych szczepów *Y. lipolytica*, pochodnych Po1h, to jest : i) *inYFP* – szczep umiarkowanie wydajny ze znacznymi ograniczeniami na poziomie transkrypcji, które nie wynikają z przeciążenia szlaku wydzielniczego, ii) *scYFP* – szczep o wysokiej ekspresji i wydajnej produkcji (HSS; *High-Secretor Strain*), iii) *SoA* – szczep o wysokiej ekspresji, wysokiej produkcji i sekrecji białka (HSS), prawdopodobnie obciążony na poziomie fałdowania oksydacyjnego związanego z tworzeniem licznych -S-S-, iv) *TIG* – szczep o wysokiej ekspresji i umiarkowanej syntezie białka; prawdopodobnie obciążony na poziomie glikozylacji (**Tabela 1**).

4.2.2. Zmiany w transkryptomie rekombinowanych szczepów *Y. lipolytica* w warunkach nadprodukcji r(s)-ProtS – analiza nieukierunkowana

W ramach funkcjonalnej analizy deregulowanego transkryptomu przeprowadzono badanie wzbogacenia puli DEGs w reprezentowane procesy biologiczne. Wśród istotnie wzbogaconych procesów biologicznych, indukowanych w warunkach produkcji *scTIG*

i *inYFP*, wskazano DEGs zaangażowane w homeostazę jonów nieorganicznych (m.in. *F15411g*; **Rycina 6**), import do komórki oraz organizację ściany komórkowej (**Rycina 7**). Warto zauważyć, iż szczepy *TIG* i *inYFP* w trakcie hodowli w warunkach chemostatu rosły głównie jako komórki owoidalne w przeciwieństwie do szczepu kontrolnego czy *SoA*, dla których dominującym morfotypem były strzępki i wydłużone owoidy (mat. uzup. Fig. S5 w **P.2**). Z uwagi na fakt, iż wszystkie szczepy wzrastały w tym samych warunkach, wydaje się, że morfotyp był zależny od wprowadzonej zmiennej – typu nadprodukcji r(s)-Prots i wynikającej z tego odpowiedzi komórki. Dodatkowo, dla *SoA* nie zarejestrowano wzbogacenia procesów biologicznych jak: organizacja ściany komórkowej, morfogeneza czy kontrola kształtu komórki, co wskazuje na podobieństwo tych procesów dla *SoA* i szczepu kontrolnego. Z drugiej strony, gen *FLO11/F19030g*, kodujący TF zaangażowany we wzrost w formie strzępek, był jednym z genów o najbardziej obniżonej ekspresji w transkryptomie *SoA* i *scYFP* (**Rycina 6**), co różnicuje szczep *SoA* i kontrolny pod względem kształtu komórki. Z drugiej strony, dla *SoA* i *scYFP* (HSS) zidentyfikowano nadekspresję genu *E02024g* kodującego białko ściany komórkowej uczestniczące w indukcji wzrostu strzępek (**Rycina 6**). Ponadto, dla *scYFP*, który wzrastał zarówno w formie owoidów jak i wydłużonych strzępek, wykazano, że regulacja kształtu komórki (*Enrichment Fold*; EF: 5.31) była istotnie wzbogaconym procesem biologicznym o obniżonej ekspresji DEGs (**Rycina 7**). W grupie tej, wyszczególniono DEGs kodujące: małe GTPazy, białka z rodziny Rho, Rac1 oraz Rax2, a także zidentyfikowano gen *SOK2/B19602g* (podobny do *MGF1*), odgrywający rolę negatywnego regulatora wzrostu pseudostrzępek, co sugeruje, że komórki w chwili poboru próbki były w trakcie przejścia dimorficznego (mat. uzup. Fig. S5 w **P.2**). Hipotezę tę, potwierdza represja genów *MGF2/B19602g* i *D02189g*, będących mediatorami polimerazy II, zaangażowanymi w regulację przejścia między owoidami a strzępkami (*SoA*: -2.41/-1.91; *scYFP*: -1.38/-2.12).

Nazwa genu/Opis	Yali0	inYFP	scYFP	scSoA	scTIG
Transporter cynkowo-żelazowy	F15411g	***	***	***	***
Hsp40 z domeną DnaJ	E27588g	**	**		
Peroksysomalny PXMP2/4	E12881g	***	***	***	***
Akwaporyna	F01210g		***		
Domena typu GATA	E31757g		***	**	
Podobny do SHQ1_Ser-Glu-rich	D13200g		**		
Zawierający domenę ankiryiny	D15070g		**		
Kwaśna fosfataza PHO2	C19866g		***	**	
zakotwiczone białko ściany kom./EUTQ	E02024g		***	***	
B01100g	B01100g		**		
Białko z rodziny Vps13	A03817g		***	***	
DAUER związany z indukcją	F22187g		***	***	
Czynnik transkrypcyjny typu DeoR	E19030g		***	***	
Enzym deubikwytujący OTU	C03091g		**		
Związany z cytoskieletem	A10945g		*		
Podobny do mannoproteiny CWP2	E33891g	***	***	***	
Transporter potasu	C10311g		***	**	
Podobny do USO1	D23947g		***	***	
Proteaza aspartylowa YPS3	D01331g		**		
Białko błonowe sygnalizacja	B02112g		**		
Białko opiekuńcze SSA8	D22352g		***	***	
Katalaza CTT1	E34265g		**	***	
Katalaza CTT1	E34749g		***	***	
E08085g	E08085g		***	***	**
Mannoproteina ściany kom. CCW12	B21450g	**	***	***	***
Białko jądrowe	D17578g		***	***	
Transporter wiążący ATP STE6	A11473g		*	***	
Glioksalaza	F00682g		***	***	***
E32263g	E32263g		***	***	
Białko podobne do syntazy ATP	F26917g		***	***	
Mrówczan dehydrogenazy FDH	E14256g		***	*	
Białko z domeną F-box	A04059g		***	***	
Metaloreduktaza żelaza/miedzi	F11825g	***	***	***	***
Endopeptydaza typu asparaginowego	B00374g		***	***	
TF z domeną Zn(2)-C6	D01353g	***	***	***	***
TF związany z filamentacją	D02189g		***	***	***
RsfA TF	E26763g		***	***	
GAP1 Permeaza aa	B19492g	***	***	***	***
Permeaza maitozy MPH2	B00396g		*		
XBP1 Represor transkrypcji	F16511g	***	***	***	***
FLO11 TF	F19030g	***	***	***	***
C23518g	C23518g		***	***	**
Dioksygenaza	A06974g		***	***	**
Konserwatywny w Yarrowia	C07744g	***	***	***	***
B19866g	B19866g		***	***	
Wiązanie chromatyny	B05698g			**	
Białko katabolizmu uracylu URC4	E20647g		***	**	
Permeaza purynowa FCY2	F30569g			***	
Transkrypt polIA	B23496r			**	
Transporter aa UGA4	D00495g		***	***	
Lipaza triacyloglicerolowa	D23419g	***	***	***	
PBP2 PAB1 białko wiążące	C19558g	**	***	***	***
Permeaza puryny UAPC	E27852g		***	***	
Cyklodrolaza GTP URC1	D08250g		***	***	**
Transporter oligopeptydów OPT2	F09691g	**	***	***	**
Białko regulatorowe azotu AreA	D20482g		***	***	
Transporter mocznika DUR3	B04202g	**	***	**	***
Transporter amonu MEP2	F12925g			**	
Wykorzystanie etanolaminy EUTQ	D02453g		***	***	
Mannozylotransferaza GPI	F25993g			***	
Transporter amonu MEP2	E27203g	***	***	***	
Metalotioneina MTP2	C20066g				***
Metalotioneina MTP2	C20060g			*	***

* p value<0.05
** p value<0.01
*** p value<0.001

log2 estimated fold change

>5.05		INDUKCJA
4.05 - 5.04		
3.05 - 4.04		
2.05 - 3.04		
1.05 - 2.04		
(-1.04) - (-1.04)		
(-1.05) - (-2.04)		
(-2.05) - (-3.04)		
(-3.05) - (-4.04)		
(-4.05) - (-5.04)		
<(-5.05)		REPRESJA

Rycina 6. Graficzne przedstawienie najbardziej deregulowanych DEGs w szczepach nadprodukujących r(s)-Prots. Indukcję/represję DEGs przedstawiono za pomocą kolorów, zgodnie z legendą, w oparciu o wartość ekspresji wyrażoną jako log₂-krotność zmiany (FC). Wszystkie zaprezentowane DEGs spełniały warunek > 2.0 log₂ Estimated_FC, w co najmniej jednym ze szczepów nadprodukującym. Skrót: TF- Czynniki transkrypcyjne. Wykres został przygotowany przy użyciu narzędzia internetowego Morpheus (<https://software.broadinstitute.org/morpheus/>).

Uzyskane dane transkryptomyczne wykazały, że ekspresja genu *CTT1/E34749g* (**Rycina 6**), kodującego peroksysosomalną katalazę znaną jako marker reakcji komórki na stres oksydacyjny, była istotnie indukowana w HSS (13- i 6-krotnie wyższa względem szczepu kontrolnego, dla *SoA* i *scYFP*) [50]. Ponadto profile transkryptomiczne HSS charakteryzowały się wysoką ekspresją niezależnej od glutationu glioksyłazy *Hsp31/F00682g*, odgrywającej kluczową rolę m.in. w ochronie komórek przed reaktywnymi formami tlenu (ROS) [149], podobnie jak *CTT1*, co dodatkowo potwierdza występowanie stresu oksydacyjnego w warunkach nadprodukcji *scSoA* i *scYFP* (**Rycina 6**). W przeciwieństwie do tego, nadprodukcja *scTIG* powodowała represję genów *C20060g/C20066g* kodujących metalotioneiny zaangażowane w wiązanie metali ciężkich, a tym samym w ochronę komórki przed toksycznością metali i stresem oksydacyjnym, jak *CTT1* (**Rycina 6**).

Jak wspomniano powyżej, wśród istotnie wzbogaconych procesów biologicznych wskazano proces organizacji ściany komórkowej, który ulegał represji w *scYFP* (EF: 1.94), w przeciwieństwie do indukcji w *inYFP* (EF:5.14) i *TIG* (EF:4.37) (**Rycina 7**). W grupie tej, zidentyfikowano m.in.: kilka glikozydaz, syntazę chityny, mutazę fosfoacetylkoglucozaminy, izomerazę glukozamino-6-fosforanową, N-acetylotransferazę, mannozylotransferazy i inne. Pomimo, że mannozylotransferazy jak: *Ktr1p*, *Ktr5p* i *Mnn2p* przypisano do tego procesu, te transferazy odgrywają rolę w glikozylacji białek. Podobnie jak gen *F25993g*, kodujący mannozylotransferazę 2 związaną z błoną AG, biorący udział w glikozylacji białek i transporcie pęcherzykowym, był jednym z genów o istotnie obniżonym poziomie ekspresji w *SoA* (**Rycina 6**). Dodatkowo, zauważono, znaczącą represję genów kodujących sekrecyjne proteazy oraz lipazy, w tym: *TGL3/D17534g*, *LIP4/E08492g*, *AXP1/B00374g*, *XPR2/A08360g* w *SoA*, a trzy z nich *XPR2*, *AXP1*, *LIP4* w *scYFP*. Białka te różnią się liczbą miejsc glikozylacji (średnio 3 N-glikozylacje, 8 O-glikozylacji), a ich represja poprawia zdolność szlaku wydzielniczego wymaganego do syntezy *scSoA/scYFP*, jednocześnie zmniejszając zapotrzebowanie na enzymy glikozylujące. Ponadto, w związku z tworzeniem i przebudową ściany komórkowej w HSS, zaobserwowano istotną indukcję genów takich jak: *CCW12/B21450g*, *STE6/A11473g*, *CWP11/E22286g* i *CWP12/E26125g*.

Profile transkrypcyjne szczepów nadprodukujących *scSoA* i *scYFP* charakteryzowały się znacząco wzbogaconymi indukowanymi procesami zaangażowanymi w: biogenezę i składanie rybosomów, transkrypcję, translację, regulację transkrypcji, modyfikację chromatyny oraz syntezę i modyfikację histonów (**Rycina 7**). Co więcej, istotna represja genu

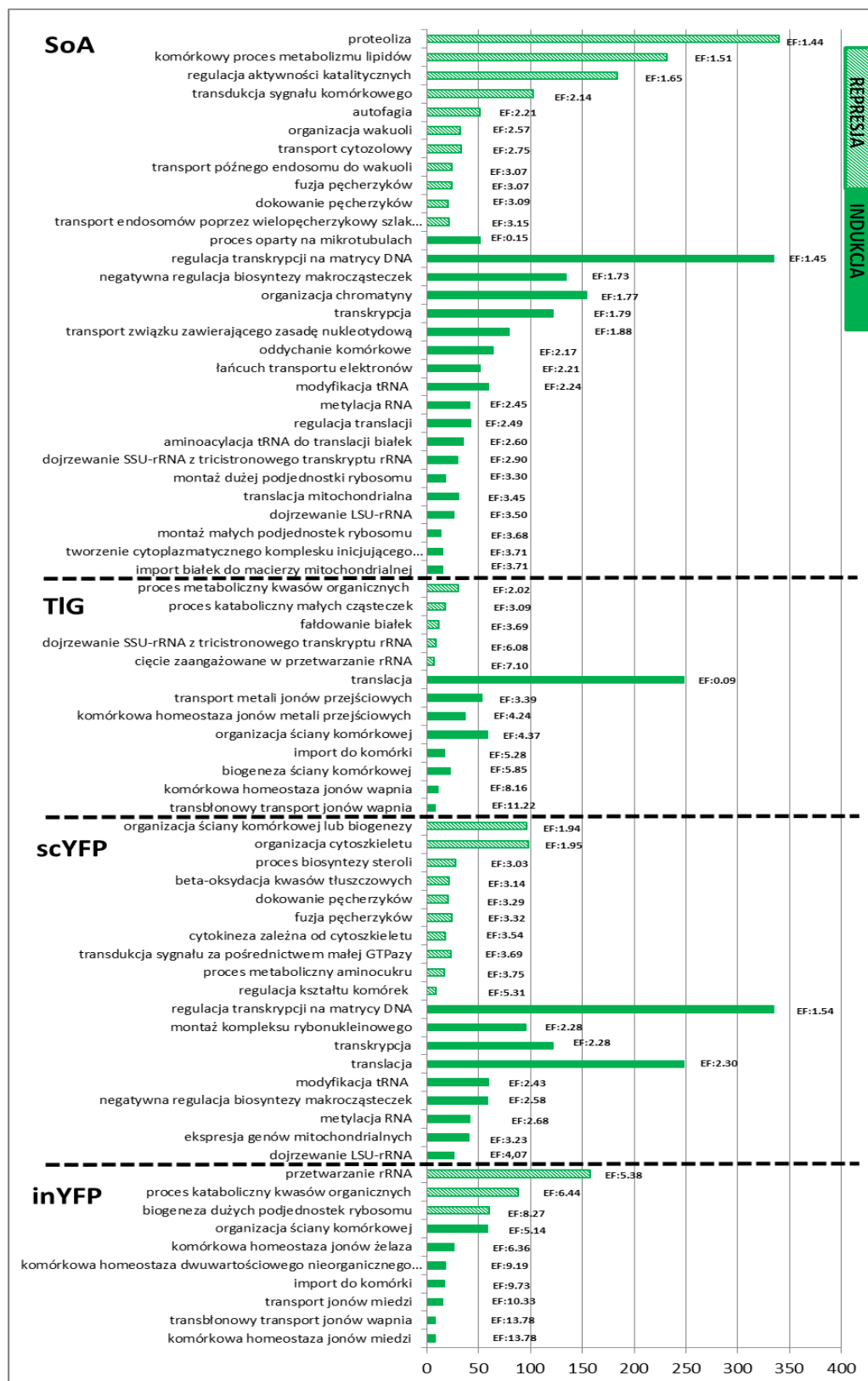
E26763g (RsfA) kodującego rybosomalny czynnik wyciszający [150] w *scYFP* (-2.14) i *SoA* (-5.39) (**Rycina 6**), potwierdza intensywnie zachodzące procesy transkrypcji i translacji. Z drugiej strony, procesy związane z biogenezą rybosomów czy przetwarzaniem rRNA znaleziono wśród wzbogaconych procesów biologicznych ulegających represji w *TIG* i *inYFP* (**Rycina 7**).

Organizacja struktur wakuolarnych (EF: 2.57), autofagia (EF: 2.21) i proteoliza (EF: 1.44) stanowiły istotnie wzbogacone procesy biologiczne ulegające represji w *SoA* (**Rycina 7**; tylko tendencja obserwowana dla *scYFP*). W grupie tej, zidentyfikowano DEGs takie jak: *ATG8/E02662g*, *ATG13/F03432g* czy *ATG15/F06358g* związane z autofagią czy proteazy wakuolarne jak *PRC1/A18810g* czy *PRB1/B16500g*. Z drugiej strony, transkryptom szczepu *SoA* został wzbogacony o indukowane geny kodujące zewnątrzkomórkowe peptydazy jak: *B20526g*, *C10135g*, *D10967g* i zakotwiczoną w GPI proteazę aspartylową *Yps3/D22957g* (także w *scYFP*), uczestniczące w proteolizie białek, o znanym negatywnym wpływie na stabilność białek wydzielniczych [52].

Wśród DEGs o obniżonym poziomie ekspresji pod wpływem nadprodukcji r(s)-Prots zidentyfikowano gen *GAP1/B19492g*, kodujący permeazę aminokwasową, we wszystkich badanych szczepach, co było zaskakujące, biorąc pod uwagę wysokie zapotrzebowanie na białkowe materiały budulcowe (**Rycina 6**). W przeciwieństwie do *GAP1*, inne permeazy AA były indukowane w transkryptomach szczepów, w tym permeazy AA: *AGP3/E20713g* i *ALP1/F01254g*. Przeciwnie, geny zaangażowane w metabolim AA były istotnie indukowane w HSS. W grupie tej wskazano m.in. DEGs kodujące: *Bat2*, *Pda2*, liazę 3-hydroksy-3-metyloglutaranowa-CoA i *Gdh2* stanowiące centralny punkt rozgałęzienia sieci obrotu AA [151,152] o podwyższonym poziomie ekspresji w *SoA*, co wskazuje na intensywny metabolizm AA w komórce.

W odniesieniu do ogólnej odpowiedzi komórkowej na nadprodukcję określonego białka, beta-oksydacja oraz komórkowe procesy metaboliczne lipidów ulegały istotnej represji w *scYFP* (EF: 3.14) i *SoA* (EF: 1.51) (**Rycina 7**). W drugiej grupie zidentyfikowano geny kodujące liczne białka związane z błoną komórkową ER: ceramidową hydroksylazę kwasów tłuszczowych o bardzo długim łańcuchu *SCS7*, O-acylotransferazę diacyloglicerolu, GPmannozylotransferazę 4, transamidazę, glikozylotransferazę dolichylowa (N-glikozylacja), transferazę *GlcNAc-1-P* (odgrywająca rolę w N-glikozylacji), syntazę difosforanu

dehydrodolichylu (która jest prenylotransferazą zaangażowaną w glikozylację białek) czy O-acylotransferazę diacyloglicerolu 1 i wiele innych. Wydaje się, że niska przewidywana liczba miejsc glikozylacji tych dwóch rs-Prots, może stanowić przyczynę tego zjawiska. Co znamienne, w komórkach nadprodukujących scSoA i w mniejszym stopniu scYFP zaobserwowano intensyfikację procesów oddychania komórkowego i wytwarzania energii, obserwowanej jako zwiększony import białek do mitochondrium czy intensywną ekspresją genów mitochondrialnych (**Rycina 7**). W tym kontekście, wśród najbardziej indukowanych DEGs obserwowano *F26917g*, kodujący białko podobne do syntazy ATP (**Rycina 6**).



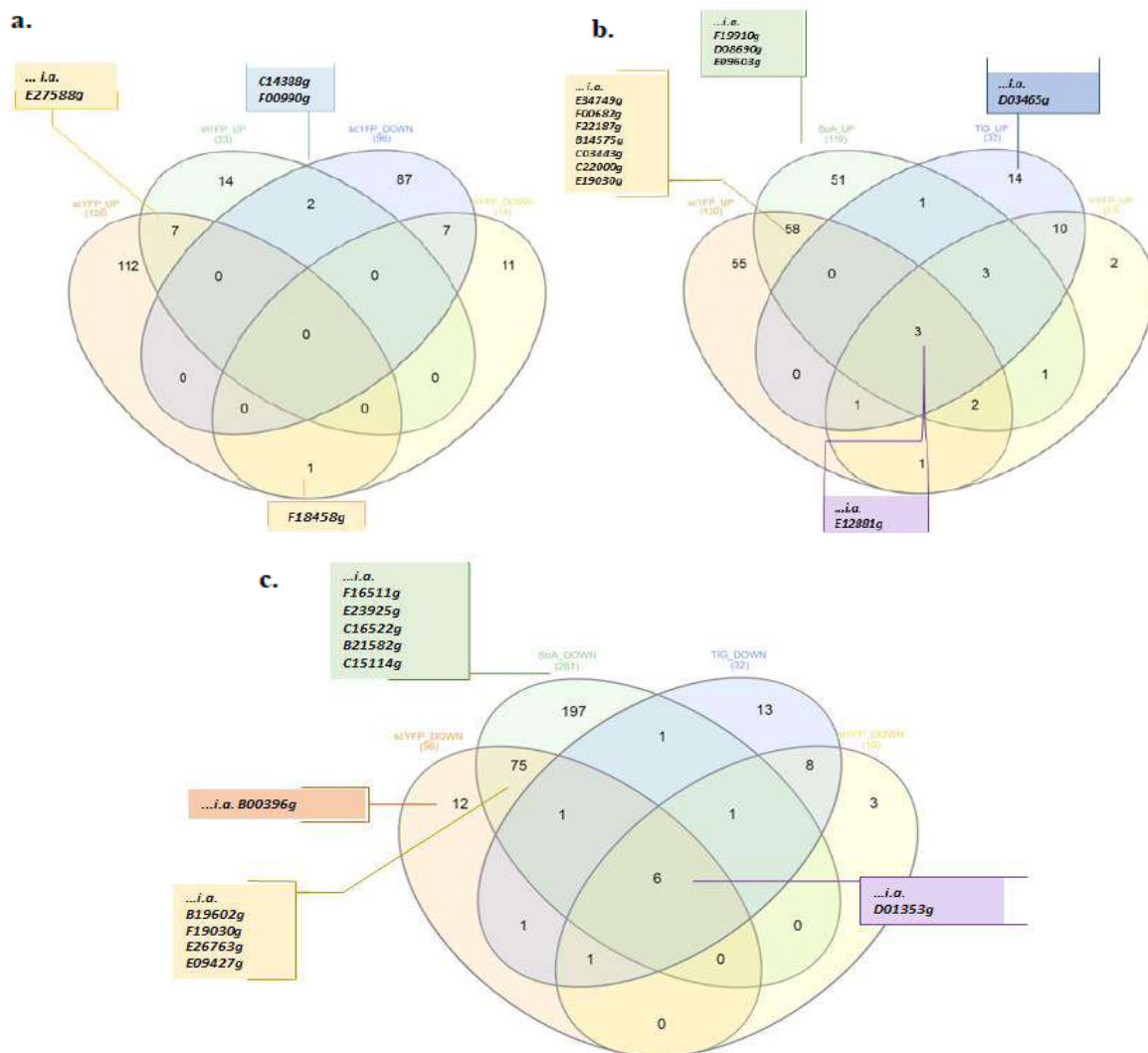
Rycina 7. Główne istotnie wzbogacone procesy biologiczne dla zestawu DEGs deregulowanych w odpowiedzi na nadprodukcję r(s)-Prots. Pokazano tylko procesy wzbogacone na poziomie istotności $p < 0.05$. Słupki reprezentują liczbę DEGs przypisanych do tych procesów (pełne – indukcja DEGs; kreski – represja DEGs) a cyfry oznaczają EF.

4.2.3. Podobieństwa i różnice w transkryptomie indukowane nadprodukcją r(s)-Prots

Identyfikację wspólnych i unikalnych DEGs w odpowiedzi na nadprodukcję biochemicznie różnych r(s)-Prots w szczepach *Y. lipolytica* wykonano poprzez analizę diagramów Venna. Przeprowadzona analiza umożliwiła zidentyfikowanie genów specyficznym: i) reagujących na wydzielanie białek (porównanie *inYFP* i *scYFP*; **Rycina 8a**); ii) oznaczających nasilone fałdowanie oksydacyjne (*SoA* vs. inne) lub iii) glikozylację (*TIG* vs. inne; **Rycina 8b.c**). Szczegółową, pełną listę tychże DEGs wraz z krótkim opisem i nazwą genu przedstawiono w mat. uzup. (nośnik elektroniczny) Fig. S6 w **P.2**.

Deregulowane transkrypty *inYFP* i *scYFP* wykazywały 7 wspólnych genów ulegających indukcji, w tym *HSP40/E27588g* kodujący białko chaperonowe zaangażowane w fałdowanie powstających białek, translokację polipeptydów przez błony, koordynację odpowiedzi komórki na stres i kierowanie białek do degradacji. Z kolei gen *F18458g* wykazywał przeciwny profil regulacji w *in-/sc-YFP* (represja w *inYFP* a indukcja w *scYFP*). Wśród genów o przeciwnym profilu regulacji (indukcja w *inYFP* a represja w *scYFP*) wyszczególniono metylotransferazę *C14388g* oraz *F00990g* kodujący białko zaangażowane w organizację ściany komórkowej i wzrost strzępek (**Rycina 8a**) co ponownie potwierdza hipotezę, że **HSS były w trakcie przejścia dimorficznego**. W kolejnym kroku, analiza diagramu Venna wykazała, że jedynym znacząco „wzbogaconym” procesem biologicznym wspólnym dla indukowanych genów w HSS (n=58; **Rycina 8b**) była „odpowiedź na nadtlenek wodoru” (EF: 66.7), w którym zidentyfikowano m.in. gen *F22187g* o dotychczas niepoznanej funkcji u *Y. lipolytica*. W grupie tej, znalazły się również: tioesteraza (*B14575g*), *HSP42/C03443g*, czynnik transkrypcyjny *E19030g*, katalaza *CTT1* czy glioksylaza *HSP31*, *C22000g* (**Rycina 8b**). Co interesujące, wśród genów ulegających znaczącej represji w *SoA* zidentyfikowano gen *XBPI/F16511g* (**Rycina 6**; **Rycina 8c**) odgrywający rolę represora transkrypcji, zaangażowany w przejście fazowe G1→S mitotycznego cyklu komórkowego, a także *CLNI/C15114g*, promujący przejście G1→S, co **wskazuje na zatrzymanie fazy wzrostu (G1→G0)** (**Rycina 8c**). Nadprodukcja *scSoA/scYFP* wpłynęła także istotnie na represję genów (**Rycina 8c**) kodujących białka uczestniczące w metabolizmie cukrów (*STL1/C16522g*, w *SoA* i *MPH2/B00396g* w *scYFP*), a także represora polimerazy II, *RSFA/E26763g* w obu szczepach. Wśród znacząco „wzbogaconych” procesów biologicznych, specyficznych dla *SoA* (n=51; **Rycina 8b**), zidentyfikowano „proces metabolizmu leucyny” (EF: 56.41) oraz „proces katabolizmu kwasu

karboksylowego (EF: 9.51), w których wskazano indukowane DEGs, odpowiedzialne za centralny metabolizm konwersji i transformacji aminokwasów (*BAT2*, *PDA2*, *MVA1*, *GDH2*, omówione powyżej). Co znamienne, zaobserwowano represję genów zaangażowanych w tworzenie strzępek w HSS (n=75): *MGF2/B19602g* oraz *FLO11/F19030g* kodujących czynniki transkrypcyjne, a także geny *E23925g* i *MHY1/B21582g* w *SoA*. (**Rycina 8c**), biorących udział w przejściu dimorficznym. Ponadto w grupie DEGs ulegających represji, wspólnych dla HSS, zidentyfikowano gen *TRX1/E09427g*, kodujący cytoplazmatyczną tioredoksynę, zaangażowaną w ochronę komórek przed działaniem stresu oksydacyjnego. Pomimo tego, że żaden proces biologiczny nie został wzbogacony w deregulowanym transkryptomie *TIG*, we wszystkich szczepach nadprodukujących sekrecyjne r-Prots (*TIG*, *SoA*, *scYFP*) zaobserwowano indukcję genu kodującego kwaśną fosfatazę Pho2 (mat. uzup. Fig.S6 w **P.2**). Co więcej, wszystkie badane szczepy *Y. lipolytica* nadprodukujące r-Prots charakteryzowały się indukcją *E12881g* (*SoA* log₂FC: 2.6; *scYFP*: 2.6; *TIG*: 2.82; *inYFP*: 2.74) (**Rycina 8b**) i represją *D01353g* (*SoA* log₂FC: -2.73; *scYFP*: -2.15; *TIG*: -2.79, *inYFP*: -2.44), genów kodujących niescharakteryzowane dotychczas białka.



Rycina 8. Indukowane i represjonowane wspólne lub unikalne geny/zestawy genów (DEGs) dla poszczególnych szczepów *Y. lipolytica* w odpowiedzi na nadprodukcję r(s) – Prots. Wskazano liczbę DEGs dla poszczególnych grup. a. *inYFP* vs. *scYFP*; b. indukowane DEGs w *SoA*, *TIG*, *scYFP* i *inYFP*; c. represjonowane DEGs w *SoA*, *TIG*, *scYFP* i *inYFP*. Analizę diagramów Venna przeprowadzono na listach DEGs z wartościami ekspresji ograniczonej do 2-krotnej zmiany.

4.2.4. Ukierunkowana analiza transkryptomu pod kątem specyficznej odpowiedzi szlaku wydzielniczego na nadprodukcję r(s)-Prots

Ukierunkowaną analizę przeprowadzono ograniczając zestaw analizowanych genów do uprzednio wytypowanych, zaangażowanych w syntezę, fałdowanie, dojrzewanie i sekrecję polipeptydów w drożdżach [26–28]. Wyróżniono 27 kategorii, według których sklasyfikowano geny (mat. uzup. (nośnik elektroniczny) Tabela S3 w **P.2**). Względny stosunek ekspresji genów w porównaniu do szczepu kontrolnego dla wszystkich genów (ok.700 genów) przypisanych do tego procesu biologicznego

zaprezentowano w mat. uzup. (nośnik elektroniczny) Tabela S4 w **P.2**. Znormalizowane odczyty ekspresji wybranych genów przedstawiono na **Rycinie 9** i użyto do ilościowej analizy poziomu transkryptów.

W pierwszej kolejności, zaobserwowano, że ekspresja genu kodującego kluczowy **czynnik transkrypcyjny Hac1**, który inicjuje UPR w drożdżach [44,55,153], była **indukowana o 23-27% wyłącznie w szczepach produkujących scTIG i inYFP**, bez zmian w poziomie ekspresji dla *SoA* czy *scYFP* (**Rycina 9a**). Pomimo tego, że ekspresja *HAC1* była stosunkowo niska, gen ten ulegał ciągłej ekspresji, zarówno w warunkach nadprodukcji rs-Prots jak i bez. Jak wiadomo, Hac1 pośredniczy w deregulacji ponad 100 genów, więc widoczny najwyższy poziom indukcji *HAC1* w *inYFP* powinien spowodować istotne zmiany w transkryptomie, co jednak było w tym szczepie mocno ograniczone. Podobnie, nie zidentyfikowano genów kodujących białka odpowiadające na UPR, zaangażowane w fałdowanie białek, degradację proteasomów czy ERAD w szczepach produkujących scTIG czy inYFP, z wyjątkiem kilku genów uczestniczących w sortowaniu białek do wakuoli: *ATG8*, *ATG13*, *VPS70/B05258g* i *CUP5/F24475g*; czy kodujących proteazy wakuolarnie jak: *PEP4/F27071g*, *PRC1*, *PRB1* (**Rycina 9b.f**, **Rycina 10**). Z kolei **ekspresja znanych regulatorów UPR**, w tym: *IRE1/A14839g*, *KAR2/E13706g*, *ERJ/C05819g* czy kluczowej izomerazy dwusiarczkowej *PDI/E03036g*, oksydazy tiolowej *ERO1/D09603g* i syntetazy glutationu *GSH/C17831g*, **pozostawała na poziomie kontrolnym we wszystkich rekombinowanych szczepach (Rycina 9c.d)**. Obserwacja ta, była szczególnie zaskakująca dla szczepu *SoA*, produkującego białko bogate w wiązania dwusiarczkowe. W przeciwieństwie do stałego poziomu ekspresji *KAR2* i *IRE1*, ekspresja genu *SLS1/E32703g*, kodującego mediator interakcji między Ire1 i Kar2 [21] niezbędny do translokacji kotranslacyjnej u *Y. lipolytica* [19,36], była represjonowana we wszystkich szczepach produkujących rs-Prots, ale nie inYFP (**Rycina 9c**). W odniesieniu do procesów fałdowania oksydacyjnego, tworzenia mostków dwusiarczkowych i związanego z tym stresu oksydacyjnego zaobserwowano obniżony poziom ekspresji *TRX1* kodującego cytoplazmatyczną tioredoksynę w szczepach HSS i zwiększony poziom *SEC61/E21912g* w *SoA* (**Rycina 9e**).

Sprawdzono również profile ekspresji genów kodujących białka, odgrywających kluczową rolę w procesach translokacji powstającego polipeptydu przez błonę ER [15,28]. Istotnie podwyższony poziom ekspresji stwierdzono dla cytozolowych białek

chaperonowych z rodziny Ssa, w szczególności *SSA5/F25289g* i *SSA8/D22352g* w HSS (**Rycina 9d**); zbliżony wzorzec ekspresji obserwowano także dla *SSE1/E13255g*.

W odniesieniu do specyficznych odpowiedzi komórki na nadprodukcję r-Protos, zaobserwowano znaczną **indukcję genów zaangażowanych w syntezę oligosacharydów i przetwarzanie glikoprotein w TIG** (*DPM/D06017g*, *PMI40/B18348g*, *SEC59/E18942g*, *ALG5/D06017g*, *SED1/E07832g*), **co było spodziewane na podstawie właściwości biochemicznych tego białka**. Ponadto, spośród 10 genów należących do rodziny Alg, 7 ulegało represji w HSS, przy jednoczesnym braku zmian ekspresji tych genów dla *TIG* i *inYFP* (mat. uzup. (nośnik elektroniczny) Tabela S4 w **P.2**). Ukierunkowana analiza transkryptomu wykazała, że w przypadku szczepu *inYFP* geny zaangażowane w powyższe procesy, takie jak: dojrzewanie glikoprotein, fałdowanie białek, tworzenie -S-S-, ERAD, degradację białek czy transport pęcherzykowy nie ulegały zmienionej ekspresji w zadanych warunkach. Spośród ok. 250 genów przypisanych do tych procesów, jedynie 4 wykazały zmieniony profil ekspresji w różnicowym transkryptomie *inYFP* (represjonowane: *FES1/F11121g*, *BFR2/D11462g*, indukowane: *HSP26/C03465g*, *HRT1/B17358g*) (**Rycina 9d.g**). Podobnie, istotnie podwyższony poziom ekspresji genu *HSP26*, a także *HSP42/E18546g* zaobserwowano dla HSS (*HSP42*: 2.26- dla *SoA* i 4.48- krotnie dla *scYFP*).

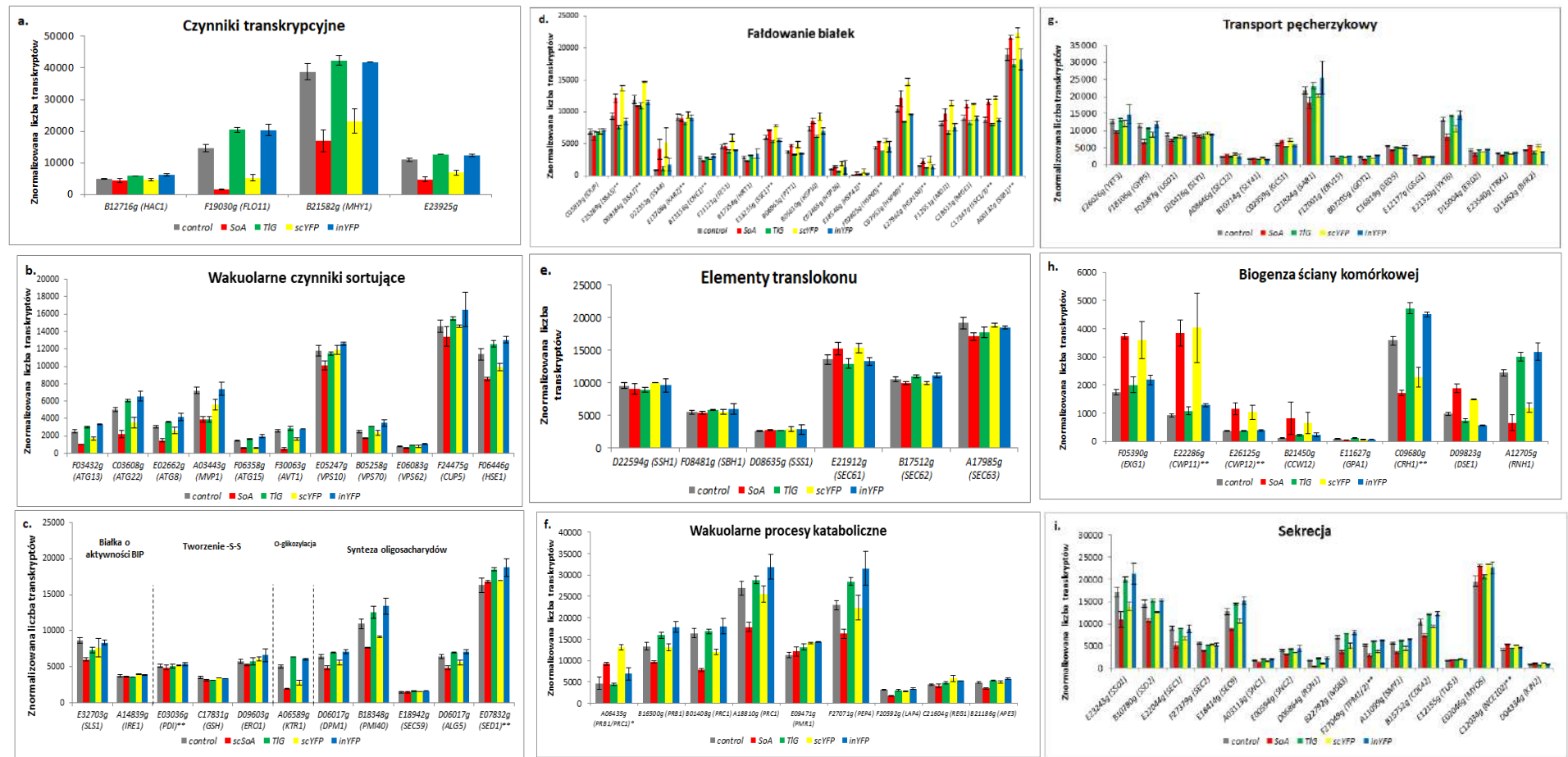
Interesujących obserwacji dokonano w odniesieniu do genów przypisanych do kategorii „wakuolarnych czynników sortujących lub wakuolarnych procesów katabolicznych” (**Rycina 9b.f; Rycina 10**). Znacząco **podwyższony poziom ekspresji tych genów stwierdzono dla TIG i inYFP**, co wskazuje na **zwiększoną degradację białek**. Wynik ten kontrastował z wartościami ekspresji obserwowanymi dla HSS. Zgodnie z przeprowadzoną analizą wzbogacenia procesu biologicznego (**Rycina 7**): degradacja białek, autofagia oraz sortowanie białek do wakuoli charakteryzowały się obniżoną reprezentacją w HSS, co było konsekwencją represji kilku określonych DEGs przypisanych do tych procesów (mat. uzup. (nośnik elektroniczny) Tabela S4 w **P.2**). Wśród nich zidentyfikowano m. in: *ATG13*, *ATG22/23/C03608g*, *ATG15* oraz *AVT1/F30063g* (**Rycina 9b, Rycina 10**). Z drugiej strony, kilka genów przypisanych do wakuolarnych procesów katabolicznych było istotnie indukowanych we wszystkich (*PMR1/E09471g*) lub w trzech (*PRB1/PRC1/A06435g*, *REG1/C21604g*) rekombinowanych szczepach *Y. lipolytica* (**Rycina 9f**).

Ponadto, nie zaobserwowano wyraźnego wpływu nadprodukcji r(s)-Prots na poziom ekspresji genów zaangażowanych w ERAD, który albo pozostał niezmienny w *inYFP*, albo nieznacznie obniżony (pięć genów w *SoA*, trzy geny w *scYFP* i jeden gen w *TIG*) (mat. uzup. (nośnik elektroniczny) Tabela S4 w **P.2**). Dodatkowo, **dla szczepów HSS stwierdzono indukcję genów zlokalizowanych w mitochondriach**, takich jak: *MDJ1/F12551g*, *MGE1/C18513g*, *HSP60/F02805g*, *HSP10/B05610g*, oraz *PRX1/A19426g* tylko w *SoA* (**Rycina 9d**). Tak wysoka nadreprezentacja indukowanych genów mitochondrialnych koresponduje ze znacznie wzbogaconym procesem biologicznym „ekspresją i translacją genów mitochondrialnych”, jak również z „oddychaniem komórkowym”, zidentyfikowanym dla HSS (**Rycina 7**).

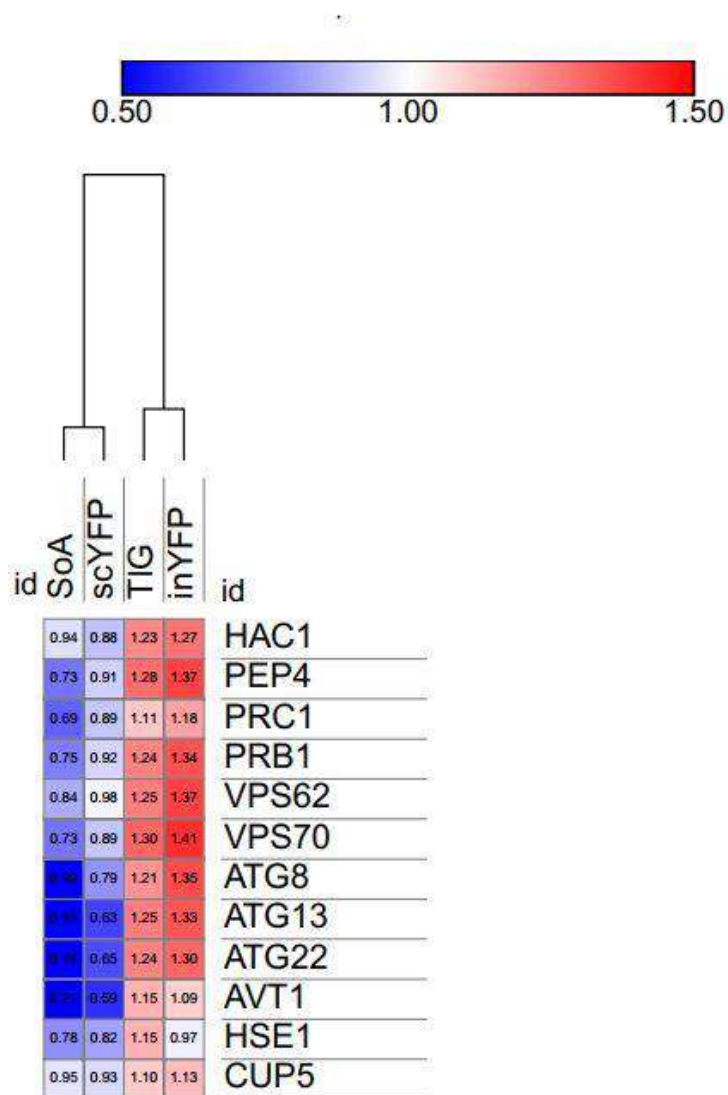
Różny poziom ekspresji genów uczestniczących w tworzeniu i przebudowie ściany komórkowej obserwowano dla wszystkich analizowanych w tym badaniu rekombinowanych szczepów *Y. lipolytica*. W związku z powyższym, nie zaobserwowano specyficznego trendu w zależności od rodzaju nadprodukowanego r(s)-Prots. Niemniej jednak, wśród genów przypisanych do tych procesów, wyszczególniono kilka ulegających silnej indukcji lub represji w zadanych warunkach. Wśród istotnie indukowanych genów w HSS, zidentyfikowano: *CCW12* (3.32/2.68-krotna zmiana w *SoA/scYFP*) kodujący mannoproteinę ściany komórkowej, *DSE1/D09823g* (2.01/1.14-krotna zmiana w *SoA/scYFP*) odgrywający rolę w separacji komórek po cytokinezie, *EXG1/F05390g* (2.23/1.93-krotna zmiana w *SoA/scYFP*) kodujący egzo-1,3-beta-glukanazę zaangażowaną w budowę ściany komórkowej, a także dwa strukturalne białka ściany komórkowej jak: *CWP11* i *CWP12*. Z drugiej strony, poziomy transkryptów genów *GPA1/E11627g* oraz *CRH1/C09680g*, kodujących białko G błony plazmatycznej oraz transglikozylazę chitynową, były obniżone w wyniku nadekspresji *SoA* i *scYFP*, natomiast nieznacznie indukowane w szczepach *TIG* i *inYFP* (**Rycina 9h**). Jak wspomniano powyżej, czynniki transkrypcyjne zaangażowane w tworzenie, przebudowę ściany komórkowej takie jak: *FLO11*, *MHY* i *E23925g* zidentyfikowano w grupie DEGs o istotnej represji w HSS i indukcji w *inYFP/TIG* (**Rycina 9a**). Co więcej, biogeneza ściany komórkowej była istotnie zmienionym procesem biologicznym dla *scYFP* i *TIG* (**Rycina 7**, **Rycina 9h**).

W odniesieniu do **transportu pęcherzykowego i procesów fuzji błon** (**Rycina 9g**) zidentyfikowano kilka DEGs istotnie deregulowanych głównie dla HSS. Wspólny wzorzec ekspresji w HSS zaobserwowano m.in. dla genów ulegających: i) indukcji:

SEC12/A08646g, *BFR2/D11462g*, *SLY41/B10714g*, *GCSI/C02959g*; czy ii) represji: *SAR1/C21814g*, *ERV15/F12001g*, *GOT1/B07205g* i *USO1/F0238g*. Wśród genów przypisanych do procesów dystalnej sekrecji (**Rycina 9i**) zidentyfikowano *SSO1,2/E23243g*, *B10780g*, *SNC1,2/A03113g*, *E00594g* czy *SEC1,2,9 /E22044g*, *F27379g*, *E18414g*, które **ulegały istotnie obniżonej ekspresji w HSS**, oraz nieznacznej indukcji bądź pozostawały na poziomie kontrolnym w *inYFP/TIG*. We wszystkich badanych szczepach zaobserwowano indukcję genów *TUS1/E12155g*, *RHO1* i *MYO5/E02046g*, jednakże zmiana to była istotna statystycznie jedynie w HSS. Co ciekawe, **ekspresja genu *NCE102/C12034g* kodującego białko zaangażowane w eksport niekonwencjonalny (białek pozbawionych sekrecyjnych sekwencji sygnałowych), była podwyższona w wszystkich trzech szczepach produkujących rs- Prots.**



Rycina 9. Poziom ekspresji wybranych genów zaangażowanych w fałdowanie, dojrzewanie i sekrecję r(s)-Prots. Znormalizowane wartości transkryptów zastosowano do ilościowego określenia poziomu transkrypcji. a. Czynniki transkrypcyjne, b. Wakuolarnie białka sortujące, c. Białka o aktywności BIP, tworzenie wiązań disiarczkowych, O-glikozylacja, synteza oligosacharydów, d. Fałdowanie białek, e. Elementy translokonu, f. Wakuolarnie procesy kataboliczne g. Transport pęcherzykowy, h. Biogeneza ściany komórkowej, i. Sekrecja. Oś X: nazwy genów Yali_, oś Y: znormalizowana liczba transkryptów. *- wartości pomnożone przez 10, ** - wartości podzielone przez 10.



Rycina 10. Graficzne przedstawienie deregulowanych DEGs zaangażowanych w wakuolarne procesy kataboliczne oraz wakuolarne sortowanie białek w szczepach nadprodukcujących r(s)-Prots. Indukcję/represję DEGs przedstawiono za pomocą kolorów, zgodnie z legendą, w oparciu o wartość ekspresji wyrażoną jako krotność zmiany (FC) ekspresji genu w próbie badanej względem kontroli. Wykres został przygotowany przy użyciu narzędzia internetowego Morpheus (<https://software.broadinstitute.org/morpheus/>).

4.3. Hodowle okresowe rekombinowanych szczepów *Y. lipolytica* z ko-nadekspresją rs-YFP i SH (P.3)

4.3.1. Klonowanie homologicznych SHs i selekcja szczepów

W oparciu o porównawczą analizę globalnego transkryptomu (analizę ekspresji wszystkich genów w genomie *Y. lipolytica*) w **P.2** oraz na podstawie przeglądu danych literaturowych [69,85,104], opracowano listę genów pomocniczych *SHs* (*Secretory Helpers*) (Tabela 1 w **P.3**). Według przyjętej definicji i założeniu, nadekspresja takich *SHs* ma prowadzić do intensyfikacji produkcji rs-Protos. Strategię klonowania dla ko-transformacji *SCYFP* i *SH* w Po1f zaprezentowano na **Rycinie 3**. W celu wyboru szczepu reprezentującego badany wariant *SCYFP-SH*, cztery subklony *Y. lipolytica*, analizowano pod kątem przyrostu biomasy, oraz wewnątrz- i zewnątrzkomórkowej FL (inYFP/exYFP) w trakcie hodowli okresowych. Wykresy przedstawiające kinetykę wzrostu, produkcji YFP (inYFP i exYFP) przedstawiono w mat. uzup. Fig. S2. w **P.3**. Jak zaobserwowano, zmienność międzyklonalna była niska, a nadekspresja *SH* nie zaburzyła wzrostu w porównaniu do szczepu kontrolnego (Control_scYFP) ($p < 0,05$). Z drugiej strony, w zależności od rodzaju ko-transformowanego *SH*, szczepy wykazywały zmienność parametrów sp_in-YFP/ex-YFP. Na tej podstawie wytypowano jeden, reprezentatywny subklon spośród każdego wariantu *SCYFP-SH*. Taki wybrany subklon wykazywał podobny przebieg krzywej wzrostu do szczepu referencyjnego, a także był najbardziej reprezentatywny dla grupy rekombinantów pod względem parametrów sp_in-YFP/ex-YFP. Nadekspresję poszczególnych *SH* w badanych szczepach potwierdzono w reakcji qPCR w czasie rzeczywistym (nie pokazano). Szczegółowym analizom poddano 12 reprezentatywnych szczepów *Y. lipolytica* (**Tabela 3**).

4.3.2. Kinetyka wzrostu, syntezy i sekrecji rs-YFP

Dwanaście reprezentatywnych szczepów *Y. lipolytica* z ko-nadekspresją *SCYFP* i jednego z badanych *SH* poddano serii hodowli okresowych, równolegle ze szczepem kontrolnym (z nadekspresją wyłącznie *SCYFP*). Hodowlę prowadzono przez 96 h w dwóch wariantach temperaturowych 25 °C i 30 °C jak opisano w podrozdziale **3.3.2**. Na Fig.1. w **P.3** przedstawiono kinetykę wzrostu (1.x.3) oraz syntezy (inYFP; 1.x.1) i sekrecji (exYFP; 1.x.2) białka reporterowego przez rekombinowane szczepy *Y. lipolytica*. Podobnie, jak wywnioskowano na podstawie wstępnych badań

przesiewowych, nadekspresja *SH* nie wpłynęła znacząco na wzrost poszczególnych szczepów rekombinowanych względem kontrolnego ($p < 0.05$). Natomiast, **obniżenie temp. hodowli o 5 °C (do 25 °C) istotnie wpłynęło na akumulację biomasy ($p < 0.05$), na korzyść niższej temp.** (Fig. 1.x.3 w **P.3**). Zgodnie z uzyskanym profilem wzrostu, niezależnie od rodzaju nadekspresjonowanego *SH*, zróżnicowanie tempa wzrostu zostało zapoczątkowane po zastosowaniu obniżonej temp, za wyjątkiem kilku punktów czasowych dla określonych wariantów SCYFP-SH np. scYFP-SSO1 w 48h w 25 °C lub scYFP-YET3 w 48h w 30 °C ($p < 0.05$).

Przeciwnie, **nadekspresja różnych *SHs* spowodowała istotne zmiany w odnotowanych wartościach fluorescencji wewnątrz-/zewnątrkomórkowej (inYFP/exYFP)** (Fig. 1.x.1., 1.x.2 w **P.3**). Niezależnie od zastosowanej temp. hodowli zaobserwowano zwiększoną syntezę i retencję YFP przy nadekspresji *RPL3*, *SSA5*, *SSA8* (Fig. 1.B/C/D.1 w **P.3**). Co więcej, poziom akumulacji YFP (inYFP) w tych trzech szczepach osiągnął najwyższe wartości spośród wszystkich analizowanych wariantów w tym badaniu. Szczególnie wysokie wartości inYFP zaobserwowano w późnej fazie stacjonarnej, gdy powyższe szczepy hodowano w obniżonej temp. (25 °C). Zależność temperaturową stwierdzono także w przypadku szczepu z nadekspresją *SSO1*, gdzie odnotowano zwiększoną akumulację YFP (inYFP) w temp. 30 °C, podczas gdy w temp. 25 °C nie zaobserwowano różnic pomiędzy szczepem rekombinowanym a kontrolnym (Fig. 1.K.1 w **P.3**). W pozostałych analizowanych wariantach nie zaobserwowano wyraźniej różnicy w parametrze inYFP między szczepem z nadekspresją *SH* a kontrolą, z wyjątkiem szczepu z nadekspresją *YET3*, dla którego obserwowano spadek w produkcji wewnątrzkomórkowej frakcji YFP w obu zastosowanych temp. (Fig.1.H.1 w **P.3**).

Najlepsze wyniki produkcji/sekrecji scYFP odnotowano dla szczepów *Y. lipolytica* z nadekspresją *RPL3*, *SSA5*, *SSA8* hodowanych w temp. 25 °C (Fig. 1.B/C/D.2 w **P.3**). Z kolei, efekt ten nie był widoczny w temp. 30 °C. Dodatkowo, niewielką poprawę w parametrze exYFP zaobserwowano również dla szczepu ze zwiększoną ekspresją *SSO1* (Fig. 1.K.2 w **P.3**) oraz w mniejszym stopniu – *HAC1*, *PDII*, *SLS1*, *CNE1*, *USO1*, *SEC1*, *CWP11* (**Tabela 4**). **W większości analizowanych wariantów do podwyższenia poziomu exYFP wymagane było jednoczesne działanie obniżonej temp. i nadekspresji *SHs* ($p < 0.05$).** W przypadku szczepów z nadekspresją *SEC1*, *YET3*, *SLS1*, *PDII* i *USO1* nie stwierdzono pozytywnego wpływu na wartości

bezwzględne exYFP, gdy szczepy hodowano w temp. 30 °C. Co więcej, nadekspresja *USO1* w połączeniu z hodowlą w 30 °C przyczyniła się do uzyskania niższych wartości exYFP ($p < 0.05$); taką zależność zaobserwowano także dla *PDI* i *SLS1* (**Tabela 4**).

Tabela 4. Wzrost (OD600 nm), bezwzględna wartość fluorescencji zewnątrzkomórkowej (exYFP [FU]) i specyficzna fluorescencja (sp_exYFP [FU OD600nm⁻¹]) kotransformantów *Y. lipolytica* w hodowlach okresowych w temp. 25°C lub 30°C w stacjonarnej fazie wzrostu (72 h). Wartości zmiany krotności (FC; *Fold change*) obliczono w odniesieniu do szczepu kontrolnego (Control_scYFP) z nadekspresją wyłącznie *SCYFP*. Liczby podane są we wskazanych jednostkach ± SD z duplikatu biologicznego, każdy mierzony w duplikacie technicznym.

Szczep <i>Y. lipolytica</i> z nadekspresją SH	72h									
	25 °C					30 °C				
	OD 600nm	exYFP [FU]	FC [exYFP]	sp_exYFP [FU OD600nm ⁻¹]	FC [sp_exYFP]	OD 600nm	exYFP [FU]	Fold change [exYFP]	sp_exYFP [FU OD600nm ⁻¹]	Fold change [sp_exYFP]
Control_scYFP	8.72 ± 0.00	11780.50 ± 252.44	1.00	1352.53 ± 26.06	1.00	5.62 ± 0.15	7572.50 ± 103.94	1.00	1348.60 ± 16.81	1.00
scYFP_HAC1	8.02 ± 0.24	14569.75 ± 264.81	1.24	1818.34 ± 20.21	1.34	5.24 ± 0.02	9610.50 ± 552.25	1.27	1834.90 ± 97.52	1.36
scYFP_RPL3	8.44 ± 0.04	22542.50 ± 2281.83	1.91	2673.97 ± 284.69	1.98	5.79 ± 0.16	11557.00 ± 141.42	1.43	1995.94 ± 80.20	1.48
scYFP_SSA5	8.83 ± 0.52	19294.25 ± 930.20	1.64	2193.26 ± 237.35	1.62	6.13 ± 0.03	10712.00 ± 869.74	1.42	1747.53 ± 132.04	1.30
scYFP_SSA8	7.69 ± 0.19	23576.25 ± 720.19	2.00	3067.31 ± 20.51	2.27	5.38 ± 0.14	10647.00 ± 379.01	1.41	1980.28 ± 123.52	1.47
scYFP_PDI	8.49 ± 0.24	15305.25 ± 579.474	1.30	1804.33 ± 15.72	1.33	5.45 ± 0.53	7235.00 ± 1350.57	0.96	1322.81 ± 119.70	0.98
scYFP_SLS1	8.19 ± 0.04	14784.50 ± 591.85	1.26	1804.26 ± 63.05	1.33	5.96 ± 0.10	7012.00 ± 397.39	0.93	1176.21 ± 46.17	0.87
scYFP_CNE1	8.75 ± 0.38	15253.50 ± 551.54	1.30	1745.04 ± 13.90	1.29	5.84 ± 0.55	8891.00 ± 12.73	1.17	1530.60 ± 147.40	1.14
scYFP_YET3	8.09 ± 0.26	13537.75 ± 172.18	1.15	1675.45 ± 32.31	1.24	5.31 ± 0.03	7154.50 ± 17.99	0.95	1346.05 ± 16.56	1.00
scYFP_USO1	8.89 ± 0.12	13901.00 ± 869.03	1.18	1564.42 ± 118.57	1.16	5.61 ± 0.43	4949.50 ± 1441.79	0.65	875.56 ± 189.38	0.65
scYFP_SEC1	8.48 ± 0.01	13626.75 ± 658.67	1.16	1607.63 ± 75.05	1.19	5.81 ± 0.03	8625.00 ± 90.51	1.14	1483.34 ± 8.89	1.10
scYFP_SSO1	7.92 ± 0.21	17724.50 ± 560.74	1.51	2243.24 ± 128.82	1.66	6.00 ± 0.11	13511.00 ± 593.97	1.78	2317.25 ± 239.95	1.67
scYFP_CWP11	8.80 ± 0.71	14741.75 ± 590.79	1.25	1681.11 ± 69.87	1.24	6.02 ± 0.06	10269.50 ± 301.94	1.36	1705.45 ± 34.12	1.27

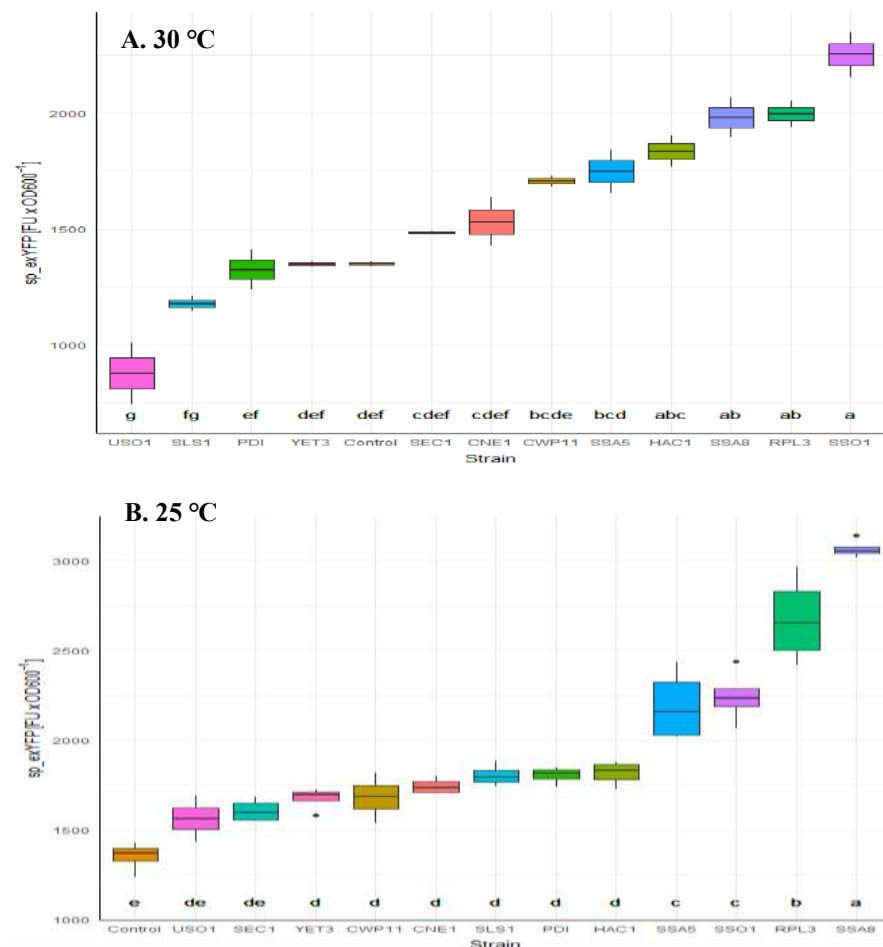
4.3.3. Szczegółowa analiza wpływu nadekspresji SHs na zdolność wydzielniczą komórek *Y. lipolytica*

Dokładna analiza profili syntezy (in-YFP) i sekrecji (ex-YFP), a także wzrostu biomasy (OD600nm) wskazała, że wszystkie procesy osiągnęły swoje plateau przy ~ 48 h (wzrost biomasy) do ~ 72 h (in-/exYFP). Dlatego wyniki z tych punktów czasowych były brane pod uwagę w dalszych analizach. Aby ocenić rzeczywisty wpływ zastosowanej temperatury i wprowadzonej modyfikacji genetycznej na zdolność wydzielniczą komórki (a nie na jej wzrost) w dalszych analizach zastosowano fluorescencję specyficzną (sekrecję specyficzną; znormalizowana zdolność wydzielnicza) sp_{exYFP} wyrażoną w [FU OD600 nm⁻¹]) ze wskazanego zakresu plateau dla dokonania porównań pomiędzy szczepami.

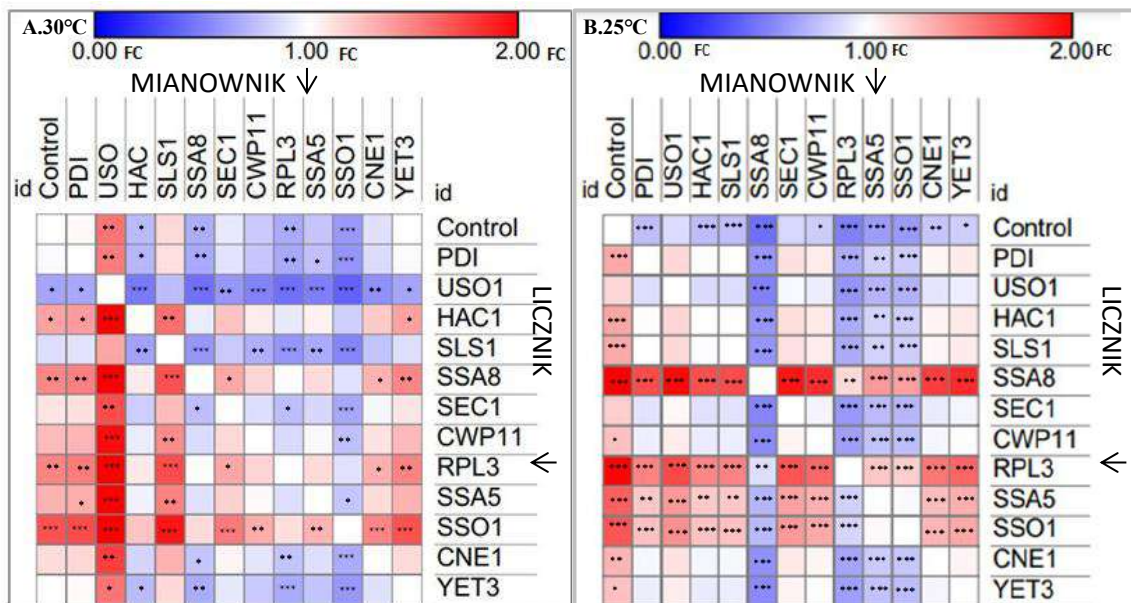
Rycina 11 ilustruje wyniki testu Tukey'a HSD klasyfikującego analizowane transformanty pod względem ich znormalizowanej zdolności wydzielniczej (sp_{exYFP}) w dwóch warunkach temperaturowych (30 °C (A.) 25 °C (B.)) w 72h hodowli. Dane liczbowe przedstawiono w **Tabeli 4**. Niezależnie od zastosowanej temp. hodowli nadekspresja *SSA8*, *SSA5*, *RPL3*, *SSO1* i *HAC1* przyczyniła się do maksymalizacji parametru sp_{exYFP} (**Rycina 11 A.B; Tabela 4**). Uzyskane wartości liczbowe sp_{exYFP} dla powyższych szczepów były wyższe w temp. 25 °C (scYFP_SSA8 1.5-krotnie, scYFP_RPL3 1.34-krotnie, scYFP_SSA5 1.25-krotnie), z kilkoma wyjątkami. W temp. hodowli 30 °C, nadekspresja *SEC1*, *CNE1* i *CWP11* wpłynęła pozytywnie na wartość sp_{exYFP} (10-27 %), natomiast poziom poprawy nie był istotny statystycznie ($p > 0.05$). Przeciwnie, nadekspresja *USO1* spowodowała prawie 40 % obniżenie parametrów exYFP oraz sp_{exYFP} podczas hodowli w temp. 30 °C. Ponadto, jest to jedyny przypadek, w którym wprowadzona modyfikacja genetyczna wywołała negatywny wpływ na zdolność wydzielniczą komórki w porównaniu do szczepu kontrolnego i do większości szczepów z nadekspresją *SH* ($p < 0.05$) (**Rycina 12**).

Hodowle zmodyfikowanych szczepów *Y. lipolytica* w obniżonej temperaturze (25 °C) spowodowały istotne zmiany w rankingu SH (**Rycina 11B**). W tym przypadku jakakolwiek nadekspresja SH miała (w określonym stopniu) pozytywny wpływ na znormalizowaną zdolność wydzielniczą *Y. lipolytica* ($p < 0.05$) (**Rycina 12**). Zastosowanie niższej temperatury wywołało swoistą polaryzację wyników, obserwowaną jako wyraźniejsze rozróżnienie grup jednorodnych, określające najkorzystniejsze modyfikacje genetyczne (**Rycina 11**). W tych warunkach można było zaobserwować jednoznaczną przewagę szczepu z nadekspresją *SSA8* nad kontrolą (~2.3-krotna poprawa w sp_{exYFP} ; **Tabela 4**) oraz pozostałymi szczepami

(**Rycina 12**). Podobnie w przypadku szczepu scYFP_RPL3 (~100% poprawa w porównaniu z kontrolą; **Tabela 4**); co skutkowało utworzeniem jednorodnej grupy dla tych dwóch szczepów. Podobnie, szczepy z nadekspresją *SSO1* i *SSA5* sklasyfikowano razem (**Rycina 11**); także jako grupę wykazującą znaczną przewagę nad szczepem kontrolnym (poprawa ~ 60%; **Tabela 4**) i pozostałymi ośmioma szczepami z nadekspresją *SH* ($p < 0.05$; **Rycina 11**; **Rycina 12**). Dalsze szczegółowe informacje na temat zmiany krotności parametru sp_exYFP dla wszystkich szczepów w dwóch punktach czasowych (48 h i 72 h) oraz dwóch temperaturach (25 °C i 30 °C) podano w mat. uzup. na Fig S3 w **P.3**.



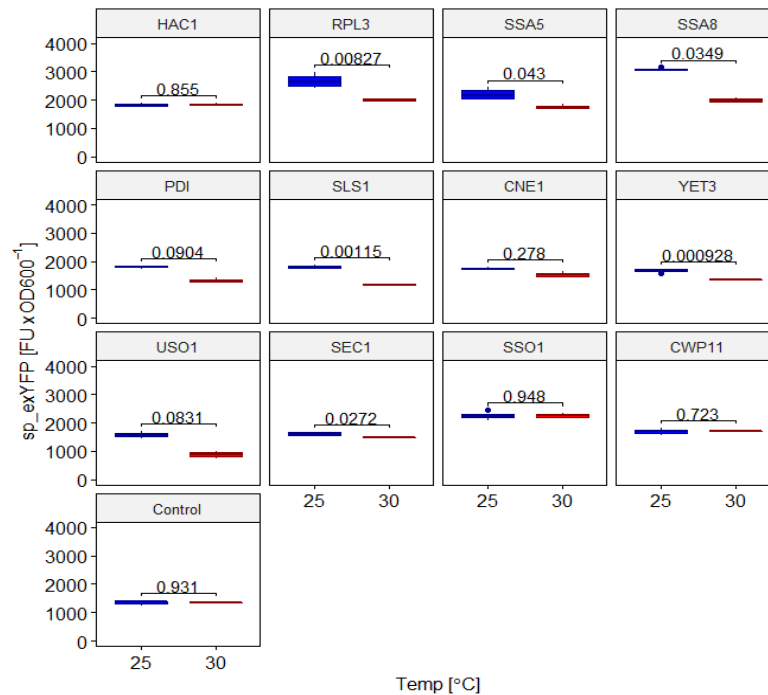
Rycina 11. Specyficzna fluorescencja sp_exYFP [FU x OD600⁻¹] kotransformantów oraz kontroli *Y. lipolytica* w hodowlach okresowych w temp. 30°C (A.) lub 25°C (B.) oznaczona w stacjonarnej fazie wzrostu (72 h). Oś X: Szczep *Y. lipolytica*, oś Y: specyficzna fluorescencja (sp_exYFP) wyrażona w [FU x OD600⁻¹]. Litery oznaczają grupy homogenne obliczone testem Tukeya HSD.



Rycina 12. Graficzne porównanie rekombinowanych szczepów *Y. lipolytica* z ko-nadekspresją SCYFP i SH oraz szczepu kontrolnego (+scYFP/-SH) zaprezentowane jako krotność zmiany (FC) pomiaru specyficznej sekrecji sp_exYFP pomiędzy szczepami w temperaturze 30 °C (A.) oraz w 25 °C (B.) w 72h hodowli. Czerwone kwadraty oznaczają przewagę licznika w przyjętej mierze (1.01–2.00 FC), a niebieskie kwadraty przewagę mianownika w przyjętej mierze (0 -0.99 FC). Istotność statystyczną porównań określono testem Tukeya HSD. *** $p \leq 0.001$ ** $p \leq 0.01$ * $p \leq 0.05$. Wykres został przygotowany przy użyciu narzędzia internetowego Morpheus (<https://software.broadinstitute.org/morpheus/>).

4.3.4. Wpływ temperatury na działanie SHs

W celu szczegółowego zbadania bezpośredniego wpływu temperatury na działanie poszczególnych SHs analizowano parametr sp_exYFP w 72h hodowli. Wyniki bezpośredniego porównania z zastosowaniem testu statystycznego t-Studenta przedstawiono na **Rycinie 13**. Co ważne, wykazano, że dla połowy badanych tutaj SH, **obniżona temperatura przyczyniła się do wzrostu potencjału wydzielniczego sp_exYFP** w sposób istotny ($p < 0.05$). Warto zauważyć, iż to interakcja między nadekspresją SH a temperaturą przyczyniła się do wzrostu w parametrze sp_exYFP, bowiem takiego efektu nie zaobserwowano dla szczepu kontrolnego ($p = 0.931$) i kilku innych szczepów z nadekspresją SH hodowanych w temperaturze 25 °C. **Tę korzystną interakcję wskazano dla wszystkich białek cytoplazmatycznych - Rpl3, Ssa5, Ssa8, ale także Sls1 rezydującego w ER, oraz dwóch SH działających w transporcie pęcherzykowym - Yet3 i Sec1.** Pewną nieistotną tendencję występowania tej korzystnej interakcji zaobserwowano również dla Pdi1 i Uso1 ($p < 0.1$).



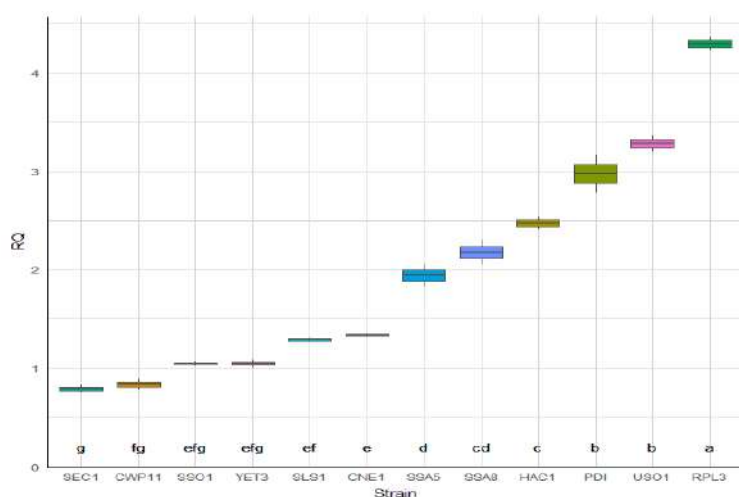
Rycina 13. Graficzne porównanie wartości specyficznej sekrecji sp_exYFP [$FU \times OD600^{-1}$] uzyskane w 72 h hodowli rekombinowanych szczepów *Y. lipolytica* (+scYFP/+SH) oraz szczepu kontrolnego (+scYFP/-SH) w temperaturze 25 °C (niebieskie boxy) oraz w 30 °C (czerwone boxy). Oś X: temperatura hodowli: 25 – 25 °C, 30 – 30 °C. Oś Y: specyficzna sekrecja YFP (sp_exYFP) wyrażona w [$FU \times OD600^{-1}$]. Liczby oznaczają wartość p testu t-Studenta.

4.3.5. Analiza poziomu ekspresji heterologicznego genu *SCYFP* w warunkach nadekspresji SHs

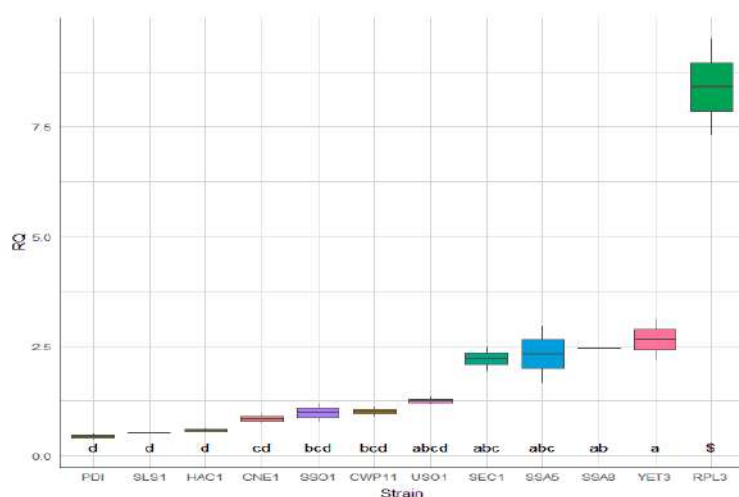
Aby uzyskać wgląd w podstawy molekularne obserwowanych makroskopowo zmian w syntezie i sekrecji scYFP przeprowadzono porównawczą analizę poziomu ekspresji genu heterologicznego dla wszystkich rekombinowanych szczepów *Y. lipolytica* hodowanych w temp. 30 °C i 25 °C (**Rycina 14A.B**). Biologiczny sens tego eksperymentu polegał na sprawdzeniu, czy podwyższona ekspresja *SH* wpłynęła na transkrypcję YFP. Przede wszystkim, nadekspresja *RPL3* znacząco wpłynęła na ekspresję *SCYFP* w obu badanych temp. hodowli (istotnie w stosunku do kontroli $p < 0.05$; nieistotnie między temp. **Rycina 15**). Ponadto, w obniżonej temp. hodowli, ekspresja *SCYFP* była dwukrotnie większa (**Rycina 14B**). Podobną zależność zaobserwowano w warunkach nadekspresji *YET3* i *SEC1* ($p < 0.05$) w obniżonej temp. (**Rycina 14B**), natomiast w temp. hodowli równej 30 °C, taki rodzaj wprowadzonej modyfikacji nie miał wpływu na ekspresję *SCYFP* (**Rycina 14A**; **Rycina 15**; nieistotny w porównaniu z kontrolą i pomiędzy temp. przy $p < 0.05$). W przypadkach genów *SSA5* i *SSA8* o największym wpływie na wartości inYFP / exYFP, ich nadekspresja

przyczyniła się do znacznie podwyższonej ekspresji *SCYFP* w obu temp. hodowli. Z drugiej strony, nadekspresja kolejnego „wydajnego SH” *SSO1* nie wpłynęła na transkrypcję genu *SCYFP* (**Rycina 14A.B**). Co ciekawe, nadekspresja *HAC1*, *PDI*, *SLS1* wywołała znaczny wzrost ekspresji *SCYFP* ($p < 0.05$), gdy szczepy hodowano w temp. 30 °C (**Rycina 14A**; tendencja obserwowana także dla *CNE1*, ale $p > 0.05$). W rzeczywistości, nadekspresja tych trzech *SHs* wywarła wpływ na poziom ekspresji *YFP* w sposób zależny od temp. (**Rycina 15**). Taką tendencję zaobserwowano także dla genu *USO1* ($p < 0.05$). W przeciwieństwie do tego, dla szczepów z nadekspresją *RPL3*, *SEC1*, *YET3*, stwierdzono wyższą ekspresję (nieistotne) i wydzielanie (istotne przy $p < 0.05$) w obniżonej temp. hodowli.

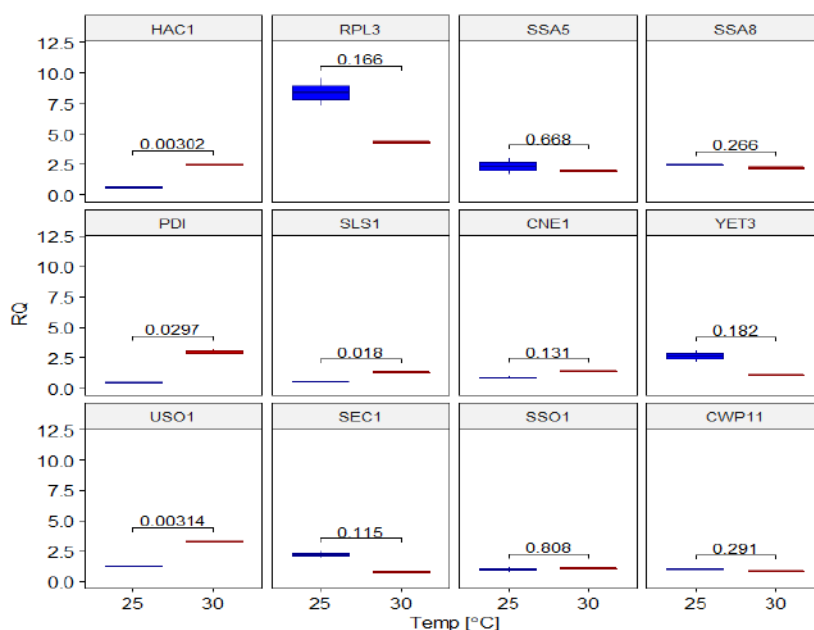
A. 30 °C



B. 25 °C



Rycina 14. Poziomu ekspresji genu *SCYFP* w szczepach *Y. lipolytica* z ko-nadekspresją *SCYFP* i *SH* hodowanych w temperaturze 30 °C (A.) lub 25 °C (B.). Oś X: Nazwa szczepu; Oś Y: RQ- poziom ekspresji genu *YFP* obliczony zgodnie z modelem ddCt. Wartość równa 1.0 wskazuje na brak zmian w poziomie ekspresji *YFP* pomiędzy szczepem kontrolnym a badanym szczepem z nadekspresją *SH*. Litery oznaczają grupy homogenne obliczone testem Tukey’a HSD.



Rycina 15. Graficzne porównanie poziomu ekspresji genu *SCYFP* w szczepach *Y. lipolytica* ko-nadekspresją *SCYFP* i *SH* hodowanych w temperaturze 30 °C (czerwone) oraz w 25 °C (niebieskie). Oś X: temperatura hodowli: 25 – 25 °C, 30 – 30 °C. Oś Y: RQ- poziom ekspresji genu YFP obliczony zgodnie z modelem ddCt. Wartość równa 1.0 wskazuje na brak zmian w poziomie ekspresji YFP pomiędzy szczepem kontrolnym a badanym szczepem z nadekspresją SH.

Aby uzyskać globalny wgląd w związek między: i) temperaturą, ii) poziomem ekspresji genu heterologicznego oraz iii) syntezą polipeptydu scYFP, przeprowadzono analizę korelacji Pearsona między całkowitą ilością wyprodukowanego YFP (suma FL dla inYFP + exYFP) i wartości RQ dla genu kodującego scYFP. Co ciekawe, stwierdzono, że taka korelacja była również zależna od temperatury, ponieważ współczynnik r dla wartości odczytanych dla szczepów hodowanych w 30 °C wynosił 0.397, podczas gdy dla szczepów hodowanych w temp. 25°C – $r = 0.855$, **wykazując dużą liniowość między transkrypcją i translacją w obniżonej temp.**

Ponadto, sprawdzono ekspresję genów kodujących białka zaangażowane w UPR (Hac1, Kar2 oraz Pdi1) dla szczepu z nadekspresją genu kodującego czynnik transkrypcyjny Hac1 (*scYFP_HAC1*) w obu temp. hodowli. Białka Kar2 i Pdi1 są uznawane za regulowane przez czynnik transkrypcyjny Hac1 w *S. cerevisiae* i *P. pastoris* [55,154–156]. Z uzyskanych danych wywnioskowano, że zwiększona ilość natywnego transkryptu *HAC1* nie miała wpływu na poziom ekspresji *KAR2* i *PDII*. Natomiast, podwyższony poziom transkryptu *HAC1* powodował zwiększoną ekspresję *YFP*, ale tylko w temp. 30 °C, zgodnie z omówionym już wzorcem ekspresji zależnym od temperatury (Fig.7 w **P.3**).

4.4. Profil transkrypcyjny rekombinowanego szczepu *Y. lipolytica* z ko-nadekspresją *HAC1* i *SCYFP* – transkryptomoczną analizą porównawczą (P.4)

4.4.1. Synteza i sekrecja YFP w szczepach *Y. lipolytica* utrzymywanych w stanie ustalonym (P.2: *inYFP*, *scYFP*, P.4: *HAC1_scYFP*)

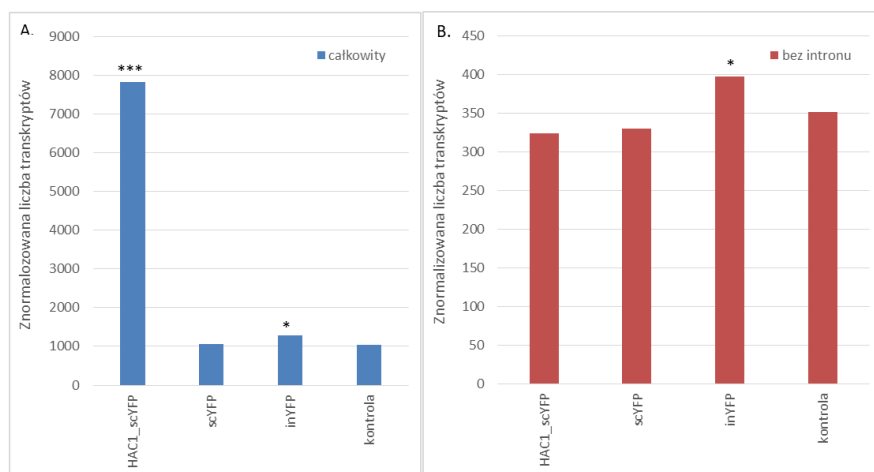
Aby uzyskać wgląd w molekularny mechanizm działania nadekspresji *HAC1* na zaobserwowaną zwiększoną produkcję polipeptydów (P.3), przeprowadzono hodowle ciągle w chemostacie komórek nadekspresjonujących *HAC1* i *SCYFP*, z których następnie izolowano całkowity RNA i sekwencjonowano (RNAseq). Kinetykę wzrostu, wykorzystania glicerolu i produkcji metabolitów w trakcie hodowli przedstawiono w mat. uzup. Fig. S3 w P.4. Ilości YFP syntetyzowane w warunkach stanu ustalonego zaprezentowano w Tabeli 5. Co znamienne, ko-nadekspresja *HAC1* spowodowała prawie 7-krotny spadek akumulowanej frakcji YFP, na korzyść wydzielania poza komórkę (*scYFP* vs. *HAC1_scYFP*). Poziom zewnątrzkomórkowej FL generowany przez *HAC1_scYFP* był prawie 2,5-krotnie wyższy niż w *scYFP* [kRFU L⁻¹].

Tabela 5. Ilości YFP syntetyzowane w warunkach stanu ustalonego przez rekombinowane szczepy *Y. lipolytica* ujęte w tym eksperymencie. Jednostki podane w nagłówkach ± SD.

Frakcja Szczep	Poziom ekspresji <i>HAC1</i> [FC]	biomasa	supernatant	biomasa	supernatant	Bibliografia
		in-YFP [kRFU L ⁻¹]	sc-YFP [kRFU L ⁻¹]	sp_in-YFP [kRFU OD ₆₀₀ ⁻¹]	sp_sc-YFP [kRFU OD ₆₀₀ ⁻¹]	
<i>inYFP</i>	1.23	130 809.60 ± 5 201.20		42 726.32 ± 3635.71		P.2
<i>scYFP</i>	1.0	812 108.30 ± 5 715.62	24 785.28 ± 187.33	257 051.30 ± 23 576.03	7 899.84 ± 1 260.26	
<i>HAC1_scYFP</i>	21.0	118 922.90 ± 3 224.39	60 123.13 ± 847.64	18 736.20 ± 923.66	9 415.35 ± 343.21	Ta praca

Ponadto, celem zbadania, jaka frakcja transkryptu nadekspresjonowanego *HAC1* uległa aktywującemu splicingowi, przeprowadzono mapowanie miejsc splicingowych na podstawie częstości występowania odczytów dla danego regionu genu. Mapowanie przeprowadzono *de novo*, ponieważ *HAC1* pochodzący z genomu referencyjnego nie posiadał tej cechy. Pomimo wysokiej nadreprezentacji znormalizowanej ilości transkryptów *HAC1* w szczepie *HAC1_scYFP* (Rycina 16A), obliczone wartości liczby transkryptów formy

splicingowanej były porównywalne z poziomem tych transkryptów w komórkach szczepu kontrolnego i szczepu *scYFP* (**Rycina 16B**).

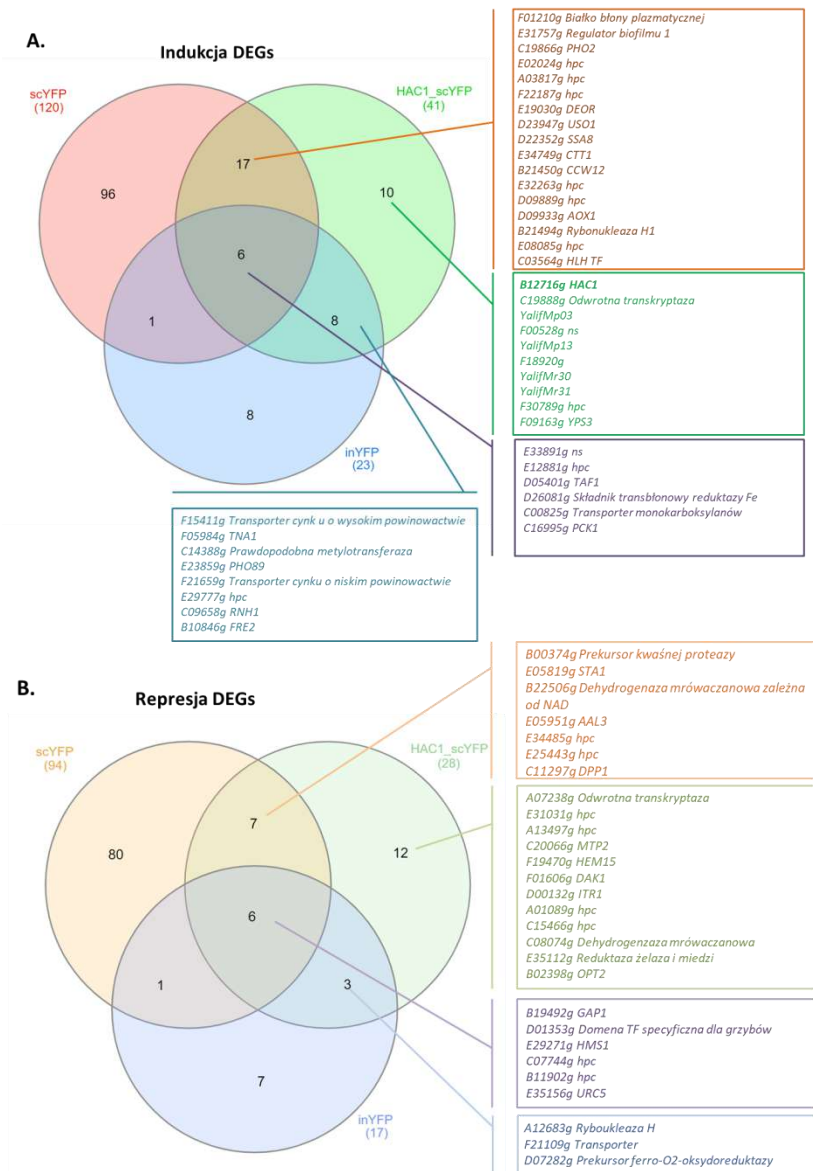


Rycina 16. Oznaczenie ilościowe całkowitej liczby transkryptów *HAC1* oraz liczby transkryptów pozbawionych intronu w transkryptomach analizowanych szczepów *Y. lipolytica*. Transkrypty *HAC1* zsekwencjonowane i określone ilościowo za pomocą analizy RNAseq zmapowano pod kątem występowania eksonów i intronów. Suma oznacza znormalizowaną liczbę wszystkich transkryptów w oparciu o liczbę eksonów. Bez intronu oznacza znormalizowaną liczbę transkryptów pozbawionych intronu. *, *** oznacza istotność statystyczną przy $p < 0.05$ i 0.001 .

4.4.2. Silnie deregulowane DEGs w warunkach nadekspresji *HAC1* i *rs-YFP*

Jednoczesna nadekspresja genów *HAC1* i *SCYFP* spowodowała istotną indukcję (316) / represję (237) genów w transkryptomie w stosunku do prototroficznego szczepu kontrolnego (mat. uzup. Fig. S6 w **P.4**). Wśród najbardziej indukowanych DEGs (> 2 FC) zidentyfikowano wiele genów zaangażowanych w transbłonowy transport w tym: *F15411g*, *TAF1/D05401*, *D26081g*, *FRE2/B10846g*, *PHO89/E23859g*, *C00825g* *F21659g*, *A07623g* i *E12881g*, a także dwie rybonukleazy (*RNH1/C09658g*, *E12331g*) i białko zawierające domenę odwrotnej transkryptazy (*C19888g*); co wskazuje na intensywny metabolizm RNA w *HAC1_scYFP*. Co ważne, wszystkie spośród w/w genów ulegały indukcji w szczepie *inYFP* z natywnie indukowaną ekspresją *HAC1* (**Rycina 17A**). Z drugiej strony, nadekspresję genów: *F15411g*, *TAF1*, *D26081g*, *FRE2*, *PHO89*, *C00825g* czy *E12881g* zaobserwowano także w szczepach HSS (*scYFP* i *SoA*), z kontrolnym poziomem ekspresji *HAC1*, ale represją *SLS1*. Z kolei, nadekspresję genów kodujących transportery (*F21659g*, *A07623g*) oraz białka uczestniczące w metabolizmie RNA (*RNH1*, *E12331g*, *C19888g*) wskazano w warunkach nadekspresji *HAC1* i braku deregulacji *SLS1* (**Rycina 17A**; **Rycina 18**). W grupie DEGs ulegających silnej indukcji w *HAC1_scYFP* znaleziono także geny kodujące: akwaporynę

(*F01210g*), regulator biofilmu (*E31757g*), fosfatazę (*PHO2/C19866g*), czynnik transkrypcyjny typu DeoR (*E19030g*), tioltransferaza rezydującą w ER (*F22187g*), alternatywną oksydazę (*AOX1/D09933g*) i katalazę (*CTT1/E34749g*) (**Rycina 17A**). Wszystkie te geny były istotnie indukowane w odpowiedzi na nadprodukcję rs-Prot, w co najmniej jednym z analizowanych szczepów HSS (*scYFP* lub *SoA*); stąd, ich indukcja w szczepie *HAC1_scYFP* wynika raczej z intensywnej syntezy *scYFP* niż z indukcji *HAC1*.



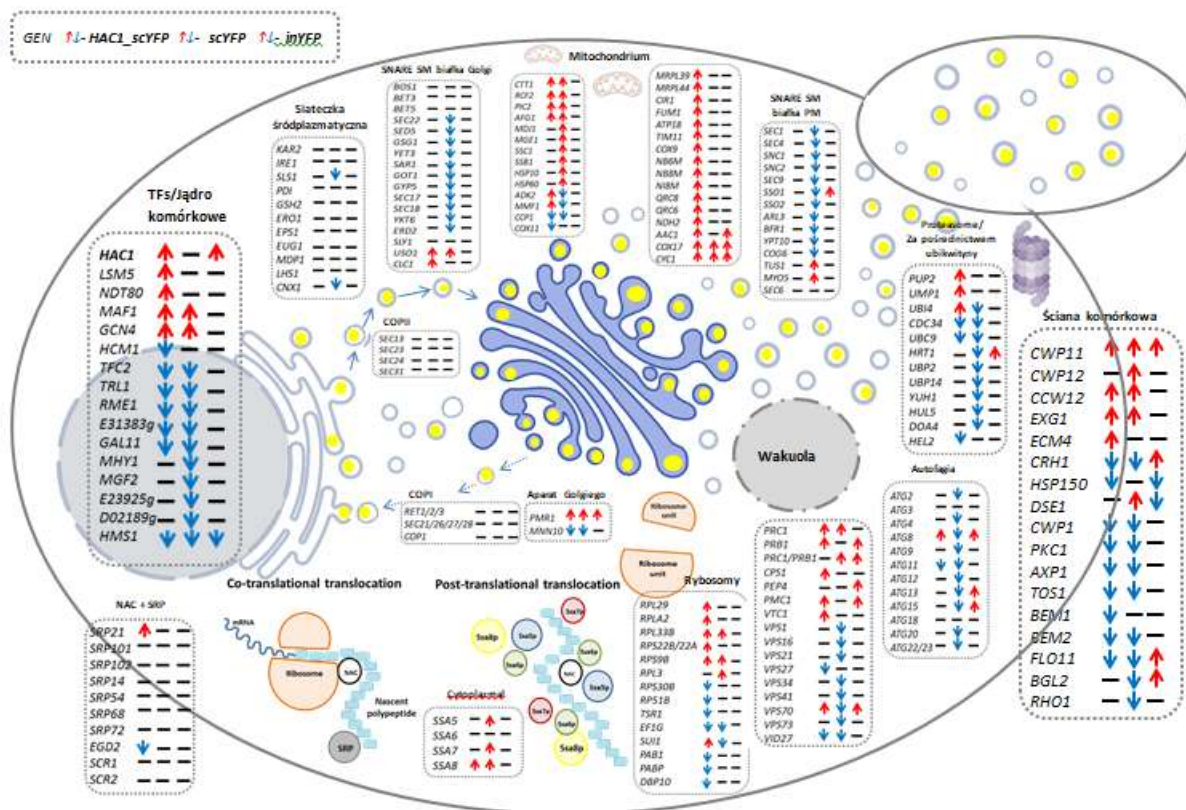
Rycina 17. Indukowane i represjonowane wspólne lub specyficzne zestawy genów (DEGs) dla poszczególnych szczepów *Y. lipolytica* (*inYFP*, *scYFP* i *HAC1_scYFP*) w odpowiedzi na nadprodukcję r(s)-Prots. Wskazano liczbę DEGs: A. Indukowanych w *inYFP*, *scYFP* i *HAC1_scYFP*, B. Represjonowanych w *inYFP*, *scYFP* i *HAC1_scYFP*. Analizę diagramów Venna przeprowadzono z listami DEGs ograniczonymi do log2-krotnej zmiany ekspresji genu próby badanej względem kontroli > 1 i wartości $p < 0.05$. Skróty: TF – czynnik transkrypcyjny, hpc – hipotetyczne białko w kładzie *Yarrowia*, ns – brak podobieństwa.

W grupie DEGs o najbardziej obniżonym poziomie ekspresji w szczepie *HAC1_scYFP* zidentyfikowano: czynnik transkrypcyjny *D01353g* o niepoznanej funkcji, a także odwrotną transkryptazę (*A07238g*; wykazującą odwrotny wzór ekspresji do *C19888g* mającego tę samą funkcję) i *URC5/E35156g* zaangażowany w katabolizm uracylu (mat. uzup. Tabela S3 w **P.4**). Ten sam czynnik transkrypcyjny *D01353g* zidentyfikowano wśród najbardziej deregulowanych DEGs ulegających represji we wszystkich szczepach *Y. lipolytica* syntetyzujących jakiegokolwiek z badanych rs-Prot. Dokładne poszukiwania funkcji *D01353g* ujawniły, że wykazuje pewne podobieństwa z represorem polimerazy RNA II. Wydaje się, więc, że jego represja uwalnia polimerazę RNA II z jej hamującej aktywności. Ograniczenie katabolizmu uracylu zaobserwowano zarówno w szczepach z nadekspresją *HAC1*, jak i w szczepie *scYFP* (**Rycina 17B**). Uzyskane dane wykazały, że transkryptomy szczepów *HAC1_scYFP* i *inYFP* zawierały wspólne zestawy genów ulegających represji, w tym: endonukleazę RNA (*A12683g*) specyficznie degradująca RNA, permeazę jonów żelaza (*FTR1/D07304g*) oraz permeazę kwasu nikotynowego (*F21109g*). Podobną tendencję zaobserwowano dla oksydoreduktazy Fe-O₂ (*D07282g*) i transportera glukozy o wysokim powinowactwie (*SNF3/C06424g*). Z drugiej strony zidentyfikowano grupę genów, które ulegały represji we wszystkich szczepach intensywnie syntetyzujących jakiegokolwiek r(s)-Prot, w tym: *AAL3/E05951g*, czynnik transkrypcyjny związany z Cup9 regulujący polimerazę RNA II (*E29271g*), permeazę aminokwasową (*GAP1/B19492g*) i flokulinę (*STAI/E05819g*).

Dla zestawu DEGs deregulowanych w warunkach ko-nadekspresji *HAC1* i *SCYFP* dokonano analizy wzbogacenia w procesy biologiczne, funkcje molekularne i składniki komórkowe. Przeprowadzona klasyfikacja funkcjonalna (Fig. 4 w **P.4**) deregulowanych DEGs wykazała, że istotnie „wzbogaconym składnikiem komórkowym” była błona plazmatyczna (EF: 2.37) i jej integralne składniki (EF: 1.49) reprezentowane przez ponad 100 genów ulegających indukcji. Z drugiej strony, we wszystkich trzech analizowanych kategoriach (proces biologiczny, funkcja molekularna, składnik komórkowy) zidentyfikowano kompleks dehydrogenazy mrówczanowej ulegający znacznej represji. Wśród istotnie wzbogaconych funkcji molekularnych indukowanych w zadanych warunkach wskazano: aktywność oksydoreduktazy (EF: 2.07) i transportery jonów (EF: 8.72/11.42). Podobnie, wiele procesów biologicznych związanych z transportem i importem jonów do komórki ulegało indukcji w szczepie *HAC1_scYFP*. Wytwarzanie energii, cykl kwasu trikarboksylowego (TCA) i jednowęglowe procesy metaboliczne także wskazano jako procesy biologiczne istotnie indukowane (Fig. 4 w **P.4**). W tej ostatniej grupie

zidentyfikowano wiele genów zaangażowanych w syntezę i konwersję AA (mat. uzup. (nośnik elektroniczny) Tabela S3 w **P.4**). Obserwacja ta dobrze koresponduje ze skoordynowaną indukcją genów zlokalizowanych w mitochondriach zaangażowanych w energetykę komórkową, co było wyróżniające dla szczepu *HAC1_scYFP* (**Rycina 18**; mat. uzup. (nośnik elektroniczny) Tabela S4 w **P.4**). Na intensywną przebudowę struktur komórkowych wskazywała indukcja ponad 30 DEGs przypisanych do organizacji i biogenezy składników komórkowych, jak również represja > 10 DEGs zaangażowanych w składanie struktur komórkowych (Fig. 4 w **P.4**). Dokładna analiza DEGs przypisanych do tych kategorii wykazała, że są to w znacznej mierze geny zlokalizowane w mitochondriach, zaangażowane w transport pęcherzykowy i restrukturyzację chromatyny (specyficzne DEGs można znaleźć w mat. uzup. (nośnik elektroniczny) Tabeli S3 w **P.4**).

W związku z istotną deregulacją DEGs kodujących białka podjednostek rybosomalnych, czy eukariotycznego czynnika inicjującego translację *SUI1/C07524g* wnioskowano o zajściu istotnych zmian w procesie syntezy białek (**Rycina 18**). Z kolei metabolizm RNA i ekspresja genów, a także biosynteza makrocząsteczek były procesami biologicznymi ulegającymi represji po ko-nadekspresji *HAC1* i *SCYFP*. Specyficzne DEGs o obniżonej ekspresji sklasyfikowane do tych kategorii to m.in: białka rybosomalne *RP51B/F14465g*, *RPS30B/F14663g*, białko przetwarzające pre-rRNA *TSR1/B08756g*, czynnik EF-1-gamma *C16049g*, a także ligaza tRNA *TRL1/D14916g* zaangażowana w biogenezę tRNA.



Rycina 18. Graficzne przedstawienie wzorców ekspresji wybranych genów istotnych dla syntezy i wydzielania r(s)-Prot. Czerwone strzałki (↑) istotnie zwiększona ekspresja genu w stosunku do szczepu kontrolnego; niebieskie strzałki (↓) istotnie obniżona ekspresja genu w stosunku do szczepu kontrolnego; czarna kreska (–) brak istotnej zmiany poziomu ekspresji w stosunku do szczepu kontrolnego. Legenda wskazuje kolejność szczepów: lewa kolumna: *HACL_scYFP*; środkowa kolumna: *scYFP*; prawa kolumna: *inYFP*.

4.4.3. Procesy biologiczne o szczególnym znaczeniu dla produkcji rs-Protów

W oparciu o dane literaturowe, jak również wygenerowane dane transkryptomyczne, a także profil dystrybucji białka modelowego YFP (**Tabela 5**), zaproponowano trzy główne procesy biologiczne które przyczyniają się do kształtowania Hac1-zależnego fenotypu: i) szybkość syntezy r(s)-Protów, ii) degradacja białek oraz iii) transport pęcherzykowy przez szlak wydzielniczy. Dane transkryptomyczne dla genów szczególnie zaangażowanych w te procesy przedstawiono na **Rycinie 18**.

4.4.3.1. Regulacja ekspresji białek

Zestaw genów istotnie deregulowanych w warunkach ko-nadekspresji *HACL* i *SCYFP* zlokalizowano w jądrze (**Rycina 18**; mat. uzup. (nośnik elektroniczny) Tabela S4 w **P.4**).

Wiele ze zidentyfikowanych DEGs kodowało TFs o różnej specyficy. Na przykład *MAF1/F10541g* (indukcja o 82%) i *TFC2/F05104g* (represja o 25%), o znanym działaniu regulacyjnym na polimerazę RNA III. Z kolei, homolog ulegającego represji genu *E31383g*, uczestniczy w pozytywnej regulacji transkrypcji przez polimerazę RNA II. Podobnie gen *GCN4/E27742g*, który był indukowany w szczepach *HAC1_scYFP* i *scYFP* (**Rycina 18**; mat. uzup. (nośnik elektroniczny) Tabela S4 w **P.4**), pełni funkcję aktywatora ponad 30 genów zaangażowanych w biosyntezę AA i puryn w odpowiedzi na ich niedobór.

4.4.3.2. Degradacja białek

Kolejnym proponowanym mechanizmem, indukowanym przez nadekspresję *HAC1* jest kierowanie zakumulowanych, nieprawidłowo sfałdowanych r(s)-Prots do degradacji. Pomimo, iż zastosowane białko reporterowe (*scYFP*) jest małym białkiem, podlegającym modyfikacjom potranslacyjnym w niewielkim stopniu, nieakumulującym nadmiernie w komórce [157] (**Tabela 1** i **Tabela 5**), **zaobserwowano intensyfikację procesów degradacji w wakuoli (Rycina 18)**. Wśród genów klasyfikowanych w tej kategorii zidentyfikowano: proteazy *PRB1* (wzrost o 22% w stosunku do kontroli prototroficznej) i *PMC1/D04873g* (wzrost o 37%), czynnik sortujący *VSP70* (wzrost o 31%), oraz kluczowego regulatora autofagii - *ATG8* (wzrost o 22%). W przeciwieństwie do szczepu *HAC1_scYFP*, szczep *scYFP* charakteryzował się globalną represją genów zaangażowanych w którykolwiek ze szlaków degradacji białek (proteasom, autofagia/wakuola) (**Rycina 18**). Aby zweryfikować to stwierdzenie, zbadano całkowitą wewnątrzkomórkową aktywność proteolityczną badanych tutaj szczepów. W oparciu o wyniki testu (Fig. 6 w **P.4**), **jakakolwiek nadprodukcja r(s)-Prot prowadziła do zwiększonej całkowitej aktywności proteolitycznej**. Z drugiej strony, **nadekspresja *HAC1* przyczyniła się do ograniczenia proteolizy w porównaniu z pozostałymi dwoma szczepami (*inYFP* i *scYFP*), ale nadal była wyższa niż w szczepie kontrolnym**.

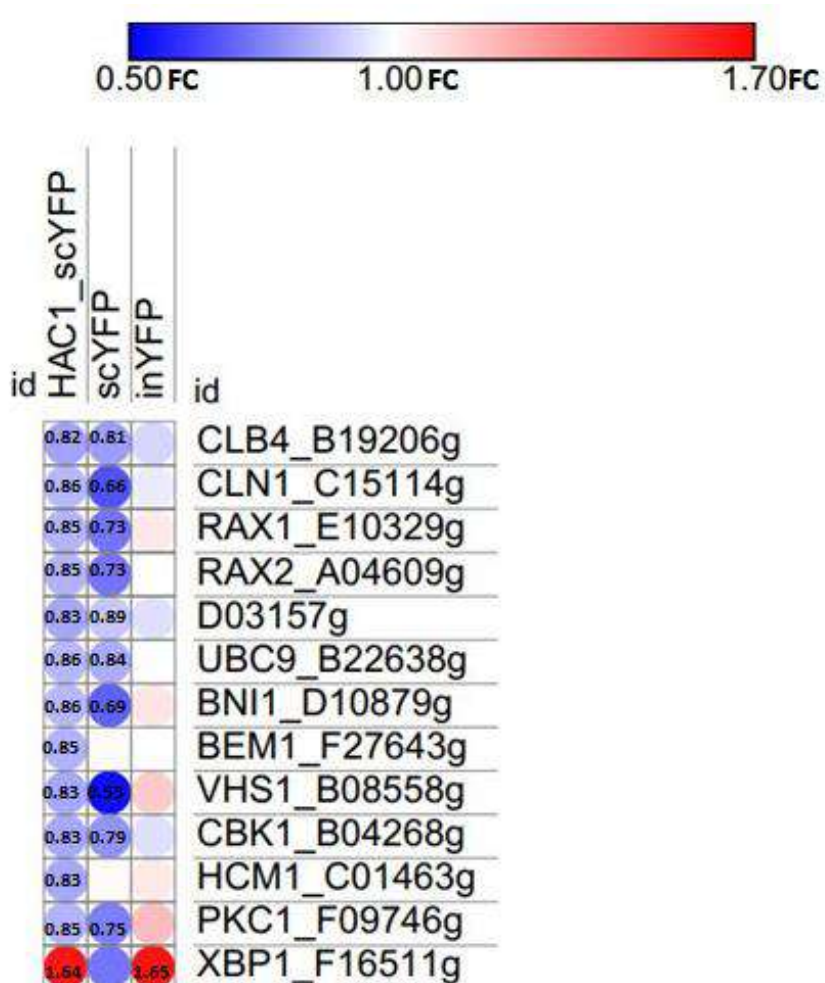
4.4.3.3. Transport pęcherzykowy

Ostatnim zaproponowanym mechanizmem, przez który *Hac1* realizuje ochronny program w stosunku do komórek *Y. lipolytica* nadprodukcujących r(s)-Prots, jest intensyfikacja transportu pęcherzykowego celem usunięcia nadmiaru polipeptydów z komórki. Choć założenie wydaje się być logiczne, to wszystkie geny z wyjątkiem dwóch (indukcja: *USO1/D23947g* i *CLC1/E33253g*) wykazywały kontrolny poziom ekspresji w szczepie

HAC1_scYFP. Natomiast, wiele genów biorących udział w tym procesie ulegało represji w szczepie wysoce nadprodukującym *scYFP* (**Rycina 18**).

4.4.4. Cykl komórkowy

Co ciekawe, w odpowiedzi na ko-nadekspresję *HAC1* i *SCYFP* zaobserwowano istotną **represję genów zaangażowanych w progresję cyklu komórkowego** (**Rycina 19**). Podobny wzorzec ekspresji tych genów zidentyfikowano dla szczepu *scYFP* (HSS). W grupie tej wskazano geny kodujące: Ser/Thr kinazy białkowe uczestniczące w podziale komórek i w morfogenezie (G2/M-specyficzne: *VHS1/B08558g*, *CBK1/B04268g*, *PKC1/F09746g*), dwie cykliny (G1/S-specyficzne: *CLB4/B19206g*, *CLN1/C15114g*), a także białko zaangażowane w specyficzną degradację cyklin fazy S i M (*UBC9/B22638g*). Ponadto, obniżony poziom ekspresji zaobserwowano dla genów: *HCM1/C01463g* (kodującego TF), *RAX1*, *RAX2*, *BNI/D10879g* i *BEM1/F27643g* biorących udział w cytokinezie, polaryzacji komórek; jak również białko wiążące GTP *D03157g* tworzące wrzeciono mitotyczne i przegrodę podziałową. Zgodnie z represją progresji cyklu komórkowego, zauważono znaczną indukcję genu *XBPI/F16511g* zarówno w szczepach *HAC1_scYFP*, jak i *inYFP*, w przeciwieństwie do jego obniżenia w szczepie *scYFP*.



Rycina 19. Graficzne przedstawienie deregulowanych DEGs zaangażowanych w progresję cyklu komórkowego w szczepach *HAC1_scYFP*, *scYFP* i *inYFP*. Indukcję/represję DEGs przedstawiono za pomocą kolorów, zgodnie z legendą, w oparciu o wartość ekspresji wyrażoną jako krotność zmiany (FC) ekspresji genu w próbce badanej względem kontroli. Wykres został przygotowany przy użyciu narzędzia internetowego Morpheus (<https://software.broadinstitute.org/morpheus/>).

5. Wykaz analiz

Analizy przeprowadzone przez Doktoranta

- 1) Przygotowanie zestawu rekombinowanych szczepów *Y. lipolytica*, pochodnych Po1g niosących heterologiczny gen kodujący białko żółtej fluorescencji (inYFP oraz scYFP) oraz szczepu kontrolnego, za pomocą komercyjnego zestawu do klonowania YLEX Kit.
- 2) Przygotowanie zestawu rekombinowanych szczepów *Y. lipolytica*, pochodnych Po1h niosących heterologiczny gen α -amylazy (*scSoA*), glukoamylazy (*scTIG*) lub YFP (*SCYFP* i *INYFP*) oraz szczepu kontrolnego, w oparciu o technologię klonowania modularnego Golden Gate,
- 3) Przygotowanie zestawu dwunastu rekombinowanych szczepów *Y. lipolytica*, pochodnych Po1f z ko-nadekspresją natywnego białka pomocniczego SH i heterologicznego genu *SCYFP*,
- 4) Hodowle rekombinowanych szczepów *Y. lipolytica* w kolbach wytrząsanych, skринing pod kątem przyrostu biomasy i nabytej aktywności amylolitycznej/zdolności do fluorescencji,
- 5) Ustalenie wartości szybkości rozcieńczania D, przy której układ osiągnął tzw. stan ustalony, poprzez wykonanie serii ciągłych hodowli bioreaktorowych szczepu kontrolnego (Po1h_Ura3),
- 6) Ciągłe hodowle bioreaktorowe rekombinowanych szczepów *Y. lipolytica* (pochodne Po1g i Po1h) w celu zbadania odpowiedzi komórkowej w warunkach nadprodukcji r(s)-Prots i przeprowadzenia globalnej analizy transkryptomu,
- 7) Hodowle okresowe w kolbach wytrząsanych rekombinowanych szczepów *Y. lipolytica* pochodnych Po1f z ko-nadekspresją SH i *SCYFP* w dwóch wariantach temperaturowych,
- 8) Ciągłe hodowle bioreaktorowe szczepu *Y. lipolytica* z ko-nadekspresją *HAC1* i *SCYFP* w celu zbadania odpowiedzi komórkowej i przeprowadzenia globalnej analizy transkryptomu,
- 9) Oznaczenie aktywności amylolitycznej i proteolitycznej za pomocą testów enzymatycznych w mikroskali: microSIT dla α -amylazy, microDNS dla glukoamylazy, microFTC-casein dla całkowitej aktywności proteolitycznej,

- 10) Oznaczenie wewnątrz-/zewnątrzkomórkowej frakcji YFP poprzez pomiar fluorescencji przy długości fali 495/527 nm,
- 11) Oznaczenie suchej masy metodą grawimetryczną oraz gęstości optycznej poprzez pomiar absorbancji przy długości fali 600nm,
- 12) Przygotowanie próbek do analizy HPLC,
- 13) Izolacja całkowitego RNA i reakcja odwrotnej transkrypcji,
- 14) Analiza poziomu ekspresji genów metodą qPCR w czasie rzeczywistym,
- 15) Analiza wyników uzyskanych z sekwencjonowania RNA przy użyciu narzędzi bioinformatycznych i baz danych.

Analizy przeprowadzone przez innych Współautorów

- 1) Przygotowanie rekombinowanych szczepów *Y. lipolytica* pochodnych Po1g niosących gen glukoamylazy (*Y. lipolytica* 2.2.1.2) i α -amylazy (*Y. lipolytica* 1.1.2.2) przez Prof. UPP dr hab. Ewelinę Celińską,
- 2) Analizy statystyczne przeprowadzone przy użyciu oprogramowania STATISTICA i Microsoft Excel 2013 dla P.1 przez Prof. UPP dr hab. Wojciecha Białasa oraz dr Piotra Kubiaka,
- 3) Analizy statystyczne przeprowadzone z wykorzystaniem programu R dla P.3 przez dr Piotra Kubiaka.

Analizy wykonane przez Pracownię Chromatograficzną Katedry Biotechnologii i Mikrobiologii Żywności UPP

- 1) Oznaczenie stężeń glicerolu i metabolitów za pomocą techniki HPLC – wykonanie analizy, identyfikacja pików i oznaczenie ilościowe metodą wzorca zewnętrznego.

Analizy zlecone

- 1) Sekwencjonowanie RNA wykonane w firmie Genomed S.A. z siedzibą w Warszawie,
- 2) Identyfikacja miejsc splicingowych w rejonie genu *HAC1* oraz oznaczenie ilościowe stosunku transkryptu *HAC1* bez intronu do transkryptu *HAC1* z intronem wykonane w firmie Genomed S.A. z siedzibą w Warszawie.

6. Dyskusja

W pierwszej części niniejszej pracy (**P.1**) heterologiczne sekwencje genów kodujących cztery r(s)-Prots wprowadzono do genomów dwóch szczepów *Y. lipolytica* (Po1h oraz Po1g), stanowiących dobrze scharakteryzowane platformy ekspresyjne [16,61,158]. Jak wspomniano wcześniej, kinetyka wytwarzania rs-Prots i odpowiedź fizjologiczna gospodarza na takie wyzwanie może być dokładnie zbadana w komórkach utrzymywanych w stanie ustalonym [110–112]. W niniejszych badaniach zastosowano ciągłe hodowle bioreaktorowe rekombinowanych szczepów *Y. lipolytica* produkujących różne biochemicznie r(s)-Prots w warunkach chemostatu. Prowadzone hodowle wykazywały **stabilność monitorowanych parametrów** (DO, pH, stężenie DCW, GLY i metabolitów) **przy wartości D równej 0.20 [h⁻¹]**. Wysoka zadana wartość D kontrastuje z niższymi wartościami (0.01 [h⁻¹], 0.009-0.31 [h⁻¹], 0.08 [h⁻¹]) wcześniej stosowanymi w trakcie hodowli drożdży *Y. lipolytica* ukierunkowanych na produkcję niskocząsteczkowych metabolitów [120–122]. Z drugiej strony, jak pokazały poprzednie badania, proces wydzielania białka jest sprzężony ze specyficznym tempem wzrostu drożdży, które jest wyższe przy wyższym tempie wzrostu, a ekspresja genów związanych z syntezą białek jest dodatnio skorelowana z tempem wzrostu komórek *S. cerevisiae* [111,159] czy *P. pastoris* [160]. Co więcej, jak pokazały poprzednie badania, przy niskich wartościach D w przedziale 0.05-0.10 [h⁻¹], zjawiska związane ze stresem są dominujące, uniemożliwiając dokładny wgląd w procesy związane z produkcją r-Prots [111].

Przedstawiona w pracy uproszczona dystrybucja C i N pomiędzy różnymi produktami pozwoliła na poczynienie ciekawych obserwacji. Szczep Po1h_scTIG kierował stosunkowo wysoką frakcją C do produkcji polioli przy obniżonych wartościach DCW i, w przeciwieństwie do innych szczepów, wytwarzał więcej MAN niż ERY. Jak ostatnio dowiedziono, specyficzny wzrost stężenia MAN odnotowano w trakcie hodowli *Y. lipolytica* poddanych działaniu czynnika stresowego (↑ Temp. 35 °C) [109]. Sugeruje to, że **nadprodukcja silnie glikozyłowanej scTIG w niniejszej pracy, również odgrywała rolę czynnika stresowego, tym samym promując produkcję MAN.**

W większości analizowanych przypadków, szczepom pochodnym Po1g utrzymywanym w stanie stacjonarnym towarzyszyła wysoka akumulacja biomasy, w porównaniu do pochodnych Po1h. Z drugiej strony, pochodne Po1h wykorzystywały większe ilości GLY na jednostkę biomasy i wytwarzały większe ilości r(s)-Prots. Otrzymane wyniki mogą sugerować, że **dla pochodnych Po1h, większy strumień C został skierowany**

na syntezę r(s)-Prots a nie produkcję biomasy. Zwiększone zużycie substratu oraz ilość wytwarzanego aktywnego reportera może wynikać z wyższej wydajności transkrypcji z promotora 4UAS-pTEF w pochodnych Po1h niż z promotora hp4d w Po1g. Aby uzyskać wgląd w molekularne podstawy obserwowanych makroskopowo reakcji fizjologicznych komórek *Y. lipolytica* w odpowiedzi na nadprodukcję r(s)-Prots, analizowano poziom ekspresji genów kodujących białka reporterowe. Relatywnie **wyższe wartości ekspresji heterologicznych genów uzyskano dla promotora 4UAS-pTEF** (pochodne Po1h) **niż hp4d** (pochodne Po1g). Oba promotory posiadają w swej strukturze 4 motywy UAS1B (*Upstream Activating Sequences*), które silnie zwiększają szybkość transkrypcji następującego genu [161]. Wydaje się, że promotor 4UAS-pTEF jest bardziej aktywny transkrypcyjnie ze względu na obecność „rdzenia” w postaci pTEF w przeciwieństwie do minimalnego pLEU2 w hp4d [17,62], co koresponduje z wynikami uzyskanymi w tym badaniu. Jednak nie można wykluczyć potencjalnego wpływu pozostałych czynników różniących te dwie platformy ekspresyjne (miejsce integracji czy rodzaj markera selekcyjnego), ponieważ wyniki **testu statystycznego** wskazują na **brak istotnego wpływu promotora na poziom ekspresji heterologicznego genu** ($p < 0.05$).

Biorąc pod uwagę bezpośrednie porównanie frakcji C i N skierowanych na syntezę r(s)-Prots, relatywnie największy strumień odnotowano dla inYFP, następnie scYFP, scTIG i scSoA. Najprawdopodobniej różnice te wynikają ze sposobu szacowania ilości bezwzględnej białek, chociaż dane te dobrze korespondują z charakterystyką biochemiczną r-Prots podaną w **Tabeli 1**. Niskie bezwzględne ilości SoA mogą być spowodowane obecnością wielu reszt Cys (11) i wynikającej z tego dużej liczby możliwych kombinacji wiązań -S-S- (35 696; **Tabela 1**); największej wśród badanych reporterów białkowych. Jak wspomniano wcześniej, proces tworzenia wiązań -S-S- jest często wskazywany jako „wąskie gardło” procesów zachodzących w świetle ER; jest to bowiem proces wymagający zarówno pod względem potrzebnej energii jak i elementów budulcowych [57,83,162]. Stąd ograniczenie na tym etapie może również odpowiadać za najmniej wyraźną różnicę w pozakomórkowej aktywności scSoA, obserwowanej między dwoma gospodarzami (Po1g i Po1h), dla których proces formowania -S-S- w ER był identyczny. Z tych samych powodów, właściwości biochemiczne reporterowego białka YFP (tj. 2 reszty Cys i 1 przewidywane miejsce glikozylacji), zadecydowały o znacznej ilości wyprodukowanych scYFP i inYFP. Z kolei, dla scTIG, stosunkowo wysoka liczba reszt Cys (8; 764 możliwe kombinacje; **Tabela 1**), ale przede wszystkim - przewidywanych miejsc glikozylacji, mogą potencjalnie

ograniczać jego produkcję (przewidywane miejsca glikozylacji: 3 N-glikozylacji, 18 O-glikozylacji). Na podstawie uzyskanych danych można przypuszczać, że **tworzenie wiązań -S-S-** jest silnie ograniczającym etapem syntezy rs-Protów w komórkach *Y. lipolytica*, ponieważ znacząco obniżało bezwzględną ilość wyprodukowanego reportera SoA (mniej glikozylowanego i o niższej masie cząsteczkowej niż TIG (52 vs. 65 kDa; **Tabela 1**)). Należy jednak pamiętać, że zastosowana metoda szacowania ilości białka na podstawie aktywności i wartości aktywności specyficznej jest niedoskonała i mogła przyczynić się do niedokładności wyników oceny ilości białka.

W niniejszej pracy (**P.1**) zaobserwowano znaczącą dysproporcję pomiędzy poziomem transkryptu a oszacowaną ilością produktu białkowego dla reportera SoA. Niezależnie od zastosowanej platformy ekspresyjnej Po1g/Po1h (i użytego promotora), poziomy ekspresji genu *SOA* były najwyższe, a ilości „pozakomórkowej aktywności” SoA były porównywalne. Jednakże, obliczone bezwzględne ilości SoA wyrażone w gramach i molach były nieoczekiwanie niskie. Może to wynikać z błędnej definicji wartości aktywności właściwej stosowanej w obliczeniach, zaczerpniętej z pracy na temat natywnego enzymu [145] i/lub różnic w aktywności białka natywnego i rekombinowanego. W badaniach Aw i wsp.[163] poczyniono podobne obserwacje, gdzie w trakcie nadprodukcji białka Hsa w komórkach *P. pastoris*, subklony wydzielające białko z małą szybkością wykazywały ekspresję genów na znacznie wyższym poziomie w porównaniu z subklonami silnie wydzielającymi Hsa. Jednakże, w przeciwieństwie do niniejszych badań, gdzie zapewniono wysoką podaż składników, tamte badania przeprowadzono w hodowlach okresowych, w których ograniczone dostarczanie energii i elementów budulcowych dla maszyneryi transkrypcyjnej i translacyjnej mogło tłumaczyć obserwowany brak liniowości w relacji mRNA:białko.

Co zaskakujące, **wewnątrzkomórkowa akumulacja YFP w szczepie Po1h_scYFP**, w którym białko było kierowane do sekrecji, **była 5-krotnie wyższa niż w Po1h_inYFP**, gdzie gen był klonowany bez peptydu sygnałowego. Sugeruje to różnice w mechanizmie translacji (na powierzchni ER dla scYFP, w cytoplazmie dla inYFP) i/lub sekwestracji nadproduktowanego białka w kontinuum ER-AG-pęcherzyk. Możliwe też jest zahamowanie procesu syntezy, przy braku jego kierowania na zewnątrz komórki. Podobnych obserwacji dokonano dla galaktozydazy CelB klonowanej w *Y. lipolytica*, która również była kierowana do przestrzeni zewnątrzkomórkowej, ale dużą część aktywności wykryto także wewnątrz komórek gospodarza [11]. W tamtym badaniu stwierdzono, że wysoka retencja CelB wynikała z przeciążenia szlaku wydzielniczego. Podobnie, w innych badaniach pokazano

stosunkowo wysoką wewnątrzkomórkową akumulację reportera fluorescencyjnego, pomimo ukierunkowania polipeptydu do szlaku sekrecyjnego w komórkach *Y. lipolytica* [16]. Na podstawie tych obserwacji można wnioskować, że **białko może być syntetyzowane na wyższym poziomie, gdy jest kierowane do sekrecji**. Z drugiej jednak strony, Theron i wsp. [8] zaobserwowali 7-krotnie niższą fluorescencję wewnątrzkomórkową w szczepach *Y. lipolytica* ekspresjonujących GFP wyposażonych w SP, w porównaniu z szczepem produkującym wewnątrzkomórkowy wariant fluoroforu.

Podsumowując, uzyskane dane sugerują, że zastosowanie bardziej aktywnego transkrypcyjnie promotora powoduje skierowanie większej ilości strumienia C w kierunku białka docelowego, dając znacznie większe ilości specyficzne i szybkości produkcji docelowego polipeptydu, ze stratą w akumulacji biomasy, bez istotnego wpływu na produkcję polioli. W zależności od właściwości biochemicznych produkowanego białka i promotora regulującego jego ekspresję, zaobserwowano specyficzne odpowiedzi w zakresie kinetyki tworzenia białek oraz poziomu transkrypcji genów. Poczynione obserwacje wskazują na silny wpływ indywidualnej kombinacji białka docelowego czy zastosowanego promotora i kilku innych, niezidentyfikowanych czynników, na odpowiedź szczepu gospodarza.

Aby odpowiedzieć na pytanie, w jaki sposób właściwości biochemiczne r(s)-Prots, wpływają na zdolności sekrecyjne drożdży *Y. lipolytica*, przeprowadzono globalną analizę transkryptomu *Y. lipolytica* (pochodnych Po1h) hodowanych w warunkach chemostatu (**P.2**). Analiza RNAseq umożliwiła identyfikację genów charakteryzujących się różnicową ekspresją (DEGs) w warunkach nadprodukcji różnych r(s)-Prots (**P.2**). Wstępnie, przypuszczano, że zastosowanie r(s)-Prots o tak odmiennych właściwościach biochemicznych uwydatni odpowiedź komórkową *Y. lipolytica* w różnych punktach szlaku translacyjno-sekrecyjnego.

Na podstawie całkowitej liczby DEGs (**Rycina 5**) uzyskanej z analizy RNAseq wskazano na wysokie zróżnicowanie profili transkryptomicznych szczepów nadprodukujących r-Prots sekrecyjne (*SoA*, *TIG*, *scYFP*) i r-Prot wewnątrzkomórkowe (*inYFP*), co sugeruje, że **kierowanie polipeptydu do przestrzeni zewnątrzkomórkowej indukuje masowe rearanżacje w komórce, przekraczające różnice związane z właściwościami biochemicznymi białek**. Wniosek ten wynika z obserwacji wskazującej na bliższe podobieństwo profili transkrypcyjnych w szczepach HSS (*scYFP* vs. *SoA*), niż w przypadku szczepów nadprodukujących to samo białko, ale w wariantach wydzielanym na zewnątrz lub nie (*scYFP* vs. *inYFP*) (**P.2**).

Dane transkryptomocne (**P.2** i **P.4**) dostarczyły dowodów na to, że **wysoka nadprodukcja rs-Prots** (scSoA i scYFP) w komórkach *Y. lipolytica* jest **nieodłącznie związana ze stresem oksydacyjnym zachodzącym w ER**, co wcześniej zaobserwowano także dla *P. pastoris* i *S. cerevisiae* [69,164]. Na podstawie analizy Gene Ontology, wskazano na istotną indukcję procesu biologicznego „odpowiedzi na nadtlentek wodoru” w HSS, reprezentowanego m.in. przez katalazę Ctt1 oraz białko Hsp31. Warto zauważyć, że **reakcja na stres oksydacyjny w komórkach *Y. lipolytica* została zainicjowana niezależnie od liczby wiązań -S-S- w strukturze drugorzędowej rs-Prots** (scYFP), co zostało już wcześniej zaobserwowane w [50,54,165]. Nieoczekiwanie, w HSS wykazano obniżony poziom cytoplazmatycznej tioredoksyny *TRX1/E09427g* (**Rycina 8c**). Aktywność tioredoksyny jest związana z procesem tworzenia wiązań -S-S-, w którym pośredniczą Pdi1 i Ero1 i opiera się na stochastycznym utlenianiu-redukcji łańcuchów bocznych Cys, które zużywają znaczne ilości czynników utleniających i redukujących (odpowiednio O₂ i glutationu). Pomimo tego, że fałdowanie oksydacyjne zachodzi w świetle ER, ani tioredoksyna ani glutation nie występują w tym kompartmentcie [165]. Zgodnie z obecnym modelem Trx1 przenosi elektrony do ER w celu redukcji utlenionego Pdi1 [68], podczas gdy glutation jest aktywnie transportowany z cytoplazmy przez specyficzne transportery [166]. W obliczu tego, obniżona ekspresja *TRX1* w HSS nie jest jasna i kontrastuje z 17% indukcją ***SEC61* w *SoA*** (**Rycina 10e**), o którym wiadomo, że jest składnikiem porów translokacyjnych i bierze udział w transbłonowym **transporcie glutationu do ER** [166]. Ta niewielka indukcja *SEC61* jest jedynym, słabym markerem **zwiększonego fałdowania oksydacyjnego** i tworzenia wiązań -S-S- w *SoA*.

Jak wcześniej wspomniano, fałdowanie białek w ER jest uważane za jedno z najbardziej ograniczających „wąskich gardeł” w szlaku wydzielania rs-Prots [85]. Nadekspresja *PDII* była często wykorzystywaną strategią wywierającą pozytywny wpływ na wydzielanie rs-Prots w *P. pastoris* [69]. Natomiast, w zależności od ilości wiązań -S-S- w rs-Prot, ko-nadekspresja *PDII* w *K. lactis* w różnym stopniu wpływała na wydzielanie białka docelowego, wywołując korzystny efekt przy rs-Prot bogatym w -S-S-, i brak istotnego wpływu dla rs-Prot wolnego od wiązań -S-S- [82]. Różnice w tych obserwacjach mogą wynikać z biochemicznych właściwości docelowego polipeptydu [52,57,85]. Warto zauważyć, że w niniejszych badaniach (**P.3**) udowodniono brak istotnej poprawy zdolności sekrecyjnej komórek *Y. lipolytica* ((sp_)exYFP) w temp. 30 °C lub istotny, lecz umiarkowany wzrost poziomu (sp_)exYFP w temp. 25 °C), spowodowany ko-nadekspresją *PDII*.

Jest prawdopodobne, że zastosowanie białka reporterowego YFP, posiadającego jedynie 2 reszty Cys (**Tabela 1**), może w pewnym stopniu wyjaśnić brak istotnego efektu nadekspresji *PDII* na wydzielanie rs-Prot.

Co ciekawe, **indukcja odpowiedzi na stres oksydacyjny** (szczepy HSS) **była niezależna od inicjacji odpowiedzi UPR** w przypadku badanych szczepów *Y. lipolytica*. Dane literaturowe wskazują, że główny regulator zjawiska UPR - Hac1 przyczynia się do wzmożonej ekspresji genów zaangażowanych w odpowiedź na stres oksydacyjny, który jest kluczowy dla zarządzania szlakiem wydzielniczym [167]. Co więcej, w niniejszej pracy **zwiększonemu poziomowi ekspresji *HAC1* w szczepach *inYFP/TIG* nie towarzyszyła reakcja na stres oksydacyjny, ale zauważalna indukcja genów zaangażowanych w sortowanie białek i degradację proteolityczną w wakuoli**. Co więcej, transkryptom szczepu *HAC1_scYFP* (**P.4**) odpowiadał profilom transkrypcyjnym szczepów *inYFP* i *TIG* z poprzedniego badania, w przeciwieństwie do szczepów HSS - *scYFP* i *SoA* (**P.2**). Wśród genów o zwiększonym poziomie ekspresji w szczepach z nadreprezentacją Hac1 zidentyfikowano *ATG8* który, bierze udział w procesie autofagii specyficznym dla ER i jest niezbędny dla przeżycia komórek poddanych silnemu stresowi ER. **Indukcja *ATG8* w szczepach *HAC1_scYFP* i *inYFP* kontrastowała z ich znaczną represją w *scYFP***. Podobnie, wiele innych genów zaangażowanych w którykolwiek ze szlaków degradacji białek (proteasom, autofagia, VPS), ulegało represji w szczepach *scYFP* i *SoA*, bez zmiany lub indukcji w szczepach z nadekspresją *HAC1* (*HAC1_scYFP* i *inYFP*) (**Rycina 18**). W oparciu o tak specyficzny wzorzec ekspresji, reprezentowany przez tak wiele genów, postawiono hipotezę, że **poziom ekspresji *HAC1* jest bezpośrednio związany z procesem sortowania i degradacją białek w wakuoli**. Aby zweryfikować to stwierdzenie, zbadano całkowitą wewnątrzkomórkową aktywność proteolityczną szczepów *scYFP*, *inYFP* i *HAC1_scYFP*. Jednak eksperymentalna weryfikacja tej hipotezy przyniosła nieoczekiwane wyniki (Fig.6 w **P.4**). Szczególnie zaskakująca obserwacja dotyczyła zwiększonej proteolizy w szczepie *scYFP*. Biologiczny sens zwiększenia całkowitej aktywności proteolitycznej przy intensywnej syntezie r(s)-Prots wydaje się zrozumiały i najprawdopodobniej jest związany z globalnie zwiększonym metabolizmem białek. Jednakże takie wyjaśnienie nie odpowiada ogólnemu profilowi danych transkryptomicznych, z których jasno wynika, że proteoliza była hamowana w *scYFP*. Szczegółowa analiza DEGs wykazała, że wiele genów o obniżonym poziomie ekspresji przypisanych do funkcjonalnej kategorii „proteoliza” to VPS, czynniki pośredniczące w ubikwitynacji i związane z autofagią, które nie posiadają aktywności

proteolitycznej *per se*, co z kolei podlegało analizie w teście enzymatycznym. Innymi słowy, **w szczepach HSS to proces kierowania białka do degradacji ulegał ograniczeniu.** Specyficzna aktywność proteolityczna była najwyższa w szczepach ze zwiększoną retencją YFP, tj. *scYFP* i *inYFP* (Tabela 5). Z drugiej strony, ko-ekspresja *HAC1* przyczyniła się do istotnego spadku aktywności proteolitycznej w stosunku do tych dwóch szczepów; która wciąż była wyższa niż w kontroli. **Zmniejszona aktywność proteolityczna w szczepie *HAC1_scYFP* może częściowo wyjaśniać wyższe ilości wydzielanego YFP (Tabela 5).**

Profile transkrypcyjne szczepów HSS były istotnie **wzbogacone** w indukcję genów zaangażowanych m.in. w **biogenezę i składanie rybosomów, transkrypcję, translację, modyfikację chromatyny**, co razem wskazuje na intensyfikację procesów transkrypcji i translacji. Represja *E26763g (RSFA)* w HSS, koresponduje z tymi obserwacjami. Przypisana RsfA funkcja wskazuje na rolę w przystosowaniu komórek do warunków powolnego wzrostu/fazy stacjonarnej poprzez represję syntezy białek; która jest jednym z najbardziej energochłonnych procesów. **Obniżenie ekspresji RsfA uwalnia zespół rybosomów spod hamującego działania**, co ma kluczowe znaczenie dla nadprodukcji r-Protów. Dane otrzymane w niniejszych badaniach kontrastują z wcześniejszymi doniesieniami, w których nadprodukcja dużego r-Prot w *Y. lipolytica* spowodowała ogólną represję procesu transkrypcji i translacji celem dostosowania komórki do ograniczonej zdolności fałdowania i dojrzewania [50]. Wyjaśniono dalej, że jeśli procesy poprzedzające fałdowanie oksydacyjne nie są spowolnione, ryzyko silnego stresu oksydacyjnego i nadmiernej utylizacji równoważników redukujących jest wysokie [50]. Biorąc pod uwagę indukcję genów biorących udział w odpowiedzi na stres oksydacyjny (np. *CTT1* w *SoA/scYFP/HAC1_scYFP* oraz *HSP31* w HSS), jest możliwe, że **wysoka transkrypcja i translacja doprowadziła do „niezrównoważonego” przepływu przez szlak wydzielniczy i przyczyniła się do wywołania stresu oksydacyjnego.** Wyniki dotyczące ograniczonego poziomu ekspresji *INYP* i *TLG* wraz z brakiem represji *RSFA* i brakiem odpowiedzi na stres oksydacyjny w tych szczepach łącznie potwierdzają to stwierdzenie.

Ograniczona częstość interakcji Kar2-Ire1 (ograniczenie ekspresji *SLS1*), brak deregulacji genów związanych z mechanizmem translacji kotranslacyjnej (receptora SRP i samego SRP), a także silna indukcja genów kodujących cytozolowe białka chaperonowe z rodziny Ssa (w szczególności *SSA5* i *SSA8*) w HSS/*HAC1_scYFP*, łącznie sugerują, że **translokacja rs-Protów odbywała się poprzez mechanizm translacji potranslacyjnej, tj. mniej wydajnym u *Y. lipolytica* [28], lub obiema drogami (P.2, P.4).** Jak wspomniano

wcześniej, białka chaperonowe Ssa należą do rodziny Hsp70 i odgrywają główną rolę w ochronie/utrzymywaniu powstającego polipeptydu w stanie kompetentnym do translokacji, kiedy polipeptyd może być przemieszczany za pośrednictwem mechanizmu potranslacyjnego. Podobny wzorzec ekspresji (znacząca indukcja ekspresji w HSS) zaobserwowano także dla innego reprezentanta cytozolowych białek opiekuńczych, *SSE1* (**Rycina 10d**). *SSE1* (i *SSE2* w *P. pastoris*; pojedynczy gen w *Y. lipolytica*) wspomaga fałdowanie białka poprzez wiązanie z powstającym peptydem i utrzymywanie go w stanie kompetentnym. Ko-ekspresja genu heterologicznego *SSE1*, jako SH, pozwoliła na 2,3-krotne zwiększenie ilości białka docelowego [69]. W badaniach własnych (**P.3**) wskazano, że **nadekspresja SHs zaangażowanych w syntezę rs-Prot w cytoplazmie (*RPL3*, *SSA5*, *SSA8*) znacząco zwiększyła ilości wewnątrzkomórkowego białka reporterowego (scYFP) niezależnie od zastosowanej temperatury; jednak dla skutecznego wydzielania, pojemność szlaku wydzielniczego musiała zostać uwolniona poprzez zastosowanie obniżonej temperatury. Innymi słowy stwierdzono, że to **łączny efekt czynnika genetycznego (nadekspresji SH) i środowiskowego (obniżona temp.) istotnie wpłynął na zdolność sekrecyjną komórek *Y. lipolytica*, a nie bezpośrednie działanie SH w sposób zależny od temp (Rycina 13)**. Co więcej, dane dotyczące poziomu ekspresji heterologicznego genu potwierdzają to stwierdzenie. Ekspresja *YFP* została istotnie zwiększona przy nadekspresji *RPL3*, *SSA5* i *SSA8* w obu zadanych temperaturach (**Rycina 14A.B**), natomiast dla wszystkich trzech badanych SH temperatura była nieistotnym czynnikiem wpływającym na ekspresję *YFP* (**Rycina 15**). Korzystny wpływ nadekspresji *SSA1* (homologa *SSA5* u *Y. lipolytica*) [81] i *SSA4* (homologa *SSA8*) [69] na syntezę i sekrecję rs-Protów potwierdzono także u *P. pastoris*. Podobnie, nadekspresja genu kodującego białko rybosomalne Rpp0 (odpowiednik Rpl3) zwiększyła wydzielanie rs-Protów u *S. cerevisiae* [104]. Zarówno badania własne jak i dane literaturowe wskazują, że geny posiadające aktywność cytozolowych białek opiekuńczych oraz zaangażowane w proces translacji to odpowiednie cele dla usprawnienia syntezy rs-Prot w systemach drożdżowych. **Kombinacja ko-nadekspresji *SSA5*, *SSA8* lub *RPL3* w warunkach obniżonej temp. przyczyniła się do maksymalizacji ilości wydzielanego rs-Protów w komórkach *Y. lipolytica*.****

Transport pomiędzy ER a AG jest często wskazywany jako kluczowy etap w prawidłowym funkcjonowaniu szlaku wydzielniczego [85,168]. Globalna analiza transkryptomu (**P.2** i **P.4**) uwidoczniała m.in. nadekspresję genu *USO1* (w HSS i *HAC1_scYFP*) kodującego białko zaangażowane w transport pęcherzykowy z ER do AG

(Rycina 6 i Rycina 18). W przeciwieństwie do pozytywnego wpływu ko-nadekspresji *USO1* w komórkach drożdży *S. cerevisiae* [169], w badaniach własnych (**P.3**) zaobserwowano negatywne oddziaływanie wprowadzonej modyfikacji na poziomy exYFP w temp. 30 °C lub brak istotnej zmiany w temp. 25 °C. Ponadto, uzyskane wyniki wskazują, że nadekspresja *USO1* nie odgrywała istotnej roli regulacyjnej w promowaniu ekspresji *YFP*. Podobnie, dla *P. pastoris* wśród testowanych SH (*COG6*, *COY1*, *IMH1*, *SEC31*) zlokalizowanych w przedziale AG nie wykazano znaczącej poprawy w wydzielaniu rs-Prots, co dobrze koresponduje z danymi otrzymanymi w niniejszej pracy [69].

Jak sugerują dane literaturowe, nadekspresja genów uczestniczących w transporcie pęcherzykowym z AG do PM czy dokowaniu i fuzji pęcherzyków egzocytarnych stanowi skuteczne podejście w intensyfikacji produkcji rs-Prots w *P. pastoris* i *S. cerevisiae* [69,77,83,93]. W niniejszej pracy (**P.2** i **P.4**) zidentyfikowano obniżony poziom ekspresji genów zaangażowanych w te procesy w HSS i niewielką indukcję lub brak zmiany ekspresji w *TIG/inYFP* (**Rycina 10g**) (**P.2**) i *HAC1_scYFP* (**Rycina 18**) (**P.4**). Represja genów transportu pęcherzykowego podczas syntezy i sekrecji rs-Prots na wysokim poziomie, wydaje się być niejasna. Stąd, w dalszej części pracy, zastosowano gen syntaksyny *SSO1* (transport z AG do PM) jako SH i określono wpływ ko-nadekspresji na produkcję scYFP (**P.3**). Zaobserwowano, że hodowla szczepu *scYFP_SSO1* w temp. 30 °C istotnie zwiększyła wewnątrzkomórkową retencję YFP, czego nie obserwowano w temp. 25 °C. Wydaje się, że zastosowanie temp. 25 °C uwalnia proces terminalnego transportu, w którym działa Sso1. Natomiast w warunkach temp. 30 °C, gdzie szlak sekrecyjny działa z pełnym obciążeniem, to działanie Sso1 uwalnia etap dystalnego transportu. Co szczególnie istotne, kumulatywne działanie nadekspresji *SSO1* i temp. 30 °C skutkowało najwyższymi odczytami sp_exYFP spośród szczepów nadekspresjonujących badane SHs w w/w temp. hodowli. W oparciu o te obserwacje można wnioskować, że celowanie/fuzja pęcherzyków wydzielniczych na etapie pomiędzy AG i PM, na którym działa *SSO1*, jest „wąskim gardłem” dystalnego szlaku wydzielniczego w *Y. lipolytica*. Jak odnotowano wcześniej, nadekspresja homologicznych *SSO1* i *SSO2* w komórkach *S. cerevisiae* spowodowała 4–6-krotny wzrost wydzielania heterologicznego enzymu bakteryjnego [170]. Kolejna modyfikacja międzygatunkowa obejmująca nadekspresję homologu *SSO1* *K. lactis* w *S. cerevisiae* [171] czy *SSO2* z *S. cerevisiae* w *P. pastoris* [172] doprowadziła do intensyfikacji produkcji wydzielanych białek w komórce gospodarza, potwierdzając jego uniwersalnie ważną rolę w procesie wydzielania

polipeptydów. Stąd, przedstawione wyniki własne, jak i dane literaturowe, uzasadniają manipulacje genetyczne na etapie ruchu pęcherzykowego AG – PM.

W odpowiedzi na intensywną syntezę różnych rs-Prots w *Y. lipolytica*, (**P.2** i **P.4**) zaobserwowano aktywację procesów biologicznych, jak: „organizacja ściany komórkowej” i „restrukturyzacja”. Procesy te były reprezentowane przez geny kodujące enzymatyczne *CCW12* oraz strukturalne *CWP11/CWP12* białka ściany komórkowej (**Rycina 10h** i **Rycina 18**). Z drugiej strony, wprowadzona modyfikacja (ko-nadekspresja *CWP11* i *SCYFP*) w sposób umiarkowany wpłynęła na nadprodukcję scYFP w porównaniu z innymi testowanymi SHs (**P.3**). Zatem, wydaje się być wysoce pożądane dalsze wyjaśnienie mechanizmu, za pomocą którego białka ściany komórkowej promują lub ograniczają wydzielanie rs-Prots u *Y. lipolytica*.

Biorąc pod uwagę fakt, że w niniejszej pracy (**P.2**) wszystkie heterologiczne geny zostały transkrypcyjnie połączone z SP o poznanej budowie i funkcji (egzo-1,3-beta-glukanaza; *B03564g*) [15] **szczególnie intrygująca wydaje się być znaczna indukcja genu *NCE102* we wszystkich szczepach syntetyzujących sekrecyjne r-Prots** bez wyjątku (brak zmian wyłącznie dla inYFP). Nce102 to białko zaangażowane w niekonwencjonalny eksport, które bierze udział w sekrecji białek pozbawionych sekrecyjnych sekwencji sygnałowych [173,174]. Wiadomo również, że ta niekonwencjonalna ścieżka eksportu działa jako alternatywna droga oczyszczania/detoksykacji w celu wyeliminowania uszkodzeń produktu białkowego, gdy podstawowa ścieżka nie jest wystarczająca [175]. Sugeruje to, że **przy bardzo wysokim wydzielaniu scSoA i scYFP, oraz wydzielaniu scTIG, *NCE102* działa jako „wyjście awaryjne” wydzielania prawidłowo sfałdowanych, funkcjonalnych białek.**

W odniesieniu do specyficznej odpowiedzi komórkowej dla poszczególnych r-Prots o odmiennych właściwościach biochemicznych, zaobserwowano znaczą indukcję genów biorących udział w syntezie oligosacharydów i przetwarzaniu glikoprotein w szczepie *TIG* (**Rycina 10c**); co dobrze koreponduje ze strukturą produkowanego polipeptydu, bogatego w miejsca glikozylacji (**Tabela 1**). Co więcej, zaobserwowano represję mannozylotransferaz przy produkcji scSoA i scYFP, które są białkami słabo glikozylowanymi (**Tabela 1**). Ostrożna analiza zestawów DEGs dla HSS, uwidoczniała także represję wszystkich głównych proteaz oraz lipaz natywnie wydzielanych przez komórki *Y. lipolytica* (*XPR2*, *AXP1*, *LIP4*, *TGL3*); białek o zróżnicowanej liczbie miejsc glikozylacji. Sugeruje to uruchomienie mechanizmu autoregulacji komórki, gdzie intensywna **nadprodukcja i sekrecja rs-Prots**

spowodowała obniżenie ekspresji natywnych genów kodujących enzymy sekrecyjne na zasadzie autoregulacji, tym samym ograniczając obciążenie szlaku wydzielniczego i poprawiając tym samym wydajność wydzielniczą komórki.

Co ciekawe, obszerna analiza danych omicznych, pozwoliła na zdobycie nowej wiedzy na temat regulacji cyklu komórkowego w *Y. lipolytica* w powiązaniu z intensywną nadprodukcją rs-Protów (**P.2** i **P.4**). Zidentyfikowano aż 10 genów (m.in. *CLNI*) o obniżonym poziomie ekspresji w szczepach *scYFP*, *HAC1_scYFP*, a także częściowo w *SoA* (**Rycina 19**), które są zaangażowane w regulację progresji cyklu komórkowego. Co uderzające w tym kontekście, stwierdzono także znaczną indukcję genu represora *XBPI/F16511g* w szczepach *HAC1_scYFP* (**P.4**) i *inYFP* (**P.2**), w przeciwieństwie do obniżenia jego ekspresji w szczepach HSS: *scYFP* i *SoA* (**P.2**) (**Rycina 19**). Udział Xbp1 w regulacji cyklu komórkowego (wiązanie z sekwencjami promotorowymi genów cyklin), ale także w odpowiedzi na stres ER został potwierdzony dla *S. cerevisiae* [176–178]. **Represja XBPI, któremu towarzyszy brak odpowiedzi transkrypcyjnej z HAC1 i obniżenie poziomu CLNI, mogą łącznie sugerować, że HSS weszły w stan spoczynku** [179]. Indukcja innego czynnika transkrypcyjnego, DeoR (*E19030g*), w obu HSS dodatkowo potwierdza to stwierdzenie. Czynniki transkrypcyjne typu DeoR biorą udział w regulacji metabolizmu węglowodanów [180], m.in. jako represory importu cukru [181]. Zatem, można spekulować, że znaczna indukcja *E19030g* związana jest z obniżeniem poziomu ekspresji permeazy maltozowej *MPH2* w *scYFP* i permeazy cukrowej *STL1* w *SoA*. Zmniejszona transkrypcja niektórych mniej istotnych genów w celu dostarczenia wystarczającej ilości energii i substratów do syntezy r-Prot była postulowana uprzednio na podstawie danych transkryptomicznych z ciągłej hodowli *P. pastoris* [66]. Z drugiej strony, **biorąc pod uwagę obserwowane wzorce ekspresji HAC1 i XBPI we wszystkich analizowanych szczepach (SoA, scYFP, HAC1_scYFP, inYFP i TIG) można postulować, że ekspresja tych dwóch genów jest skorelowana**, natomiast wymaga to dalszych analiz.

W dalszej części pracy (**P.3**) szczegółowe analizy przeprowadzone w **P.2** posłużyły jako przesłanki do wytypowania zestawu dwunastu SHs, które w zadanych warunkach ulegały istotnej deregulacji. Badania rozpoczęto od konstrukcji szeregu transformantów *Y. lipolytica*, z ko-nadekspresją *SH* z *SCYFP*. Rekombinowane szczepy testowano w dwóch różnych temperaturach (25 °C i 30 °C) pod kątem ich zdolności wydzielniczej, retencji białka i ekspresji genów kodujących SHs i rs-YFP. Uzasadnieniem takiego podejścia były wcześniejsze badania dotyczące istotnego wpływu temperatury hodowli na działanie SHs ko-

nadeksprimowanych w *P. pastoris* [55] i *S. cerevisiae* [104] badanych w zakresie od 20 °C do 37 °C. Ponadto, Gasser i wsp. [55] wykazali, że obniżenie temp. hodowli o 5 °C (z 25 °C do 20 °C) zwiększyło produkcję białka reporterowego o 40%. W niniejszych badaniach zastosowano temp. 25 °C, wcześniej stosowaną u *P. pastoris* [69] do testowania SHs oraz temp. 30 °C jako odniesienie dla typowej temp. hodowli *Y. lipolytica*. Na podstawie przedstawionych danych w **P.3**, wykazano, że **nadekspresja któregokolwiek z SH nie miała wpływu na tempo wzrostu rekombinowanych szczepów *Y. lipolytica*, natomiast obniżona temp. hodowli sprzyjała akumulacji biomasy**. Obserwacje te, doskonale korespondują z ostatnimi doniesieniami, dotyczącymi wpływu obniżonej temp. hodowli (20 °C) na przyrost biomasy, a także wydajną syntezę i sekrecję owadziej α -amylazy w komórkach *Y. lipolytica* [109]. Podobnie, Gasser i wsp. [69] nie zaobserwowali negatywnego wpływu nadekspresji SH na wzrost *P. pastoris*. Efekty nadekspresji poszczególnych SHs omówiono w innych częściach Dyskusji, przy okazji omawiania ich profilu transkrypcyjnego w hodowlach ciągłych.

W ostatniej części pracy (**P.4**) podjęto próbę wyjaśnienia działania, nadekspresji jednego z uprzednio badanych SH, kodującego TF Hac1 u *Y. lipolytica*, na funkcjonowanie szlaku sekrecyjnego w trakcie produkcji rs-YFP. Analizie porównawczej poddano trzy szczepy *Y. lipolytica* syntetyzujące heterologiczne białko reporterowe YFP: i) szczep *scYFP* z bazowym poziomem ekspresji genu *HAC1*, pomimo wysokiego poziomu nadsyntezy rs-YFP (**P.2**), ii) szczep *inYFP* syntetyzujący wewnątrzkomórkowy YFP, z natywnie podwyższoną ekspresją *HAC1* o 23% w porównaniu z kontrolą (oznaczoną metodą RNAseq; **P.2**), iii) szczep *HAC1_scYFP* z syntetycznie indukowaną ekspresją genu *HAC1* (6-krotnie) w stosunku do kontroli, i o ~30% wyższą syntezą YFP (zgodnie z wynikami z hodowli okresowych w **P.3**). Transkryptomy szczepów *inYFP* i *scYFP* posłużyły jako referencje, umożliwiające dokładniejszą charakterystykę transkryptomu *HAC1*.

Wbrew wynikom uprzednich doniesień literaturowych [111,182], w niniejszej pracy zaobserwowano stały poziom ekspresji znanych regulatorów UPR (*IRE1*, *KAR2*, *PDI*, *ERJ*, *ERO1*, *GSH*) we wszystkich rekombinowanych szczepach (**P.2**) jak i *HAC1_scYFP* (**P.4**). Wyjątek stanowił gen *CNE1* ulegający represji w szczepie *scYFP*. W przeciwieństwie do stałego poziomu ekspresji *KAR2* i *IRE1*, ekspresja genu *SLS1/E32703g*, kodującego mediator interakcji między Ire1 i Kar2 [21] oraz białko niezbędne do translokacji kotranslacyjnej u *Y. lipolytica* [19,36], była obniżona we wszystkich szczepach produkujących rs-Protos, ale nie *inYFP* (**Rycina 10c**) i *HAC1_scYFP* (**Rycina 18**). Jak wykazano wcześniej, fałdowanie

białek w ER jest najbardziej ograniczającym procesem w szlaku translacyjno-sekrecyjnym [85]. Stwierdzenie to sformułowano na podstawie pozytywnego wpływu nadekspresji *PDI* czy *HAC1* na sekrecję rs-Protów w *P. pastoris* [69], czy nadekspresji *CNE1* u *H. polymorpha* [183]. W niniejszym badaniu (**P.3**), wprowadzona nadekspresja regulatorów UPR - *HAC1*, *SLS1* i *CNE1* (PDI – omówiono powyżej) przyczyniała się do zwiększonego poziomu ekspresji *SCYFP* w temp. 30 °C, ale bez istotnego wpływu na jego syntezę i sekrecję. Z drugiej strony, w tych szczepach hodowanych w temp. 25 °C, obniżony poziom ekspresji *SCYFP* występował wraz ze zwiększoną sekrecją rs-Prot. Podobnych obserwacji dokonano dla *P. pastoris*, gdzie obniżenie temp. hodowli o 5 °C spowodowało 1,4-krotny wzrost szybkości wydzielania produktu, pomimo tego, że poziom ekspresji heterologicznego genu był znacząco obniżony [55]. Na podstawie doniesień literaturowych oraz danych otrzymanych w niniejszej pracy, postuluje się, że **ograniczenie transkrypcji heterologicznego genu na skutek obniżonej temperatury, przyczynia się do usprawnienia działania maszynerii faldującej w ER**, co w konsekwencji przyczynia się do zwiększenia **efektywności syntezy i wydzielania rs-Protów**. Jednakże, wyniki uzyskane dla szczepów z ko-nadekspresją cytozolowych białek chaperonowych, dla których obserwowano indukcję poziomu ekspresji heterologicznego genu niezależnie od temp., wymykają się temu stwierdzeniu. Tam, pomimo istotnie zwiększonego poziomu syntezy białka, poziom sekrecji był także istotnie zwiększony; ale z kolei to zjawisko było temperaturo-zależne.

Lista skutecznych strategii obejmujących wykorzystanie TF *HAC1*, jako SH, jest bardzo długa [50,54,55,69,72,184]. W odniesieniu do ko-nadekspresji *HAC1* w niniejszej pracy (**P.3**), wprowadzona modyfikacja wywołała istotną poprawę poziomów *sp_exYFP* (ok. 30%) w stosunku szczepu kontrolnego w obu badanych temperaturach. Porównywalny poziom poprawy w zakresie nadprodukcji rs-Prot zaobserwowano również u *P. pastoris* (ok. 50%) [172]. Na podstawie otrzymanych wyników stwierdzono indukcję ekspresji genu *SCYFP* wyłącznie w temp. 30°C, ale ostateczne wartości *sp_exYFP* były porównywalne w obu temperaturach ($1818,34 \pm 20,21$ vs. $1834,9 \pm 97,52$ [FU OD600 nm⁻¹]) odpowiednio dla 25°C i 30°C; **Tabela 4**). Zgodnie z argumentacją przedstawioną w [104], można wnioskować, że zastosowane w tym badaniu temperatury hodowli pozwoliły na uwidocznienie efektów ko-nadekspresji *HAC1*. Technicznie żadna z zaimplementowanych temperatur nie odciążała/obciążała biologicznych procesów, w których zaangażowany jest Hac1, a raczej – **zakres działania Hac1 jest na tyle szeroki, że wywierał pozytywny wpływ przy ograniczeniu transkrypcji (25 °C) lub sekrecji (30 °C)**.

Warto zauważyć, że zwiększony poziom ekspresji *HAC1* obserwowany w *inYFP* (**P.2**) nie może wynikać z kanonicznej indukcji ER-zależnego UPR, ponieważ białko *inYFP* nie jest celowane do sekrecji, a zatem nie wchodzi do kompartmentu ER. Z drugiej strony, indukcja *HAC1* w szczepie *inYFP* nie była odosobnionym zjawiskiem spowodowanym nadmierną syntezą *inYFP*, ale rzeczywiście towarzyszyło jej wiele innych zjawisk (DEGs) typowych dla UPR. Podobnie zbliżony profil obserwowano w szczepie nadprodukującym glukoamylazę *TIG* (**P.2**), która powinna być transportowana zgodnie z klasycznym wzorcem, z racji N-terminalnego peptydu sygnałowego, i tym samym szczep mógł uaktywnić typowy ER-zależny UPR. Transkryptomy obu szczepów charakteryzowały się wysokim poziomem podobieństwa, co sugeruje podobną odpowiedź biologiczną. Możliwe, że **indukcja *HAC1* w szczepie *inYFP* wynika z pośredniego działania tzw. stresu „cytozolowego UPR” [185] regulowanego przez *Hsf1*** lub z innego mechanizmu. Taka ścisła zależność między „cytozolowym-” a ER-zależnym UPR jest dobrze opisana dla roślin [186], dla których implikacja czynników transkrypcyjnych bZIP60 i bZIP28/17 (40% podobieństwa do *Hac1*, na podstawie blastp) została szczegółowo zbadana. W niniejszym badaniu transkryptom szczepu *Y. lipolytica HAC1_scYFP* był podobny do transkryptomów opisanych dla *S. cerevisiae*, jako (i) indukowany przez „regulator specyficzny dla ER”, czyli *Hac1* [156], oraz (ii) przez „regulator specyficzny dla cytozolu”, czyli *Hsf1* [54]. Na przykład, geny zaangażowane w metabolizm kwasów tłuszczowych (*MGA2/D24101g* FC: 1.18), biosyntezę hemu (*HEM12/C01716g* FC: 0.61, *HEM13/E29975g* FC: 0.78, *HEM15/F19470g* FC: 0.33), biogenezę ściany komórkowej (*EXG2/F05390g* FC: 1.29, *PKC1/F09746g* FC: 0.85) lub biosyntezę fosfolipidów (*OPI3/E12441g* FC: 1.27) mają ten sam wzorzec ekspresji, co regulom indukowany przez *HAC1* *S. cerevisiae* [156]. Z drugiej strony inne geny ulegające deregulacji w szczepie *HAC1_scYFP*, odgrywały rolę w ubiquitynacji i proteolizie (*UBI4/F18403g* FC: 1.27, *PIB1/D15510g* FC: 0.78), metabolizmie węglowodanów (*UGP1/A02310g* FC: 1.16), transdukcji sygnału [54], jak typowo indukowane przez *Hsf1*. Tak silne podobieństwo obserwowanego tutaj transkryptomu szczepu *HAC1_scYFP* na odpowiedniki *Hsf1*-/*Hac1*-specyficzne z *S. cerevisiae* sugeruje, że rzeczywiście **transkryptom szczepu *inYFP* mógł się rozwinąć w oparciu o akumulację *inYFP* w cytoplazmie, co spowodowało zwiększenie aktywności *HSF1* i towarzyszącą mu indukcję *HAC1* (przez nieznaną mechanizm), po którym następuje aktywacja jego określonych celów na dalszych etapach.** Wysokie podobieństwo kanonicznego transkryptomu UPR szczepu *TIG* do wymienionych genów markerowych z transkryptem *S. cerevisiae*

specyficznym dla *HAC1*, a tylko jednego wspólnego genu z transkryptem specyficznym dla *HSF1*, stanowi poparcie tej hipotezy.

Geny zaangażowane w odpowiedź na stres oksydacyjny, progresję cyklu komórkowego (**Rycina 19**) i kilka represorów polimerazy RNA II i III były istotnie deregulowane w szczepie *HAC1_scYFP*. Wydaje się jednak, że ta deregulacja miała miejsce w odpowiedzi na intensywną syntezę rs-YFP, a nie wyniku specyficznego działania *Hac1*. Podobnie, w oparciu o profile DEG zilustrowane na **Rycinie 18** wydaje się, że geny zlokalizowane w jądrze komórkowym (głównie TFs) były bardziej wrażliwe na nadmierną syntezę rs-YFP niż na zwiększone ilości transkryptu *HAC1*. Podobnie, indukcja wielu genów zlokalizowanych w mitochondriach (**Rycina 18**) i zaangażowanych w energetykę komórkową w szczepie *HAC1_scYFP* jest konsekwencją jednoczesnej nadekspresji dwóch heterologicznych genów i wysokoenergetycznego procesu wydzielania rs-Prot. Pod tym względem, szczepy HSS - *scYFP* i *SoA* (**Rycina 7**), wykazują częściowe podobieństwo.

Pomimo wysokiej nadreprezentacji transkryptów *HAC1* (**Rycina 16A**), liczebność formy splicingowanej w szczepie *HAC1_scYFP* były porównywalne z poziomem obserwowanym w kontroli prototroficznej i w szczepie *scYFP* bez UPR (**Rycina 16B**). Powszechnie przyjmuje się, że dla skutecznego działania transkrypt *HAC1* wymaga niekonwencjonalnego splicingu, w celu usunięcia hamującej translację krótkiej sekwencji intronu [44,46,154]. Wyniki przedstawione na (**Rycina 16A.B**) sugerują, że **sama akumulacja substratu** (tj. mRNA *HAC1*) **jest niewystarczająca dla tworzenia efektywnego = splicingowanego transkryptu *HAC1* w *Y. lipolytica***. Wcześniejsze badania pokazały, że akumulacja mRNA *XBPI* (homolog *HAC1* z drożdży; analiza sekwencji Fig. S5 w **P.4**) była warunkiem wstępnym i wystarczającym dla następującej reakcji splicingu [155]. Warto tu zaznaczyć, że dojrzewanie transkryptu *XBPI* przebiega bardzo podobnie, jak *HAC1*, i w obu pośredniczy rezydująca w ER nukleaza Ire1 [187]. Na podstawie obserwowanych tutaj wzorców akumulacji dojrzałych form transkryptu *HAC1* (**Rycina 16B**) i dystrybucji białka YFP (**Tabela 5**), można stwierdzić, że **wysoka akumulacja YFP wewnątrz komórki** (najwyższa w szczepie *scYFP*) **nie promuje zajścia splicingu transkryptu *HAC1*** (najniższej w szczepach *scYFP* i *HAC1_scYFP*). **Akumulacja prawidłowo sfalowanego YFP** (bo fluorescencyjnie aktywne) **nie inicjuje niezbędnego sygnału** (ekspozycji domeny hydrofobowej) **dla zaistnienia UPR**, nawet gdy białko zostało nagromadzone na wysokim poziomie. Stąd też pytanie o mechanizm leżącego u podstaw obserwowanej poprawy wydzielania rs-YFP pomimo porównywalnych ilości splicingowanego transkryptu *HAC1*

w szczepie *HAC1_scYFP* i szczepach kontrolnych (**Rycina 16A.B.**). Jak wykazano dla Xbp1 [187], obie formy, bez intronu i z intronem, mogą ulegać translacji i modulować komórkową aktywność UPR poprzez bezpośrednie wiązanie z komplementarnymi elementami cis; natomiast wariant poprawnie złożonego transkryptu jest znacznie wydajniejszy. Obecnie są to jedynie spekulacje, które wymagają dalszych dogłębnych badań.

Co ciekawe, wśród genów o **obniżonej ekspresji** w szczepach nadprodukujących *scYFP* zidentyfikowano **ligazę tRNA Trl1** zaangażowaną w biogenezę tRNA. Interesująco, jedną z funkcji Trl1 jest ligacja eksonów po procesie niekonwencjonalnego splicingu transkryptu *HAC1/XBP1* [188]. Oprócz *HAC1* i *IRE1*, *TRL1* (*RLG1* u *S. cerevisiae*) jest uważany za **kluczowy element odpowiedzi UPR** u *S. cerevisiae* [187]. Represję *TRL1* zaobserwowano we wszystkich szczepach *Y. lipolytica* z nadekspresją wydzielniczego r-Prot (o 14%, 33%, 14% i 24% odpowiednio dla szczepu *HAC1_scYFP*, *scYFP*, *TIG* i *SoA*), ale nie dla *inYFP*, który jest jedynym przypadkiem, w którym zaobserwowano znaczny wzrost liczby transkryptów *HAC1* po splicingu (**Rycina 16B**). Chociaż nie wyjaśnia, w jaki sposób nadekspresja *HAC1* przyczyniła się do zwiększonego wydzielania rs-Prot, sugeruje, że **ligaza Trl1 jest rzeczywiście zaangażowana w niekonwencjonalny splicing *HAC1* / *XBP1* w *Y. lipolytica*** i jest homologiem Rlg1.

Podsumowując, deregulacja ekspresji genów w szczepie *HAC1_scYFP* i wyżej wymienione korelacje między unikalnymi DEGs a wygenerowanymi fenotypami mogą wynikać z wielu czynników jak: intensywna synteza *scYFP*, obciążenie metaboliczne spowodowane ekspresją dwóch genów (*HAC1* i *SCYFP*) lub nadreprezentacją transkryptu *HAC1* (a zwłaszcza - towarzyszący stały poziom *SLS1*, co jest ważnym odkryciem tego badania) lub być konsekwencją innych niezidentyfikowanych dotąd zjawisk.

7. Podsumowanie i kluczowe wnioski

1. Porównanie reakcji fizjologicznej komórek *Y. lipolytica* utrzymywanych w stanie ustalonym dostarczyły wyników popierających następujące stwierdzenia:
 - 1.1. Skierowanie białka do szlaku sekrecyjnego pozwala na uzyskanie większych jego ilości (scYFP vs. inYFP),
 - 1.2. Małe polipeptydy o niskim poziomie modyfikacji potranslacyjnych są syntetyzowane na wyższym poziomie,
 - 1.3. Szczepy skonstruowane w podłożu genetycznym Po1h okazały się być lepszymi producentami r(s)-Prots. Obserwacja ta nie wynika wyłącznie z zastosowania w tych szczepach nieco bardziej aktywnego transkrypcyjnie promotora (4UAS_pTEF vs. hp4d; różnica nieistotna statystycznie), ale inne czynniki także wpłynęły na taki wynik. W tym kontekście, szczepy konstruowane na bazie szczepu Po1h generowały mniej biomasy, przy jednoczesnym większym tempie utylizacji substratu. Uproszczona analiza bilansu masowego wykazała, że w tych szczepach większy strumień węgla był kierowany na potrzeby syntezy badanych r(s)-Prots.
 - 1.4. W kontekście różnic białko-zależnych, szczepy nienależące do grupy HSS („wydajnych producentów”), wykazywały niższe tempo utylizacji substratu i niższy przyrost biomasy w porównaniu z odpowiednikami HSS, konstruowanymi w danym podłożu genetycznym. Wydaje się zasadne stwierdzenie, że to właściwości biochemiczne nadprodukowanego białka, jako jedyna kontrolowana zmienna, przyczyniły się do takiego wyniku.
2. Ciekawych obserwacji dokonano w kontekście zachowania/braku liniowości następujących po sobie procesów biologicznych w szlaku syntezy i sekrecji białka, w zależności od właściwości biochemicznych r(s)-Prot i temperatury prowadzenia procesu:
 - 2.1. Ekspresja genu kodującego białko zasobne w wiązania disiarczkowe (-S-S-) była najwyższa spośród badanych, co jednak nie przyczyniło się do uzyskania wysokich ilości finalnego białka, aktywnego enzymatycznie. Tak wysoka dysproporcja nie była obserwowana dla innych białek badanych. Wydaje się zasadne stwierdzenie, że tak, jak i w przypadku innych systemów drożdżowych, energo- i materiałochłonny proces oksydacyjnego fałdowania stanowi wąskie gardło szlaku translacyjno-sekrecyjnego i przyczynia się do dysproporcji pomiędzy poziomem transkrypcji i syntezy polipeptydu.

- 2.2. Zaobserwowano, że obniżenie temperatury prowadzenia hodowli szczepów ko-nadprodukujących białko reporterowe i białko pomocnicze SH, przyczynia się do ograniczenia transkrypcji genu reporterowego przy jednoczesnym zwiększeniu syntezy dojrzałego polipeptydu. Należy nadmienić, że obserwacji tej dokonano na przykładzie małego białka o niskim stopniu modyfikacji posttranslacyjnych. W tym wypadku, zgodnie z klasycznym modelem, szlak sekrecyjny nie jest przeciążony, a w konsekwencji znacznie zwiększona jest liniowość pomiędzy następującymi po sobie procesami w szlaku syntezy i sekrecji białka.
- 2.3. Co ważne, efekt ten nie jest widoczny w przypadku ko-nadekspresji *SH* zaangażowanych w syntezę i fałdowanie białek w cytoplazmie. W tym wypadku, pomimo intensyfikacji transkrypcji genu, ilość wytworzonego białka została zwiększona, co było szczególnie widoczne przy zastosowaniu obniżonej temperatury. Jednocześnie, obserwacja ta wskazuje, że w tym przypadku, intensywna transkrypcja nie stanowi czynnika zaburzającego liniowość procesów w szlaku syntezy i sekrecji białek, ale to działanie samego szlaku wydzielniczego, jest uwalniane przy niższej temperaturze.
3. Właściwości biochemiczne i/lub intensywność procesu syntezy r(s)-Prots wpływa na regulację morfotypu komórki *Y. lipolytica*. Szczepy wysoko-produkujące (HSS) w skali populacji charakteryzowały się morfotypem mieszanym, natomiast szczepy nisko-produkujące wykazywały morfotyp owoidalny. Biorąc pod uwagę, że morfotyp filamentowy jest markerem inicjacji reakcji na stres w komórkach *Y. lipolytica*, wydaje się zasadne stwierdzenie, że populacja komórek HSS w momencie przeprowadzenia badania była w trakcie przejścia dimorficznego w kierunku morfotypu filamentowego. Za tym stwierdzeniem przemawia profil transkrypcyjny genu zaangażowanego w indukcję filamentacji *SOK1* i genów zaangażowanych w stres oksydacyjny. Jednakże profil czynników indukcji filamentacji (*FLO11*, *MHY1*, *MGF2*) przeczy temu stwierdzeniu. Dlatego dalsze badania w tym zakresie są niezbędne, aby jednoznacznie stwierdzić rodzaj przejścia dimorficznego w szczepach HSS.
4. Profile transkryptomiczne wygenerowane w stanie ustalonym przez badane szczepy jednoznacznie wskazują, że kierowanie polipeptydu do przestrzeni zewnątrzkomórkowej indukuje masowe rearanżacje w komórce, przekraczające różnice związane z właściwościami biochemicznymi białek. Innymi słowy, sekrecja ma większy wpływ na liczbę DEGs niż właściwości biochemiczne syntetyzowanego rs-Prot.

5. Szczepy wysoko-produkujące (HSS) charakteryzowały się istotną intensyfikacją procesów transkrypcji i translacji, co wnioskowano na podstawie masowej indukcji genów zaangażowanych w transkrypcję, składanie rybosomów i regulację tych procesów, oraz represji genu *RsfA* kodującego czynnik transkrypcyjny wyciszający te procesy komórkowe.
6. Wysoka nadprodukcja rs-Protów w komórkach *Y. lipolytica* jest nieodłącznie związana z inicjacją stresu oksydacyjnego zachodzącego w ER, jak potwierdza profil ekspresji licznych genów. Co ważne, reakcja na stres oksydacyjny w komórkach *Y. lipolytica* jest zainicjowana niezależnie od liczby wiązań -S-S- w strukturze drugorzędowej rs-Protów. Wydaje się, że omówiona w punkcie 5. intensyfikacja procesów transkrypcji i translacji, przyczyniła się do „niezrównoważonej” syntezy białka, a w konsekwencji inicjację stresu oksydacyjnego.
7. Zaobserwowano, że wysokiemu poziomowi syntezy rs-Protów (HSS) towarzyszy zatrzymanie progresji cyklu komórkowego w fazie G₀. To stwierdzenie nie współgra z obserwowanym wyższym przyrostem biomasy szczepów HSS w porównaniu do szczepów nisko-produkujących, ale profil transkrypcyjny wyraźnie wskazuje na zajście tego zjawiska. Co należy zaznaczyć, zatrzymanie cyklu komórkowego w szczepie z syntetycznie indukowanym UPR *HAC1-scYFP* wynika z nadprodukcji rs-Protów, a nie jest efektem działania Hac1.
8. Jak wykazano, ani syntetycznie indukowana, ani natywnie zainicjowana indukcja ekspresji *HAC1* nie prowadzi do deregulacji genów, typowo wskazywanych jako elementy regulowane przez Hac1 (*KAR2 IRE1, ERJ, ERO1, GSH* czy *PDII*). Natomiast, jak zaobserwowano profil ekspresji genu *HAC1* jest powiązany z profilami ekspresji genów *SLS1, XBP1, TRL1*. Wydaje się, że geny te są zależne od aktywności Hac1 jako czynnika transkrypcyjnego w komórkach *Y. lipolytica*. W tym kontekście należy wspomnieć, że indukcja ekspresji genu *SLS1* każdorazowo współwystępowała z nadekspresją genu kodującego białko sekrecyjne.
9. Cechą charakterystyczną dla UPR powiązaną z *HAC1* w komórkach *Y. lipolytica* jest wzmożone kierowanie białek do wakuoli w celu degradacji oraz indukcja autofagii. Natomiast, sama proteoliza jest promowana przy nadprodukcji jakiegokolwiek białka.
10. Obserwowany w szczepie *inYFP HAC1*-zależny UPR nie jest wynikiem klasycznego ER-zależnego procesu, bowiem białko to nie jest kierowane do sekrecji. Jak pokazały analizy porównawcze profili transkryptomicznych badanych szczepów i profili Hac1- i Hsf1-zależnych z *S. cerevisiae*, indukcja UPR w szczepie *inYFP* wynika z interakcji pomiędzy

odpowiedzią Hsf1-zależnego cytozolowego UPR i właściwej dla ER Hac1-zależnej odpowiedzi. Mechanizm tej interakcji nie jest poznany.

11. Dane transkryptomyczne nieoczekiwanie wykazały ograniczenie ekspresji genów zaangażowanych w dystalne etapy sekrecji polipeptydów w szczepach HSS. Jednakże, zastosowanie jednego z nich, *Sso1*, jako białka pomocniczego pozwoliło na zwiększenie możliwości wydzielniczych rekombinowanego szczepu, szczególnie w warunkach, gdy szlak sekrecyjny działał z pełnym obciążeniem (30 vs. 25 °C).
12. Uzyskane dane sugerują, że przynajmniej część zsyntetyzowanych polipeptydów rs-Prot była translokowana na drodze translokacji post-translacyjnej i ulegała sekrecji na drodze niekonwencjonalnego eksportu. Wniosek ten sformułowano na podstawie obserwacji dotyczących: (i) ograniczonej częstości interakcji Kar2-Ire1 (ograniczenie ekspresji *SLS1*), czyli kompleksu niezbędnego dla translokacji ko-translacyjnej, (ii) braku deregulacji genów związanych z mechanizmem translokacji ko-translacyjnej (receptora SRP i samego SRP), a także (iii) silnej indukcji genów kodujących cytozolowe białka chaperonowe z rodziny *Ssa*. W tym kontekście niezwykle ciekawy był wzrost ekspresji genu *NCE102* kodującego białko zaangażowane w eksport niekonwencjonalny.
13. Kluczowym aplikacyjnym osiągnięciem niniejszej pracy jest opracowanie kombinacji modyfikacji genetycznej (ko-nadekspresji *SSA5*, *SSA8*, *RPL3*) oraz warunków środowiskowych (temp. 25 °C), które łącznie przyczyniają się do bardzo wysokiej intensyfikacji produkcji rs-Prot w komórkach *Y. lipolytica*.

8. Literatura

- [1] A. Milton, V. Gomes, T.S. Carmo, L.S. Carvalho, Comparison of Yeasts as Hosts for Recombinant Protein Production, *Microorganisms*. 6 (2018) 38. <https://doi.org/10.3390/microorganisms6020038>.
- [2] A.L. Demain, P. Vaishnav, Production of recombinant proteins by microbes and higher organisms, *Biotechnol. Adv.* 27 (2009) 297–306. <https://doi.org/10.1016/j.biotechadv.2009.01.008>.
- [3] C. Madzak, *Yarrowia lipolytica*: recent achievements in heterologous protein expression and pathway engineering, *Appl. Microbiol. Biotechnol.* 99 (2015) 4559–4577. <https://doi.org/10.1007/s00253-015-6624-z>.
- [4] C. Madzak, *Yarrowia lipolytica* Strains and Their Biotechnological Applications: How Natural Biodiversity and Metabolic Engineering Could Contribute to Cell Factories Improvement, *J. Fungi*. 7 (2021) 548. <https://doi.org/10.3390/jof7070548>.
- [5] E. Celińska, J.-M. Nicaud, Filamentous fungi-like secretory pathway strayed in a yeast system: peculiarities of *Yarrowia lipolytica* secretory pathway underlying its extraordinary performance, *Appl. Microbiol. Biotechnol.* 103 (2019) 39–52. <https://doi.org/10.1007/s00253-018-9450-2>.
- [6] C. Madzak, C. Gaillardin, J. Beckerich, Heterologous protein expression and secretion in the non-conventional yeast *Yarrowia lipolytica*: a review, *J. Biotechnol.* 109 (2004) 63–81. <https://doi.org/10.1016/j.jbiotec.2003.10.027>.
- [7] M. Groenewald, T. Boekhout, C. Neuvéglise, C. Gaillardin, P.W.M. van Dijck, M. Wyss, *Yarrowia lipolytica*: Safety assessment of an oleaginous yeast with a great industrial potential, *Crit. Rev. Microbiol.* 40 (2014) 187–206. <https://doi.org/10.3109/1040841X.2013.770386>.
- [8] C.W. Theron, M. Vandermies, S. Telek, S. Steels, P. Fickers, Comprehensive comparison of *Yarrowia lipolytica* and *Pichia pastoris* for production of *Candida antarctica* lipase B, *Sci. Rep.* (2020) 1–9. <https://doi.org/10.1038/s41598-020-58683-3>.
- [9] E. Celińska, M. Borkowska, W. Białas, Enhanced production of insect raw-starch-digesting alpha-amylase accompanied by high erythritol synthesis in recombinant *Yarrowia lipolytica* fed-batch cultures at high-cell-densities, *Process Biochem.* 52

- (2017) 78–85. <https://doi.org/10.1016/j.procbio.2016.10.022>.
- [10] M. Kubiak, M. Borkowska, P. Korpys, Feeding strategy impacts heterologous protein production in *Yarrowia lipolytica* fed - batch cultures — Insight into the role of osmolarity, *Yeast*. 36 (2019) 305–318. <https://doi.org/10.1002/yea.3384>.
- [11] P. Swietalski, F. Hetzel, I. Seitzl, L. Fischer, Secretion of a low and high molecular weight β - glycosidase by *Yarrowia lipolytica*, *Microb. Cell Fact.* 19 (2020) 1–13. <https://doi.org/10.1186/s12934-020-01358-5>.
- [12] S. Muller, T. Sandal, P. Kamp-Hansen, D. H., Comparison of Expression Systems in the Yeasts *Kluyveromyces lactis*, *Schizosaccharomyces pombe* and *Yarrowia lipolytica*. Cloning of Two Novel Promoters from *Yarrowia lipolytica*, *Yeast*. 1283 (1998) 1267–1283. [https://doi.org/https://doi.org/10.1002/\(SICI\)1097-0061\(1998100\)14:14<1267::AID-YEA327>3.0.CO;2-2](https://doi.org/https://doi.org/10.1002/(SICI)1097-0061(1998100)14:14<1267::AID-YEA327>3.0.CO;2-2).
- [13] E. Celińska, R. Ledesma-Amaro, M. Larroude, T. Rossignol, C. Pauthenier, J. Nicaud, Brief report Golden Gate Assembly system dedicated to complex pathway manipulation in *Yarrowia lipolytica*, *Microb. Biotechnol.* 10 (2017) 450–455. <https://doi.org/10.1111/1751-7915.12605>.
- [14] C. Holkenbrink, M.I. Dam, K.R. Kildegaard, J. Beder, J. Dahlin, D. Dom, I. Borodina, EasyCloneYALI: CRISPR / Cas9-Based Synthetic Toolbox for Engineering of the Yeast *Yarrowia lipolytica*, *Biotechnol. J.* 13 (2018) 1–8. <https://doi.org/10.1002/biot.201700543>.
- [15] E. Celińska, M. Borkowska, W. Białas, P. Korpys, J.M. Nicaud, Robust signal peptides for protein secretion in *Yarrowia lipolytica*: identification and characterization of novel secretory tags, *Appl. Microbiol. Biotechnol.* 102 (2018) 5221–5233. <https://doi.org/10.1007/s00253-018-8966-9>.
- [16] E. Celińska, M. Borkowska, P. Korpys-Woźniak, M. Kubiak, J. Nicaud, P. Kubiak, M. Gorczyca, W. Białas, Optimization of *Yarrowia lipolytica*- based consolidated biocatalyst through synthetic biology approach : transcription units and signal peptides shuffling, *Appl. Microbiol. Biotechnol.* 104 (2020) 5845–5859. <https://doi.org/https://doi.org/10.1007/s00253-020-10644-6>.
- [17] Y. Park, P. Korpys, M. Kubiak, P. Soudier, P. Tr, M. Larroude, T. Rossignol, J. Nicaud, Engineering the architecture of erythritol-inducible promoters for regulated and enhanced gene expression in *Yarrowia lipolytica*, *FEMS Yeast Res.* 19 (2019) 1–

13. <https://doi.org/10.1093/femsyr/foy105>.
- [18] M. Kabani, J.-M. Beckerich, C. Gaillardin, Sls1p Stimulates Sec63p-Mediated Activation of Kar2p in a Conformation-Dependent Manner in the Yeast Endoplasmic Reticulum, *Mol. Cell. Biol.* 20 (2000) 6923–6934. <https://doi.org/10.1128/mcb.20.18.6923-6934.2000>.
- [19] A. Boisramé, J.M. Beckerich, C. Gaillardin, Sls1p, an endoplasmic reticulum component, is involved in the protein translocation process in the yeast *Yarrowia lipolytica*, *J. Biol. Chem.* 271 (1996) 11668–11675. <https://doi.org/10.1074/jbc.271.20.11668>.
- [20] A. Boisramé, M. Chasles, A. Babour, J. Beckerich, C. Gaillardin, Sbh1p, a subunit of the Sec61 translocon, interacts with the chaperone calnexin in the yeast *Yarrowia lipolytica*, *J. Cell Sci.* 115 (2002) 4947–4956. <https://doi.org/10.1242/jcs.00187>.
- [21] A. Babour, M. Kabani, A. Boisramé, J. Beckerich, Characterization of Ire1 in the yeast *Yarrowia lipolytica* reveals an important role for the Sls1 nucleotide exchange factor in unfolded protein response regulation, *Curr. Genet.* 53 (2008) 337–346. <https://doi.org/10.1007/s00294-008-0190-1>.
- [22] J. Beckerich, A. Boisramé-Baudevin, C. Gaillardin, *Yarrowia lipolytica*: a model organism for protein secretion studies, *Internati. Microbiol.* 1 (1998) 123–130. <https://doi.org/doi:10.1016/j.jbiotec.2003.10.027>.
- [23] A. Boisramé, C. Gaillardin, Heterologous expression and characterization of a β -1, 6-glucanase from *Aspergillus fumigatus*, (2009) 663–669. <https://doi.org/10.1007/s00253-008-1780-z>.
- [24] C. Madzak, L. Otterbein, M. Chamkha, S. Moukha, M. Asther, C. Gaillardin, J. Beckerich, Heterologous production of a laccase from the basidiomycete *Pycnoporus cinnabarinus* in the dimorphic yeast *Yarrowia lipolytica*, 5 (2005) 635–646. <https://doi.org/10.1016/j.femsyr.2004.10.009>.
- [25] K.R. Love, T.J. Politano, V. Panagiotou, B. Jiang, T.A. Stadheim, C. Love, Systematic Single-Cell Analysis of *Pichia pastoris* Reveals Secretory Capacity Limits Productivity, *PLoS One.* 7 (2012) 1–11. <https://doi.org/10.1371/journal.pone.0037915>.
- [26] D. Swennen, C. Henry, J.M. Beckerich, Folding proteome of *Yarrowia lipolytica* targeting with uracil permease mutants, *J. Proteome Res.* 9 (2010) 6169–6179.

<https://doi.org/10.1021/pr100340p>.

- [27] D. Swennen, J. Beckerich, *Yarrowia lipolytica* vesicle-mediated protein transport pathways, *BMC Evol. Biol.* 19 (2007) 1–19. <https://doi.org/10.1186/1471-2148-7-219>.
- [28] M. Delic, M. Valli, A.B. Graf, M. Pfeffer, D. Mattanovich, B. Gasser, The secretory pathway: exploring yeast diversity, *FEMS Microbiol. Rev.* 37 (2013) 872–914. <https://doi.org/10.1111/1574-6976.12020>.
- [29] C.M. Dobson, Principles of protein folding, misfolding and aggregation, *Semin. Cell&Developmental Biol.* 15 (2004) 3–16. <https://doi.org/10.1016/j.semcdb.2003.12.008>.
- [30] T. Anelli, R. Sitia, Protein quality control in the early secretory pathway, *EMBO J.* 27 (2008) 315–327. <https://doi.org/10.1038/sj.emboj.7601974>.
- [31] E. Celińska, M. Borkowska, W. Białas, P. Korpys, J. Nicaud, Robust signal peptides for protein secretion in *Yarrowia lipolytica*: identification and characterization of novel secretory tags, *Appl. Microbiol. Biotechnol.* 102 (2018) 5221–5233. <https://doi.org/10.1007/s00253-018-8966-9>.
- [32] F. Darvishi, A. Zarei, C. Madzak, In silico and in vivo analysis of signal peptides effect on recombinant glucose oxidase production in nonconventional yeast *Yarrowia lipolytica*, *World J. Microbiol. Biotechnol.* 34 (2018) 1–8. <https://doi.org/10.1007/s11274-018-2512-x>.
- [33] N. Martineau, D. Sharma, Function of SSA Subfamily of Hsp70 Within and Across Species Varies Widely in Complementing *Saccharomyces cerevisiae* Cell Growth and Prion Propagation, *PLoS One.* 4 (2009) e6644. <https://doi.org/10.1371/journal.pone.0006644>.
- [34] D.T.W. Ng, J.D. Brown, P. Walter, Signal Sequences Specify the Targeting Route to the Endoplasmic Reticulum Membrane, *J. Cell Biol.* 134 (1996) 269–278. <https://doi.org/https://doi.org/10.1083/jcb.134.2.269>.
- [35] D.S. Yaver, S. Matoba, D.M. Ogrydziak, A Mutation in the Signal Recognition Particle 7S RNA of the Yeast *Yarrowia lipolytica* Preferentially Affects Synthesis of the Alkaline Extracellular Protease: In vivo Evidence for Translational Arrest, *J. Cell Biol.* 116 (1992) 60–616. <https://doi.org/https://doi.org/10.1083/jcb.116.3.605>.
- [36] A. Boisramé, M. Kabani, J.M. Beckerich, E. Hartmann, C. Gaillardin, Interaction of

- Kar2p and Sls1p is required for efficient co- translational translocation of secreted proteins in the yeast *Yarrowia lipolytica*, *J. Biol. Chem.* 273 (1998) 30903–30908. <https://doi.org/10.1074/jbc.273.47.30903>.
- [37] F. Hatahet, L.W. Ruddock, R. Editors, K. Ahn, A. Benham, D. Craik, L. Ellgaard, D. Ferrari, S. Ventura, Protein Disulfide Isomerase : A Critical Evaluation of Its Function in Disulfide Bond Formation, *Antioxid. Redox Signal.* 11 (2009) 2807–2850. <https://doi.org/https://doi.org/10.1089/ars.2009.2466>.
- [38] M. Delic, C. Rebnegger, F. Wanka, V. Puxbaum, C. Haberhauer-troyer, S. Hann, G. Köllensperger, D. Mattanovich, B. Gasser, Oxidative protein folding and unfolded protein response elicit differing redox regulation in endoplasmic reticulum and cytosol of yeast, *Free Radic. Biol. Med.* 52 (2012) 2000–2012. <https://doi.org/10.1016/j.freeradbiomed.2012.02.048>.
- [39] N. Berner, K. Reutter, D.H. Wolf, Protein Quality Control of the Endoplasmic Reticulum and Ubiquitin – Proteasome- Triggered Degradation of Aberrant Proteins : Yeast Pioneers the Path, *Annu. Rev. Biochem.* 87 (2018) 751–782. <https://doi.org/https://doi.org/10.1146/annurev-biochem-062917-012749>.
- [40] G.M. Preston, C.J. Guerriero, M.B. Metzger, S. Michaelis, J.L. Brodsky, G.M. Preston, C.J. Guerriero, M.B. Metzger, S. Michaelis, J.L. Brodsky, Substrate Insolubility Dictates Hsp104-Dependent Article Substrate Insolubility Dictates Hsp104-Dependent Endoplasmic-Reticulum-Associated Degradation, *Mol. Cell.* 70 (2018) 242–253. <https://doi.org/10.1016/j.molcel.2018.03.016>.
- [41] C. Patil, P. Walter, Intracellular signaling from the endoplasmic reticulum to the nucleus : the unfolded protein response in yeast and mammals, *Curr. Opin. Cell Biol.* 13 (2001) 349–356. [https://doi.org/https://doi.org/10.1016/S0955-0674\(00\)00219-2](https://doi.org/https://doi.org/10.1016/S0955-0674(00)00219-2).
- [42] M.H. Oh, S.A. Cheon, H.A. Kang, J. Kim, Functional characterization of the unconventional splicing of *Yarrowia lipolytica* HAC1 mRNA induced, *Yeast.* 1 (2010) 443–452. <https://doi.org/10.1002/yea>.
- [43] N.L.M. Callewaert, W. Verecken, K. Jacqueline, M. De Pourcq, M. Be, S.C. Jozef, *YARROWIA LIPOLYTICA AND PICHIA PASTORIS* HAC1 NUCLEiC ACIDS, *United Strat. Pat.* 2 (2011).
- [44] M. Guerfal, S. Ryckaert, P.P. Jacobs, P. Ameloot, K. Van Craenenbroeck, R. Derycke, N. Callewaert, The HAC1 gene from *Pichia pastoris* : characterization and effect of its

- overexpression on the production of secreted , surface displayed and membrane proteins, *Microb. Cell Fact.* 9 (2010) 1–12. <https://doi.org/https://doi.org/10.1186/1475-2859-9-49>.
- [45] M. Hernandez-Elvira, F. Torres-Quiroz, A. Escamilla-Ayala, E. Dominguez-Martin, R. Escalante, L. Kawasaki, L. Ongay-Larios, R. Coria, The Unfolded Protein Response Pathway in the Yeast *Kluyveromyces lactis* . A Comparative View among Yeast Species, *Cells.* 7 (2018) 106. <https://doi.org/10.3390/cells7080106>.
- [46] G. Whyteside, R. Mat, M.J.C. Alcocer, D.B. Archer, Activation of the unfolded protein response in *Pichia pastoris* requires splicing of a HAC1 mRNA intron and retention of the C-terminal tail of Hac1p, *FEBS Lett.* 585 (2011) 1037–1041. <https://doi.org/10.1016/j.febslet.2011.02.036>.
- [47] H. Raschmanová, A. Weninger, Z. Knejzlik, K. Melzoch, K. Kovar, Engineering of the unfolded protein response pathway in *Pichia pastoris*: enhancing production of secreted recombinant proteins, *Appl. Microbiol. Biotechnol.* 105 (2021) 4397–4414. <https://doi.org/https://doi.org/10.1007/s00253-021-11336-5>.
- [48] H. Kim, S.J. Yoo, H.A. Kang, Yeast synthetic biology for the production of recombinant therapeutic proteins, *FEMS Yeast Res.* 15 (2014). <https://doi.org/10.1111/1567-1364.12195>.
- [49] K. De Pourcq, W. Vervecken, I. Dewerte, A. Valevska, A. Van Hecke, N. Callewaert, Engineering the yeast *Yarrowia lipolytica* for the production of therapeutic proteins homogeneously glycosylated with Man 8 GlcNAc 2, *Microb. Cell Fact.* 11 (2012) 1–12. <https://doi.org/https://doi.org/10.1186/1475-2859-11-53>.
- [50] K.E.J. Tyo, Z. Liu, D. Petranovic, J. Nielsen, Imbalance of heterologous protein folding and disulfide bond formation rates yields runaway oxidative stress, *BMC Biol.* 10 (2012) 1–14.
- [51] P. Xu, A.S. Robinson, Decreased Secretion and Unfolded Protein Response Up-Regulation Are Correlated With Intracellular Retention for Single-Chain Antibody Variants Produced in Yeast, *Biotechnol. Bioeng.* 104 (2009) 20–29. <https://doi.org/10.1002/bit.22376>.
- [52] V. Puxbaum, D. Mattanovich, B. Gasser, Quo vadis? The challenges of recombinant protein folding and secretion in *Pichia pastoris*, *Appl. Microbiol. Biotechnol.* 99 (2015) 2925–2938. <https://doi.org/10.1007/s00253-015-6470-z>.

- [53] J. Hou, T. Österlund, Z. Liu, D. Petranovic, J. Nielsen, T. Osterlund, Z. Liu, D. Petranovic, J. Nielsen, Heat shock response improves heterologous protein secretion in *Saccharomyces cerevisiae*, *Appl. Microbiol. Biotechnol.* 97 (2013) 3559–3568. <https://doi.org/10.1007/s00253-012-4596-9>.
- [54] J. Hou, H. Tang, Z. Liu, T. Osterlund, J. Nielsen, D. Petranovic, Management of the endoplasmic reticulum stress by activation of the heat shock response in yeast, *FEMS Yeast Res.* 14 (2014) 481–494. <https://doi.org/10.1111/1567-1364.12125>.
- [55] B. Gasser, M. Maurer, J. Rautio, M. Sauer, A. Bhattacharyya, M. Saloheimo, M. Penttilä, D. Mattanovich, Monitoring of transcriptional regulation in *Pichia pastoris* under protein production conditions, *BMC Genomics.* 8 (2007) 1–18. <https://doi.org/10.1186/1471-2164-8-179>.
- [56] R.J. Zahrl, D.A. Peña, D. Mattanovich, B. Gasser, Systems biotechnology for protein production in *Pichia pastoris*, *FEMS Yeast Res.* 17 (2017) 1–31. <https://doi.org/10.1093/femsyr/fox068>.
- [57] R.J. Zahrl, B. Gasser, D. Mattanovich, P. Ferrer, Detection and elimination of cellular bottlenecks in protein-producing yeasts, *Methods Mol. Biol.* 1923 (2019) 75–95. https://doi.org/10.1007/978-1-4939-9024-5_2.
- [58] E.J. Thak, S.J. Yoo, H.Y. Moon, H.A. Kang, Yeast synthetic biology for designed cell factories producing secretory recombinant proteins, *FEMS Yeast Res.* 20 (2020) foaa009. <https://doi.org/https://doi.org/10.1093/femsyr/foaa009>.
- [59] X. Xu, Y. Liu, G. Du, R. Ledesma-amaro, Microbial Chassis Development for Natural Product Biosynthesis, *Trends Biotechnol.* 38 (2020) 779–796. <https://doi.org/10.1016/j.tibtech.2020.01.002>.
- [60] D. Petranovic, J. Nielsen, Z. Liu, K.E.J. Tyo, Different Expression Systems for Production of Recombinant Proteins in *Saccharomyces cerevisiae*, *Biotechnol. Bioeng.* 109 (2012) 1259–1268. <https://doi.org/10.1002/bit.24409>.
- [61] E. Celinska, W. Bialas, M. Borkowska, Cloning , expression, and purification of insect (*Sitophilus oryzae*) alpha-amylase, able to digest granular starch, in *Yarrowia lipolytica* host, *Appl. Microbiol. Biotechnol.* 99 (2015) 2727–2739. <https://doi.org/10.1007/s00253-014-6314-2>.
- [62] R. Dulermo, F. Brunel, T. Dulermo, R.L. Amaro, J. Vion, M. Trassaert, S. Thomas,

- J.M. Nicaud, C. Leplat, Using a vector pool containing variable - strength promoters to optimize protein production in *Yarrowia lipolytica*, *Microb. Cell Fact.* (2017) 1–11. <https://doi.org/10.1186/s12934-017-0647-3>.
- [63] L. Liu, P. Otoupal, A. Pan, H.S. Alper, Increasing expression level and copy number of a *Yarrowia lipolytica* plasmid through regulated centromere function, *FEMS Yeast Res.* 14 (2014) 1124–1127. <https://doi.org/10.1111/1567-1364.12201>.
- [64] Y. Chen, S. Partow, G. Scalcinati, V. Siewers, J. Nielsen, Enhancing the copy number of episomal plasmids in *Saccharomyces cerevisiae* for improved protein production, *FEMS Yeast Res.* 12 (2012) 598–607. <https://doi.org/10.1111/j.1567-1364.2012.00809.x>.
- [65] P. Pyati, E. Fitches, J.A. Gatehouse, Optimising expression of the recombinant fusion protein biopesticide ω - hexatoxin - Hv1a / GNA in *Pichia pastoris*: sequence modifications and a simple method for the generation of multi - copy strains, *J. Ind. Microbiol. Biotechnol.* 41 (2014) 1237–1247. <https://doi.org/10.1007/s10295-014-1466-8>.
- [66] M. Kubiak, S. Bozena, S. Rabian, U. Jankowska, E. Celińska, Hyperosmolarity adversely impacts recombinant protein synthesis by *Yarrowia lipolytica* — molecular background revealed by quantitative proteomics, *Appl. Microbiol. Biotechnol.* (2022) 349–367. <https://doi.org/10.1007/s00253-021-11731-y>.
- [67] M. Gorczyca, J. Kaźmierczak, P. Fickers, E. Celińska, Synthesis of Secretory Proteins in *Yarrowia lipolytica*: Effect of Combined Stress Factors and Metabolic Load, *J. Ind. Microbiol. Biotechnol.* 23 (2022) 3602. <https://doi.org/https://doi.org/10.3390/ijms23073602>.
- [68] J. Hou, K. Tyo, Z. Liu, D. Petranovic, J. Nielsen, Engineering of vesicle trafficking improves heterologous protein secretion in *Saccharomyces cerevisiae*, *Metab. Eng.* 14 (2012) 120–127. <https://doi.org/10.1016/j.ymben.2012.01.002>.
- [69] B. Gasser, M. Sauer, M. Maurer, B. Gasser, M. Sauer, M. Maurer, G. Stadlmayr, D. Mattanovich, Transcriptomics-Based Identification of Novel Factors Enhancing Heterologous Protein Secretion in Yeasts, *Appl. Environ. Microbiol.* 73 (2007) 6499–6507. <https://doi.org/10.1128/AEM.01196-07>.
- [70] T. Lee, Y. Bae, M. Kim, J. Seo, Overexpression of HAC1 gene increased levels of both intracellular and secreted human kringle fragment in *Saccharomyces cerevisiae*,

- Process Biochem. 47 (2012) 2300–2305. <https://doi.org/10.1016/j.procbio.2012.09.006>.
- [71] M. Wu, Q. Shen, Y. Yang, S. Zhang, W. Qu, J. Chen, H. Sun, S. Chen, Disruption of YPS1 and PEP4 genes reduces proteolytic degradation of secreted HSA / PTH in *Pichia pastoris* GS115, J. Ind. Microbiol. Biotechnol. 40 (2013) 589–599. <https://doi.org/10.1007/s10295-013-1264-8>.
- [72] R.J. Zahrl, R. Prielhofer, J. Burgard, D. Mattanovich, B. Gasser, Synthetic activation of yeast stress response improves secretion of recombinant proteins, N. Biotechnol. 73 (2023) 19–28. <https://doi.org/10.1016/j.nbt.2023.01.001>.
- [73] L. Marsalek, C. Gruber, F. Altmann, M. Aleschko, B. Gasser, V. Puxbaum, T. Tulln, B. Gasser, Disruption of genes involved in CORVET complex leads to enhanced secretion of heterologous carboxylesterase only in protease deficient *Pichia pastoris*, Biotechnol. J. 12 (2017) 1600584. <https://doi.org/10.1002/biot.201600584>.
- [74] H.L. Zhao, C. Xue, Y. Wang, Increasing the cell viability and heterologous protein expression of *Pichia pastoris* mutant deficient in PMR1 gene by culture condition optimization, Appl. Microbiol. Biotechnol. 81 (2008) 235–241. <https://doi.org/10.1007/s00253-008-1666-0>.
- [75] M.M. Harmsen, M.I. Bruyne, H.A. Raue, J. Maat, Overexpression of binding protein and disruption of the PMRI gene synergistically stimulate secretion of bovine prochymosin but not plant Thaumatin in yeast, Appl. Environ. Microbiol. 46 (1996) 365–370. <https://doi.org/https://doi.org/10.1007/BF00166231>.
- [76] Y. Chiba, H. Akeboshi, New Era of Glycoscience : Intrinsic and Extrinsic Functions Performed by Glycans Glycan Engineering and Production of ‘ Humanized ’ Glycoprotein in Yeast Cells, Biol. Pharm. Bull. 32 (2009) 786–795. <https://doi.org/https://doi.org/10.1248/bpb.32.786>.
- [77] H. Tang, X. Bao, Y. Shen, M. Song, S. Wang, Engineering Protein Folding and Translocation Improves Heterologous Protein Secretion in *Saccharomyces cerevisiae*, Biotechnol. Bioeng. 112 (2015) 1872–1882. <https://doi.org/10.1002/bit.25596>.
- [78] R. Zahrl, R. Prielhofer, O. Ata, K. Baumann, D. Mattanovich, B. Gasser, Pushing and pulling proteins into the yeast secretory pathway enhances recombinant protein secretion, Metab. Eng. 74 (2022) 36–48. <https://doi.org/10.1016/j.ymben.2022.08.010>.
- [79] J.D. Smith, B.C. Tang, A.S. Robinson, Protein Disulfide Isomerase , But Not Binding

- Protein , Overexpression Enhances Secretion of a Non-Disulfide-Bonded Protein in Yeast, *Biotechnol. Bioeng.* 85 (2004) 340–350. <https://doi.org/10.1002/bit.10853>.
- [80] T. Payne, C. Finnis, L.R. Evans, D.J. Mead, S. V Avery, D.B. Archer, D. Sleep, Modulation of Chaperone Gene Expression in Mutagenized *Saccharomyces cerevisiae* Strains Developed for Recombinant Human Albumin Production Results in Increased Production of Multiple Heterologous Proteins, *Appl. Environ. Microbiol.* 74 (2008) 7759–7766. <https://doi.org/10.1128/AEM.01178-08>.
- [81] W. Zhang, H.L. Zhao, C. Xue, X.H. Xiong, X.Q. Yao, X.Y. Li, H.P. Chen, Z.M. Liu, Enhanced secretion of heterologous proteins in *Pichia pastoris* following overexpression of *Saccharomyces cerevisiae* chaperone proteins, *Biotechnol. Prog.* 22 (2006) 1090–1095. <https://doi.org/10.1021/bp060019r>.
- [82] T. Lodi, B. Neglia, C. Donnini, Secretion of Human Serum Albumin by *Kluyveromyces lactis* Overexpressing KIPDI1 and KIERO1, *Appl. Environ. Microbiol.* 71 (2005) 4359–4363. <https://doi.org/10.1128/AEM.71.8.4359>.
- [83] M. Delic, R. Go, D. Mattanovich, B. Gasser, Engineering of Protein Folding and Secretion — Strategies to Overcome Bottlenecks for Efficient Production of Recombinant Proteins, *Biotechnol. Bioeng.* 112 (2014) 1872–1882. <https://doi.org/10.1089/ars.2014.5844>.
- [84] J. Hou, K.E.J. Tyo, Z. Liu, D. Petranovic, J. Nielsen, Metabolic engineering of recombinant protein secretion by *Saccharomyces cerevisiae*, *FEMS Yeast Res.* 12 (2012) 491–510. <https://doi.org/10.1111/j.1567-1364.2012.00810.x>.
- [85] A. Idiris, H. Tohda, H. Kumagai, Engineering of protein secretion in yeast : strategies and impact on protein production, *Appl. Microbiol. Biotechnol.* 86 (2010) 403–417. <https://doi.org/10.1007/s00253-010-2447-0>.
- [86] J.C. De Ruijter, A.D. Frey, Analysis of antibody production in *Saccharomyces cerevisiae* : effects of ER protein quality control disruption, *Appl. Microbiol. Biotechnol.* 99 (2015) 9061–9071. <https://doi.org/10.1007/s00253-015-6807-7>.
- [87] R.J. Zahrl, D. Mattanovich, B. Gasser, The impact of ERAD on recombinant protein secretion in *Pichia pastoris* (syn *Komagataella spp.*), *Microbiology.* 164 (2018) 453–463. <https://doi.org/10.1099/mic.0.000630>.
- [88] M. Valkonen, M. Ward, H. Wang, M. Penttila, Improvement of Foreign-Protein

- Production in *Aspergillus niger* var . *awamori* by Constitutive Induction of the Unfolded-Protein Response, *Appl. Environ. Microbiol.* 69 (2003) 6979–6986. <https://doi.org/10.1128/AEM.69.12.6979>.
- [89] M. Valkonen, M. Penttilä, Effects of Inactivation and Constitutive Expression of the Unfolded- Protein Response Pathway on Protein Production in the Yeast *Saccharomyces cerevisiae*, *Appl. Environ. Microbiol.* 69 (2003) 2065–2072. <https://doi.org/10.1128/AEM.69.4.2065>.
- [90] M. Valkonen, M. Penttilä, M. Saloheimo, M. Valkonen, M. Penttilä, Effects of Inactivation and Constitutive Expression of the Unfolded- Protein Response Pathway on Protein Production in the Yeast *Saccharomyces cerevisiae*, *Appl. Environ. Microbiol.* 69 (2003) 2065–2072. <https://doi.org/10.1128/AEM.69.4.2065>.
- [91] J. Bao, M. Huang, D. Petranovic, J. Nielsen, Moderate expression of SEC16 increases protein secretion by *Saccharomyces cerevisiae*, *Appl. Environ. Microbiol.* 83 (2017). <https://doi.org/10.1128/AEM.03400-16>.
- [92] M. Huang, G. Wang, J. Qin, D. Petranovic, J. Nielsen, Engineering the protein secretory pathway of *Saccharomyces cerevisiae* enables improved protein production, *Proc. Natl. Acad. Sci. U. S. A.* 115 (2018) E11025–E11032. <https://doi.org/10.1073/pnas.1809921115>.
- [93] H. Tang, M. Song, Y. He, J. Wang, S. Wang, Y. Shen, J. Hou, Biotechnology for Biofuels Engineering vesicle trafficking improves the extracellular activity and surface display efficiency of cellulases in *Saccharomyces cerevisiae*, *Biotechnol. Biofuels.* (2017) 1–13. <https://doi.org/10.1186/s13068-017-0738-8>.
- [94] A. Idiris, K. Bi, H. Tohda, H. Kumagai, Y. Giga-hama, Construction of a protease-deficient strain set for the fission yeast *Schizosaccharomyces pombe* , useful for effective production of protease-sensitive heterologous proteins, *Yeast.* 23 (2006) 83–99. <https://doi.org/10.1002/yea.1342>.
- [95] I. Fitzgerald, B.S. Glick, Secretion of a foreign protein from budding yeasts is enhanced by cotranslational translocation and by suppression of vacuolar targeting, *Microb. Cell Fact.* 13 (2014) 1–12. <https://doi.org/10.1186/s12934-014-0125-0>.
- [96] M. Agaphonov, N. Romanova, S. Sokolov, A. Iline, T. Kalebina, G. Gellissen, M. Ter-avanessian, Defect of vacuolar protein sorting stimulates proteolytic processing of human urokinase-type plasminogen activator in the yeast *Hansenula polymorpha*,

- FEMS Yeast Res. 5 (2005) 1029–1035. <https://doi.org/10.1016/j.femsyr.2005.07.003>.
- [97] L. Xu, Y. Shen, J. Hou, B. Peng, H. Tang, X. Bao, Secretory pathway engineering enhances secretion of cellobiohydrolase I from *Trichoderma reesei* in *Saccharomyces cerevisiae*, J. Biosci. Bioeng. 117 (2013) 45–52. <https://doi.org/10.1016/j.jbiosc.2013.06.017>.
- [98] A. Idiris, H. Tohda, M. Sasaki, Enhanced protein secretion from multiprotease-deficient fission yeast by modification of its vacuolar protein sorting pathway, Appl. Microbiol. Biotechnol. 85 (2010) 667–677. <https://doi.org/10.1007/s00253-009-2151-0>.
- [99] K. Tomimoto, Y. Fujita, T. Iwaki, Y. Chiba, Y. Jigami, K. Nakayama, Y.N. Jima, H. Abe, K.N. Akayama, Y.N. Akajima, H.A. Be, Protease-Deficient *Saccharomyces cerevisiae* Strains for the Synthesis of Human-Compatible Glycoproteins, Biosci. Biotechnol. Biochem. 77 (2013) 2461–2466. <https://doi.org/10.1271/bbb.130588>.
- [100] L. Jønson, J.F. Rehfeld, A.H. Johnsen, Enhanced peptide secretion by gene disruption of CYM1, a novel protease in *Saccharomyces cerevisiae*, Eur. J. Biochem. 271 (2004) 4788–4797. <https://doi.org/10.1111/j.1432-1033.2004.04443.x>.
- [101] Y. Liu, W.J. Mo, T.F. Shi, M.Z. Wang, J.G. Zhou, Y. Yu, W.S. Yew, Mutational Mtc6p attenuates autophagy and improves secretory expression of heterologous proteins in *Kluyveromyces marxianus*, Microb. Cell Fact. (2018) 1–14. <https://doi.org/10.1186/s12934-018-0993-9>.
- [102] G.J. De Nobel, J.A. Barnett, Passage of Molecules Through Yeast Cell Walls: a Brief Essay-Review, Yeast. 7 (1991) 313–323. <https://doi.org/https://doi.org/10.1002/yea.320070402>.
- [103] J. Li, M. Zhang, C. Wan, R. Den Haan, F. Bai, X. Zhao, J.B.I.B. Ioeng, Improved cellulase production in recombinant *Saccharomyces cerevisiae* by disrupting the cell wall protein-encoding gene CWP2, J. Biosci. Bioeng. 129 (2019) 165–171. <https://doi.org/10.1016/j.jbiosc.2019.08.012>.
- [104] A.E. Wentz, E. V. Shusta, A novel high-throughput screen reveals yeast genes that increase secretion of heterologous proteins, Appl. Environ. Microbiol. 73 (2007) 1189–1198. <https://doi.org/10.1128/AEM.02427-06>.
- [105] P. Korpys-Woźniak, E. Celińska, Global transcriptome profiling reveals genes

- responding to overproduction of a small secretory, a high cysteine- and a high glycosylation-bearing protein in *Yarrowia lipolytica*, *Biotechnol. Reports.* 31 (2021) e00646. <https://doi.org/10.1016/j.btre.2021.e00646>.
- [106] I. Vigentini, Heterologous protein production in *Zygosaccharomyces bailii*: physiological effects and fermentative strategies, 5 (2005) 647–652. <https://doi.org/10.1016/j.femsyr.2004.11.006>.
- [107] J.H.D. Van Zyl, R. Den Haan, W.H. Van Zyl, Overexpression of native *Saccharomyces cerevisiae* ER-to-Golgi SNARE genes increased heterologous cellulase secretion, *Appl. Microbiol. Biotechnol.* 100 (2016) 505–518. <https://doi.org/10.1007/s00253-015-7022-2>.
- [108] M.C.. Hensing, R.J. Rouwenhorst, J.J. Heijnen, J.P. Van Dijken, J.T. Pronk, Physiological and technological aspects of large-scale heterologous-protein production with yeasts, *Antonie Van Leeuwenhoek.* 67 (1995) 261–279. <https://doi.org/https://doi.org/10.1007/BF00873690>.
- [109] M. Kubiak, W. Białas, E. Celińska, Thermal treatment improves a process of crude glycerol valorization for the production of a heterologous enzyme by *Yarrowia lipolytica*, *Biotechnol. Reports.* 31 (2021) e00648. <https://doi.org/10.1016/j.btre.2021.e00648>.
- [110] M. Arvas, T. Pakula, B. Smit, J. Rautio, H. Koivistoinen, P. Jouhten, E. Lindfors, M. Wiebe, M. Penttilä, M. Saloheimo, Correlation of gene expression and protein production rate - a system wide study, 12 (2011) 1–25. <https://doi.org/https://doi.org/10.1186/1471-2164-12-616>.
- [111] Z. Liu, J. Hou, J.L. Martínez, D. Petranovic, J. Nielsen, Correlation of cell growth and heterologous protein production by *Saccharomyces cerevisiae*, *Appl. Microbiol. Biotechnol.* 97 (2013) 8955–8962. <https://doi.org/10.1007/s00253-013-4715-2>.
- [112] K. Peebo, P. Neubauer, Application of Continuous Culture Methods to Recombinant Protein Production in Microorganisms, *Microorganisms.* 6 (2018) 56. <https://doi.org/10.3390/microorganisms6030056>.
- [113] A. Novick, L. Szilard, Description of the Chemostat, *Science* (80-.). 112 (1950) 715–716. <https://doi.org/DOI: 10.1126/science.112.2920.715>.
- [114] J. Monod, Technique, theory and applications, *Ann. Inst. Pasteur.* 79 (1950) 390–410.

- [115] P.A. Hoskisson, G. Hobbs, Continuous culture – making a comeback?, *Microbiology*. 151 (2005) 3153–3159. <https://doi.org/10.1099/mic.0.27924-0>.
- [116] O. Daniela, Delneri, Francesco L, Brancia, Stephen G, Towards a truly integrative biology through the functional genomics of yeast, *Curr. Opin. Biotechnol.* 12 (2001) 87–91.
- [117] H. Pedersen, M. Beyer, Glucoamylase production in batch, chemostat and fed-batch cultivations by an industrial strain of *Aspergillus niger*, *Appl. Microbiol. Biotechnol.* 53 (2000) 272–277.
- [118] E. Rodríguez-carmona, O. Cano-garrido, M. Dragosits, M. Maurer, A. Mader, R. Kunert, D. Mattanovich, A. Villaverde, F. Vázquez, Recombinant Fab expression and secretion in *Escherichia coli* continuous culture at medium cell densities : Influence of temperature, *Process Biochem.* 47 (2012) 446–452. <https://doi.org/10.1016/j.procbio.2011.11.024>.
- [119] V. Looser, B. Bruhlmann, F. Bumbak, C. Stenger, M. Costa, A. Camattari, D. Fotiadis, K. Kovar, Cultivation strategies to enhance productivity of *Pichia pastoris* : A review, *Biotechnol. Adv.* 33 (2015) 1177–1193. <https://doi.org/10.1016/j.biotechadv.2015.05.008>.
- [120] A. Rywińska, P. Juszczak, M. Wojtatowicz, W. Rymowicz, Chemostat study of citric acid production from glycerol by *Yarrowia lipolytica*, *J. Biotechnol.* 152 (2011) 54–57. <https://doi.org/10.1016/j.jbiotec.2011.01.007>.
- [121] A. Ochoa-estopier, S.E. Guillouet, D-stat culture for studying the metabolic shifts from oxidative metabolism to lipid accumulation and citric acid production in *Yarrowia lipolytica*, *J. Biotechnol.* 170 (2014) 35–41. <https://doi.org/10.1016/j.jbiotec.2013.11.008>.
- [122] M. Rakicka, A.M. Mirończuk, L. Tomaszewska-Hetman, A. Rywińska, W. Rymowicz, An Effective Method of Continuous Production of Erythritol from Glycerol by *Yarrowia lipolytica* MK1, *Food Technol. Biotechnol.* 55 (2017) 125–130. <https://doi.org/https://doi.org/10.17113/ftb.55.01.17.4812>.
- [123] E. Celińska, W. Białas, M. Borkowska, W. Grajek, Cloning, expression, and purification of insect (*Sitophilus oryzae*) alpha-amylase, able to digest granular starch, in *Yarrowia lipolytica* host, *Appl. Microbiol. Biotechnol.* 99 (2015) 2727–2739. <https://doi.org/10.1007/s00253-014-6314-2>.

- [124] T.S. Thorsen, A.H. Johnsen, K. Josefsen, B. Jensen, Identification and characterization of glucoamylase from the fungus *Thermomyces lanuginosus*, *Biochim. Biophys. Acta - Proteins Proteomics*. 1764 (2006) 671–676. <https://doi.org/10.1016/j.bbapap.2006.01.009>.
- [125] C. Madzak, B. Tréton, Strong Hybrid Promoters and Integrative Expression / Secretion Vectors for Quasi-Constitutive Expression of Heterologous Proteins in the Yeast *Yarrowia lipolytica* Further Reading, *J. Mol. Microbiol. Biotechnol.* 2 (2000) 207–216.
- [126] G. Pignede, H. Wang, F. Fudalej, C. Gaillardin, M. Seman, J. Nicaud, Characterization of an Extracellular Lipase Encoded by LIP2 in *Yarrowia lipolytica*, *J. Bacteriol.* 182 (2000) 2802–2810. <https://doi.org/10.1128/AEM.66.8.3283-3289.2000>.
- [127] J.M. Nicaud, C. Madzak, P. Van Den Broek, C. Gysler, P. Duboc, P. Niederberger, C. Gaillardin, Protein expression and secretion in the yeast *Yarrowia lipolytica*, *FEMS Yeast Res.* 2 (2002) 371–379. [https://doi.org/10.1016/S1567-1356\(02\)00082-X](https://doi.org/10.1016/S1567-1356(02)00082-X).
- [128] J. Sambrook, D. Russell, *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory Press, New York, 2001.
- [129] G. Barth, C. Gaillardin, *Yarrowia lipolytica*, in: K. Wolf (Ed.), *Nonconv. Yeasts Biotechnol. A Handb.*, Springer Berlin Heidelberg, 1996: pp. 313–388. <https://doi.org/10.1007/978-3-642-79856-6>.
- [130] M. Borkowska, W. Białas, M. Kubiak, E. Celinska, Rapid micro-assays for amylolytic activities determination: customization and validation of the tests, *Appl. Microbiol. Biotechnol.* 103 (2019) 2367–2379. <https://doi.org/10.1007/s00253-018-09610-0>.
- [131] R.M. Goncalves Cristiana, Rodriguez-jasso, N. Gomes, A. Teixeira, I. Belo, Adaptation of dinitrosalicylic acid method to microtiter plates, *Anal Methods*. (2010) 2046–2048. <https://doi.org/10.1039/c0ay00525h>.
- [132] K.J. Livak, T.D. Schmittgen, Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method, *Methods*. 25 (2001) 402–408. <https://doi.org/10.1006/meth.2001.1262>.
- [133] M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads, *EMBnet. J.* 17 (2011) 10–12. <https://doi.org/10.14806/ej.17.1.200>.
- [134] D. Kim, J.M. Paggi, C. Park, C. Bennett, S.L. Salzberg, Graph-based genome

- alignment and genotyping with HISAT2 and HISAT-genotype, *Nat. Biotechnol.* 37 (2019). <https://doi.org/10.1038/s41587-019-0201-4>.
- [135] S. Anders, P.T. Pyl, W. Huber, Genome analysis HTSeq — a Python framework to work with high-throughput sequencing data, *Bioinformatics.* 31 (2015) 166–169. <https://doi.org/10.1093/bioinformatics/btu638>.
- [136] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, *Genome Biol.* 15 (2014) 1–21. <https://doi.org/10.1186/s13059-014-0550-8>.
- [137] H. Heberle, G.V. Meirelles, F.R. Silva, G.P. Telles, R. Minghim, InteractiVenn : a web-based tool for the analysis of sets through Venn diagrams, *BMC Bioinformatics.* 16 (2015) 1–7. <https://doi.org/10.1186/s12859-015-0611-3>.
- [138] H. Mi, A. Muruganujan, X. Huang, X. Guo, P.D. Thomas, Protocol Update for large-scale genome and gene function analysis with the PANTHER, *Nat. Protoc.* 14 (2016) 703–721. <https://doi.org/10.1038/s41596-019-0128-8>.
- [139] L.D. Goldstein, Y. Cao, G. Pau, M. Lawrence, T.D. Wu, S. Seshagiri, R. Gentleman, Prediction and Quantification of Splice Events from RNA-Seq Data., *PLoS One.* 11 (2016) e0156132. <https://doi.org/10.1371/journal.pone.0156132>.
- [140] J.T. Robinson, H. Thorvaldsdóttir, W. Winckler, M. Guttman, E.S. Lander, G. Getz, J.P. Mesirov, Integrative genomics viewer., *Nat. Biotechnol.* 29 (2011) 24–26. <https://doi.org/10.1038/nbt.1754>.
- [141] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The Sequence Alignment/Map format and SAMtools., *Bioinformatics.* 25 (2009) 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
- [142] H. T., *BioEdit_a_user_friendly_biological_seque.pdf*, (1999) 95–98.
- [143] F. Corpet, Multiple sequence alignment with hierarchical clustering, *Nucleic Acids Res.* 16 (1988) 10881–10890. <https://doi.org/https://doi.org/10.1093/nar/16.22.10881>.
- [144] V.B. Rao, N.V.S. Sastri, P.V.S. Rao, Purification and characterization of a thermostable glucoamylase from the thermophilic fungus *Thermomyces lanuginosus*, *Biochem. J.* 193 (1981) 379–387. <https://doi.org/https://doi.org/10.1042/bj1930379>.
- [145] J.E. Baker, S.M. Woo, Purification, partial characterization, and postembryonic levels of amylases from *Sitophilus oryzae* and *Sitophilus granarius*, *Arch. Insect Biochem.*

- Physiol. (1985). <https://doi.org/10.1002/arch.940020409>.
- [146] K.P. Wall, M.K. Knowles, R. Dillon, M.K. Knowles, Fluorescence Quantum Yield Measurements of Fluorescent Proteins : A Laboratory Experiment for a Biochemistry or Molecular Biophysics Laboratory Course w, *Biochem. Mol. Biol. Educ.* 43 (2015) 52–59. <https://doi.org/10.1002/bmb.20837>.
- [147] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, A. Bairoch, Protein Identification and Analysis Tools on the ExPASy Server, Humana Press. (2005) 571–607. <https://doi.org/https://doi.org/10.1385/1-59259-890-0:571>.
- [148] E. Celinska, M. Borkowska, W. Białas, Enhanced production of insect raw-starch-digesting alpha-amylase accompanied by high erythritol synthesis in recombinant *Yarrowia lipolytica* fed-batch cultures at high-cell-densities, *Process Biochem.* 52 (2017) 78–85. <https://doi.org/10.1016/j.procbio.2016.10.022>.
- [149] A. Skoneczna, A. Micia, M. Skoneczny, *Saccharomyces cerevisiae* Hsp31p , a stress response protein conferring protection against reactive oxygen species, *Free Radic. Biol. Med.* 42 (2007) 1409–1420. <https://doi.org/10.1016/j.freeradbiomed.2007.01.042>.
- [150] M. Pech, J. Kijek, H. Yamamoto, F. Naeve, A. Tovchigrechko, K. Yamamoto, W. Szaflarski, N. Takeuchi, T. Stellberger, M.E. Diefenbacher, K.H. Nierhaus, P. Uetz, RsfA (YbeB) Proteins Are Conserved Ribosomal Silencing Factors, *Plos Genet.* 8 (2012) e1002815. <https://doi.org/10.1371/journal.pgen.1002815>.
- [151] E. Celinska, M. Borkowska, W. Białas, M. Kubiak, P. Korpys, M. Archacka, R. Ledesma-amaro, J.Nicaud, Genetic engineering of Ehrlich pathway modulates production of higher alcohols in engineered *Yarrowia lipolytica*, *FEMS Yeast Res.* 19 (2019) foy122. <https://doi.org/10.1093/femsyr/foy122>.
- [152] E. Celinska, M. Olkowicz, W. Grajek, L-Phenylalanine catabolism and 2-phenylethanol synthesis in *Yarrowia lipolytica* — mapping molecular identities through whole-proteome quantitative mass spectrometry analysis, *FEMS Yeast Res.* 15 (2015) fov041. <https://doi.org/10.1093/femsyr/fov041>.
- [153] A. Graf, B. Gasser, M. Dragosits, M. Sauer, G.G. Leparc, T. Tüchler, D.P. Kreil, D. Mattanovich, Novel insights into the unfolded protein response using *Pichia pastoris* specific DNA microarrays, *BMC Genomics.* 9 (2008) 1–13. <https://doi.org/10.1186/1471-2164-9-390>.

- [154] J.S. Cox, P. Walter, A Novel Mechanism for Regulating Activity of a Transcription Factor That Controls the Unfolded Protein Response, *Cell*. 87 (1996) 391–404. [https://doi.org/https://doi.org/10.1016/S0092-8674\(00\)81360-4](https://doi.org/https://doi.org/10.1016/S0092-8674(00)81360-4).
- [155] K. Mori, T. Kawahara, H. Yoshida, H. Yanagi, T. Yura, Signalling from endoplasmic reticulum to nucleus : transcription factor with a basic-leucine zipper motif is required for the unfolded protein-response pathway, *Genes to Cells*. 1 (1996) 803–817.
- [156] K.J. Travers, C.K. Patil, L. Wodicka, D.J. Lockhart, J.S. Weissman, P. Walter, S. Francisco, Functional and Genomic Analyses Reveal an Essential Coordination between the Unfolded Protein Response and ER-Associated Degradation, *Cell*. 101 (2000) 249–258. [https://doi.org/https://doi.org/10.1016/S0092-8674\(00\)80835-1](https://doi.org/https://doi.org/10.1016/S0092-8674(00)80835-1).
- [157] P. Korpys-Woźniak, P. Kubiak, W. Białas, E. Celińska, Impact of overproduced heterologous protein characteristics on physiological response in *Yarrowia lipolytica* steady-state-maintained continuous cultures, *Appl. Microbiol. Biotechnol.* 104 (2020) 9785–9800. <https://doi.org/10.1007/s00253-020-10937-w>.
- [158] E. Celinska, M. Borkowska, W. Białas, Evaluation of heterologous α -amylase production in two expression platforms dedicated for *Yarrowia lipolytica* : commercial Po1g – pYLSC (php4d) and custom-made A18 – pYLTEF (pTEF), *Yeast*. 33 (2016) 165–181. <https://doi.org/10.1002/yea>.
- [159] J.I. Castrillo, L.A. Zeef, D.C. Hoyle, N. Zhang, A. Hayes, D.C.J. Gardner, M.J. Cornell, J. Petty, L. Hakes, L. Wardleworth, B. Rash, M. Brown, W.B. Dunn, D. Broadhurst, K.O. Donoghue, S.S. Hester, T.P.J. Dunkley, S.R. Hart, N. Swainston, P. Li, S.J. Gaskell, N.W. Paton, K.S. Lilley, D.B. Kell, S.G. Oliver, BioMed Central Research article Growth control of the eukaryote cell : a systems biology study in yeast, *J. Biol.* 6 (2007) 1–25. <https://doi.org/https://doi.org/10.1186/jbiol54>.
- [160] M. Buchetics, M. Dragosits, M. Maurer, C. Rebnegger, D. Porro, M. Sauer, B. Gasser, D. Mattanovich, Reverse Engineering of Protein Secretion by Uncoupling of Cell Cycle Phases From Growth, *Biotechnol. Bioeng.* 108 (2011) 2403–2412. <https://doi.org/10.1002/bit.23198>.
- [161] C. Madzak, S. Blanchin-Roland, R.R.C. Otero, C. Gaillardin, Functional analysis of upstream regulating regions from the *Yarrowia lipolytica* XPR2 promoter, *Microbiology*. 145 (1999) 75–87. <https://doi.org/https://doi.org/10.1099/13500872-145-1-75>.

- [162] A. Read, M. Schröder, The Unfolded Protein Response : An Overview, *Nat. Rev. Mol. Cell Biol.* 10 (2021) 384. <https://doi.org/10.3390/biology10050384>.
- [163] R. Aw, G.R. Barton, D.J. Leak, Insights into the prevalence and underlying causes of clonal variation through transcriptomic analysis in *Pichia pastoris*, *Appl. Microbiol. Biotechnol.* 101 (2017) 5045–5058. <https://doi.org/10.1007/s00253-017-8317-2>.
- [164] J.L. Martínez, E. Meza, D. Petranovic, J. Nielsen, The impact of respiration and oxidative stress response on recombinant α -amylase production by *Saccharomyces cerevisiae*, *Metab. Eng. Commun.* 3 (2016) 205–210. <https://doi.org/10.1016/j.meteno.2016.06.003>.
- [165] D.G.J.A. Mora-lorca, B.S.F.J. Naranjo-galindo, F.M. Cristina, P.J. Goikolea, C.D. Link, C. Neri, M. Dolores, R.P. Vázquez-manrique, E.F. Veit, G. Roser, P. Elisa, C. Peter, A. Juan, Loss of glutathione redox homeostasis impairs proteostasis by inhibiting autophagy-dependent protein degradation, *Cell Death Differ.* (2019) 1545–1565. <https://doi.org/10.1038/s41418-018-0270-9>.
- [166] A.J. Ponsero, A. Igarria, M.A. Darch, B.D. Autre, M.B. Toledano, S. Miled, C.E. Outten, J.R. Winther, G. Palais, Endoplasmic Reticulum Transport of Glutathione by Sec61 Is Regulated by Ero1 and Bip, *Mol. Cell.* 67 (2017) 962–973. <https://doi.org/10.1016/j.molcel.2017.08.012>.
- [167] M. Guerfal, S. Ryckaert, P.P. Jacobs, P. Ameloot, K. Van Craenenbroeck, R. Derycke, N. Callewaert, The HAC1 gene from *Pichia pastoris*: Characterization and effect of its overexpression on the production of secreted, surface displayed and membrane proteins, *Microb. Cell Fact.* 9 (2010) 1–12. <https://doi.org/10.1186/1475-2859-9-49>.
- [168] J.R. Shuster, Gene expression in yeast: protein secretion, *Curr. Opin. Biotechnol.* 2 (1991) 685–690. [https://doi.org/10.1016/0958-1669\(91\)90035-4](https://doi.org/10.1016/0958-1669(91)90035-4).
- [169] M. Huang, Y. Bai, S.L. Sjostrom, B.M. Hallström, Z. Liu, D. Petranovic, M. Uhlén, H.N. Joensson, H. Andersson-Svahn, J. Nielsen, Microfluidic screening and whole-genome sequencing identifies mutations associated with improved protein secretion by yeast, *Proc. Natl. Acad. Sci. U. S. A.* 112 (2015) E4689–E4696. <https://doi.org/10.1073/pnas.1506460112>.
- [170] L. Ruohonen, J. Toikkanen, V. Tieaho, M. Outola, H. Soderlund, S. Keranen, Enhancement of Protein Secretion in *Saccharomyces cerevisiae* by Overproduction of Sso Protein, a Late-acting Component of the Secretory Machinery, *Yeast.* 13 (1997)

337–351. [https://doi.org/10.1002/\(sici\)1097-0061\(19970330\)13:4<337::aid-yea98>3.3.co;2-b](https://doi.org/10.1002/(sici)1097-0061(19970330)13:4<337::aid-yea98>3.3.co;2-b).

- [171] J.H. Toikkanen, L. Sundqvist, S. Keranen, *Kluyveromyces lactis* SSO1 and SEB1 genes are functional in *Saccharomyces cerevisiae* and enhance production of secreted proteins when overexpressed, *Yeast*. 2 (2004) 1045–1055. <https://doi.org/10.1002/yea.1151>.
- [172] B. Gasser, M. Sauer, M. Maurer, G. Stadlmayr, D. Mattanovich, Transcriptomics-based identification of novel factors enhancing heterologous protein secretion in yeasts, *Appl. Environ. Microbiol.* 73 (2007) 6499–6507. <https://doi.org/10.1128/AEM.01196-07>.
- [173] C. Nombela, C. Gil, W.L. Chaffin, Non-conventional protein secretion in yeast, *Trends Microbiol.* 14 (2006) 15–21. <https://doi.org/10.1016/j.tim.2005.11.009>.
- [174] A.. Cleves, D.N.. Cooper, S.. Barondes, R.. Kelly, A New Pathway for Protein Export in *Saccharomyces cerevisiae*, *J. Cell Biol.* 133 (1996) 1017–1026. <https://doi.org/https://doi.org/10.1083/jcb.133.5.1017>.
- [175] L. Desmyter, J. Verstraelen, S. Dewaele, C. Libert, R. Contreras, C. Chen, Nonclassical export pathway : overexpression of NCE102 reduces protein and DNA damage and prolongs lifespan in an SGS1 deficient *Saccharomyces cerevisiae*, *Biogerontology*. 8 (2007) 527–535. <https://doi.org/10.1007/s10522-007-9095-5>.
- [176] B. Mai, L. Breeden, Xbp1 , a Stress-Induced Transcriptional Repressor of the *Saccharomyces cerevisiae* Swi4 / Mbp1 Family, *Mol. Cell. Biol.* 17 (1997) 6491–6501. <https://doi.org/https://doi.org/10.1128/MCB.17.11.6491>.
- [177] B. Mai, L. Breeden, CLN1 and Its Repression by Xbp1 Are Important for Efficient Sporulation in Budding Yeast, *Mol. Cell. Biol.* 20 (2000) 478–487. <https://doi.org/https://doi.org/10.1128/MCB.20.2.478-487.2000>.
- [178] T. Iwawaki, M. Tokuda, Biochemical and Biophysical Research Communications Function of yeast and amphioxus tRNA ligase in IRE1alpha-dependent XBP1 mRNA splicing, *Biochem. Biophys. Res. Commun.* 413 (2011) 527–531. <https://doi.org/10.1016/j.bbrc.2011.08.129>.
- [179] S. Liang, B. Wang, L. Pan, Y. Ye, M. He, S. Han, S. Zheng, X. Wang, Y. Lin, Comprehensive structural annotation of *Pichia pastoris* transcriptome and the response to various carbon sources using deep paired-end RNA sequencing, *BMC Genomics*. 13

- (2012) 738. <https://doi.org/https://doi.org/10.1186/1471-2164-13-738>.
- [180] L. Aravind, V. Anantharaman, S. Balaji, M.M. Babu, L.M. Iyer, The many faces of the helix-turn-helix domain: Transcription regulation and beyond, *Microbiol. Rev.* 29 (2005) 231–262. <https://doi.org/10.1016/j.femsre.2004.12.008>.
- [181] L. Gaigalat, J. Schlüter, M. Hartmann, S. Mormann, A. Tauch, A. Pühler, J. Kalinowski, The DeoR-type transcriptional regulator SugR acts as a repressor for genes encoding the phosphoenolpyruvate: sugar phosphotransferase system (PTS) in *Corynebacterium glutamicum*, *BMC Mol. Biol.* 8 (2007) 1–20. <https://doi.org/10.1186/1471-2199-8-104>.
- [182] A.R. Hesketh, J.I. Castrillo, T. Sawyer, D.B. Archer, S.G. Oliver, Investigating the physiological response of *Pichia (Komagataella) pastoris* GS115 to the heterologous expression of misfolded proteins using chemostat cultures, *Appl. Microbiol. Biotechnol.* 97 (2013) 9747–9762. <https://doi.org/10.1007/s00253-013-5186-1>.
- [183] J. Klabunde, G. Kunze, G. Gellissen, C.P. Hollenberg, Integration of heterologous genes in several yeast species using vectors containing a *Hansenula polymorpha*-derived rDNA-targeting element, *FEMS Yeast Res.* 4 (2003) 185–193. [https://doi.org/10.1016/S1567-1356\(03\)00148-X](https://doi.org/10.1016/S1567-1356(03)00148-X).
- [184] J. Liu, Q. Han, Q. Cheng, Y. Chen, R. Wang, X. Li, Y. Liu, D. Yan, Efficient Expression of Human Lysozyme Through the Increased Gene Dosage and Co-expression of Transcription Factor Hac1p in *Pichia pastoris*, *Curr. Microbiol.* 77 (2020) 846–854. <https://doi.org/10.1007/s00284-019-01872-9>.
- [185] K.J. Kauffman, E.M. Pridgen, F.J. Doyle, P.S. Dhurjati, A.S. Robinson, Decreased protein expression and intermittent recoveries in BiP levels result from cellular stress during heterologous protein expression in *Saccharomyces cerevisiae*, *Biotechnol. Prog.* 18 (2002) 942–950. <https://doi.org/10.1021/bp025518g>.
- [186] K. Löchli, E. Torbica, M. Haile, W. Deborah, B. Aatika, A. Daniela, Crosstalk between endoplasmic reticulum and cytosolic unfolded protein response in tomato, *Cell Stress Chaperones.* (2022) 1–18. <https://doi.org/10.1007/s12192-022-01316-7>.
- [187] H. Yoshida, T. Matsui, A. Yamamoto, T. Okada, K. Mori, XBP1 mRNA Is Induced by ATF6 and Spliced by IRE1 in Response to ER Stress to Produce a Highly Active Transcription Factor, *Cell.* 107 (2001) 881–891. [https://doi.org/https://doi.org/10.1016/S0092-8674\(01\)00611-0](https://doi.org/https://doi.org/10.1016/S0092-8674(01)00611-0).

- [188] K.B. Hooks, S. Griffiths-jones, Conserved RNA structures in the non-canonical Hac1/Xbp1 intron, *RNA Biol.* 8 (2011) 552–556.
<https://doi.org/10.4161/rna.8.4.15396>.

Wykaz Tabel

Tabela 1. Właściwości r-Prots wykorzystanych w tej pracy.

Tabela 2. Szczepy bakterii wykorzystane w tej pracy.

Tabela 3. Szczepy drożdży wykorzystane w tej pracy.

Tabela 4. Wzrost (OD_{600 nm}), bezwzględna wartość fluorescencji zewnątrzkomórkowej (exYFP [FU]) i specyficzna fluorescencja (sp_exYFP [FU OD_{600nm}⁻¹]) kotransformantów *Y. lipolytica* w hodowlach okresowych w temp. 25°C lub 30°C w stacjonarnej fazie wzrostu (72 h). Wartości zmiany krotności (FC; *Fold change*) obliczono w odniesieniu do szczepu kontrolnego (Control_scYFP) z nadekspresją wyłącznie *SCYFP*. Liczby podane są we wskazanych jednostkach ± SD z duplikatu biologicznego, każdy mierzony w duplikacie technicznym.

Tabela 5. Ilości YFP syntetyzowane w warunkach stanu ustalonego przez rekombinowane szczepy *Y. lipolytica*

Wykaz Rycin

Rycina 1. Szlak wydzielniczy w *Yarrowia lipolytica*. Szczegółowe informacje o roli poszczególnych elementów aparatu wydzielniczego przedstawionych na rycinie znajdują się w tekście. Wszystkie białka przedstawione na rysunku zostały zidentyfikowane w *Y. lipolytica* za pomocą podejścia proteomicznego lub genomiki porównawczej [26–28]. Kolorem ciemnoczerwonym zaznaczono białka odróżniające *Y. lipolytica* od drożdży modelowych (przekreślenia – brak genu kodującego daną aktywność; duża czcionka – unikalne występowanie genu kodującego daną aktywność w *Y. lipolytica* lub jego dominująca rola, np. Sls1 NEF). Nieznana funkcja Lhs1 została zaznaczona cudzysłowem. Dominującą rolę szlaku translokacji kotranslacyjnej wskazuje grubsza czerwona strzałka w porównaniu z cienką szarą translokacją oznaczającą translokację potranslacyjną. Powstający polipeptyd jest zaznaczony na pomarańczowo. Trójkąty w linii – N-glikany. Białka opiekuńcze, czynniki wiążące, SNARE, białka SM zostały wymienione jako zidentyfikowane przez analizę proteomiczną w [27]. Rycina z Celińska i Nicaud [5].

Rycina 2. Schematyczne przedstawienie strategii klonowania kasety do nadekspresji zastosowanej w tej pracy. A. Komercyjny system klonowania YLEX. B. Klonowanie

modularne Golden Gate. Skróty: M – Marker auksotroficzny (LEU2, URA3), P – promotor drożdżowy (php4d/p4UASpTEF), SP - peptyd sygnałowy wydajnie kierujący r-Prot do przestrzeni zewnątrzkomórkowej (preXPR2/SP1), G – gen kodujący r-Prot (SoA, TIG, YFP), T– terminator drożdżowy (tXPR2/tLip2), Ins: zetaUP/zetaDOWN – sekwencje kierujące zeta umożliwiające losową insercję kasety do genomu gospodarza (rekombinacja niehomologiczna) lub pBR-platforma dokująca (rekombinacja homologiczna), A-J – litery oznaczające poszczególne 4 nt sekwencje flankujące, AmpR – oporność na ampicylinę, Ori – bakteryjne miejsce replikacji, *NotI* – miejsce rozpoznawane przez endonukleazę restrykcyjną *NotI* (miejsce trawienia przed transformacją drożdży).

Rycina 3. Schematyczne przedstawienie strategii klonowania do wektora JMP62 zastosowanej w tej pracy. Kolorami oznaczono elementy wektorów: region zeta (pomarańczowy), *LoxP* i *LoxR* (ciemnoniebieski), konstytutywny promotor pTEF (zielony), heterologiczne geny: scYFP (A. żółty) lub SH (B. niebieski) wyposażone we fragmenty *BamHI-AvrII*, markery auksotroficzne URA3ex (A. czerwony) lub LEU2ex (B. jasnoniebieski). Bakteryjne miejsca replikacji (różowe) i gen oporności na kanamycynę KanR (fioletowe), zawarte w części „bakteryjnej” wektorów, odrzucone przed transformacją komórek drożdży przez trawienie restrykcyjne *NotI*.

Rycina 4. Poziom ekspresji genów kodujących rs-Protów analizowany metodą qPCR. Próbki do analizy ekspresji genów pobrano po osiągnięciu stanu ustalonego przez komórki *Y. lipolytica* w ciągłych hodowlach w warunkach chemostatu. Oś X: geny reporterowe (*INYFP*, *SCYFP*, *SOA*, *TLG*); oś Y: LOG₁₀RQ (zlogarytmowana wartość względnego poziomu ekspresji). Względny poziom ekspresji (RQ) określany metodą $\Delta\Delta Ct$.

Rycina 5. Całkowita liczba genów różnicowo ekspresjonowanych (DEGs) określona dla każdego szczepu rekombinowanego w porównaniu ze szczepem kontrolnym. Oś X: liczba DEG: czerwone słupki – geny o zwiększonej ekspresji vs. kontroli; niebieskie słupki – geny o obniżonej ekspresji vs. kontroli. Oś Y: rekombinowane szczepy wytwarzające różne białka reporterowe.

Rycina 6. Graficzne przedstawienie najbardziej deregulowanych DEGs w szczepach nadprodukcujących r(s)-Protów. Indukcję/represję DEGs przedstawiono za pomocą kolorów, zgodnie z legendą, w oparciu o wartość ekspresji wyrażoną jako log₂-krotność zmiany (FC). Wszystkie zaprezentowane DEGs spełniały warunek $> 2.0 \log_2_Estimated_FC$, w co najmniej jednym ze szczepów nadprodukcującym. Skróty: TF- Czynniki transkrypcyjne.

Wykres został przygotowany przy użyciu narzędzia internetowego Morpheus (<https://software.broadinstitute.org/morpheus/>).

Rycina 7. Główne istotnie wzbogacone procesy biologiczne dla zestawu DEGs deregulowanych w odpowiedzi na nadprodukcję r(s)-Prots. Pokazano tylko procesy wzbogacone na poziomie istotności $p < 0.05$. Słupki reprezentują liczbę DEGs przypisanych do tych procesów (pełne – indukcja DEGs; kreski – represja DEGs) a cyfry oznaczają EF.

Rycina 8. Indukowane i represjonowane wspólne lub unikalne geny/zestawy genów (DEGs) dla poszczególnych szczepów *Y. lipolytica* w odpowiedzi na nadprodukcję r(s) – Prots. Wskazano liczbę DEGs dla poszczególnych grup. a. *inYFP* vs. *scYFP*; b. indukowane DEGs w *SoA*, *TlG*, *scYFP* i *inYFP*; c. represjonowane DEGs w *SoA*, *TlG*, *scYFP* i *inYFP*. Analizę diagramów Venna przeprowadzono na listach DEGs z wartościami ekspresji ograniczonej do 2-krotnej zmiany.

Rycina 9. Graficzne przedstawienie deregulowanych DEGs zaangażowanych w wakuolarny procesy kataboliczne oraz wakuolarny sortowanie białek w szczepach nadprodukcujących r(s)-Prots. Indukcję/represję DEGs przedstawiono za pomocą kolorów, zgodnie z legendą, w oparciu o wartość ekspresji wyrażoną jako krotność zmiany (FC) ekspresji genu w próbie badanej względem kontroli. Wykres został przygotowany przy użyciu narzędzia internetowego Morpheus (<https://software.broadinstitute.org/morpheus/>).

Rycina 10. Poziom ekspresji wybranych genów zaangażowanych w fałdowanie, dojrzewanie i sekrecję r(s)-Prots. Znormalizowane wartości transkryptów zastosowano do ilościowego określenia poziomu transkrypcji. a. Czynniki transkrypcyjne, b. Wakuolarny białka sortujące, c. Białka o aktywności BIP, tworzenie wiązań disiarczkowych, O-glikozylacja, synteza oligosacharydów, d. Fałdowanie białek, e. Elementy translokony, f. Wakuolarny procesy kataboliczne g. Transport pęcherzykowy, h. Biogeneza ściany komórkowej, i. Sekrecja. Oś X: nazwy genów *Yali_*, oś Y: znormalizowana liczba transkryptów. *- wartości pomnożone przez 10, ** - wartości podzielone przez 10.

Rycina 11. Specyficzna fluorescencja *sp_exYFP* [FU OD600^{-1}] kotransformantów oraz kontroli *Y. lipolytica* w hodowlach okresowych w temp. 30°C (A.) lub 25°C (B.) oznaczona w stacjonarnej fazie wzrostu (72 h). Oś X: Szczep *Y. lipolytica*, oś Y: specyficzna fluorescencja (*sp_exYFP*) wyrażona w [FUxOD600^{-1}]. Litery oznaczają grupy homogeniczne obliczone testem Tukeya HSD.

Rycina 12. Graficzne porównanie rekombinowanych szczepów *Y. lipolytica* z ko-nadekspresją SCYFP i SH oraz szczepu kontrolnego (+scYFP/-SH) zaprezentowane jako krotkość zmiany (FC) pomiaru specyficznej sekrecji sp_exYFP pomiędzy szczepami w temperaturze 30 °C (A.) oraz w 25 °C (B.) w 72h hodowli. Czerwone kwadraty oznaczają przewagę licznika w przyjętej mierze (1.01–2.00 FC), a niebieskie kwadraty przewagę mianownika w przyjętej mierze (0 - 0.99 FC). Istotność statystyczną porównań określono testem Tukeya HSD. *** $p \leq 0.001$ ** $p \leq 0.01$ * $p \leq 0.05$. Wykres został przygotowany przy użyciu narzędzia internetowego Morpheus (<https://software.broadinstitute.org/morpheus/>).

Rycina 13. Graficzne porównanie wartości specyficznej sekrecji sp_exYFP [FU x OD600⁻¹] uzyskane w 72 h hodowli rekombinowanych szczepów *Y. lipolytica* (+scYFP/+SH) oraz szczepu kontrolnego (+scYFP/-SH) w temperaturze 25 °C (niebieskie boxy) oraz w 30 °C (czerwone boxy). Oś X: temperatura hodowli: 25 – 25 °C, 30 – 30 °C. Oś Y: specyficzna sekrecja YFP (sp_exYFP) wyrażona w [FU x OD600⁻¹]. Liczby oznaczają wartość p testu t-Studenta.

Rycina 14. Poziomu ekspresji genu YFP w szczepach *Y. lipolytica* z ko-nadekspresją SCYFP i SH hodowanych w temperaturze 30 °C.(A.) lub 25 °C (B.). Oś X: Nazwa szczepu; Oś Y: RQ- poziom ekspresji genu YFP obliczony zgodnie z modelem ddCt. Wartość równa 1.0 wskazuje na brak zmian w poziomie ekspresji YFP pomiędzy szczepem kontrolnym a badanym szczepem z nadekspresją SH. Litery oznaczają grupy homogeniczne obliczone testem Tukey'a HSD.

Rycina 15. Graficzne porównanie poziomu ekspresji genu YFP w szczepach *Y. lipolytica* ko-nadekspresją SCYFP i SH hodowanych w temperaturze 30 °C (czerwone) oraz w 25 °C (niebieskie). Oś X: temperatura hodowli: 25 – 25 °C, 30 – 30 °C. Oś Y: RQ- poziom ekspresji genu YFP obliczony zgodnie z modelem ddCt. Wartość równa 1.0 wskazuje na brak zmian w poziomie ekspresji YFP pomiędzy szczepem kontrolnym a badanym szczepem z nadekspresją SH.

Rycina 16. Oznaczenie ilościowe całkowitej liczby transkryptów HAC1 oraz liczby transkryptów pozbawionych intronu w transkryptomach analizowanych szczepów *Y. lipolytica*. Transkrypty HAC1 zsekwencjonowane i określone ilościowo za pomocą analizy RNAseq zmapowano pod kątem występowania eksonów i intronów. Suma oznacza znormalizowaną liczbę wszystkich transkryptów w oparciu o liczbę eksonów. Bez intronu

oznacza znormalizowaną liczbę transkryptów pozbawionych intronu. *, *** oznacza istotność statystyczną przy $p < 0.05$ i 0.001 .

Rycina 17. Indukowane i represjonowane wspólne lub specyficzne zestawy genów (DEGs) dla poszczególnych szczepów *Y. lipolytica* (*inYFP*, *scYFP* i *HAC1_scYFP*) w odpowiedzi na nadprodukcję r(s)-Prots. Wskazano liczbę DEGs: A. Indukowanych w *inYFP*, *scYFP* i *HAC1_scYFP*, B. Represjonowane w *inYFP*, *scYFP* i *HAC1_scYFP*. Analizę diagramów Venna przeprowadzono z listami DEGs ograniczonymi do log2-krotnej zmiany ekspresji genu próby badanej względem kontroli > 1.0 i wartości $p < 0.05$. Skróty: TF – czynnik transkrypcyjny, hpc – hipotetyczne białko w kładzie *Yarrowia*, ns – brak podobieństwa.

Rycina 18. Graficzne przedstawienie wzorców ekspresji wybranych genów istotnych dla syntezy i wydzielania r(s)-Prot. Czerwone strzałki (↑) znacząca zwiększona ekspresja genu w stosunku do szczepu kontrolnego; niebieskie strzałki (↓) obniżona ekspresja genu w stosunku do szczepu kontrolnego; czarna kreska (–) brak znaczącej zmiany poziomu ekspresji w stosunku do szczepu kontrolnego. Legenda wskazuje kolejność szczepów: lewa kolumna: *HAC1_scYFP*; środkowa kolumna: *scYFP*; prawa kolumna: *inYFP*.

Rycina 19. Graficzne przedstawienie deregulowanych DEGs zaangażowanych w progresję cyklu komórkowego w szczepach *HAC1_scYFP*, *scYFP* i *inYFP*. Indukcję/represję DEGs przedstawiono za pomocą kolorów, zgodnie z legendą, w oparciu o wartość ekspresji wyrażoną jako krotność zmiany (FC) ekspresji genu w próbie badanej względem kontroli. Wykres został przygotowany przy użyciu narzędzia internetowego Morpheus (<https://software.broadinstitute.org/morpheus/>).

Dorobek naukowy

Pozostałe publikacje wchodzące w skład dorobku naukowego

1. **Korpys-Woźniak P.**, Kubiak M., Borkowska M., Celińska E., Construction and Assembly of Standardized Biobricks for Synthetic Pathways Engineering in Yeasts [in:] Synthetic Biology of Yeasts: Tools and Applications, Springer (2022) (80 pkt. MEiN).
2. Celińska E., Borkowska M., **Korpys-Woźniak P.**, Kubiak M., Nicaud J.-M., Kubiak P., Gorczyca M., Białas W., Optimization of *Yarrowia lipolytica*-based consolidated biocatalyst through synthetic biology approach: transcription units and signal peptides shuffling, Applied Microbiology and Biotechnology (2020) 104:5845–5859 (IF: 4,813; 100 pkt. MEiN).
3. Kubiak M., Borkowska M., Białas W., **Korpys P.**, Celińska E., Feeding strategy impacts heterologous protein production in *Yarrowia lipolytica* fed-batch cultures—Insight into the role of osmolarity, Yeast. (2019) 36(5): 305–318 (IF: 3,239; 70 pkt. MEiN).
4. Park Y.-K., **Korpys P.**, Kubiak M., Celińska E., Soudier P., Trebulle P., Larroude M., Rossignol T., Nicaud J.-M., Engineering the architecture of erythritol-inducible promoters for regulated and enhanced gene expression in *Yarrowia lipolytica*, FEMS Yeast Research (2018) 19(1) (IF: 2,796; 100 pkt. MEiN).
5. Celińska E., Borkowska M., Białas W., **Korpys P.**, Nicaud JM., Robust signal peptides for protein secretion in *Yarrowia lipolytica*: identification and characterization of novel secretory tags. Applied Microbiology and Biotechnology (2018) 102: 5221-5233 (IF: 3,67; 100 pkt. MEiN).
6. Celińska E., Borkowska M., Białas W., Kubiak M., **Korpys P.**, Archacka M., LedesmaAmaro R., Nicaud J.-M., Genetic engineering of Ehrlich pathway modulates production of higher alcohols in engineered *Yarrowia lipolytica*, FEMS Yeast Research (2018) 19(2) (IF: 2,796; 100 pkt. MEiN).
7. Celińska E., Bonikowski R., Białas W., Dobrowolska A., Słoma B., Borkowska M., Kubiak M., **Korpys P.**, Grajek W., *Pichia cactophila* and *Kluyveromyces lactis* are Highly Efficient Microbial Cell Factories of Natural Amino Acid-Derived Aroma Compounds. Molecules (2018) 23(97) (IF: 4,412; 100 pkt. MEiN).
8. Hasiów-Jaroszewska B., Budzyńska D., Rymelska N., **Korpys P.**, Borodynko-Filas N., Phylogenetic evidence of natural reassortants in the Cucumber mosaic virus

population in Poland. Canadian Journal of Plant Pathology (2018). 40: 587-593 (IF:1.939; 70 pkt. MEiN).

Projekty badawcze

1. Diamentowy Grant MNiSW nr DI2017 000947: Intensyfikacja produkcji heterologicznych białek w komórkach niekonwencjonalnych drożdży *Yarrowia lipolytica* poprzez zastosowanie inżynierii molekularnych mechanizmów dojrzewania i sekrecji polipeptydów (przyznany na lata 2018-2024), **kierownik projektu**.
2. Preludium-21 NCN nr 2022/45/N/NZ9/00017: Inżynieria szlaku sekrecyjnego u *Yarrowia lipolytica* projektowana na podstawie wyników badań omicznych dla usprawnienia produkcji białek sekrecyjnych -charakterystyka funkcjonalna nowych genów (przyznany na lata 2023-2026), **kierownik projektu**.
3. Projekt MNiSW Iuventus Plus V nr IP2015 011074: Produkcja heterologicznych enzymów amylolitycznych w rekombinowanych komórkach drożdży i ich aplikacja w procesach biotechnologicznych (w latach 2017-2020), **wykonawca projektu**.
4. Projekt NCN OPUS 2015/17/B/NZ8/02407: Dynamika ewolucyjna wirusa czarnej pierścieniowej plamistości pomidora (*Tomato black ring virus*, TBRV) i jego adaptacja do różnych gospodarzy (w latach 2014-2016), **wolontariusz**.

Staż naukowe

1. Staż w ramach projektu pt.: „Studiujesz – praktykuj. Program stażowy dla studentów Wydziału Rolnictwa i Bioinżynierii Uniwersytetu Przyrodniczego w Poznaniu” (UDAPOWR.03.01.00-00-S237/15) – zrealizowany w terminie I-III 2018 r. w INRA (Institut National de la Recherche Agronomique), MICALIS INSTITUT (Domaine de Vilvert, F78350 Jouy-en-Josas, Francja) pod kierunkiem Prof. Jean-Marca Nicaud. Temat: Engineering the architecture of erythritol-inducible promoters for regulated and enhanced gene expression in *Yarrowia lipolytica*.

Doniesienia pokonferencyjne

1. **Korpys-Woźniak P., Celińska E.** Omics-guided engineering of secretory pathway for enhanced synthesis of secretory proteins in *Yarrowia lipolytica*, międzynarodowa konferencja 1st Polish Yeast Conference (Rzeszów, 22-24.06.2022r.), prezentacja ustna.

2. **Korpys-Woźniak P.**, Celińska E. Insight into *Yarrowia lipolytica*'s molecular response to overproduction of heterologous proteins with different biochemical characteristics, międzynarodowa konferencja ICY15 International Congress on Yeasts (Wiedeń, online, 23-27.08.2021 r.), prezentacja ustna.
3. **Celińska E.**, Kubiak-Szymendera M., **Korpys-Woźniak P.**, Borkowska M., Białas W., Kubiak P., Gorczyca M., Kaźmierczak J., Pryszcz L., Steels S., Fickers P., Nicaud J.-M. Molecular and bioprocess studies for superior production of recombinant secretory proteins in *Yarrowia lipolytica*, międzynarodowa konferencja ICY15 International Congress on Yeasts (Wiedeń, online, 23-27.08.2021 r.), poster.
4. **Korpys P.**, Borkowska M., Białas W., Kubiak M., Nicaud JM., Celińska E. Secretory protein production in nonconventional yeast *Yarrowia lipolytica*, międzynarodowa konferencja 46th Annual Conference on Yeasts (Smolenice, Słowacja, 7-10.05.2019 r.), nagroda za najlepszy poster.
5. **Kubiak M.**, Borkowska M., Białas W., **Korpys P.**, Celińska E. Insight into the role of osmolarity in heterologous protein production in *Yarrowia lipolytica*, międzynarodowa konferencja 46th Annual Conference on Yeasts (Smolenice, Słowacja, 7-10.05.2019 r.), prezentacja ustna.
6. **Celińska E.**, Borkowska M., **Korpys P.**, Kubiak M., Białas W., Gorczyca M., Nicaud J.-M. Serious study on secretory protein synthesis and optimization of multi-gene expression constructs for *Yarrowia lipolytica*, międzynarodowa konferencja 46th Annual Conference on Yeasts (Smolenice, Słowacja, 7-10.05.2019 r.), prezentacja ustna.
7. **Korpys P.**, Kubiak M., Borkowska M., Białas W., Nicaud J.-M., Celińska E. Nowe kierunki modyfikacji genetycznych drożdży z gatunku *Yarrowia lipolytica* mające na celu intensyfikację produkcji heterologicznych białek, I Ogólnopolska Konferencja Biotechnologia niejedno ma imię (Poznań, 24-25.11.2018 r.), poster.
8. **Kubiak M.**, **Korpys P.**, Białas W., Borkowska M., Celińska E. *Yarrowia lipolytica* – drożdże do zadań specjalnych – inżynieryjne strategie doskonalenia, I Ogólnopolska Konferencja Biotechnologia niejedno ma imię (Poznań, 24-25.11.2018r.), poster, wyróżnienie.
9. Celińska E., **Korpys P.**, Borkowska M., Białas W., Kubiak M., Nicaud J.-M. Identification and evaluation of new signal peptides for secretory protein production in *Yarrowia lipolytica*, międzynarodowa konferencja Advances in Microbiology and

- Biotechnology (Lwów, Ukraina, 29-31.10.2018r.), nagroda za najlepszy poster (II miejsce).
10. Kubiak M., Borkowska M., Białas W., **Korpys P.**, Celińska E. Effect of feeding frequency on recombinant protein production in *Yarrowia lipolytica* fed-batch culture - preliminary indications of osmotic stress influence, międzynarodowa konferencja Advances in Microbiology and Biotechnology (Lwów, Ukraina, 29-31.10.2018 r.), poster.
 11. Celińska E., Borkowska M., Białas W., **Korpys P.**, Kubiak M., Nicaud J.-M. Production of heterologous secretory proteins in *Yarrowia lipolytica*, międzynarodowa konferencja Advances in Microbiology and Biotechnology (Lwów, Ukraina, 29-31.10.2018 r.), prezentacja ustna.
 12. Park Y.-K., **Korpys P.**, Soudier P., Kubiak M., Celińska E., Trebulle P., Larroude M., Nicaud J.-M., Rossignol T., Engineering architecture of inducible promoters for regulated and enhanced gene expression in *Yarrowia lipolytica*, międzynarodowa konferencja NonConventional Yeasts Conference w Rzeszowie (15-18.05.2018 r.), poster.
 13. Kubiak M., Borkowska M., Białas W., **Korpys P.**, Celińska E., Manipulation of feeding mode as a strategy for enhancing heterologous protein production in *Yarrowia lipolytica*, międzynarodowa konferencja Non-Conventional Yeasts Conference (Rzeszów, 15- 18.05.2018 r.), poster.
 14. Celińska E., Borkowska M., Białas W., **Korpys P.**, Kubiak M., Nicaud J.-M., New secretory tags for protein secretion in *Yarrowia lipolytica*, międzynarodowa konferencja Non-Conventional Yeasts Conference (Rzeszów, 15-18.05.2018r.), prezentacja ustna.
 15. **Korpys P.**, Kubiak M., Borkowska M., Celińska E., Klonowanie modularne Golden Gate jako cenne narzędzie do nadekspresji białek w komórkach niekonwencjonalnych drożdży *Yarrowia lipolytica*, II Ogólnopolskie Sympozjum Nauk Przyrodniczo-Rolniczych w Poznaniu (7-8.04.2018r.), prezentacja ustna, zajęcie I miejsca.
 16. Kubiak M., **Korpys P.**, Borkowska M., Białas W., Celińska E., Dobór warunków hodowli okresowo-dolewowych w procesie produkcji enzymów amylolitycznych przez rekombinowane szczepy drożdży *Yarrowia lipolytica*, II Ogólnopolskie Sympozjum Nauk Przyrodniczo-Rolniczych (Poznań, 7-8.04.2018r.), prezentacja ustna.

17. **Korpys P.**, Kubiak M., Celińska E., Wykorzystanie techniki Golden Gate do konstrukcji kaset ekspresyjnych dedykowanych dla drożdży *Yarrowia lipolytica*, IV Poznańskie Sympozjum Młodych Naukowców (Poznań, 18.11.2017 r.), poster.
18. **Kubiak M.**, **Korpys P.**, Białas W., Celińska E., Produkcja amylazy w hodowlach okresowo-dolewowych rekombinowanych drożdży *Yarrowia lipolytica*, IV Poznańskie Sympozjum Młodych Naukowców (Poznań, 18.11.2017 r.), poster.
19. **Korpys P.**, Kubiak M., Celińska E., Potencjał przemysłowy enzymów otrzymywanych w wyniku heterologicznej ekspresji w *Yarrowia lipolytica*, XV Konferencja „Biotechnologia: Dziś na Uniwersytecie Technologiczno-Przyrodniczym, jutro w regionie kujawsko-pomorskim” (Bydgoszcz, 9.06.2017 r.), poster.
20. **Korpys P.**, Celińska E., Biomasa drożdży *Yarrowia lipolytica* jako alternatywne źródło tłuszczu, białka i witamin w dodatkach do żywności, międzynarodowa II Konferencja Młodych Naukowców w Warszawie (25-26.05.2017), prezentacja ustna, wyróżnienie.
21. **Korpys P.**, Kubiak M., Celińska E., *Yarrowia lipolytica* – unikatowy metabolizm w procesach biotechnologicznych, I Ogólnopolskie Sympozjum Nauk Przyrodniczo Rolniczych (Poznań, 22-23.04.2017 r.), poster.
22. **Kubiak M.**, **Korpys P.**, Borkowska M., Celińska E., Produkcja owadziej alfa-amylazy w rekombinowanych komórkach *Yarrowia lipolytica* – kinetyka procesu i korelacja aktywności z liczbą kopii genu, I Ogólnopolskie Sympozjum Nauk Przyrodniczo Rolniczych w (Poznań, 22-23.04.2017), poster, wyróżnienie.

Monografie pokonferencyjne

1. **Korpys P.**, Kubiak M., Borkowska M., Celińska E., pt. Klonowanie modularne Golden Gate jako cenne narzędzie do nadekspresji białek w komórkach niekonwencjonalnych drożdży *Yarrowia lipolytica*. Monografia pokonferencyjna II Ogólnopolskiego Sympozjum Nauk Przyrodniczo-Rolniczych w Poznaniu pt. *Rozwój nauk przyrodniczo – rolniczych* (red. K. Mikołajczak i M. Kubiak), ISBN 978-83-943832-2-0 (Wyd. Uniwersytetu Przyrodniczego w Poznaniu).
2. Kubiak M., **Korpys P.**, Borkowska M., Białas W., Celińska E. Dobór warunków hodowli okresowo-dolewowych w procesie produkcji enzymów amylolitycznych przez rekombinowane szczepy drożdży *Yarrowia lipolytica*. Monografia pokonferencyjna II Ogólnopolskiego Sympozjum Nauk Przyrodniczo-Rolniczych w

Poznaniu pt. *Rozwój nauk przyrodniczo – rolniczych* (red. K. Mikołajczak i M. Kubiak), ISBN 978-83-943832-2-0 (Wyd. Uniwersytetu Przyrodniczego w Poznaniu).

OŚWIADCZENIE

Oświadczam, że jestem współautorem pracy będącej częścią monotematycznego cyklu publikacji, wchodzącego w skład rozprawy doktorskiej Pani mgr inż. Pauliny Korpys-Woźniak:

Korpys-Woźniak P., Kubiak P., Białas W., Celińska E. (2020) Impact of overproduced heterologous proteins characteristics on physiological response in *Yarrowia lipolytica* steady-state maintained continuous cultures. *Applied Microbiology and Biotechnology*. 104:9785-9800.

Tytuł, imię i nazwisko	Udział %	Wkład w przygotowanie pracy	Miejscowość i data	Podpis
mgr inż. Paulina Korpys-Woźniak	70	Wykonanie prac laboratoryjnych, w tym: konstrukcja rekombinowanych szczepów <i>Y. lipolytica</i> , screening szczepów, prowadzenie ciągłych hodowli bioreaktorowych i analiz pobranych prób; Zebranie i przygotowanie danych do analiz statystycznych; Opracowanie otrzymanych wyników, w tym: przygotowanie rycin, tabel; Udział w dyskusji wyników; Napisanie manuskryptu; Zapewnienie finansowania	Poznań 22.05.23.	Paulina Korpys-Woźniak
dr inż. Piotr Kubiak	5	Nadzór nad prowadzeniem ciągłych hodowli bioreaktorowych; Wykonanie analizy statystycznej wyników	Poznań 22.05.2023	Kubiak
Prof. UPP dr hab. Wojciech Białas	5	Nadzór nad prowadzeniem ciągłych hodowli bioreaktorowych; Wykonanie analizy statystycznej wyników	Poznań 22.05.2023	Białas
Prof. UPP dr hab. Ewelina Celińska	20	Stworzenie koncepcji i zaprojektowanie badań; Koordynowanie badań laboratoryjnych; Analiza otrzymanych danych; Interpretacja wyników i dyskusja z danymi literaturowymi; Merytoryczne redagowanie manuskryptu	Poznań 22-05-2023	Celińska

OŚWIADCZENIE

Oświadczam, że jestem współautorem pracy będącej częścią monotematycznego cyklu publikacji, wchodzącego w skład rozprawy doktorskiej Pani mgr inż. Pauliny Korpys-Woźniak:

Korpys-Woźniak P., Celińska E. (2021) Global transcriptome profiling reveals genes responding to overproduction of a small secretory, a high cysteine- and a high glycosylation-bearing protein in *Yarrowia lipolytica*. *Biotechnology Reports*. 31.

Tytuł, imię i nazwisko	Udział %	Wkład w przygotowanie pracy	Miejscowość i data	Podpis
mgr inż. Paulina Korpys-Woźniak	70	Zaprojektowanie doświadczeń; Wykonanie eksperymentów, w tym: prowadzenie ciągłych hodowli bioreaktorowych i analiz pobranych próbek, izolacja RNA; Przygotowanie i zabezpieczenie prób do sekwencjonowania RNA; Analiza bioinformatyczna danych omicznych; Przygotowanie rycin, tabel zawartych w artykule; Napisanie artykułu; Zapewnienie finansowania	Poznań 22.05.23.	Paulina Korpys- Woźniak
Prof. UPP dr hab. Ewelina Celińska	30	Stworzenie koncepcji i zaprojektowanie badań; Koordynowanie badań laboratoryjnych; Analiza otrzymanych danych; Analiza bioinformatyczna; Interpretacja wyników i dyskusja z danymi literaturowymi; Merytoryczne redagowanie manuskryptu	Poznań 22.05.2023	Celińska

OŚWIADCZENIE

Oświadczam, że jestem współautorem pracy będącej częścią monotematycznego cyklu publikacji, wchodzącego w skład rozprawy doktorskiej Pani mgr inż. Pauliny Korpys-Woźniak:

Korpys-Woźniak P., Kubiak P., Celińska E. (2021) Secretary helpers for enhanced production of heterologous proteins in *Yarrowia lipolytica*. *Biotechnology Reports*. 32.

Tytuł, imię i nazwisko	Udział %	Wkład w przygotowanie pracy	Miejscowość i data	Podpis
mgr inż. Paulina Korpys-Woźniak	75	Zaprojektowanie doświadczeń; Wykonanie eksperymentów, w tym: przygotowanie biblioteki rekombinowanych szczepów <i>Y. lipolytica</i> , przeprowadzenie hodowli w kolbach wytrząsanych; Wykonanie większości analiz; Zebranie i przygotowanie danych do analiz statystycznych; Opracowanie i interpretacja otrzymanych wyników, w tym: przygotowanie rycin, tabel; Napisanie artykułu; Zapewnienie finansowania	Poznań 22.05.23.	Paulina Korpys- Woźniak
dr inż. Piotr Kubiak	5	Wykonanie analizy statystycznej wyników i ich wizualizacja	Poznań 22.05.2023	Kubiak
Prof. UPP dr hab. Ewelina Celińska	20	Stworzenie koncepcji i zaprojektowanie badań; Koordynowanie badań laboratoryjnych; Przeanalizowanie otrzymanych danych; Interpretacja wyników i dyskusja z danymi literaturowymi; Merytoryczne redagowanie manuskryptu	Poznań 22.05.2023	Celińska

OŚWIADCZENIE

Oświadczam, że jestem współautorem pracy będącej częścią monotematycznego cyklu publikacji, wchodzącego w skład rozprawy doktorskiej Pani mgr inż. Pauliny Korpys-Woźniak:

Korpys-Woźniak P., Celińska E. (2023) Molecular background of *HAC1*-driven improvement in the secretion of recombinant protein in *Yarrowia lipolytica* based on comparative transcriptomics. *Biotechnology Reports*.

Tytuł, imię i nazwisko	Udział %	Wkład w przygotowanie pracy	Miejscowość i data	Podpis
mgr inż. Paulina Korpys-Woźniak	70	Zaprojektowanie doświadczeń; Wykonanie większości eksperymentów, w tym: prowadzenie ciągłych hodowli bioreaktorowych i analiz pobranych próbek, izolacja RNA; Przygotowanie i zabezpieczenie prób do sekwencjonowania RNA; Analiza bioinformatyczna danych omicznych; Opracowanie i interpretacja otrzymanych wyników, w tym: przygotowanie rycin, tabel; Napisanie artykułu; Zapewnienie finansowania	Poznań 22.05.23r.	Paulina Korpys- Woźniak
Prof. UPP dr hab. Ewelina Celińska	30	Stworzenie koncepcji i zaprojektowanie badań; Koordynowanie badań laboratoryjnych; Przeanalizowanie otrzymanych danych; Analiza bioinformatyczna; Interpretacja wyników i dyskusja z danymi literaturowymi; Merytoryczne redagowanie manuskryptu	Poznań 22.05.2023	Celińska



Impact of overproduced heterologous protein characteristics on physiological response in *Yarrowia lipolytica* steady-state-maintained continuous cultures

Paulina Korpys-Woźniak¹ · Piotr Kubiak¹ · Wojciech Białas¹ · Ewelina Celińska¹

Received: 10 July 2020 / Revised: 23 September 2020 / Accepted: 27 September 2020 / Published online: 6 October 2020
© The Author(s) 2020

Abstract

Overproduction of recombinant secretory proteins triggers numerous physiological perturbations. Depending on a given heterologous protein characteristics, the producer cell is faced with different challenges which lead to varying responses in terms of its physiology and the target protein production rate. In the present study, we used steady-state-maintained *Yarrowia lipolytica* cells to investigate the impact of different heterologous proteins on the physiological behavior of the host cells. Such an approach allowed to uncouple the impact of the overproduction of a particular protein from the phenomena that result from growth phase or are caused by the heterogeneity of the analyzed populations. Altogether, eight variants of recombinant strains, individually overproducing heterologous proteins of varying molecular weight (27–65 kDa) and reporting activity (enzymatic and fluorescent) were subjected to chemostat cultivations. The steady-state-maintained cells were analyzed in terms of the substrate utilization, biomass and metabolites production, as well as the reporter protein synthesis. Simplified distribution of carbon and nitrogen between the respective products, as well as expression analysis of the heterologous genes were conducted. The here-obtained data suggest that using a more transcriptionally active promoter results in channeling more C flux towards the target protein, giving significantly higher specific amounts and production rates of the target polypeptide, at the cost of biomass accumulation, and with no significant impact on the polyols production. The extent of the reporter protein's post-translational modifications, i.e., the number of disulfide bonds and glycosylation pattern, strongly impacts the synthesis process. Specific responses in terms of the protein formation kinetics, the gene expression levels, and transcript-to-protein linearity were observed.

Key Points

- Eight expression systems, producing different reporter proteins were analyzed.
- The cells were maintained in steady-state by continuous chemostat culturing.
- Protein- and promoter-specific effects were observed.

Keywords Heterologous protein · Recombinant strain · Overexpression · Overproduction · Secretion · Chemostat · Steady state

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00253-020-10937-w>) contains supplementary material, which is available to authorized users.

✉ Ewelina Celińska
ewelina.celinska@up.poznan.pl

¹ Department of Biotechnology and Food Microbiology, Poznan University of Life Sciences, ul. Wojska Polskiego 48, 60-627 Poznań, Poland

Introduction

Growing demand for industrial enzymes and biopharmaceutical proteins triggers increasing interest in microbial cell factories that serve as protein production platforms. It is now well recognized that yeast-based expression systems offer a unique combination of the ease of manipulation with complex folding and maturation of the target polypeptide. Among several industrially relevant hosts, *Yarrowia lipolytica* emerges as a highly attractive system for heterologous protein production (Madzak 2018), which was enabled by the expansion of the synthetic biology toolbox designed for this species (Wagner and Alper 2016; Larroude et al. 2018; Park et al. 2019).

Moreover, *Y. lipolytica* was found to perform well in bioreactor cultures in terms of cell growth, secretory protein titer, and productivity (Celińska et al. 2017a; Theron et al. 2020). It was evidenced that *Y. lipolytica* produces high levels of active enzyme, generates stable isotope-labelled variants of a secretory protein, or that it is suited for the production of complex high molecular weight proteins targeted for secretion (Madzak et al. 2005; Boonvitthya et al. 2013; Nars et al. 2014; Theron et al. 2020). Recently, *Y. lipolytica* was shown to manage secretory production of a heterologous protein of molecular weight > 120 [kDa] (Swietalski et al. 2020).

Overproduction of recombinant secretory proteins triggers numerous physiological perturbations to the host cell, frequently leading to decreased growth rate and underperformance in terms of the desired polypeptide production (Graf et al. 2009; Puxbaum et al. 2015). It is now well recognized that depending on the characteristics of a given heterologous protein, the producer cell is faced with different challenges, and that eventual bottlenecks and limitations may occur at different levels of the protein production/secretory pathway (Gasser et al. 2007; Nocon et al. 2014). Omics-assisted approaches allow to get deeper insight into actual processes taking place in the cell upon overproduction of a specific protein. It was demonstrated that over 60 genes were transcriptionally changed when either of two different heterologous proteins were overproduced in *Saccharomyces cerevisiae* (Liu et al. 2013). Correspondingly, different genetic engineering strategies allow to alleviate limitations encountered upon overproduction of different heterologous proteins (Hou et al. 2012; Puxbaum et al. 2015; Celińska and Nicaud 2019). It was highlighted that larger proteins, even if produced at moderate level, impose much higher metabolic burden to the producer cells than smaller ones (Liu et al. 2013). In that study, it was shown that upon overproduction of a larger secretory protein, the genes involved in the amino acid and energy metabolisms were significantly upregulated when compared with even more intensive production of a smaller protein. This leads to a conclusion that, depending on the overproduced protein characteristics, the response from the host cell in terms of its physiology and the production rate will be different.

The kinetics of production of a heterologous secretory protein and the host's response to such a challenge can be most accurately studied in steady-state-maintained cells (Arvas et al. 2011; Liu et al. 2013; Peebo and Neubauer 2018). During a steady state, the concentration of intra- and extracellular molecules produced or metabolized by cultured cells remain constant over prolonged time which is counted in culture residence times (Hoskisson and Hobbs 2005). The steady state can be reached by synchronized feeding of nutrients and withdrawal of culture from a bioreactor vessel at a specified dilution rate (D ; [h^{-1}]). It was observed that a culture reaches the steady state by somehow matching its growth rate exactly

to the dilution rate over a wide range of dilution rate values. One of the key advantages of such a mode of cultivation is that under the steady state, the population is reasonably homogeneous and synchronized in a specific growth phase. Such approach eliminates a number of molecular phenomena related to growth phase, stress response, or the other types of heterogeneities occurring in a typical batch culture, which, if not controlled, may lead to incorrect interpretation of physiological changes occurring in the cell. Chemostat cultivations are relevant for both applied and basic research and are frequently used in studies on recombinant protein production in a variety of microbial expression platforms (e.g., *Escherichia coli*, *Komagataella phaffii*, *S. cerevisiae*, and *Aspergillus niger*) (Pedersen et al. 2000; Rodríguez-Carmona et al. 2012; Liu et al. 2013; Looser et al. 2014). So far, most of the research involving chemostats and *Y. lipolytica* was focused on the production of small molecular weight metabolites, such as citric acid, erythritol, and lipids (Rywińska et al. 2011; Ochoa-Estopier and Guillouet 2014; Rakicka et al. 2017), and no investigations into heterologous protein production kinetics under steady state have been conducted to date.

In the present study, we used continuous chemostat cultures to develop steady-state-maintained populations of *Y. lipolytica* recombinant strains producing four variants of reporter proteins—a fluorescent reporter in an intracellular and secreted form, and two enzymatic reporters targeted for secretion. The strains were constructed in two expression platforms, comprising the most popular *Y. lipolytica* expression hosts. Altogether, eight variants of recombinant strains, overproducing individually the heterologous proteins, and two reference strains were subjected to chemostat cultivations. Steady-state-maintained cells were analyzed in terms of the biomass and metabolite production, substrate utilization kinetics, and reporter protein production. Simplified distribution of carbon (C) and nitrogen (N) between the respective products and expression of the heterologous genes were analyzed as well.

Materials and methods

Strains and small-scale cultivations

All the strains and vectors used in this study are listed in Table S1. Regarding *Y. lipolytica* host strains used in this study, both Po1h and Po1g originated from the same wild-type strain (W29) and are deleted for both extracellular proteases (AEP and AXP). On the other hand, the strains differ in the auxotrophy selection marker (Leu2 for Po1g and Ura3 for Po1h), integration site of the recombinant cassette (pBR-platform and random integration via zeta elements), and the promoter-governing expression of the cloned genes (hp4d and 4UAS-pTEF, respectively). *Y. lipolytica* and *E. coli*

strains were routinely maintained as described in Barth and Gaillardin (1996) and Sambrook and Russell (2001). Briefly, *Escherichia coli* strain JM109 was grown in LB medium ([g L⁻¹]: yeast extract (BTL, Lodz, Poland), 5; bacto peptone (BTL), 10; NaCl (POCH, Gliwice, Poland), 5) supplemented with appropriate antibiotic (ampicillin at 100 [mg L⁻¹] or kanamycin at 40 [μg L⁻¹]) and agar ((Biomaxima, Lublin, Poland), 15 [g L⁻¹]) at 37 °C and with 250 rpm shaking. *Y. lipolytica* strains were cultured in yeast nitrogen base (YNB; [g L⁻¹]: YNB (Sigma-Aldrich, Merck KGaA, St. Louis, USA), 1.7; (NH₄)₂SO₄ (PoCh), 5; glucose (PoCh), 20), yeast extract-peptone-dextrose (YPD; [g L⁻¹]: yeast extract, 10; bacto peptone, 20; glucose, 20), or yeast extract-peptone-starch (YPS; [g L⁻¹]: yeast extract, 10; bacto peptone, 20; starch (Sigma-Aldrich), 10; glucose, 20) solidified with agar (15 [g L⁻¹]), at 30 °C and with 250 rpm shaking.

Molecular biology protocols

Standard molecular biology techniques were used throughout the research (Barth and Gaillardin 1996; Sambrook and Russell 2001). Manipulations with DNA fragments were conducted using appropriate kits from A&A Biotechnology (Gdynia, Poland). DNA was amplified using Phire DNA polymerase (Thermo Fisher Scientific, Waltham, MA USA) or RUN DNA polymerase (A&A Biotechnology). Two *Y. lipolytica* strains were used as hosts in this study: Po1g and Po1h (Table S1). Three genes encoding recombinant proteins were expressed individually in each host (Table S2), as follows: (i) enzymatic reporter: α-amylase (SoA) derived from rice weevil *Sitophilus oryzae* (52 [kDa]), (ii) enzymatic reporter glucoamylase (TIG) from fungus *Thermomyces lanuginosus* (65 [kDa]), and (iii) fluorescent reporter: YFP (Yellow Fluorescent Protein, YFP) (27 [kDa]). The last reporter was cloned as either intracellular (inYFP) or secreted protein (scYFP). Altogether, four variants of the reporters were used: scSoA, scTIG, scYFP, inYFP. According to our previous studies, neither SoA nor TIG activity could be detected inside the *Y. lipolytica* host cells (unpublished), making inSoA/inTIG variants pointless.

In *Y. lipolytica* Po1g host, the genes were cloned using a commercial cloning system (YLEX™ kit; Yeastern Biotech Co. Ltd; Fig. 1a), integrating the expression cassettes in a pBR322 docking platform. For the secretion, the genes were transcriptionally fused with preXPR2 signal sequence. Expression of the heterologous genes was governed by a strong, semi-constitutive promoter hp4d (Madzak et al. 2000). In *Y. lipolytica* Po1h host strain, the genes were cloned using Golden Gate expression cassettes bearing a single transcription unit (TU) (Celińska et al. 2017b) (Fig. 1b), under the control of a strong, semi-constitutive 4UAS-pTEF promoter (Dulermo et al. 2017). The reactions were conducted using T4 ligase and *Bsa*I restriction endonuclease from New England

Biolabs (NEB, Ipswich, MA, USA). For the secretion, the genes were transcriptionally fused with a SP1 signal sequence, native for exo-1,3-beta-glucanase (YALI0B03564g) (Celińska et al. 2018). The expression cassette assembly was conducted as described previously (Celińska et al. 2017b). All the clonings were conducted using primers listed in Table S3.

The expression cassettes were sequenced (Genomed, Warsaw, Poland) to verify their correctness, linearized with *Not*I (Thermo Fisher Scientific) and transformed into Po1h and Po1g *Y. lipolytica* strains. Clones appearing after 48 h were replicated on a fresh YNB plate and verified for the presence of the introduced heterologous genes through colony PCR with specific primers (Table S3). Acquired amylolytic phenotypes (scSoA, scTIG) were screened via starch-iodine drop test on YPS agar plates, as described previously (Celińska et al. 2016). Acquired fluorescence phenotypes (inYFP, scYFP) was screened using AxioVert fluorescence microscope (ZEISS, Oberkochen, Germany) and quantitatively assessed in microcultures using an automatic plate reader/fluorimeter (Tecan Infinite M200; Tecan Group Ltd., Männedorf, Switzerland). Positive strains were deposited as glycerol stocks at - 80 °C.

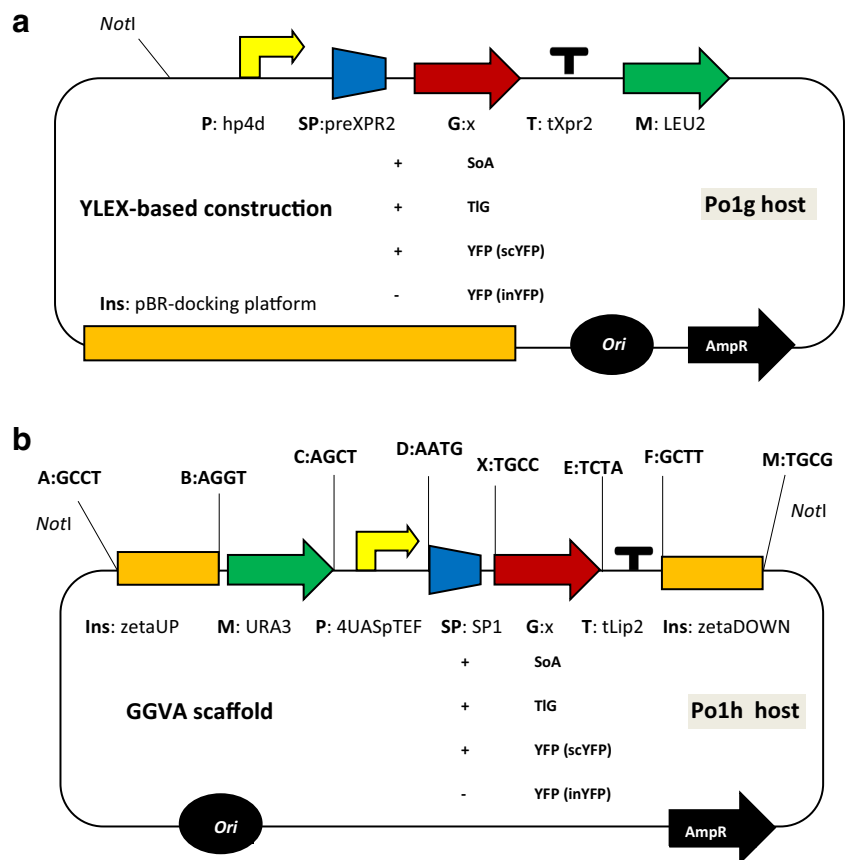
Shake-flask cultures for selection of representative sub-clones

Three sub-clones were randomly selected from amongst the positive *Y. lipolytica* transformants. The sub-clones were subjected to primary screening to select representative strains for further studies. Screening cultures were conducted in 50 mL shake flasks containing 5 mL of YPG₂₀ medium ([g L⁻¹]: yeast extract, 10; peptone, 20, glycerol, 20) at 28 °C and 250 rpm shaking for 48 h. Samples were periodically withdrawn from the cultures and, after centrifugation, the biomass and the supernatant samples were stored separately, at - 20 °C, until analyzed. After thawing, the biomass was washed twice in sterile saline solution (0.85% NaCl) and appropriately diluted cell suspensions were subjected to spectrophotometric measurements at 600 nm wavelength (Tecan Infinite M200).

Inoculum preparation

Selected, representative sub-clones were revived by transferring the biomass from the glycerol stock onto YPD agar plate and incubating at 30 °C for 24 h. Single colonies were transferred into 100 mL shake flasks containing 30 mL of YPG₂₀ medium and cultured for 22 h, at 28 °C and 250 rpm in an orbital shaker incubator (BIOSAN, ES-20, Riga, Latvia). Such pre-cultures (30 mL) were aseptically transferred into a bioreactor vessel (10% of the bioreactor working volume).

Fig. 1 Cloning vectors used in this study. **a** YLEX commercial cloning system-based vectors. **b** Golden Gate assembly-based vectors. The GGVA was assembled using a set of 4-nt overhangs indicated in the scheme (A, B, C, D, X, E, F, and M). Abbreviations: Ins, regions for non-homologous recombination (GGVA) or directing integration with pBR-docking platform (YLEX); M, selection marker LEU2/URA3; P, promoter 4UAS-pTEF/hp4d; SP, signal peptide (derived from spYALI0B03564g for GGA or XPR2 gene for YLEX-based expression); G, target gene (one of the three SoA/TIG/YFP devoid of their native SP); T, terminator tLip2/tXpr2; ori, bacterial origin of replication; AmpR, bacterial selection marker. The expression cassettes were digested with *NotI* restriction endonuclease prior to transformation



Bioreactor cultivations

Batch and continuous cultures were performed in a 0.5-L stirred tank Multifors 2 bioreactor (Infors HT, Bottmingen-Basel, Switzerland) equipped with two flat-bladed turbines with six blades (Rushton turbine). Throughout the cultures, the temperature, pH and dissolved oxygen (DO; [%]) were stabilized at 28 °C, pH 5.5 by automatic addition of 20% NaOH, 10% H₂SO₄, and at 20 [%] (oxygen saturation) by using cascade-control approach with stirring from 100 up to 1200 rpm, and constant aeration at 2.0 vvm. Foaming was minimized by the addition of a defoaming agent AntiFoam 204 (Sigma-Aldrich). The cultures were conducted at a working volume of 0.3 L and were initiated with a batch culture stage in YPG₁₀₀ medium ([g L⁻¹]: yeast extract, 10; bactopectone, 20; glycerol, 100). Continuous cultivation stage was initiated when glycerol was nearly completely consumed, which was indicated by an increase in the DO parameter. Continuous stage was executed by constantly feeding YPG₁₀₀ medium and withdrawing an equal volume of the culture at a working dilution rate [*D*; h⁻¹] of approximately 0.20 [h⁻¹]. The *D* value was determined in preliminary studies. The steady state, marked by constant concentration of the biomass, residual glycerol, and metabolites, was reached after at least six residence times [*nV*]. Two milliliter samples were

collected periodically, centrifuged at 12,045×*g* for 6 min (Eppendorf MiniSpin; Hamburg, Germany) and stored at -20 °C until analyzed, except for the YFP-containing samples. For the YFP-producing strains, the biomass was washed in sterile saline solution and resuspended in the equal volume of the same solution prior to storage. The supernatants and the washed biomass were stored at -80 °C for further analyses. All the cultures were conducted in at least two biological replicates.

Analytical methods

Biomass concentration and cell viability

The biomass samples were defrosted and washed twice with a sterile saline solution. Dry cellular weight (DCW; [g_{DCW} L⁻¹]) was determined gravimetrically through drying the biomass in a laboratory dryer at 105 °C for 2 - 3 days, until a stable readout was reached. The dry biomass concentration was expressed in grams of the cell dry mass per liter [g_{DCW} L⁻¹]. The strains viability during the bioreactor cultivations was assessed by surface-plating and colony counts. Upon sampling, the cultures were decimally diluted in a sterile saline solution in MTP up to 10¹⁰-fold, followed by plating on YPD agar plates using a stainless steel replicator (Sigma-Aldrich).

The plates were incubated for 24 h at 30 °C. The results were expressed as surface colony forming units [scfu] reflecting the number of colony counts at the highest decimal dilution, where the colony growth was observed.

Concentration of the substrate and metabolites determined by HPLC

The culture supernatants were thawed on ice, diluted, and passed through 0.45 µm membrane syringe filters (Millipore; Merck-Millipore, Burlington, MA, USA). A high-performance liquid chromatograph, Agilent Technologies 1200 series (Agilent Technologies, Santa Clara, CA, USA), was used to determine the concentration of glycerol (GLY; [g L⁻¹]) and metabolites (erythritol, ERY; mannitol, MAN; citric acid, CA; α-ketoglutaric acid, α-KG [g L⁻¹]) contained in the culture liquid. The apparatus was equipped with a refractive-index detector (G1362A) and a Rezex ROA-Organic Acid H+ column (Phenomenex, Torrance, CA, USA). Operating conditions were as follows: 0.005 N H₂SO₄ as eluent at a flow rate of 0.6 [mL min⁻¹]; the column temperature was set at 40 °C. External standards (purchased from Sigma-Aldrich) were used for identification and quantification of the peaks areas in chromatograms, which were analyzed using ChemStation for LC 3D software (Agilent).

Fluorescence analysis

All the fluorescence measurements were conducted in flat-bottomed MTP plates (Corning; Sigma-Aldrich) in Tecan Infinite M200 automatic plate reader at the following wavelength settings: excitation at 495 nm/emission at 527 nm. The extracellular YFP fluorescence was measured in 200 µL of the supernatant samples defrosted on ice. Fresh culture medium (also stored at -80 °C) was used as the background reference. The intracellular YFP fluorescence was measured using the pre-washed biomass samples, thawed on ice and resuspended in sterile saline solution. Serial dilutions were conducted prior to the measurements. The reference strains (Po1g_Leu2+ and Po1h_Ura3+) were used for the background normalization. Importantly, the pre-washed biomass of the reference strains did not exhibit intrinsic, background fluorescence under the adopted assaying conditions. All the measurements were conducted in technical triplicate out of each biological duplicate. After normalization, the fluorescence results were expressed in relative fluorescence units [RFU]. RFU was defined as the sample median fluorescence value minus the background fluorescence value (fresh medium or the reference strains biomass). Additionally, the fluorimetry results were randomly verified through observations under fluorescence microscope (ZEISS AxioVert, AxioCam 350 color; filterset: 09; ZEISS).

Enzymatic assays—SoA-microSIT and TIG-microDNS

Miniaturized colorimetric assays were used to determine the extracellular activity of both enzymatic reporters. The SoA amylolytic activity was examined in the supernatant samples using microSIT assay, described previously (Borkowska et al. 2019). One activity unit (α-amylase activity unit [AAU]) refers to the amount of enzyme that decomposes 1 mg of starch per 1 mL, during 1 min, at pH 5.0 and 40 °C, under the applied experimental conditions. The TIG glucoamylase activity was examined in the supernatant samples using microDNS (3,5-dinitrosalicylic acid) method according to (Goncalves et al. 2010). One activity unit (glucoamylase activity unit [GAU]) was defined as the amount of enzyme that releases 1 µg of reducing sugar equivalents per 1 mL, during 1 h, at pH 5.0 and 40 °C, under the adopted assay conditions. In both enzymatic assays, the reference strains Po1h_Ura3+/Po1g_Leu2+ were assayed under the same conditions to assess any possible background activity. All the readouts were normalized per blank reaction with distilled water. All the colorimetric enzymatic assays were read using the Tecan Infinite M200 plate reader at 580 nm (microSIT) and 540 nm (microDNS) wavelength. Each sample was analyzed in technical duplicates.

Heterologous gene expression analysis through RTqPCR

Expression level of the target genes (SoA, TIG, YFP) was analyzed in the samples withdrawn from the cultures in the steady state. The biomass from 2 mL of the culture was used for isolation of total RNA. RNA isolation was performed using Bead-Beat Total RNA Mini Kit (A&A Biotechnology). Isolated RNA was verified for quantity and quality through gel electrophoresis and spectrophotometric measurement (NanoDrop, Thermo Fisher Scientific). cDNA synthesis was conducted using SuperScript III Reverse Transcriptase and oligo(dT) primer, according to the manufacturer's protocol (Thermo Fisher Scientific). cDNA preparations were used as templates in RTqPCR, carried out in an Applied Biosystems 7500 device (Applied Biosystems, Foster City, USA). The reactions were set up using RT HSPCR Mix SYBR® B (A&A Biotechnology) according to the manufacturer's specifications. LoROX dye was used as a passive reference. The expression level of the genes was normalized to the expression level of the actin gene (*ACT1*), used as internal calibrator. All the primers for RTqPCR were designed with Primer Expert Software (Applied Biosystems) and are listed in Table S3. The following thermal profile was used: 95 °C 3 min, (95 °C 15 s, 60 °C 30 s, 72 °C 30 s) × 40, 72 °C 1 min, Melt Curve 94 °C 15 s, 60 °C 60 s, 95 °C 30 s, 60 °C 15 s. Fluorescence from SYBR®Green was measured at the end of the elongation step, and obtained data were processed according to ΔΔCT method (Livak and Schmittgen 2001). cDNA preparations of the reference strains (*Y. lipolytica*

Po1h_Ura3+ and Po1g_Leu2+) were used as the external calibrators, to which 1.0 expression level was assigned. All the samples were analyzed in technical duplicates.

Data processing and statistical analysis

Concentration of the reporter proteins

The absolute concentration of enzymatic reporters was calculated based on their specific activity values, equal to 80 [GAU mg⁻¹] for TIG (Basaveswara Rao et al. 1981) and 478 [AAU mg⁻¹] for SoA (Baker and Woo 1985). The absolute concentration of eYFP reporter was calculated based on a calibration curve for eGFP provided by (BioTek Instruments, Inc., Winooski, VT, USA). For the purpose of re-calculation, the following data were used to estimate the relative brightness of eYFP with respect to eGFP: $\epsilon_{eYFP} = 83,400$ [cm⁻¹M⁻¹], $\Phi_{eYFP} = 0.61$, $\epsilon_{eGFP} = 55,000$ [cm⁻¹M⁻¹], $\Phi_{eGFP} = 0.60$ (Wall et al. 2015).

Simplified carbon and nitrogen distribution between the products

Distribution of C and N was calculated based on molecular weight and elementary composition of the major products synthesized by *Y. lipolytica* strains in the chemostat cultures. The following formulas and molecular weight values were used to calculate the content of carbon and nitrogen in the target protein: TIG: C₂₉₁₀H₄₃₆₅N₇₈₁O₉₀₉S₁₀, and 65,154.96 [Da]; SoA: C₂₂₉₉H₃₄₃₃N₆₂₉O₇₃₀S₁₈, and 52,140.40 [Da]; eYFP: C₁₂₁₅H₁₈₆₃N₃₁₇O₃₆₄S₈, and 26,991.54 [Da]. All these data were withdrawn from ExPASy server, using ProtParam tool (Gasteiger et al. 2005). Elementary composition of *Y. lipolytica* biomass was defined previously (Celińska et al. 2017a). A sum of C and N contained in all the products considered in this analysis was defined as 100%, and a fraction of each particular product was calculated based on determined concentration of DCW and the metabolites and the amount of the target protein, calculated as indicated above.

Statistical analysis

Statistical analysis was performed using Statistica 13 software (Tibco, CA, USA). After confirming homogeneity of variance, Tukey's HSD test was used to identify statistically homogenous groups of data at $p < 0.05$. Error bars indicate \pm SD of at least three replicates.

Results

Construction and selection of representative clones

Several positive sub-clones received after transformation of the parental strains were grown in shake-flasks to investigate inter-clonal variation and to identify possible growth impairment due to the genetic modification. As shown in Fig. S1, in the majority of cases, the observed variation in growth was negligible between the sub-clones and the reference strains (Po1h_Ura3+ or Po1g_Leu2+). In all the cases, it was possible to select a sub-clone being representative for each of the variants, demonstrating identical growth as the corresponding reference strain. These strains were subjected to the following studies.

Dilution rate adjustment

Prior to setting the bioreactor chemostat cultivations with the selected representative sub-clones, preliminary cultivations of the reference strain Po1h_Ura3+ were conducted to determine appropriate dilution rate of the culture. Three dilution rate settings were tested: 0.06, 0.12, 0.20 [h⁻¹]. Two parameters were used as determinants of the steady state: concentration of DCW [g_{DCW} L⁻¹] and concentration of GLY [g L⁻¹]. The results are shown in Fig. S2. At the lowest D , none of the determinants could be stabilized, demonstrating fluctuations from 56.37 to 75.53 and 14.40 to 23.85 [g L⁻¹] in GLY and DCW concentrations, respectively. Such high variation in DCW and GLY indicate high heterogeneity of the cultured cells population in terms of growth phases. Increasing D to 0.12 [h⁻¹] resulted in similar lack of stabilization of DCW, which fluctuated from 16.90 to 26.65 [g L⁻¹]. Accordingly, GLY concentration varied between 78.32 and 94.34 [g L⁻¹], which was still not satisfactory. The highest D tested, equal to 0.20 [h⁻¹], allowed to reach sufficiently stable readouts in both steady-state determinants, 88.71 to 91.05 of GLY [g L⁻¹] and 8.15 to 10.35 DCW [g_{DCW} L⁻¹]. Additionally, all the remaining metabolites were present in the culturing media at stable concentrations. Moreover, based on the results of this experiment it can be stated that oxygen availability was the growth-limiting factor. At D 0.20 [h⁻¹], the DO [%] was maintained near 20 [%] without extreme changes in the stirring rate. At dilution rates of 0.06 and 0.12 [h⁻¹], the readouts from the DO probe indicated complete oxygen depletion, making it impossible to precisely control DO in the bioreactor which might have triggered the lack of long-term stability of the cultures. The outcome indicates that at the D value of 0.2 [h⁻¹] the cultured population was homogenous in terms of growth phase and that the steady state was reached. Therefore, this value of D was used in the following chemostat cultivations of the selected representative *Y. lipolytica* sub-clones and the reference strains.

Chemostat cultivations—general observations

All the chemostat cultivations were conducted according to the same scheme—feeding of the medium and withdrawal of the culture were initiated immediately at the end of the batch stage (exhaustion of GLY marked by an increase in the DO [%] parameter). The steady state was reached typically after 6 to 9 residence times [nV]. The samples withdrawn from the cultures were analyzed for the concentration of the main carbon source, DCW and the key metabolites, viability of the cells, as well as the activity or fluorescence [AAU/GAU/RFU] resulting from expression of the respective heterologous genes. In Fig. 2, Table 1, and Table S4, these values were presented for all the strains maintained in the steady state, while Fig. S3 depicts kinetics of the whole cultures, including both the initial batch stage and the following chemostat stage. In the steady-state stage, GLY concentration was maintained at approx. 90 [g L⁻¹] ensuring satisfactory provision of the carbon source for intensive production of the heterologous proteins (Fig. 2a; Fig. S3). While generally, the differences in GLY [g L⁻¹] concentration were minor between the cultures of different strains, statistically important ($p < 0.05$) higher values were observed for Po1g- and Po1h-derived strains producing scTIG and Po1g-descendant producing inYFP, which was accompanied by a 2.82-, 2.71, and 1.96-fold lower glycerol consumption rate (Q_{GLY}) than in the corresponding control strains (Table 1). Po1g derivatives producing scTIG, inYFP, and, additionally, scSoA were also characterized by a 3.09-fold, 2.00-, and 1.70-fold lower specific GLY utilization rate than in control strain ($p < 0.05$). As observed, the Po1g derivatives, producing heterologous proteins, were in majority characterized by lower specific GLY consumption ($Y_{s/x} < 1.53$ [g g_{DCW}⁻¹]) and specific GLY consumption rate ($q_{GLY} < 0.20 \pm 0.03$ [g g_{DCW}⁻¹ h⁻¹]) than their Po1h counterparts (except Po1h_scYFP). Alongside, the steady-state biomass concentration in those Po1g derivatives was higher in than in their Po1h-derived counterparts (Fig. 2d; Table 1; $p < 0.05$). Indeed, the highest biomass concentration and production rate was observed for Po1g-descendants producing scSoA and scYFP ($p < 0.05$). Significantly lower biomass concentration [g_{DCW} L⁻¹] and production rate [g_{DCW} L⁻¹ h⁻¹] were observed for Po1h_scTIG (1.93- and 2.18-fold than the reference) and Po1h_inYFP (1.55- and 1.97-fold than the reference) ($p < 0.05$). Considering the absolute concentrations of the polyols produced by *Y. lipolytica* strains, their levels remained low (< 1.30 [g L⁻¹]; Fig. 2b, c; Table 1), but some of the strains differed significantly from the remaining in terms of the polyols production, e.g., the highest ERY [g L⁻¹] production was observed for Po1g_scSoA, while the lowest—for Po1h_scTIG, Po1h_inYFP, and Po1g_Leu2+ reference strain. As observed, the ERY concentrations were correlated with DCW values, as illustrated by highly corresponding specific ERY production rate (the highest for Po1h_scTIG 0.033 ±

0.002 [g g_{DCW}⁻¹ h⁻¹], and the lowest for Po1g_Leu2+ 0.015 ± 0.003 [g g_{DCW}⁻¹ h⁻¹]; all the remaining strains were grouped in the same homogenous group in HSD Tukey test at $p < 0.05$. MAN synthesis in the steady state was significantly different for two recombinant strains expressing TIG enzymatic reporter, for which the highest (Po1h_scTIG 0.99 ± 0.04 [g L⁻¹], 0.054 ± 0.001 [g g_{DCW}⁻¹ h⁻¹]) and the lowest (Po1g_scTIG 0.09 ± 0.08 [g L⁻¹], 0.003 ± 0.002 [g g_{DCW}⁻¹ h⁻¹]) MAN concentrations and specific production rates were observed. In this regard, Po1g_scTIG showed significantly different ratio of ERY to MAN formation reaching 12-fold higher production of the former ($p < 0.05$).

Production of the reporter proteins expressed as total AU or FU per culture volume is presented in Fig. 2e–h and Table 1. No enzymatic activity AAU or GAU was detected in the yeast biomass (done for randomly selected samples), but only in the supernatants of scTIG (3.52-fold higher [GAU L⁻¹] for Po1h than for Po1g derivatives) and scSoA (comparable [AAU L⁻¹] for Po1h than for Po1g derivatives) producing strains. Some minor fluorescence level was observed in the culture medium of the Po1g strain producing inYFP, which was not the case for Po1h-derivative-producing inYFP without the signal peptide. On the other hand, a very high level of intracellular fluorescence was observed for Po1h_scYFP strain, producing YFP equipped with an operable signal peptide, which was 5-fold higher than for Po1h_inYFP, and 10-fold higher than for Po1g_inYFP. Nevertheless, the extracellular fluorescence for this strain was also the highest from those observed in this study; giving more than 2.5-fold higher RFU than the one by po1g_scYFP. For the latter strain, also the intracellular fluorescence was relatively low. In terms of specific activity or specific fluorescence measures [(kRFU/AAU/GAU) g_{DCW}⁻¹], Po1h derivatives exhibited superiority over Po1g derivatives, ranging from 1.51-fold [AAU g_{DCW}⁻¹], through 2.62- and 4.59-fold [RFU g_{DCW}⁻¹] for inYFP and scYFP, up to 6.69-fold [GAU g_{DCW}⁻¹] higher specific measure for the former (Table 1). Similar trends were observed for the volumetric productivity [(kRFU/AAU/GAU) L⁻¹ h⁻¹] and the specific productivity [(kRFU/AAU/GAU) g_{DCW}⁻¹ h⁻¹] parameters (Table 1). The only example escaping this conclusion is the volumetric productivity of SoA which was comparable for Po1g- and Po1h-derived strains.

Production of the reporter proteins in the chemostat cultures—comparison of different reporters

In order to directly compare production of different reporter proteins studied here, it was necessary to unify expression of their amounts to a uniform unit. Therefore, based on available specific activities of SoA and TIG, as well as calibration curve, molar extinction coefficient, brightness, and quantum yield of eGFP and eYFP, the amounts of the reporter proteins were estimated and expressed in grams and moles (Table 2).

Table 1 The main bioprocessing parameters of the chemostat cultivations of *Yarrowia lipolytica*-derivatives producing the reporter proteins (in YFP/scYFP/scSoA/scTIG)

Strains	Q_{GLY} [$g L^{-1} h^{-1}$]		q_{GLY} [$g g_{DCW}^{-1} h^{-1}$]		Q_{DCW} [$g L^{-1} h^{-1}$]		$Y_{s,x}$ [$g g_{DCW}^{-1}$]		$Y_{protein,x}$ [(kRFU/AAU/GAU) g_{DCW}^{-1}]		$Q_{protein}$ [(kRFU/AAU/GAU) $L^{-1} h^{-1}$]		$q_{protein}$ [(kRFU/AAU/GAU) $g_{DCW}^{-1} h^{-1}$]	
	Mean	$\pm SD$	Mean	$\pm SD$	Mean	$\pm SD$	Mean	$\pm SD$	Mean	$\pm SD$	Mean	$\pm SD$	Mean	$\pm SD$
Polg_Leu2+	2.14	0.35	0.34	0.05	1.43	0.03	1.50	0.23						
Polh_Ura3+	3.28	0.07	0.44	0.01	1.81	0.03	1.86	0.11						
Polg_inYFP	1.09*	0.52*	0.17*	0.07*	1.14*	0.09*	0.94	0.38	10,561.29 [kRFU g_{DCW}^{-1}]	1997.53	13,708.69 [kRFU $L^{-1} h^{-1}$]	3359.42	2119.68 [kRFU $g_{DCW}^{-1} h^{-1}$]	370.76
Polh_inYFP	1.95*	0.41*	0.40	0.08	0.92*	0.07*	2.13	0.43	27,680.92 [kRFU g_{DCW}^{-1}]	2893.94	24,343.02 [kRFU $L^{-1} h^{-1}$]	4520.57	5163.48 [kRFU $g_{DCW}^{-1} h^{-1}$]	541.18
Polg_scYFP	2.88	0.08	0.30	0.02	1.88*	0.07*	1.53	0.08	1390.96 [kRFU g_{DCW}^{-1}]	102.33	5286.58 [kRFU $L^{-1} h^{-1}$]	220.50	540.81 [kRFU $g_{DCW}^{-1} h^{-1}$]	35.49
Polh_scYFP	1.86*	0.32*	0.24*	0.03*	1.52*	0.16*	1.27	0.06	6387.61 [kRFU g_{DCW}^{-1}]	377.80	9440.37 [kRFU $L^{-1} h^{-1}$]	540.61	1220.64 [kRFU $g_{DCW}^{-1} h^{-1}$]	58.86
Polg_scSoA	2.18	0.39	0.20	0.03	2.24*	0.13*	0.97	0.16	71.58 [AAU g_{DCW}^{-1}]	4.15	162.07 [AAU $L^{-1} h^{-1}$]	6.06	14.9 [AAU $g_{DCW}^{-1} h^{-1}$]	1.24
Polh_scSoA	2.49	0.53	0.32	0.05	1.44*	0.07*	1.73	0.04	108.64 [AAU g_{DCW}^{-1}]	3.63	155.64 [AAU $L^{-1} h^{-1}$]	0.91	19.92 [AAU $g_{DCW}^{-1} h^{-1}$]	0.54
Polg_scTIG	0.76*	0.00*	0.11*	0.01*	1.65	0.10	0.47*	0.03*	22.73 [GAU g_{DCW}^{-1}]	5.65	57.99 [GAU $L^{-1} h^{-1}$]	5.94	7.96 [GAU $g_{DCW}^{-1} h^{-1}$]	1.06
Polh_scTIG	1.21*	0.22*	0.31	0.06	0.83*	0.04*	1.47	0.29	152.02 [GAU g_{DCW}^{-1}]	17.32	119.78 [GAU $L^{-1} h^{-1}$]	18.24	32.08 [GAU $g_{DCW}^{-1} h^{-1}$]	4.07

$\pm SD$, standard deviation from at least two biological replicates and a specified number of technical replicates

* $p < 0.05$, significantly different from the references

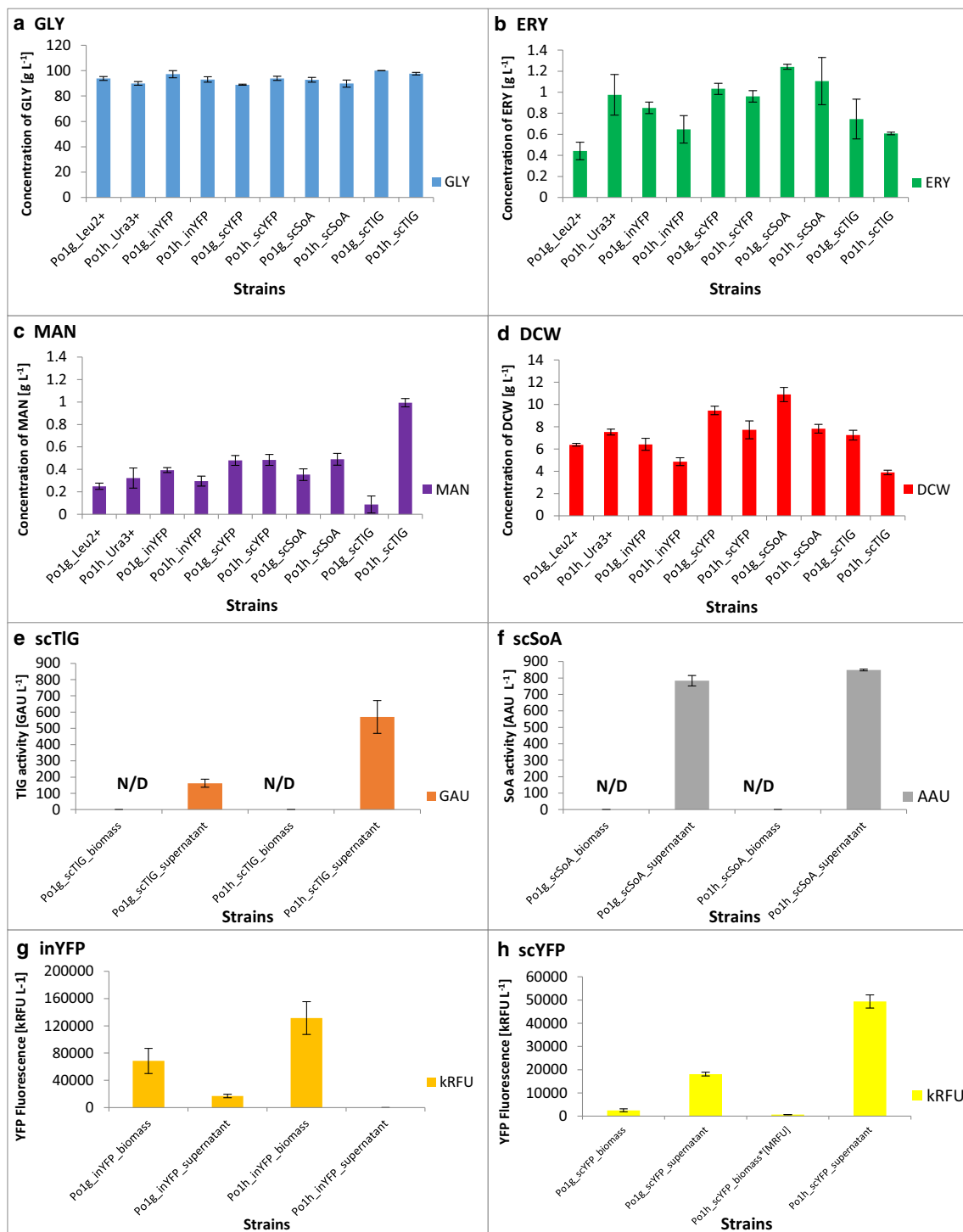


Fig. 2 Glycerol consumption, concentration of major small molecular weight metabolites, biomass, secretory protein production (scTIG/scSoA), and yellow fluorescence (inYFP/scYFP) during steady-state chemostat cultivations of Po1g- and Po1h-derived *Yarrowia lipolytica* strains. *x*-axis: strains. *y*-axis: **a** concentration of glycerol (GLY) in [g L⁻¹], **b** concentration of erythritol (ERY) in [g L⁻¹], **c** concentration of mannitol (MAN) in [g L⁻¹], **d** biomass accumulation: dry cell weight

(DCW) in [g_{DCW} L⁻¹], **e** TIG activity in culture medium in [GAU L⁻¹], **f** SoA activity in culture medium in [AAU L⁻¹], **g** YFP fluorescence of Po1g_inYFP/Po1h_inYFP recombinant strains in the cellular pellet and culture medium in [kRFU], **h** YFP fluorescence of Po1g_scYFP/Po1h_scYFP recombinant strains in the cellular pellet and culture medium in [kRFU]. Error bars indicate \pm SD from at least biological duplicates in two technical replicates. N/D, non-detected

Noteworthy, it needs to be stressed that the following measures are just an approximation, and accurate comparisons can

be done only for the variants with the same reporter protein (inYFP and scYFP together, separately SoA, and separately

Table 2 Estimated amounts and related production parameters of the reporter proteins (inYFP/scYFP/scSoA/scTIG) in continuous cultures of *Yarrowia lipolytica* derivatives

Strains	Protein [$\mu\text{g L}^{-1}$]		Protein [pmol L^{-1}]		Q_{Protein} [$\mu\text{g L}^{-1} \text{h}^{-1}$]		Q_{Protein} [$\text{pmol L}^{-1} \text{h}^{-1}$]		Y_{Protx} [$\mu\text{g gDCW}^{-1}$]		Y_{Protx} [pmol g^{-1}]		q_{Protx} [$\mu\text{g g}^{-1} \text{h}^{-1}$]		q_{Protx} [$\text{pmol g}^{-1} \text{h}^{-1}$]	
	Mean	$\pm\text{SD}$	Mean	$\pm\text{SD}$	Mean	$\pm\text{SD}$	Mean	$\pm\text{SD}$	Mean	$\pm\text{SD}$	Mean	$\pm\text{SD}$	Mean	$\pm\text{SD}$	Mean	$\pm\text{SD}$
Po1g_inYFP	375.93	105.14	13,927.33	3895.70	66.54	18.61	2465.00	689.59	58.06	12.19	2150.95	451.69	10.28	2.16	380.72	79.95
Po1h_inYFP	722.43	135.59	26,766.00	5024.37	136.54	25.63	5058.67	949.87	147.69	17.52	5471.85	649.13	27.91	3.31	1034.18	122.69
Po1g_scYFP	99.43	6.68	3684.67	248.47	19.79	1.33	733.33	49.33	10.53	0.94	390.00	34.74	2.10	0.18	77.61	6.92
Po1h_scYFP	271.30	18.37	10,051.67	680.18	53.17	3.60	1970.00	133.59	35.18	1.71	1303.27	63.39	6.90	0.33	255.44	12.43
Po1g_scSoA	28.87	1.19	553.00	22.72	5.92	0.24	113.33	4.73	2.65	0.23	50.90	4.50	0.54	0.05	10.44	0.92
Po1h_scSoA	31.23	0.15	599.33	3.51	5.75	0.03	110.33	0.58	4.00	0.21	76.68	4.02	0.74	0.04	14.11	0.74
Po1g_scTIG	33.77	5.15	518.00	78.94	7.67	1.17	118.00	17.69	4.69	1.01	72.02	15.48	1.06	0.23	16.35	3.52
Po1h_scTIG	118.83	21.03	1824.00	323.01	25.19	4.46	386.67	68.12	30.37	4.10	466.16	62.98	6.44	0.87	98.83	13.35

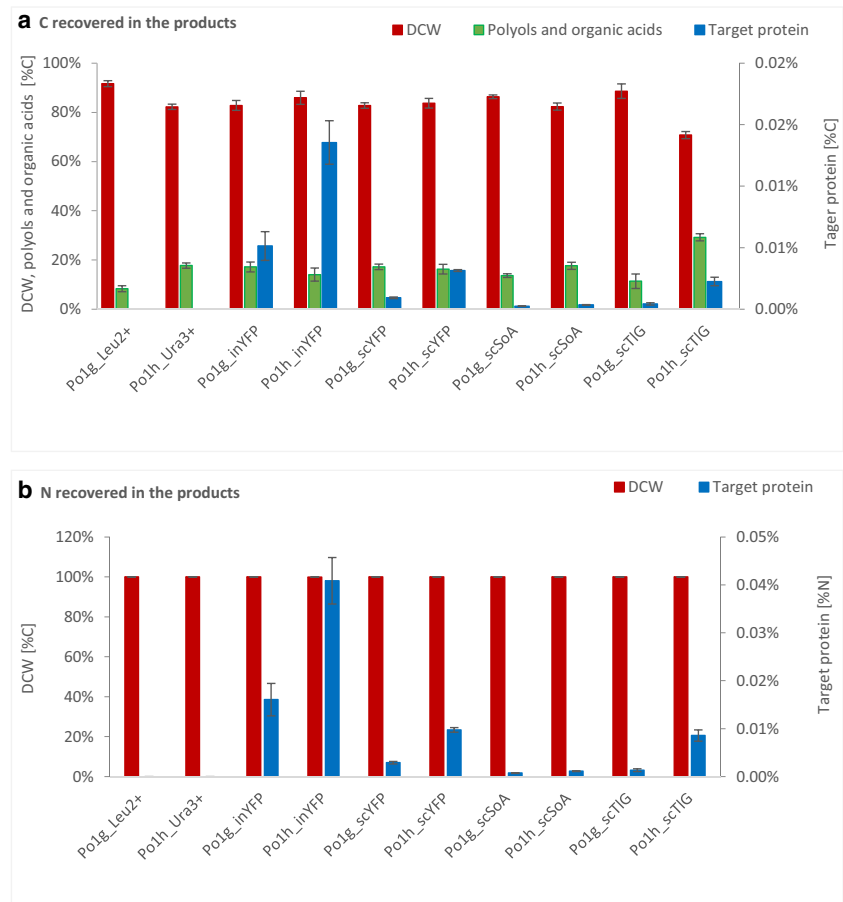
$\pm\text{SD}$, standard deviation from at least two biological replicates and a specified number of technical replicates

TIG). Irrespective of the adopted measure ($[\mu\text{g L}^{-1}]$, $[\text{pmol}]$, $[\mu\text{g gDCW}^{-1}]$, $[\text{pmol g}^{-1} \text{h}^{-1}]$, etc.), Po1h_inYFP was superior in terms of the amount of produced reporter protein ($p < 0.05$), immediately followed by Po1g_inYFP. It was also reflected by the % of carbon (C) and nitrogen (N) channeled to the specific products synthesized by the *Y. lipolytica* strains under steady state (Fig. 3). In terms of the protein amounts expressed in molar values— $[\text{pmol}]$ or $[\text{pmol h}^{-1}]$ or in the amount of C/N channeled to these products, it was observed that Po1h_scYFP produced significantly higher amounts of the reporter than any other strain producing secretory reporters ($p < 0.05$). Interestingly, in terms of the absolute molar amount $[\text{pmol}]$ and volumetric productivity in moles $[\text{pmol h}^{-1}]$, a clear separation between the efficiency in production of the small reporter proteins (inYFP, scYFP) from the larger secretory reporters (SoA, TIG) was observed ($p < 0.05$). In the majority of measures, production of SoA by either of the host strains was the lowest; and usually, these two strains, Po1h_scSoA and Po1g_scSoA, were classified as the least efficient. Nevertheless, since all of the strains giving the highest/lowest absolute amounts of the reporter proteins synthesize YFP/SoA, it cannot be excluded that this observation results from the mode of the proteins amounts quantitation (calibration curve for YFP and specific activity for SoA and TIG). Only purification of the reporter proteins, experimental determination of their activity/fluorescence and amount in all the samples could provide accurate data in this regard, which is beyond the scope of this work. In every parameter and for all the reporter proteins studied here, Po1h turned out to be a superior host for the protein reporter production compared with Po1g ($p < 0.05$). This difference was the least clear for SoA.

Gene expression

To verify if the differences in the amounts of the produced reporter proteins result from the differences in the transcription efficiency, we analyzed the expression level of genes encoding the reporter proteins in the steady-state-maintained cells (Fig. 4). As observed, the expression level of SoA-encoding gene was the highest, irrespective of the host strain ($> 4.0 \log_{10}\text{RQ}$). No significant effect of the expression system (4UAS-pTEF in Po1h derivatives or hp4d for Po1g-derived strains) could be observed ($p < 0.05$), although each time expression of a respective gene was at least slightly higher in the Po1h host. Significantly higher expression from 4UAS-pTEF was observed for scTIG and scYFP, while comparable expression level from both promoters was observed for SoA and inYFP, which corresponds with the data on the absolute amounts of the proteins based on their reporting activity (Fig. 2e–h) and the calculated amounts of the proteins (Table 2).

Fig. 3 Simplified carbon and nitrogen distribution between the products. C and N amounts were determined based on either elementary analysis or chemical formulas. Total of C and N in all the considered products was defined as 100%. Error bars indicate \pm SD from at least biological duplicates in two technical replicates



Discussion

In this study, we used chemostat cultivations as a tool to characterize heterologous protein production in *Y. lipolytica* expressing four protein-encoding genes. Two of the most widely

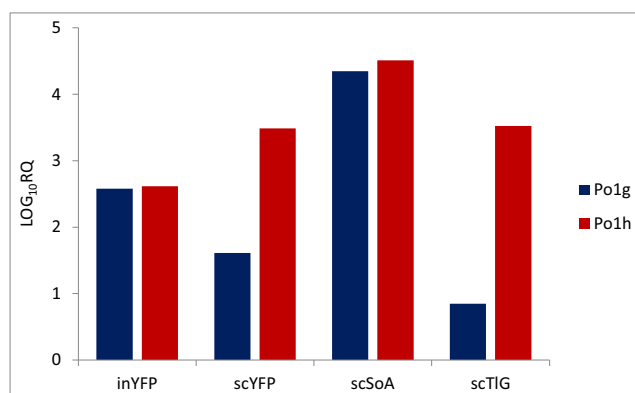


Fig. 4 Relative expression levels of the heterologous genes encoding the reporter proteins in steady-state chemostat cultivations of *Y. lipolytica* recombinant strains. *x*-axis, reporter genes (scSoA, scTIG, inYFP, and scYFP); *y*-axis, LOG₁₀RQ - Relative Quantitation values obtained in RTqPCR transformed logarithmically for more convenient presentation. Red: Po1h-derived *Y. lipolytica* strains. Blue: Po1g-derived *Y. lipolytica* strains. Transcript levels were determined by RTqPCR, from at least biological duplicates in two technical replicates

studied and exploited *Y. lipolytica* expression platforms were used (Po1h- and Po1g-based), differing in the auxotrophy selection marker, integration site of the recombinant cassette, the promoter-governing expression of the cloned genes, and the signal peptide governing the proteins secretion (spYALI0B03564g and spXPR2pre, respectively). While indeed, the type of used auxotrophy selection marker (i.e., Leu2) was shown to impact lipids accumulation level in *Y. lipolytica* (Blazek et al. 2014; Kerkhoven et al. 2017), no such findings were reported in terms of protein overproduction, which was the key response studied here. On the other hand, the key uncertainty of Po1h and Po1g comparison relies on the mode of the cassette integration in the genome. While the former expression system is dedicated for single-copy integrations (pBR-platform; Fig. 1) the latter one relies on random integrations using zeta elements in zeta-less strain (Pignede et al. 2000). In both cases, it was expected to obtain low-copy integrants, as non-defective selection markers were used (Le Dall et al. 1994). To account for any differences resulting from the number of the cassettes integrated with the host genome or the site of integration, several sub-clones of each type were pre-screened for growth and production of the reporters. Such strategy was previously shown to assure identification of low-copy integrants, bearing the cassette integrated at a neutral site

in the genome of *Y. lipolytica* (Vogl et al. 2018; Holkenbrink et al. 2018; Theron et al. 2020). The sub-clones selected for further studies did not differ importantly in growth rate (Fig. S1), which suggests low variability in this regard coming from selection of low-copy integrants, bearing the cassette integrated at neutral site in the genome. Nevertheless, although not very probable, integration of some additional copies of the randomly integrated GGA in the Po1h genome cannot be excluded.

The stability of developed steady state in the chemostat cultures was monitored through online and off-line measurements, including DO, pH, concentration of DCW, GLY, and the major metabolites. As discussed in Arvas et al. (2011) and Liu et al. (2013), under carbon limitation, the cell growth rate, modulated by D [h^{-1}] parameter, strongly impacts the overall physiology of a cell. In the present study, only at the highest adopted D 0.20 [h^{-1}], the system exhibited satisfactory chemo-stability; although, our cultures were conducted without carbon limitation, which was the case in the referenced studies. Such D seems to be very high, when compared with the other reports on continuous cultures of *Y. lipolytica*—0.01 [h^{-1}] (Rakicka et al. 2017), 0.009–0.031 [h^{-1}] (Rywińska et al. 2011), and 0.08 [h^{-1}] (Ochoa-Estopier and Guillouet 2014). However, it was evidenced that expression of the genes related to the protein synthesis is positively correlated with growth rate in *S. cerevisiae* (Castrillo et al. 2007; Liu et al. 2013) and *K. phaffii* (Buchetics et al. 2011). On the other hand, high growth rate was detrimental for production of laccase in *Y. lipolytica* (Madzak et al. 2005). In contrast, it was shown that the protein secretion process is coupled to specific growth rate in yeast, being higher at a higher growth rate. Molecular events underlying this phenomenon are endoplasmic reticulum processing, protein turnover, cell cycle, and global stress response (Liu et al. 2013). Similarly, transcriptional activity was shown to be positively correlated with increasing dilution rate in *S. cerevisiae* under carbon-limitation (Liu et al. 2013), but this observation was gene-dependent. Moreover, in *S. cerevisiae*, at lower D (0.05–0.10 [h^{-1}]) stress-related phenomena were shown to play an important role, disabling accurate insight into the processes related to production of the heterologous proteins, which, on the other hand, played the main role at $D > 0.10$ –0.20 [h^{-1}] (Liu et al. 2013).

Typically, cultivations under steady state enable accurate tracking of C and N distribution between different products. While here-conducted tracking and recovery of C and N contained in the products was greatly simplified (compared with what can be achieved using C13 labelled substrate), it allowed to make some observations in this regard. Po1h_scTIG strain channeled a relatively high fraction of C to polyols at a loss in DCW, and, unlike the other strains, it produced more MAN than ERY. This strain, and the other Po1h-derivative-producing inYFP, channeled relatively high fraction of C and N to the heterologous proteins. In general,

steady state of Po1g derivatives was accompanied with high accumulation of the biomass (DCW), in comparison with Po1h-derived counterparts. On the other hand, in the majority of cases, Po1h derivatives utilized higher amounts of GLY per biomass unit and uniformly produced higher amounts of the target proteins, based on their reporting activity per volume [(AAU/GAU/RFU) L^{-1}]. Such an outcome could suggest that higher C flux was channeled to synthesis of the target proteins rather than biomass in Po1h derivatives. Hence, it was tempting to state that the increased specific substrate consumption and higher amount of the produced active reporter results from higher transcription efficiency of 4UAS-pTEF (Po1h) over hp4d (Po1g); although, a potential effect of the other factors differing between the two expression platforms cannot be excluded. Considering direct comparison of C and N fractions recovered in the different reporter proteins studied here, the highest amount of the elements was found in inYFP, followed by scYFP, TIG, and SoA at the end. While these differences most probably result from the mode of the proteins' absolute amount estimation, these results well correspond with biochemical characteristics of the protein, given in Table S2. The low absolute amounts of SoA may result from high number of Cys (11), and the resulting high number of possible S–S combinations (35696; Table S2), which is the highest among the studied protein reporters. It is thus plausible, that synthesis of SoA protein is limited at disulfide bond formation in the ER, which is a stochastic process, highly demanding in terms of energy and building blocks. Limitation at this stage, could also account for the least clear difference in the extracellular SoA activity, seen between the two hosts, as ER-resident S-S forming process was the same for the two platforms, and its contribution could be higher than by the other factors. For the same reasons, i.e., only 2 Cys residues and a single predicted glycosylation site, the amounts of scYFP and inYFP reporter were the highest in terms of recovered C and N. TIG reporter also bears relatively high number of Cys residues (8; 764 possible combinations), but it is additionally highly glycosylated, which could potentially further limit its production (predicted glycosylation sites: 3 N-glycosylation, 18 O-glycosylation). On the other hand, glycosylation is a targeted process, relying on specific motives within the polypeptide structure, and thus could be less energy-consuming. Based on current results, it can be speculated that disulfide bonds formation is the most limiting step in *Y. lipolytica* host, as it strongly decreased the absolute amount of the SoA reporter, being less glycosylated and having lower molecular weight than TIG (52 vs. 65 kDa).

To get an insight into the background behind the observed lower biomass accumulation and higher production of the reporter proteins per volume in Po1h-derivative strains, expression level of the reporter-encoding genes was analyzed. Both promoters used in this study are preceded by 4 upstream activating sequences (UAS1B) (Madzak et al. 1999) which

were repeatedly shown to strongly enhance transcription rate of the following gene (Blazeck et al. 2011; Blazeck et al. 2013; Dulermo et al. 2017). Upon direct comparison (Dulermo et al. 2017), 4UAS-pTEF was more transcriptionally active than hp4d, as also the core pTEF promoter is more transcriptionally active than the minimal Leu2 promoter, which is the core of hp4d (Blazeck et al. 2012). In the present study, the expression of all the heterologous genes was at least slightly higher from 4UAS-pTEF than hp4d, which corresponds to the previous literature reports and our findings on the heterologous protein production.

Linear relationships between mRNA level and the protein product are rarely seen for the secretory reporters. Such relationship was found for a relatively small CelB protein (55 [kDa]) produced in *Y. lipolytica* but was not valid when a secretory protein M1 with a higher molecular weight (120 [kDa]) was expressed in the same expression system (Swietalski et al. 2020). In the latter case, the mRNA level was disproportionately high when compared with the protein amount indicating insufficient capacity at the translation level and possible bottlenecks in folding and/or secretion. On the other hand, it was observed that highly transcribed and translated CelB could not be efficiently secreted due to overloading of the secretory pathway, as high proportion of the CelB activity was detected inside the host cells. Positive relationship between expression level and protein production and activity was identified for a panel of genes encoding both intracellular and secreted polypeptides in *Y. lipolytica*, i.e., RedStar2, glucoamylase, YFP, and α -amylase, but such linearity was unexpectedly not observed for the other secretory reporter XlnC (Dulermo et al. 2017). Consistently, production of an insulin precursor (IP) and of an amylase in *S. cerevisiae* was lower under the TEF1 promoter than under the TPI promoter even though their transcription was greater from the former (Liu et al. 2013). In all these cases, it was concluded that excessive protein production negatively affects protein folding because of the titration of chaperons and the saturation of the secretory pathway. Considering the present study, such a type of disproportion between the transcript levels and the estimated amount of the protein product was observed for the SoA reporter. Irrespective of the host strain (and the promoter used), SoA transcripts were the most abundant based on their relative quantitation. Correspondingly, no differences were observed in the amount of extracellular SoA activity between the hosts, indicating that the data corroborate themselves mutually. No intracellular activity was detected in randomly selected samples, but it is probable that SoA is active only if secreted, and cannot be detected when retained inside the cells, as found previously (Celińska et al. 2016). However, when “translated into” the absolute measures of the protein (expressed in grams and moles), these amounts were unexpectedly low. This may result from erroneous definition of a specific activity value used in the calculation, taken from a

paper on native insect enzyme (Baker and Woo 1985), and/or differences in the specific activity of the native and recombinant protein. On the other hand, similar outcome was observed upon overproduction of HSA protein in *K. phaffii*, where sub-clones secreting the protein at low rate expressed the genes at significantly higher levels than the high-secreting sub-clones (Aw et al. 2017). In that study, it was evidenced that the cells overproducing secretory proteins primarily suffer from starvation, due to increased demand for nutrients upon excessive production of the heterologous protein. However, these studies were conducted in batch cultivations, where limited provision of energy and building blocks for the transcriptional and translational machinery could account for the observed lack of linearity in mRNA:protein relationship. Considering the here-adopted mode of the strains maintenance (high nutrients provision in steady state), it seems unlikely that starvation was the reason for limited production of SoA upon its efficient transcription. Aw et al. (2017) also demonstrated that high secretors exhibited decreased viability, which was also observed for the best producer and secretor (Po1h_scYFP) in this study. Furthermore, based on Po1h_scYFP example (high expression level, \log_{10} RQ of 3.49 vs. 4.4 for SoA, and high estimated protein production, 271.3 [$\mu\text{g L}^{-1}$] vs. approx. 31 [$\mu\text{g L}^{-1}$] for SoA), it seems that still sufficient capacity in translational-secretory machinery and abundance of nutrients remained available for *Y. lipolytica* host cells.

Previously, high linearity between mRNA quantity and a mean fluorescence was reported for intracellularly localized small proteins (GFP and RedStar2) expressed under a wide range of promoters of highly different strength (Blazeck et al. 2011; Dulermo et al. 2017). In this study, the levels of transcription and intracellular fluorescence were highly corresponding in both inYFP-producing strains (\log_{10} RQ of approx. 2.6; total [kRFU] in Po1g_inYFP, 85.278 vs. 131.509 [kRFU L^{-1}] in Po1h_inYFP; [RFU] for the biomass and the supernatant is given as total). Still, as in the case of SoA, it was surprising that the expression from the weaker promoter (hp4d) is comparable with the expression governed by the stronger promoter (4UAS-pTEF). On the other hand, the expression of scYFP differed significantly, depending on the promoter type and the host strain. Higher expression of scYFP from 4UAS-pTEF was accompanied by a very high production of the protein (total in [kRFU] in the supernatant and the biomass, 722.386 vs. 20.615 for Po1h_ and Po1g_scYFP, respectively). The lower level of scYFP expression and the protein production by Po1g derivative was accompanied by expected ratio between the scYFP:inYFP (7.2), indicating efficient secretion of the reporter equipped with spXPR2, and only minor retention inside. On the other hand, the extremely high production of YFP in Po1h_scYFP was accompanied by very high retention of the YFP intended for secretion (674.000 [kRFU] in the biomass vs. 49.386 [kRFU] in the supernatant; the scYFP:inYFP ratio of 0.07). Surprisingly,

intracellular accumulation of YFP in Po1h_scYFP was 5-fold higher than in Po1h_inYFP designed to achieve high intracellular [RFU] (no extracellular YFP detected). Nevertheless, it seems somehow reasonable that the strains with high secretion of a recombinant protein have high intracellular flux of this protein and thus high inFL signals than low producers. Similar observations were reported for CelB cloned in *Y. lipolytica*, which was also intended for secretion but a high proportion of the activity was detected inside the host cells (Swietalski et al. 2020). It was concluded that this high retention of CelB resulted from overloading of the secretory pathway. Corresponding conclusion could explain here observed high intracellular [RFU] of Po1h_scYFP. In our recent studies, we also observed relatively high intracellular accumulation of the fluorescent reporter targeted for secretion upon high overproduction (Gorczyca et al. 2020). On the other hand, (Theron et al. 2020) observed 7-fold lower intracellular fluorescence in the *Y. lipolytica* strains expressing GFP equipped with a signal peptide, as compared with that of strain expression intracellular variant of the fluorophore.

Still it remains puzzling what is the reason for the differences in the expression level of the genes cloned under the same promoter, or lack of significant differences in the transcripts RQ expressed from different promoters (as for SoA and inYFP). It was also observed previously that even if different genes were embedded in exactly the same expression cassettes, platforms and genetic environment, their expression level differed significantly. Such an observation was done for genes cloned in *Y. lipolytica* encoding CelB and M1 galactosidases (Swietalski et al. 2020) (3-fold higher expression level for M1 gene, at significantly lower extracellular activity), or for IP and amylase cloned in *S. cerevisiae*, where significantly higher relative transcript levels were observed for the former although both were cloned according to the same strategy (Liu et al. 2013). Nevertheless, no explanation has been postulated. It can be only speculated that either mRNA stability of the two transcripts is highly different or that transcription is somehow reversibly regulated by the translation efficiency and/or the traffic in the secretory pathway. Considering our current results, depending on the target protein characteristics, the metabolic burden imposed on the producer cell will differ (gene-to-gene variation), but also, different promoters would be differently susceptible to such feedback regulation (promoter-to-promoter variation). Definitely, more in-depth studies are required to answer these questions.

As previously postulated, depending on the characteristics of the overproduced heterologous protein, the response from the host cell in terms of its physiology and the polypeptide production rate is different. In the present study, we used steady-state maintained *Y. lipolytica* cells to investigate the impact of different heterologous proteins on the physiological behavior of the host cells. Such approach allowed to uncouple the impact of a particular protein overproduction from phenomena resulting from growth phase or caused by heterogeneity of the analyzed

population. The here obtained data suggest that, using the more transcriptionally active promoter results in channeling more C flux towards the target protein, giving significantly higher specific amounts and production rates of the target polypeptide, at a loss of the biomass accumulation, with no significant impact on the polyols production. Depending on the target protein traits and the promoter governing its expression, specific responses in terms of the protein formation kinetics and the gene transcription level were observed. The same was observed previously for *S. cerevisiae* or *K. phaffii*, for which no generalizable conclusions could be withdrawn, indicating strong impact of the individual combination of the target protein, promoter and some other, unidentified factors, on the host strain response. In the present study, some relationships between complexity of the reporter protein's post-translation modifications and the absolute amounts of produced protein could be seen. Based on our current results, it can be suggested that the disulfide bonds formation in the ER is the most limiting step in *Y. lipolytica* protein production platform.

Authors' contributions PKW conducted all the laboratory work, processed the data and prepared them for publication, secured financial support, and wrote the manuscript under supervision of EC; PK and WB served with their expertise in continuous cultures, conducted data processing and statistical analyses; EC conceived the concept of the study and wrote and revised the manuscript.

Funding This study was funded by the Ministry of Science and Higher Education in Poland, Project DI 2017 000947 and partially also by No. IP2015 011074 (MSHE).

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests Ethics approval N/A

Consent to participate N/A

Consent for publication All the authors express consent for publication of the manuscript.

Availability of data and material All data generated or analyzed during this study are included in this published article [and its electronic supplementary material]

Code availability N/A

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Arvas M, Pakula T, Smit B, Rautio J, Koivistoinen H, Jouhten P, Lindfors E, Wiebe M, Penttilä M, Saloheimo M (2011) Correlation of gene expression and protein production rate - a system wide study. *BMC Genomics* 12:616. <https://doi.org/10.1186/1471-2164-12-616>
- Aw R, Barton GR, Leak DJ (2017) Insights into the prevalence and underlying causes of clonal variation through transcriptomic analysis in *Pichia pastoris*. *Appl Microbiol Biotechnol* 101:5045–5058. <https://doi.org/10.1007/s00253-017-8317-2>
- Baker JE, Woo SM (1985) Purification, partial characterization, and postembryonic levels of amylases from *Sitophilus oryzae* and *Sitophilus granarius*. *Arch Insect Biochem Physiol* 2:415–428. <https://doi.org/10.1002/arch.940020409>
- Barth G, Gaillardin C (1996) *Yarrowia lipolytica*. In: Wolf K (ed) Nonconventional yeasts in biotechnology: a handbook. Springer, Berlin, pp 313–388
- Basaveswara Rao V, Sastri NVS, Subba Rao PV (1981) Purification and characterization of a thermostable glucoamylase from the thermophilic fungus *Thermomyces lanuginosus*. *Biochem J* 193:379–387. <https://doi.org/10.1042/bj1930379>
- Blazcek J, Liu L, Redden H, Alper H (2011) Tuning gene expression in *Yarrowia lipolytica* by a hybrid promoter approach. *Appl Environ Microbiol* 77:7905–7914. <https://doi.org/10.1128/AEM.05763-11>
- Blazcek J, Garg R, Reed B, Alper HS (2012) Controlling promoter strength and regulation in *Saccharomyces cerevisiae* using synthetic hybrid promoters. *Biotechnol Bioeng* 109:2884–2895. <https://doi.org/10.1002/bit.24552>
- Blazcek J, Reed B, Garg R, Gerstner R, Pan A, Agarwala V, Alper HS (2013) Generalizing a hybrid synthetic promoter approach in *Yarrowia lipolytica*. *Appl Microbiol Biotechnol* 97:3037–3052. <https://doi.org/10.1007/s00253-012-4421-5>
- Blazcek J, Hill A, Liu L, Knight R, Miller J, Pan A, Otoupal P, Alper HS (2014) Harnessing *Yarrowia lipolytica* lipogenesis to create a platform for lipid and biofuel production. *Nat Commun*. <https://doi.org/10.1038/ncomms4131>
- Boonvitthya N, Bozonnet S, Burapatana V, O'Donohue MJ, Chulalaksananukul W (2013) Comparison of the heterologous expression of *Trichoderma reesei* endoglucanase II and cellobiohydrolase II in the yeasts *Pichia pastoris* and *Yarrowia lipolytica*. *Mol Biotechnol* 54:158–169. <https://doi.org/10.1007/s12033-012-9557-0>
- Borkowska M, Białas W, Kubiak M, Celińska E (2019) Rapid microassays for amylolytic activities determination: customization and validation of the tests. *Appl Microbiol Biotechnol* 103:2367–2379. <https://doi.org/10.1007/s00253-018-09610-0>
- Buchetics M, Dragosits M, Maurer M, Rebnegger C, Porro D, Sauer M, Gasser B, Mattanovich D (2011) Reverse engineering of protein secretion by uncoupling of cell cycle phases from growth. *Biotechnol Bioeng* 108:2403–2412. <https://doi.org/10.1002/bit.23198>
- Castrillo JI, Zeef LA, Hoyle DC, Zhang N, Hayes A, Gardner DCJ, Cornell MJ, Petty J, Hakes L, Wardleworth L, Rash B, Brown M, Dunn WB, Broadhurst D, O'Donoghue K, Hester SS, Dunkley TPJ, Hart SR, Swainston N, Li P, Gaskell SJ, Paton NW, Lilley KS, Kell DB, Oliver SG (2007) Growth control of the eukaryote cell: A systems biology study in yeast. *J Biol*. <https://doi.org/10.1186/jbiol54>
- Celińska E, Nicaud JM (2019) Filamentous fungi-like secretory pathway strayed in a yeast system: peculiarities of *Yarrowia lipolytica* secretory pathway underlying its extraordinary performance. *Appl Microbiol Biotechnol* 103:39–52. <https://doi.org/10.1007/s00253-018-9450-2>
- Celińska E, Borkowska M, Białas W (2016) Evaluation of heterologous α -amylase production in two expression platforms dedicated for *Yarrowia lipolytica*: commercial Polg-pYLSC (php4d) and custom-made A18-pYLTEF (pTEF). *Yeast* 33:165–181. <https://doi.org/10.1002/yea.3149>
- Celińska E, Borkowska M, Białas W (2017a) Enhanced production of insect raw-starch-digesting alpha-amylase accompanied by high erythritol synthesis in recombinant *Yarrowia lipolytica* fed-batch cultures at high-cell-densities. *Process Biochem* 52:78–85. <https://doi.org/10.1016/j.procbio.2016.10.022>
- Celińska E, Ledesma-Amaro R, Larroude M, Rossignol T, Pauthenier C, Nicaud JM (2017b) Golden Gate assembly system dedicated to complex pathway manipulation in *Yarrowia lipolytica*. *Microb Biotechnol* 10:450–455. <https://doi.org/10.1111/1751-7915.12605>
- Celińska E, Borkowska M, Białas W, Korpys P, Nicaud JM (2018) Robust signal peptides for protein secretion in *Yarrowia lipolytica*: identification and characterization of novel secretory tags. *Appl Microbiol Biotechnol* 102:5221–5233. <https://doi.org/10.1007/s00253-018-8966-9>
- Dulermo R, Brunel F, Dulermo T, Ledesma-Amaro R, Vion J, Trassaert M, Thomas S, Nicaud JM, Leplat C (2017) Using a vector pool containing variable-strength promoters to optimize protein production in *Yarrowia lipolytica*. *Microb Cell Factories*. <https://doi.org/10.1186/s12934-017-0647-3>
- Gasser B, Sauer M, Maurer M, Stadlmayr G, Mattanovich D (2007) Transcriptomics-based identification of novel factors enhancing heterologous protein secretion in yeasts. *Appl Environ Microbiol* 73:6499–6507. <https://doi.org/10.1128/AEM.01196-07>
- Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed) The proteomics protocols handbook. Humana Press Inc., pp 571–608
- Goncalves C, Rodriguez-Jasso RM, Gomes N, Teixeira JA, Belo I (2010) Adaptation of dinitrosalicylic acid method to microtiter plates. *Anal Methods* 2:2046–2048. <https://doi.org/10.1039/c0ay00525h>
- Gorczyca M, Kaźmierczak J, Steels S, Fickers P, Celińska E (2020) Impact of oxygen availability on heterologous gene expression and polypeptide secretion dynamics in *Yarrowia lipolytica*-based protein production platforms. *Yeast*. <https://doi.org/10.1002/yea.3499>
- Graf A, Dragosits M, Gasser B, Mattanovich D (2009) Yeast systems biotechnology for the production of heterologous proteins. *FEMS Yeast Res* 9:335–348. <https://doi.org/10.1111/j.1567-1364.2009.00507.x>
- Holkenbrink C, Dam MI, Kildegaard KR, Beder J, Dahlin J, Domenech Belda D, Borodina I (2018) EasyCloneYALI: CRISPR/Cas9-based synthetic toolbox for engineering of the yeast *Yarrowia lipolytica*. *Biotechnol J*. <https://doi.org/10.1002/biot.201700543>
- Hoskisson PA, Hobbs G (2005) Continuous culture - making a comeback? *Microbiology* 151:3153–3159. <https://doi.org/10.1099/mic.0.27924-0>
- Hou J, Tyo KEJ, Liu Z, Petranovic D, Nielsen J (2012) Metabolic engineering of recombinant protein secretion by *Saccharomyces cerevisiae*. *FEMS Yeast Res* 12:491–510. <https://doi.org/10.1111/j.1567-1364.2012.00810.x>
- Kerkhoven EJ, Kim YM, Wei S, Nicora CD, Fillmore TL, Purvine SO, Webb-Robertson BJ, Smith RD, Baker SE, Metz TO, Nielsen J (2017) Leucine biosynthesis is involved in regulating high lipid accumulation in *Yarrowia lipolytica*. *MBio*. <https://doi.org/10.1128/mBio.00857-17>
- Larroude M, Rossignol T, Nicaud JM, Ledesma-Amaro R (2018) Synthetic biology tools for engineering *Yarrowia lipolytica*. *Biotechnol Adv* 36:2150–2164. <https://doi.org/10.1016/j.biotechadv.2018.10.004>
- Le Dall MT, Nicaud JM, Gaillardin C (1994) Multiple-copy integration in the yeast *Yarrowia lipolytica*. *Curr Genet* 26:38–44. <https://doi.org/10.1007/BF00326302>

- Liu Z, Hou J, Martínez JL, Petranovic D, Nielsen J (2013) Correlation of cell growth and heterologous protein production by *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* 97:8955–8962. <https://doi.org/10.1007/s00253-013-4715-2>
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 25:402–408. <https://doi.org/10.1006/meth.2001.1262>
- Looser V, Bruhlmann B, Bumbak F, Stenger C, Costa M, Camattari A, Fotiadis D, Kovar K (2014) Cultivation strategies to enhance productivity of *Pichia pastoris*: a review. *Biotechnol Adv* 33:1177–1193. <https://doi.org/10.1016/j.biotechadv.2015.05.008>
- Madzak C (2018) Engineering *Yarrowia lipolytica* for use in biotechnological applications: a review of major achievements and recent innovations. *Mol Biotechnol* 60:621–635. <https://doi.org/10.1007/s12033-018-0093-4>
- Madzak C, Blanchin-Roland S, Cordero Otero RR, Gaillardin C (1999) Functional analysis of upstream regulating regions from the *Yarrowia lipolytica* XPR2 promoter. *Microbiology* 145:75–87. <https://doi.org/10.1099/13500872-145-1-75>
- Madzak C, Tréton B, Blanchin-Roland S (2000) Strong hybrid promoters and integrative expression/secretion vectors for quasi-constitutive expression of heterologous proteins in the yeast *Yarrowia lipolytica*. *J Mol Microbiol Biotechnol* 2:207–216
- Madzak C, Otterbein L, Chamkha M, Moukha S, Asther M, Gaillardin C, Beckerich JM (2005) Heterologous production of a laccase from the basidiomycete *Pycnoporus cinnabarinus* in the dimorphic yeast *Yarrowia lipolytica*. In: *FEMS Yeast Research*. Elsevier, pp 635–646
- Nars G, Saurel O, Bordes F, Saves I, Rемаud-Siméon M, André I, Milon A, Marty A (2014) Production of stable isotope labelled lipase Lip2 from *Yarrowia lipolytica* for NMR: investigation of several expression systems. *Protein Expr Purif* 101:14–20. <https://doi.org/10.1016/j.pep.2014.05.007>
- Nocon J, Steiger MG, Pfeffer M, Sohn SB, Kim TY, Maurer M, Rußmayer H, Pflügl S, Ask M, Haberhauer-Troyer C, Ortmayr K, Hann S, Koellensperger G, Gasser B, Lee SY, Mattanovich D (2014) Model based engineering of *Pichia pastoris* central metabolism enhances recombinant protein production. *Metab Eng* 24:129–138. <https://doi.org/10.1016/j.ymben.2014.05.011>
- Ochoa-Estopier A, Guillouet SE (2014) D-stat culture for studying the metabolic shifts from oxidative metabolism to lipid accumulation and citric acid production in *Yarrowia lipolytica*. *J Biotechnol* 170:35–41. <https://doi.org/10.1016/j.jbiotec.2013.11.008>
- Park YK, Vandermies M, Soudier P, Telek S, Thomas S, Nicaud JM, Fickers P (2019) Efficient expression vectors and host strain for the production of recombinant proteins by *Yarrowia lipolytica* in process conditions. *Microb Cell Factories* 18:167. <https://doi.org/10.1186/s12934-019-1218-6>
- Pedersen H, Beyer M, Nielsen J (2000) Glucoamylase production in batch, chemostat and fed-batch cultivations by an industrial strain of *Aspergillus niger*. *Appl Microbiol Biotechnol* 53:272–277. <https://doi.org/10.1007/s002530050020>
- Peebo K, Neubauer P (2018) Application of continuous culture methods to recombinant protein production in microorganisms. *Microorganisms* 6:56. <https://doi.org/10.3390/microorganisms6030056>
- Pignede G, Wang HJ, Fudalej F, Seman M, Gaillardin C, Nicaud JM (2000) Autocloning and amplification of LIP2 in *Yarrowia lipolytica*. *Appl Environ Microbiol* 66:3283–3289. <https://doi.org/10.1128/AEM.66.8.3283-3289.2000>
- Puxbaum V, Mattanovich D, Gasser B (2015) Quo vadis? The challenges of recombinant protein folding and secretion in *Pichia pastoris*. *Appl Microbiol Biotechnol* 99:2925–2938
- Rakicka M, Biegalska A, Rymowicz W, Dobrowolski A, Mironczuk AM (2017) Polyol production from waste materials by genetically modified *Yarrowia lipolytica*. *Bioresour Technol* 243:393–399. <https://doi.org/10.1016/j.biortech.2017.06.137>
- Rodríguez-Carmona E, Cano-Garrido O, Dragosits M, Maurer M, Mader A, Kunert R, Mattanovich D, Villaverde A, Vázquez F (2012) Recombinant Fab expression and secretion in *Escherichia coli* continuous culture at medium cell densities: influence of temperature. *Process Biochem* 47:446–452. <https://doi.org/10.1016/j.procbio.2011.11.024>
- Rywińska A, Juszczak P, Wojtatowicz M, Rymowicz W (2011) Chemostat study of citric acid production from glycerol by *Yarrowia lipolytica*. *J Biotechnol* 152:54–57. <https://doi.org/10.1016/j.jbiotec.2011.01.007>
- Sambrook J, Russell D (2001) *Molecular cloning: a laboratory manual*, 3rd edn. Cold Spring Harbor Laboratory Press, New York
- Swietalski P, Hetzel F, Seitz I, Fischer L (2020) Secretion of a low and high molecular weight β -glycosidase by *Yarrowia lipolytica*. *Microb Cell Factories* 19:1–13. <https://doi.org/10.1186/s12934-020-01358-5>
- Theron CW, Vandermies M, Telek S, Steels S, Fickers P (2020) Comprehensive comparison of *Yarrowia lipolytica* and *Pichia pastoris* for production of *Candida antarctica* lipase B. *Sci Rep* 10:1–9. <https://doi.org/10.1038/s41598-020-58683-3>
- Vogl T, Gebbie L, Palfreyman RW, Speight R (2018) Effect of plasmid design and type of integration event on recombinant protein expression in *Pichia pastoris*. *Appl Environ Microbiol*. <https://doi.org/10.1128/AEM.02712-17>
- Wagner JM, Alper HS (2016) Synthetic biology and molecular genetics in non-conventional yeasts: current tools and future advances. *Fungal Genet Biol* 89:126–136. <https://doi.org/10.1016/j.fgb.2015.12.001>
- Wall KP, Dillon R, Knowles MK (2015) Fluorescence quantum yield measurements of fluorescent proteins: a laboratory experiment for a biochemistry or molecular biophysics laboratory course. *Biochem Mol Biol Educ* 43:52–59. <https://doi.org/10.1002/bmb.20837>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Applied Microbiology and Biotechnology

Impact of overproduced heterologous proteins characteristics on physiological response in *Yarrowia lipolytica* steady-state maintained continuous cultures

Korpys-Woźniak Paulina, Kubiak Piotr, Białas Wojciech, Celińska Ewelina*

¹Department of Biotechnology and Food Microbiology, Poznan University of Life Sciences, ul. Wojska Polskiego 48, 60-627 Poznań, Poland

*Corresponding author: Ewelina Celińska, e-mail: ewelina.celinska@up.poznan.pl, phone: +48618466007, ORCID 0000-0001-8372-8459

Fig.S1 Growth of Po1g-derived and Po1h-derived *Yarrowia lipolytica* sub-clones in shake-flask cultures of. X axis: culturing time [h], Y axis: biomass accumulation (OD600nm).

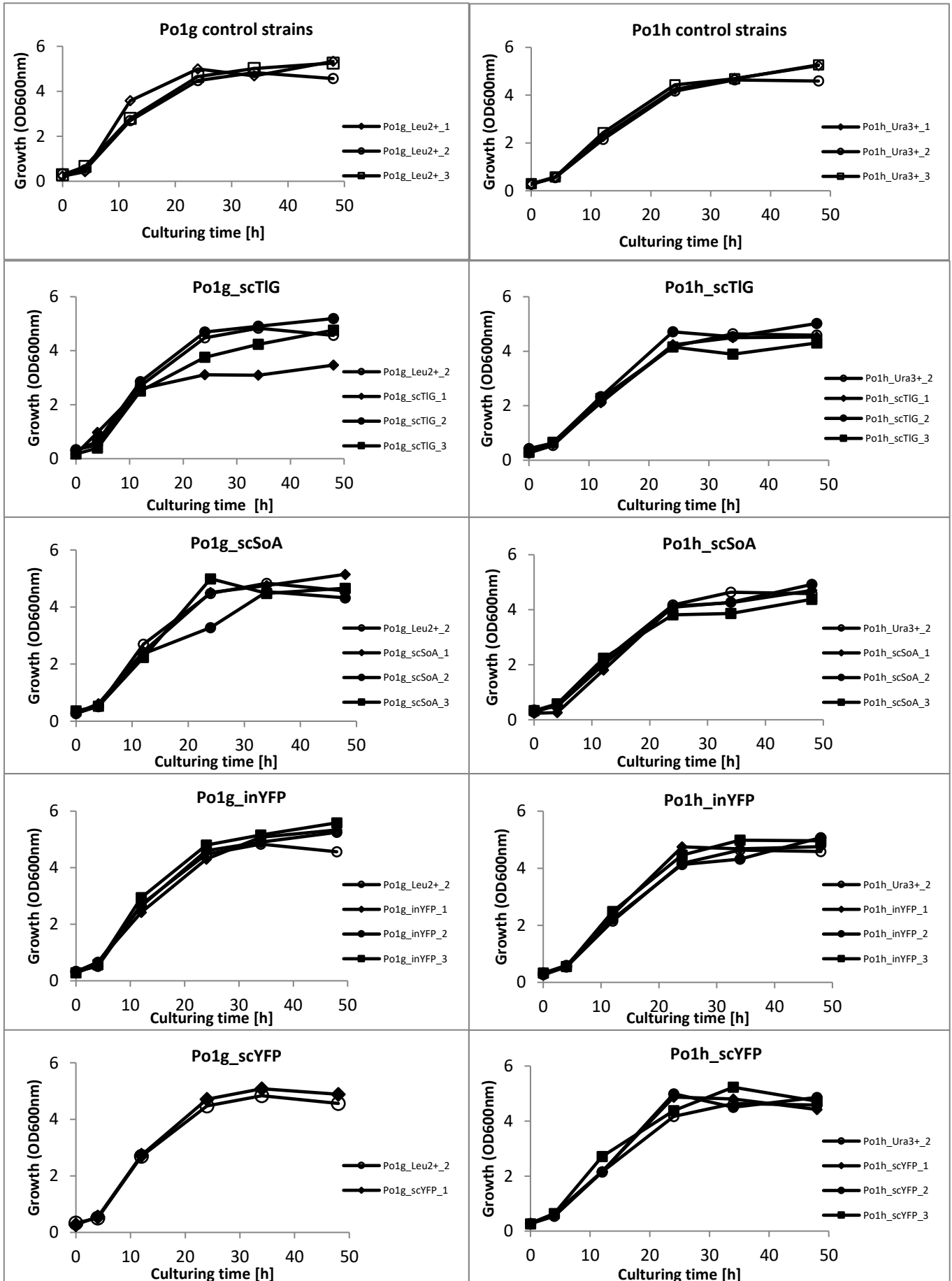


Fig.S2. Concentration of glycerol (GLY), biomass (DCW) and small molecular weight metabolites: erythritol (ERY), mannitol (MAN) and citric acid (CA) for continuous culture of *Yarrowia lipolytica* control strain (Po1h_Ura3+) grown in continuous cultures at: (A) 0.06, (B) 0.12 and (C) 0.20 [h⁻¹] dilution rates. X axis: residence time [nV], Y axis: concentration of GLY in [g L⁻¹] and DCW in [g_{DCW} L⁻¹] and ERY, MAN, CA in [g L⁻¹]. The error bars indicate $\pm SD$ at least duplicate.

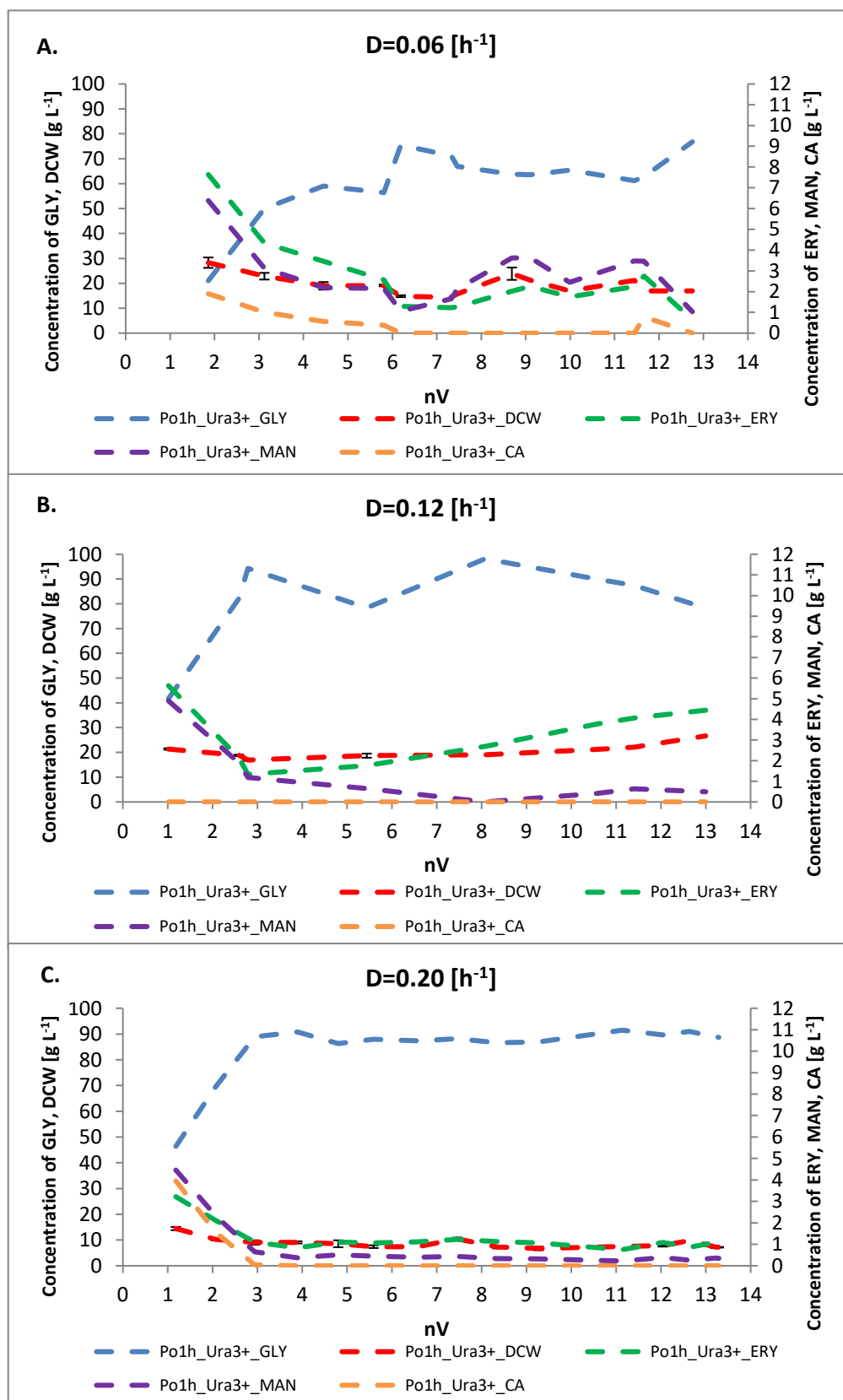
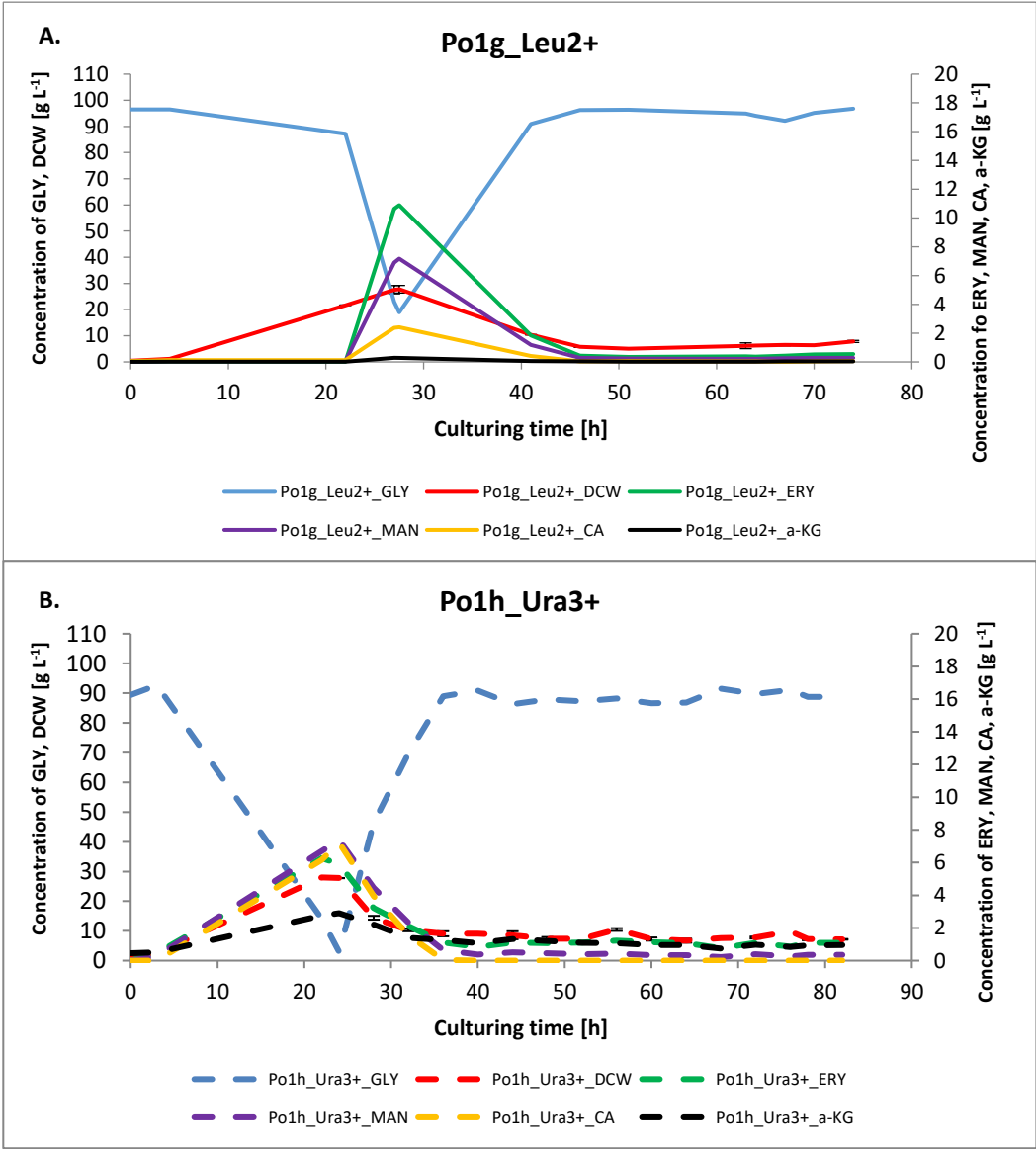
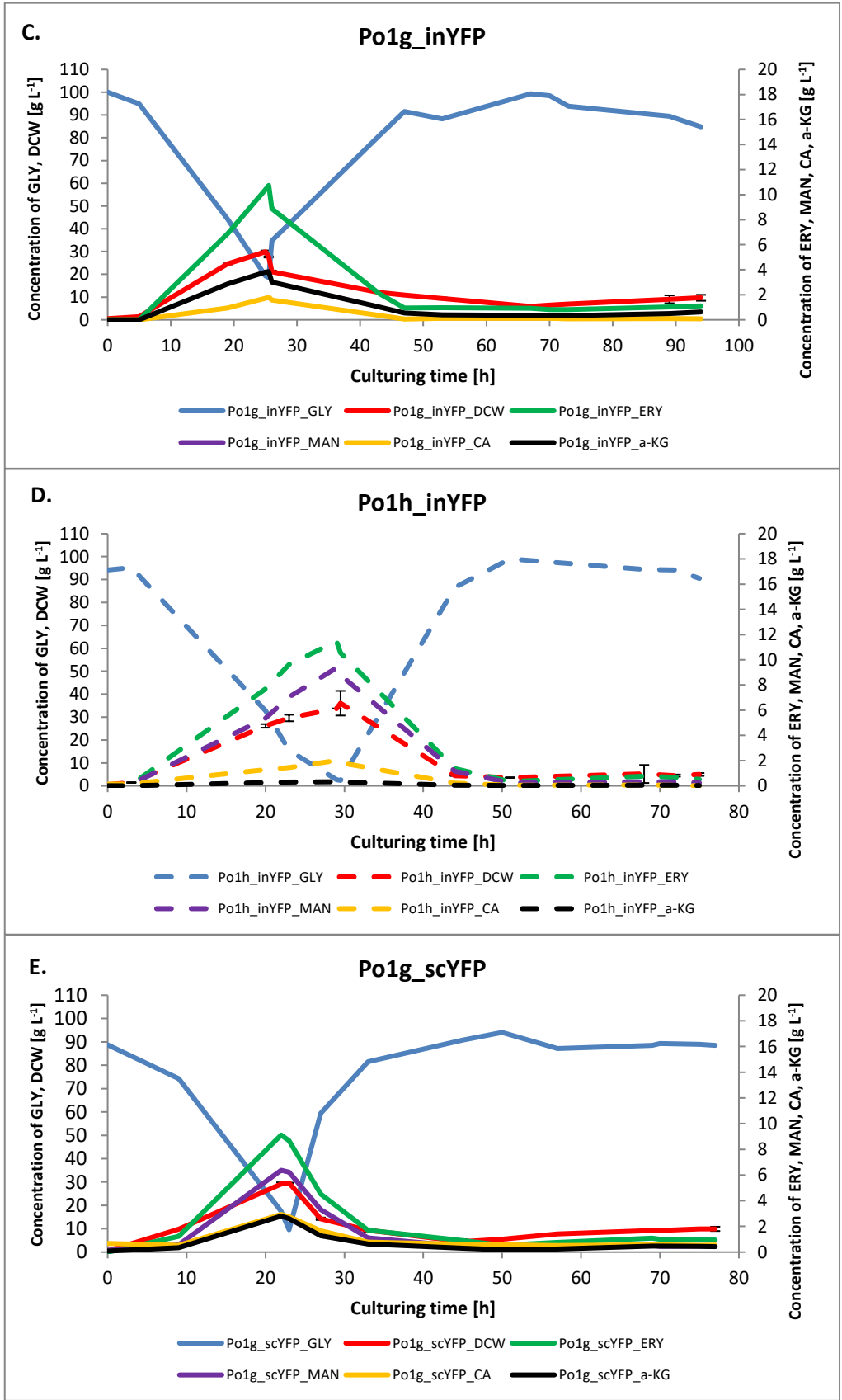
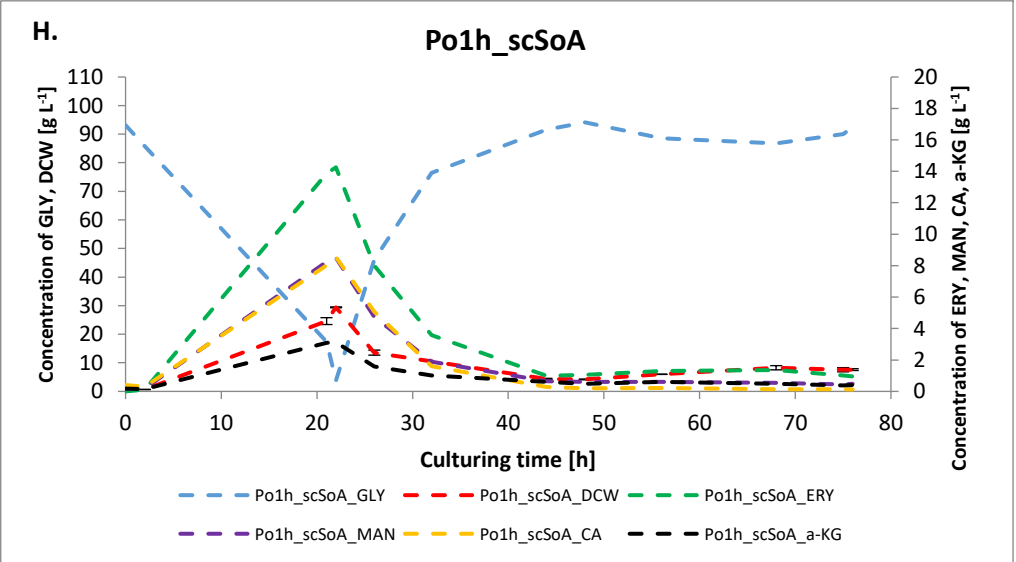
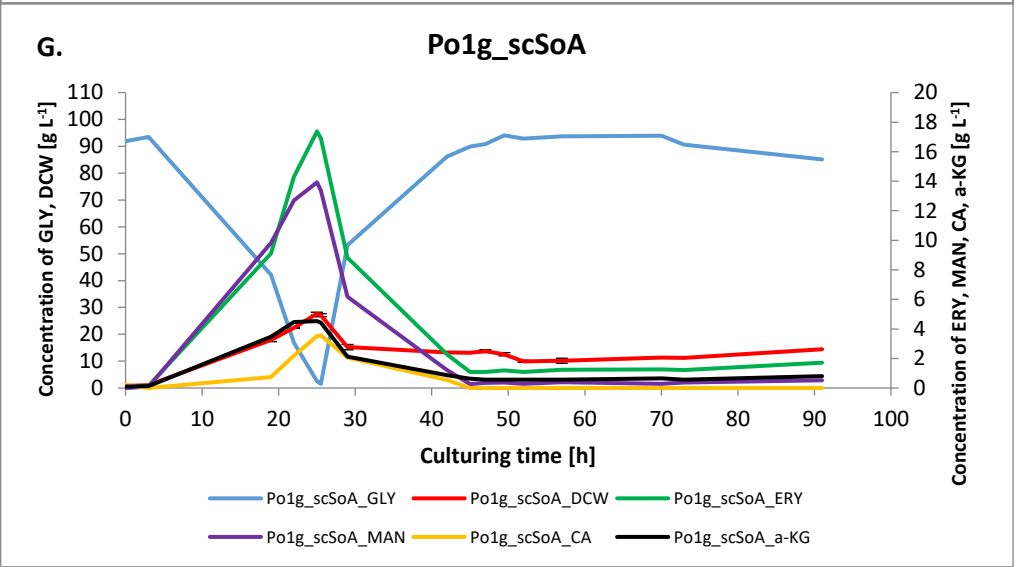
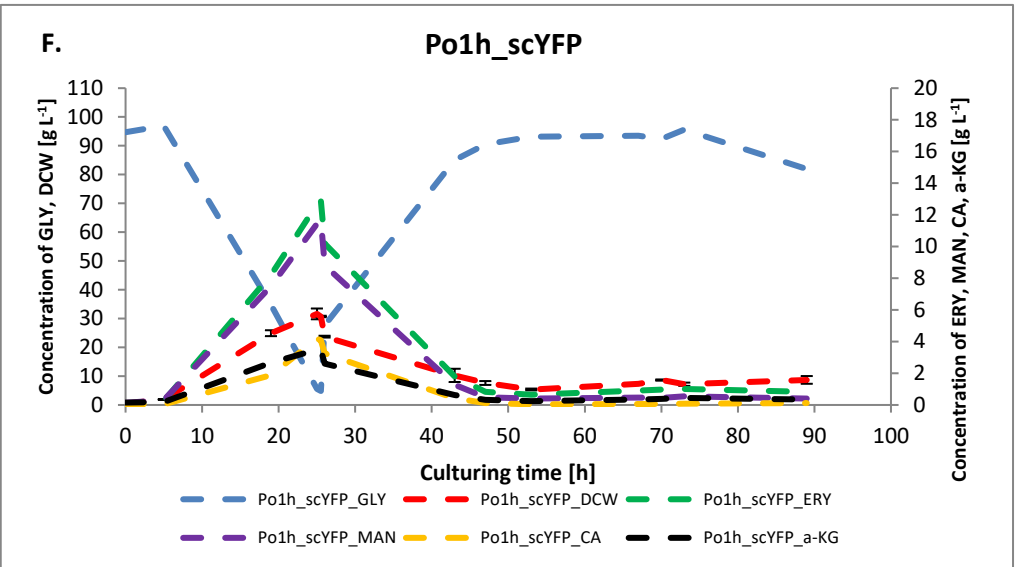


Fig.S3. Kinetics of Po1g-derived and Po1h-derived *Y. lipolytica* strains cultures, including the batch and continuous stages; X axis: culturing time (h); Y axis: glycerol concentration ($[g L^{-1}]$; GLY); biomass growth ($[g_{DCW} L^{-1}]$; DCW); erythritol ($[g_{DCW} L^{-1}]$; ERY), mannitol ($[g L^{-1}]$; MAN), citric acid ($[g L^{-1}]$; CA), α -ketoglutaric acid ($[g L^{-1}]$; a-KG) concentrations in (A) Po1g_Leu2+, (B) Po1h_Ura3+, (C) Po1g_inYFP, (D) Po1h_inYFP, (E) Po1g_scYFP, (F) Po1h_scYFP, (G) Po1g_scSoA, (H) Po1h_scSoA, (I) Po1g_scTIG, (J) Po1h_scTIG strains during batch and continuous bioreactor cultures; pO_2 20%, pH 5.5; D 0.20 $[h^{-1}]$; Error bars indicate $\pm SD$ from at least biological duplicates.







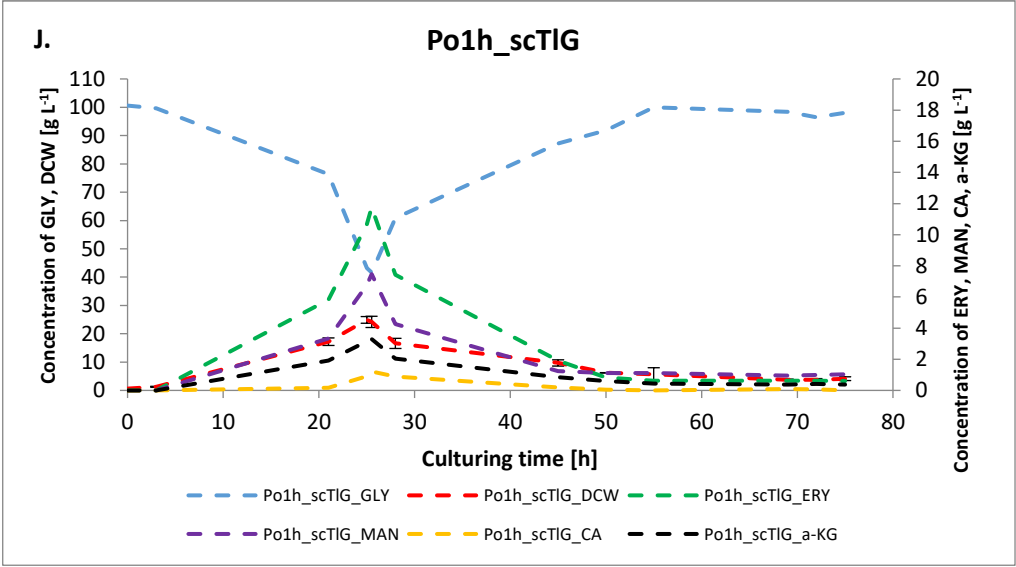
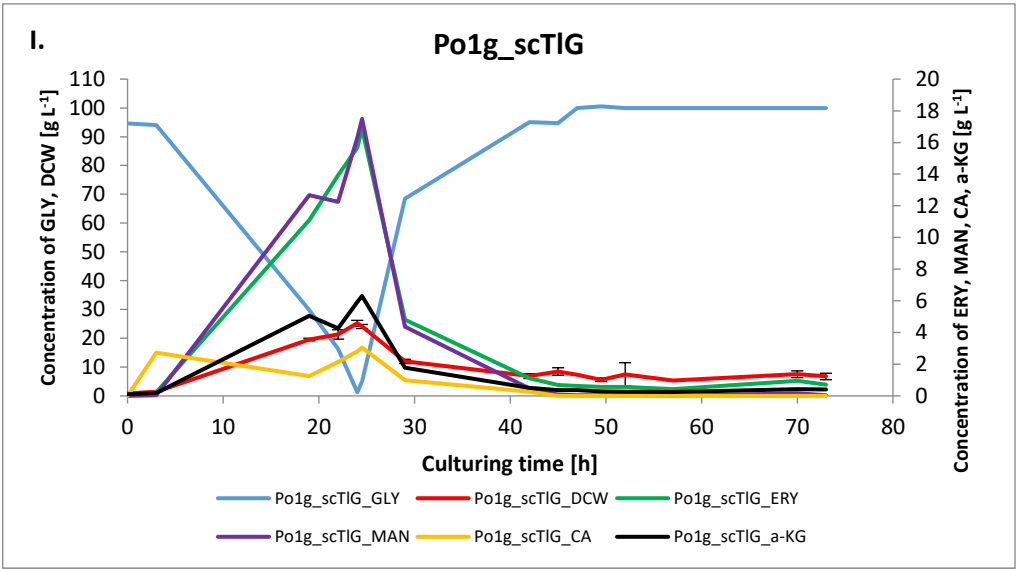


Table S1. Strains and vectors used in this study

Promoter	Name	Characteristics	Use	Supplier/reference
	Strains			
	<i>Yarrowia lipolytica</i> Po1g	Genotype: <i>MatA, leu2-270, ura3-302::URA3, xpr2-332, axp-2</i> Phenotype: <i>Leu⁻, ΔAEP, ΔAXP, Suc⁺, pBR platform</i>	Parental strain - host for expression of recombinant SoA,TIG,YFP gene	Yeastern Biotech Co.
<i>hp4d</i> 4UASpLeu2	<i>Yarrowia lipolytica</i> Po1g_Leu2+_2 (ECY32)	Po1g derivative transformed with Leu2 cassette	Control for Po1g derivatives	This study
	<i>Yarrowia lipolytica</i> Po1g_scSoA_2 (1.1.2.2)	Po1g derivative with expression cassette pYLSC_SoA	Expression of heterologous α -amylase from <i>S. oryzae</i>	Celińska et al. 2015
	<i>Yarrowia lipolytica</i> Po1g_scTIG_2 (2.2.1.2)	Po1g derivative with expression cassette pYLSC_TIG	Expression of heterologous glucoamylase from <i>T.lanuginosus</i>	This study
	<i>Yarrowia lipolytica</i> Po1g_scYFP_1 (ECY205)	Po1g derivative with expression cassette pYLSC_YFP	Expression of secreted heterologous YFP	This study
	<i>Yarrowia lipolytica</i> Po1g_inYFP_2 (ECY18)	Po1g derivative with expression cassette pYLEX_YFP	Expression of cytosolic heterologous YFP	This study
	<i>Yarrowia lipolytica</i> Po1h	Genotype: <i>Mata, ura3-302, xpr2-322, axp1-2, leu2-270::LEU2</i> Phenotype: <i>Ura⁻, ΔAEP, ΔAXP, Suc⁺</i>	Parental strain - host for expression of heterologous SoA,TIG,YFP gene	Celińska et al. 2020
<i>4UASpTEF</i>	<i>Yarrowia lipolytica</i> Po1h_Ura3+_2 (GGY231)	Po1h derivative transformed with Ura3 cassette	Control for Po1h derivatives	This study
	<i>Yarrowia lipolytica</i> Po1h_scSoA_1 (GGY237)	Po1h derivative with expression cassette bearing single TU with SP1 upstream fused to gene encoding α -amylase	Expression of secreted heterologous α -amylase from <i>S. oryzae</i>	This study
	<i>Yarrowia lipolytica</i> Po1h_scTIG_2 (GGY228)	Po1h derivative with expression cassette bearing single TU with SP1 upstream fused to gene encoding glucoamylase	Expression of secreted heterologous glucoamylase from <i>T.lanuginosus</i>	This study
	<i>Yarrowia lipolytica</i> Po1h_scYFP_3 (GGY251)	Po1h derivative with expression cassette bearing single TU with SP1 upstream fused to gene encoding YFP	Expression of secreted heterologous YFP	This study

	<i>Yarrowia lipolytica</i> Po1h_inYFP_3 (GGY245)	Po1h derivative with expression cassette bearing single TU with gene encoding YFP, without SP1	Expression of cytosolic heterologous YFP	This study
	JM109 <i>Escherichia coli</i>	F'(traD36,proAB+,lacIq,Δ(lacZ)M15), endA1,recA1,hsdR17 (rk-,mk+), mcrA,supE44,λ-gyrA96, relA1, Δ(lacproAB),thi-1	Host for routine cloning, vector assembly and propagation	Sigma-Aldrich
	Vectors			
<i>hp4d</i> 4UASpLeu2	pYLSC	pBR322 backbone, hybrid promoter (hp4d) XPR2 pre-region, leucine gene (LEU2) – selection marker Ltd, Taiwan Integration site at pBR-docking platform Prior to transformation the vector was linearized with NotI endonuclease http://www.yeastern.com/Products.php?pid=273	Secretion vector used for transformation of <i>Y. lipolytica</i> Po1g strain	Yeastern Biotech Co.
<i>hp4d</i> 4UASpLeu2	pYLEX	pBR322 backbone, hybrid promoter (hp4d), leucine gene (LEU2) – selection marker Ltd, Taiwan Integration site at pBR-docking platform Prior to transformation the vector was linearized with NotI endonuclease http://www.yeastern.com/Products.php?pid=273	Expression vector used for transformation of <i>Y. lipolytica</i> Po1g strain	Yeastern Biotech Co.

Table S2. Biochemical properties of the reporter proteins used in this study

Protein	Molecular weight [Da][#]	Number of aa[#]	Theoretical pI[#]	Predicted N-glycosylation sites[*]	Predicted O-glycosylation sites^{**}	GRAVY[#]	Number of Cys[§]	Total different combinations S-S[§]
YFP	26991.54	239	5.58	-	1	-0.464	2	2
SoA	52140,40	474	4.88	1	8	-0.395	11	35696
TIG	65154.96	600	4.90	3	18	-0.266	8	764

^{*}<http://www.cbs.dtu.dk/services/NetNGlyc/> N-glycosylation

^{**}<http://www.cbs.dtu.dk/services/NetOGlyc/> O-glycosylation

[#]<https://web.expasy.org/protparam/> - MW, aa, GRAVY, pI

[§]https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_cysteines.html disulfide bonds

Table S3. Primers used in this study

Primer	Sequence 5'→3'	Reference	
GGP_G1_SP1_eYFP_D_Fw	GGTCTCTAATGAAGTTCACATTTGCTGCCGTTACCGCCGCGCTGGCCTCGTCCG	This study	
GGP_G1_eYFP_D_Fw	GGTCTCTAATGGTGAAGGGCGAGGAGCTGTTACCC		
GGP_G1_eYFP_E_Rev	GGTCTCTTAGATTACTTGTACAGCTCGTCCATGCCGAG		
GGP_G1_SP1_D_Fw	GGTCTCTAATGAAGTTCACATTTGCTGCCG	Celińska et al. 2017	
GGP_G1_SoA_E_Rev	GGTCTCTTAGACTAGTGGTGGTGGTGG	This study	
GGP_G1_TIG_E_Rev	GGTCTCTTAGATCAATGATGATGATGATGATG		
GGP_T1_LIP2_E_Fw	GGTCTCTTAGTGTCTGTGGTATCTAAGCTATTTATC		
GGP_T1_LIP2_F_Rev	GGTCTCTAAGCCGATTTGTCTTAGAGGAACGC		
GGP_P1_4UASpTEF_GS_C_Fw	GGGGTCTCTAGCTGCATGCTGAGGTGTCTCACAAG		
GGP_P1_4UASpTEF_GS_D_Rev	GGTCTCTCATTCTCGGGTGTGAGTTGACAAGG		
GGP_M_Ura3_B_Fw	GGTCTCTAGGTATAACTTCGTATAGCATAACATTATACGAAG		
GGP_M_Ura3_C_Rev	GGTCTCTAGCTTCGCTTCGGATAACTCC		
GGP_InsertUP_zeta_A_Fw_NotI	GGTCTCTGCCTGCGGCCGCTGTGCGGAACCCG		
GGP_InsertUP_zeta_B_Rev	GGTCTCTACCTTCTAGCAAAGTGCTTTGTGC		
GGP_InsertDOWN_zeta_L_Fw	GGTCTCTGAGTCATGTGAACACTCGCTCTG		
GGP_InsertDOWN_zeta_M_Rev_NotI	GGTCTCTGCAGCGGCCGCACTGAAGGGCTTTGTGAG		
6560 F	GAT CCG GCA TGC ACT GAT C		Yeastern Biotech Co.
6904 R	AAC ACC GGT GTT GGA CTC AG		
pYLSC_YFP_KpnI_Rev	AAAGGTACCTTACTTGTACAGCTCGTCCATGCCG	This study	
pYLSC_YFP_SfiI_Rev	GGCCGTTCTGGCCCATGGTGAAGGGCGAGGAGCTG		
pYLEX_YFP_Fw	AATGGTGAGCAAGGGCGAGGAGCTGTTACCCG		
pYLSC/EX_YFP_KpnI_Rev	AAAGGTACCTTACTTGTACAGCTCGTCCATGCCG	Celińska et al. 2015	
pYLSC_SoA_SfiI_Fw	AGGCCGTTCTGGCCAGAACGACCCCACTTCC		
pYLSC_SoA_KpnI_Rev	AAGGTACC CTAGTGGTGGTGGTGGTGGTGG		
pYLSC_TIG_SfiI_Fw	AGGCCGTTCTGGCCAGAACGACCCCACTTCC		
pYLSC_TIG_KpnI_Rev	AAGGTACC TCAATGATGATGATGATGATG	This study	
r-t_SoA_Fw	GTAACAACGTGGGAATCCGAAT		
r-t_SoA_Rev	CCCTGGCCGTTCCAAGTAG		
r-t_TIG_Fw	TCTGATGTTGTTGGGAATCTGA		
r-t_TIG_Rev	TGATGTCTCCACCAAGAGTTTACTG		
r-t_YFP_Fw	CAACAGCCACAACGTCTATATCATG		
r-t_YFP_Rev	ATGTTGTGGCGGATCTTGAAGT		
r-t_ACT_Fw	CGAGCGAATGCACAAGGA		
r-t_ACT_Rev	GAGCGGTGATCTTGACCTTGA		

Table S4. Viability of Po1g- and Po1h-derived *Y. lipolytica* strains maintained in steady-state in chemostat cultivations. ([*]- surface colony forming unit).

Strains	[scfu*]
Po1g_Leu2+	8.30*10 ⁴
Po1h_Ura3+	1.93*10 ⁵
Po1g_inYFP	1.43*10 ⁵
Po1h_inYFP	1.77*10 ⁵
Po1g_scYFP	1.10*10 ⁵
Po1h_scYFP	9.70*10 ⁴
Po1g_scSoA	1.00*10 ⁵
Po1h_scSoA	1.67*10 ⁵
Po1g_scTIG	1.70*10 ⁵
Po1h_scTIG	1.07*10 ⁵



Global transcriptome profiling reveals genes responding to overproduction of a small secretory, a high cysteine- and a high glycosylation-bearing protein in *Yarrowia lipolytica*

Paulina Korpys-Woźniak, Ewelina Celińska *

Department of Biotechnology and Food Microbiology, Poznan University of Life Sciences, ul. Wojska Polskiego 48, 60-637, Poznań, Poland

ARTICLE INFO

Keywords:

Transcriptomics
Heterologous proteins
Secretory pathway
Yeast
Metabolic burden

ABSTRACT

Investigation of the yeast cell's response to recombinant secretory protein (rs-Prot) overproduction is relevant for both basic and applied research. Imbalance, overloading or stress within this process impacts the whole cell. In the present study, by using steady-state cultures and transcriptomics, we investigated the cellular response of *Yarrowia lipolytica* challenged with high-level expression of genes encoding proteins with significantly different biochemical characteristics: a small protein retained within the cell i) or secreted ii), a medium size secretory protein with a high number of disulfide bonds iii), or glycosylation sites iv). Extensive analysis of omics data, supported by careful manual curation, led to some anticipated observations on oxidative and unfolded protein stress (*CTT1*, *PXMP2/4*, *HAC1*), glycosylation (*ALGs*, *KTRs*, *MNTs*, *MNNs*), folding and translocation (*SSAs*, *SSEs*) but also generated new exciting knowledge on non-conventional protein secretion (*NCE102*), transcriptional regulators (*FLO11*, *MHY1*, *DO1353* g, *RSFA*, *E23925g* or *MAF1*), vacuolar proteolysis targets in *Y. lipolytica* (*ATGs*, *VPSs*, *HSE1*, *PRB1*, *PRC1*, *PEP4*) or growth arrest (*CLN1*) upon rs-Prots overproduction.

1. Introduction

The investigation of the yeast cell's response to recombinant secretory protein (rs-Prot) overproduction is relevant for both basic and applied research. It enables discovery of basic principles governing the cell, and identification of new targets for genetic engineering of the host to improve its performance. Synthesis of rs-Prot is a complex process taking place in multiple compartments of the cell. To be executed successfully, multiple processes must be coordinated, starting from transcription and translation, translocation into endoplasmic reticulum (ER) lumen, through folding, specific proteolytic cleavage, glycosylation, and disulfide bond formation, followed by a complex net of vesicular transit between compartments, all with an extensive system of quality control mechanisms and check points [1]. The cell must adjust the chaperone capacity, assure availability of resources, like amino acids, redox equivalents, ATP, membranes and glycans, as well as to adjust trafficking patterns, based on the collection of proteins that are

synthesized at a given time. It has been estimated, that one-third of all the proteins is processed by the secretory pathway [2]. As such, substantial cellular resources are expended to maintain fluent operation of the secretory pathway, and consequently, imbalance, overloading or stress within this process severely impact the whole cellular metabolism. Due to high complexity of the pathway, dysfunctions may occur at many different levels, and dissecting the causative agent can be challenging [2]. Hence, the design of an efficient rs-Prot's production platform requires an in-depth knowledge on its biology when challenged with this metabolically demanding task.

The world rs-Prot market constitutes one of the key branches of current industrial biotechnology, earning billion dollar (USD) revenues each year [3]. Within that field, yeast-based production platforms share the leadership with bacteria and fungi. Significant effort has gone into engineering yeast for increasing protein secretion (revised in [4–6]). Aside from engineering strategies targeting the r-Prot-encoding sequence itself, like shuffling with signal peptides or design of

Abbreviations: DEG(s), differentially expressed gene(s); ER, endoplasmic reticulum; UPR, unfolded protein response. Genes names are given as shortened codes without preceding YALi0.; HSS(s), high-secretor strain(s); *Pichia pastoris*, traditional name for the renamed *Komagataella phaffii*; r(s)-Prot, recombinant (secretory) heterologous protein; *SoA/TIG/scYFP/inYFP*, strains overproducing a given heterologous protein (small and italicized letters); *SOA/TLG/SCYFP/INYFP*, genes encoding the target heterologous proteins (capitalized and italicized); *SoA/TIG/scYFP/inYFP*, heterologous proteins (small and non-italicized).

* Corresponding author.

E-mail address: ewelina.celinska@up.poznan.pl (E. Celińska).

<https://doi.org/10.1016/j.btre.2021.e00646>

Received 2 April 2021; Received in revised form 8 June 2021; Accepted 9 June 2021

Available online 11 June 2021

2215-017X/© 2021 The Author(s).

Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

completely synthetic targeting elements [7–9], native mechanisms of the secretory pathway have been modified as well. The latter strategies targeted oxidative folding [10–12], elements of intracellular protein trafficking [13,14], the main transcription factor governing response to unfolded proteins (UPR) [15–17] or abolished excessive proteolysis [18–20].

Being aware of the complexity of the rs-Prot synthesis process, on a quest for new engineering targets the researchers harness omics analytical platforms. Global profiling of the metabolite, protein or transcript pools is used to scrutinize physiological response of a cell challenged with heterologous overproduction. Such a strategy was found very potent in identification of unobvious targets for cell engineering [21–23] or cultivation conditions favoring rs-Prot production [24,25]. The new targets were identified as the most responsive individual genes (differentially expressed genes; DEGs) or being a part of a highly responsive, up- or down-regulated biological process. In several studies, transcriptomics was used to comparatively study a global yeast cell's response to overproduction of several different r(s)-Prots. In a series of papers on *Saccharomyces cerevisiae* perturbed with secretion of a small not glycosylated protein with even number of cysteines, and a larger protein with an odd number of Cys and a glycosylation site, numerous interesting observations were reported [2,5,26–28]. Recently, a comparative transcriptomics study on *Yarrowia lipolytica* challenged with overproduction of a small (55 kDa) and large (>100 kDa) beta-glycosidases has been reported [29]. All these studies provide substantial amounts of comprehensive, detailed and valuable knowledge, with high applicatory potential.

In the present study, we investigated cellular response of *Y. lipolytica* strains challenged with high-level expression of different r(s)-Prot-encoding genes. *Y. lipolytica* is one of the yeast strains considered an attractive protein production platform [4]. One of the biggest strengths of this species is an atypical, as for yeast, secretory pathway, resembling filamentous fungi [30]. We've been particularly interested in how *Y. lipolytica* cells respond to perturbations imposed by synthesis of proteins with significantly different biochemical characteristics. Hence, for our experimental set we've chosen: i) a small intracellular protein (inYFP), as a reference to its secretory counterpart (scYFP), ii) a small secretory protein with negligible posttranslational modifications (scYFP), iii) a medium size secretory protein with a high number of disulfide bonds (SoA;), iv) a medium size secretory protein with a high number of glycosylation sites (TIG; Table 1). The set of reporter r-Prots was composed in order to highlight cellular responses of *Y. lipolytica* to burdening the translational-secretory machinery at different points/processes. To get a thorough insight into the cell's molecular response independently from the environmental conditions and the growth phase, the strains were cultured under the same conditions, in continuous mode to reach a steady-state. The following, global transcriptome profiling allowed to grasp expected, but here experimentally confirmed for *Y. lipolytica* phenomena, like an outburst of oxidative stress upon high rs-Prot synthesis; but also to observe many new, previously unreported for *Y. lipolytica*, responses. New targets for further cell engineering strategies have been identified and our knowledge on *Y. lipolytica* has been largely broadened. This knowledge can be used as a guidance for *Y. lipolytica* strains improvement towards more effective production of heterologous secretory proteins.

Table 1

Biochemical properties of the r-Prot reporter proteins used in this study. Bioinformatic Tools used for determination of glycosylation sites and disulfide bonds are given in section 2.7.5.

Protein	Molecular weight [Da] [#]	Number of aa [#]	Theoretical pI [#]	Predicted N-glycosylation sites*	Predicted O-glycosylation sites**	Number of Cys [§]	Total different combinations S-S [§]
YFP	26991.54	239	5.58	–	1	2	2
SoA	5,214,040	474	4.88	1	8	11	35,696
TIG	65154.96	600	4.90	3	18	8	764

2. Materials and methods

2.1. Strains and culturing

All the yeast strains used in this study are derivatives of *Y. lipolytica* Po1h (Table S.1). The control strain (Po1h_Ura3) was transformed with a solo *URA3* marker cassette to generate a prototroph. Recombinant *Y. lipolytica* strains were transformed with Golden Gate Assembly cassettes bearing a single transcription unit [TU] each with one of the target genes [31]. The target genes encoded either enzymatic (alpha-amylase – SoA, glucoamylase - TIG) or fluorescent (intracellular in- /secretory sc-form of YFP) proteins, expressed individually under the control of a synthetic hybrid promoter 4UASpTEF (4 direct repetitions of UAS and a minimal promoter of pTEF with CA environment). Secretory enzymatic reporters were transcriptionally fused with a signal peptide sequence native for *Y. lipolytica* exo-1,3-beta-glucanase (B03564g) [7]. Yeast strains were routinely cultured in YPD medium (YPD [g L⁻¹]: yeast extract (BTL, Lodz, Poland), 10; bactopectone (BTL), 20; glucose, 20 (POCh, Gliwice, Poland)) liquid or solidified with agar at 15 [g L⁻¹] (Biomaxima, Lublin, Poland), at 28 °C. For long-term storage, the strains were kept in 50 % glycerol stocks at –80 °C.

2.2. Chemostat cultivation

Continuous chemostat cultivations were performed according to previously described methodology [32] with some minor modifications. Briefly, the biomass from the glycerol stocks was transferred onto YPD agar plate and incubated at 30 °C for 24 h. The precultures were prepared by inoculating 30 mL of YPG₂₀ medium ([g L⁻¹]: yeast extract, 10; bactopectone, 20; glycerol (POCh), 20) with a single colony, followed by cultivation at 28 °C for 22 h with agitation (250 rpm). Bioreactor cultures were conducted in fully instrumented and automatically controlled 0.5 L stirred tank bioreactors Multifors 2 (Infors HT, Bottmingen Basel, Switzerland) under the following conditions: temperature 28 °C, pO₂ was kept at 20 % using cascade-control approach with stirring from 100 up to 1200 rpm, and constant aeration at 2.0 vvm. pH was maintained at 5.5 by automatic addition of 20 % NaOH and 10 % H₂SO₄. Foaming was minimized by automatic addition of a defoaming agent AntiFoam 204 (Sigma-Aldrich). The culturing medium YPG₁₀₀ ([g L⁻¹]: yeast extract, 10; bactopectone, 20; glycerol, 100) in 300 mL volume was inoculated at 10 %. The culture was continued in a batch mode until glycerol was consumed, which was indicated by an increase in the pO₂ parameter. Then the chemostat cultures were initiated at a dilution rate [D] of 0.20 [h⁻¹]. Steady state was reached after at least 6 residence times [τ]. The experiments were performed in biological duplicate.

2.3. RNA extraction

Total RNA was extracted from the biomass samples collected from the steady-state culture. The RNA was isolated using Bead-Beat Total RNA Mini Kit (A&A Biotechnology, Gdynia, Poland) according to the manufacturer's protocol. Extracts were obtained by mechanical disruption of the cells using Mixer Mill MM400 (Retsch, Germany) at 30 Hz for 30 s. The RNA extracts were verified for quantity using spectrophotometer (NanoDrop, Thermo Fisher Scientific) and for quality through agarose gel electrophoresis [33].

2.4. RTqPCR

The total RNA samples were reverse transcribed to the complementary DNA (cDNA) using SuperScript III Reverse Transcriptase and oligo (dT) primer, according to the manufacturer's protocol (Thermo Fisher Scientific). The received cDNA samples were used as templates for RTqPCR reaction using RT HSPCR Mix SYBR® B (A&A Biotechnology) and gene-specific primers listed in **Table S.2**. The RTqPCR was carried out in an Applied Biosystems 7500 apparatus (Applied Biosystems, Foster City, USA) using the following thermal profile: 95 °C 3 min, (95 °C 15 s, 60 °C 30 s., 72 °C 30 s.) × 40, 72 °C 1 min, Melt Curve 94 °C 15 s, 60 °C 60 s, 95 °C 30 s., 60 °C 15 s.. The expression level of the genes was normalized to the expression level of actin- and translocation protein Sec62-encoding genes (*ACT1*; *SEC62*), used as internal calibrators, and to the reference strain (*Y. lipolytica* Po1h_Ura3+), used as the external calibrator, to which 1.0 expression level was assigned. All the samples were analyzed in biological duplicate and in technical triplicate. The results were presented as Log₁₀ddCT calculated according to the $\Delta\Delta$ CT method [34].

2.5. RNA sequencing

Library construction and NGS sequencing were performed at Genomed S.A. (Warsaw, Poland) using HiSeq 4000 next-generation sequencer (Illumina) in a paired end mode. Reads with a length of 150 base pairs were generated. The raw sequencing reads are available on the NCBI Sequence Read Archive under the BioProject submission **PRJNA701856** (see Data Availability section).

2.6. RNAseq data analysis

2.6.1. Raw data processing and data filtration

NGS adapters were removed using Cutadapt [35]. Initial standard filtering was done using *q25* quality parameter and *m15* the minimum length of the reads. Re-filtering was used to discard reads <20 bp. Quality control was conducted using FastQC [36]. Post-QC data were aligned and mapped to the *Y. lipolytica* CLIB122 reference genome (GenBank assembly identifier: GCF_000002525.2) using Hisat2 [37]. The number of readings pairs, mapped to an individual gene, was counted using the HTseq program [38]. Gene Ontology (GO) term annotations were assigned using the reference genome template (GCF_000002525.2). Comparative gene expression analysis was conducted using the DESeq2 package [39]. To compare the expression of different genes across the samples, raw read counts were converted into normalized RNAseq counts. Alternatively, relative gene expression level in reference to the control strain (Po1h_Ura3) was expressed as fold change. All Differentially Expressed Genes (DEGs) meet the requirement of $FDR \leq 0.05$ (adjusted p value) tested *via* Wald test in DESeq2. Only DEGs meeting all those conditions were considered as responsive to introduced variable (overproduction of a specific heterologous protein) and subjected to further analyses. A complete list of all DEGs is attached as **Table S.3**.

2.6.2. Gene sets analysis

GO term statistical overrepresentation test was carried out by running DEGs list analysis with PANTHER v14.0-based online tool (rel. 20,200,728) [40]. First, GO terms were assigned to YALI DEGs according to GO Ontology database DOI: 10.5281/zenodo.4437524 Released 2021-01-01, specific for *Y. lipolytica* species. The test was carried out for biological process GO complete annotation data set, using Fisher's exact test and False Discovery Rate for multiple testing. GSA was conducted on complete gene lists of all DEGs, as indicated by DESeq2. Only results with $p < 0.05$ are displayed.

2.6.3. DEGs analysis through venn diagrams

Analysis of DEGs lists through Venn diagrams was carried out using

InteractiVenn online tool [41]. Up to 6 sets of up or down regulated gene lists were analyzed. Generated Venn Diagrams were later adjusted using inkscape (<https://inkscape.org/release/inkscape-1.0.2/>), according to the InteractiVenn's instruction. For the analysis using Venn Diagrams, DEGs predicted by DESeq2 were further filtered for 2-fold up- or down-regulation. Assignment of genes' names and genes' functions was manually curated by cross-referencing several databases – Panther, UniProt, GRYC and NCBI and blasting (blastp) the sequences against the databases' collections. The most relevant matches were chosen to finally describe a given gene. All the identified DEGs sets, common or specific for a given strain, were subjected to Gene Sets Analysis (according to methodology in point 2.6.2.). Identified enriched biological processes were indicated, where found.

2.7. Analytical methods

2.7.1. Biomass concentration

Dry cellular weight (DCW; [g L⁻¹]) was determined gravimetrically through drying the biomass in a laboratory dryer at 105 °C for 2–3 days, until a stable readout was reached. The dry biomass concentration was expressed in grams of the cell dry mass per liter [g L⁻¹].

2.7.2. HPLC analysis

High-performance liquid chromatography was conducted using Agilent Technologies 1200 series chromatograph (Agilent Technologies, Santa Clara, CA, USA) equipped with a refractive-index detector (G1362A) and a Rezex ROA-Organic Acid H + column (Phenomenex, Torrance, CA, USA). Operating conditions were as follows: 0.005 N H₂SO₄ as eluent at a flow rate of 0.6 [mL min⁻¹]; the column temperature was set at 40 °C.

2.7.3. Enzymatic assays

The activity of alpha-amylase (SoA) was measured in the culture supernatant samples using microSIT assay, according to [42]. One activity unit (alpha-amylase activity unit [AAU]) refers to the amount of enzyme that decomposes 1 mg of starch per 1 mL, during 1 min, at pH 5.0 and 40 °C, under the applied experimental conditions. The glucoamylase (TIG) activity was examined in the supernatant samples using microDNS (3,5-dinitrosalicylic acid) method, described previously [43]. One activity unit (glucoamylase activity unit [GAU]) was defined as the amount of enzyme that releases 1 µg of reducing sugar equivalents per 1 mL, during 1 h, at pH 5.0 and 40 °C, under the adopted assay conditions. The reference strain Po1h_Ura3+ was assayed in parallel to determine background level. All the readouts were normalized per blank reaction with distilled water. The colorimetric enzymatic assays were read using the Tecan Infinite M200 plate reader at 580 nm (microSIT) and 540 nm (microDNS) wavelength. Each sample was analyzed in technical duplicate.

2.7.4. Fluorescence determination

Fluorescence of the YFP reporter was measured in flat-bottomed MTP plates (Corning; Sigma-Aldrich) using Tecan Infinite M200 automatic plate reader. scYFP was assayed in 200 µl supernatant samples. inYFP was assayed in ddH₂O-washed biomass re-suspended in the initial volume of ddH₂O. The following wavelength settings were used: excitation 495 nm/emission 527 nm. The fluorescence was expressed as Relative Fluorescence Unit [RFU] defined as the sample median fluorescence value minus the background fluorescence value. scYFP was normalized vs fresh culture medium. inYFP was normalized vs Po1h_Ura3+ reference strain biomass. Each sample was analyzed in technical duplicate.

2.7.5. Basic bioinformatic tools

Nucleotide or protein sequences were withdrawn and analyzed using NCBI (<https://www.ncbi.nlm.nih.gov/>), GRYC (<http://gryc.inra.fr/>) and UniProt (<https://www.uniprot.org/>) platforms. Gene function

classification and basic information on the identified DEGs or their homologs was obtained from either GRYC, SGD (<https://www.yeastgenome.org/>) or UniProt. The search for homologs was conducted using blastp tool using default settings. Basic biochemical properties of the target proteins were determined using ProtParam Tool from ExPASy service (<https://web.expasy.org/protparam/>). Prediction of N- and O-glycosylation sites, and the disulfide bonds formation combinations was conducted using NetNGlyc and NetOGlyc tools at <http://www.cbs.dtu.dk/services> and NPSA tool at <https://npsa-prabi.ibcp.fr>, respectively.

3. Results and discussion

3.1. Basic characteristics of the strains and the reporter proteins

The experimental set of strains in this study covers four *Y. lipolytica* Po1h derivatives overproducing the following heterologous reporter proteins (r-Prot): i) small intracellular protein (inYFP), ii) small secretory protein with negligible posttranslational modifications (scYFP), iii) medium secretory protein with a high number of disulfide bonds (SoA), iv) medium secretory protein with a high number of glycosylation sites (TIG; Table 1). It was presumed that overproduction of such a set of varying proteins would highlight the cellular response of *Y. lipolytica* to burdening the translational-secretory machinery at different points.

To gain an insight into the cell's molecular response, independently from the environmental conditions and the growth phase, the strains were cultured in continuous mode to reach a steady state (Fig. S.1). The amounts of the target proteins produced under steady states are given in Table 2. Interestingly, the amount of biomass-accumulated YFP in [kRFU/g_{DCW}] was ~10-fold higher for scYFP (intended for secretion) than inYFP (intracellular), even without accounting for the fraction secreted outside the scYFP cells. This observation suggests that due to the difference in the translation mechanism (on ER for scYFP, in cytoplasm for inYFP) and/or sequestration of the overproduced protein within the ER-Golgi-vesicle continuum, more of the protein could be synthesized when targeted for secretion. The practical significance of this observation is that targeting a protein for secretion allows more product to be obtained, at least in this particular case.

To determine how the amounts of the heterologous proteins correspond with gene expression levels, RTqPCR targeting the heterologous genes was carried out (Fig. 1). The expression level was comparable for the genes encoding secretory reporters – scYFP, SoA and TIG (in descending ranking). On the other hand, the inYFP-encoding gene's expression was significantly lower ($p < 0.05$), even though all the genes were cloned under the same promoter. Such a trend was identical irrespective of the internal calibrator used (*ACT1* or *SEC62*).

The conducted basic characterization indicated that our set of strains covers: i) scYFP – a highly expressing and efficiently producing strain, with no burden on disulfide bond formation or glycosylation (high-secretor strain; HSS), ii) inYFP – a moderately-producing strain, with significant limitations at the transcription level; the limitation cannot be caused by overloading the secretory pathway but some other mechanisms, iii) SoA – a strain with a high expression level and very high

Table 2

Amounts of r-Prototypes synthesized by *Y. lipolytica* strains under steady-state conditions in continuous cultures.

Strain	Fraction	Amount of Protein in [AU/L] or [RFU/L]	Amount of Protein in [AU/g] or [RFU/g _{DCW}]
SoA	Supernatant	729.42 ± 11.20 [AAU/L]	96.33 ± 6.55 [AAU/g _{DCW}]
TIG	Supernatant	1 069.83 ± 5.54 [GAU/L]	75.43 ± 6.38 [GAU/g _{DCW}]
inYFP	Biomass	130 809.60 ± 5201.20 [kRFU/L]	10 872.85 ± 810.41 [kRFU/g _{DCW}]
	Biomass	812 108.30 ± 5715.62 [kRFU/L]	109 519.42 ± 7 523.53 [kRFU/g _{DCW}]
scYFP	Supernatant	24 785.28 ± 187.33 [kRFU/L]	3 342.31 ± 53.80 [kRFU/g _{DCW}]

production and secretion (HSS) of the protein, possibly burdened at the oxidative folding level due to -S-S- bond formation; iv) TIG – a strain with decent expression and production levels; burdened at the glycosylation level due to the protein's biochemical traits. How the perturbations introduced by the overproduction of the heterologous proteins impact the cellular response at the molecular level was further investigated through global transcriptional profiling. The analysis was conducted in both explorative and targeted mode, to study a general, global reaction of the cell, as well as to gain an insight into the processes taking place in the translational-secretory machinery.

3.2. Global transcriptome profiling – general observations

Quantification of raw counts for individual genes in the analyzed total RNA samples showed slightly different distributions, but none of the samples differed significantly from the rest (Fig. S.2). Moreover, biological replicates were characterized by high similarity in this regard. In terms of the overall similarity of the transcription profiles compared in terms of normalized read counts, strains expressing SCYFP and SOA were clustered together, while those expressing TLG and INYFP were more distant, generating a subcluster with some similarity to the control strain, located in the middle of the heatmap (Fig. S.3). The shape and local densities of volcano plots generated for the four strains (Fig. S.4) well illustrated those observations.

Primarily, the high dissimilarity of scYFP's and inYFP's transcriptomic profiles depicted in Fig. S.3 and Fig. S.4.a.b indicates that targeting a polypeptide for secretion induces massive rearrangements within the cell, exceeding the differences associated with biochemical properties of the proteins. Intracellular overproduction of YFP resulted in 327 DEGs vs the control strain, while its secretion resulted in 2199 DEGs, which is almost 7-fold more (Fig. 2). The high total number of DEGs in the SoA-producing strain (3127 DEGs; Fig. 2) suggests that high overproduction of a protein with a large number of Cys triggers substantial, multidirectional changes in the host cell. Expectedly, the number of DEGs was the lowest for inYFP-producing strains. The similarity of TIG's transcriptional profile with that of inYFP may suggest that the maturation and secretion of such a polypeptide – highly glycosylated with low number of disulfide bonds – it less demanding. To gain more detailed insight into the specifics of transcriptional profiles generated by each of the strains, we aimed to test these hypotheses.

3.3. Global response of the cell – explorative analysis

3.3.1. Biological processes and genes significantly responding to the reporter proteins' overproduction

First we searched for biological processes that were up- and down-regulated upon overproduction of a specific protein type from our set. The transcriptional profiles of TIG- and inYFP-producing strains were enriched in upregulated genes involved in inorganic ion homeostasis (copper, calcium, zinc and iron; with the most up-regulated gene across the whole transcriptome – zinc-iron transporter *F15411* g; Fig. 3), import into the cell, and fungal-type cell wall organization (Fig. 4.c.g). Microscopic observations demonstrated that the strains inYFP and TIG grew in the chemostat culture predominantly as ovoid-shaped cells, in contrast to the control and SoA strains, for which the dominant morphotype was filamentous and elongated ovoids (Fig. S.5). Since all the strains were maintained under the same conditions, it seems that the morphotype was dependent on the introduced variable – the type of overproduced protein and the resulting cellular response. No significant enrichment in the cell wall organization, morphogenesis, or cell shape control in SoA was recorded in the transcriptomic data, which indicates similarity in these biological processes for this and the control strain (corresponding to microscopic images). On the other hand, *FLO11* (*F19030g*), involved in filamentous and invasive growth, was amongst the most down-regulated genes in the SoA and scYFP strains' transcriptome Fig. 3, which differentiates SoA and control strains in terms of the cell shape

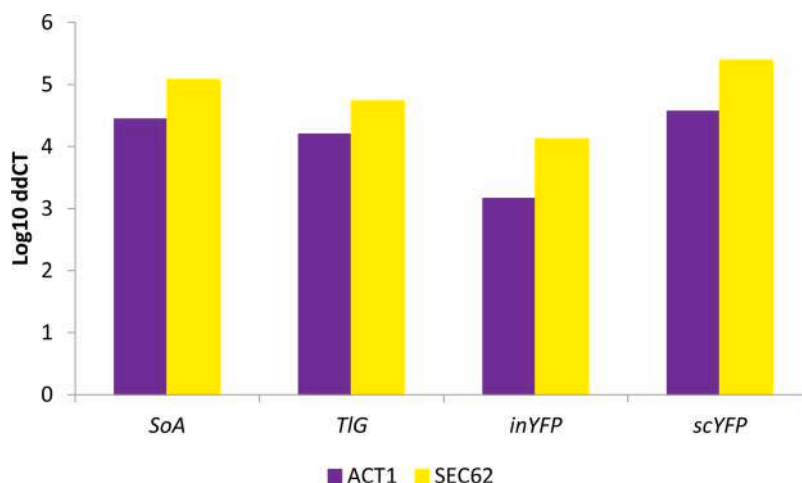


Fig. 1. Relative expression level of the genes encoding heterologous r(s)-Prots in *Y. lipolytica* strains. X axis: Reporter gene (*SOA*, *TLG*, *SCYFP* and *INYFP*). Y axis: Log_{10} ddCt – relative quantitation value normalized to *ACT1* (purple) or *SEC62* (yellow) used as internal calibrators. (For interpretation of the references to colour in the Figure, the reader is referred to the web version of this article).

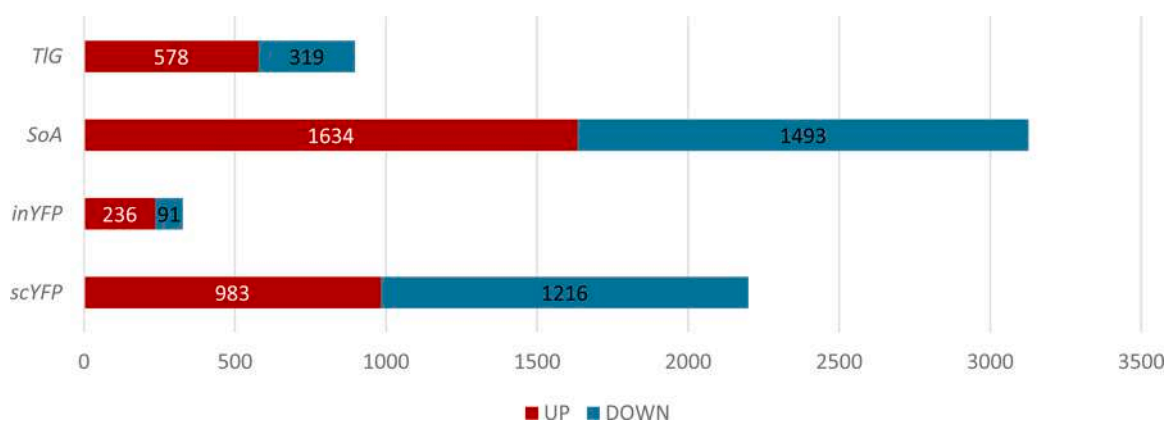


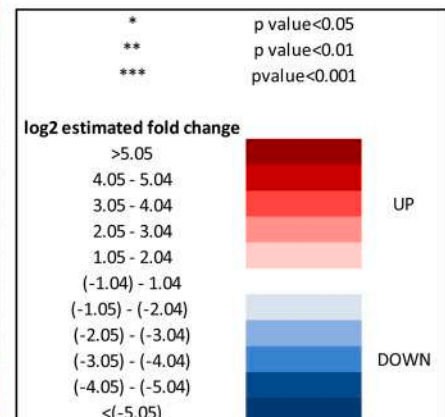
Fig. 2. A total number of DEGs determined for each recombinant strain vs the control strain. X axis: number of DEGs: red fields – upregulated DEGs; blue fields – downregulated DEGs. Y axis: recombinant strains producing different r-Prots. (For interpretation of the references to colour in the Figure, the reader is referred to the web version of this article).

control. It thus was an open question whether this high downregulation of *FLO11* is actually a hallmark of the onset of dimorphic transition to ovoid morphotype. In the second most active HSS (*scYFP*), which grew in both ovoid and very long filamentous forms (Fig. S.5), the cell shape regulation was the most enriched downregulated biological process (Enrichment fold 5.31; Fig. 4.b). This process was enriched in small GTPases, members of the Rho family (*RHO1*, *RHO2*, *RHO3*, *RHO4*), or *RAC1*, and *RAX2*, responsible for bud site selection. Interestingly, amongst the downregulated genes assigned to this biological process in the *scYFP*, we identified *SOK2* (*B19602g*; similar to *MGF1*), which is a negative regulator of pseudohyphal growth, suggesting that at the moment of sample collection the cell population could be under dimorphic transition towards filaments. In addition, one of the most upregulated genes in *scYFP* and *SoA* codes for a GPI-anchored cell wall protein involved in cell wall organization and induction of hyphal growth (*E02024g*) (Fig. 3, Fig. 5.b). On the other hand, mycelial growth factors *MGF2* (*B19602g*) and *D02189g*, mediators of RNA polymerase II involved in the regulation of the transition between yeast and filamentous forms, were downregulated in both *SoA* (-2.41/-1.91) and *scYFP* (-1.38/-2.12) (Fig. 5.c), which supports the hypothesis, conceived based on *FLO11* downregulation and the gene set enrichment analysis, that the highly producing cells, representing both filamentous and ovoid morphotype, were under dimorphic transition towards ovoid shape. A high response of the “cell wall biogenesis” biological process was previously

observed for *S. cerevisiae* overproducing r-Prot [2]; however, in that study, the researchers reported upregulation of invasive, filamentous growth, which was associated by the authors with increase in the traffic of secretory vesicles to the membrane. According to those authors' argumentation, filamentous or pseudohyphal cells have a high surface-to-volume ratio and inherently require higher Golgi-to-cell membrane trafficking rates to supply cell membranes and cell wall components for growth. Our observations concerning *Y. lipolytica* suggest that upon high r-Prot overproduction, the cells reprogrammed the morphotype to ovoid. Additionally, in terms of combining *Y. lipolytica* morphotype and r-Prot overproduction, our recent study evidenced that ovoid cells are more active and efficient in secretory r-Prot synthesis compared to filamentous forms [44], but the molecular background of that observation was not known. Expression of *SOK2* is regulated by RNA polymerase II, and *MGF2*, *FLO11* and *D02189g* are the mediators of RNA polymerase II activity, which altogether provides a direct link between the gene overexpression and morphotype in *Y. lipolytica*. Those genes' operation makes an important contribution to our understanding of the relationship between heterologous gene overexpression and morphology in *Y. lipolytica*.

Morphological transition to a filamentous form is a known marker of stress in *Y. lipolytica* cells [45,46]. In the present study, filamentation of the control strain is not clearly understood, as the cells were maintained at high carbon/nutrient provision and at a pH that should not promote

Gene name	<i>Yali0_</i>	inYFP	scYFP	scSoA	scTIG
Zinc-iron transporter	<i>F15411g</i>	***	***	***	***
DnaJ domain-bearing Hsp40	<i>E27588g</i>	**	**		
Peroxisomal PXMP2/4	<i>E12881g</i>	***	***	***	***
Aquaporin	<i>F01210g</i>		***		
GATA-type domain	<i>E31757g</i>		***	**	
Similar to <i>SHQ1_Ser</i> -Glu-rich	<i>D13200g</i>		**		
Ankyrin domain-containing	<i>D15070g</i>		**		
Acid phosphatase <i>PHO2</i>	<i>C19866g</i>		***	***	
GPI -anchored cell wall protein/EUTQ	<i>E02024g</i>		***	***	
<i>B01100g</i>	<i>B01100g</i>		**		
<i>VPS13</i> family protein	<i>A03817g</i>		***	***	
DAUER up-regulated-related	<i>F22187g</i>		***	***	
DeoR-type TF	<i>E19030g</i>		***	***	
OTU deubiquitinating enzyme	<i>C03091g</i>		**		
Cytoskeleton-associated	<i>A10945g</i>		*		
Similar to <i>CWP2</i> mannoprotein	<i>E33891g</i>	***	***	***	
Potassium transporter	<i>C10311g</i>		***	**	
Similar to <i>USO1</i>	<i>D23947g</i>		***	***	
Aspartyl protease <i>YPS3</i>	<i>D01331g</i>		**		
Membrane protein signalling	<i>B02112g</i>		**		
<i>SSA8</i> chaperone	<i>D22352g</i>		***	***	
Catalase <i>CTT1</i>	<i>E34265g</i>		**	***	
Catalase <i>CTT1</i>	<i>E34749g</i>		***	***	
<i>E08085g</i>	<i>E08085g</i>		***	***	**
Cell wall mannoprotein <i>CCW12</i>	<i>B21450g</i>	**	***	***	***
Fungal. Nuclear Protein	<i>D17578g</i>		***	***	
ATP-binding transporter <i>STE6</i>	<i>A11473g</i>		*	***	
Glyoxalase	<i>F00682g</i>		***	***	***
<i>E32263g</i>	<i>E32263g</i>		***	***	
ATP synthase F0-like protein	<i>F26917g</i>		***	***	
Formate dh <i>FDH</i>	<i>E14256g</i>		***	*	
F-box domain-containing protein	<i>A04059g</i>		***	***	
Ferric/cupric metalloredutase	<i>F11825g</i>	***	***	***	***
Aspartic-type endopeptidase	<i>B00374g</i>		***	***	
Zn(2)-C6 fungal-type domain TF	<i>D01353g</i>	***	***	***	***
Filamentation TF	<i>D02189g</i>		***	***	***
RsfA TF	<i>E26763g</i>		***	***	
<i>GAP1</i> General AA permease	<i>B19492g</i>	***	***	***	***
Maltose permease <i>MPH2</i>	<i>B00396g</i>		*		
<i>XBPI</i> TR	<i>F16511g</i>	***		***	***
<i>FLO11</i> TF	<i>F19030g</i>	***	***	***	***
<i>C23518g</i>	<i>C23518g</i>		***	***	**
Dioxygenase	<i>A06974g</i>		***	***	**
Conserved in <i>Yarrowia</i> clade	<i>C07744g</i>	***	***	***	***
<i>B19866g</i>	<i>B19866g</i>		***	***	
Chromatin binding	<i>B05698g</i>			**	
Uracil catabolism protein <i>URC4</i>	<i>E20647g</i>		***	**	
Purine permease <i>FCY2</i>	<i>F30569g</i>			***	
polyA transcript	<i>B23496r</i>			**	
Amino acid transporter <i>UGA4</i>	<i>D00495g</i>		***	***	
Triacylglycerol lipase	<i>D23419g</i>	***	***	***	
PBP2 PAB1 binding protein	<i>C19558g</i>	**	***	***	***
Purine permease <i>UAPC</i>	<i>E27852g</i>		***	***	
GTP cyclohydrolase <i>URC1</i>	<i>D08250g</i>		***	***	**
Oligopeptide transporter <i>OPT2</i>	<i>F09691g</i>	**	***	***	**
Nitrogen regulatory protein <i>AreA</i>	<i>D20482g</i>		***	***	
Urea transporter <i>DUR3</i>	<i>B04202g</i>	**	***	***	***
Ammonium transporter <i>MEP2</i>	<i>F12925g</i>			**	
Ethanolamine utilization <i>EUTQ</i>	<i>D02453g</i>		***	***	
GPI mannosyltransferase	<i>F25993g</i>			***	
Ammonium transporter <i>MEP2</i>	<i>E27203g</i>	***	***	***	
Metallothionein <i>MTP2</i>	<i>C20066g</i>				***
Metallothionein <i>MTP2</i>	<i>C20060g</i>			*	***



(caption on next page)

Fig. 3. Heatmap of the most responsive DEGs in the TIG-, inYFP-, scYFP- and SoA-overproducing strains. Shortened YALI_ code and a short description of a given gene are shown. Assigned names and functions were manually curated for all the ambiguous hits by blasting the sequence against UniProt and NCBI non-redundant protein database, and inferring the most accurate match. Red (up-regulated) and blue (down-regulated) squares were used to represent the transcription profiles of genes expressed as \log_2 Estimated fold change. Indicated genes meet the requirement of > 2.0 fold change in at least one of the overproducing strains. Abbreviations: dh - dehydrogenase, TF - Transcription Factor, TR - Transcription Repressor. (For interpretation of the references to colour in the Figure, the reader is referred to the web version of this article).

filamentation (pH 5.5; filament formation is induced at pH close to 7.0; Gorczyca, Kaźmierczak et al., to be published). On the other hand, both HSSs were overexpressing *CTT1* (*E34749g*) encoding a peroxisomal catalase, which is a known, ubiquitous marker of an ongoing cellular stress response (Fig. 3, Fig. 5.b). Its expression was over 13- and 6-fold lower in a filamenting control strain compared to HSSs. Hence, it suggests that filamentation in the control strain was not associated with the general stress response but was probably caused by some other factors. Formation of filaments by the SoA-overproducing strain and the above discussed dimorphic transition ongoing in *scYFP* are more straightforward to explain. In this regard, *CTT1*, which plays a key role in the oxidative stress response and in protecting proteins against oxidative inactivation, was one of the most responsive, overexpressed DEGs in HSSs. It is known that overproduction of rs-Prot is inherently associated with oxidative stress due to energy- and material-consuming oxidative folding taking place in the ER [30]. Additionally, transcriptomic profiles of *SoA* and *scYFP* were characterized by high overexpression of glutathione-independent glyoxalase Hsp31-related (*F00682g*; Fig. 3; Fig. 5.b), which plays a role in detoxifying endogenously produced glyoxals [47,48] and is involved in protection against reactive oxygen species (ROS), similarly to *CTT1*. This further supports the statement on induction of oxidative stress in the HSSs. Notably, it has been demonstrated that oxidative stress is induced in the overproducing strains, even if the overproduced protein bears only a limited number of Cys residues [2,12,49], which relates to the *scYFP* case.

In contrast, upon TIG overproduction, the most downregulated genes were metallothioneins (*C20060g* and *C20066g*; Fig. 3; Fig. 5.c), which are cysteine-rich heavy-metal-binding proteins, localized to the membrane of the Golgi apparatus and cytoplasm, serving as the primary regulators of trace metal homeostasis in yeast [50]. Their principal role is protection against metal toxicity and oxidative stress (as *CTT1*), and regulation of zinc and copper pools. Biosynthesis of metallothioneins in yeast increases by several-fold throughout oxidative stress. Their downregulation is in high contrast to the most upregulated zinc-iron transporter (*F15411g*); and such a co-expression pattern is not fully understood. Likewise, one of the most upregulated genes in *SoA* (*F11825g*; Fig. 3) encodes a ferric/cupric metalloredutase responsible for reducing extracellular iron and copper prior to import. In each of the studied transcriptomes, genes involved in “metal ion homeostasis and import” head the list of upregulated DEGs, highlighting the significance of this biological process upon enhanced protein production.

The Rax2 protein mentioned above, responsible for bud site selection, was also assigned to several significantly enriched downregulated biological processes in *scYFP* (Fig. 4.b), including cytoskeleton organization (Enrichment fold: 1.95), and cytoskeleton-dependent cytokinesis (Enrichment fold: 3.54), which was composed of cytoskeletal proteins, actin, tubulins, actin binding motor proteins, GTPase-activating protein and tropomyosins. Their downregulation in the intensively secreting strain is not currently clear, except for the hypothesis on the HSSs' transition to ovoid morphotype. On the other hand, high upregulation of *USO1* (*D23947g*) in both HSSs suggests enhancement of the intracellular protein transport (Fig. 3; Fig. 5.b) and *USO1*'s dominant role in that biological process. *Uso1* is an essential protein involved in vesicle-mediated ER to Golgi transport; it binds membranes, acts during vesicle docking to the Golgi, and is required for assembly of the ER-to-Golgi SNARE complex.

The cell wall organization process (Enrichment fold: 1.94) was significantly downregulated in the *scYFP* strain. Manual inspection of

the genes assigned to this process revealed several glycosidases, chitin synthase, phosphoacetylglucosamine mutase, glucosamine-6-phosphate isomerase and N-acetyl transferase. These genes were clustered together with mannose-6-phosphate isomerase, cell wall mannoprotein cis3-related (note that fungal-type cell wall organization was upregulated in the ovoid-shaped *inYFP* and *TIG* strains) and mannosyltransferases, including *KTR1*, *KTR5*, and *MNN2* (Fig. 4.b). Although the mannosyltransferases were assigned to a cell wall organization process, it is known that apart from being involved in cell wall macromolecule biosynthesis, they play a role in glycosylation of proteins. Likewise, the individual gene (not a part of an enriched biological process) *F25993g*, encoding Golgi membrane-bound mannosyltransferase 2, involved in protein glycosylation and vesicular trafficking, was one of the most downregulated DEGs in the *SoA* strain (Fig. 3; Fig. 5.c). Downregulation of the secretory pathway constituents in the HSSs requires additional comment. Both the proteins are scarcely glycosylated (Table 1). It may be speculated that the prevalence of the non-glycosylated *scYFP/SoA* as cargos in the secretory pathway resulted in relative downregulation of mannosyltransferases versus the control strain, which were involved in the native polypeptide maturation in the latter. Indeed, manual inspection of the gene sets revealed that all the main secretory proteases and lipases, including *TGL3* (*D17534g*), *LIP4* (*E08492g*), *AXP1* (*B00374g*, *B05654g*) and *XPR2* (*A08360g*), were downregulated in *SoA*, and three of them – *XPR2*, *AXP1*, *LIP4* – were downregulated in *scYFP* (Fig. 5.a.c). These proteins differ in terms of number of glycosylation sites, but they bear on average 3 N-glycosylations and 8 O-glycosylations. Their downregulation improves the capacity of the secretory pathway, required for *SoA/scYFP* synthesis, but at the same time decreases the requirement for glycosylating enzymes. Furthermore, in relevance to the cell wall restructuring in HSSs, a gene encoding cell wall mannoprotein, *CCW12* (*B21450g*), was identified amongst the most upregulated genes, together with *STE6* (*A11473g*), which is an ABC-type peptide-transporting ATPase involved in the export of methylated and prenylated α -factor pheromone [51] (for *scYFP* – only upregulated; not “the most responsive”; Fig. 3). While it has been postulated that *Ste6* is specific towards α -pheromone, its closest mammalian homologue encodes a multidrug resistance (*MDR1*) channel, involved in excretion of numerous toxic compounds. It can be speculated that upon high oxidative stress in the HSS overproducing cells, many toxic compounds are generated and are pumped out of the cell via the *Mdr1/Ste6* membrane channel. More details on cell wall biogenesis-related genes are given hereafter.

Another significantly downregulated biological process in TIG-/inYFP-producing strains was associated with ribosome biogenesis, rRNA processing and organic acid metabolic (catabolic) processes (Fig. 4.d.h), which was again rather surprising, considering the ongoing overproduction of heterologous proteins. Still, those biological processes were only slightly down-regulated, since when conducting the enrichment analysis with the gene sets filtered for 2-fold change, they were not identified as significant. In contrast, transcriptional profiles generated upon overproduction of *SoA* and *scYFP* were significantly enriched in upregulated genes involved in ribosome biogenesis and assembly, transcription, translation, regulation of DNA-templated transcription, chromatin modification, histones synthesis and modification, which is straightforward to explain by intensive ongoing transcription and translation (Fig. 4.a.e). High downregulation of *E26763g* (*RSFA*) in *scYFP* (–2.14) and *SoA* (–5.39) strains (Fig. 3; Fig. 5.c), which is a ribosomal silencing factor [52], well corresponds with those observations.

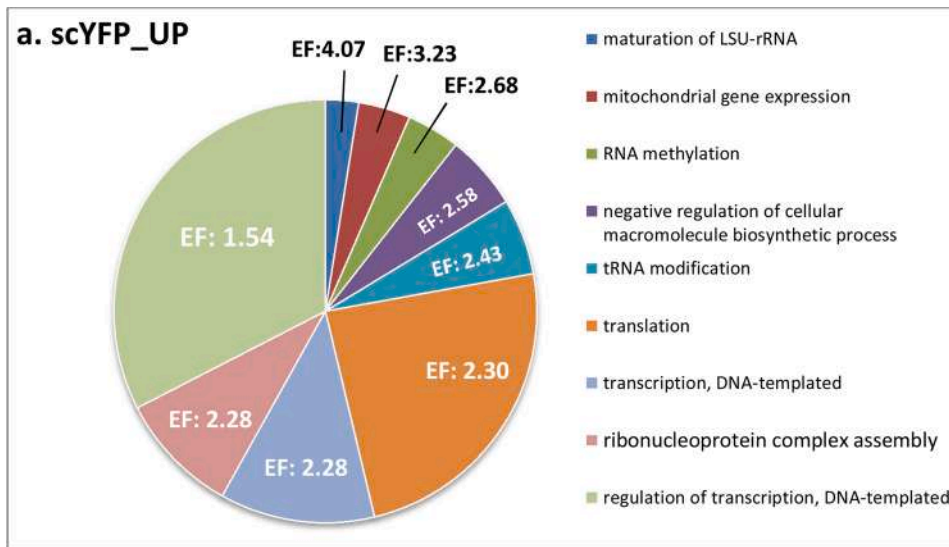


Fig. 4. Pie charts showing significantly enriched biological processes determined via statistical overrepresentation test conducted using Panther tool on the complete DEGs lists. scYFP_UP (a.), scYFP_DOWN (b.), inYFP_UP (c.), inYFP_DOWN (d.), SoA_UP (e.), SoA_DOWN (f.), TIG_UP (g.) and TIG_DOWN (h.). Up- and down-regulated DEGs identified for each strain were analyzed separately to reveal up- and downregulated biological processes. Only biological processes enriched at significance level $p < 0.05$ are shown. The results were manually curated to reduce redundancy of several biological processes. EF – Enrichment fold. (For more convenient visualization of the Figure in color, the reader is referred to the web version of this article).

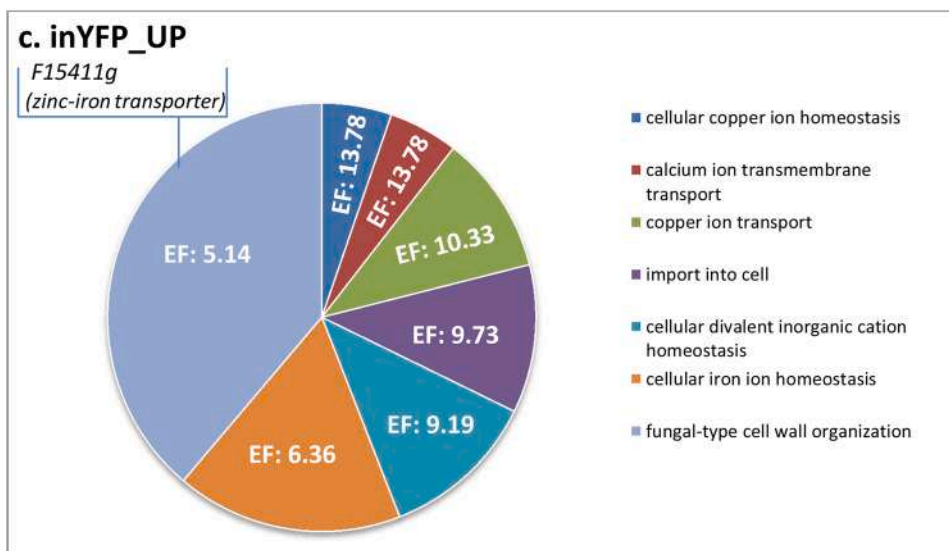
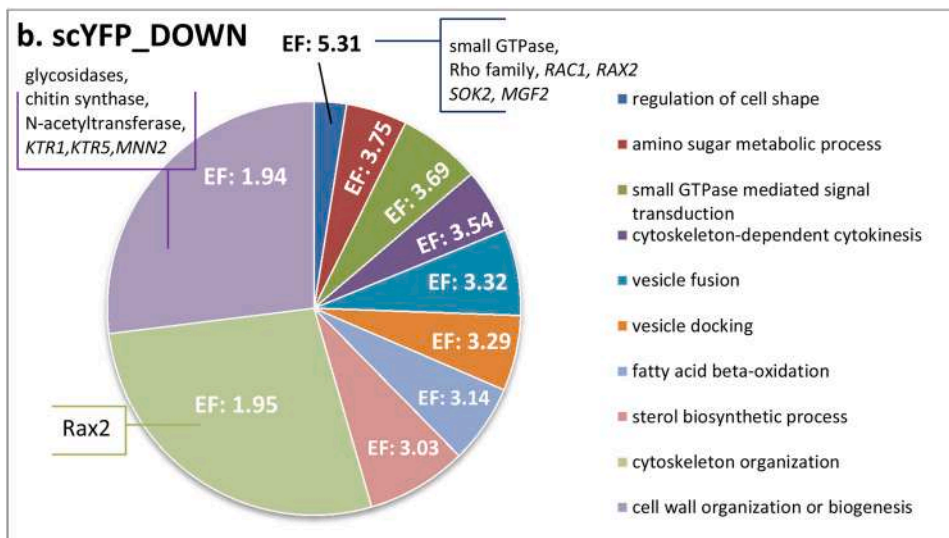


Fig. 4. (continued).

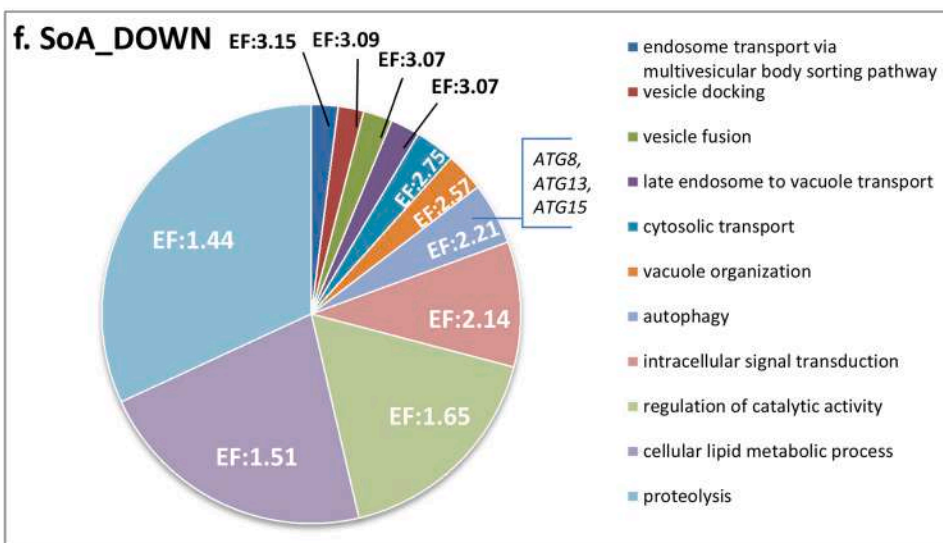
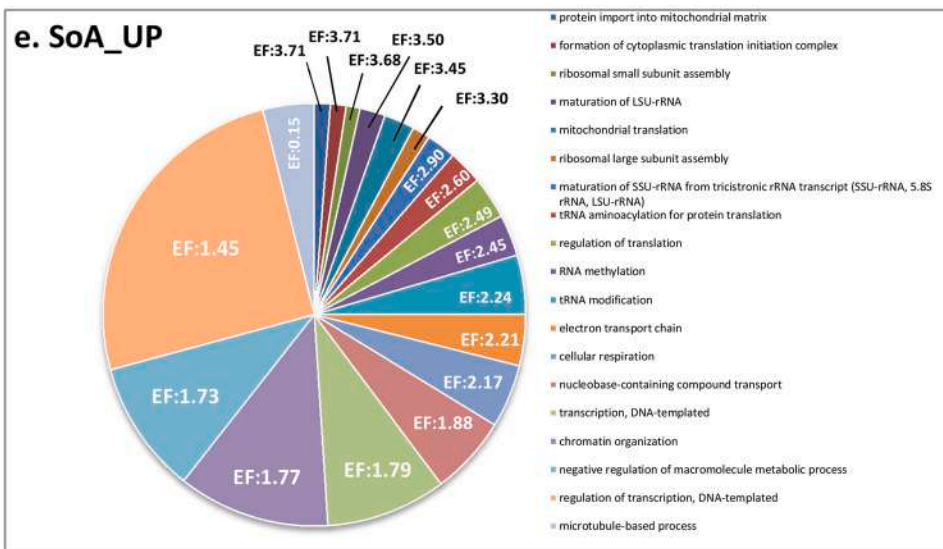
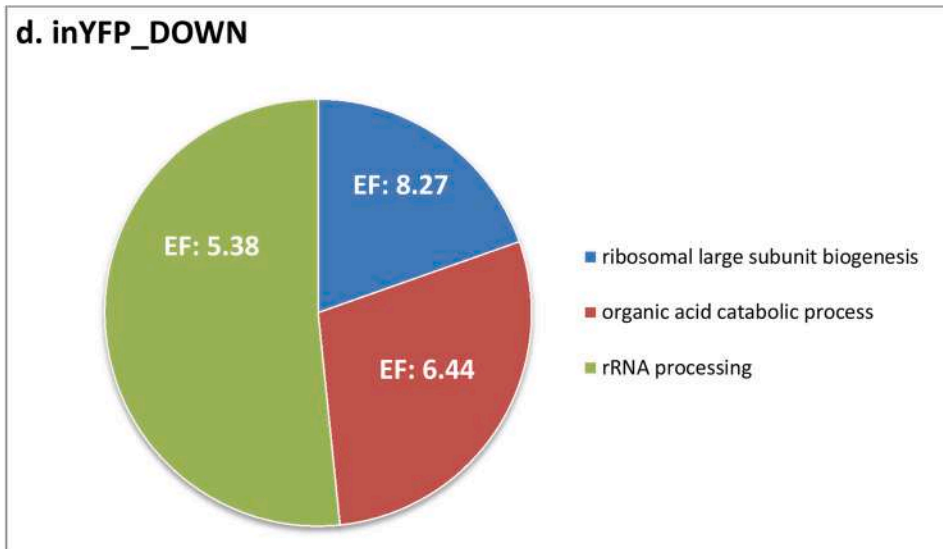
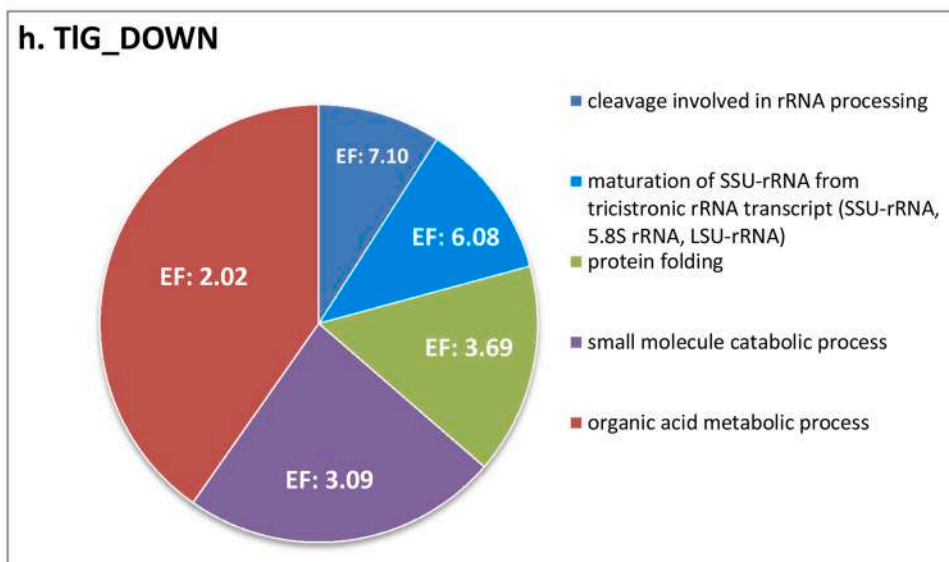
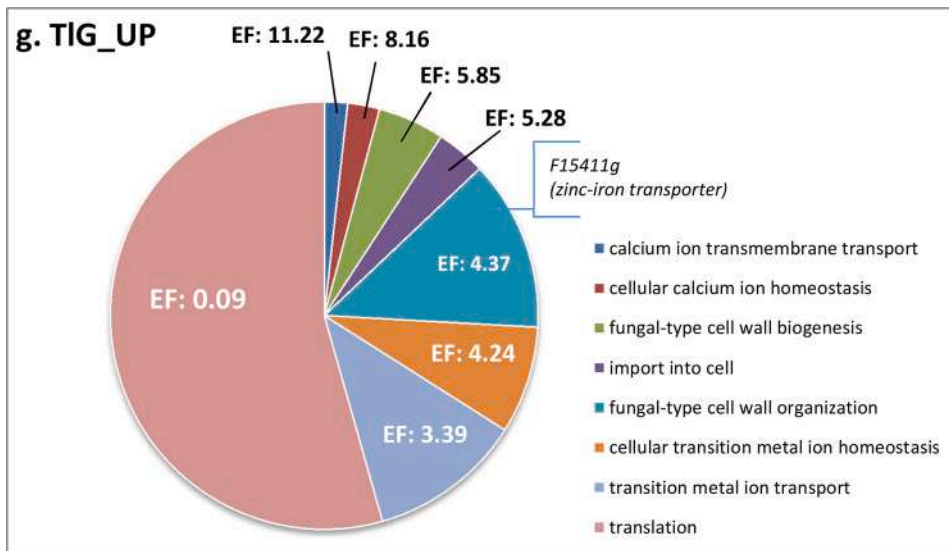


Fig. 4. (continued).



RsfA helps cells to adapt to slow-growth/stationary phase conditions by down-regulating protein synthesis, which is one of the most energy-consuming processes. Its downregulation releases ribosome assembly from inhibitory action of RsfA, which is of key importance for the overproduction of heterologous proteins. Intriguingly, *RSFA* was not significantly downregulated in *TIG* and *inYFP*, which were also overproducing r-Prot. In this regard, it was previously observed that overproduction of a large r-Prot resulted in repression of overall transcription and translation [2]. That observation contrasts with our results on *RSFA* and high enrichment in ribosome assembly and biogenesis process upregulation in HSSs. The authors explained that the repression is a general, broad-spectrum response used to adjust the expression rates to reduced capacity of folding and maturation. It was further explained that if the processes upstream oxidative folding are not downregulated, the risk of severe oxidative stress and excessive consumption of reducing equivalents is high [2]. Considering the high upregulation of oxidative stress response genes in HSSs studied here, it is possible that unrestricted high transcription and translation led to imbalanced flow through the secretory pathway. Results on limited expression level of *INYP* and *TLG* (Fig. 1) together with lack of downregulation of *RSFA* and no oxidative stress response outburst in these strains collectively support that statement. However, it remains

unclear why the defensive processes were not implemented in HSSs; and it requires further deepened studies.

Within this subject, amongst significantly upregulated biological processes in HSSs, we identified “negative regulation of cellular macromolecule biosynthesis” (Fig. 4.a.e). In our previous work we postulated that upon high overproduction of a rs-Prot, transcription of its gene is repressed to avoid overloading the secretory pathway [32], as was previously reported for *S. cerevisiae* [2]. In the present research, we observed that the biological process “transcription” is upregulated in the HSSs, but one of the key genes assigned to that process is specific towards polymerase III transcribing small RNAs required for translation, cell growth and the cell cycle (and not mRNAs). The responsible gene (*F10541g*; *MAF1*) is a conserved repressor of RNA polymerase III, so its upregulation results in decreased transcription of the governed genes (Fig. 5.b); which, in this case, corresponds to previous observations [2] made for *S. cerevisiae*.

Vacuole organization, autophagy and proteolysis (Enrichment fold: 2.57, 2.21, 1.44, respectively) were significantly downregulated biological processes in *SoA* (Fig. 4.f; only a tendency observed for *scYFP*); including such genes as *ATG8*, *ATG13*, and *ATG15* (autophagy-related proteins). Autophagy, through the sequestration and delivery of cargo to the lysosomes, is the major route for degrading long-lived proteins and

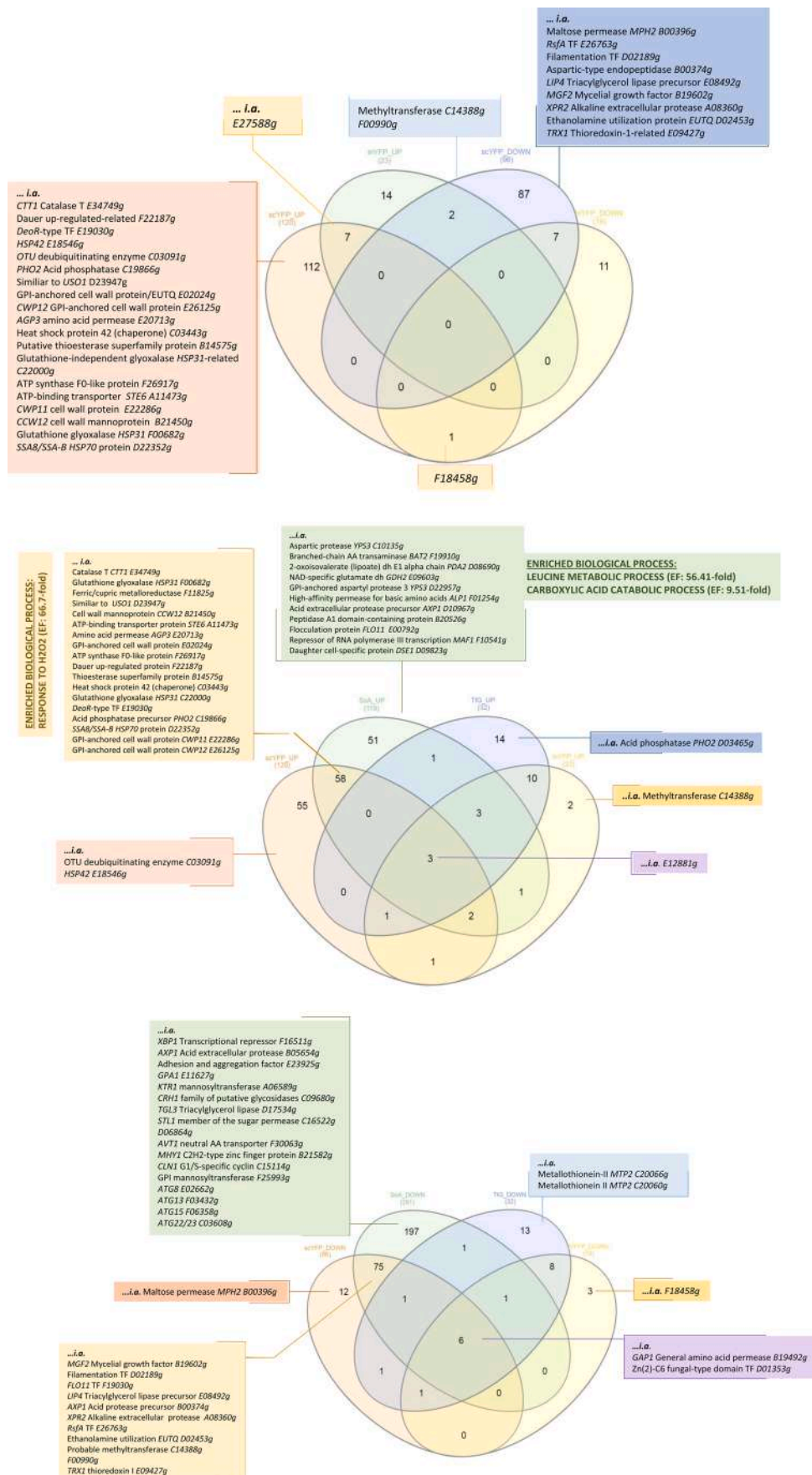


Fig. 5. Venn Diagrams analysis illustrating shared and dissimilar genes/gene sets for individual strains. The numbers indicate the number of shared or strain-specific genes, up/down-regulated in scYFP- and inYFP- producing strains (a.); up-regulated genes in SoA-, TIG-, scYFP- and inYFP-producing strains (b.); down-regulated genes in SoA-, TIG-, scYFP- and inYFP-producing strains (c.). Venn diagrams analysis was done with the DEGs lists restricted to 2- fold change. Shortened YALI code of the gene and a short description is given only for selected genes – full lists are given in **Figure S.6**. Significantly enriched biological processes within generated gene groups identified *via* GSA are given in **BOLD** and underlined (found only for two gene sets). Abbreviations: dh - dehydrogenase, TF - Transcription factor, TR - Transcription Repressor, AA- Amino Acid. (For more convenient visualization of the Figure in color, the reader is referred to the web version of this article).

cytoplasmic organelles [53–55]. Formation of the autophagic vesicles requires the recruitment of Atg8 (downregulated in HSSs), a ubiquitin-like protein, to the membrane of the nascent autophagosomes. Their downregulation suggests limited protein degradation in comparison to the control strain. Additionally, one of the most upregulated genes in *scYFP* was *C03091g*, associated with protein degradation. It bears high similarity to OTU domain-containing protein acting as thiol-dependent deubiquitinating enzyme [56] (Fig. 3; Fig. 5.a). The transcriptome of SoA-overproducing strain was enriched in an upregulated panel of peptidase A1 domain-containing proteins, encoded by *B20526g*, *C10135g*, *D10967g*, and *D22957g* (Fig. 5.b). The former three peptidases are extracellular, membrane-unbound, aspartic-type endopeptidases that play a role in extracellular proteolysis. The latter peptidase is Yps3, a GPI-anchored aspartyl protease which is involved in fungal-type cell wall organization and may also play a role in the maturation of GPI-mannoproteins associated with the cell wall. *YPS3* was also one of the most upregulated DEGs in *scYFP* (Fig. 3). Yps3 belongs to the yapsin family of known, negative impact on the stability of secretory proteins [57]. Disruption of yapsin-encoding genes [58] by combined knock-out of *YPS1/PEP4* increased the titer of the non-degraded product from ~40 up to 80 % [59]. In *Y. lipolytica*, the most popular expression platform strains, belonging to the Po1-series, typically bear disruption of one or both major extracellular proteases – *XPR2* and *AXP1*, but yapsins have not been considered for disruption to date. It would be interesting to test a similar strategy in *Y. lipolytica*, since now it is clear that the genes are highly expressed upon r-Prot production.

Additionally, within the scope of protein and amino acid turnover and transportation, we observed common significant downregulation of *GAP1* encoding a general amino acid permease (*B19492g*; Fig. 3; Fig. 5.c), which was surprising considering the high demand for the protein building blocks. In contrast, some other amino acid permeases were identified as upregulated within the strains' transcriptomes, including the amino acid permease *AGP3* and the high-affinity permease for basic amino acid import *ALP1*. On the other hand, genes involved in amino acid turnover and transformation were significantly upregulated in HSSs. *BAT2* (branched chain amino acid aminotransferase), *PDA2* (2-oxoisovalerate dehydrogenase), 3-hydroxy-3-methylglutarate-CoA lyase, and *GDH2* (glutamate dehydrogenase) constitute the central branch point in the amino acid turnover net [60,61]. They were all upregulated in the SoA-overproducing strain, indicating intensive amino acid metabolism in the cell (Fig. 5.b). Taken together, these results suggest that upon limited autophagy in the HSSs, extracellular proteolysis of peptone and YE components is catalyzed by the upregulated set of extracellular/membrane-bound proteases and results in high amino acid provision in the medium. Amino acids are taken up by the upregulated amino acid transporters (e.g. *AGP3* and *ALP1*) and the imported moieties are subjected to transformations by the amino acid turnover net, according to the requirements of the ongoing overproduction.

Finally, in terms of the general cellular response to a specific protein's overproduction, beta-oxidation and cellular lipid metabolic processes were downregulated in *scYFP* and SoA (Enrichment fold: 3.14 and 1.51), respectively (Fig. 4.b.f). Modulation of the lipid metabolism under heterologous secretory protein's overproduction could be caused by increased requirement for phospholipids for the intracellular membrane trafficking. It has been postulated that limitation of membrane availability may be the key limiting factor for efficient operation of the secretory pathway [62]. High upregulation of the *EUTQ* gene (*E02024g*; ethanolamine utilization protein) in the SoA producer and downregulation of the putative ethanolamine utilization protein *D02453g* (also in *scYFP*) are direct indications of intensive turnover of phospholipids ongoing in the cell (Fig. 3). Manual inspection of the biological processes' representatives showed that the downregulated lipid metabolism gene set in SoA consists mainly of numerous ER-located membrane-bound proteins, including ceramide very long chain fatty acid hydroxylase *SCS7*, diacylglycerol O-acyltransferase, GPI

mannosyltransferase 4, GPI anchor transamidase, dolichyl glycosyltransferase (N-glycosylation), GlcNAc-1-P transferase (playing a role in N-linked glycosylation), dehydrolipoyl diphosphate synthase (which is a prenyltransferase involved in protein glycosylation), or diacylglycerol O-acyltransferase 1. Although being assigned to the downregulated lipid metabolism, these genes encode ER-/Golgi-membrane proteins involved in protein maturation. The low number of predicted glycosylation sites and the presumed abundance of SoA/*scYFP* as cargos in the secretory pathway may account for such observations. Still, some fatty acid metabolic process-related genes were included in that downregulated group (in SoA), including peroxisomal acyl-CoA oxidase, and enoyl-CoA isomerase 2 (Fig. 4.f). Likewise, the *scYFP*-producing strain showed downregulation in lipid metabolism, including peroxisomal multifunctional enzyme type 2, acyl-coenzyme A oxidase 2 and 3, enoyl-CoA hydratase and isomerase (Fig. 4.b). Furthermore, the cells overproducing SoA (and *scYFP*, but to lesser extent) were also particularly active in cellular respiration and generation of energy associated with upregulated protein import into the mitochondrion, and mitochondrial gene expression (Fig. 4.a.e). In this regard, amongst the most responsive upregulated DEGs we identified *F26917g*, which encodes ATP synthase F0-like protein. Increased energy demand for the r-Prot overproduction easily explains that observation (Fig. 3). As demonstrated previously [2], the secretory stress shifts metabolism to increased oxygen and ATP requirements.

3.3.2. Similarities and dissimilarities in the global gene response profiles of *Y. lipolytica* to different proteins' overproduction

Venn diagram analysis is useful for identification of common and distinct genes and processes characterizing a given transcriptomic profile. Depending on the input gene sets, it is possible to identify molecular objects "hallmarking" a given biological process. In this research we aimed to identify genes specifically responding to protein secretion (comparison of *inYFP* vs *scYFP*; Fig. 5.a) and marking intensified oxidative folding (SoA vs other) or glycosylation (*TIG* vs other) (Fig. 5.b.c). A complete list of genes (accompanied by a short description and YALI name) used for Venn diagrams analysis is available in the Supplementary material (Fig. S.6).

Systematic comparison of DEGs sets of *inYFP* and *scYFP* was expected to reveal markers of high overproduction of intracellular and secretory r-Prot in *Y. lipolytica*. Such genes were initially searched within the groups of strain-specific and differential expression direction-specific genes ($n = 112$, $n = 14$, $n = 87$, $n = 11$; Fig. 5.a), but no enrichment in any specific biological process could be found there. Interestingly, transcriptomes of *inYFP* and *scYFP* shared 7 common upregulated genes, including *E27588g*, which at the same time is the second most upregulated gene in *inYFP* (Fig. 3, Fig. 5.a). The gene encodes a DnaJ-domain-bearing Hsp40 chaperone regulating the ATPase activity of Hsp70, involved in folding of nascent proteins, translocation of polypeptides across organelle membranes, coordinating responses to stress, and targeting selected proteins for degradation. Its specific, high upregulation in both YFP-producing strains may indicate its specific role in folding of non-post-translationally modified proteins. Then we sought for markers of high overproduction of the intracellular and secretory r-Prot in the common gene groups but of opposite regulation in *in-/sc-YFP*. A single gene, *F18458g*, was downregulated in *inYFP* and upregulated in *scYFP*. The *F18458g* gene after the blastp search showed weak similarity to the bifunctional transcriptional activator/DNA repair enzyme Ada, bearing an assigned function of methylguanine-DNA cytosine methyltransferase in *Clavispora lusitanae*. Due to the lack of a more specific functional annotation it is impossible to infer its role in the studied processes. Amongst the genes of the opposite pattern of regulation (up in *inYFP*, down in *scYFP*) we identified a methyltransferase (*C14388g*) and a GPI-anchored cell wall protein (*F00990g*) involved in cell wall organization and hyphal growth that is induced specifically in response to hyphal growth (Fig. 5.a). The latter again confirms our hypothesis that the HSSs were undergoing a dimorphic transition to ovoid morphotype.

The conducted Venn diagram analysis revealed that the only significantly enriched biological process common for upregulated genes in HSSs ($n = 58$; Fig. 5.b) was the “response to hydrogen peroxide” (Enrichment fold: 66.7). Aside from that, manual inspection of the list of common genes allowed us to identify the uncharacterized gene *F22187g*, which was amongst the most upregulated genes in both HSSs (Fig. 3). Basic search demonstrated its similarity to another uncharacterized gene (*NCU09057*) from *Neurospora crassa*. Establishing its putative function was difficult, indicating lack of interest in this protein to date. A blastp search against the GRYC database revealed *F22187g*'s similarity with (depending on the region of similarity): 1) glutaredoxin (thioltransferase; belonging to the LEA type 4 family; 28 % within a range of 18–634 amino acids) from *Planoprotostelium fungivorum*, 2) protein LEA-1, isoform j from *P. fungivorum*, 3) LEA protein DR_1172 from *Deinococcus radiodurans*, or 4) CsbD domain-containing protein, being an integral membrane component in *Williamsia* spp. On the other hand, in the Panther database, *F22187g* was assigned to a dauer upregulated-related protein family. The term “dauer” denotes a growth arrest stage in a model nematode (*C. elegans*) intensively studied in relation to longevity and aging processes. The dauer (growth arrest) gene family includes a number of different heat shock proteins (e.g. Hsp20, smHsps) and oxidoreductases (e.g. cytochrome P450, glutathionyl-, glucuronosyl/glucosyl-, short-chain- oxidoreductases), as well as UDP-glucuronosyltransferases and glutathione S-transferases [63]. It was proposed that broad spectrum detoxification and repair of damaged proteins constitute a molecular basis of the growth arrest and the consequent prolonged viability. Interestingly, the smooth ER was proposed to be a “longevity organelle”, where detoxification and repair take place. In relation to the current findings and the blast search results, it could be postulated that *F22187g* is an ER-resident thioltransferase, upregulated in SoA- and scYFP-overproducing strains, supporting homeostasis maintenance; however, it must be kept in mind that the function was assigned only based on the blast similarity search. In reference to the current study, growth arrest and upregulation of detoxification processes may be the fundamental biological processes accompanying intensive r-Prot overproduction, with the key role of ER. The other upregulated genes supporting this statement, and shared by scYFP and SoA, were thioesterase (*B14575g*), *HSP42* (*C03443g*) and glyoxalase *HSP31* (*F00682g*, *C22000g*) (Fig. 5.b).

Within the scope of the proposed mechanism of growth arrest associated with r-Prot overproduction, we observed intriguing, high downregulation of *F16511g* (–302; aka *YAL12_F00325g*; Fig. 3) specific for SoA (Fig. 5.c). *F16511g/F00325g* has an assigned role as an *XBPI* transcriptional repressor involved in G1/S transition of the mitotic cell cycle. In *S. cerevisiae* (*YIL101C*) *Xbp1* binds to RNA polymerase II-specific genes' promoter sequences of the cyclin genes. Manual inspection of SoA's DEGs showed significant downregulation of *CLN1* G1/S-specific cyclin (*C15114g*; promoting G1 to S transition), indicating growth arrest (G1→G0) (Fig. 5.c). That observation contradicts the possible involvement of *XBPI* in *CLN1*'s downregulation, at least, according to the known mechanism (its downregulation should increase expression of *CLN1*, and promote G1 → S transition). On the other hand, it is known that *CLN1* expression is also controlled by other transcriptional regulators – *Swi4/Swi6* and *Mbp1/Swi6*. However, *Y. lipolytica*'s homologs were not found within SoA-induced DEGs. Nevertheless, *CLN1*'s downregulation is a hallmark of cell transition to quiescence, maintaining G1 arrest. *XBPI*'s downregulation upon SoA overproduction is also interesting due to its other interactions, namely, its role in the unfolded protein response (UPR). Upon ER stress, *XBPI* mRNA is spliced by ER-resident *Ire1* (like *HAC1*, discussed below), thereby generating functional spliced *XBPI*, translocating into the nucleus to initiate transcriptional programs that regulate a subset of UPR- and non-UPR-associated genes. It is also known that *XBPI* expression is induced by different stress conditions. Its high downregulation in SoA is unclear, considering the high enrichment fold in cellular “response to hydrogen peroxide” in both HSSs. Taken together, downregulation of

XBPI accompanied by no transcriptional response from *HAC1* (discussed below), and downregulation of *CLN1*, could collectively suggest that HSSs entered the quiescence state rather than UPR. Upregulation of another transcription factor, *DEOR* (*E19030g*), in both HSSs (Fig. 3; Fig. 5.b) further supports that statement. *DeoR*-type transcription factors bearing a sugar-binding domain, which is also present in the enzymes of the phosphosugar isomerase family, are involved in regulation of sugar metabolism in bacteria [64], e.g. as a repressor of sugar import [65]. It can be speculated that significant *E19030g* upregulation is associated with downregulation of maltose permease *MPH2* (*B00396g*) in scYFP and sugar permease *STL1* (*C16522g*) in SoA (Fig. 5.c). Reduced transcription of some less important genes in order to supply enough energy and substrates for synthesis of r-Prot has also been suggested based on transcriptomics data from continuous cultivation of *P. pastoris* overproducing r-Prot [66].

No scYFP-specific up- or downregulated biological process was identified. On the other hand, SoA-specific upregulated genes ($n = 51$; Fig. 5.b) accounted for enrichment of the “leucine metabolic process” (Enrichment fold: 56.41) and carboxylic acid catabolic process (Enrichment fold: 9.51). Manual inspection demonstrated that the central metabolism of amino acid conversions and transformations was grouped in those enriched processes (including *BAT2*, *PDA2*, *MVA1*, *GDH2*, discussed above). While no particular biological process was enriched by specific ($n = 12/n = 197$) or common ($n = 75$) downregulated genes in SoA and scYFP (Fig. 5.c), what attracted our attention was a common downregulation *MGF2* transcription factor, backed up with SoA-specific downregulation of *MHY1*, *FLO11*, “adhesion and aggregation factor” (*E23925g*) involved in filamentation. Intriguing downregulation of the cytoplasmic thioredoxin *TRX1* (*E09427g*) (within a group $n = 75$), shared by HSSs, is discussed below.

While no *TIG*-specific biological process was enriched, we observed common upregulation of the *PHO2* acid phosphatase-encoding gene in all the strains overproducing rs-Prot (scYFP, SoA, *TIG*). *PHO2* is involved in scavenging of inorganic phosphorus required for energy storage and transfer. Detailed analysis of Fig. 5.b (and Fig. S.6) showed however that the same *PHO2* activity was encoded by different genes: 1) *C19866g* (highly similar to *D03465g*) in HSSs, and 2) *D03465g* in *TIG*. Importantly, *PHO2* was amongst the top most responsive upregulated DEGs in both HSSs (Fig. 3), which is logical in the face of high-energy demanding r-Prot overproduction. The difference between the two isoforms, and the reason for strain-specific upregulation, is not understood. It would be interesting to know whether enhanced burden on the secretory pathway is the actual causative agent for *PHO2* upregulation.

Interestingly, all the transcriptomes studied here were characterized by uniform high upregulation of an uncharacterized protein *E12881g* (SoA log₂ fold change: 2.6, scYFP: 2.6, *TIG*: 2.82, inYFP: 2.74) (Fig. 3). Blasting the polypeptide sequence of *E12881g* against the protein database indicated its ~30 % similarity to a putative membrane protein belonging to the peroxisomal *PXMP2/4* family in *Metschnikowia aff. pulcherrima*, *Ogataea parapolyomorpha* and *Komagataella phaffii*. *PXMP2/4s* are putatively involved in pore-forming activity and may contribute to the unspecific permeability of the peroxisomal membrane. Their uniform upregulation makes them an interesting target in further studies on DEGs significantly responding to heterologous protein overproduction, in general. On the other hand, all the strains studied here bear a common pattern of high, significant downregulation of the uncharacterized gene *D01353g*, encoding Zn(2)C6 fungal-type domain-containing protein (Fig. 3). The gene was downregulated to a very high degree in all tested strains: SoA (log₂ fold change: -273) scYFP (-215), *TIG* (-279), inYFP (-244). While the primary search in the GRYC database showed that *D01353g* is an uncharacterized zinc-binding transcription factor, blasting it against the UniProt database showed some similarities with a *ZCF32* transcription factor from *C. albicans*, which is involved in negative regulation of transcription by RNA polymerase II (specifically). Hence, its uniform downregulation in all the overproducing strains suggests release from transcription inhibition by the “ZCF32-like”

homolog under ongoing overexpression of inYFP, scYFP, SoA and TIG proteins in *Y. lipolytica*. Such a mechanism, of downregulation of the RNA polymerase II repressor, is additionally enhanced by the above-discussed E26763g (*RSFA*) (Fig. 3; Fig. 5.c) in scYFP (-2.14) and SoA (-5.39) strains.

3.4. Specific response of the secretory pathway to the heterologous secretory proteins' overproduction

The targeted analysis was conducted on a set of pre-selected genes of known involvement in protein maturation and the secretory pathway. The list of genes was developed based on previous evidence of their role in those biological processes in yeast [1,22,23]. We distinguished 27 categories, according to which the genes were classified (Table S.3). The normalized expression level of selected genes from that set is given in Fig. 6, whereas the relative expression ratio vs the control strain for all the genes assigned to this biological process (~ 700 genes) is shown in Table S.4. Such focused analysis allowed us to gain a deeper insight into fine changes in the secretory machinery triggered by overproduction of different rs-Protos. As previously discussed, the changes in mRNA levels are subtle and specific when induced solely by overproduced r-Protos [17].

Primarily, we saw that the key transcription factor *HAC1* (B12716g), which initiates UPR in yeast [15–17], was upregulated by 23–27 % solely in *TIG* and *inYFP*, without any significant changes in the expression level in *SoA/scYFP* vs the control strain (Fig. 6.a). Its expression was relatively low, but the gene was continuously expressed, either when challenged with r-Prot or not. The same was observed in *P. pastoris* continuous cultures [67]. However, in that study, continuous expression of *HAC1* was interpreted as a hallmark of continuous activation of UPR and ERAD. In *S. cerevisiae*, *HAC1* was continuously upregulated, together with its downstream targets, when the cells were used as r-Prot production platforms [27]. On one hand, our observation was highly surprising, as we expected to observe signs of high UPR in the HSSs. Moreover, it is known that Hac1 regulates transcription of several hundred genes, so its upregulation should induce substantial transcriptional changes as well [16]; this was limited for *inYFP*, having the

largest increase in *HAC1* expression. Likewise, the typical downstream processes responding to UPR (including protein folding and proteasomal degradation, ERAD; Table S.4) were not activated in *TIG/inYFP* strains, except for several genes involved in vacuolar protein sorting: *ATG8* (E02662g), *ATG13* (F03432g), *VPS70* (B05258g) and *CUP5* (F24475g); and vacuolar proteases *PEP4* (F27071g), *PRC1* (A18810g), *PRB1* (B16500g) (Fig. 6.b.f). It was impossible to determine whether upregulation of those vacuolar protein sorting/degradation genes already signifies activation of UPR in *TIG/inYFP*. In contrast, the expression level of the UPR's regulators, including *IRE1* (A14839g), *KAR2* (BIP; E13706g), *ERJ* (C05819g), and the key protein disulfide isomerase *PDI1* (E03036g), together with its reactivator *ERO1* (D09603g) and glutathione synthetase *GSH* (C17831g), all remained at the control level in all the recombinant strains studied here (Fig. 6.c.d). Such an observation was particularly surprising for *SoA*, rich in disulfide bonds. Lack of their upregulation was previously interpreted as lack of an actual *HAC1*-triggered response [2]. In relation to oxidative folding, disulfide bond formation and, associated with it, oxidative stress, we observed downregulation of the cytoplasmic thioredoxin *TRX1* (E09427g) shared by HSSs (within a group with n = 75; Fig. 5.c). Thioredoxin activity is intrinsically associated with the process of disulfide bond formation, mediated by Pdi1 and Ero1, and relies on stochastic oxidation-reduction of cysteine side chains, which consumes considerable amounts of oxidizing and reducing agents (O₂ and glutathione, respectively). Although oxidative folding takes place within the ER lumen, neither thioredoxin nor glutathione redox systems were identified within the ER [49]. According to the current model, a cytoplasmic thioredoxin (e.g. *TRX1*; E09427g) shuttles electrons into the ER to reduce oxidized Pdi [68], while glutathione is actively transported from the cytoplasm by specific transporters [69]. In the face of that, the downregulation of *TRX1* in HSSs is not clear, and contrasted by 17 % upregulation of *SEC61* (E21912g) in *SoA* (Fig. 6.e), which is known to transport glutathione, required for oxidative folding, into the ER lumen [69]. That slight upregulation of the *SEC61* channel is the only, faint marker of enhanced oxidative folding and disulfide bond formation in *SoA*.

In contrast to a constant level of *KAR2* and *IRE1* expression, the expression of *SLS1* (E32703g), encoding a mediator of interaction

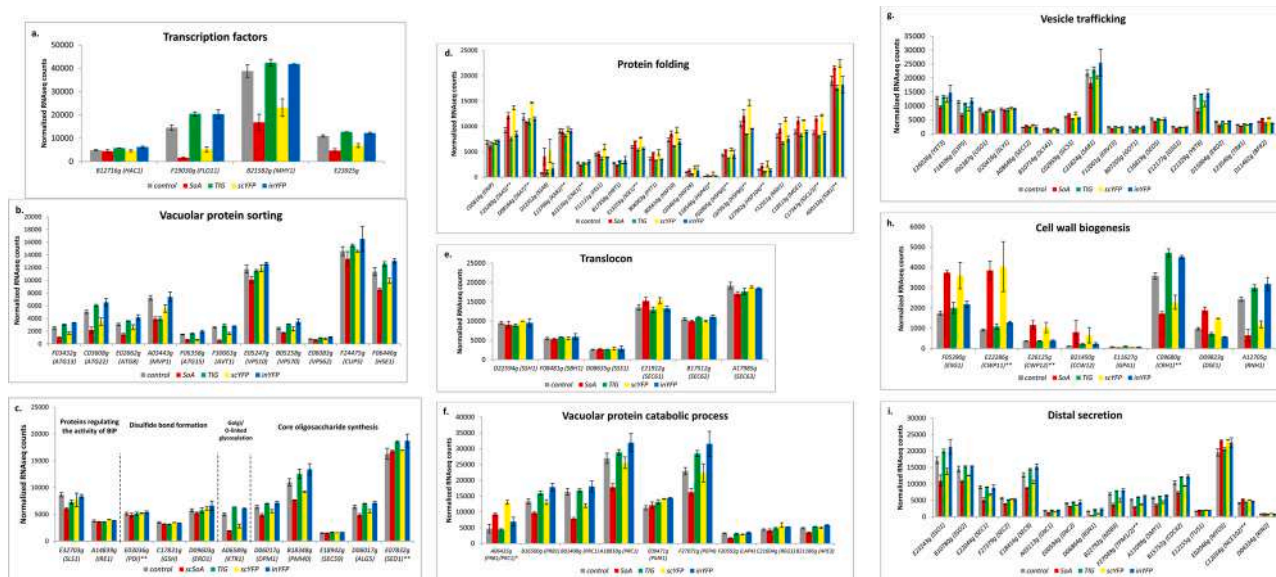


Fig. 6. Expression level of selected genes involved in the r-Prot folding, maturation and secretion. Normalized transcript counts were used for quantification of the transcript levels. Transcription factors (a), Vacuolar protein sorting (b), disulfide bonds formation, O- glycosylation, core oligosaccharides proteins (c), Protein folding (d), Translocon's elements (e), Vacuolar protein catabolic process (f), Vesicle trafficking (g), Cell wall biogenesis (h), Distal secretion (i). Targeted transcriptome analysis. X axis: shortened YALI code of a gene with a three letter name. Y axis: Normalized Counts of transcript. For clarity of the Figure some quantification values were * multiplied by 10, ** divided by 10. (For more convenient visualization of the Figure in color, the reader is referred to the web version of this article).

between Ire1 and Kar2 [70], essential for co-translational translocation in *Y. lipolytica* [71,72], was downregulated upon overproduction of all the rs-ProtS (SoA, TIG and scYFP), but not for inYFP (Fig. 6.c). Sls1 operates as an ADP to ATP exchange factor for Kar2 [72]. In an ATP-bound form, Kar2 is in an Ire1-interaction-favoring state. In that state, Ire1 cannot execute its endonucleolytic activity towards *HAC1/XBP1* mRNAs, which is required for their translation. Hence, it could be speculated that by downregulation of *SLS1*, the interaction of Kar2-Ire1 is limited, and consequently, Ire1 is released for processing *HAC1* mRNA to a translationally active form. In that way, although not upregulated in HSSs, *HAC1* could be translated and initiate massive rearrangements in its transcriptomes (Fig. 2) even from its basic transcription level (Fig. 6.a). Moreover, in the case of HSSs, systematic changes in protein folding, vesicular transportation and vacuolar protein catabolic processes were observed, which could support the proposed model. Definitely, those hypotheses require further deepened biochemical investigations. Additionally, in view of the lack of significant change of *HAC1* expression in HSSs, a comment on the high *CTT1* upregulation is required. While indeed *HAC1* may induce expression of genes involved in the oxidative stress response, which is important for managing the secretory pathway, it was previously demonstrated that the oxidative stress response is not directly managed through the UPR [2], which explains the current observations made for HSSs.

Furthermore, the postulated limited rate of Kar2-Ire1 interaction would suggest that the translocation takes place *via* the post-translational translocation mechanism, *i.e.* the less efficient mode in *Y. lipolytica* [1], or *via* both pathways. This hypothesis is supported by the highly elevated expression of cytosolic chaperones of the Ssa family, in particular *SSA5* (*F25289g*) and *SSA8* (*D22352g*), in HSSs (Fig. 6.d; Table S.4). While the expression of the other Ssa-family members, *SSA5* and *SSA7* (*D08184g*) was very high in all the analyzed strains, including the control (Fig. 6.d), the expression level was most changed for *SSA8* (5.5- and 6.2-fold higher expression in HSSs vs the control) (Table S.4). The Ssa chaperones belong to the Hsp70 family with the primary role in securing the nascent polypeptide in a translocation competent state, when the polypeptide is to be translocated *via* a post-translational mechanism. A similar expression pattern (significant upregulation in HSSs) was observed for another member of the cytosolic chaperones, *SSE1* (*E13255g*) (Fig. 6.d). *SSE1* (and *SSE2* in *P. pastoris*; a single gene in *Y. lipolytica*) assists in protein folding by binding to the nascent peptide and maintaining it in a folding-competent state, but cannot actively promote folding reaction. Co-expression of a heterologous gene with *SSE1* as a helper protein allowed the target protein amount to be increased by 2.3-fold [21]. Additionally, we did not observe any substantial changes in the expression of the genes encoding the co-translational translocation mechanism-associated activities, SRP receptor or SRP itself. Again, this supports the hypothesis that the rs-ProtS studied here were translocated *via* a post-translational mechanism, or, possibly, *via* both in SoA.

With respect to r-Prot-specific cellular responses, we observed systematic, substantial upregulation of the genes involved in synthesis of core oligosaccharides (*DPM1*: *D06017g*, *PMI40*: *B18348g*, *SEC59*: *E18942g*, *ALG5*: *D06017g*) and in glycoprotein processing in *TIG* (Fig. 6.c; Table S.4). Seven in ten genes belonging to the Alg family, involved in glycoprotein processing, were downregulated in HSSs, with concomitant lack of change for *TIG* and *inYFP*. A similar expression profile was also observed for many genes assigned to the GPI-anchoring process (Table S.4). Only a single gene involved in that process, *SED1* (*E07832g*), was upregulated by 20 %, and that upregulation was observed solely in *TIG*. All these data indicate that the selection of *TIG* as a model of highly glycosylated protein was relevant.

Likewise, none of the genes involved in glycoproteins' maturation or processing was significantly up-/downregulated in *inYFP*. The same observation relates to protein folding, disulfide bond formation, protein degradation, ERAD, and vesicular transportation, with only four exceptions in ~250 genes, assigned to those processes; these were:

downregulation of *FES1* (*F11121g*) and *BFR2* (*D11462g*), upregulation of *HSP26* (*C03465g*) and *HRT1* (*B17358g*) (Fig. 6.d.g). *HSP26* encodes a chaperone suppressing unfolded proteins' aggregation, also possessing mRNA-binding activity. The gene was also upregulated in SoA (1.47-fold change) and *scYFP* (1.84-fold change), and was significantly downregulated in *TIG* (0.66-fold change). While differential expression of *HSP26* in all the studied strains demonstrates its importance for r-Prot synthesis, the pattern of expression is unclear. Likewise, the expression level of another cytoplasmic chaperone, *HSP42* (*E18546g*), having unfolded protein-binding and protein-sequestering activity, was significantly elevated, by 2.26- and 4.48-fold, for SoA and *scYFP*, respectively. It is plausible that, except for Ssas, it assists nascent polypeptides to the ER surface for translocation.

Several other interesting observations were made based on the results of the targeted analysis (Fig. 6; Table S.4), namely: 1) vacuolar protein sorting and protein degradation-associated genes' expression was on average 2-fold lower in HSSs; in contrast 2) several above-mentioned vacuolar protein sorting-involved genes (Fig. 6.b), *ATG13*, *ATG8*, *CUP5*, *VPS70*, but also *HSE1* (*F06446g*) and *VPS62* (*E06083g*), were significantly upregulated in *TIG*- and *inYFP*-producing strains, suggesting increased protein degradation occurring in these strains. Whether it is a result of *HAC1* upregulation and induction of UPR remains an open question; 3) in all the strains, expression level of the genes involved in ERAD either remained unchanged or was decreased (five genes in SoA, three genes in *scYFP* and a single gene in *TIG*); 4) within the protein folding biological process we identified several genes significantly upregulated in HSSs, that were all localized to mitochondria, including *MDJ1* (*F12551g*), *MGE1* (*C18513g*), *HSP60* (*F02805g*), and *HSP10* (*B05610g*) (Fig. 6.d), and SoA-specific upregulation of *PRX1* (*A19426g*), with high similarity to mitochondrial peroxiredoxin with thioredoxin peroxidase activity, reactivated by glutathione (involved in cell redox homeostasis and the cellular response to oxidative stress). Such high overrepresentation of upregulated mitochondrial genes corresponds well with the significantly enriched biological process "mitochondrial gene expression and translation" as well as with "cellular respiration", identified for HSSs (Fig. 4.a.e).

According to the conducted biological process enrichment analysis (Fig. 4), protein degradation, autophagy and vacuolar protein sorting were significantly downregulated in HSSs, which was a consequence of many minor downregulations of specific DEGs assigned to those processes, as shown in Table S.4. In approximately 80 genes assigned to the vacuolar protein sorting process, 18 and 30 remained at an unchanged expression level, 2 and 1 were upregulated and the remaining ~60 and ~50 were significantly downregulated in SoA and *scYFP* strains, with concomitant lack of changes in the expression level for *TIG* and *inYFP* (with few exceptions). Amongst the most downregulated DEGs, we found *ATG13*, *ATG22/23* (*C03608g*), *ATG15* (*F06358g*), and *AVT1* (*F30063g*), and the level of their downregulation was correlated for the two HSSs (Fig. 6.b; Table S.4). Functionally associated with "vacuolar protein sorting", the "vacuolar protein catabolic process" was the most DEG-enriched category out of the 27 distinguished. All the genes assigned to that process were somehow affected by the introduced variable. Several genes showed prevailing upregulation, *e.g.* *PMR1* (*E09471g*), significantly upregulated in all the strains, or *PRB1/PRC1* (*A06435g*) and *REG1* (*C21604g*), upregulated in three strains and unchanged in the fourth strain (Fig. 6.f; Table S.4). Several genes showed a differential pattern of expression depending on the strain, such as *PEP4* (*F27071g*), *PRC1* (*A18810g*), *PRB1* (*B16500g*) and *APE3* (*B21186g*), which were downregulated in HSSs and upregulated in *TIG/inYFP*. The remaining genes assigned to that process were either downregulated or at an unchanged expression level, depending on the strain (Table S.4). The uniformly upregulated *PMR1* encodes Ca²⁺-ATPase, responsible for transport of Ca²⁺ and Mn²⁺ inside the Golgi lumen, which are required for glycosylation, sorting and ERAD. Due to the latter implication, *PMR1* is assigned to a vacuolar protein catabolic process, although it localizes to the Golgi membrane. Its disruption was shown to increase secretion of

heterologous proteins in *S. cerevisiae*, *Kluyveromyces lactis* and *Hansenula polymorpha* [73–75]. Its upregulation observed in the present study and the previous results obtained for the other yeast species could be interpreted as undesired targeting and degradation of some portion of correctly folded proteins which could be successfully secreted. It was previously pointed out that mis-sorting to the vacuole causes severe losses of rs-Prot in *P. pastoris* [20]. Likewise, it was shown that significant amounts of r-Prots accumulate in the vacuole in *S. cerevisiae*, and material and energy expenses for their production and degradation were incurred [76]. In relevance to this, we observed high (and undesired) upregulation of the key, vacuolar serine protease *PRB1/PRC1* (A06435g) in *SoA* (2.03-fold change), *scYFP* (2.58-fold change) and *inYFP* (1.46-fold change) (Fig. 6.f), indicating increased protein degradation processes in these strains. Deletion of this gene was a prerequisite for successful modification ($\Delta vps21$) of the early steps of the vacuolar sorting and degradation pathway targeting improved heterologous protein secretion in *P. pastoris* [20]. In the view of its currently documented response to r-Prot overproduction in *Y. lipolytica*, and the previous data for *P. pastoris*, it would be relevant and interesting to investigate the impact of *Aprb1* genotype (accompanied by another modification) on the secretory capacity of *Y. lipolytica*. Likewise, *PEP4*, *PRC1* (A18810g) and *PRB1* (B16500g) genes, encoding vacuolar aspartyl protease and broad-specificity C-terminal exopeptidase involved in non-specific protein degradation in the vacuole, were upregulated in *TIG* and *inYFP*. Together with the previously discussed upregulation of vacuolar protein sorting markers, including *ATG8*, *ATG13*, *CUP5*, and *VPS70*, they represent interesting targets for a genetic engineering strategy aiming at decreased degradation of heterologous protein production in *Y. lipolytica* (at least for strains overproducing *TIG* and *inYFP*). Downregulation of the above-mentioned vacuolar protein sorting markers as well as *PEP4*, *PRC1* (A18810g), and *PRB1* (B16500g) could account for, at least part of, the predominance of HSSs over *TIG/inYFP* in terms of the produced protein amounts (Table 2). It would well align with the previous observations and conclusions [27,67,77] that even though UPR activates foldases, chaperones and other folding and secretory helpers, its continuous activation reduces the final yields of r-Prot.

Cell wall biogenesis- and distal secretion-related biological processes were enriched in the genes responsive to the introduced variable. However, no general trend could be observed depending on the overproduced protein (as could be done in the case of protein folding or glycoprotein processing). This means that the final set of significant DEGs was built by many up-/and down-regulated genes in a given strain. In nearly 170 genes assigned to cell wall biogenesis, only a few were highly responsive, while the majority were up-/down-regulated by not more than 20–30 %. As mentioned above, regulation of cell wall biogenesis was conducted by several highly responsive transcription factors, including *FLO11*, *MHY1*, and *E23925g*, that apparently finetuned the repertoire of the downstream genes, as required for adequate cell wall modifications (Fig. 6.a). Amongst the most responsive genes in HSSs (for which the cell wall biogenesis was a significantly changed biological process; Fig. 3; Fig. 4; Fig. 6.h), we identified highly upregulated *CCW12* (B21450g) (3.32/2.68-fold change in *SoA/scYFP*), encoding a cell wall mannoprotein that plays a role in maintenance of newly synthesized areas of the cell wall; *DSE1* (D09823g; 2.01/1.14-fold change in *SoA/scYFP*), involved in cell separation after cytokinesis; *EXG1* (F05390g; 2.23/1.93-fold change in *SoA/scYFP*), the major exo-1,3-beta-glucanase of the cell wall, involved in cell wall beta-glucan assembly; and two structural cell wall proteins *CWP11/CWP12* (E22286g/E26125g). Amongst the most responsive downregulated genes in HSSs we identified *GPA1* (E11627g), a plasma membrane/endosome G-protein involved in signaling, and *CRH1* (C09680g), encoding chitin transglycosylase, transferring chitin to beta-glucans, both being slightly upregulated in *TIG/inYFP* strains and downregulated in HSSs.

Within the vesicular transportation and membrane fusion processes (Fig. 6.g), we identified several up- and downregulated genes, but

mainly for HSSs. A common expression pattern in both HSSs was observed for: 1) upregulated *SEC12* (A08646g), *BFR2* (D11462g), *SLY41* (B10714g), *GCS1* (C02959g), and downregulated *SAR1* (C21824g), *ERV15* (F12001g), *GOT1* (B07205g), and *USO1* (F02387g) involved in COPII anterograde transport; 2) as well as *SED5* (C16819g), several representatives of Sec and Trs families, *GSG1* (E12177g), *YKT6* (E21329g), playing a role in membrane fusion; 3) and also *ERD2* (D15004g) and *TRX1* (E23540g), involved in retrograde Golgi to ER transport. It is not possible to discuss the specifics of all the above-mentioned genes here, but the responsive genes may be potentially considered as targets in genetic engineering strategies enhancing rs-Prot production. Such a strategy was successfully implemented [14]. Interestingly, *BFR1* (D11462g), upregulated here, was previously shown to be a useful secretory helper in *P. pastoris*, increasing the target r-Prot production by 1.5-fold [21], which demonstrates plausible similarities in its operation in the two yeast species.

Similarly, in 65 genes assigned to the distal secretion process (Fig. 6.i), only one (and only in the *SoA* strain) met the requirement of fold change >2 for DEG (*RSN1*; D06864g; -2.25), while the remaining genes' expression was changed to a smaller degree. We paid attention to the genes that were previously indicated as interesting targets for improving r-Prot overproduction in yeast, e.g. *SSO1/2* (E23243g, B10780g), encoding a syntaxin homolog, acting in late stages of secretion post-Golgi t-SNARE fusion [14,78], *SNC1/2* (A03113g, E00594g), vesicle membrane receptor protein (v-SNARE); involved in the fusion between Golgi-derived secretory vesicles with the plasma membrane, and *SEC1/2/9* (E22044g, F27379g, E18414g), encoding an SM-like protein involved in docking and fusion of exocytic vesicles, which binds to assembled SNARE complexes at the membrane (reviewed by Celińska et al. [30]). They all shared a common expression pattern of downregulation in HSSs and slight upregulation or no change in *TIG/inYFP*. Their downregulation upon very high overproduction and secretion, with no evident upregulation of any redundant activities, is unclear. The same relates to *KIN2* (D22770g), which was previously used as a very efficient secretion helper, increasing the target protein production by 2.2-fold upon co-expression [21]. However, in the present research, production of either of the studied r-Prot did not trigger changes in its expression profile. Moreover, exactly the same expression pattern was observed for many other genes assigned to the distal secretion process, but exerting different functions, including the kinesin-like myosin passenger-protein *SMY1* (A11099g), tropomyosin *TPM1/2* (F27049g), Rab GTPase-activating protein *MSB3* (B22792g) important for polarized growth, and rho-like GTPase *CDC42* (B15752g), having a key role in establishment and maintenance of cell polarity and regulating actin assembly (Fig. 6.i; Table S.4). Their downregulation could result from the action of the filamentation factors, and previously postulated transition into ovoid forms, and not be related to r-Prot transportation and secretion. Only a few genes assigned to the distal secretion process escaped that general expression pattern. These were, for example, the rho1-GDP exchange factor *TUS1* (E12155g), required for signaling of cell wall defects to Rho1, and the type-I myosin *MYO5* (E02046g) that promotes actin assembly, which were upregulated in HSSs. Interestingly, another gene, *NCE102* (C12034g), was upregulated in all the secretory protein-overproducing strains (no change for *inYFP* solely). *Nce102* stands for non-classical export protein 2, which is involved in secretion of proteins that lack classical secretory signal sequences. While its upregulation upon overproduction of the secretory heterologous proteins is logical, it was surprising considering the fact that all the heterologous genes were transcriptionally fused to a signal peptide of typical structure and function (exo-1,3-beta-glucanase; B03564g) [7]. *Nce102* is implicated in export of proteins which lack a cleavable signal sequence. However, the non-classical export pathway is known to operate as an alternative clearance/detoxification pathway to eliminate damaged material, when the basic pathway is not sufficient. Thus, we postulate that upon the very high secretion of *SoA* and *scYFP*, *NCE102* works as an alternative route of secretion of correctly folded, functional proteins, as

we observed in the activity/fluorescence tests (Table 2).

4. Conclusions

The conducted research improved our understanding of phenomena taking place inside *Y. lipolytica* cells upon r-Prot overproduction. First of all, based on global comparison of the expression profiles generated by SoA, scYFP, TIG and inYFP, we conclude that targeting a protein for secretion (as opposed to its retention within a cell) induces massive transcriptional changes, greatly exceeding those resulting from secreting a protein with different biochemical characteristics. Interestingly, if the same protein (YFP) was targeted for secretion, its yield was > 10-fold higher. Current data, confirmed for *Y. lipolytica*, as is already known for *P. pastoris* and *S. cerevisiae*, show that high overproduction of r-Prot is inherently associated with oxidative stress. Here it was indicated by high upregulation of a “response to hydrogen peroxide” biological process, and upregulation of catalases, as the most responsive DEGs. Importantly, the phenomenon does not rely on the number of disulfide bonds within the r-Prot’s secondary structure, at least for the two HSSs studied here. Overproduction of r-Prot triggered massive rearrangements within the cell by acting through transcriptional regulators, such as *DOI1353g*, downregulated upon overproduction of any r-Prot, irrespective of its biochemical characteristics, or specific for HSSs – *RSFA* or *MAF1*. Based on their putative function, their downregulation in HSSs led to release of RNA polymerase II and III, respectively, from their inhibitory activity. On the other hand, the key transcription factor, typically associated with the response to overproduction of r-Prots, *HAC1*, was upregulated solely in the strains producing lower yields of r-Prots (inYFP/TIG). Those strains were also characterized by upregulation in vacuolar protein sorting and protein degradation. Whether it was caused by the actual onset of UPR could not be definitely established, as the other processes, typically upregulated under UPR, remained at the control level. Likewise, whether lower yields of TIG/inYFP were caused by the UPR remains to be experimentally verified.

The global approach undertaken here allowed us to perceive unobvious relationships, such as overproduction of r-Prot, with given characteristics (or at a given rate) and the cell’s morphotype. Genes putatively linking these phenomena were indicated here. The unexpected downregulation of *SLS1* (an interaction mediator of *IRE1* and *KAR2*), shared by all the secretory r-Prot producers, led us to form conclusions on the mechanism of the nascent polypeptides’ translocation and to speculations on its regulation. Based on cytosolic chaperones’ expression pattern, we postulate that the polypeptides were in the majority of cases translocated into the ER lumen via a post-translational translocation mechanism. Traditionally, the interaction between a polypeptide’s signal peptide and SRP was considered the decision-making point [79]. Based on the current data, we speculate that the ER occupancy also participates in the decision making. However, this hypothesis requires experimental testing. We also report implication of *NCE102* in rs-Prot secretion, which was highly upregulated in all the rs-Prot overproducers. While its assigned function is secretion of proteins that lack typical signal peptides (not true for SoA, TIG, or scYFP), we postulate its significant role in secretion of functional r-Prots, equipped with a typical signal peptide, under overproduction conditions. And finally, we hypothesize that the cells enter into a growth arrest phase (G1 phase) upon high overproduction of r-Prots, as a strategy undertaken by the cells to withstand the high burden imposed on them by genetic engineering. The hypothesis is supported by downregulation of a G1/S-specific cyclin gene, activation of detoxification processes (common upregulation of peroxisomal *PXMP2/4*), and transition into ovoid morphotype in HSSs.

Author statement

PKW designed the experiments, performed experimental work, analyzed omics data and wrote the manuscript. EC conceived the study,

analyzed omics data and wrote the manuscript.

Funding

This study was financially supported by the Ministry of Science and Higher Education (currently, Ministry of Education and Science) project number: DI 2017 000947.

Data availability

All data accompanying this research are presented directly in the manuscript, supplementary materials or are available in Sequence Read Archive, NCBI database under BioProject number: PRJNA701856.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The Authors would like to thank Dr Piotr Kubiak for guidance and assistance in chemostat cultures. DBFM’s chromatography team is acknowledged for great help in HPLC analyses.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.btre.2021.e00646>.

References

- [1] M. Delic, M. Valli, A.B. Graf, M. Pfeffer, D. Mattanovich, B. Gasser, The secretory pathway: exploring yeast diversity, *FEMS Microbiol. Rev.* 37 (2013) 872–914, <https://doi.org/10.1111/1574-6976.12020>.
- [2] K.E.J. Tyo, Z. Liu, D. Petranovic, J. Nielsen, Imbalance of heterologous protein folding and disulfide bond formation rates yields runaway oxidative stress, *BMC Biol.* 10 (2012), <https://doi.org/10.1186/1741-7007-10-16>.
- [3] E. Çelik, P. Çalik, Production of recombinant proteins by yeast cells, *Biotechnol. Adv.* 30 (2012) 1108–1118, <https://doi.org/10.1016/j.biotechadv.2011.09.011>.
- [4] C. Madzak, Engineering *Yarrowia lipolytica* for use in biotechnological applications: a review of major achievements and recent innovations, *Mol. Biotechnol.* 60 (2018) 621–635, <https://doi.org/10.1007/s12033-018-0093-4>.
- [5] J. Hou, K.E.J. Tyo, Z. Liu, D. Petranovic, J. Nielsen, Metabolic engineering of recombinant protein secretion by *Saccharomyces cerevisiae*, *FEMS Yeast Res.* 12 (2012) 491–510, <https://doi.org/10.1111/j.1567-1364.2012.00810.x>.
- [6] R.J. Zahrl, D.A. Peña, D. Mattanovich, B. Gasser, Systems biotechnology for protein production in *Pichia pastoris*, *FEMS Yeast Res.* 17 (2017) 1–15, <https://doi.org/10.1093/femsyr/fox068>.
- [7] E. Celińska, M. Borkowska, W. Białas, P. Korpys, J.M. Nicaud, Robust signal peptides for protein secretion in *Yarrowia lipolytica*: identification and characterization of novel secretory tags, *Appl. Microbiol. Biotechnol.* 102 (2018) 5221–5233, <https://doi.org/10.1007/s00253-018-8966-9>.
- [8] G. Duan, L. Ding, D. Wei, H. Zhou, J. Chu, S. Zhang, J. Qian, Screening endogenous signal peptides and protein folding factors to promote the secretory expression of heterologous proteins in *Pichia pastoris*, *J. Biotechnol.* (2019), <https://doi.org/10.1016/j.jbiotec.2019.06.297>.
- [9] T. Yarimizu, M. Nakamura, H. Hoshida, R. Akada, Synthetic signal sequences that enable efficient secretory protein production in the yeast *Kluyveromyces marxianus*, *Microb. Cell Fact.* 14 (2015), <https://doi.org/10.1186/s12934-015-0203-y>.
- [10] J.D. Smith, B.C. Tang, A.S. Robinson, Protein disulfide isomerase, but not binding protein, overexpression enhances secretion of a non-disulfide-bonded protein in yeast, *Biotechnol. Bioeng.* 85 (2004) 340–350, <https://doi.org/10.1002/bit.10853>.
- [11] H. Tang, X. Bao, Y. Shen, M. Song, S. Wang, C. Wang, J. Hou, Engineering protein folding and translocation improves heterologous protein secretion in *Saccharomyces cerevisiae*, *Biotechnol. Bioeng.* 112 (2015) 1872–1882, <https://doi.org/10.1002/bit.25596>.
- [12] J. Hou, H. Tang, Z. Liu, T. Österlund, J. Nielsen, D. Petranovic, Management of the endoplasmic reticulum stress by activation of the heat shock response in yeast, *FEMS Yeast Res.* 14 (2014) 481–494, <https://doi.org/10.1111/1567-1364.12125>.
- [13] J. Bao, M. Huang, D. Petranovic, J. Nielsen, Moderate expression of SEC16 increases protein secretion by *Saccharomyces cerevisiae*, *Appl. Environ. Microbiol.* 83 (2017), <https://doi.org/10.1128/AEM.03400-16>.
- [14] J.H.D. Van Zyl, R. Den Haan, W.H. Van Zyl, Overexpression of native *Saccharomyces cerevisiae* ER-to-Golgi SNARE genes increased heterologous cellulase

- secretion, *Appl. Microbiol. Biotechnol.* 100 (2016) 505–518, <https://doi.org/10.1007/s00253-015-7022-2>.
- [15] A. Graf, B. Gasser, M. Dragosits, M. Sauer, G.G. Leparc, T. Tüchler, D.P. Kreil, D. Mattanovich, Novel insights into the unfolded protein response using *Pichia pastoris* specific DNA microarrays, *BMC Genomics* 9 (2008) 1–13, <https://doi.org/10.1186/1471-2164-9-390>.
- [16] M. Guerfal, S. Ryckaert, P.P. Jacobs, P. Ameloot, K. Van Craenenbroeck, R. Derycke, N. Callewaert, The HAC1 gene from *Pichia pastoris*: characterization and effect of its overexpression on the production of secreted, surface displayed and membrane proteins, *Microb. Cell Fact.* 9 (2010) 1–12, <https://doi.org/10.1186/1475-2859-9-49>.
- [17] B. Gasser, M. Maurer, J. Rautio, M. Sauer, A. Bhattacharyya, M. Saloheimo, M. Penttilä, D. Mattanovich, Monitoring of transcriptional regulation in *Pichia pastoris* under protein production conditions, *BMC Genomics* 8 (2007) 1–18, <https://doi.org/10.1186/1471-2164-8-179>.
- [18] W.A. Rodriguez-Limas, V. Tannenbaum, K.E.J. Tyo, Blocking endocytotic mechanisms to improve heterologous protein titers in *Saccharomyces cerevisiae*, *Biotechnol. Bioeng.* 112 (2015) 376–385, <https://doi.org/10.1002/bit.25360>.
- [19] A. Idiris, H. Tohda, M. Sasaki, K. Okada, H. Kumagai, Y. Giga-Hama, K. Takegawa, Enhanced protein secretion from multiprotease-deficient fission yeast by modification of its vacuolar protein sorting pathway, *Appl. Microbiol. Biotechnol.* 85 (2010) 667–677, <https://doi.org/10.1007/s00253-009-2151-0>.
- [20] L. Marsalek, C. Gruber, F. Altmann, M. Aleschko, D. Mattanovich, B. Gasser, V. Puxbaum, Disruption of genes involved in CORVET complex leads to enhanced secretion of heterologous carboxylesterase only in protease deficient *Pichia pastoris*, *Biotechnol. J.* 12 (2017), <https://doi.org/10.1002/biot.201600584>.
- [21] B. Gasser, M. Sauer, M. Maurer, G. Stadlmayr, D. Mattanovich, Transcriptomics-based identification of novel factors enhancing heterologous protein secretion in yeasts, *Appl. Environ. Microbiol.* 73 (2007) 6499–6507, <https://doi.org/10.1128/AEM.01196-07>.
- [22] D. Swennen, J.M. Beckerich, *Yarrowia lipolytica* vesicle-mediated protein transport pathways, *BMC Evol. Biol.* 7 (2007) 219, <https://doi.org/10.1186/1471-2148-7-219>.
- [23] D. Swennen, C. Henry, J.M. Beckerich, Folding proteome of *Yarrowia lipolytica* targeting with uracil permease mutants, *J. Proteome Res.* 9 (2010) 6169–6179, <https://doi.org/10.1021/pr100340p>.
- [24] C.B. Matthews, A. Kuo, K.R. Love, J.C. Love, Development of a general defined medium for *Pichia pastoris*, *Biotechnol. Bioeng.* 115 (2018) 103–113, <https://doi.org/10.1002/bit.26440>.
- [25] M.A. Nieto-Taype, J. Garrigós-Martínez, M. Sánchez-Farrando, F. Valero, X. Garcia-Ortega, J.L. Montesinos-Seguí, Rationale-based selection of optimal operating strategies and gene dosage impact on recombinant protein production in *Komagataella phaffii* (*Pichia pastoris*), *Microb. Biotechnol.* 13 (2020) 315–327, <https://doi.org/10.1111/1751-7915.13498>.
- [26] J. Hou, T. Österlund, Z. Liu, D. Petranovic, J. Nielsen, T. Osterlund, Z. Liu, D. Petranovic, J. Nielsen, Heat shock response improves heterologous protein secretion in *Saccharomyces cerevisiae*, *Appl. Microbiol. Biotechnol.* 97 (2013) 3559–3568, <https://doi.org/10.1007/s00253-012-4596-9>.
- [27] Z. Liu, J. Hou, J.L. Martínez, D. Petranovic, J. Nielsen, Correlation of cell growth and heterologous protein production by *Saccharomyces cerevisiae*, *Appl. Microbiol. Biotechnol.* 97 (2013) 8955–8962, <https://doi.org/10.1007/s00253-013-4715-2>.
- [28] J.L. Martínez, E. Meza, D. Petranovic, J. Nielsen, The impact of respiration and oxidative stress response on recombinant α -amylase production by *Saccharomyces cerevisiae*, *Metab. Eng. Commun.* 3 (2016) 205–210, <https://doi.org/10.1016/j.meteno.2016.06.003>.
- [29] P. Swietalski, F. Hetzel, I. Seitz, L. Fischer, Secretion of a low and high molecular weight β -glycosidase by *Yarrowia lipolytica*, *Microb. Cell Fact.* 19 (2020) 1–13, <https://doi.org/10.1186/s12934-020-01358-5>.
- [30] E. Celińska, J.-M. Nicaud, Filamentous fungi-like secretory pathway strayed in a yeast system: peculiarities of *Yarrowia lipolytica* secretory pathway underlying its extraordinary performance, *Appl. Microbiol. Biotechnol.* 103 (2019) 39–52, <https://doi.org/10.1007/s00253-018-9450-2>.
- [31] E. Celińska, R. Ledesma-Amaro, M. Larroude, T. Rossignol, C. Pauthenier, J. M. Nicaud, Golden gate assembly system dedicated to complex pathway manipulation in *Yarrowia lipolytica*, *Microb. Biotechnol.* 10 (2017) 450–455, <https://doi.org/10.1111/1751-7915.12605>.
- [32] P. Korpys-Woźniak, P. Kubiak, W. Białas, E. Celińska, Impact of overproduced heterologous protein characteristics on physiological response in *Yarrowia lipolytica* steady-state-maintained continuous cultures, *Appl. Microbiol. Biotechnol.* (2020), <https://doi.org/10.1007/s00253-020-10937-w>.
- [33] J. Sambrook, D. Russell, *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory Press, New York, 2001.
- [34] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻ $\Delta\Delta$ CT method, *Methods.* 25 (2001) 402–408, <https://doi.org/10.1006/meth.2001.1262>.
- [35] M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads, *EMBnet. J.* 17 (2011) 10–12, <https://doi.org/10.14806/ej.17.1.200>.
- [36] S. Andrews, A Quality Control Tool for High Throughput Sequence Data, 2010. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- [37] D. Kim, J. Paggi, C. Park, C. Bennett, S. Salzberg, Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype, *Nat. Biotechnol.* 37 (2019) 907–915, <https://doi.org/10.1038/s41587-019-0201-4>.
- [38] S. Anders, P.T. Pyl, W. Huber, HTSeq-A python framework to work with high-throughput sequencing data, *Bioinformatics.* 31 (2015) 166–169, <https://doi.org/10.1093/bioinformatics/btu638>.
- [39] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, *Genome Biol.* 15 (2014) 1–21, <https://doi.org/10.1186/s13059-014-0550-8>.
- [40] H. Mi, A. Muruganujan, X. Huang, D. Ebert, C. Mills, X. Guo, P. Thomas, Protocol update for large-scale genome and gene function analysis with PANTHER classification system (v.14.0), *Nat. Protoc.* 14 (2019) 703–721, <https://doi.org/10.1038/s41596-019-0128-8>.
- [41] H. Heberle, V.G. Meirelles, F.R. da Silva, G.P. Telles, R. Minghim, InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams, *BMC Bioinformatics* 16 (2015) 1–7, <https://doi.org/10.1186/s12859-015-0611-3>.
- [42] M. Borkowska, W. Białas, M. Kubiak, E. Celińska, Rapid micro-assays for amyolytic activities determination: customization and validation of the tests, *Appl. Microbiol. Biotechnol.* 103 (2019) 2367–2379, <https://doi.org/10.1007/s00253-018-09610-0>.
- [43] C. Goncalves, R.M. Rodriguez-Jasso, N. Gomes, J.A. Teixeira, I. Belo, Adaptation of dinitrosalicylic acid method to microtiter plates, *Anal. Methods* 2 (2010) 2046–2048, <https://doi.org/10.1039/c0ay00525h>.
- [44] M. Gorczyca, J. Kaźmierczak, S. Steels, P. Fickers, E. Celińska, Impact of oxygen availability on heterologous gene expression and polypeptide secretion dynamics in *Yarrowia lipolytica*-based protein production platforms, *YEAST.* 37 (2020) 559–568, <https://doi.org/10.1002/yea.3499>.
- [45] J. Ruiz-Herrera, R. Sentandreu, Different effectors of dimorphism in *Yarrowia lipolytica*, *Arch. Microbiol.* 178 (2002) 477–483, <https://doi.org/10.1007/s00203-002-0478-3>.
- [46] F.M. Kawasse, P.F. Amaral, M.H.M. Rocha-Leão, A.L. Amaral, E.C. Ferreira, M.A. Z. Coelho, Morphological analysis of *Yarrowia lipolytica* under stress conditions through image processing, *Bioprocess Biosyst. Eng.* 25 (2003) 371–375, <https://doi.org/10.1007/s00449-003-0319-z>.
- [47] Q. Zhao, Y. Su, Z. Wang, C. Chen, T. Wu, Y. Huang, Identification of glutathione (GSH)-independent glyoxalase III from *Schizosaccharomyces pombe*, *BMC Evol. Biol.* 14 (2014) 1–18, <https://doi.org/10.1186/1471-2148-14-86>.
- [48] S. Hasim, N.A. Hussin, F. Alomar, K.R. Bidasee, K.W. Nickerson, M.A. Wilson, A glutathione-independent glyoxalase of the DJ-1 superfamily plays an important role in managing metabolically generated methylglyoxal in *Candida albicans*, *J. Biol. Chem.* 289 (2014) 1662–1674, <https://doi.org/10.1074/jbc.M113.505784>.
- [49] D. Guerrero-Gómez, J.A. Mora-Lorca, B. Sáenz-Narciso, F.J. Naranjo-Galindo, F. Muñoz-Lobato, C. Parrado-Fernández, J. Goikolea, A. Cedazo-Minguez, C. D. Link, C. Neri, M.D. Sequeda, R.P. Vázquez-Manrique, E. Fernández-Suárez, V. Goder, R. Pané, E. Cabisco, P. Askjaer, J. Cabello, A. Miranda-Vizuete, Loss of glutathione redox homeostasis impairs proteostasis by inhibiting autophagy-dependent protein degradation, *Cell Death Differ.* 26 (2019) 1545–1565, <https://doi.org/10.1038/s41418-018-0270-9>.
- [50] C. Buchman, P. Skroch, J. Welch, S. Fogel, M. Karin, The CUP2 gene product, regulator of yeast metallothionein expression, is a copper-activated DNA-binding protein, *Mol. Cell. Biol.* 9 (1989) 4091–4095, <https://doi.org/10.1128/mcb.9.9.4091>.
- [51] S. Michaelis, J. Barrowman, Biogenesis of the *Saccharomyces cerevisiae* pheromone a-Factor, from yeast mating to human disease, *Microbiol. Mol. Biol. Rev.* 76 (2012) 626–651, <https://doi.org/10.1128/mmb.00010-12>.
- [52] R. Häuser, M. Pech, J. Kjek, H. Yamamoto, B. Titz, F. Naeve, A. Tovchigrechko, K. Yamamoto, W. Szaflarski, N. Takeuchi, T. Stellberger, M.E. Diefenbacher, K. H. Nierhaus, P. Uetz, RsfA (YbeB) proteins are conserved ribosomal silencing factors, *PLoS Genet.* 8 (2012) 1–12, <https://doi.org/10.1371/journal.pgen.1002815>.
- [53] P. Fabrizio, S. Hoon, M. Shamalnasab, A. Galbani, M. Wei, G. Giaever, C. Nislow, V. D. Longo, Genome-wide screen in *Saccharomyces cerevisiae* identifies vacuolar protein sorting, autophagy, biosynthetic, and tRNA methylation genes involved in life span regulation, *PLoS Genet.* 6 (2010) 1–14, <https://doi.org/10.1371/journal.pgen.1001024>.
- [54] S. Alers, S. Wesselborg, B. Stork, ATG13: Just a companion, or an executor of the autophagic program? *Autophagy.* 10 (2014) 1481, <https://doi.org/10.4161/aut.29717>.
- [55] V. Ramya, R. Rajasekharan, ATG15 encodes a phospholipase and is transcriptionally regulated by YAP1 in *Saccharomyces cerevisiae*, *FEBS Lett.* 590 (2016) 3155–3167, <https://doi.org/10.1002/1873-3468.12369>.
- [56] F.E. Reyes Turcu, K.H. Ventii, K.D. Wilkinson, Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes, *Annu. Rev. Biochem.* 78 (2009) 363–397, <https://doi.org/10.1146/annurev.biochem.78.082307.091526>.
- [57] V. Puxbaum, D. Mattanovich, B. Gasser, Quo vadis? The challenges of recombinant protein folding and secretion in *Pichia pastoris*, *Appl. Microbiol. Biotechnol.* 99 (2015) 2925–2938, <https://doi.org/10.1007/s00253-015-6470-z>.
- [58] C.I.F. Silva, H. Teles, A.P.H.A. Moers, G. Eggink, F.A. de Wolf, M.W.T. Werten, Secreted production of collagen-inspired gel-forming polymers with high thermal stability in *Pichia pastoris*, *Biotechnol. Bioeng.* 108 (2011) 2517–2525, <https://doi.org/10.1002/bit.23228>.
- [59] M. Wu, Q. Shen, Y. Yang, S. Zhang, W. Qu, J. Chen, H. Sun, S. Chen, Disruption of YPS1 and PEP4 genes reduces proteolytic degradation of secreted HSA/PTH in *Pichia pastoris* GS115, *J. Ind. Microbiol. Biotechnol.* 40 (2013) 589–599, <https://doi.org/10.1007/s10295-013-1264-8>.
- [60] E. Celińska, M. Olkowicz, W. Grajek, L-Phenylalanine catabolism and 2-phenylethano synthesis in *Yarrowia lipolytica*-mapping molecular identities through whole-proteome quantitative mass spectrometry analysis, *FEMS Yeast Res.* 15 (2015), <https://doi.org/10.1093/femsyr/fov041>.
- [61] E. Celińska, M. Borkowska, W. Białas, M. Kubiak, P. Korpys, M. Archacka, R. Ledesma-Amaro, J.J.-M. Nicaud, Genetic engineering of Ehrlich pathway

- modulates production of higher alcohols in engineered *Yarrowia lipolytica*, FEMS Yeast Res. 19 (2019) 1–13, <https://doi.org/10.1093/femsyr/foyl22>.
- [62] A. Idiris, H. Tohda, H. Kumagai, K. Takegawa, Engineering of protein secretion in yeast: strategies and impact on protein production, Appl. Microbiol. Biotechnol. 86 (2010) 403–417, <https://doi.org/10.1007/s00253-010-2447-0>.
- [63] D. Gems, J.J. McElwee, Broad spectrum detoxification: the major longevity assurance process regulated by insulin/IGF-1 signaling? Mech. Ageing Dev. 126 (2005) 381–387, <https://doi.org/10.1016/j.mad.2004.09.001>.
- [64] L. Aravind, V. Anantharaman, S. Balaji, M.M. Babu, L.M. Iyer, The many faces of the helix-turn-helix domain: transcription regulation and beyond, FEMS Microbiol. Rev. 29 (2005) 231–262, <https://doi.org/10.1016/j.femsre.2004.12.008>.
- [65] L. Gaigalat, J.P. Schlüter, M. Hartmann, S. Mormann, A. Tauch, A. Pühler, J. Kalinowski, The DeoR-type transcriptional regulator SugR acts as a repressor for genes encoding the phosphoenolpyruvate: sugar phosphotransferase system (PTS) in *Corynebacterium glutamicum*, BMC Mol. Biol. 8 (2007) 1–20, <https://doi.org/10.1186/1471-2199-8-104>.
- [66] S. Liang, B. Wang, L. Pan, Y. Ye, M. He, S. Han, S. Zheng, X. Wang, Y. Lin, Comprehensive structural annotation of *Pichia pastoris* transcriptome and the response to various carbon sources using deep paired-end RNA sequencing, BMC Genomics 13 (2012), <https://doi.org/10.1186/1471-2164-13-738>.
- [67] A.R. Hesketh, J.I. Castrillo, T. Sawyer, D.B. Archer, S.G. Oliver, Investigating the physiological response of *Pichia (Komagataella) pastoris* GS115 to the heterologous expression of misfolded proteins using chemostat cultures, Appl. Microbiol. Biotechnol. 97 (2013) 9747–9762, <https://doi.org/10.1007/s00253-013-5186-1>.
- [68] G.J. Poet, O.B. Oka, M. Lith, Z. Cao, P.J. Robinson, M.A. Pringle, E.S. Arnér, N. J. Bulleid, Cytosolic thioredoxin reductase 1 is required for correct disulfide formation in the ER, EMBO J. 36 (2017) 693–702, <https://doi.org/10.15252/embj.201695336>.
- [69] A.J. Ponsoero, A. Igarria, M.A. Darch, S. Miled, C.E. Outten, J.R. Winther, G. Palais, B. D'Autreaux, A.A. Delaunay-Moisan, M.B. Toledano, B. D'Autreaux, A. A. Delaunay-Moisan, M.B. Toledano, Endoplasmic reticulum transport of glutathione by Sec61 is regulated by Ero1 and bip, Mol. Cell 67 (2017) 962–973, <https://doi.org/10.1016/j.molcel.2017.08.012>.
- [70] A. Babour, M. Kabani, A. Boisramé, J.M. Beckerich, Characterization of Ire1 in the yeast *Yarrowia lipolytica* reveals an important role for the Sls1 nucleotide exchange factor in unfolded protein response regulation, Curr. Genet. 53 (2008) 337–346, <https://doi.org/10.1007/s00294-008-0190-1>.
- [71] A. Boisramé, M. Kabani, J.M. Beckerich, E. Hartmann, C. Gaillardin, Interaction of Kar2p and Sls1p is required for efficient co-translational translocation of secreted proteins in the yeast *Yarrowia lipolytica*, J. Biol. Chem. 273 (1998) 30903–30908, <https://doi.org/10.1074/jbc.273.47.30903>.
- [72] A. Boisramé, J.M. Beckerich, C. Gaillardin, Sls1p, an endoplasmic reticulum component, is involved in the protein translocation process in the yeast *Yarrowia lipolytica*, J. Biol. Chem. 271 (1996) 11668–11675, <https://doi.org/10.1074/jbc.271.20.11668>.
- [73] H.L. Zhao, C. Xue, Y. Wang, X.Q. Yao, Z.M. Liu, Increasing the cell viability and heterologous protein expression of *Pichia pastoris* mutant deficient in PMR1 gene by culture condition optimization, Appl. Microbiol. Biotechnol. 81 (2008) 235–241, <https://doi.org/10.1007/s00253-008-1666-0>.
- [74] M.M. Harmsen, M.I. Bruyne, H.A. Raué, J. Maat, Overexpression of binding protein and disruption of the PMR1 gene synergistically stimulate secretion of bovine prochymosin but not plant Thaumatin in yeast, Appl. Microbiol. Biotechnol. 46 (1996) 365–370, <https://doi.org/10.1007/BF00166231>.
- [75] Y.S. Sohn, C.S. Park, S.B. Lee, D.D.Y. Ryu, Disruption of PMR1, encoding a Ca²⁺-ATPase homolog in *Yarrowia lipolytica*, affects secretion and processing of homologous and heterologous proteins, J. Bacteriol. 180 (1998) 6736–6742, <https://doi.org/10.1128/jb.180.24.6736-6742.1998>.
- [76] K.E.J. Tyo, Z. Liu, Y. Magnusson, D. Petranovic, J. Nielsen, Impact of protein uptake and degradation on recombinant protein secretion in yeast, Appl. Microbiol. Biotechnol. 98 (2014) 7149–7159, <https://doi.org/10.1007/s00253-014-5783-7>.
- [77] K.J. Kauffman, E.M. Pridgen, F.J. Doyle, P.S. Dhurjati, A.S. Robinson, Decreased protein expression and intermittent recoveries in BiP levels result from cellular stress during heterologous protein expression in *Saccharomyces cerevisiae*, Biotechnol. Prog. 18 (2002) 942–950, <https://doi.org/10.1021/bp025518g>.
- [78] J.H. Toikkanen, L. Sundqvist, S. Keränen, *Kluyveromyces lactis* SSO1 and SEB1 genes are functional in *Saccharomyces cerevisiae* and enhance production of secreted proteins when overexpressed, Yeast 21 (2004) 1045–1055, <https://doi.org/10.1002/yea.1151>.
- [79] D.T.W. Ng, J.D. Brown, P. Walter, Signal sequences specify the targeting route to the endoplasmic reticulum membrane, J. Cell Biol. 134 (1996) 269–278, <https://doi.org/10.1083/jcb.134.2.269>.

SUPPLEMENTARY

Global transcriptome profiling reveals genes responding to overproduction of a small secretory, a high cysteine- and a high glycosylation-bearing protein in *Yarrowia lipolytica*

List of Supplementary Materials

Figure S.1. Kinetics of Po1h-derived *Y. lipolytica* strain cultures maintained in steady-state; *X* axis: culturing time (h) or residence time (τ); *Y* axis: glycerol concentration ($[g L^{-1}]$; GLY); biomass growth ($[g L^{-1}]$; DCW); glucoamylase activity ($[GAU L^{-1}]$); α -amylase activity ($[AAU L^{-1}]$); YFP fluorescence ($[kRFU L^{-1}]$); in (A) *TIG*, (B) *SoA*, (C) *scYFP*, (D) *inYFP*; pO₂ 20%, pH 5.5; D 0.20 [h⁻¹]; Error bars indicate \pm SD from at least biological duplicates.

Figure S.2. Box plot illustrating distribution of the number of gene readings (logarithmically scaled) in the samples subjected to RNAseq. Bottom and top edge of the box is the 25th and 75th percentile respectively. The middle line of the box is the median.

Figure S.3. Heatmap illustrating Pearson correlation of the number of specific gene readings in individual samples subjected to RNAseq analysis.

Figure S.4. Volcano plots showing the dependence of the negative logarithm values of *p_value_not_adjusted* from the logarithmic change in expression fold *log2_Estimated_FoldChange* of the specific recombinant strain sample vs the control sample. *inYFP* vs control (a.), *scYFP* vs control (b.), *scTIG* vs control (c.), *scSoA* vs control (d.). Genes with *p_value_not_adjusted* < 0.05 are marked as blue dots, genes with *Estimated_FoldChange* > 2.0 are marked in green. Genes meeting both of the above conditions are marked in red, and their names are also included in the chart. To prepare the chart, the *EnhancedVolcano* function of the *EnhancedVolcano* package (R program) was used.

Figure S.5. Cell morphology of the *Y. lipolytica* strains maintained in steady-state chemostat cultivations. A – Po1h_Ura3 (control strain; filaments); B – *TIG* (ovoids); C – *SoA* (filaments and elongated ovoids); D – *inYFP* (ovoids); E – *scYFP* (very long filaments and ovoids). Observations were performed under 1000 \times magnification.

Figure S.6. Venn Diagrams analysis illustrating shared and dissimilar genes/gene sets for individual strains. The numbers indicate the number of shared or strain-specific genes, up-/down-regulated in *scYFP*- and *inYFP*- producing strains (a); up-regulated genes in *SoA*-, *TIG*-, *scYFP*- and *inYFP*-producing strains (b); down-regulated genes in *SoA*-, *TIG*-, *scYFP*- and *inYFP*-producing strains (c). Venn diagrams analysis was performed on DEGs list restricted to 2- fold change. Boxes give shortened YALI_ code of the gene and a short description for all the genes within a gene set. Abbreviations: dh - dehydrogenase, TF - Transcription Factor, TR – Transcription Repressor, AA- Amino Acid.

Table S.1. Strains used in this study.

Table S.2. RTqPCR primers used in this study.

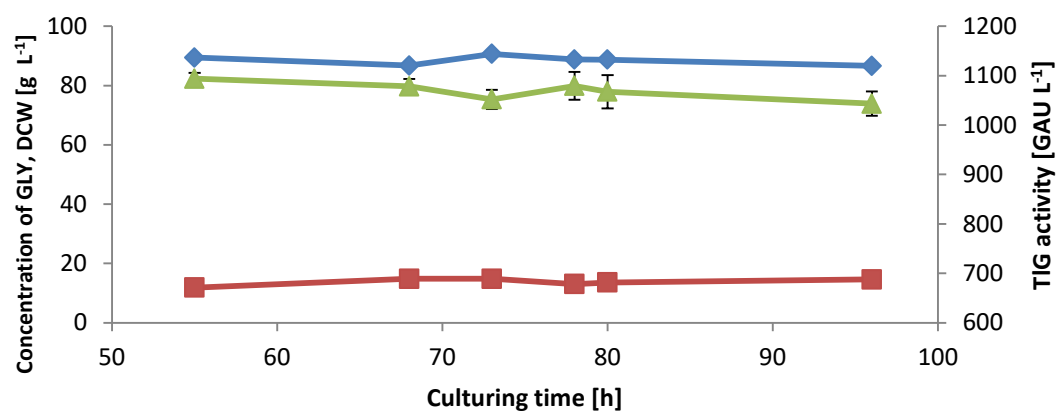
Table S.3. A complete list of DEGs with their basic characteristics as determined using DESeq2.- Excel file

Table S.4. Fold Change in expression level relative to the control strain for all genes assigned to the secretory pathway during the overproduction of heterologous secretory proteins by *Y. lipolytica*

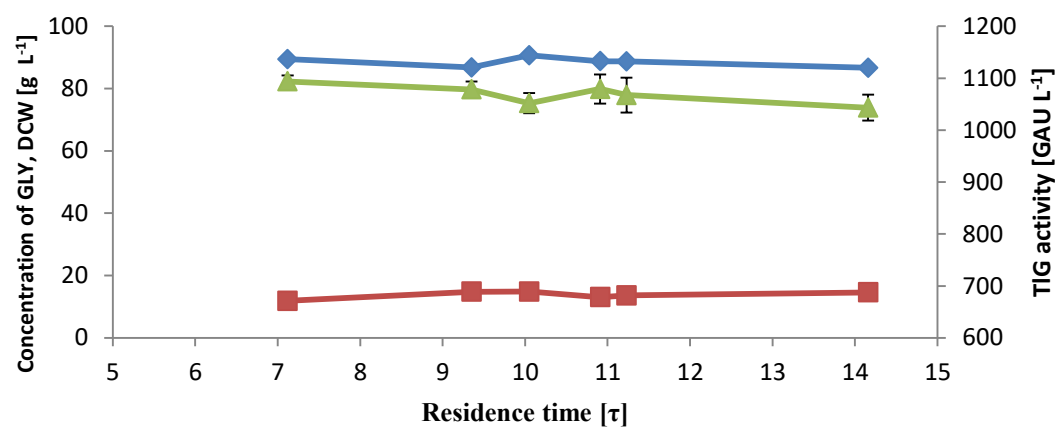
strains. The genes were categorized according to 27 categories distinguished for translocation, folding, maturation and secretion of r-Prots – Excel file

Fig.S.1

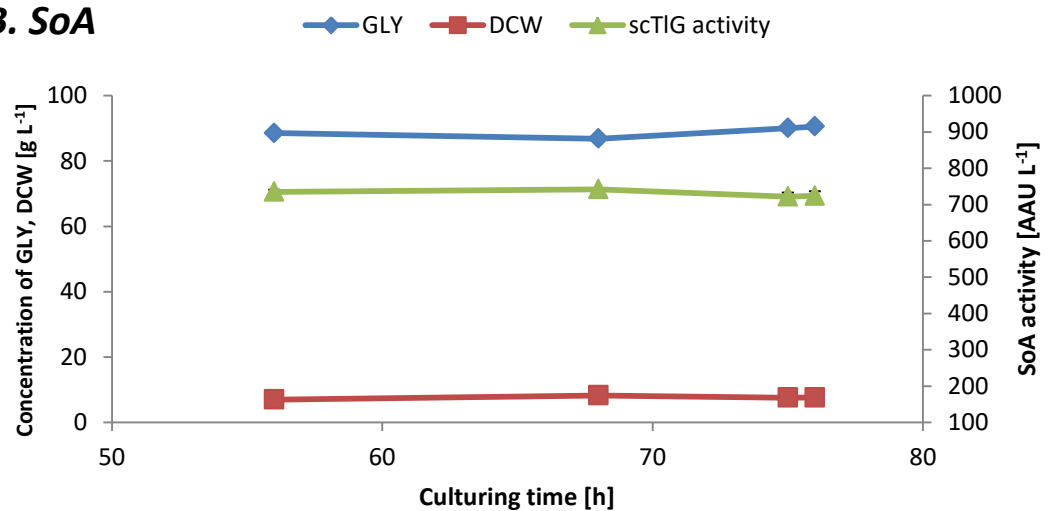
A. TIG



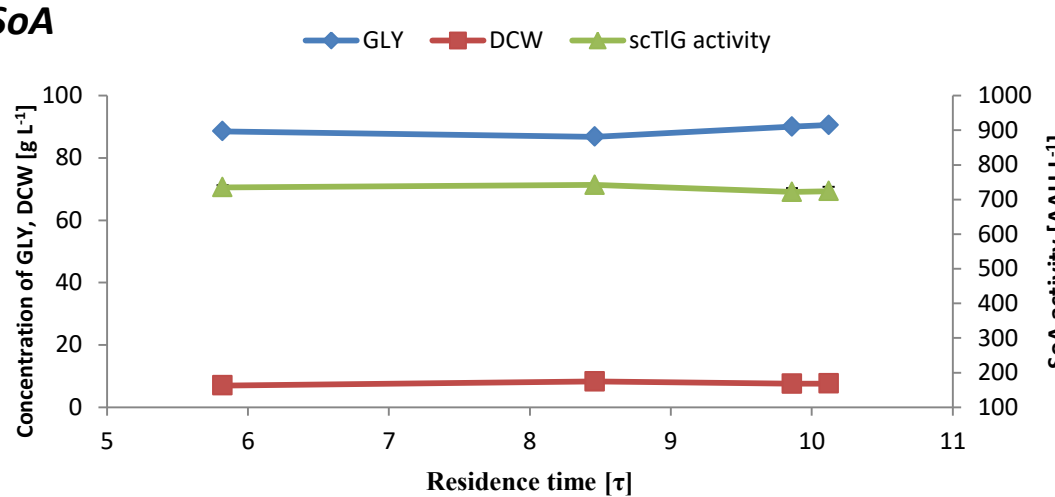
TIG



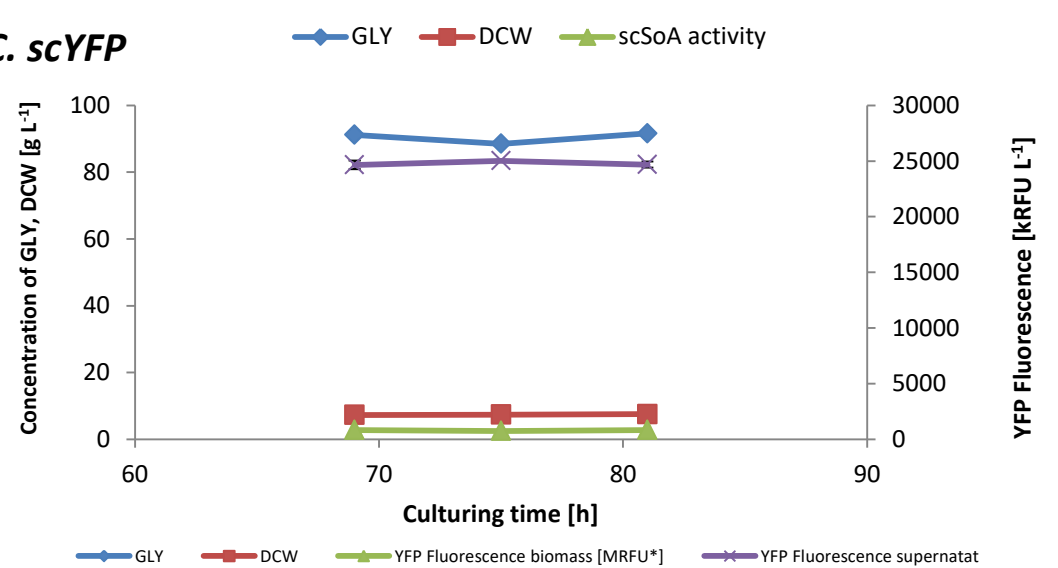
B. SoA



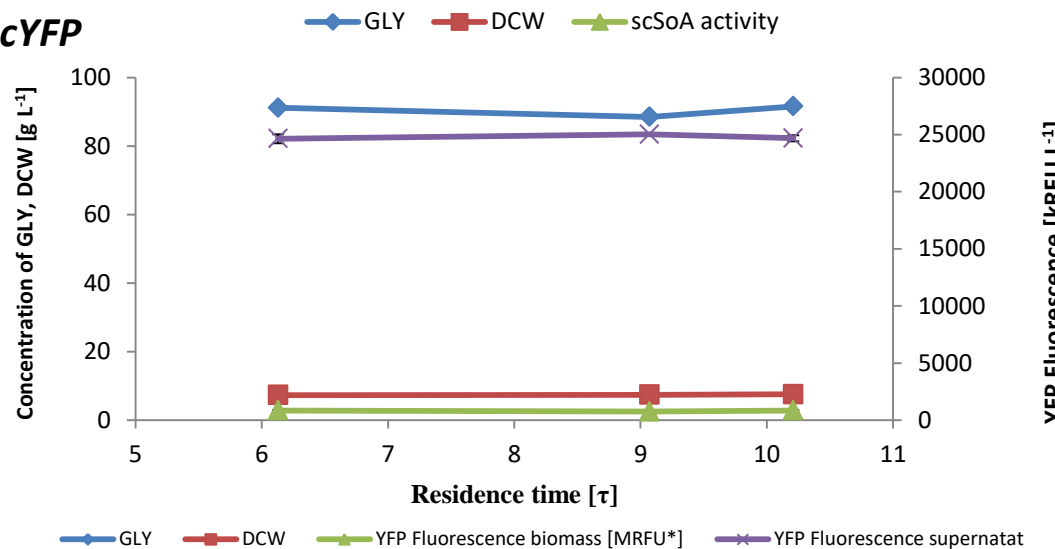
SoA



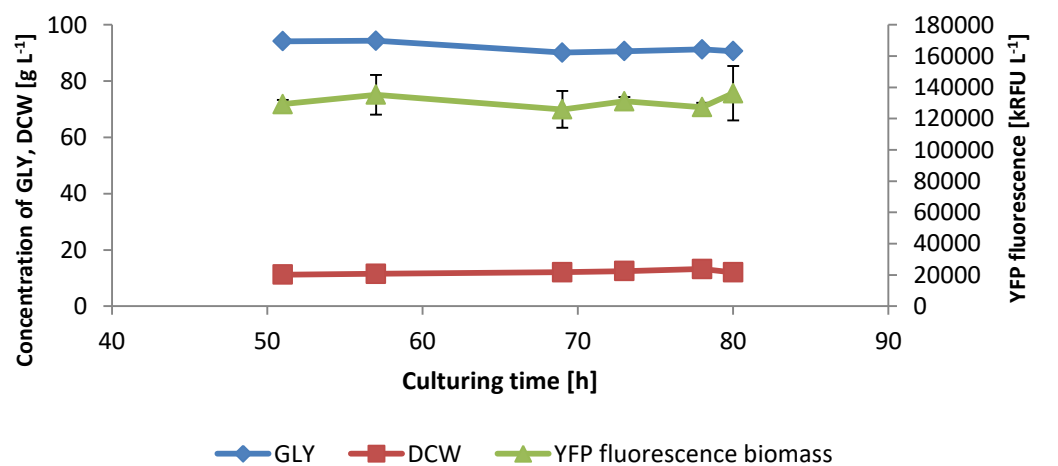
C. scYFP



scYFP



D. inYFP



inYFP

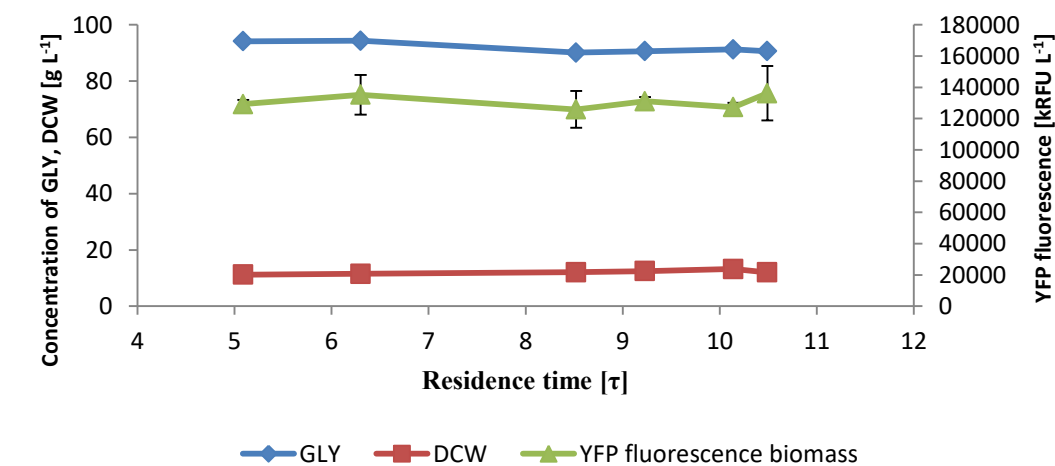


Fig.S.2



Fig.S.3

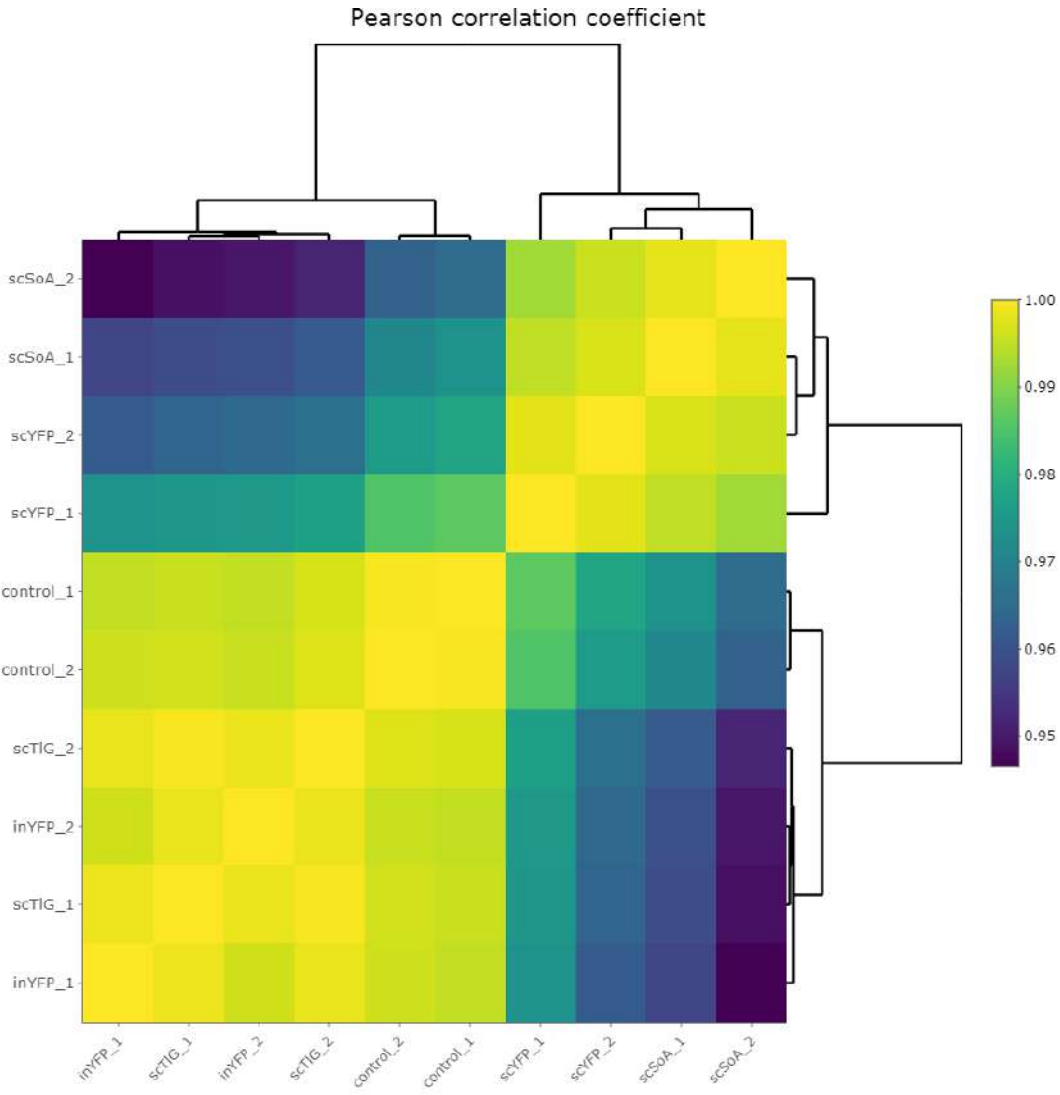
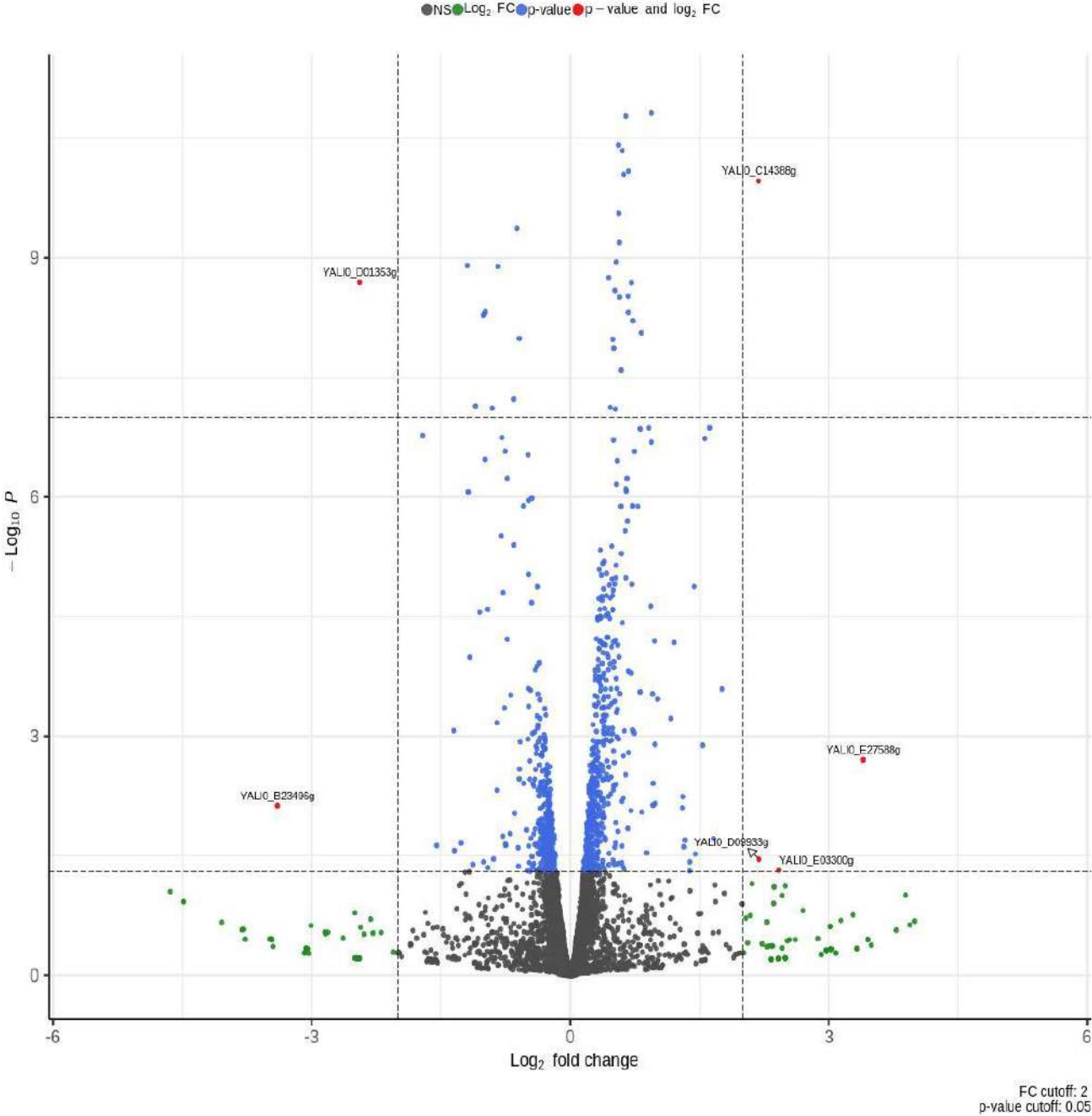


Fig.S.4

a. inYFP vs control

Differential Expression Volcano Plot

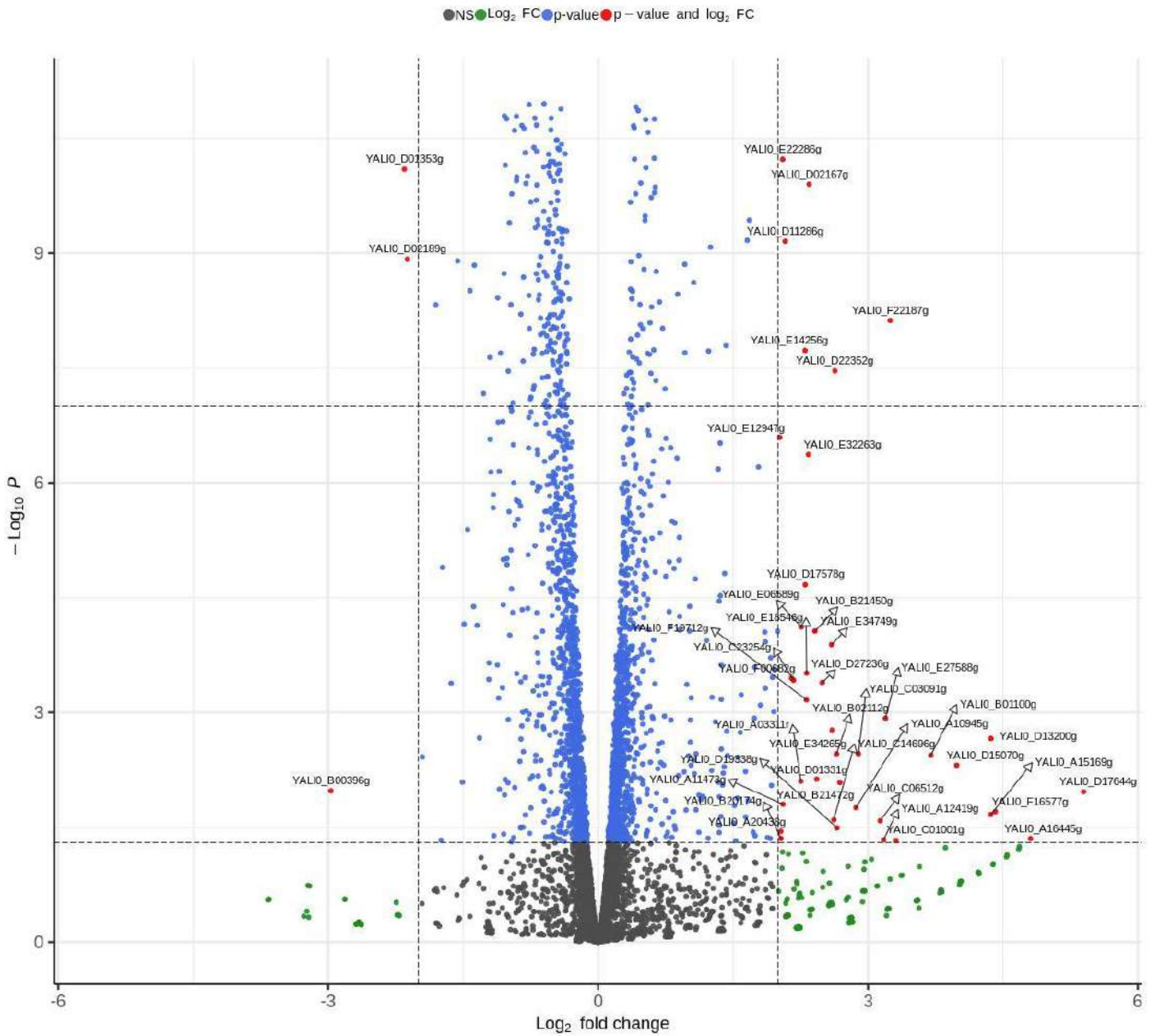
inYFP vs control



b. scYFP vs control

Differential Expression Volcano Plot

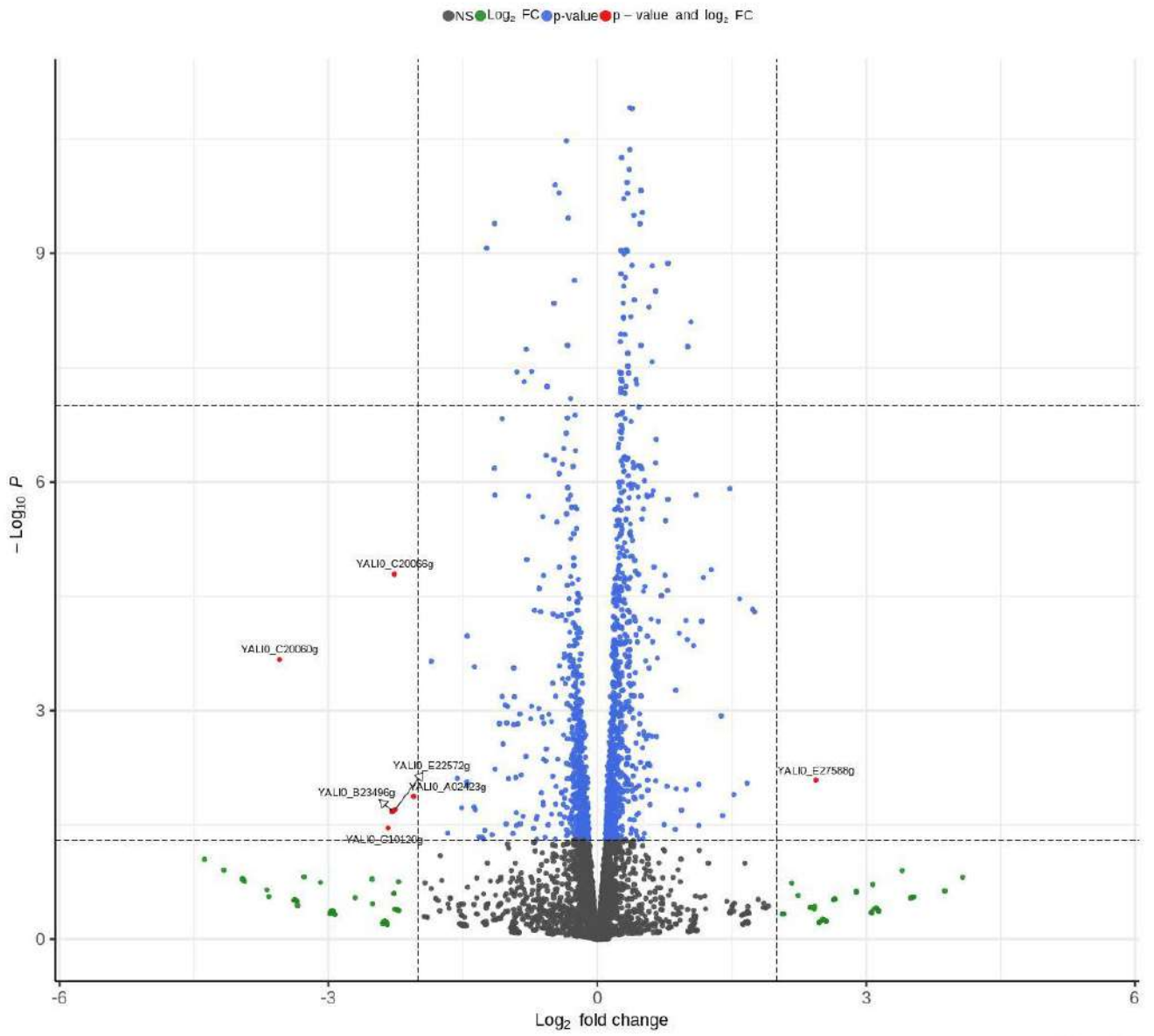
scYFP vs control



c. scTIG vs control

Differential Expression Volcano Plot

scTIG vs control

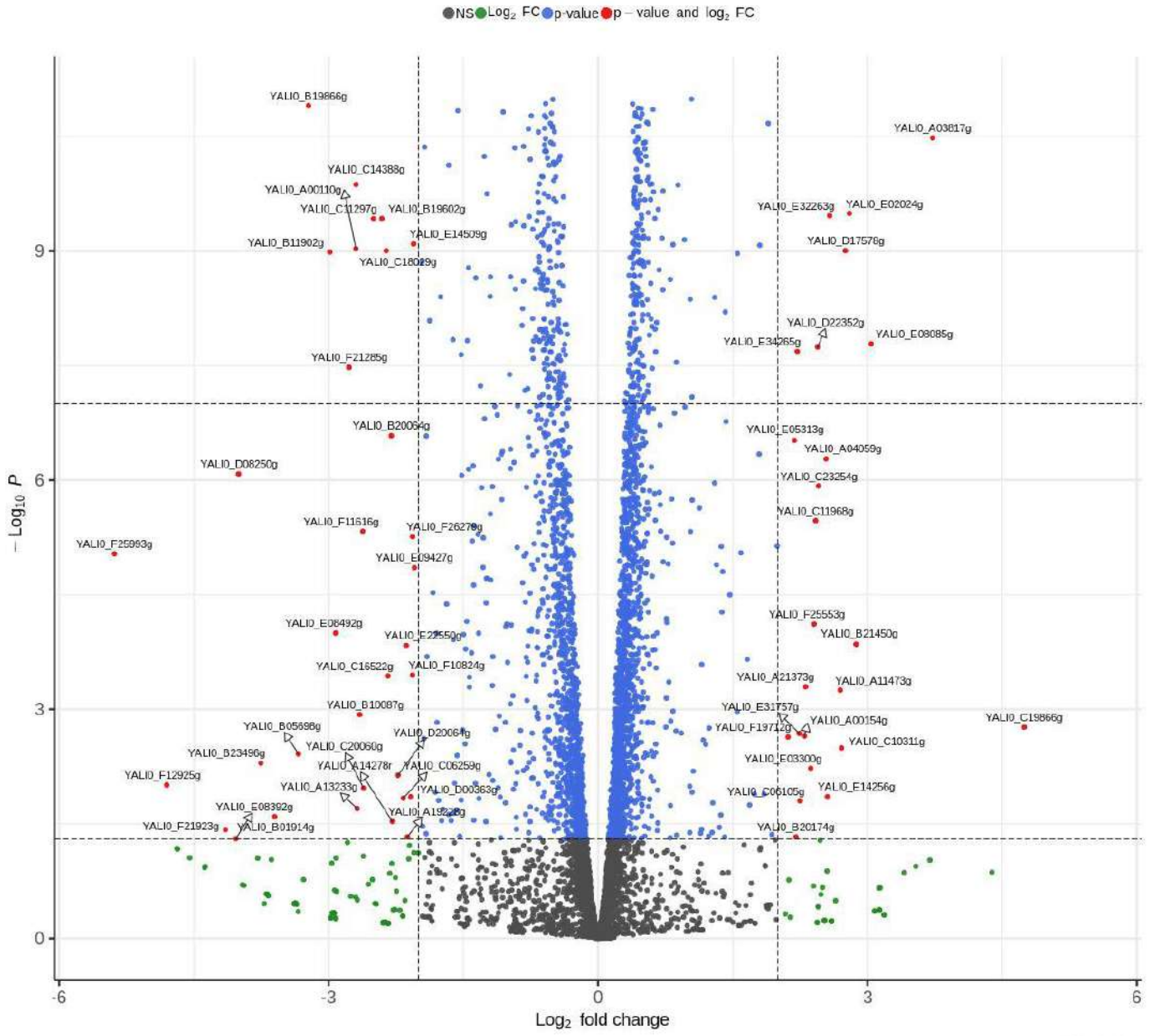


FC cutoff: 2
p-value cutoff: 0.05

d. scSoA vs control

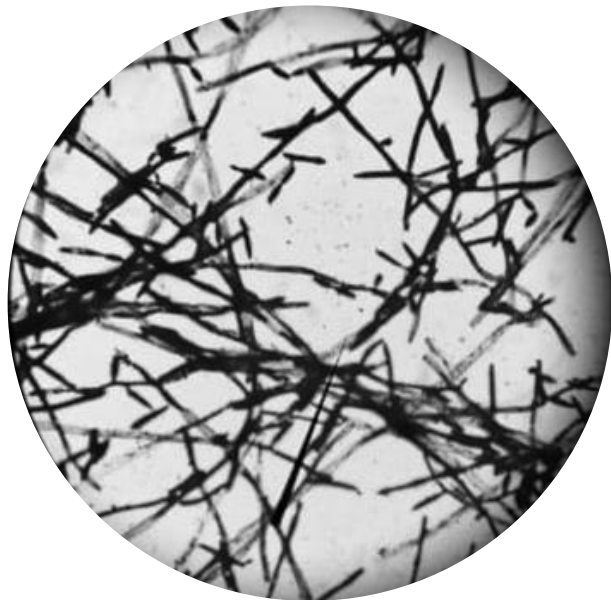
Differential Expression Volcano Plot

scSoA vs control

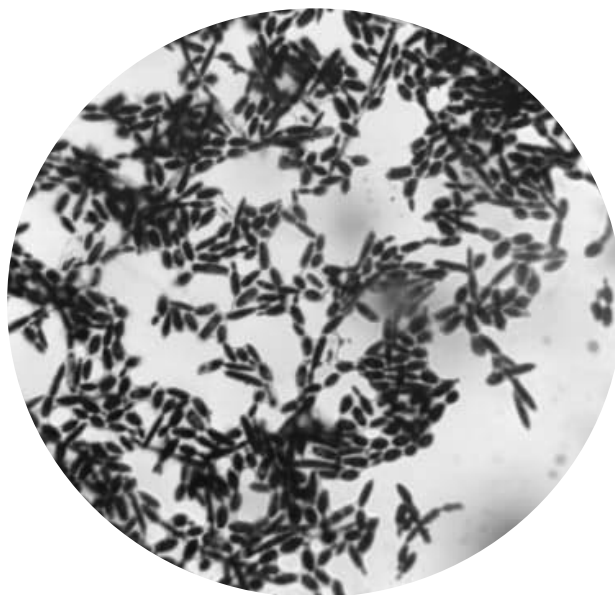


FC cutoff: 2
p-value cutoff: 0.05

Fig.S.5. A.



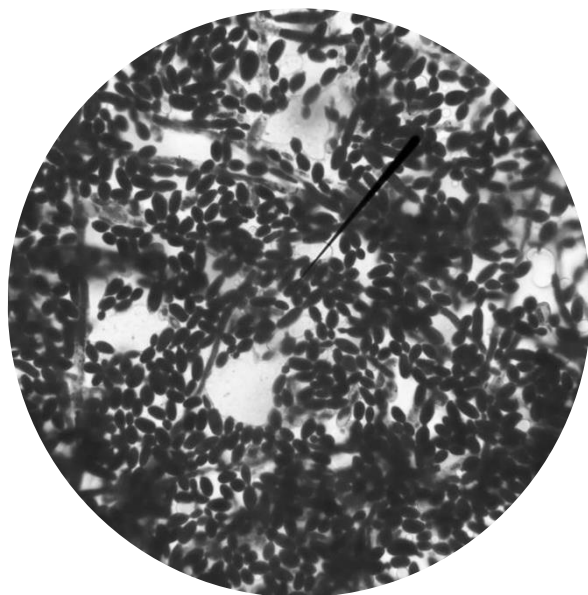
B.



C.



D.



E.

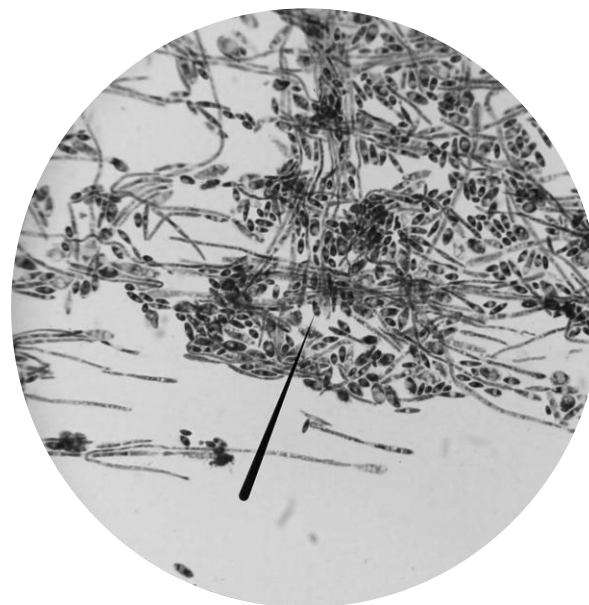


Fig.S.6.a

Aquaporin *F01210g*
 GATA-type domain *E31757g*
 Similar to *SHQ1_Ser-Glu-rich D13200g*
 Ankyrin domain-containing *D15070g*
PHO2 Acid phosphatase *C19866g*
 GPI-anchored cell wall protein *E02024g B01100g*
VPS13 family protein *A03817g*
 Dauer up-regulated-related *F22187g*
DeoR-type TF *E19030g*
OTU deubiquitinating enzyme *C03091g*
 Cytoskeleton-associated *A10945g*
 Potassium transporter *C10311g*
 Similar to *USO1 D23947g*
YPS3 aspartyl protease *3 D01331g*
 Membrane protein_signalling *B02112g*

SSA8/SSA-B HSP70 protein *D22352g*
CTT1 catalase T *E34265g*
CTT1 Catalase T *E34749g*
D27236g
 Alpha-ketoglutarate-dependent sulfonate dioxygenase *E14707g*
CCW12 cell wall mannoprotein *B21450g*
ADH5 alcohol dehydrogenase V *D02167g*
E32263g
HSP42 E18546g
AA3 ADP/ATP carrier protein *F19712g*
 Fungal, Nuclear protein *D17578g*
FDH Formate dh *E14256g*
E03938g
 Polyketide cyclase/dehydratase *E06589g*

Sodium/bile acid cotransporter 7 *F18744g*
 N-acetyltransferase domain-containing protein *E34221g*
 Ferric/cupric metalloredutase *F11825g*
A02783g
HSP31 glyoxalase *A00682g*
CWP12 GPI-anchored cell wall protein *E26125g*
AGP3 amino acid permease *E20713g*
E34441g
 Hemagglutinin/hemolysin-related protein *B17028g*
LIP1 Carboxylesterase/lipase type B *E10659g*
 YalifMt53
PRB1 protease B, vacuolar *A06435g*
D24816g
D13618g
ATPase_2 domain-containing protein *E33429g*
 Related to transcription activator amyR *C09482g*
YPS3 GPI-anchored aspartyl protease 3 *A13013g*
 Secreted aspartic protease 2 *C20273g*
 Ribonuclease H *B21494g*
A08195g
E28259g
B06160g
E22396g
A17020g
 Heat shock protein 42 (chaperone) *C03443g*
 Putative thioesterase superfamily protein *B14575g*
E08085g
A00154g
F19492g
D00451g
 Glutathione-independent glyoxalase *HSP31*-related *C22000g*

tRNA Thr (TGT) *A03311r*
C23254g
HSP31 glyoxalase *F00682g*
 MFS domain-containing protein *D00407g*
D11286g
 ATP-binding transporter *STE6 A11473g*
CWP11 cell wall protein *E22286g*
 Formate transporter 1 *E12947g*
D09889g
 Putative xylose and arabinose reductase *A19910g*
C03586g
 F-box domain-containing protein *A04059g*
 Alternative oxidase *AOX1 D09933g*
 Protein *ECM13 C11968g*
DDR48 heat shock protein *F10373g*
 Probable sugar transporter *F25553g*
 TauD domain-containing protein *F22825g*
C21252g
 Oligopeptide transporter *OPT2 F13959g*
 Related to transcription activator a *E34925g*
A12259g
C19822g
 ATP synthase F0-like protein *F26917g*
D07656g
 MFS efflux transporter of unknown specificity *C06105g*
 F-box domain-containing protein *E01562g*
GIN3 unknown function *F29909g*
C21274g
 DNA damage-responsive protein 48 *C17567g*
D04356g
E17083g

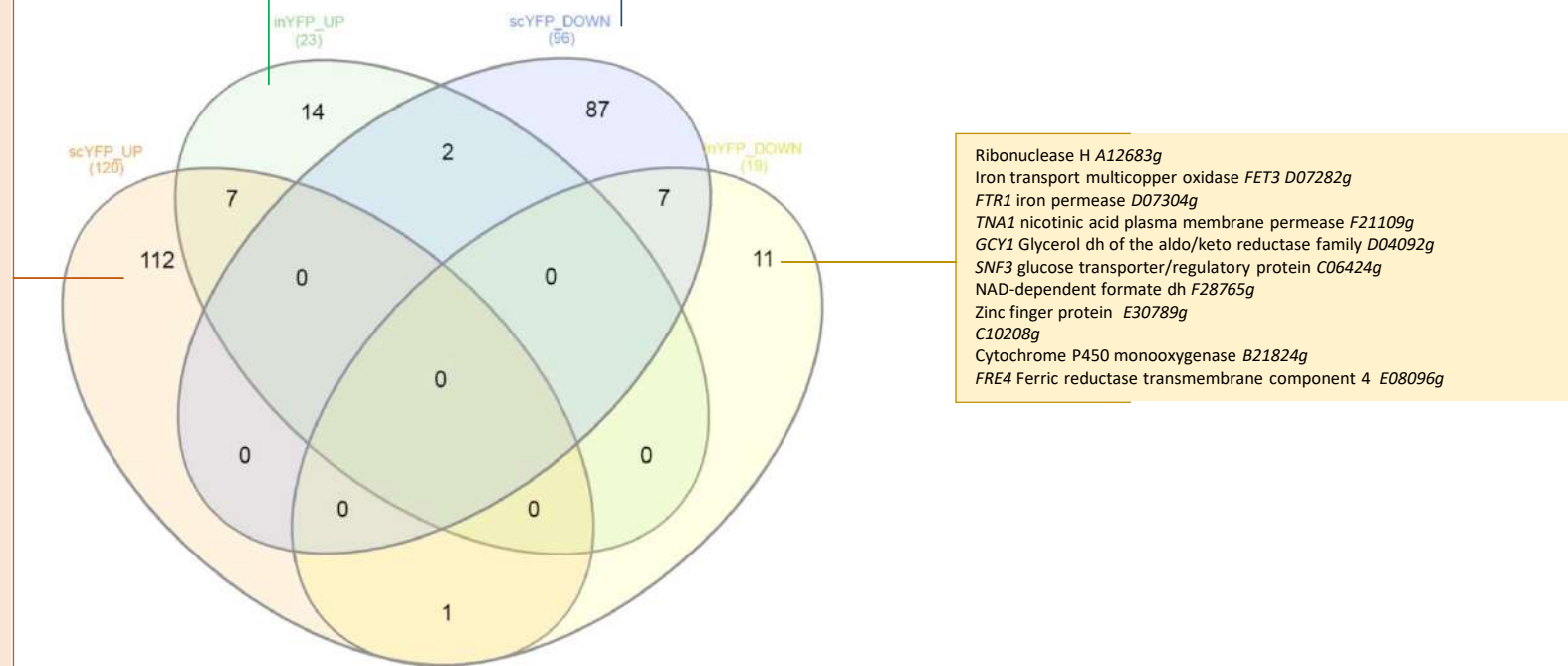
B11792g
AAL6 Acyl/aryl-CoA ligase, putative *C05885g*
PIC2 Mitochondrial phosphate carrier protein *C05753g*
C22143g
 Cleavage and polyadenylation specificity factor subunit *D08646g*
 Transcription factor *TT8 C03564g*
D17248g
OYE2 NADPH dh *B06413g*
STA1 extracellular alpha-1, 4-glucan glucosidase *C22924g*
OPT1 oligopeptide transporter *C20823g*
 Meiotically up-regulated gene 191 protein *A19888g*
 MFS domain-containing protein *E20427g*
GPH1 glycogen phosphorylase *F04169g*
 MFS domain-containing protein *D20196g*
SPS19 peroxisomal 2,4-dienoyl-CoA reductase *F01650g*
D02695g
 Acid phosphatase *PHOA F25773g*
C06534g
C17215g

Zinc-iron transporter *F15411g*
C06908g
PTH11-like integral membrane protein *A07623g*
PHO89 Na⁺-coupled phosphate transport protein *E23859g*
E29777g
FRE2 ferric (and cupric) reductase *B10846g*
TNA1 nicotinic acid plasma membrane permease *F05984g*
E12331g
 Putative transcriptional activator *D14872g*
RNH1 ribonuclease H *C09658g*
F26565g
C19844g
ZRT2 zinc transporter of the plasma membrane *F21659g*
OAR1 3-oxoacyl-[acyl-carrier protein] reductase *F29975g*

Maltose permease *MPH2 B00396g*
RsfA TF *E26763g*
 Filamentation TF *D02189g*
 Aspartic-type endopeptidase *B00374g*
C00869g
F15653g
STA1 extracellular alpha-1,4-glucan glucosidase *E05819g*
PBP2 PAB1 binding protein *C19558g*
F21285g
 PolyA transcript *F11616r*
 Caretonoid isomeroxygenase *C18029g*
AMD2 amidase *F16401g*
B06688g
DeoR-type TF *F19030g*
C08473g
 3-hydroxyisobutyrate dh, mitochondrial *F02607g*
LIP4 Triacylglycerol lipase precursor *E08492g*
 Putative Acyl CoA acyltransferase *C23705g*
 1, 4-butanediol diacrylate esterase *C22748g*
 Secreted aspartic protease 4 *C08899g*
 Triacylglycerol lipase *D23419g*
 General amino-acid permease *GAP1 B16522g*
D09185g
MUC1 GPI-anchored cell surface glycoprotein *A00176g*
MGF2 Mycelial growth factor *B19602g*
MEP2 ammonia permease *E27203g*
 Ribonuclease H1 *F26543g*
FAT3 Peroxisomal protein *B05456g*
D16181g
ZRT2 zinc transporter of the plasma membrane *F21659g*
 Oligopeptide transporter *OPT2 A00110g*
C23452g

XPR2 Alkaline extracellular protease *A08360g*
 Aquaporin *F00462g*
JEN1 carboxylic acid transporter protein *C15488g*
 NAD-dependent formate dh *B22506g*
CIT1 Mitochondrial 2-methyl citrate synthase *E00638g*
AAL3 Acyl/aryl-CoA ligase *E05951g*
 Ethanalamine utilization protein *EUTQ D02453g*
C23518g
coPTR2 Peptide transport protein *E04510g*
C09240g
ARA1 Protein of aldo/keto reductase family *F06974g*
 Formate dh 1 *B19976g*
 Acid phosphatase *PHO2 B19866g*
F22847g
MGH1 Methylglutaconyl-CoA hydratase *F22121g*
 BTB domain-containing protein *B19228g*
 Serum paraoxonase/arylesterase 3 *C21516g*
C01199g
FAT4 Peroxisomal protein *C09284g*
 Related to myo-inositol transport protein *ITR1 F18084g*
C04389g
 MARVEL-domain containing protein *C04455g*
B10384g
 PolyA transcript *C20801r*
C20713g
 Acid phosphatase *PHOA F23067g*
PRY1 protein precursor *D06149g*
 Ribonuclease H1 *A12705g*
 ARYL-alcohol dh AAD16-related *C12595g*
C12518g
 Protein *ECM13 B06028g*

DIP5 dicarboxylic amino acid permease *B20064g*
F30437g
A19360g
STE18 GTP-binding protein gamma subunit *A06699g*
C02167g
B05434g
C15268g
 F-box domain-containing protein *D14366g*
 Uncharacterized protein *F30107g*
 Marvel domain-containing protein *D06479g*
F18920g
ADH4 alcohol dh IV *E15818g*
A19250g
 Metallophos domain-containing protein *E34485g*
E25443g
A21120g
CWP1 Cell wall protein *E11517g*
GPR1 G-protein coupled receptor *E29315g*
E28548g
UGA4 GABA-specific permease *D00495g*
SPS191 Peroxisomal 2,4-dienoyl-CoA reductase *D01694g*
D10945g
D01716g
DPP1 diacylglycerol pyrophosphate phosphatase *C11297g*
TRX1 Thioredoxin-1-related *E09427g*



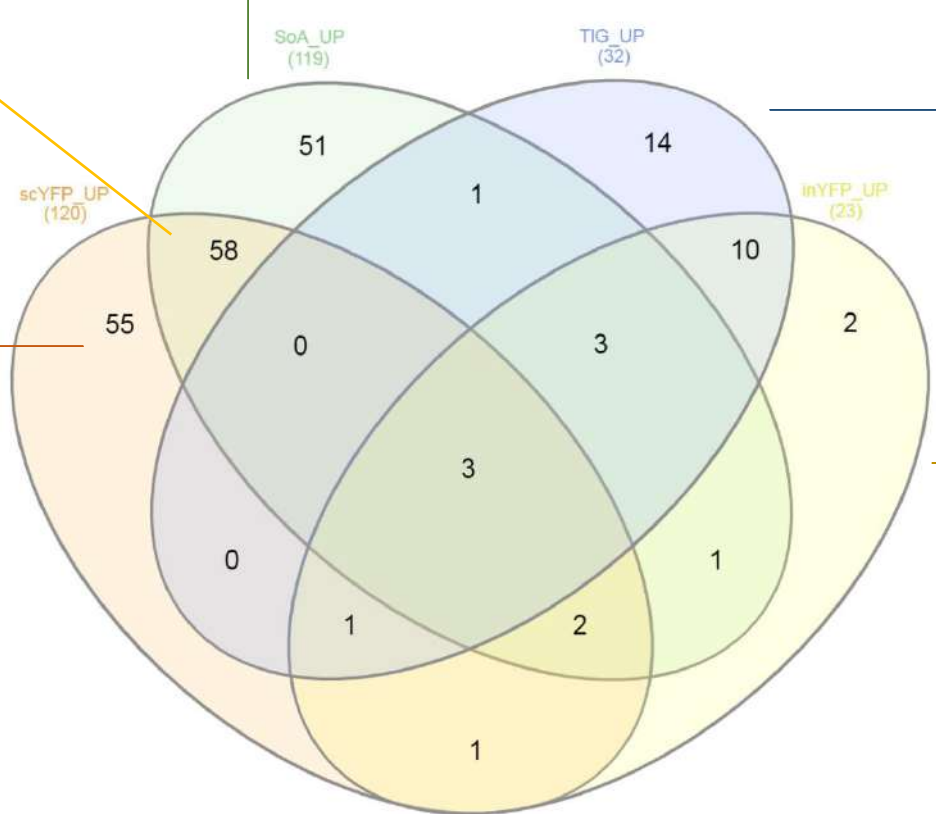
Ribonuclease H *A12683g*
 Iron transport multicopper oxidase *FET3 D07282g*
FTR1 iron permease *D07304g*
TNA1 nicotinic acid plasma membrane permease *F21109g*
GCY1 Glycerol dh of the aldo/keto reductase family *D04092g*
SNF3 glucose transporter/regulatory protein *C06424g*
 NAD-dependent formate dh *F28765g*
 Zinc finger protein *E30789g*
C10208g
 Cytochrome P450 monooxygenase *B21824g*
FRE4 Ferric reductase transmembrane component 4 *E08096g*

Fig.S.6.b

GATA-type domain-containing protein <i>E31757g</i>	Sugar transporter <i>F25553g</i>
Acid phosphatase precursor <i>PHO2 C19866g</i>	ATP synthase F0-like protein <i>F26917g</i>
<i>E02024g</i>	<i>C06105g</i>
<i>A03817g</i>	<i>F29909g</i>
Dauer up-regulated protein <i>F22187g</i>	<i>E17083g</i>
<i>DeoR</i> -type TF <i>E19030g</i>	FAD-binding FR-type domain-containing protein <i>F11825g</i>
Potassium transporter <i>C10311g</i>	GPI-anchored cell wall protein <i>CWP12 E26125g</i>
Similar to <i>USO1 D23947g</i>	<i>F18458g</i>
<i>SSA8/SSA-B HSP70</i> protein <i>D22352g</i>	Amino acid permease <i>AGP3 E20713g</i>
Catalase T, cytosolic <i>CTT1 E34265g</i>	<i>D24816g</i>
Catalase T <i>CTT1 E34749g</i>	ATP-binding protein <i>E33429g</i>
<i>E14707g</i>	<i>B21494g</i>
Cell wall mannoprotein <i>CCW12 B21450g</i>	<i>E28259g</i>
PKS_ER domain-containing protein <i>D02167g</i>	<i>A17020g</i>
<i>E32263g</i>	Heat shock protein 42 (chaperone) <i>C03443g</i>
ADP/ATP carrier protein <i>AAC3 F19712g</i>	Thioesterase superfamily protein <i>B14575g</i>
Fungal, Nuclear protein <i>D17578g</i>	<i>E08085g</i>
Formate dh <i>FDH E14256g</i>	<i>A00154g</i>
<i>E03938g</i>	Glutathione glyoxalase <i>HSP31 C22000g</i>
<i>C23254g</i>	Acyl/aryl-CoA ligase <i>AAL6 C05885g</i>
Glutathione glyoxalase <i>HSP31 F00682g</i>	Mitochondrial phosphate carrier protein <i>PIC2 C05753g</i>
ATP-binding transporter protein <i>STE6 A11473g</i>	Transcription factor <i>HLH C03564g</i>
GPI-anchored cell wall protein <i>CWP11 E22286g</i>	NADPH dh <i>OYE2 B06413g</i>
<i>E12947g</i>	<i>E20427g</i>
<i>D09889g</i>	Peroxisomal 2,4-dienoyl-CoA reductase <i>SPS19 F01650g</i>
Xylose and arabinose reductase <i>A19910g</i>	<i>D02695g</i>
F-box domain-containing protein <i>A04059g</i>	Acid phosphatase <i>PHOA F25773g</i>
Alternative oxidase <i>AOX1 D09933g</i>	<i>C06534g</i>
<i>C11968g</i>	<i>C17215g</i>

<i>E03300g</i>	GPI-anchored aspartyl protease 3 <i>YPS3 D22957g</i>
<i>A21373g</i>	<i>D21802g</i>
Monocarboxylate transporter <i>E05313g</i>	Mitochondrial peroxiredoxin with thioredoxin peroxidase activity <i>PRX1 A19426g</i>
<i>E04026g</i>	Multidrug resistance protein <i>PGP1 F20174g</i>
Glycerol dh of the aldo/keto reductase <i>GCY14 E18348g</i>	<i>EXG1 F05390g</i>
Arrestin motif containing protein <i>B21318g</i>	Glucan 1, 3-beta-glucosidase precursor <i>EXG1 B12078g</i>
Plasma membrane iron permease <i>FTR1 D06688g</i>	Methylenetetrahydrofolate dh [NAD+] <i>MTD1 F28765g</i>
Cytochrome P450 monooxygenase <i>B21824g</i>	<i>B18392g</i>
Glycerol dh of the aldo/keto reductase <i>GCY1 D04092g</i>	Glyoxylate pathway regulator <i>GPR1 C23617g</i>
<i>C08393g</i>	Ribonuclease H1 <i>E10725g</i>
<i>A21461g</i>	Hydroxymethylglutaryl-CoA lyase <i>MVA1 B22550g</i>
Hydroxytrimethyllysine aldolase <i>CNT2 A21417g</i>	Methylcrotonoyl-CoA carboxylase, subunit beta <i>MCC2 E12595g</i>
Metallophos domain-containing protein <i>E34463g</i>	High-affinity permease for basic amino acids <i>ALP1 F01254g</i>
<i>A12441g</i>	Ribonuclease H1 <i>RNH1 F25729g</i>
Aspartic protease <i>YPS3 C10135g</i>	Multidrug resistance protein <i>SCR1 F25597g</i>
<i>E09625g</i>	NADPH-dependent aldehyde reductase <i>ARI1 D07062g</i>
Iron permease <i>FTR1 D07304g</i>	Acid extracellular protease precursor <i>AXP1 D10967g</i>
Phosphodiesterase <i>E31581g</i>	Peptidase A1 domain-containing protein <i>B20526g</i>
<i>PDB2</i> 3-methyl-2-oxobutanoate dh <i>F05038g</i>	Flocculation protein <i>FLO11 E00792g</i>
Sulfoquinovosyl transferase <i>SQD2 E26279g</i>	Repressor of RNA polymerase III transcription <i>MAF1 F10541g</i>
Ribose 5-phosphate isomerase <i>F01628g</i>	<i>A18964g</i>
Component of membrane <i>B18986g</i>	Mitochondrial carrier protein <i>RIM2 F05500g</i>
Mitochondrial cytochrome c oxidase <i>F24409g</i>	Formate dh 1 <i>F13937g</i>
Branched-chain AA transaminase <i>BAT2 F19910g</i>	Heat shock protein <i>HSP12 D20526g</i>
2-oxoisovalerate (lipoate) dh E1 alpha chain <i>PDA2 D08690g</i>	Daughter cell-specific protein <i>DSE1 D09823g</i>
NAD-specific glutamate dh <i>GDH2 E09603g</i>	

Sodium/bile acid cotransporter 7 <i>F18744g</i>	Aquaporin <i>F01210g</i>
N-acyltransferase domain-containing protein <i>E34221g</i>	<i>D13200g</i>
<i>A02783g</i>	Ankurin domain-containing protein <i>D15070g</i>
<i>A00682g</i>	<i>B01100g</i>
<i>E34441g</i>	OTU deubiquitinating enzyme <i>C03091g</i>
Hemagglutinin/hemolysin-related protein <i>B17028g</i>	Kinetochore protein <i>NUF2 A10945g</i>
Carboxylesterase/lipase type B <i>LIP1 E10659g</i>	GPI-anchored aspartyl protease 3 <i>YPS3 D01331g</i>
<i>YalifMt53</i>	Membrane protein_signalling <i>B02112g</i>
Vacuolar protease B <i>PRB1 A06435g</i>	<i>D27236g</i>
<i>D13618g</i>	Alpha-ketoglutarate-dependent sulfonate dioxygenase <i>JLP1 B21472g</i>
Related to transcription activator amyR <i>C09482g</i>	<i>HSP42 E18546g</i>
GPI-anchored aspartyl protease 3 <i>YPS3 A13013g</i>	Polyketide cyclase/dehydratase domain-containing protein <i>E06589g</i>
Secreted aspartic protease 2 <i>C20273g</i>	tRNA Thr (TGT) <i>A03311r</i>
<i>A08195g</i>	Transmembrane transporter activity <i>D00407g</i>
<i>B06160g</i>	<i>D11286g</i>
<i>E22396g</i>	<i>C03586g</i>
<i>F19492g</i>	Heat shock protein <i>DDR48 F10373g</i>
<i>D00451g</i>	Alpha-ketoglutarate-dependent sulfonate dioxygenase <i>F22825g</i>
<i>B11792g</i>	<i>C21252g</i>
<i>C22143g</i>	Oligopeptide transporter <i>OPT2 F13959g</i>
<i>D08646g</i>	Related to transcription activator a <i>E34925g</i>
<i>D17248g</i>	<i>A17259g</i>
alpha-1,4-glucan glucosidase <i>STA1 C22924g</i>	<i>C19822g</i>
Oligopeptide transporter <i>OPT1 C20823g</i>	<i>D07656g</i>
Meiotically up-regulated gene 191 gene <i>A19888g</i>	F-box domain-containing protein <i>E01562g</i>
Glycogen phosphorylase <i>GPH1 F04169g</i>	<i>C21274g</i>
<i>D20196g</i>	DNA damage-responsive protein 48 <i>C17567g</i>
	<i>D04356g</i>



<i>C08965g</i>
Polyketide cyclase/hydratase <i>C05236g</i>
Carboxylic acid transporter <i>JEN1 C15488g</i>
Acyl-CoA acyltransferase <i>C23705g</i>
<i>F30789g</i>
Malate synthase <i>MLS2 D19140g</i>
Acid extracellular protease <i>AXP1 E33363g</i>
<i>C08473g</i>
Acid phosphatase <i>PHO2 D03465g</i>
Iron transport multicopper oxidase <i>FET3 D06754g</i>
A-agglutinin anchor subunit <i>AGA1 D09185g</i>
Cell wall protein <i>CWP1 E11517g</i>
<i>F20548g</i>
<i>F21285g</i>

Methyltransferase <i>C14388g</i>
<i>TNA1</i> nicotinic acid permease <i>F05984g</i>

Fig.S.6.c

GPI mannosyltransferase *F25993g*
 Urea transport protein *DUR3 B04202g*
 Ammonia permease *MEP2 F12925g*
 Nitrogen regulatory protein areA *D20482g*
 Oligopeptide transporter *OPT2 F09691g*
UAPC Purine permease *E27852g*
URC1 GTP cyclohydrolase *D08250g*
 Purine-cytosine permease *FCY2 F30569g*
 PolyA transcript *B23496r*
URC7 Uracil catabolism protein *E20647g*
B05698g
XBP1 Transcriptional repressor *F16511g*
 Xanthine oxidase *A06974g*
D02046g
 GABA-specific permease *UGA4 E07205g*
URC6 UMP pyrophosphorylase *E20625g*
B10087g
A13233g
 General amino acid permease *GAP1 C17237g*
ARG8 acetylornithine aminotransferase *C05709g*
 Oligopeptide transporter *OPT2 E19294g*
MHO2 protein unknown function *C03696g*
STL1 member of the sugar permease *C16522g*
D06864g
AVT1 neutral AA transporter *F30063g*
 Putative MSF transporter *D20064g*
E22506g
 Putative transcriptional activator *D14872g*
PRC1 carboxypeptidase *D08052g*
 Putative dh *D01738g*
 Carboxypeptidase Y precursor *C00803g*
E30701g
 F-box protein *COS111 B01166g*
AMD2 Putative amidase *E11847g*
 GLOBIN domain-containing protein *D15928g*
C19239g
B14685g
C23837g
 Transcriptional regulator *F25113g*
 Oxidation-reduction process *E19789g*
C03806g
AXP1 Acid extracellular protease *B05654g*
 General alpha-glucoside permease *D15488g*
 [PSI⁺] induction protein 2 *E06809g*
 Serine/threonine- protein kinase *SKS1 B08558g*
 L-ornithine n(5)-monooxygenase *F26983g*
RNY1 ribonuclease T2 *F25839g*
FRE2 ferric (and cupric) reductase *B13046g*
KTR1 mannosyltransferase *A06589g*

RsfA TF *E26763g*
 Filamentation TF *D02189g*
 Acid protease precursor *AXP1 B00374g*
C00869g
STA1 alpha-1,4-glucan glucosidase *F15653g*
STA1 alpha-1,4-glucan glucosidase in frame stop codon *E05819g*
PBP2 PAB1 binding protein *C19558g*
F21285g
 PolyA transcript *F11616g*
 Carotenoid isomeroxygenase *C18029g*
AMD2 amidase *F16401g*
B06688g
FLO11 TF *F19030g*
 Big_8 domain-containing protein *C08473g*
LIP4 Triacylglycerol lipase precursor *E08492g*
 Putative Acyl CoA acyltransferase *C23705g*
 Serum paraoxonase/arylesterase 3 *C21516g*
 Triacylglycerol lipase *D23419g*

D00363g
 N- carbamyl-L-cysteine amidohydrolase *A02607g*
 Aa_trans domain-containing protein *F07106g*
E28831g
 Ribonuclease T2-like *RNY1 C23991g*
D19602g
SGA1 sporulation specific factor 1,4-alpha-glucosidase *E05203g*
E11330g
B02200g
E31772g
 NADP- glutamate dh *GDH3 F17820g*
 Fe2OG dioxygenase domain-containing protein *E04070g*
 WD_REPEATS_REGION *E32879g*
 PENDE Acyl-coenzyme A:6-aminopenicillanic-acid acyltransferase *E30833g*
 Microsomal glutathione S-transferase 3 *E03586g*
C04301g
B01210g
PEP12 syntaxin *E11825g*
A14608g
 Endochitinase *CTS2 F04532g*
E06105g
 Palmitoyltransferase *PFA5 E01804g*
 Methyltransferase *A02673g*
B09955g
F25641g
 Ribosomal N-acetyltransferase *D06391g*
 Guanine nucleotide-binding protein alpha subunit *GPA1 E11627g*
C06259g
PIB1 phosphatidylinositol(3)-phosphate *D15510g*
 Putative C2H2 zinc finger transcriptionalregulator *F19822g*
AMO2 Amine oxidase *C18315g*
B03542g
MUM2 meiotic protein *B17160g*
E31559g
 Related to lincomyacin-condensing protein lmbA *C20669g*
YCF1 glutathione S-conjugate transporter *A19580g*
 Beta-lactamase domain-containing protein *E31889g*
YPS3 GPI-anchored aspartyl protease 3 *F10549g*
 Multidrug-efflux transporter *C08228g*
 Oligopeptide transporter *OPT2 B02398g*
 26S proteasome non-ATPase regulatory subunit *D15092g*
 Oligopeptide transporter *OPT2 F18964g*
 Transport and Golgi organization protein *D01760g*
SPS192 Peroxisomal 2,4-dienoyl-CoA reductase *E14322g*
SEC7 domain-containing protein *D01221g*
ATG8 E02662g
CWP1 Cell wall protein *C22836g*
 Vesicle intracellular protein transport (SNARE complex) *F26279g*

B10384g
 PolyA transcript *C20801*
C20713g
 Putative acid phosphatase *F23067g*
PRY1 protein precursor *D06149g*
 Ribonuclease H *A12705g*
 Putative aryl-alcohol dh *C12595g*
C12518g
B06028g
 DIPS dicarboxylic amino acid permease *B20064g*
F30437g
STE18 GTP-binding protein *A06699g*
B05434g
C15268g
D14366g
F30107g
 Secreted aspartic protease 4 *C08899g*
C01199g
 Arrestin domain-containing protein *F18920g*

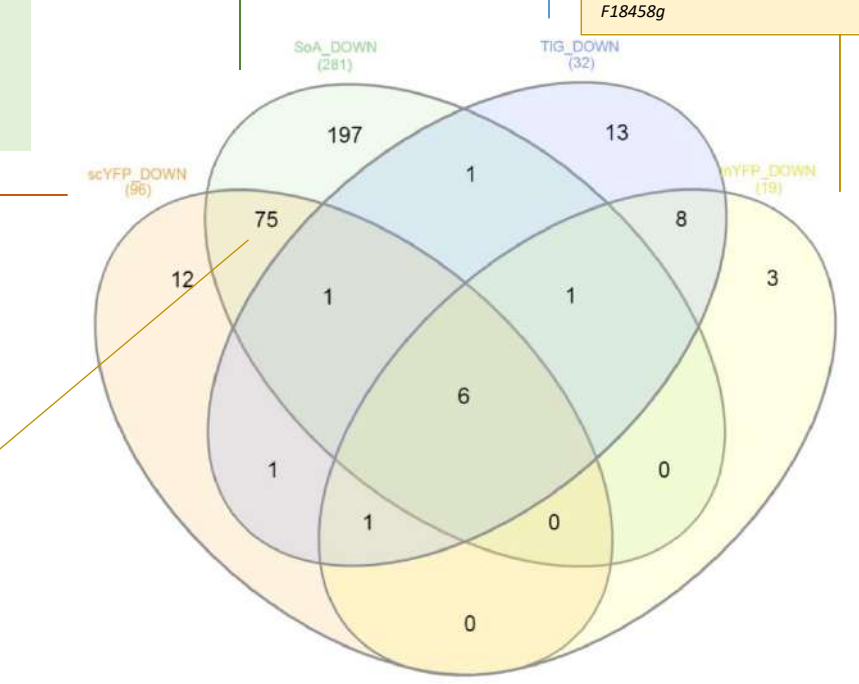
NADPH dh *OYE3 D01243g*
 Phosphoinositide phosphatase *E01408g*
F09031g
 Ser/thr protein kinase *IME2 F03113g*
A04323g
 Xenobiotic compound monooxygenase *B10604g*
E25806g
STB3 SIN3 protein-binding protein *C09232g*
PPS1 protein tyrosine phosphatase *B04290g*
 AB hydrolase-1 domain-containing protein *E19899g*
D26323g
B09867g
B09867g
 Uricase (Urate oxidase) *F09251g*
DUF202 F14685g
 Putative lipase *ATG15 F06358g*
 Hydroxyisourate hydrolase *F22605g*
 YeeE/YedE family *F01408g*
 Poly [ADP-ribose] polymerase *C17061g*
PRC1 carboxypeptidase *C23661g*
 F-box motif containing protein *D17512g*
UPF0183 protein *D17666g*
C06710g
BIO2 Biotin synthase *D15400g*
FTR1 iron permease *A04917g*
RHO3 GTPase *B13662g*
*GEF1*voltage-gated chloride channel protein *C20053g*
 Carboxypeptidase Y precursor *B01408g*
GLT1 glutamate synthase (NAPDPH) *B19998g*
AHC1 A10703g
E18722g
E20229g
A14564g
F03619g
F00220g
B09031g
C11165g

Maltose permease *MPH2 B00396g*
 3-hydroxyisobutyrate dh *F02607g*
 Peroxisomal protein *FAT3 B05456g*
 Carboxylic acid transporter protein *JEN1 C15488g*
 Formate dh *B19976g*
 Methylglutaconyl-CoA hydratase *MGH1 F22121g*
 Peroxisomal protein *FAT4 C09284g*
 Related to myo-inositol transport protein *ITR1 F18084g*
A19360g
C02167g
 Alcohol dh IV *ADH4 E15818g*
 Metallophos-domain-containing protein *E34485g*

ATG13 F03432g
B21626g
 GPI-anchored cell surface glycoprotein *STA4/FLO11 E20251g*
D25586g
E08382g
F00946g
F26961g
C03135g
 Septin-type G domain-containing protein *F09009g*
C09372g
ERG2 C-8 sterol isomerase *B17204g*
D21593g
YHM2 Citrate and oxoglutarate carrier *B10736g*
B21714g
D13266g
CYB2 lactate dh cytochrome b2 *D12661g*
 OHCU_decarbox domain-containing protein *A01111g*
 Histidine acid phosphatase *E27093g*
 Adhesion and aggregation factor *E23925g*
 STAS domain-containing protein *E13442g*
E11319g
TGL3 Triacylglycerol lipase *D17534g*
ATG22/23 C03608g
MHY1 C2H2-type zinc finger protein *B21582g*
HOL1 member *F03795g*
FCY1 Cytosine *F22517g*
CLN1 G1/S-specific cyclin *C15114g*
STA1 alpha-1, 4-glucan glucosidase *E22550g*

Metallothionein-II *MTP2 C20066g*
 Zn(2)-C6 fungal type domain-containing protein *F18326g*
 Dihydroxyacetone kinase *DAK1 E20691g*
 Related to transcription activator amyR *C07821g*
 Acid extracellular protease precursor *AXP1 D10967g*
 N-alkane inducible cytochrome P450 *ALK1 E25982g*
 Dihydroxyacetone kinase *DAK3 F01606g*
 Ferric/cupric metalloredutase *F11825g*
 Urea transporter *DUR3 B04202g*
E34441g
A01089g
 NADPH dh 1 *OYE1 A01485g*
 Xylose and arabinose reductase *A19910g*

NAD-dependent formate dh *F28765g*
 Cytochrome P450-like protein *B21824g*
F18458g



Promoter	Strains	Characteristics	Use	Supplier/reference
4UASpTEF	<i>Yarrowia lipolytica</i> Po1h	Genotype: <i>Mata</i> , <i>ura3-302</i> , <i>xpr2-322</i> , <i>axp1-2</i> , <i>leu2-270::LEU2</i> Phenotype: <i>Ura⁻</i> , Δ <i>AEP</i> , Δ <i>AXP</i> , <i>Suc⁺</i>	Parental strain - host for expression of heterologous <i>SoA</i> , <i>TIG</i> , <i>YFP</i> gene	Celińska et al. 2020
	Po1h_URA3+ (GGY231)	Po1h derivative transformed with Ura3 cassette	Control for Po1h derivatives	Korpys-Woźniak et al. 2020
	Po1h_scSoA (GGY237)	Po1h derivative with expression cassette bearing single TU with SP1 upstream fused to gene encoding α -amylase	Expression of secreted heterologous α -amylase from <i>S. oryzae</i>	Korpys-Woźniak et al. 2020
	Po1h_scTIG (GGY228)	Po1h derivative with expression cassette bearing single TU with SP1 upstream fused to gene encoding glucoamylase	Expression of secreted heterologous glucoamylase from <i>T. lanuginosus</i>	This study
	Po1h_scYFP (GGY251)	Po1h derivative with expression cassette bearing single TU with SP1 upstream fused to gene encoding YFP	Expression of secreted heterologous YFP	Korpys-Woźniak et al. 2020
	Po1h_inYFP (GGY245)	Po1h derivative with expression cassette bearing single TU with gene encoding YFP, without SP1	Expression of cytosolic heterologous YFP	Korpys-Woźniak et al. 2020
	JM109 <i>Escherichia coli</i>	F'(traD36,proAB+,lacIq, Δ (lacZ)M15), endA1,recA1,hsdR17 (rk-,mk+), mcrA,supE44, λ -gyrA96, relA1, Δ (lacproAB),thi-1	Host for routine cloning, vector assembly and propagation	Sigma-Aldrich

Table S.1.

Primer	Sequence 5'→3'	Reference
r-t_SoA_Fw	GTAACAACGTGGGAATCCGAAT	Korpys-Woźniak et al. 2020
r-t_SoA_Rev	CCCTGGCCGTTCGAAGTAG	
r-t_TIG_Fw	TCTGATGTTGTTTGGGAATCTGA	
r-t_TIG_Rev	TGATGTCTCCACCAAGAGTTTACTG	
r-t_YFP_Fw	CAACAGCCACAACGTCTATATCATG	
r-t_YFP_Rev	ATGTTGTGGCGGATCTTGAAGT	
r-t_ACT_Fw	CGAGCGAATGCACAAGGA	
r-t_ACT_Rev	GAGCGGTGATCTTGACCTTGA	
r-t_Sec62_Fw	CTACCTGCGGAGCCACAAG	This study
r-t_Sec62_Rev	TCGGATCGCTTACCGTTGAG	

Table S.2



Secretory helpers for enhanced production of heterologous proteins in *Yarrowia lipolytica*

Paulina Korpys-Woźniak, Piotr Kubiak, Ewelina Celińska*

Department of Biotechnology and Food Microbiology, Poznan University of Life Sciences, ul. Wojska Polskiego 48, 60-637 Poznań, Poland

ARTICLE INFO

Keywords:

Yeast
Temperature
Transcription
Secretion
Protein trafficking

ABSTRACT

Depending on the suboptimal factor, the target protein secretion can be over 1000-fold below the theoretical maximum. The bottlenecks may be alleviated by co-overexpression of “secretory helpers” (*SHs*). Here we proposed twelve *SHs*, functionally spanning the whole transcription-translation-translocation-folding-maturation-excretion pipeline. The genes were co-transformed with an easy-to-track reporter, and tested less than two temperatures. Our results indicated a clear distinction in the effects triggered by *SHs* involved in either synthesis or trafficking of the heterologous polypeptides. For superior operation of synthesis-related *SHs*, namely *RPL3*, *SSA5* and *SSA8*, the secretory pathway’s capacity must be released by applying decreased temperature (25 °C). The other *SHs* considered (e.g. *SSO1*, *CWP11*) did not give such spectacular results in the amounts of the target heterologous polypeptide, but allowed to maintain secretory capacity under unfavorable thermal conditions. This study provides generalizable guidelines for cloning/culturing strategies aiming at enhancement of heterologous protein secretion in *Y. lipolytica*.

Abbreviations

(sp)_{ex}YFP – (specific) extracellular fluorescence of YFP expressed in (FU OD600nm⁻¹) FU
(sp)_{in}YFP – (specific) intracellular fluorescence of YFP expressed in (FU OD600nm⁻¹) FU
SCYFP/SH(s) – genes encoding the target proteins: secretory YFP/secretory helper (s) (capitalized and italicized)
scYFP/SH (s) – proteins: secretory YFP/secretory helper (s) (non-capitalized and non-italicized)
FL – Fluorescence

1. Introduction

While yeast cells offer a multitude of benefits when used as heterologous protein production platforms (*i.a.* ease of genetic engineering and cultivation, eukaryotic post-translational modifications and secretion, large quantities of the product in grams per liter, if optimized), they are also subjected to numerous limitations. It has been estimated, that, depending on the suboptimal factor, the target protein secretion titers are 100- to 1000-fold lower than the theoretical maximum [1]. To account for this, plenty of inventive engineering strategies have been

adopted to reach the desired product gain, including process and genetics optimization [2–9]. One of the possible approaches is to co-overexpress a gene involved in the process of protein synthesis, folding, maturation or trafficking, as it has been shown that the stresses imposed by heterologous protein synthesis and secretion may be diminished by overexpression of genes encoding products involved in the secretory pathway. The so-called secretory helpers (*SHs*), secretion enhancers or secretion helper factors have been adopted for assisting the process of heterologous proteins synthesis in several yeast species, including *Komagataella phaffii* (traditional name *Pichia pastoris* will be used hereafter) [10,11], *Saccharomyces cerevisiae* [12–16], *Kluyveromyces lactis* [17], or *Hansenula polymorpha* [18] (for comprehensive reviews on the subject see [2,3,8,15]). The biggest challenge in such approaches is to first accurately identify the bottleneck from amongst a multitude of possibilities found across the translational-secretory machinery, and then—to identify an operable and efficient *SH* that would alleviate the limitation.

When it comes to identification of the rate-limiting step within the transcription-translation-translocation-folding-maturation and secretion pipeline, what is already known for sure, is that it depends on both—the host cell and the biochemical characteristics of the overproduced polypeptide [12,19–22]. All these studies, clearly

* Corresponding author.

E-mail address: ewelina.celinska@up.poznan.pl (E. Celińska).

<https://doi.org/10.1016/j.btre.2021.e00669>

Received 23 July 2021; Received in revised form 19 August 2021; Accepted 3 September 2021

Available online 11 September 2021

2215-017X/© 2021 The Author(s).

Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

demonstrated that the host cell reacts differently to overproduction of biochemically-different polypeptides, and that different *SHs* are thus efficient in enhancing the target protein synthesis. Nevertheless, several *SHs* with high potential of generalizability have been discovered and described for *P. pastoris* and *S. cerevisiae*.

As previously demonstrated, transcriptomics, functional genomics and metabolic models are valuable and reliable tools for fishing out and selection of genetic engineering targets for enhancing synthesis of heterologous secretory proteins. Microarrays-based transcriptomics enabled identification of completely new, efficient *SHs* operating as chaperones, ATPase, or exocytosis-associated kinase [10]. Whole genome metabolic model with incorporated synthesis of heterologous protein was proved to be functional for prediction of targets (*i.a.* dehydrogenases, decarboxylases, transferases) for genetic engineering enhancing synthesis of heterologous proteins [11]. High-throughput genomic library screens allowed to identify unexpected secretion enhancers within the genes involved in cell wall biogenesis, but also expectedly-ribosomal elements and foldases [12]. Microfluidic screening combined with high-throughput whole-genome sequencing resulted in many insightful observations and allowed to identify new engineering targets within genes involved in genome maintenance, trafficking of polypeptides, respiration, stress response *etc.* [23]. In our previous study we conducted comparative profiling of global transcriptomes in chemostat-maintained *Yarrowia lipolytica* cells overexpressing several different heterologous proteins [24]. The following careful analysis of the expression profiles, revealed genes that were differentially regulated upon overexpression of specific, or any heterologous secretory protein. These genes were considered as natural targets for further modifications with the aim to improve secretory capacity of this yeast species, which has not been attempted, to date.

Therefore, in the present study we used literature premises and our own experimental indications for selection of twelve potential *SHs*, and co-cloned them in *Y. lipolytica* with an easy-to-track reporter. The strains were tested in batch cultivations for their secretory capacity, synthesized proteins retention, and the target gene expression. Cultivations were conducted under two temperatures, based on the previous literature data. The research yielded some interesting conclusions in the interplay between function of the co-overexpressed *SH* and the temperature. The most beneficial cloning-cultivation engineering strategy for enhancing the heterologous secretory protein production in *Y. lipolytica* was indicated.

2. Materials and methods

2.1. Strains, growth and culture conditions

All strains used in this study are listed in Tab.S1. *Y. lipolytica* and *E. coli* strains were maintained as described in [25,26]. *E. coli* JM109 strain, used for sub-cloning, and its derivatives were grown at 37 °C with 250 rpm shaking in LB medium ((g L⁻¹): yeast extract (BTL, Lodz, Poland), 5; bacto-peptone (BTL), 10; NaCl (POCh, Gliwice, Poland), 5), supplemented with kanamycin (Sigma Aldrich, Merck KGaA, St. Louis, USA, 40 (µg mL⁻¹)) and agar (Biomaxima, Lublin, Poland; 15 (g L⁻¹)), when required. *Y. lipolytica* Po1f strain (*ura- leu-*, ATCC MYA2613) was used as a parental strain for co-transformations. *Y. lipolytica* Po1h strain (*ura-*, *leu+*) transformed with a URA3 marker cassette was used as a negative control strain for assessment of background fluorescence. Yeast strains were routinely maintained in a minimal yeast nitrogen base medium (YNB; (g L⁻¹): YNB (Sigma-Aldrich, Merck KGaA), 1.7; (NH₄)₂SO₄ (PoCh), 5; glucose (PoCh), 20), or in a rich yeast extract-peptone-dextrose medium (YPD; (g L⁻¹): yeast extract, 10; bacto-peptone, 20; glucose, 20), solidified with agar (15 (g L⁻¹)) when required, at 30 °C and with 250 rpm shaking, for liquid cultures.

2.2. Molecular biology techniques and reagents

Transformations of *E. coli* (heat-shock) and *Y. lipolytica* (lithium acetate heat-shock) strains were performed as described in [25,26]. Restriction digestion of DNA fragments was done using *Bam*HI, *Avr*II, *Not*I enzymes (New England Biolabs, Ipswich, USA; Thermo Fisher Scientific, Waltham, USA), as indicated. Ligation of DNA fragments into pCR Blunt II TOPO vector (Thermo Scientific, Waltham, USA) or JMP62 vectors was conducted using 200 U T4 DNA ligase (NEB) and T4 DNA ligase buffer (NEB). PCR was performed using Phire DNA polymerase (Thermo Fisher Scientific, Waltham, MA USA) or RUN DNA polymerase (A&A Biotechnology, Gdynia, Poland). DNA plasmids isolation, DNA fragment extraction from agarose gel, and purification of DNA fragments were all conducted using appropriate kits from A&A Biotechnology (Gdynia, Poland). All the reactions and protocols were used in accordance with the manufacturer's recommendations.

2.3. Cloning strategy

A DNA sequence encoding YFP (26 kDa) was transcriptionally fused with a signal sequence SP1, native for exo-1,3-beta-glucanase (*YALI0B03564g*) to assure its efficient secretion [27]. The *SH* genes were amplified directly on *Y. lipolytica* genomic DNA using Phire DNA polymerase with specific oligonucleotides listed in Tab.S2. *CNE1* gene DNA sequence was purchased from SYNGEN (Wroclaw, Poland) as a completely synthetic DNA element, due to the presence of numerous *Bam*HI recognition sites, that were later required as unique for cloning. The oligonucleotides contained restriction sites *Bam*HI and *Avr*II, which enabled cloning in JMP62 plasmids [28]. JMP62-URA3ex was used as a vector for the model reporter protein (scYFP), and JMP62-LEU2ex was used as a vector for the *SH* genes Fig.S1. All the genes were cloned under the control of a constitutive pTEF promoter. The presence of expression cassettes in recombinant *E. coli* strains was verified by colony PCR. The expression cassettes were sequenced (Genomed, Warsaw, Poland) to verify correctness of the DNA constructions. Positive bacterial strains were deposited as glycerol stocks at -80 °C.

Prior to the transformation into *Y. lipolytica* Po1f strains, the JMP62-based constructions were digested with *Not*I endonuclease to remove bacterial elements. The JMP62-LEU2ex empty expression vector and the JMP62-URA3ex expression vector containing the *SCYFP* were co-transformed into *Y. lipolytica* Po1f strain, that was used as a reference for fluorescence evaluation and external calibrator for relative quantification of the expression level. Recombinant fluorescent strains were obtained by co-transformation with the vector JMP62-URA3ex-*SCYFP* and JMP62-LEU2ex-*SHs* into *Y. lipolytica* Po1f parental strain (Fig.S1). All transformants were selected on YNB medium at 30 °C for 48 h, and then, replica-plated on fresh YNB and YPD agar plates. The fluorescence phenotype (scYFP) was verified via fluorescence microscopy observations (ZEISS AxioVert, AxioCam 350 color; filterset: 09) and quantitatively assessed in microcultures using an automatic plate reader/fluorimeter (Tecan Infinite M200; Tecan Group Ltd., Männedorf, Switzerland). The presence of genes encoding *SHs* in the recombinant strains was verified by colony PCR using specific primers. Positive yeast strains were deposited as glycerol stocks at -80 °C.

2.4. Screening for representative sub-clones

Four sub-clones of each variant were selected from amongst all *Y. lipolytica* positive transformants and screened to select a representative strain for further studies. The screening was conducted in 50 mL shake flasks containing 5 mL YNB medium at 30 °C and 220 rpm shaking for 72 h. Samples were periodically withdrawn from the cultures and centrifuged. The biomass was washed twice in a sterile saline solution (0.85% NaCl), resuspended in equal volume of the same solution and appropriately diluted, to assure readout within the method linearity range. The supernatant and the washed biomass were subjected to

fluorimetric measurements. The biomass was also analyzed via spectrophotometric measurements for biomass growth determination. All the cultures were performed in biological duplicate.

2.5. Batch cultivations

Four milliliters of YPD medium were inoculated with a single colony of the representative strain (YPD agar plates) and grown at 30 °C with 250 rpm shaking. After 22 h, three milliliters (10% of working volume) of these pre-cultures were transferred into 30 mL of the YPD medium in 250 mL shake flasks and incubated at 30 °C for 24 h with 220 rpm shaking. For the cultures conducted at 25 °C, the temperature downshift was executed at 24 h time-point and continued until the end of culturing. All the cultures were conducted for 96 h in batch culture mode. Samples (1 mL) were periodically collected, centrifuged and analyzed as described in 2.4 and 2.6. All the cultures were conducted in at least two biological replicates.

2.6. Analytical methods

2.6.1. Optical density determination (OD600)

The growth of *Y. lipolytica* strains was monitored by measuring optical density at 600 nm (OD 600 nm). The yeast biomass was suspended and diluted in a sterile saline solution (0.85% NaCl) and measured in flat-bottomed MTP plates (Corning; Sigma-Aldrich) in Tecan Infinite M200 automatic plate reader in 200 µL. All the measurements were conducted in technical duplicate out of each biological replicate.

2.6.2. Fluorimetry (in / exYFP)

The fluorescence (FL) measurements of biomass and supernatant samples were performed in flat-bottomed MTP plates (Corning; Sigma-Aldrich) in Tecan Infinite M200 automatic plate reader at the wavelengths (excitation/emission) 495/527 nm. The extracellular FL was measured in 200 µL in supernatant samples. The intracellular FL was measured in 200 µL pre-washed culture pellet samples, resuspended and diluted in the saline solution (0.85% NaCl). Each fluorimetric measurement was normalized vs background fluorescence, either biomass of the Po1h_Ura3 negative control strain or fresh YNB/YPD media for inYFP and exYFP readouts, respectively. FL results were expressed as: i) inYFP-raw FL readout of the washed biomass in [FU], ii) sp_inYFP-specific fluorescence of the washed biomass in [FU OD600nm⁻¹], iii) exYFP-raw FL readout of the supernatant in [FU], iv) sp_exYFP-specific fluorescence of the supernatant in [FU OD600 nm⁻¹]. All the measurements were conducted in technical duplicate out of each biological duplicate. Additionally, the fluorimetry results were randomly verified through observations under fluorescence microscope.

2.6.3. Gene expression analysis (RTqPCR)

Determination of the *SH*- and *YFP*-encoding genes expression level in the recombinant *Y. lipolytica* strains was conducted by RTqPCR. The samples were collected after 24 h of growth in batch cultures under 30 °C or 25 °C. Biomass from 1 mL of the culture was used for isolation of total RNA using Bead-Beat Total RNA Mini Kit (A&A Biotechnology). After qualitative (agarose gel electrophoresis) and quantitative (spectrophotometry) verification of total RNA preparations, the material was reverse transcribed to cDNA using SuperScript III Reverse Transcriptase and oligo(dT) primer, according to the manufacturer's instructions (Thermo Fisher Scientific). cDNA preparations were used as templates in RTqPCR, carried out in an Applied Biosystems 7500 device (Applied Biosystems, Foster City, USA). The reactions were set up using RT HSPCR Mix SYBR® B (A&A Biotechnology) according to the manufacturer's specifications. LoROX dye was used as a passive reference. Actin-encoding gene (*ACT1*) was used as the internal calibrator for normalization of the expression level. The gene-specific primer pairs are listed in Tab.S2. The following thermal profile was used: 95 °C 3 min, (95 °C 15 s, 60 °C 30 s, 72 °C 30 s) × 40, 72 °C 1 min, Melt Curve 94 °C 15 s, 60 °C 60

s, 95 °C 30 s, 60 °C 15 s. The target genes' expression was quantified using the 2^{-ΔΔCt} method [29]. cDNA preparations of the control strain (expressing *scYFP*, without *SH*'s overexpression) was used as the external calibrator, to which 1.0 expression level was assigned. All the samples were analyzed in technical duplicates.

2.6.4. Data analysis

For convenience of comparison, some data were additionally expressed as fold change (FC) values calculated by making a fraction of two specific values expressed in the same units. Such FC values prepared using sp_exYFP were fed into Morpheus (<https://software.broad-institute.org/morpheus/>) online tool to draw illustrative heatmaps. Kinetic graphics representing growth and evolution of (sp)in/exYFP in batch cultures were prepared using Microsoft Excel. Likewise, bar charts illustrating gene expression data were prepared using Excel graphic tools. Statistical analysis was performed with R (version 4.1.0 <https://www.R-project.org/>). Tukey HSD (p-value of 0.05) was performed using the agricolae package (<https://cran.r-project.org/package=agricolae>). The rstatix package was used to perform t-tests (<https://CRAN.R-project.org/package=rstatix>). Boxplots were prepared using ggplot2 (<https://ggplot2.tidyverse.org>) and ggpubr (<https://CRAN.R-project.org/package=ggpubr>) packages.

3. Results

3.1. Cloning of homologous secretory helpers in *Y. lipolytica*

A list of *SH*s studied here is given in Table 1. The selection was based on our previous indications inferred from comparative global transcriptome profiling in *Y. lipolytica* [24] and literature data demonstrating benefits of overexpression of a given *SH* in *S. cerevisiae*, *K. phaffii*, *H. polymorpha*, *K. lactis* or *Aspergillus oryzae*, e.g. [2,10,12]. Cloning strategy for co-transformation of the reporter protein (*scYFP*) and the *SH*s is shown in Fig.S1. For each *SCYFP-SH* combination, four *Y. lipolytica* sub-clones were initially screened in micro-cultures for biomass growth, as well as intra- and extra-cellular fluorescence, to select a single representative strain. As can be observed in Fig.S2.x.1 (where x is a letter assigned to cloned *SH*), the inter-clonal variation was reasonably low and overexpression of *SH*s did not impair growth of the transformants, when compared to the control strain ($p < 0.05$). On the other hand, depending on the co-transformed *SH*, the strains exhibited variability in sp_inYFP/exYFP parameters Fig.S2.x.1 and S2.x.2. Based on that initial screening, we selected one sub-clone from each co-transformation type, that demonstrated the most similar growth curve to the reference strain, and was the most representative for the group of recombinants in terms of sp_inYFP/exYFP parameters (avoiding outliers). Enhanced expression of the *SH*-encoding genes in these selected double-transformants over the *YFP*-expressing control was confirmed by RTqPCR (not shown). Twelve representative strains were subjected to further, more detailed analyses.

3.2. Batch cultivations of *Y. lipolytica* strains with elevated dosage of *SH*s

3.2.1. Kinetics of growth, heterologous protein synthesis and secretion

The twelve representative *Y. lipolytica* strains, overexpressing both the secretory reporter (*scYFP*) and one of the *SH*s, were subjected to a series of batch cultivations (scaled up vs the preliminary screens) in parallel with the control strain (expressing solely *SCYFP*). The cultures were continued over 96 h under two thermal conditions—25 °C and 30 °C. Kinetics of biomass growth, as well as inYFP and exYFP evolution is presented in Fig. 1.x.3, 0.1.x.1 and 0.1.x.2, respectively. As inferred from the preliminary micro-culture screens, overexpression of *SH*s did not significantly affect the growth of either the recombinant and control strain ($p < 0.05$). In contrast, the decrease in the temperature by 5 °C (to 25 °C) significantly impacted biomass accumulation ($p < 0.05$), in favor of the lower (Fig. 1.x.3). The course of growth curves clearly indicates

Table 1
Secretory helpers analyzed in this study with short description of their function.

<i>Y. lipolytica</i>	Component	Cellular localization	Description
<i>YALIOB12716g</i>	<i>HAC1</i>	Nucleus	Transcription factor that regulates the unfolded protein response, via UPRE binding, and membrane biogenesis. Hac1p mediates activation of hundreds of molecular events, including increased provision of chaperones and membranes, to concertedly relieve burdened secretory pathway
<i>YALIOC21560g</i>	<i>RPL3</i>	Ribosome	Protein component of the large (60S) ribosomal subunit
<i>YALIOF25289g</i>	<i>SSA5</i>	Cytosolic	Cyttoplasmic members of the HSP70 family; play a key role in folding, targeting and post-translational translocation. The protection of the released polypeptide in extended conformation (translocation competent state) is secured by the action of cytosolic chaperones.
<i>YALIOD22352g</i>	<i>SSA8</i>		
<i>YALIOE03036g</i>	<i>PDI</i>	ER lumen	Protein disulfide isomerase essential for disulfide bond formation, which relies on stochastic oxidation-reduction of cysteine side chains; consumes considerable amounts of oxidating and reducing agents (O ₂ and GSH, respectively) during the folding of secretory proteins
<i>YALIOE32703g</i>	<i>SLS1</i>	ER lumen	Sls1, a nucleotide exchange factor for Kar2/BiP, has important functions in regulating ER stress and the interaction of Kar2/BiP and Ire1. Sls1 regulates this interaction, by stimulating the conversion of BiP from the ADP-bound to the ATP-bound state, which favors its interaction with Ire1. Sls1p acts in the protein translocation process, interacting directly with translocating polypeptides to facilitate their transfer and/or help their folding in the ER.
<i>YALIOB13156g</i>	<i>CNE1</i>	ER membrane	Calnexin involved in glycosylation and monitoring the folding state of the nascent glycosylated polypeptides. Key component of the quality control mechanism in the ER
<i>YALIOE26026g</i>	<i>YET3</i>	ER membrane / ER-Golgi transport	Yet3 involved in transport between the ER membrane and Golgi (homologous to the mammalian Bap31); its traffic is dependent on vesicular transport and is associated with new synthesized membrane polypeptide
<i>YALIOD23947g</i>	<i>USO1</i>	ER-Golgi transport	Essential protein involved in vesicle-mediated ER to Golgi transport; binds membranes and functions during vesicle docking to the Golgi; required for assembly of the ER-to-Golgi SNARE complex
<i>YALIOE22044g</i>	<i>SEC1</i>	Golgi-Plasma membrane	Sm-like protein involved in docking and fusion of exocytic vesicles; binds to assembled

Table 1 (continued)

<i>Y. lipolytica</i>	Component	Cellular localization	Description
<i>YALIOE23243g</i>	<i>SSO1</i>	Plasma membrane	SNARE complexes at the membrane and stimulates membrane fusion. SEC1 is required for transport of polypeptides from Golgi to plasma membrane
<i>YALIOE22286g</i>	<i>CWP11</i>	Cell wall	Plasma membrane t-SNARE functioning at the targeting/fusion of the Golgi-derived secretory vesicles to the plasma membrane GPI-anchored cell wall protein. Essential protein involved in the cell wall remodeling during production of the secretory protein

that the differentiation in growth rate was initiated once the temperature downshift was applied. This general observation holds valid for all the twelve strains, irrespective of the *SH* cloned. Only at some specific time-points and for specific gene combinations, minor temporary differences in the biomass accumulation between the *SH* recombinants and the control strain could be observed (e.g. strain *SCYFP-SSO1* at 48 h, 25 °C; or *SCYFP-YET3* at 48 h, 30 °C; $p < 0.05$).

In contrast, overexpression of different *SHs* triggered significant and highly variable changes in inYFP and exYFP evolution over the culturing time (Fig. 1.x.1 and Fig. 1.x.2). Primarily, a clear impact of increased dosage of *RPL3*, *SSA5* and *SSA8* on enhanced synthesis and retention of YFP was observed, irrespectively of the adopted temperature (Fig. 1.B/C/D.1). For all these strains, the inYFP accumulation level reached the highest values from amongst all the strains (compare Y axis ranges in Fig. 1.x.1). Interestingly, in the late stationary phase of growth, particularly high inYFP accumulation was observed when the *RPL3*-, *SSA5*- and *SSA8*-overexpressing strains were cultured under 25 °C. Temperature-dependence was also observed in the case of *SSO1*-overexpressing strain, where significantly increased inYFP accumulation was recorded under 30 °C, while no difference in inYFP was found between the *SH*-overproducing strain and control cultured under 25 °C (Fig. 1.K.1). For the remaining cases studied here, no clear difference in the inYFP parameter were observed between the modified and the control strains, except for *YET3*-overexpressing strain, for which the *SH*'s dose increase caused a decrease in intracellular YFP under both temperatures applied (Fig. 1.H.1).

When it comes to exYFP absolute values (Fig. 1.x.2; Table 2), we observed three spectacular improvements caused by overexpression of *RPL3*, *SSA5*, *SSA8* in *Y. lipolytica* strains cultured at 25 °C (Fig. 1.B/C/D.2). The effect of these genetic modifications was not that remarkable, when these strains were grown at 30 °C. Some minor improvements were also observed due to increased dosage of *SSO1* (Fig. 1.K.2), and to a lesser extent of *HAC1*, *PDI1*, *SLS1*, *CNE1*, *USO1*, *SEC1*, *CWP11* (Table 2). In majority of those cases, combined action of the lowered temperature and the genetic modification was required for the enhancement in exYFP level ($p < 0.05$). *SEC1*, *YET3*, *SLS1*, *PDI1*, and *USO1* overexpression did not exert positive effect on exYFP levels when the strains were cultured under 30 °C. In fact, overexpression of *USO1* combined with cultivation at 30 °C contributed to lower exYFP levels ($p < 0.05$); such a tendency was also observed for *PDI1* and *SLS1*.

Careful analysis of the inYFP and exYFP evolution, as well as biomass growth curves (Fig. 1) showed that all the processes reached their plateau at ~ 48 h (biomass growth) to ~ 72 h (in/exYFP). At that stage, the effects of the temperature and the genetic modification could be observed macroscopically, by reading absorbance or fluorimetry. In addition, to evaluate the actual impact of the two given factors (*SH* overexpression and temperature) on the cell's secretory capacity and not

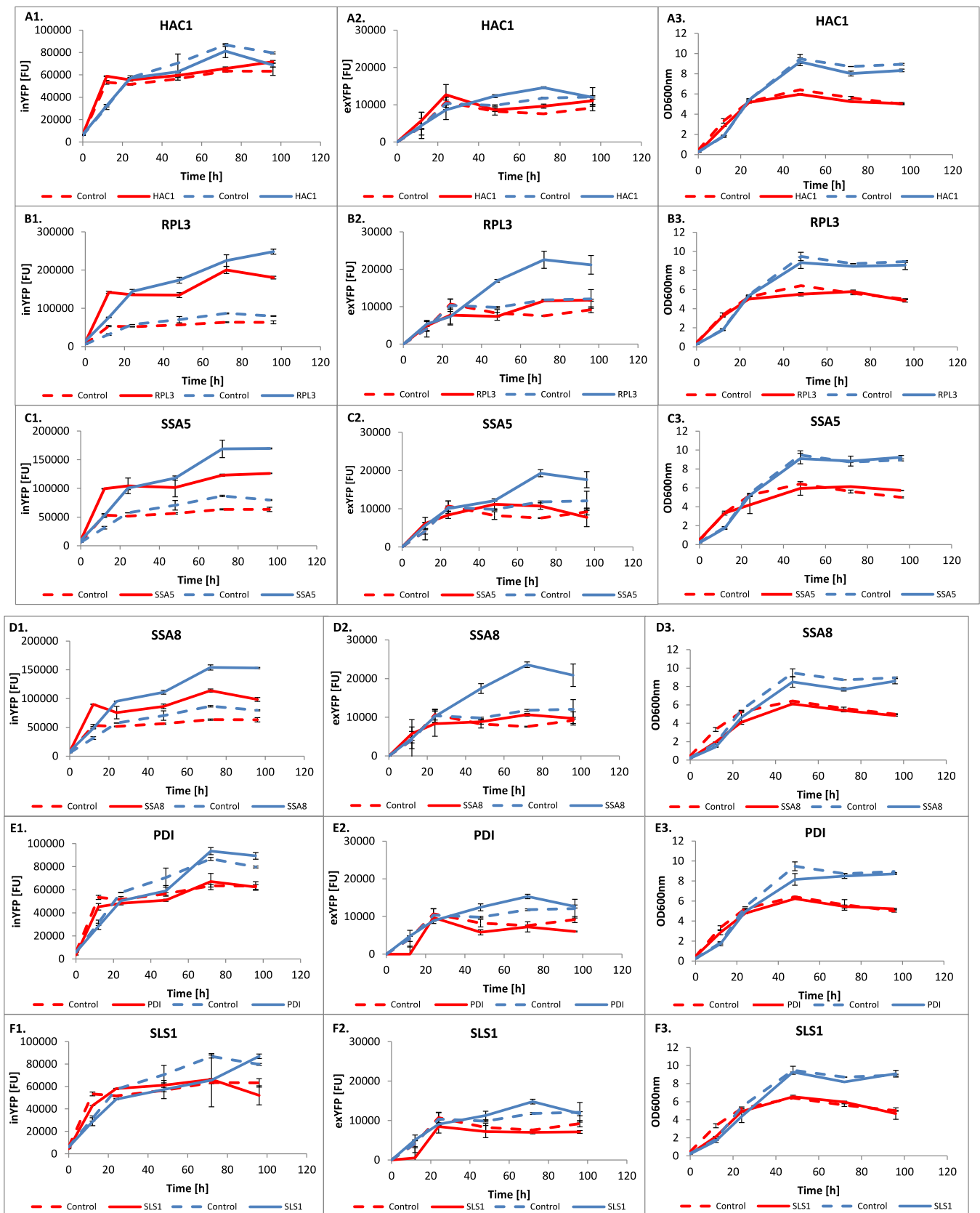


Fig. 1. Kinetics of inYFP (0.1 [FU]) exYFP (0.2 [FU]) and biomass growth (0.3 [OD600 nm]) in batch cultures of *Y. lipolytica* co-transformants (solid lines) and the reference strain (dashed lines) under 25 °C (blue lines) and 30 °C (red lines). X axis: culturing time [h]; Y axis: inYFP (0.1 [FU]), exYFP (0.2 [FU]) and biomass growth (0.3 [OD600 nm]). In the case of sub-figures 0.1, please mind Y axis range. Error bars indicate mean values \pm SD from biological duplicate, each measured in technical duplicate.

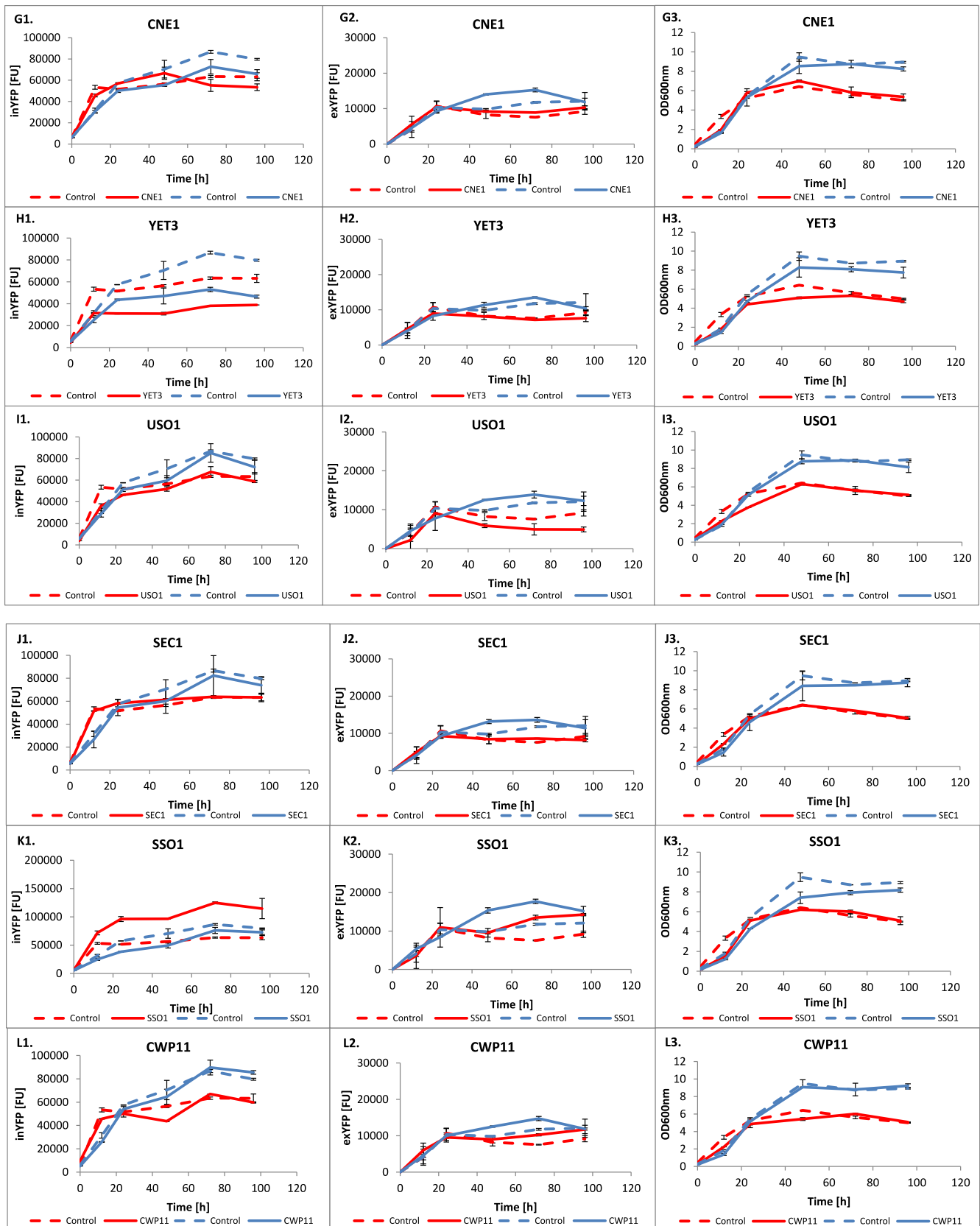


Fig. 1. (continued).

its growth, in our further analyses we use specific fluorescence, understood as specific secretion (sp_exYFP in $[FU OD600\text{ nm}^{-1}]$) parameter as the one enabling most accurate description of the studied biological

process. Since the kinetics of growth was corresponding in all of the processes, we could compare different cultures using sp_exYFP $[FU OD600\text{ nm}^{-1}]$ at a specified time point from the indicated plateau range.

Table 2

Numerical data for growth (OD 600 nm), extracellular fluorescence (exYFP [FU]) and extracellular fluorescence normalized per biomass accumulation (sp_exYFP [FU OD600nm⁻¹]) of *Y. lipolytica* co-transformants at stationary phase of growth (72 h) in batch cultures under 25 °C or 30 °C. Fold change values were calculated in reference to a control strain overexpressing solely *YFP*. Numbers are given in the indicated units ± SD from biological duplicate, each measured in technical duplicate.

Helper gene	25 °C					72h					30 °C				
	OD 600nm	exYFP [FU]	Fold change [exYFP]	sp_exYFP [FU OD600 nm ⁻¹]	Fold change [sp_exYFP]	OD 600nm	exYFP [FU]	Fold change [exYFP]	sp_exYFP [FU OD600 nm ⁻¹]	Fold change [sp_exYFP]	OD 600nm	exYFP [FU]	Fold change [exYFP]	sp_exYFP [FU OD600 nm ⁻¹]	Fold change [sp_exYFP]
Control_scYFP	8.72 ± 0.00	11,780.50 ± 252.44	1.00	1352.53 ± 26.06	1.00	5.62 ± 0.15	7572.50 ± 103.94	1.00	1348.60 ± 16.81	1.00	5.62 ± 0.15	7572.50 ± 103.94	1.00	1348.60 ± 16.81	1.00
scYFP_HAC1	8.02 ± 0.24	14,569.75 ± 264.81	1.24	1818.34 ± 20.21	1.34	5.24 ± 0.02	9610.50 ± 552.25	1.27	1834.90 ± 97.52	1.36	5.24 ± 0.02	9610.50 ± 552.25	1.27	1834.90 ± 97.52	1.36
scYFP_RPL3	8.44 ± 0.04	22,542.50 ± 2281.83	1.91	2673.97 ± 284.69	1.98	5.79 ± 0.16	11,557.00 ± 141.42	1.43	1995.94 ± 80.20	1.48	5.79 ± 0.16	11,557.00 ± 141.42	1.43	1995.94 ± 80.20	1.48
scYFP_SSA5	8.83 ± 0.52	19,294.25 ± 930.20	1.64	2193.26 ± 237.35	1.62	6.13 ± 0.03	10,712.00 ± 869.74	1.42	1747.53 ± 132.04	1.30	6.13 ± 0.03	10,712.00 ± 869.74	1.42	1747.53 ± 132.04	1.30
scYFP_SSA8	7.69 ± 0.19	23,576.25 ± 720.19	2.00	3067.31 ± 20.51	2.27	5.38 ± 0.14	10,647.00 ± 379.01	1.41	1980.28 ± 123.52	1.47	5.38 ± 0.14	10,647.00 ± 379.01	1.41	1980.28 ± 123.52	1.47
scYFP_PDI	8.49 ± 0.24	15,305.25 ± 579.474	1.30	1804.33 ± 15.72	1.33	5.45 ± 0.53	7235.00 ± 1350.57	0.96	1322.81 ± 119.70	0.98	5.45 ± 0.53	7235.00 ± 1350.57	0.96	1322.81 ± 119.70	0.98
scYFP_SLS1	8.19 ± 0.04	14,784.50 ± 591.85	1.26	1804.26 ± 63.05	1.33	5.96 ± 0.10	7012.00 ± 397.39	0.93	1176.21 ± 46.17	0.87	5.96 ± 0.10	7012.00 ± 397.39	0.93	1176.21 ± 46.17	0.87
scYFP_CNE1	8.75 ± 0.38	15,253.50 ± 551.54	1.30	1745.04 ± 13.90	1.29	5.84 ± 0.55	8891.00 ± 12.73	1.17	1530.60 ± 147.40	1.14	5.84 ± 0.55	8891.00 ± 12.73	1.17	1530.60 ± 147.40	1.14
scYFP_YET3	8.09 ± 0.26	13,537.75 ± 172.18	1.15	1675.45 ± 32.31	1.24	5.31 ± 0.03	7154.50 ± 17.99	0.95	1346.05 ± 16.56	1.00	5.31 ± 0.03	7154.50 ± 17.99	0.95	1346.05 ± 16.56	1.00
scYFP_USO1	8.89 ± 0.12	13,901.00 ± 869.03	1.18	1564.42 ± 118.57	1.16	5.61 ± 0.43	4949.50 ± 1441.79	0.65	875.56 ± 189.38	0.65	5.61 ± 0.43	4949.50 ± 1441.79	0.65	875.56 ± 189.38	0.65
scYFP_SEC1	8.48 ± 0.01	13,626.75 ± 658.67	1.16	1607.63 ± 75.05	1.19	5.81 ± 0.03	8625.00 ± 90.51	1.14	1483.34 ± 8.89	1.10	5.81 ± 0.03	8625.00 ± 90.51	1.14	1483.34 ± 8.89	1.10
scYFP_SSO1	7.92 ± 0.21	17,724.50 ± 560.74	1.51	2243.24 ± 128.82	1.66	6.00 ± 0.11	13,511.00 ± 593.97	1.78	2317.25 ± 239.95	1.67	6.00 ± 0.11	13,511.00 ± 593.97	1.78	2317.25 ± 239.95	1.67
scYFP_CWP11	8.80 ± 0.71	14,741.75 ± 590.79	1.25	1681.11 ± 69.87	1.24	6.02 ± 0.06	10,269.50 ± 301.94	1.36	1705.45 ± 34.12	1.27	6.02 ± 0.06	10,269.50 ± 301.94	1.36	1705.45 ± 34.12	1.27

3.2.2. Effect of the SHs elevated dosage on secretory capacity

Fig. 2 illustrates the results of Tukey HSD test classifying the double-transformants in terms of their normalized secretory capacity (sp_exYFP; [FU OD600 nm⁻¹]) under two thermal conditions at 72 h of the culturing. Irrespective of the culturing temperature, increased dose of *SSA8*, *SSA5*, *RPL3*, *SSO1* and *HAC1* turned the strains into most efficient secretors of the model protein (Fig. 2.A.B.; Fig. 3.A.B.; numerical data can be found in Table 2). Yet, the absolute numbers for sp_exYFP were in general higher for the strains grown under 25 °C (scYFP_SSA8 1.5-fold, scYFP_RPL3 1.34-fold, scYFP_SSA5 1.25-fold), with some minor exceptions. Under 30 °C, *SEC1*, *CNE1* and *CWP11* overexpression positively impacted sp_exYFP levels, but the level of improvement (10–27%) was not statistically significant ($p > 0.05$). *PDI1* and *YET3* overexpression had no observable effect on the normalized secretory capacity of the strain when cultured at 30 °C. On the other hand, *USO1* overexpression led to nearly 40% decrease in both exYFP and sp_exYFP measures when the strain was grown under 30 °C. It was also the only case, where the genetic modification triggered inferior outcome compared to the control strain ($p < 0.05$) and to majority of the strains bearing the other *SH*-overexpression (Fig. 3.A).

Culturing the modified strains under decreased temperature (25 °C) caused significant changes in the ranking of *SHs* (Fig. 2.B). In this case, any *SH* overexpression had some positive effect on the normalized secretory capacity of *Y. lipolytica* ($p < 0.05$; a tendency for *USO1* and *SEC1*; see also Fig. 3.B). Applying the lower temperature triggered a kind of polarization of the results, observed as a more clear distinction of the homogeneous groups, clearly defining the most beneficial genetic modifications (Fig. 2.B). Under these conditions, the unambiguous prevalence of the *SSA8*-overexpression strain over the control (~2.3-fold improvement in sp_exYFP; Table 2), and all the remaining strains, could be seen (Fig. 3.B). The leader strain was directly followed by the “*RPL3* overexpressor” (~100% improvement vs the control; Table 2), making a homogeneous group of its own. The *SSO1* and *SSA5*-overexpressing strains were categorized together (Fig. 2.B), demonstrating

significant superiority over the control strain (~60% improvement; Table 2) and the remaining eight *SH*-overexpressing strains ($p < 0.05$; Fig. 2.B and Fig. 3.B). Further specific details on the fold change in the sp_exYFP measures by all the strains at two time points (48 h and 72 h) and the two temperatures (25 °C and 30 °C) are given in Fig.S3.

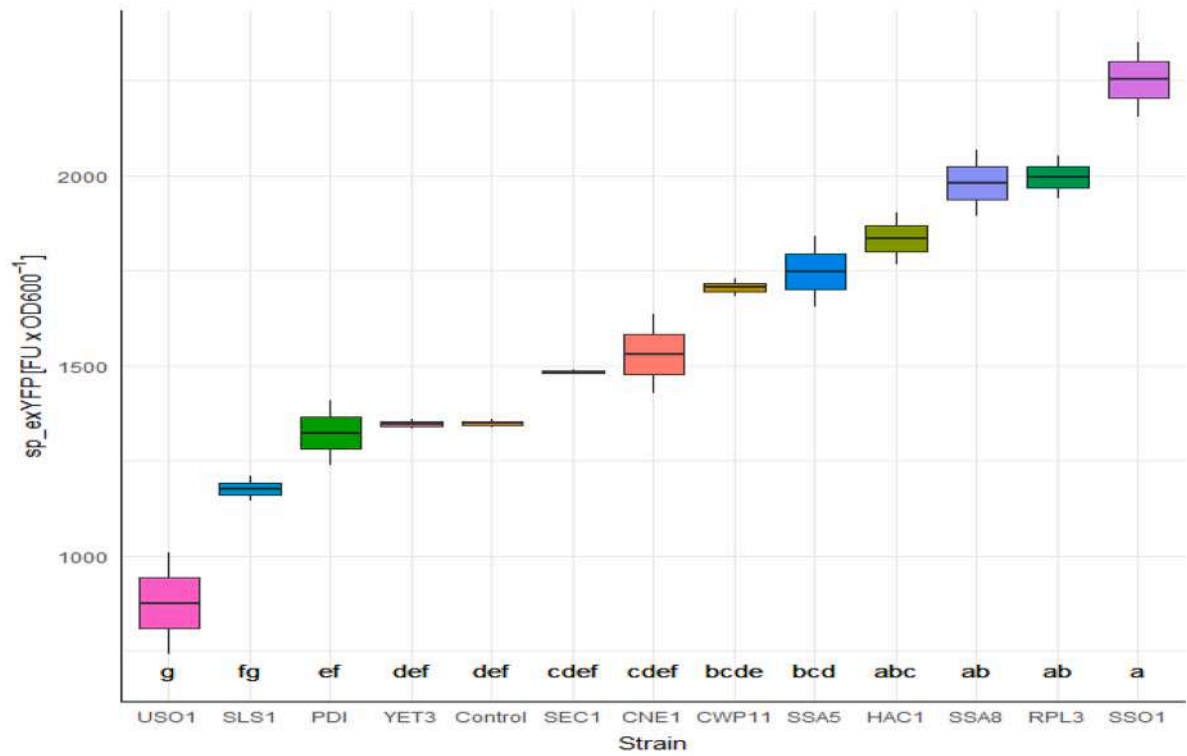
3.2.3. Effect of the temperature on SHs action in *Y. lipolytica*

The results presented in Fig. 1.B.2, C.1, C.2, D.1, D.2 K.1 indicate a clear modulating effect of the temperature on the YFP accumulation and/or secretion. This indication led us to the hypothesis that maybe the action of the specific *SHs* is somehow temperature-dependent. To test this, we again selected a specific time point (72 h) and sp_exYFP parameter, describing synthesis and secretion potential without growth-related phenomena, and run *t*-test statistical comparison. Results of this direct comparison are shown in Fig. 4. Indeed, for half of the *SH* studied here, the decreased temperature had a significant positive impact in terms of the resultant secretory potential ($p < 0.05$). Notably, it was actually the interaction between *SH* overexpression and the temperature that contributed to elevated sp_exYFP, as such effect could not be seen for the control strains ($p = 0.931$) and several other *SH*-overexpressing strains cultured under 25 °C. That beneficial interaction could be seen for all the cytoplasmic helpers–*RPL3*, *SSA5*, *SSA8*, but also ER-resident *SLS1*, and two *SHs* operating within vesicular transportation–*YET3* and *SEC1*. Some insignificant tendency of this beneficial interaction occurrence was also observed for *PDI1* and *USO1* ($p < 0.1$), which was significant at the earlier time point for *USO1* (Fig.S4).

3.2.4. Expression level of the heterologous gene upon SHs overexpression

To investigate whether the macroscopically observed changes in synthesis and secretion of YFP were underlain by the *YFP* gene expression level, we conducted comparative gene expression analysis for all the representative strains cultured under 25 °C and 30 °C (Fig. 5.A.B; Fig. 6). Biological sense of this experiment was to see whether the elevated dose of the *SHs* impacted transcription of the *YFP*, as previous

A. 30°C



B. 25°C

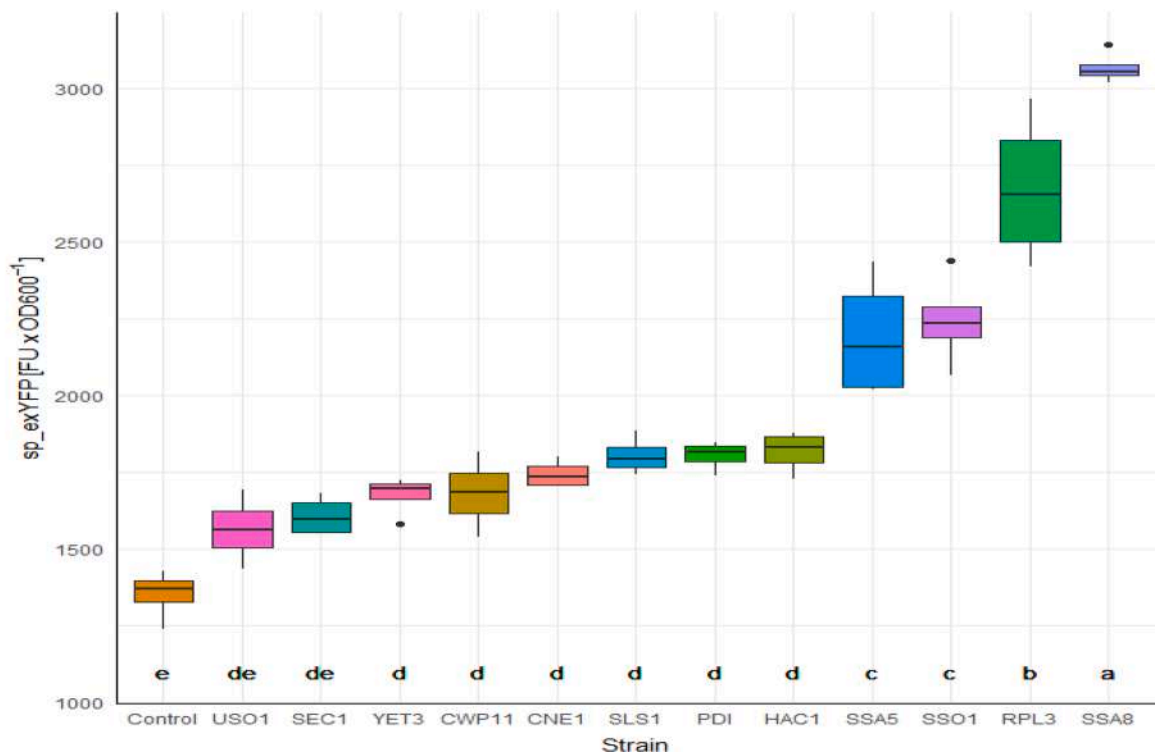


Fig. 2. Specific secretion (fluorescence) of the heterologous reporter protein (scYFP) by *Y. lipolytica* co-transformants and the control strain in stationary phase of growth (72 h) in batch cultures under 30 °C (A) and 25 °C (B). X axis: co-overexpressed *SH*; letters indicate homogenous groups calculated by Tukey HSD test. Y axis: specific secretion (sp_exYFP) of YFP in [FU OD600 nm⁻¹]. The boxes cover the 50% of input data points and a median indicated as solid line.

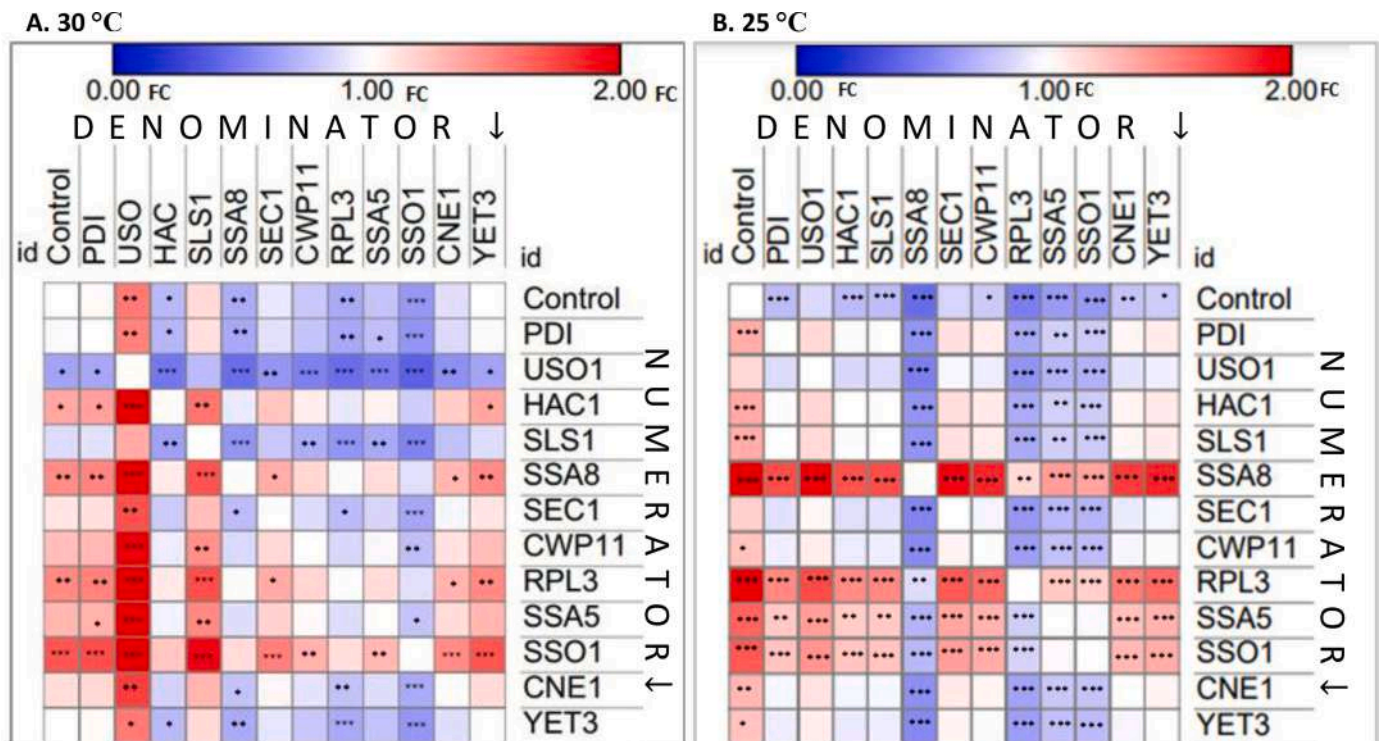


Fig. 3. Multiple comparisons of *Y. lipolytica* co-transformants and the control strain presented as fold change of specific secretion measures (sp_exYFP; $\frac{\text{NUMERATOR}_{FU\ OD600nm-1}}{\text{DENOMINATOR}_{FU\ OD600nm-1}}$) between the strains, cultured under 30 °C (A) and 25 °C (B). To use the heatmap correctly, first identify the strain on the left and use it as a NUMERATOR in fold change calculation, then, select a strain from the top of the heatmap, to be used as DENOMINATOR in the fold change calculation. Red squares represent prevalence of the numerator in the adopted measure (1.01–2.00 FC) and blue squares represent prevalence of the denominator in the adopted measure (0–0.99 FC). Statistical significance of the comparisons was determined by a Tukey HSD test. *** $P \leq 0.001$ ** $P \leq 0.01$ * $P \leq 0.05$.

research demonstrated that, in specific cases, transcription of heterologous genes can be reversibly modulated by the downstream processes like translation and folding events [24]. If no specific mechanism of reverse regulation occurs, the rate of transcription initiation events should be decreased under decreased temperature, according to the classic Arrhenius model. While in majority of cases the common model held valid, for several specific cases we observed a slight tendency for increased *YFP* expression level at elevated temperature (not significant at $p < 0.05$; Fig. 6). Primarily, the increased dosage of *RPL3* had a tremendous effect on the gene of interest overexpression ($p < 0.05$), which was, in addition, doubled when the strain was grown at 25 °C (significant over the control at both temperatures, $p < 0.05$; Fig. 5.A.B; but not significant between the temperatures, $p = 0.166$; Fig. 6). Likewise, doubling the *YFP* RQ value under 25 °C was observed for the strains overexpressing *YET3* and *SEC1* (significant over the control at $p < 0.05$; Fig. 5.B., but not significant between the temperatures, $p \sim 0.15$; Fig. 6), but those genes overexpression under 30 °C had no significant effect on RQ (not significant over the control at $p < 0.05$; Fig. 5.A). The other genes with the highest impact on inYFP/exYFP synthesis (as illustrated in Fig. 1, Fig. 2, Fig. 3, Fig. 4), *SSA5* and *SSA8*, contributed to significant upregulation of the *YFP* gene under both temperatures (Fig. 5.A.B). On the other hand, that was not the case for the other “efficient *SH*”, namely *SSO1*, as its overexpression had no impact on *YFP* gene transcription (Fig. 5.A.B). Intriguingly, overexpression of *HAC1*, *SLS1*, *PDI1* triggered a significant increase in the heterologous gene expression when the strains were grown under 30 °C (Fig. 5.B; a tendency observed for *CNE1*, but $p > 0.05$). In fact, all these genes exerted their impact on *YFP* gene expression level in a temperature-dependent manner (Fig. 6). Such a mode and direction of the temperature-dependent action was also observed for a one more gene, namely *USO1* ($p < 0.05$). An opposite trend was observed for the strains co-overexpressing *RPL3*, *SEC1*, *YET3*, for which coordinately higher

expression (not significant) and secretion (significant at $p < 0.05$) were seen under 25 °C.

To get a global insight into the relationship between: i) the temperature, ii) the heterologous gene expression level and iii) synthesis of scYFP polypeptide, we conducted Pearson correlation analysis between total amount of produced YFP (sum of FU for inYFP+exYFP) and RQ values for *YFP*-encoding gene. Strikingly, we noted, that such correlation was also temperature-dependent, as the *r* factor for the values read in 30 °C-maintained strains was 0.397, while for those cultured under 25 °C – $r = 0.855$, showing high linearity between transcription and translation under the lowered temperature.

Finally, to get an insight into the relationship between *Hac1p* transcription factor and its known downstream “interaction partners”, we analyzed gene expression level of *HAC1*, *PDI1*, *KAR2*, as well as scYFP, in the *HAC1*-overexpressing strain (Fig. 7). It was expected that upon elevated provision of *HAC1* transcript, its downstream targets will be expressed at higher rates as well. According to a known model, an Ire1p-spliced *HAC1* mRNA is translated into a functional transcription factor and binds to the UPR motifs in the promoter regions of UPR-regulated genes, activating expression of a multitude of ER-resident chaperones and foldases (e.g. *KAR2*, *PDI1*) [30,31]. From the obtained data, we inferred that the increased dose of native *HAC1* transcript had no effect on its typical regulon elements, as neither *KAR2* nor *PDI1* expression was enhanced. Elevated *HAC1* transcript level triggered enhanced expression of *YFP*, but only under 30 °C, according to its already discussed temperature-dependent expression pattern.

4. Discussion

In the present study, the choice of the *SHs* (the engineering targets) was based on our previous transcriptomics study conducted with *Y. lipolytica* [24], as well as the literature data [2,10,12]. The targets

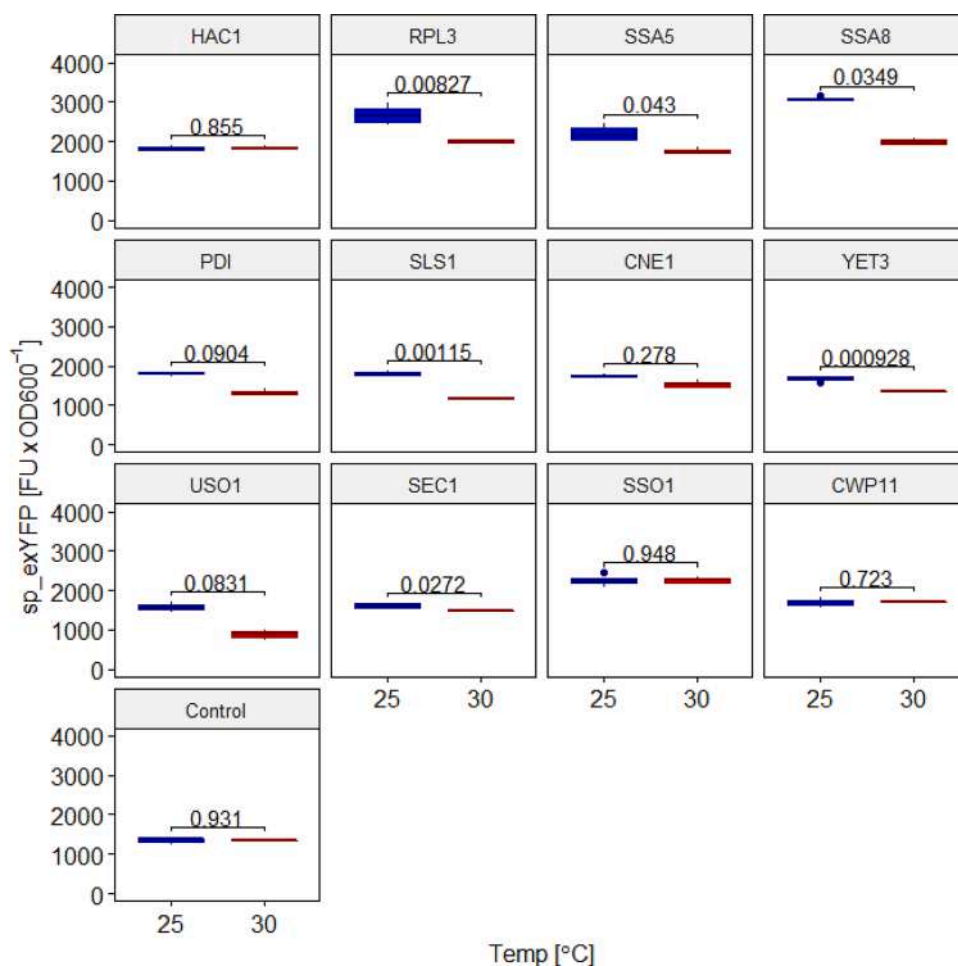


Fig. 4. Direct comparison of specific secretion values [FU OD600 nm⁻¹] reached at 72 h with an indicated co-transformants (top of each panel) under two cultivation temperatures 25 °C (blue box) and 30 °C (red box). Numbers indicate p value of t-Student's test. X axis: cultivation temperature 25–25 °C, 30–30 °C. Y axis: specific secretion (sp_exYFP) of YFP in [FU OD600 nm⁻¹]. The boxes cover the 50% of input data points and a median indicated as solid line.

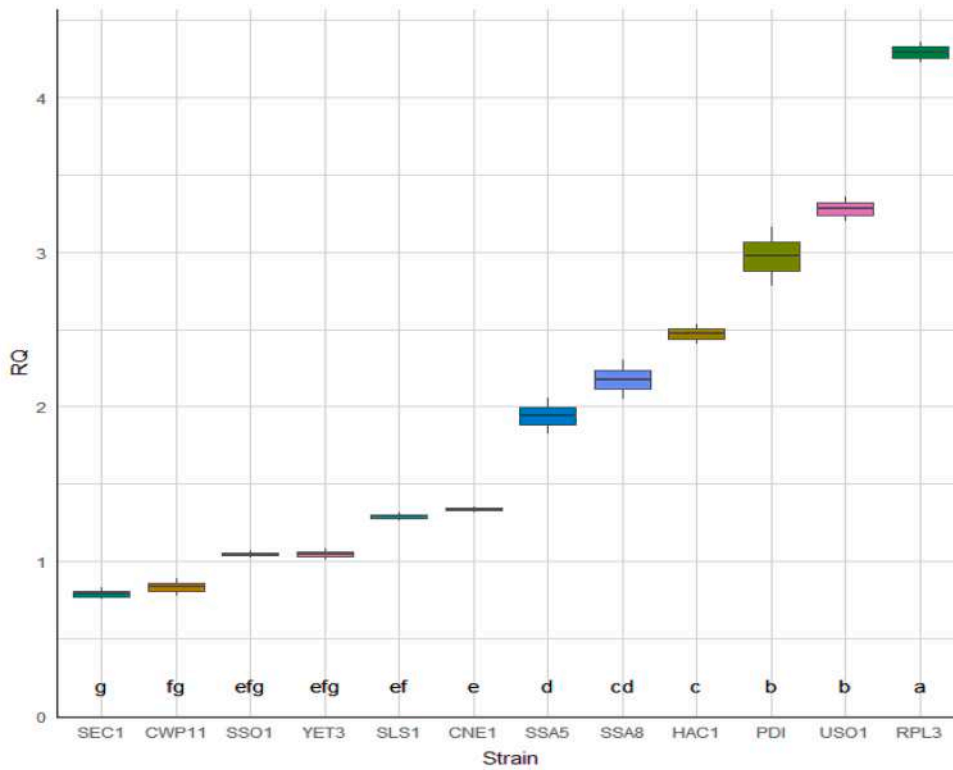
were previously identified as differently regulated genes (DEGs; either up- or downregulated) in response to overexpression of one of three different heterologous reporter proteins targeted for secretion – a small fluorescent protein (scYFP), an alpha-amylase rich in cysteines (SoA), and a glucoamylase rich in glycosylation sites (TIG), as well as one intracellular reporter (inYFP). Hence, according to a rationale and strategy adopted previously by Gasser et al. (2007), different polypeptides were used across the study (from identification of the targets to studying of the effects of their overexpression), which precludes consideration of protein-specific effects, and leads to generalizable conclusions. The set of SHs studied here (Table 1) covers those operating throughout a transcription-translation-folding-maturation-secretion pipeline, including a major transcription factor governing UPR activation (HAC1), a ribosome element (RPL3), cytosolic chaperones (SSA5 and SSA8), ER-residents involved in folding and stress signaling (PDI1, SLS1, CNE1), as well as components of proximal (YET3, USO1) and distal (SEC1, SSO1, CWP11) vesicular transportation of the cargo protein outside the cell. As shown by the transcriptomics analysis [24], the genes included in the set, responded differently to the imposed metabolic burden, e.g. *CWP11* was highly upregulated when SoA and scYFP reporters were intensively synthesized and secreted, which also applies to SSAs and *USO1*; on the other hand, *SEC1* and *SSO1* were significantly downregulated under these conditions. Yet, in the present study, we attempted solely overexpression (and not deletion) of the selected SHs, irrespective of the previously observed regulation direction. The approach was based on former literature reports on beneficial outcomes of the SHs overexpression in different fungal production platforms

(specific references are given hereafter).

In addition, we decided to implement two different temperatures for testing the SHs operation. The rationale driving this approach were the findings by [32] and [12] on significant temperature-dependency of multiple secretory enhancers co-overexpressed in *P. pastoris* and *S. cerevisiae*, respectively, studied in a range from 20 °C to 37 °C. In the former study, 40% improvement in the target protein production was achieved by the reduction of the temperature from 25 °C to 20 °C. On the other hand, the latter research demonstrated that, for example, *ERO1* was the only SH from amongst tested, that exhibited beneficial impact on synthesis and secretion of a specific reporter at 20 °C. In contrast, most significant effects of several other SHs overexpression (including cell wall genes *CCW12/CWP2* and ribosomal *RPP0*) towards secretion of specific reporter proteins, were elicited under elevated temperatures 30–37 °C [12]. That study indirectly indicated that under regular/elevated temperatures, the secretory pathway is burdened due to excessive transcripts provision, and hence, only in those conditions the effect of several SHs overexpression was apparent. Still, that previous research founded a solid rationale behind testing the double transformants constructed here under different temperatures. In the present study, we adopted the induction temperature of 25 °C, as used previously for testing different SHs in *P. pastoris* [10]. In addition, since the former transcriptomics study, that provided data for the SHs selection [24], was conducted under 30 °C, which is a typical temperature for *Y. lipolytica* cultivations, that thermal condition was used as a reference.

Kinetic studies (Fig. 1.x.3) showed that while overexpression of SHs did not have any impact on growth rate of the recombinant strains, the

A. 30°C



B. 25°C

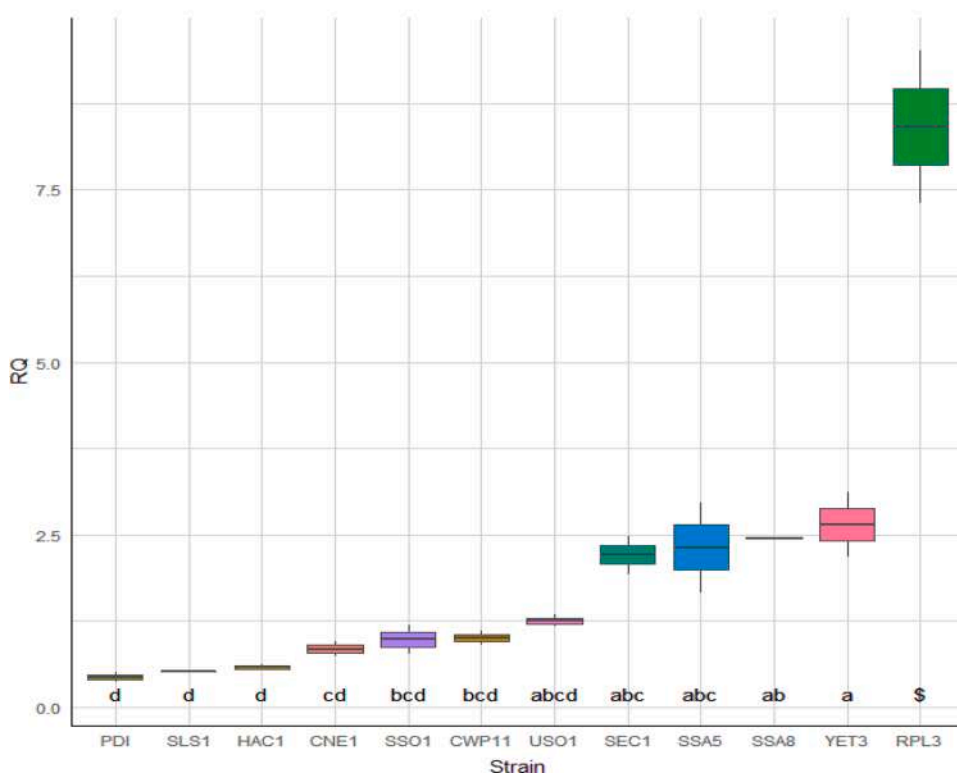


Fig. 5. Expression level of the heterologous gene encoding YFP in the *Y. lipolytica* co-transformants under two cultivation temperatures 30 °C (A) and 25 °C (B). X axis: co-overexpressed *SH*; letters indicate homogeneous groups calculated by Tukey HSD test. Y axis: relative quantitation value calculated according to ddCt model. The boxes cover the 50% of input data points and a median indicated as solid line. Value 1.0 indicates lack of change in the YFP expression level between the control strains and the indicated co-transformants.

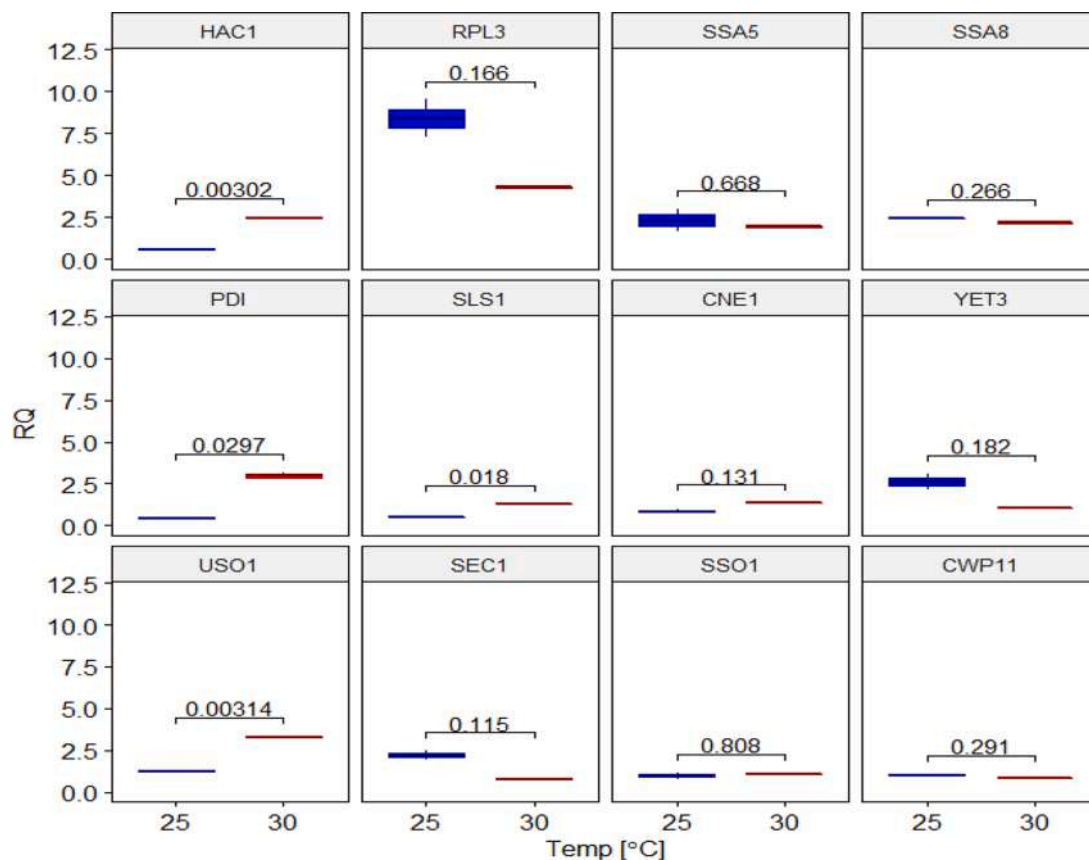


Fig. 6. Direct comparison of the expression level of the heterologous gene (YFP) in the indicated co-transformants (top of each panel) under two cultivation temperatures 25 °C (blue box) and 30 °C (red box). Numbers indicate p value of t-Student's test. X axis: cultivation temperature 25–25 °C, 30–30 °C. Y axis: relative quantitation value calculated according to ddCt model. The boxes cover the 50% of input data points and a median indicated as solid line. Value 1.0 indicates lack of change in the YFP expression level between the control strains and the indicated co-transformants.

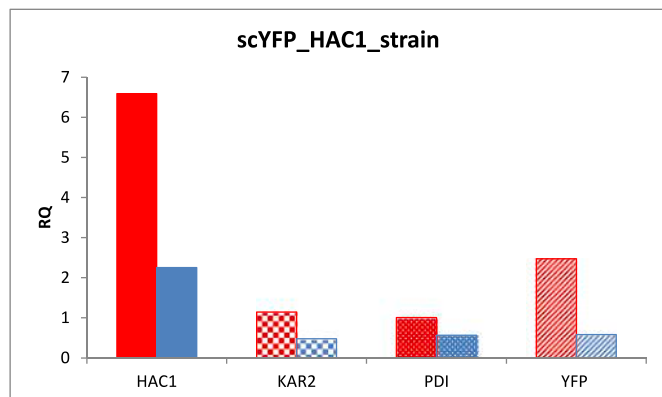


Fig. 7. Expression level of *HAC1* gene, its two direct downstream targets – *KAR2* and *PDI*, and the heterologous gene encoding YFP in the *Y. lipolytica* strain overexpressing *YFP* and *HAC1* under two cultivation temperatures 30 °C (red bars) and 25 °C (blue bars). X axis: RTqPCR-targeted gene. Y axis: relative quantitation value calculated according to ddCt model. Value 1.0 indicates lack of change in the YFP expression level between the control strains and the indicated co-transformants.

decreased temperature uniformly promoted higher biomass accumulation. These observations stay in agreement with the findings by [10] on lack of detrimental impact of *SHs* overexpression on *P. pastoris*' growth; but also, with our recent results showing that decreased temperature in fact promotes biomass growth (and heterologous protein synthesis and secretion) in *Y. lipolytica* [33]. Analysis of kinetic data presented in

Fig. 1.x.1 and 1.x.2, allowed to withdraw several interesting conclusions. Based on Fig. 1.B/C/D.1 we inferred that overexpression of *SHs* involved in translation (*RPL3*) and chaperoning activity executed in cytoplasm, directly after translation, (*SSA5* and *SSA8*) contribute to significantly increased intracellular accumulation of the reporter (inYFP), irrespectively of the adopted induction temperature. In contrast, their beneficial impact was elicited for the extracellular fraction of YFP (exYFP) solely under decreased temperature (Fig. 1. B/C/D.2). Hence, we postulate, that overexpression of *SHs*: *RPL3*, *SSA5* and *SSA8* enhanced the synthesis of the reporter protein in temperature-independent manner (as the inYFP fold change over the control at 25 °C and 30 °C was similar), but the decreased temperature was permissive for the release of the secretory pathway's capacity. Therefore, it is the combined effect of both the *SHs* overexpression and permissive temperature (25 °C) that contributes to the increases in exYFP visualized in Fig. 1.B/C/D.2 (numerical values 1.98-, 1.62- and 2.27-fold over the control [sp_exYFP] in Table 2 and Fig. 2.B). Hence, the observation visualized in Fig. 4, on statistical significance of the temperature on sp_exYFP parameter, relates to such a combined effect of genetic and environmental factors, rather than direct operation of the *SHs* in a temperature-dependent manner. Our data on the target gene expression levels corroborate this statement, as *YFP* expression was significantly enhanced due to overexpression of all these three *SHs* under both adopted temperatures (even slightly higher at 30 °C for *RPL3*), (Fig. 5.A.B), and for all three *SHs* the temperature was a not-significant factor affecting *YFP* expression (Fig. 6). However, the macroscopic outcomes, seen as significantly elevated exYFP, were observed solely under 25 °C. On top of that, since the effect in inYFP (Fig. 1.B/C/D.1) was seen under both temperatures, it is a direct

evidence on relieved secretory pathway under 25 °C. In the literature, overexpression of *SSA8*'s homologue (*SSA4*) brought 40% improvement in heterologous protein synthesis and secretion in *P. pastoris* [10]. *S. cerevisiae*'s *SSA1* overexpression in *P. pastoris* resulted in 4-fold enhancement in a recombinant protein synthesis [34]. In that latter study, it was postulated and experimentally evidenced that combined overexpression of chaperones and folding helpers, such as *KAR2* or *PDII*, can further improve the secretion levels. Overexpression of a ribosomal protein *RPPO* (as *RPL3* in this study) in *S. cerevisiae* enhanced final yields of secretory reporters by over 2.5-fold [12]. Based on our and the literature data, it can be concluded that the genes involved in translation and/or having chaperoning activity are useful as *SHs*, enabling substantial enhancement of the heterologous secretory protein synthesis. Combination of the *SHs* co-overexpression and the temperature downshift, allows to reach even better results in terms of extracellular yields of the target protein.

Significant temperature-dependency was also observed in intracellular retention of YFP upon overexpression of syntaxin *SSO1* (Fig. 1.K.1). In this case, accumulation (inYFP) of the reporter was significantly higher at 30 °C, but the protein's secretion was more efficient under 25 °C (Fig. 1.K.2). Careful analysis of the data presented in Fig. 1.K.1/2, Fig. 2.A.B and Table 2, combined with argumentation presented by [12] on burdened secretory pathway under elevated temperatures, imply that 30 °C was the permissive environmental condition that allowed demonstration of the *SSO1*-overexpression phenotype. In fact, the combined action of the *SSO1* overexpression and 30 °C resulted in the highest sp_exYFP readout under this temperature, from those studied here. It can be thus further inferred that the targeting/fusion of the Golgi-derived secretory vesicle to the cell membrane, at which *SSO*'s operates, is the bottleneck of the secretory pathway under regular temperature of *Y. lipolytica* cultivation, which can be alleviated by either *SSO1* overexpression OR decreased temperature (exYFP: 13,511 ± 593.97 vs 11,780.5 ± 252.44 [FU]±SD, for the 25 °C and *SSO1* overexpression, respectively). As reported previously, overexpression of homologous syntaxins *SSO1* and *SSO2* in *S. cerevisiae* resulted in 4–6-fold increase in secretion of heterologous bacterial enzyme [35], which well aligns with our observations and finds the rationale behind the manipulations within late-Golgi to cell membrane traffic. Overexpression of *S. cerevisiae*'s *SSO2* in *P. pastoris* triggered 20% increase in the secretory reporter synthesis when induced under 25 °C [10]. Another inter-species modification including *K. lactis*' *SSO1* homologue's overexpression in *S. cerevisiae* led to enhanced production of secreted proteins in the host cell [13], confirming its universally important role in the process of polypeptides secretion.

It was very surprising to see the negative effect of *USO1* co-overexpression on exYFP levels under 30 °C (reduced by nearly 40%; Fig. 1.I.2, Fig. 2.A, Table 2) and lack of impact of increased *USO1* dose on exYFP under 25 °C (Fig. 1.I.1/2, Fig. 2.B, Table 2). *USO1* is involved in vesicle mediated ER to Golgi transport, which is most frequently pointed as a key bottleneck in the secretory pathway [2,36]. Previously, *USO1*-encoding gene was identified as one of the most up-regulated DEGs upon high synthesis and secretion of two secretory reporters (SoA and scYFP; [24]), which well aligned with the claimed, key limitation at this stage of the secretory pathway. Interestingly, *USO1* was also identified as a promising genetic engineering target, based on microfluidic screening and whole genome sequencing in *S. cerevisiae* [23]. Our current data suggest that *USO1* co-overexpression results in no significant change (25 °C, Fig. 3.B) or inferior levels of sp_exYFP (30 °C, Fig. 3.A) in *Y. lipolytica*, when compared to the control strain. Results shown in (Fig. 5.A, Fig. 6) indicate that *USO1* had no regulatory role in promoting *YFP*'s expression (which was significantly higher at 30 °C, as expected), but that elevated expression was accompanied by significantly reduced (sp)_exYFP levels (Fig. 1.I.2, Fig. 2.A). It is thus suggested, that excessive provision of *YFP* and *USO1* transcripts triggered substantial stress for ER-localized foldases / translocon elements, which awaken UPR leading to the YFP degradation. Under 25 °C, its impact on

YFP expression, accumulation and secretion was negligible (Fig. 5.B, Fig.1.1.1/2, Fig. 2.B, Table 2). In previous studies, several targets localized to Golgi compartment were tested for their efficiency as *SHs* [10], including *COG6*, *COY1*, *IMH1*, and *SEC31*. Their overexpression showed no significant improvement in the target protein secretion, which well corresponds with our current observations. It also leaves an open question on both—mode of these genes regulation / operation, which contribute to such unexpected outcomes; but also—on correctness of assigning the key limiting role of this specific stage of the secretory pathway (as commonly postulated). In contrast, moderate overexpression of *SEC16* (involved in protein translocation from ER to Golgi) increased the secretion of a heterologous enzymatic reporter, and two other proteins [37].

The set of *SHs* studied in the present research, also included also the most straightforward targets, frequently adopted as secretion enhancers in fungi, namely transcription factor *HAC1*, and ER resident chaperones *PDII*, *SLS1*, and *CNE1*. Overexpression of these genes is considered one of most useful approaches in the yeast secretion engineering, as ER-localized protein folding is claimed, by many authors, the most rate-limiting bottleneck in heterologous protein secretion [2]. Overexpression of *PDII* and *HAC1* had a significant positive impact on secretion of a target protein in *P. pastoris*, yielding 1.5–2.3-fold improvement, depending on the culturing mode and measure [10]. Co-overexpression of calnexin *CNE1* with any of four reporters (three glycosylated and one glycosylation-free) triggered uniformly positive impact on the reporters secretion in *H. polymorpha* [18]. On the other hand, the effect of *PDII* co-overexpression is known to be strongly dependent on biochemical characteristics of the target polypeptide (for excellent reviews see: [2,3,8]). For example, in *K. lactis* the effect was variable, depending on presence of disulfide bonds in the target proteins [17]. In that study, the beneficial effect was observed solely, when the target protein was “disulfidebonded”, while secretion of disulfide-free interleukin remained unaffected. It could, at least to some extent, explain our observations on lack of any substantial improvement in (sp)_exYFP levels upon *PDII* co-overexpression under 30 °C (Fig. 1.F. 2, Fig. 2.A, Fig. 3.A, Table 2). In contrast, significant but moderate improvement (~30%) in (sp)_exYFP due to *PDII* overexpression could be seen after the temperature downshift (Fig. 1.F. 2, Fig. 2.B, Fig. 3.B, Table 2). The mechanism, by which this temperature dependency was executed, seems similar to what was postulated above for *USO1*, especially considering the expression profile of the target gene (Fig. 5.A.B). Exactly the same relationship between the gene of interest expression level, secretion of the reporter and the temperature, was observed for the remaining *SHs* from this group: *HAC1*, *SLS1*, *CNE1* (Fig. 1.A/E/F/G.2, Fig. 2.A.B, Fig. 3.AB, Fig. 4, Fig. 5.A.B, Fig. 6, Table 2). For all these genes, any improvement in expression level of the *YFP* gene over the control was seen under 30 °C (Fig. 5.A), but it was associated with no improvements in its synthesis and secretion (Fig. 2.A). Decreased expression of the reporter under 25 °C (Fig. 5.B) was accompanied by significantly improved secretion of the heterologous protein (all grouped in a distinct group “d”; Fig. 2.B). The same observation was reported previously for *P. pastoris*, for which, the change in the cultivation temperature from 25 °C to 20 °C led to a 1.4-fold increase of specific product secretion rate, although the transcriptional levels of the product genes (Fab light and heavy chain) were significantly reduced [32]. Therefore, based on our current and the literature data, it is postulated that without any specific reverse regulation mechanism, the transcription of the heterologous genes is decreased under the decreased temperature, which gives sufficient capacity to the folding machinery to correctly process the nascent polypeptides. Under these conditions more correctly folded proteins could be synthesized and secreted.

Due to its regulatory role, exploitation of *HAC1* as the secretion enhancer seems to be a particularly interesting approach. *HAC1* mediates activation of hundreds of molecular events, including increased provision of chaperones and membranes, to concertedly relieve

burdened secretory pathway [38,39]. The list of successful strategies comprising exploitation of *HAC1* as *SH* is very long, including strategies executed in *P. pastoris* [10] and *S. cerevisiae* [16,20,32,40]. In our study, co-overexpression of *HAC1* triggered significant, ~30% improvement in (sp_)exYFP levels over the control strain, under both temperatures (Fig. 2.A,B, Fig. 3.A,B, Table 2). Comparable level of improvement, reaching 50%, was previously observed in *P. pastoris* [10]. Interestingly, in the present study, the % improvements were similar under both temperatures, but the level of exYFP was > 50% higher under 25 °C. Expression of YFP-encoding gene was elevated due to *HAC1* co-overexpression solely under increased temperature, but the final sp_exYFP values remained strikingly similar under both temperatures (1818.34 ± 20.21 vs 1834.9 ± 97.52 [FU OD600 nm⁻¹] for 25 °C and 30 °C, respectively; Table 2). By overtaking the argumentation by [12], it could be stated that both temperatures adopted here allowed to visualize the effects of *HAC1* co-overexpression. Mechanistically, none of implemented temperatures specifically relieved / burdened biological processes in which *HAC1* is involved, or rather *HAC1*'s scope of activity is so broad that it exerts its positive effect when either transcription (under 25 °C) or secretion (under 30 °C) are limited. Considering the key role of *HAC1* in managing the proteins synthesis and secretion process, and our previous, unexpected findings on lack or very slight upregulation of *HAC1* upon overproduction heterologous proteins [24], we paid particular attention to the strain co-overexpressing this specific *SH* and conducted gene expression analysis of the two genes *PDI1* and *KAR2*, known to be the direct downstream targets of *HAC1* (Fig. 7). It was again very surprising to see, that neither of the genes was upregulated due to upregulation of *HAC1*. While it well corroborates our previous transcriptomics data, earlier studies conducted with *P. pastoris* [41], and *S. cerevisiae* [22] reported that upregulation of *HAC1* is accompanied by enhanced expression of its downstream targets. It could be speculated that since we co-overexpressed native *HAC1*, and not its spliced variant, the level of transcript was significantly increased (Fig. 7), but the level of operable protein was not that substantially improved to induce massive UPR (marked by *PDI1* and *KAR2* upregulation). For being translated, *HAC1* requires unconventional splicing by dimerized IRE1. Under balanced conditions, IRE1 is bounded with *KAR2*, mediated by *SLS1*, and is not available for the *HAC1* transcript processing [42–44]. Elevated dose of *SLS1* strengthens the interaction of *KAR2*-IRE1, disallowing *HAC1* processing, while Δ *sls1* genotype displays a constitutive UPR at a “maximal” level [43]. The previously postulated key role of *SLS1* in regulation of *HAC1* translation in *Y. lipolytica* [42,44], was recently corroborated by transcriptomics data showing its uniform downregulation upon overproduction of any heterologous secretory protein [24], which probably promoted *HAC1* splicing and fine-tuning of secretion. Yet, *SLS1* is a multifunctional protein, which, in addition to mediating IRE1-*KAR2* interaction, promotes the SEC63-mediated activation of *KAR2*'s ATPase activity [44,45]. It is plausible that this specific molecular function of *SLS1* underlies the observed ~30% enhancement in (sp_)exYFP under *SLS1* co-overexpression (Table 2). It also well aligns with a previous observation that Δ *sls1* genotype decreased the level of secretory proteins production [42,44]. Interestingly, the beneficial effect of *SLS1* co-overexpression required implementation of permissive, lowered temperature, releasing the secretory pathway from burden. Definitely, more insightful studies merging approaches addressing gene expression, protein abundance and the reporter protein secretion assessment is required to understand this mechanism in *Y. lipolytica*.

5. Conclusions

In summary, the present study allowed to withdraw several interesting conclusions on the impact of selected *SHs* overexpression and the modulating role of temperature in synthesis and secretion of heterologous proteins in *Y. lipolytica*. Here we demonstrated that under decreased temperature, the biological processes transcription and

translation are balanced, as the amount of transcript and protein products are then highly correlated. We also observed a clear distinction in the effects of adopting as the *SHs*, the genes involved in either SYNTHESIS or TRANSPORTATION of the heterologous polypeptides. The former (*RPL3*, *SSA5*, *SSA8*), significantly enhance synthesis of the protein irrespective of culturing temperature, however, for efficient secretion of the protein accumulated in large amounts, the secretory pathway's capacity must be released by applying decreased temperature (25 °C). Exploitation of the genes involved in the protein trafficking as the *SHs*, does not give such spectacular results in terms of the amounts of synthesized target polypeptide, however, their overexpression allows to assist the secretory pathway in maintaining the capacity under not favorable thermal conditions (*SSO1*, *CWP11*). What underlies that relationship between the secretory pathway's capacity and the temperature, is, either increased or decreased, rate of transcription, which, if excessive - induces stress in the secretory pathway and causes loss in the final product, and if balanced - gives sufficient capacity to the folding machinery to correctly process the nascent polypeptides. This study provides generalizable guidelines for application in cloning/culturing strategies aiming at enhancement of heterologous protein secretion in *Y. lipolytica*.

Funding

This study was financially supported by the Ministry of Science and Higher Education (currently, Ministry of Education and Science) project number: DI 2017 000 947.

Data availability

All data generated within the study are available directly in the article or in the supplementary online materials.

Supplementary Materials

Fig.S1. Schematic representation of cloning strategy followed in this study. Two vectors from JMP62 series were co-transformed into *Y. lipolytica* recipient strain. Elements of the vectors were indicated in colors: zeta region (orange), *LoxP* and *LoxR* (dark blue), constitutive *pTEF* promoter (green), genes of interest: scYFP (A. yellow) or *SH* (B. blue) equipped in the *Bam*HI-*Avr*II fragments, the *URA3ex* (A. red) or *LEU2ex* (B. light blue) excisable auxotrophic markers. Bacterial *ori* of replication (pink) and kanamycin resistance gene *KanR* (purple), contained in the “bacterial” part of the vectors, were discarded before the yeast cell transformation by *NotI* restriction digestion.

Fig.S2. Kinetics of growth (0.1) and specific fluorescence of intracellular (sp_inYFP; 0.1) and extracellular (exYFP; 0.2) fractions of the heterologous protein YFP in randomly selected four sub-clones of each co-transformation type. The strains were cultures in shake flask cultures. Sub-charts “x.1”: X axis: Time [h]. Y axis: biomass accumulation (continuous lines) [OD 600 nm]; Y axis auxiliary: Specific intracellular fluorescence of YFP (sp_inYFP; dashed lines) in [FU OD600nm⁻¹]. Sub-charts “x.2”: X axis: sub-clones and time of culturing [h]; Y axis: extracellular fluorescence of YFP (exYFP) in [FU]. Given values are means \pm SD from at least biological duplicate.

Fig.S3. Multiple comparisons of *Y. lipolytica* co-transformants and the control strain presented as fold change of specific secretion measures (sp_exYFP; $\frac{\text{NUMERATOR}_{\text{FU OD600nm}^{-1}}}{\text{DENOMINATOR}_{\text{FU OD600nm}^{-1}}}$) between the strains, cultured under 30 °C (A, C) and 25 °C (B, D) for 48 h (A, B) and 72 h (C, D). To use the heatmap correctly, first identify the strain on the left and use it as a NUMERATOR in fold change calculation, than, select a strain from the top of the heatmap, to be used as DENOMINATOR in the fold change calculation. Red squares represent prevalence of the numerator in the adopted measure (1.01–2.00 FC) and blue squares represent prevalence of the denominator in the adopted measure (0–0.99 FC). Statistical

significance of the comparisons was determined by a Tukey HSD test. Numbers within the comparison squares indicate the exact value of FC.

Tab.S1. *E. coli* and *Y. lipolytica* strains used in this study.

Tab.S2. Oligonucleotides used in this study. Bold sequences correspond to the part of targeted gene. Underlined sequences correspond to introduced restriction sites (*Bam*HI/*Avr*II). Oligonucleotides used for RTqPCR are preceded by “r-” suffix.

CRedit authorship contribution statement

Paulina Korpys-Woźniak: Investigation, Methodology, Visualization, Data curation, Writing – original draft, Writing – review & editing. **Piotr Kubiak:** Formal analysis, Software, Visualization. **Ewelina Celińska:** Conceptualization, Supervision, Writing – original draft, Writing – review & editing.

Declarations of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Graphical Abstract was prepared with BioRender.com.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.btre.2021.e00669](https://doi.org/10.1016/j.btre.2021.e00669).

References

- [1] M. Schröder, The cellular response to protein unfolding stress, *Exploit. Fungi Symp. Br. Mycol. Soc. Held Univ. Manchester Sept. 2005 (2008)* 117–139, <https://doi.org/10.1017/CBO9780511902451.009>.
- [2] A. Idiris, H. Tohda, H. Kumagai, Engineering of protein secretion in yeast: strategies and impact on protein production, (2010) 403–417. 10.1007/s00253-010-2447-0.
- [3] R.J. Zahrl, B. Gasser, D. Mattanovich, P. Ferrer, Detection and elimination of cellular bottlenecks in protein-producing yeasts, *Methods Mol. Biol.* 1923 (2019) 75–95, https://doi.org/10.1007/978-1-4939-9024-5_2.
- [4] R.J. Zahrl, D.A. Peña, D. Mattanovich, B. Gasser, Systems biotechnology for protein production in *Pichia pastoris*, (n.d.) 1–31. 10.1093/femsyr/fox068. 2019.
- [5] E. Celińska, J.-M. Nicaud, Filamentous fungi-like secretory pathway strayed in a yeast system: peculiarities of *Yarrowia lipolytica* secretory pathway underlying its extraordinary performance, *Appl. Microbiol. Biotechnol.* 103 (2019) 39–52, <https://doi.org/10.1007/s00253-018-9450-2>.
- [6] C. Madzak, *Yarrowia lipolytica* strains and their biotechnological applications: how cellular biodiversity and metabolic engineering could contribute to cell factories improvement, *J. Fungi.* 7 (2021) 548, <https://doi.org/10.3390/jof7070548>.
- [7] C. Madzak, Engineering *Yarrowia lipolytica* for use in biotechnological applications: a review of major achievements and recent innovations, *Mol. Biotechnol.* 60 (2018) 621–635, <https://doi.org/10.1007/s12033-018-0093-4>.
- [8] V. Puxbaum, D. Mattanovich, B. Gasser, Quo vadis? the challenges of recombinant protein folding and secretion in *Pichia pastoris*, *Appl. Microbiol. Biotechnol.* 99 (2015) 2925–2938, <https://doi.org/10.1007/s00253-015-6470-z>.
- [9] E. Celińska, J.-M. Nicaud, W. Białas, Hydrolytic secretome engineering in *Yarrowia lipolytica* for consolidated bioprocessing on polysaccharide resources: review on starch, cellulose, xylan, and inulin, *Appl. Microbiol. Biotechnol.* 105 (2021) 975–989, <https://doi.org/10.1007/s00253-021-11097-1>.
- [10] B. Gasser, M. Sauer, M. Maurer, G. Staldmayr, D. Mattanovich, Transcriptomics-based identification of novel factors enhancing heterologous protein secretion in yeasts, *Appl. Environ. Microbiol.* 73 (2007) 6499–6507, <https://doi.org/10.1128/AEM.01196-07>.
- [11] J. Nocon, M.G. Steiger, M. Pfeffer, S.B. Sohn, T.Y. Kim, M. Maurer, H. Rußmayer, S. Pflügl, M. Ask, C. Haberhauer-Troyer, K. Ortmayr, S. Hann, G. Koellensperger, B. Gasser, S.Y. Lee, D. Mattanovich, Model based engineering of *Pichia pastoris* central metabolism enhances recombinant protein production, *Metab. Eng.* 24 (2014) 129–138, <https://doi.org/10.1016/j.ymben.2014.05.011>.
- [12] A.E. Wentz, E.V. Shusta, A novel high-throughput screen reveals yeast genes that increase secretion of heterologous proteins, *Appl. Environ. Microbiol.* 73 (2007) 1189–1198, <https://doi.org/10.1128/AEM.02427-06>.
- [13] J.H. Toikkanen, L. Sundqvist, S. Ker, *Kluyveromyces lactis* SSO1 and SEB1 genes are functional in *Saccharomyces cerevisiae* and enhance production of secreted proteins when overexpressed, *Yeast* 2 (2004) 1045–1055, <https://doi.org/10.1002/yea.1151>.
- [14] J. Hou, T. Österlund, Z. Liu, D. Petranovic, J. Nielsen, T. Osterlund, Z. Liu, D. Petranovic, J. Nielsen, Heat shock response improves heterologous protein secretion in *Saccharomyces cerevisiae*, *Appl. Microbiol. Biotechnol.* 97 (2013) 3559–3568, <https://doi.org/10.1007/s00253-012-4596-9>.
- [15] J. Hou, K.E.J. Tyo, Z. Liu, D. Petranovic, J. Nielsen, Metabolic engineering of recombinant protein secretion by *Saccharomyces cerevisiae*, *FEMS Yeast Res* 12 (2012) 491–510, <https://doi.org/10.1111/j.1567-1364.2012.00810.x>.
- [16] J. Hou, H. Tang, Z. Liu, T. Österlund, J. Nielsen, D. Petranovic, Management of the endoplasmic reticulum stress by activation of the heat shock response in yeast, *FEMS Yeast Res* 14 (2014) 481–494, <https://doi.org/10.1111/1567-1364.12125>.
- [17] T. Lodi, B. Neglia, C. Donnini, Secretion of human serum albumin by *Kluyveromyces lactis*, Overexpressing KIPDII and KIERO1, *Appl. Environ. Microbiol.* 71 (2005) 4359–4363, <https://doi.org/10.1128/AEM.71.8.4359>.
- [18] J. Klabunde, G. Kunze, G. Gellissen, C.P. Hollenberg, Integration of heterologous genes in several yeast species using vectors containing a *Hansenula polymorpha*-derived rDNA-targeting element, *FEMS Yeast Res* 4 (2003) 185–193, [https://doi.org/10.1016/S1567-1356\(03\)00148-X](https://doi.org/10.1016/S1567-1356(03)00148-X).
- [19] P. Korpys-Woźniak, P. Kubiak, W. Białas, E. Celińska, Impact of overproduced heterologous protein characteristics on physiological response in *Yarrowia lipolytica* steady-state-maintained continuous cultures, *Appl. Microbiol. Biotechnol.* 104 (2020) 9785–9800, <https://doi.org/10.1007/s00253-020-10937-w>.
- [20] K.E.J. Tyo, Z. Liu, D. Petranovic, J. Nielsen, Imbalance of heterologous protein folding and disulfide bond formation rates yields runaway oxidative stress, *BMC Biol* 10 (2012) 16, <https://doi.org/10.1186/1741-7007-10-16>.
- [21] J.L. Martínez, E. Meza, D. Petranovic, J. Nielsen, The impact of respiration and oxidative stress response on recombinant α -amylase production by *Saccharomyces cerevisiae*, *Metab. Eng. Commun.* 3 (2016) 205–210, <https://doi.org/10.1016/j.meteno.2016.06.003>.
- [22] Z. Liu, J. Hou, J.L. Martínez, D. Petranovic, J. Nielsen, Correlation of cell growth and heterologous protein production by *Saccharomyces cerevisiae*, *Appl. Microbiol. Biotechnol.* 97 (2013) 8955–8962, <https://doi.org/10.1007/s00253-013-4715-2>.
- [23] M. Huang, Y. Bai, S.L. Sjöström, B.M. Hallström, Z. Liu, D. Petranovic, M. Uhlén, H. N. Joensson, H. Andersson-Svahn, J. Nielsen, Microfluidic screening and whole-genome sequencing identifies mutations associated with improved protein secretion by yeast, *Proc. Natl. Acad. Sci. U. S. A.* 112 (2015) E4689–E4696, <https://doi.org/10.1073/pnas.1506460112>.
- [24] P. Korpys-Woźniak, E. Celińska, Global transcriptome profiling reveals genes responding to overproduction of a small secretory, a high cysteine- and a high glycosylation-bearing protein in *Yarrowia lipolytica*, *Biotechnol. Reports.* 31 (2021) e00646, <https://doi.org/10.1016/j.btre.2021.e00646>.
- [25] G. Barth, C. Gaillardin, *Yarrowia lipolytica*, in: K. Wolf (Ed.), *Nonconv. Yeasts Biotechnol. A Handb.*, Springer Berlin Heidelberg, 1996, pp. 313–388, <https://doi.org/10.1007/978-3-642-79856-6>.
- [26] J. Sambrook, D. Russell, *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory Press, New York, 2001.
- [27] E. Celińska, M. Borkowska, W. Białas, P. Korpys, J.M. Nicaud, Robust signal peptides for protein secretion in *Yarrowia lipolytica*: identification and characterization of novel secretory tags, *Appl. Microbiol. Biotechnol.* 102 (2018) 5221–5233, <https://doi.org/10.1007/s00253-018-8966-9>.
- [28] J.M. Nicaud, C. Madzak, P. Van Den Broek, C. Gysler, P. Duboc, P. Niederberger, C. Gaillardin, Protein expression and secretion in the yeast *Yarrowia lipolytica*, *FEMS Yeast Res* 2 (2002) 371–379, [https://doi.org/Pii%20S1567-1356\(02\)00082-X](https://doi.org/Pii%20S1567-1356(02)00082-X).
- [29] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻ $\Delta\Delta$ CT method, *Methods* 25 (2001) 402–408, <https://doi.org/10.1006/meth.2001.1262>.
- [30] J.S. Cox, P. Walter, A novel mechanism for regulating activity of a transcription factor that controls the unfolded protein response, *Cell* 87 (1996) 391–404, [https://doi.org/10.1016/S0092-8674\(00\)81360-4](https://doi.org/10.1016/S0092-8674(00)81360-4).
- [31] K. Mori, T. Kawahara, H. Yoshida, H. Yanagi, T. Yura, Signalling from endoplasmic reticulum to nucleus: transcription factor with a basic-leucine zipper motif is required for the unfolded protein-response pathway, *Genes cells* (1996) 803–817. 10.1046/j.1365-2443.1996.d01-274.x.
- [32] B. Gasser, M. Maurer, J. Rautio, M. Sauer, A. Bhattacharyya, M. Saloheimo, M. Penttilä, D. Mattanovich, Monitoring of transcriptional regulation in *Pichia pastoris* under protein production conditions, *BMC Genomics* 8 (2007) 1–18, <https://doi.org/10.1186/1471-2164-8-179>.
- [33] M. Kubiak, W. Białas, E. Celińska, Thermal treatment improves a process of crude glycerol valorization for the production of a heterologous enzyme by *Yarrowia lipolytica*, *Biotechnol. Reports.* 31 (2021) e00648, <https://doi.org/10.1016/j.btre.2021.e00648>.
- [34] W. Zhang, H.L. Zhao, C. Xue, X.H. Xiong, X.Q. Yao, X.Y. Li, H.P. Chen, Z.M. Liu, Enhanced secretion of heterologous proteins in *Pichia pastoris* following overexpression of *Saccharomyces cerevisiae* chaperone proteins, *Biotechnol. Prog.* 22 (2006) 1090–1095, <https://doi.org/10.1021/bp060019r>.
- [35] L. Ruohonen, J. Toikkanen, V. Tieaho, M. Outola, H. Soderlund, S. Keranen, Enhancement of protein secretion in *Saccharomyces cerevisiae* by overproduction of Sso protein, a late-acting component of the secretory machinery, *Yeast* 13 (1997) 337–351, 10.1002/(sici)1097-0061(19970330)13:4<337::aid-yea98>3.3.co;2-b.
- [36] J.R. Shuster, Gene expression in yeast: protein secretion, *Curr. Opin. Biotechnol.* 2 (1991) 685–690, [https://doi.org/10.1016/0958-1669\(91\)90035-4](https://doi.org/10.1016/0958-1669(91)90035-4).
- [37] J. Bao, M. Huang, D. Petranovic, J. Nielsen, Moderate expression of SEC16 increases protein secretion by *Saccharomyces cerevisiae*, *Appl. Environ. Microbiol.* 83 (2017), <https://doi.org/10.1128/AEM.03400-16>.

- [38] K.J. Kauffman, E.M. Pridgen, F.J. Doyle, P.S. Dhurjati, A.S. Robinson, Decreased protein expression and intermittent recoveries in BiP levels result from cellular stress during heterologous protein expression in *Saccharomyces cerevisiae*, *Biotechnol. Prog.* 18 (2002) 942–950, <https://doi.org/10.1021/bp025518g>.
- [39] M. Guerfal, S. Ryckaert, P.P. Jacobs, P. Ameloot, K. Van Craenenbroeck, R. Derycke, N. Callewaert, The HAC1 gene from *Pichia pastoris*: characterization and effect of its overexpression on the production of secreted, surface displayed and membrane proteins, *Microb. Cell Fact.* 9 (2010) 1–12, <https://doi.org/10.1186/1475-2859-9-49>.
- [40] J. Liu, Q. Han, Q. Cheng, Y. Chen, R. Wang, X. Li, Y. Liu, D. Yan, Efficient expression of human lysozyme through the increased gene dosage and Co-expression of transcription factor hac1p in *Pichia pastoris*, *Curr. Microbiol.* 77 (2020) 846–854, <https://doi.org/10.1007/s00284-019-01872-9>.
- [41] A.R. Hesketh, J.I. Castrillo, T. Sawyer, D.B. Archer, S.G. Oliver, Investigating the physiological response of *Pichia (Komagataella) pastoris* GS115 to the heterologous expression of misfolded proteins using chemostat cultures, *Appl. Microbiol. Biotechnol.* 97 (2013) 9747–9762, <https://doi.org/10.1007/s00253-013-5186-1>.
- [42] A. Boisramé, J.M. Beckerich, C. Gaillardin, Sls1p, an endoplasmic reticulum component, is involved in the protein translocation process in the yeast *Yarrowia lipolytica*, *J. Biol. Chem.* 271 (1996) 11668–11675, <https://doi.org/10.1074/jbc.271.20.11668>.
- [43] A. Babour, M. Kabani, A. Boisramé, J. Beckerich, Characterization of Ire1 in the yeast *Yarrowia lipolytica* reveals an important role for the Sls1 nucleotide exchange factor in unfolded protein response regulation, (2008) 337–346. 10.1007/s00294-008-0190-1.
- [44] A. Boisramé, M. Kabani, J.M. Beckerich, E. Hartmann, C. Gaillardin, Interaction of Kar2p and Sls1p is required for efficient co- translational translocation of secreted proteins in the yeast *Yarrowia lipolytica*, *J. Biol. Chem.* 273 (1998) 30903–30908, <https://doi.org/10.1074/jbc.273.47.30903>.
- [45] M. Kabani, J.-M. Beckerich, C. Gaillardin, Sls1p stimulates Sec63p-mediated activation of Kar2p in a conformation-dependent manner in the yeast endoplasmic reticulum, *Mol. Cell. Biol.* 20 (2000) 6923–6934, <https://doi.org/10.1128/mcb.20.18.6923-6934.2000>.

SUPPLEMENTARY

Secretary helpers for enhanced production of heterologous proteins in *Yarrowia lipolytica*

List of Supplementary Materials

Fig.S1. Schematic representation of cloning strategy followed in this study. Two vectors from JMP62 series were co-transformed into *Y. lipolytica* recipient strain. Elements of the vectors were indicated in colors: zeta region (*orange*), LoxP and LoxR (*dark blue*), constitutive pTEF promoter (*green*), genes of interest: scYFP (A. *yellow*) or SH (B. *blue*) equipped in the BamHI-AvrII fragments, the URA3ex (A. *red*) or LEU2ex (B. *light blue*) excisable auxotrophic markers. Bacterial *ori* of replication (*pink*) and kanamycin resistance gene *KanR* (*purple*), contained in the “bacterial” part of the vectors, were discarded before the yeast cell transformation by *NotI* restriction digestion.

Fig.S2. Kinetics of growth (0.1) and specific fluorescence of intracellular (sp_inYFP; 0.1) and extracellular (exYFP; 0.2) fractions of the heterologous protein YFP in randomly selected four sub-clones of each co-transformation type. The strains were cultures in shake flask cultures. Sub-charts “x.1”: X axis: Time [h]. Y axis: biomass accumulation (continuous lines) [OD 600 nm]; Y axis auxiliary: Specific intracellular fluorescence of YFP (sp_inYFP; dashed lines) in [FU OD600nm⁻¹]. Sub-charts “x.2”: X axis: sub-clones and time of culturing [h]; Y axis: extracellular fluorescence of YFP (exYFP) in [FU]. Given values are means ±SD from at least biological duplicate.

Fig. S3. Multiple comparisons of *Y. lipolytica* co-transformants and the control strain presented as fold change of specific secretion measures (sp_exYFP; $\frac{NUMERATOR_FU\ OD600nm^{-1}}{DENOMINATOR_FU\ OD600nm^{-1}}$) between the strains, cultured under 30 °C (A, C) and 25 °C (B, D) for 48 h (A, B) and 72 h (C, D). To use the heatmap correctly, first identify the strain on the left and use it as a NUMERATOR in fold change calculation, than, select a strain from the top of the heatmap, to be used as DENOMINATOR in the fold change calculation. Red squares represent prevalence of the numerator in the adopted measure (1.01 – 2.00 FC) and blue squares represent prevalence of the denominator in the adopted measure (0 - 0.99 FC). Statistical significance of the comparisons was determined by a Tukey HSD test. Numbers within the comparison squares indicate the exact value of FC.

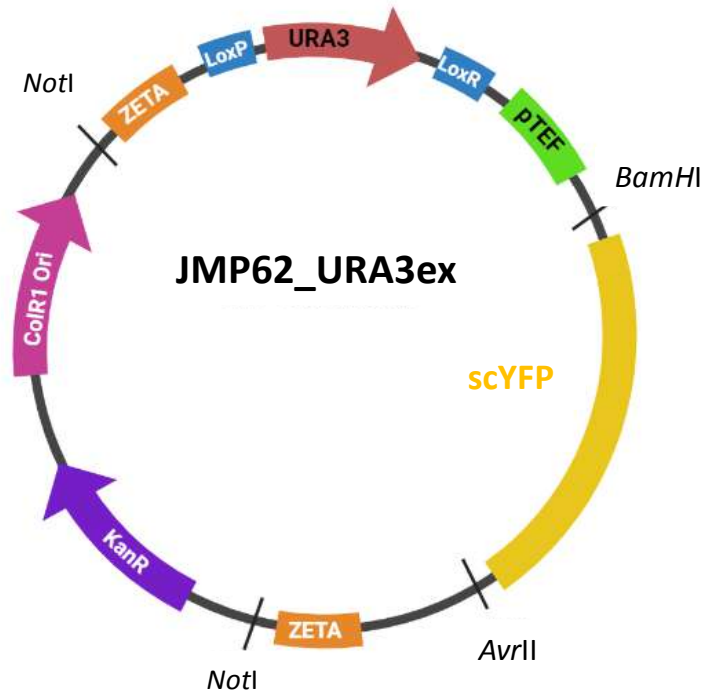
Fig. S4. Direct comparison of specific secretion values [FU OD600 nm⁻¹] reached at 48 h with an indicated co-transformants (top of each panel) under two cultivation temperatures 25 °C (blue box) and 30 °C (red box). Numbers indicate p value of t-Student's test. X axis: cultivation temperature 25–

25 °C, 30–30 °C. Y axis: specific secretion (sp_exYFP) of YFP in [FU OD600 nm⁻¹]. The boxes cover the 50% of input data points and a median indicated as solid line.

Tab.S1. *E. coli* and *Y. lipolytica* strains used in this study.

Tab.S2. Oligonucleotides used in this study. Bold sequences correspond to the part of targeted gene. Underlined sequences correspond to introduced restriction sites (*Bam*HI/*Avr*II). Oligonucleotides used for RTqPCR are preceded by “r-t” suffix.

Fig.S1.A.



B.

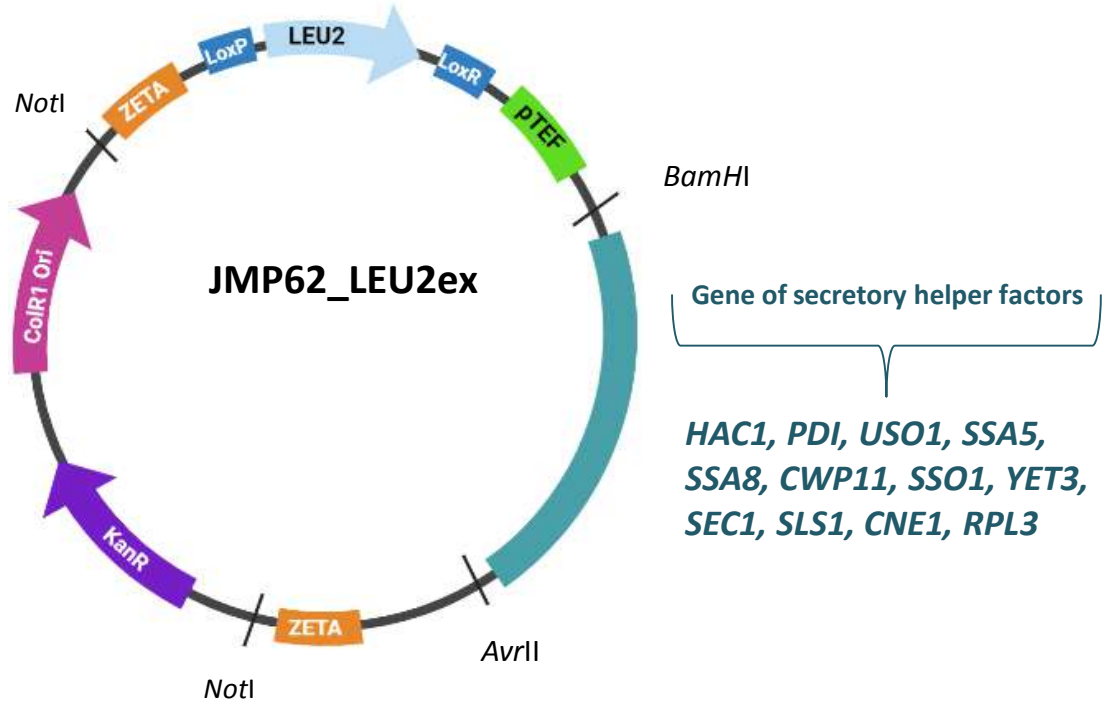
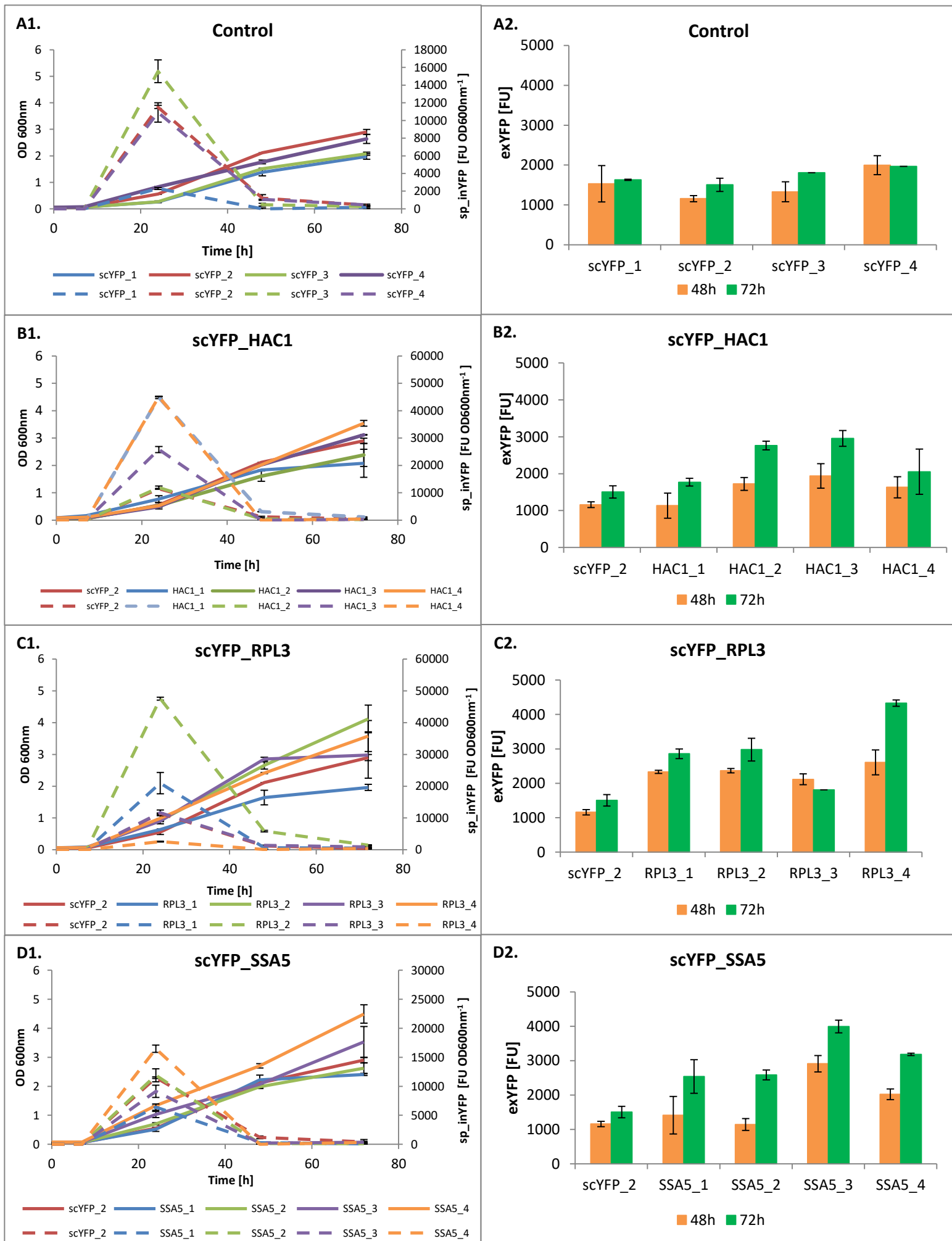
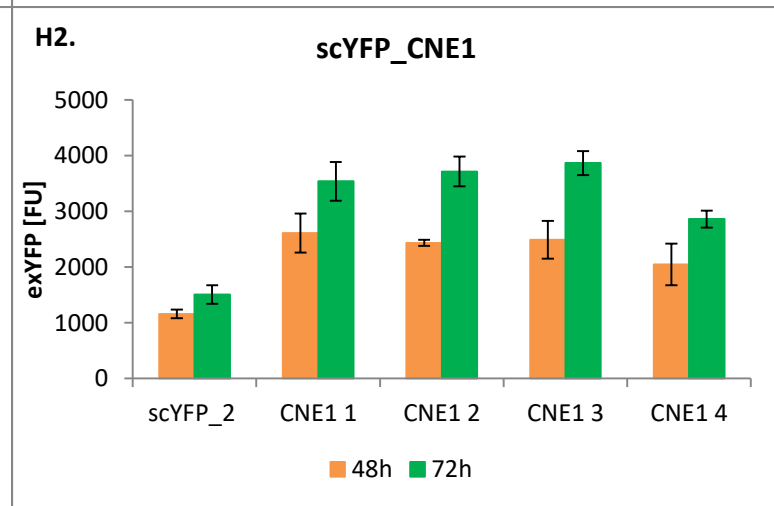
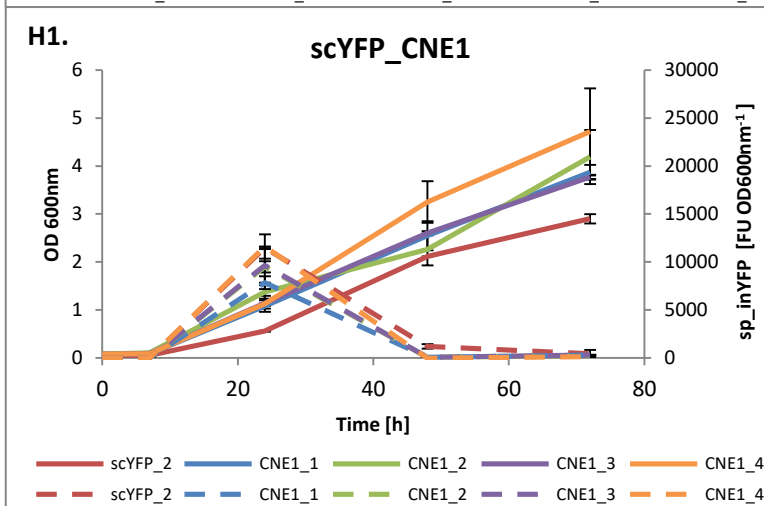
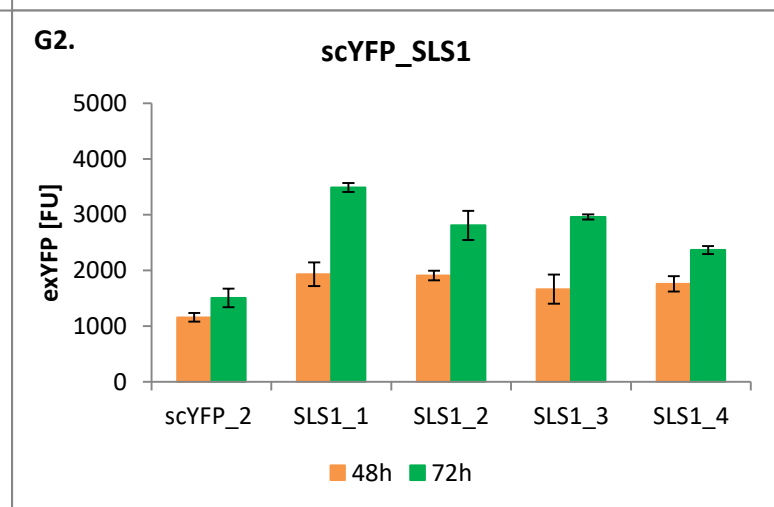
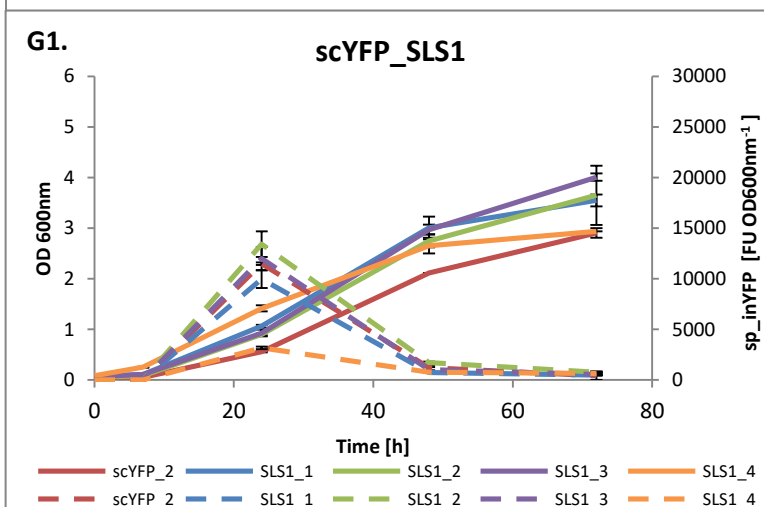
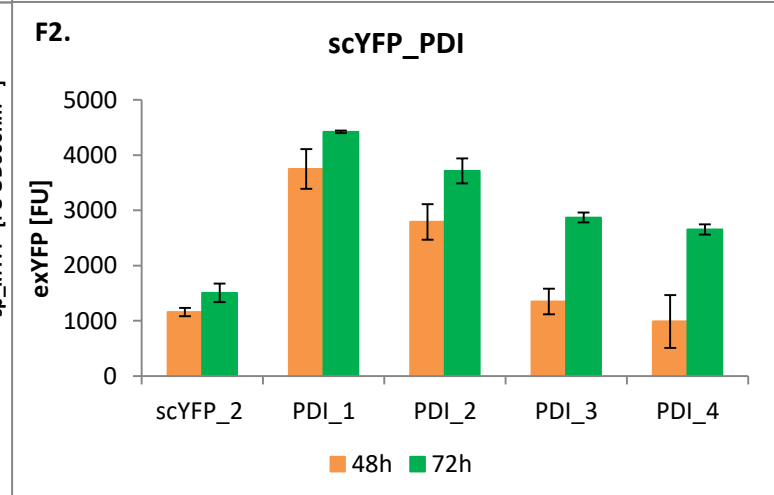
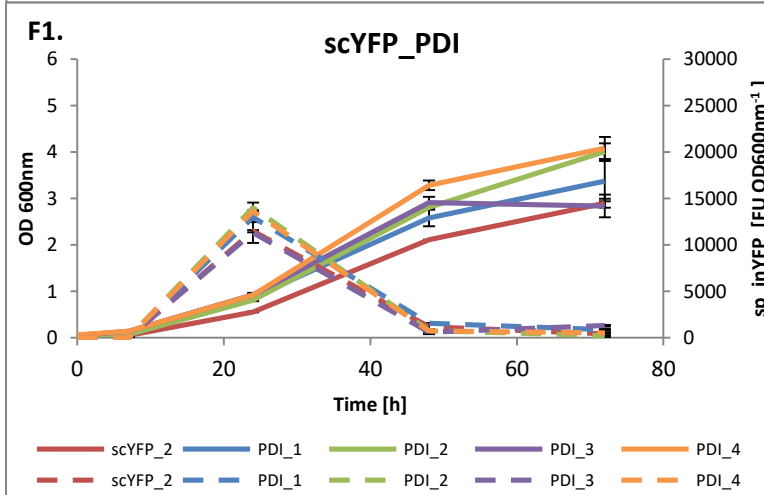
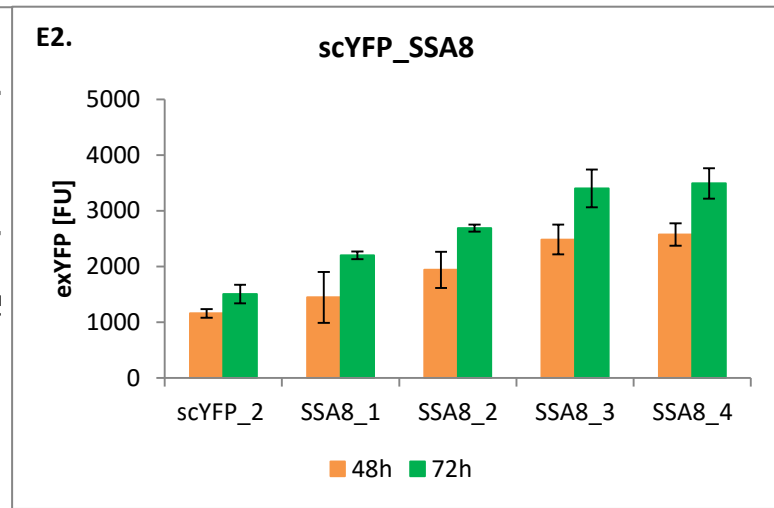
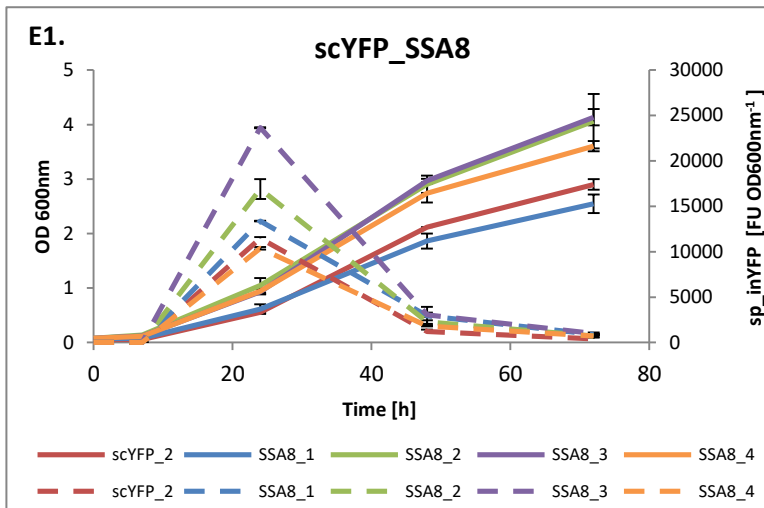
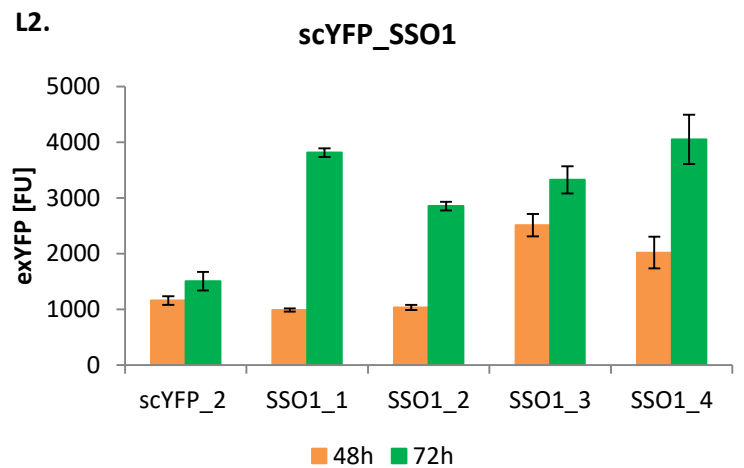
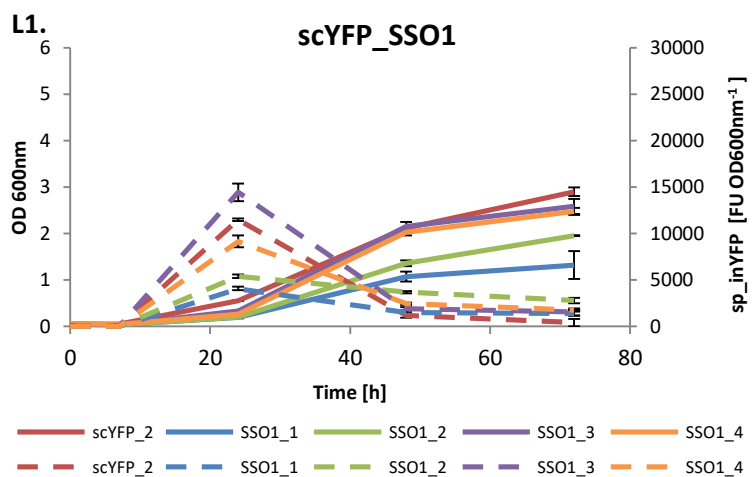
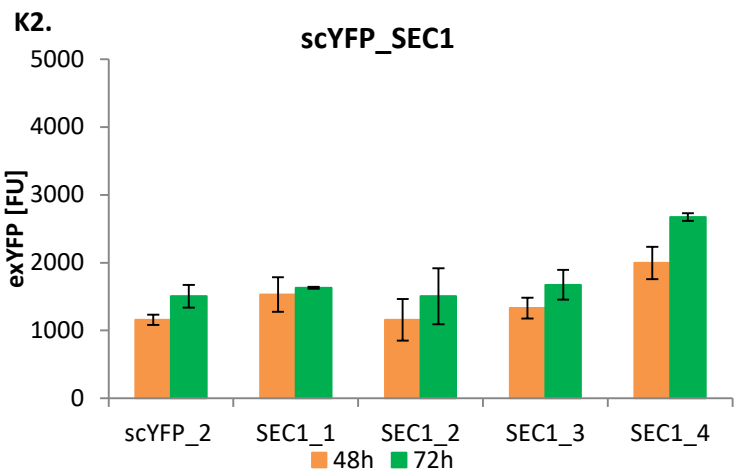
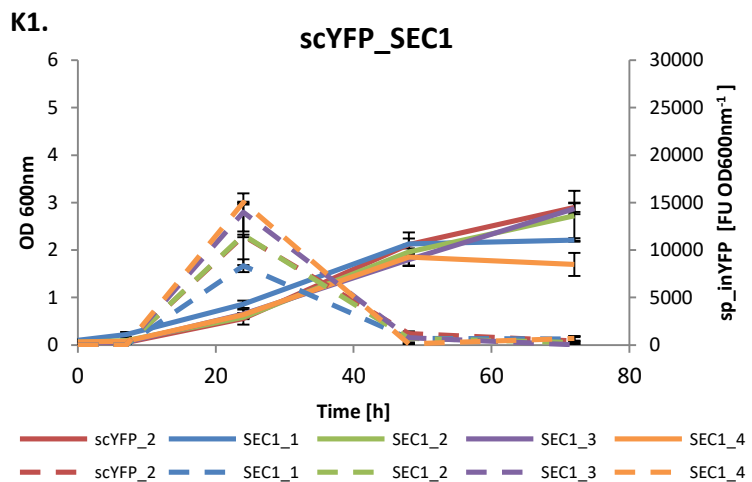
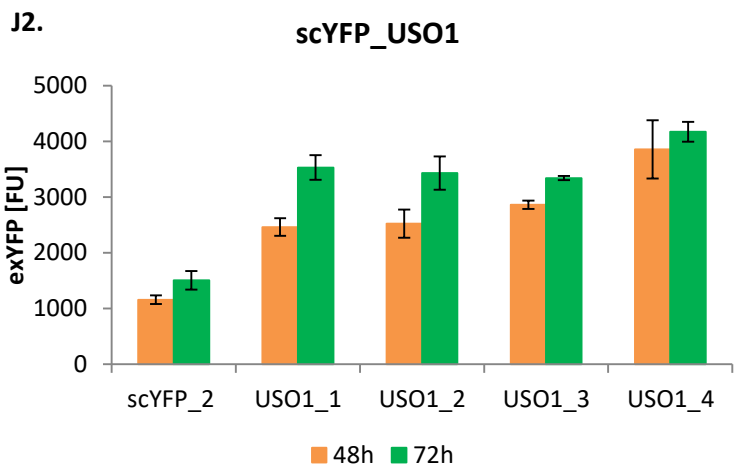
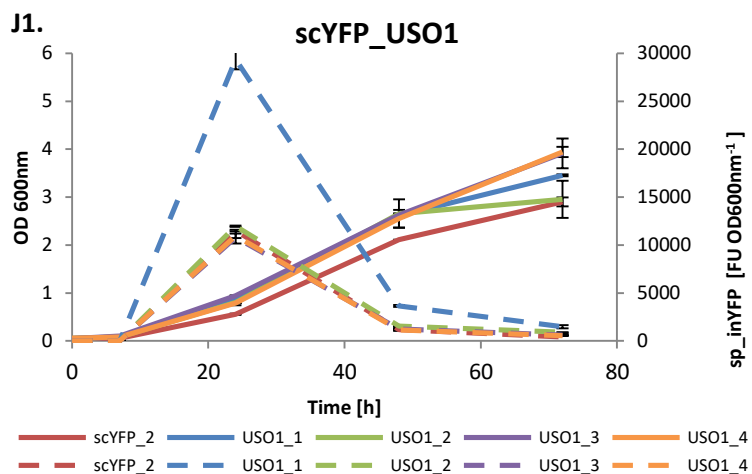
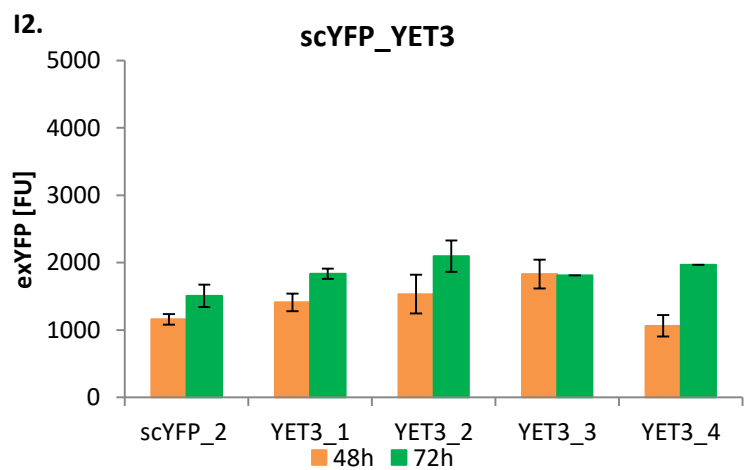
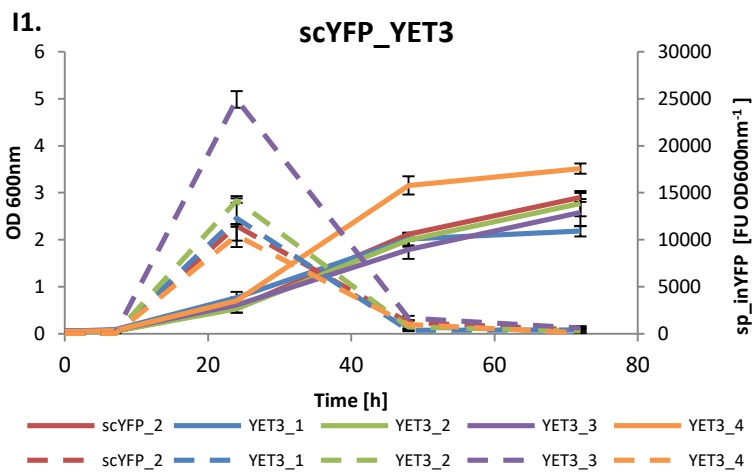


Fig.S2







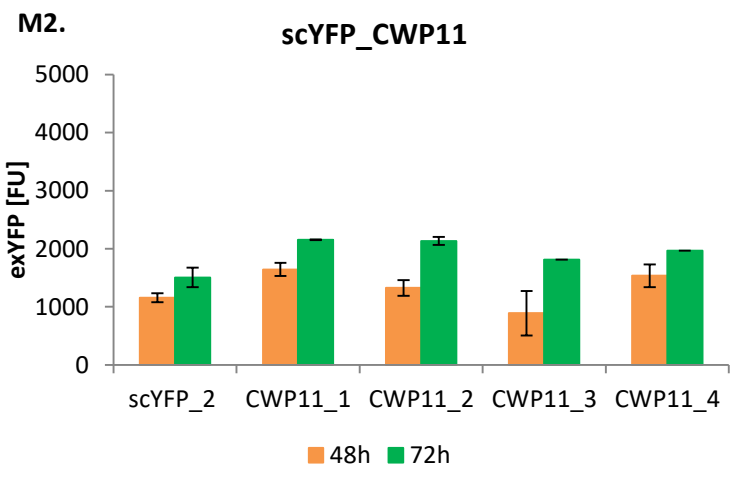
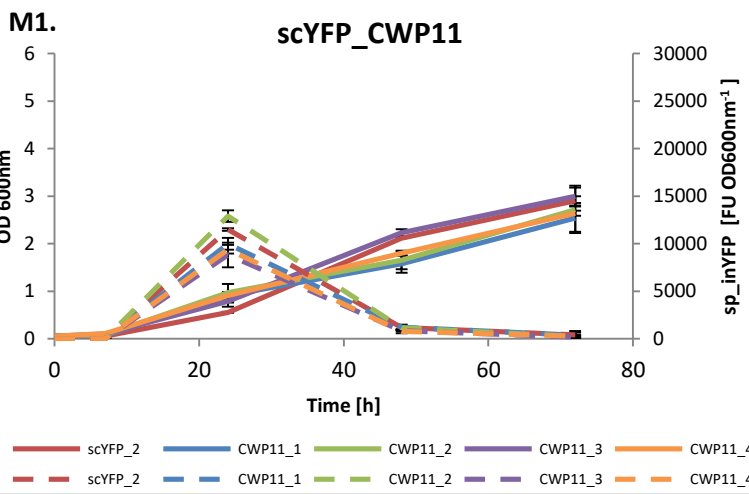
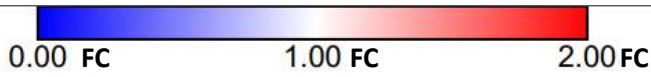


Fig.S3



A.

48h 30 °C

id	Control	PDI	USO1	HAC1	SLS1	SSA8	SEC1	CWP11	RPL3	SSA5	SSO1	CNE1	YET3
Control	1.00	1.38	1.21	0.89	1.17	0.83	0.97	0.77	0.96	0.57	0.83	0.98	0.84
PDI	0.72	1.00	0.87	0.65	0.85	0.60	0.70	0.56	0.69	0.41	0.60	0.71	0.61
USO1	0.83	1.14	1.00	0.74	0.97	0.69	0.80	0.64	0.79	0.47	0.69	0.81	0.70
HAC1	1.12	1.55	1.35	1.00	1.31	0.93	1.09	0.87	1.07	0.63	0.93	1.10	0.94
SLS1	0.86	1.18	1.03	0.76	1.00	0.71	0.83	0.66	0.82	0.48	0.71	0.84	0.72
SSA8	1.20	1.66	1.45	1.07	1.41	1.00	1.17	0.93	1.15	0.68	1.00	1.18	1.01
SEC1	1.03	1.42	1.24	0.92	1.21	0.86	1.00	0.80	0.99	0.58	0.86	1.01	0.87
CWP11	1.29	1.79	1.56	1.15	1.51	1.08	1.25	1.00	1.24	0.73	1.08	1.27	1.09
RPL3	1.05	1.44	1.26	0.93	1.22	0.87	1.01	0.81	1.00	0.59	0.87	1.02	0.88
SSA5	1.77	2.44	2.13	1.58	2.07	1.47	1.72	1.37	1.69	1.00	1.47	1.73	1.49
SSO1	1.20	1.66	1.45	1.07	1.40	1.00	1.16	0.93	1.15	0.68	1.00	1.18	1.01
CNE1	1.02	1.41	1.23	0.91	1.19	0.85	0.99	0.79	0.98	0.58	0.85	1.00	0.86
YET3	1.19	1.64	1.43	1.06	1.39	0.99	1.15	0.92	1.14	0.67	0.99	1.16	1.00

B.

DENOMINATOR ↓

48h 25 °C

id	Control	PDI	USO1	HAC1	SLS1	SSA8	SEC1	CWP11	RPL3	SSA5	SSO1	CNE1	YET3
Control	1.00	0.71	0.76	0.80	0.89	0.53	0.63	0.78	0.56	0.81	0.52	0.66	0.79
PDI	1.41	1.00	1.07	1.13	1.25	0.74	0.88	1.10	0.79	1.14	0.73	0.92	1.11
USO1	1.32	0.94	1.00	1.06	1.18	0.70	0.83	1.04	0.74	1.07	0.69	0.87	1.04
HAC1	1.25	0.88	0.94	1.00	1.11	0.66	0.78	0.98	0.70	1.01	0.65	0.82	0.98
SLS1	1.12	0.80	0.85	0.90	1.00	0.59	0.70	0.88	0.63	0.91	0.58	0.74	0.88
SSA8	1.90	1.35	1.44	1.52	1.69	1.00	1.19	1.49	1.07	1.54	0.99	1.25	1.49
SEC1	1.60	1.13	1.21	1.28	1.42	0.84	1.00	1.25	0.90	1.30	0.83	1.05	1.26
CWP11	1.28	0.91	0.97	1.02	1.14	0.67	0.80	1.00	0.72	1.04	0.66	0.84	1.00
RPL3	1.78	1.26	1.35	1.43	1.58	0.94	1.11	1.39	1.00	1.44	0.93	1.17	1.40
SSA5	1.23	0.87	0.93	0.99	1.10	0.65	0.77	0.96	0.69	1.00	0.64	0.81	0.97
SSO1	1.92	1.36	1.45	1.54	1.71	1.01	1.20	1.51	1.08	1.56	1.00	1.26	1.51
CNE1	1.53	1.08	1.15	1.22	1.36	0.80	0.95	1.19	0.86	1.24	0.79	1.00	1.20
YET3	1.27	0.90	0.96	1.02	1.13	0.67	0.80	1.00	0.71	1.03	0.66	0.83	1.00

NUMERATOR ↓

C.

72h 30 °C

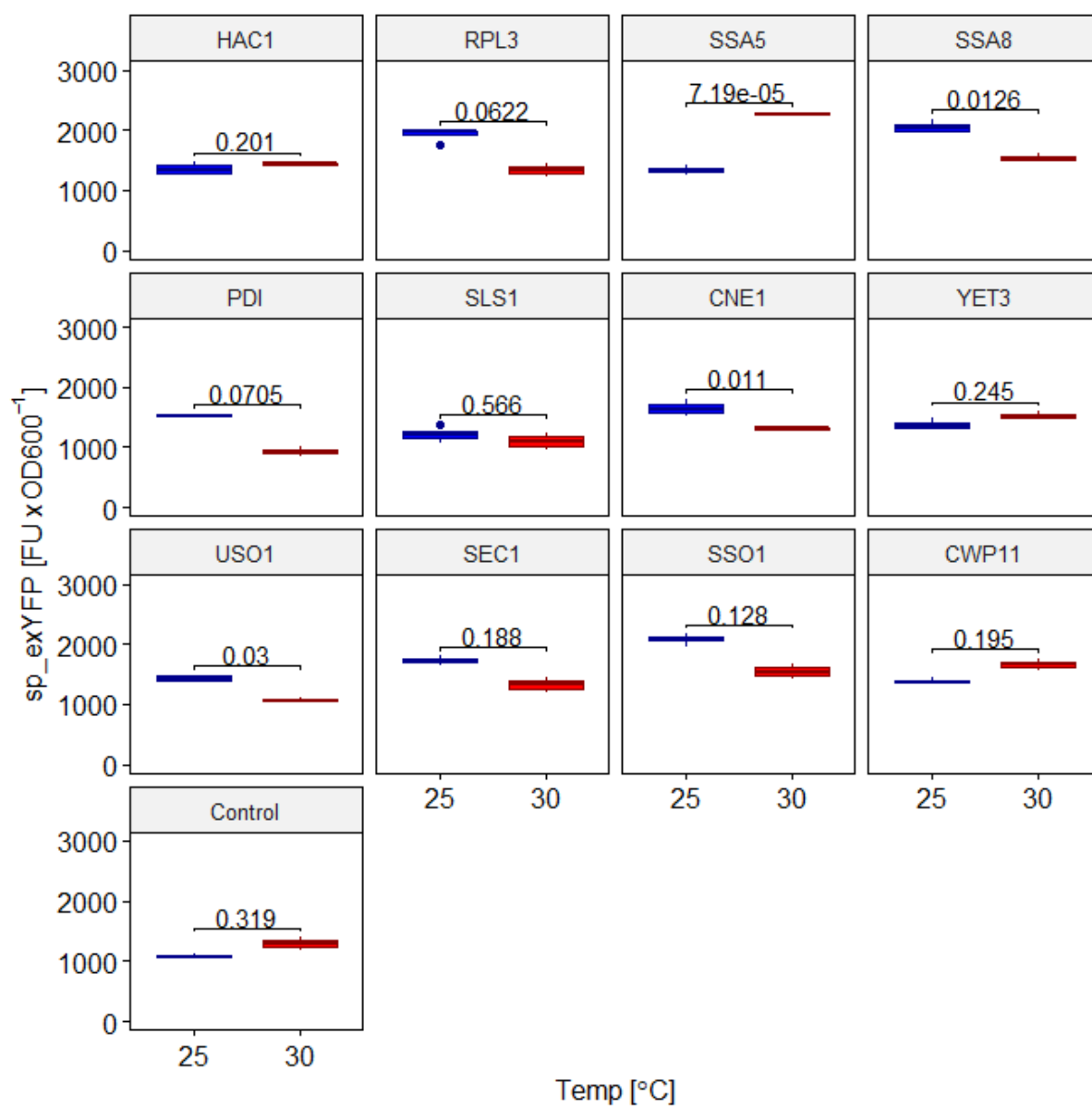
id	Control	PDI	USO1	HAC1	SLS1	SSA8	SEC1	CWP11	RPL3	SSA5	SSO1	CNE1	YET3
Control	1.00	1.02	1.54	0.73	1.15	0.68	0.91	0.79	0.68	0.77	0.60	0.88	1.00
PDI	0.98	1.00	1.51	0.72	1.12	0.67	0.89	0.78	0.66	0.76	0.59	0.86	0.98
USO1	0.65	0.66	1.00	0.48	0.74	0.44	0.59	0.51	0.44	0.50	0.39	0.57	0.65
HAC1	1.36	1.39	2.10	1.00	1.56	0.93	1.24	1.08	0.92	1.05	0.81	1.20	1.36
SLS1	0.87	0.89	1.34	0.64	1.00	0.59	0.79	0.69	0.59	0.67	0.52	0.77	0.87
SSA8	1.47	1.50	2.28	1.08	1.68	1.00	1.34	1.16	0.99	1.13	0.88	1.29	1.47
SEC1	1.10	1.12	1.69	0.81	1.26	0.75	1.00	0.87	0.74	0.85	0.66	0.97	1.10
CWP11	1.26	1.29	1.95	0.93	1.45	0.86	1.15	1.00	0.85	0.98	0.76	1.11	1.27
RPL3	1.48	1.51	2.28	1.00	1.70	1.01	1.35	1.17	1.00	1.14	0.80	1.30	1.48
SSA5	1.30	1.32	2.00	0.95	1.49	0.88	1.18	1.02	0.88	1.00	0.78	1.14	1.30
SSO1	1.67	1.70	2.57	1.23	1.91	1.14	1.52	1.32	1.13	1.29	1.00	1.47	1.67
CNE1	1.13	1.16	1.75	0.83	1.30	0.77	1.03	0.90	0.77	0.88	0.68	1.00	1.14
YET3	1.00	1.02	1.54	0.73	1.14	0.68	0.91	0.79	0.67	0.77	0.60	0.88	1.00

D.

72h 25 °C

id	Control	PDI	USO1	HAC1	SLS1	SSA8	SEC1	CWP11	RPL3	SSA5	SSO1	CNE1	YET3
Control	1.00	0.75	0.86	0.74	0.75	0.44	0.64	0.80	0.51	0.62	0.60	0.78	0.81
PDI	1.33	1.00	1.15	0.99	1.00	0.59	1.12	1.07	0.67	0.82	0.80	1.03	1.08
USO1	1.16	0.87	1.00	0.86	0.87	0.51	0.97	0.93	0.59	0.71	0.70	0.90	0.93
HAC1	1.34	1.01	1.16	1.00	1.01	0.59	1.13	1.08	0.68	0.83	0.81	1.04	1.09
SLS1	1.33	1.00	1.15	0.99	1.00	0.59	1.12	1.07	0.67	0.82	0.80	1.03	1.08
SSA8	2.27	1.70	1.96	1.69	1.70	1.00	1.91	1.82	1.15	1.40	1.37	1.76	1.83
SEC1	1.19	0.89	1.03	0.88	0.89	0.52	1.00	0.96	0.60	0.73	0.72	0.92	0.96
CWP11	1.24	0.93	1.07	0.92	0.93	0.55	1.05	1.00	0.63	0.77	0.75	0.96	1.00
RPL3	1.98	1.48	1.71	1.47	1.48	0.87	1.66	1.59	1.00	1.22	1.19	1.53	1.60
SSA5	1.62	1.22	1.40	1.21	1.22	0.72	1.36	1.30	0.82	1.00	0.98	1.26	1.31
SSO1	1.68	1.24	1.43	1.23	1.24	0.73	1.40	1.33	0.84	1.02	1.00	1.29	1.34
CNE1	1.29	0.97	1.12	0.96	0.97	0.57	1.09	1.04	0.65	0.80	0.78	1.00	1.04
YET3	1.24	0.93	1.07	0.92	0.93	0.55	1.04	1.00	0.63	0.76	0.75	0.96	1.00

Fig.S4



Microorganism	Name	Relevant genotype	Source
<i>Escherichia coli</i>	<i>E. coli</i> JM109	<i>F'</i> (<i>traD36, proAB+, lacIq, Δ(lacZ)M15</i>), <i>endA1, recA1, hsdR17 (rk-,mk+), mcrA, supE44, λ-gyrA96, relA1, Δ(lacproAB), thi-1</i>	Sigma-Aldrich
	JM109_ <i>scYFP</i> (ECB71)	JMP62-pTEF- <i>scYFP</i> -URA3ex	This study
	JM109_ <i>HAC1</i> (ECB75)	JMP62-pTEF- <i>HAC1</i> -LEU2ex	
	JM109_ <i>RPL3</i> (ECB84)	JMP62-pTEF- <i>RPL3</i> -LEU2ex	
	JM109_ <i>SSA5</i> (ECB74)	JMP62-pTEF- <i>SSA5</i> -LEU2ex	
	JM109_ <i>SSA8</i> (ECB72)	JMP62-pTEF- <i>SSA8</i> -LEU2ex	
	JM109_ <i>PDI</i> (ECB70)	JMP62-pTEF- <i>PDI</i> -LEU2ex	
	JM109_ <i>SLS1</i> (ECB79)	JMP62-pTEF- <i>SLS1</i> -LEU2ex	
	JM109_ <i>CNE1*</i> (ECB89)	JMP62-pTEF- <i>CNE1*</i> -LEU2ex	
	JM109_ <i>YET3</i> (ECB80)	JMP62-pTEF- <i>YET3</i> -LEU2ex	
	JM109_ <i>USO1</i> (ECB82)	JMP62-pTEF- <i>USO1</i> -LEU2ex	
	JM109_ <i>SEC1</i> (ECB73)	JMP62-pTEF- <i>SEC1</i> -LEU2ex	
	JM109_ <i>SSO1</i> (ECB78)	JMP62-pTEF- <i>SSO1</i> -LEU2ex	
	JM109_ <i>CWP11</i> (ECB85)	JMP62-pTEF- <i>CWP11</i> -LEU2ex	
<i>Yarrowia lipolytica</i>	<i>Y. lipolytica</i> Po1h	<i>MatA, ura3-302, xpr2-322, axp1-2, leu2-270::LEU2</i>	Celińska et al. 2020
	Po1h_ <i>Ura3</i> (GGY231)	<i>MatA, ura3-302::URA3, xpr2-322, axp1-2, leu2-270::LEU2</i>	Korpys-Woźniak et al. 2020
	<i>Y. lipolytica</i> Po1f	<i>MatA, leu2-270, ura3-302, xpr2-322, axp-2</i>	ATCC MYA2613
	Po1f <i>scYFP</i> _empty_vector (ECY220_2)	<i>MatA, leu2-270, ura3-302, xpr2-322, axp-2, pTEF-LEU2ex, pTEF-scYFP-URA3ex</i>	This study
	Po1f <i>scYFP</i> _ <i>HAC1</i> (ECY212_3)	<i>MatA, leu2-270, ura3-302, xpr2-322, axp-2, pTEF-HAC1-LEU2ex, pTEF-scYFP-URA3ex</i>	
	Po1f <i>scYFP</i> _ <i>RPL3</i> (ECY225_2)	<i>MatA, leu2-270, ura3-302, xpr2-322, axp-2, pTEF-RPL3-LEU2ex, pTEF-scYFP-URA3ex</i>	
	Po1f <i>scYFP</i> _ <i>SSA5</i> (ECY215_4)	<i>MatA, leu2-270, ura3-302, xpr2-322, axp-2, pTEF-SSA5-LEU2ex, pTEF-scYFP-URA3ex</i>	
	Po1f <i>scYFP</i> _ <i>SSA8</i> (ECY226_3)	<i>MatA, leu2-270, ura3-302, xpr2-322, axp-2, pTEF-SSA8-LEU2ex, pTEF-scYFP-URA3ex</i>	
	Po1f <i>scYFP</i> _ <i>PDI</i> (ECY222_2)	<i>MatA, leu2-270, ura3-302, xpr2-322, axp-2, pTEF-PDI-LEU2ex, pTEF-scYFP-URA3ex</i>	
	Po1f <i>scYFP</i> _ <i>SLS1</i> (ECY224_1)	<i>MatA, leu2-270, ura3-302, xpr2-322, axp-2, pTEF-SLS1-LEU2ex, pTEF-scYFP-URA3ex</i>	
	Po1f <i>scYFP</i> _ <i>CNE1*</i> (ECY227_2)	<i>MatA, leu2-270, ura3-302, xpr2-322, axp-2, pTEF-CNE1*-LEU2ex, pTEF-scYFP-URA3ex</i>	
	Po1f <i>scYFP</i> _ <i>YET3</i> (ECY221_3)	<i>MatA, leu2-270, ura3-302, xpr2-322, axp-2, pTEF-YET3-LEU2ex, pTEF-scYFP-URA3ex</i>	
	Po1f <i>scYFP</i> _ <i>USO1</i> (ECY223_3)	<i>MatA, leu2-270, ura3-302, xpr2-322, axp-2, pTEF-USO1-LEU2ex, pTEF-scYFP-URA3ex</i>	
	Po1f <i>scYFP</i> _ <i>SEC1</i> (ECY218_3)	<i>MatA, leu2-270, ura3-302, xpr2-322, axp-2, pTEF-SEC1-LEU2ex, pTEF-scYFP-URA3ex</i>	
	Po1f <i>scYFP</i> _ <i>SSO1</i> (ECY214_4)	<i>MatA, leu2-270, ura3-302, xpr2-322, axp-2, pTEF-SSO1-LEU2ex, pTEF-scYFP-URA3ex</i>	
Po1f <i>scYFP</i> _ <i>CWP11</i> (ECY219_1)	<i>MatA, leu2-270, ura3-302, xpr2-322, axp-2, pTEF-CWP11-LEU2ex, pTEF-scYFP-URA3ex</i>		

* *YALIO_B13156g* sequence encoding CNE1 was modified by SYNGEN company (Wroclaw, Poland) to removed *Bam*HI restriction sites located within the sequence.

Primer	Sequence 5'→3'
SP1_BamHI_Fw	GCGCGGATCCATGAAGTTCACATTT
YFP_BamHI_Fw	GCGCGGATCCATGGTGAGCAAGGGC
YFP_AvrII_Rev	GCGCCCTAGGTTACTTGTACAGCTC
PDI_BamHI_Fw	GCGCGGATCCATGAAGTTCAGTCCCTCA
PDI_AvrII_Rev	GCGCCCTAGGTTAAAGCTCATCATCAATCT
CWP11_BamHI_Fw	GCGCGGATCCATGAAGTACTCCGCCGT
CWP11_AvrII_Rev	GCGCCCTAGGTTAGAGAAGAGCAGCGACGG
USO1_BamHI_Fw	GCGCGGATCCATGAGCAAAGCCGTGAAAAT
USO1_AvrII_Rev	GCGCCCTAGGTTACTTTCAAAGACGGAG
SSO1_BamHI_Fw	GCGCGGATCCATGTCCAACACTACAATTAC
SSO1_AvrII_Rev	GCGCCCTAGGCTAGTTGTTGTTGAGACCC
SLS1_BamHI_Fw	GCGCGGATCCATGAAGTTCAGCAAGACT
SLS1_AvrII_Rev	GCGCCCTAGGTTAAAGCTCATCACGAGCGTC
SSA5_BamHI_Fw	GCGCGGATCCATGCCTAACACATCTGTA
SSA5_AvrII_Rev	GCGCCCTAGGTTAGTCGACCTCCTCGAC
SSA8_BamHI_Fw	GCGCGGATCCATGTCTAAAGCCGTCGG
SSA8_AvrII_Rev	GCGCCCTAGGTTAGTCAACCTCCTCAAC
HAC1_BamHI_Fw	GCGCGGATCCATGTCTATCAAGCGAG
HAC1_AvrII_Rev	GCGCCCTAGGTTACTCCAATCCCCCAA
CNE1_BamHI_Fw	GCGCGGATCCATGCGACTTTCCAAA
CNE1_AvrII_Rev	GCGCCCTAGGCTAAGCCTTTGCTTCTTG
RPL3_BamHI_Fw	GCGCGGATCCATGTCCCACCGAAAGT
RPL3_AvrII_Rev	GCGCCCTAGGTTAAAGGTCCTTCTTGAG
YET3_BamHI_Fw	GCGCGGATCCATGACACTCTACTACA
YET3_AvrII_Rev	GCGCCCTAGGTTAGGCGGTCTTCTTGTC
SEC1_BamHI_Fw	GCGCGGATCCATGTCCGACAACCTTGCTT
SEC1_AvrII_Rev	GCGCCCTAGGCTACTTGAAGAATTTGCCG
r-t_YFP_Fw	CAACAGCCACAACGTCTATATCATG
r-t_YFP_Rev	ATGTTGTGGCGGATCTTGAAGT
r-t_ACT_Fw	CGAGCGAATGCACAAGGA
r-t_ACT_Rev	GAGCGGTGATCTTGACCTTGA
r-t_HAC1_Fw	CGTCCGAAAACAACGATCTACA
r-t_HAC1_Rev	AGTGTGTTCCATGAGGTGCAT
r-t_PDI_Fw	CGAGGCAATTGTCCAGTACTTG
r-t_PDI_Rev	GCTCTTCTCGTTGGCGAAT
r-t_SSA5_Fw	TGCTGAGAAGTACAAGGAGCAAGA
r-t_SSA5_Rev	GCGGTAGGACTCGAGACCATT
r-t_SSA8_Fw	GCTCCCTGTCTCTCGGTATC
r-t_SSA8_Rev	GGTAGGGATGGTGGAGTTTCG
r-t_SLS1_Fw	TGAAAGTGACCACCAGACCTTCT
r-t_SLS1_Rev	TGGCTGCCTGTCCATTGA
r-t_SSO1_Fw	CTCGTACACGCAGGACCAGAT
r-t_SSO1_Rev	GCGGTCTTGAGCGACTGT
r-t_SEC1_Fw	GGGCGAAGTCCGCTCAGT
r-t_SEC1_Rev	ACCACCGGCCACAAACAC
r-t_YET3_Fw	GCAGCAGAAGGATAAGGACATTG
r-t_YET3_Rev	CGCTGATTCGGTTGACTCAGT
r-t_USO1_Fw	GCTGCCGAAAAGTGGAGTGA
r-t_USO1_Rev	ATCCATAGCCTGGTTCACCTTC
r-t_RPL3_Fw	TGCACCGTTGTGCGAGTT
r-t_RPL3_Rev	GGTGAGCCTTCTTCTGCTTAAGAG
r-t_CWP11_Fw	CTGTCTACCCCAACTCCTCTCT
r-t_CWP11_Rev	TCGTGCGAGTCAGTGTGAGTAA
r-t_CNE1_Fw	GAAGATTCCTCAACCCCAACTTC
r-t_CNE1_Rev	GAAGCCAAGACCACCAATGG



Research Article

Molecular background of HAC1-driven improvement in the secretion of recombinant protein in *Yarrowia lipolytica* based on comparative transcriptomics

Paulina Korpys-Woźniak, Ewelina Celińska *

Department of Biotechnology and Food Microbiology, Poznan University of Life Sciences, ul. Wojska Polskiego 48, 60-637 Poznań, Poland

ARTICLE INFO

Keywords:

Secretion
Heterologous protein
Yeast
Endoplasmic reticulum (ER) stress
Recombinant protein production

ABSTRACT

While the unfolded protein response (UPR) and its major regulator – transcription factor Hac1 are well-conserved across *Eukarya*, species-specific variations are repeatedly reported. Here we investigated molecular mechanisms by which co-over-expression of *HAC1* improves secretion of a recombinant protein (r-Prot) in *Yarrowia lipolytica*, using comparative transcriptomics. Co-over-expression of *HAC1* caused an >2-fold increase in secreted r-Prot, but its intracellular levels were decreased. The unconventional splicing rate of the *HAC1* mRNA was counted through transcript sequencing. Multiple biological processes were affected in the *HAC1*-and-r-Prot co-over-expressing strain, including ribosome biogenesis, nuclear and mitochondrial events, cell cycle arrest, attenuation of gene expression by RNA polymerase III and II, as well as modulation of proteolysis and RNA metabolism; but whether the *HAC1* co-over-expression/induction was the actual causative agent for these changes, was not always clear. We settled that the expression of the “conventional” *HAC1* targets (*KAR2* and *PDI1*) is not affected by its over-expression.

1. Introduction

Research on yeast protein synthesis and secretion is relevant to basic and applied studies. Knowledge of the molecular background of this biological process provides direct indications for its improvement through molecular and bioprocessing approaches. In our previous studies, we investigated the global transcriptome of *Yarrowia lipolytica* upon the over-synthesis of several biochemically different recombinant (secretory) proteins (r(s)-Prots) [1]. Both general and protein type-specific molecular responses were noted. Knowledge gained from that study prompted the design of genetic engineering strategies enhancing the production of rs-Prots in this species [2]. Genes that were significantly deregulated due to r(s)-Prot over-synthesis were used as “secretory helpers” assisting the synthesis and secretion of rs-Prot. In both those previous studies, we paid particular attention to the expression pattern of a key gene involved in the modulation of protein synthesis and secretion, namely *HAC1* (*YALIOB12716g*; hereafter all genes names are abbreviated – without *YALIO* prefix) and its regulome.

HAC1 encodes a transcription factor mediating unfolded protein response (UPR), which is triggered by the accumulation of incorrectly folded polypeptides in the over-loaded endoplasmic reticulum (ER)

[3–7]. The biological sense of UPR is to restore cellular homeostasis by debottlenecking the synthesis-folding-secretion pathway. Hac1 mediates the deregulation of hundreds of genes involved in polypeptide formation, folding, and maturation, but also in lipid synthesis, membrane expansion, and many others [4,8]. It is thus one of the most frequently targeted genes to improve the production of rs-Prots in yeasts [4,7,9–15], recently reviewed in [16]. *HAC1* homologs were described in different yeast species, and details of their operation in *Saccharomyces cerevisiae* and several other species are well-known [4,5,7,9,17]. While the Hac1-mediated UPR pathway is phylogenetically conserved in *Eukarya*, some variation, even amongst different yeast species, occurs [4,10,18–20]. The variation relates to the structure of the *HAC1* transcript, its expression, and splicing pattern, as well as the sequence of DNA binding motif, and the Hac1-driven regulome [4,10,16,18]. Hence, it becomes obvious that for the successful engineering of rs-Prot production in a given yeast species, more insight into the specificities of the UPR operation is needed.

In our previous studies on rs-Prot synthesis and secretion in *Y. lipolytica*, we found that *HAC1* up-regulation was dependent on biochemical characteristics of r(s)-Prot synthesized, and the rate of its synthesis [1]. High-level over-synthesis of a small secretory fluorescent

* Corresponding author.

E-mail address: ewelina.celinska@up.poznan.pl (E. Celińska).

protein (sc-YFP) or highly-disulfide-bonded amylase did not lead to the induction of *HAC1* expression. In contrast, the formation of “difficult-to-synthesize” highly-glycosylated secretory glucoamylase (TIG) and intracellular YFP (in-YFP) triggered *HAC1* up-regulation and initiation of a spectrum of downstream effects. Notably, it was surprising to see that the intracellular r-Prot (in-YFP) over-synthesis triggered the up-regulation of *HAC1* and the following downstream events, which are typically associated with ER compartment. But the transcriptomics profiles induced by the rs-Prot (TIG) and r-Prot (in-YFP) over-expression were highly correlated. In addition, the amounts of synthesized r (s)-Prots were generally lower; which was associated with the induction of massive vacuolar proteolysis. Such an observation could incorrectly suggest that up-regulation of *HAC1* did not bring any improvement in the rs-Prots synthesis in *Y. lipolytica*. However, in the following study, we evidenced that co-over-expression of *HAC1* with a secretory YFP (*SCYFP* gene) [2] exerted a positive effect on the rs-Prot productivity, increasing it by 30%.

It was surprising to see that the *HAC1* up-regulation, either caused by the accumulation of incorrectly folded polypeptides [1] or by over-expression from a synthetic construct [2], did not induce expression of its known direct downstream targets, i.e. protein disulfide isomerase *PDI1* (*E03036g*) and a multifunctional chaperone *KAR2* (*E13706g*). Previous studies conducted with *Pichia pastoris* and *S. cerevisiae* [5,11,18,21] reported that up-regulation of *HAC1* was accompanied by enhanced expression of these downstream targets. Moreover, upregulation of *KAR2* gene expression was also seen previously in *Y. lipolytica* upon overexpression of a spliced variant of *HAC1*, as assayed by real-time qPCR analysis [19]. We previously speculated that in the case of *ylHAC1* co-over-expression, the outcome was a consequence of using a native, non-spliced variant of the *HAC1* gene [2], which accumulated without being fully processed. For being translated, *HAC1* requires unconventional splicing by dimerized protein Ire1 (*A14839g*), which is a sensor of unfolded polypeptides and a nuclease. Under balanced conditions, Ire1 is bound to Kar2 and is not available for the *HAC1* transcript processing. Accumulation of incorrectly folded polypeptides causes Kar2 dissociation from its interaction partner and releases Ire1 for mRNA splicing. The key role of nucleotide exchange factor Sls1 (*E32703g*) in the Kar2-Ire1 interaction in *Y. lipolytica* was evidenced [22–24] and later observed in the context of the global transcriptome [1]. Nonetheless, the lack of transcriptional response from *KAR2* and *PDI1* was observed in both cases – (i) when the native *HAC1* gene was synthetically over-expressed and thus could accumulate due to insufficient processing, and (ii) when it was natively up-regulated and presumably fully processed due to complete onset of UPR. Hence, the question of the specificity of *HAC1* operation in *Y. lipolytica* remained open.

This study aimed to further characterize the secretory pathway in *Y. lipolytica* with a particular focus on *HAC1* operation. To this end, we conducted comparative transcriptomics analysis of steady-state-maintained *Y. lipolytica* cells: (i) over-expressing native *HAC1* and sc-YFP reporter protein from synthetic constructs, (ii) over-synthesizing sc-YFP at high levels, and (iii) subjected to (cytosolic) UPR and up-regulation of *HAC1* triggered by over-synthesis of intracellular reporter protein (in-YFP). Using such an approach we were able to dissect the effects triggered by different conditions (*HAC1* / high-level synthesis of rs-Prot) and evaluate their practical consequences in terms of rs-Prot production.

2. Materials and methods

2.1. Yeast strains

All the strains analyzed in this study were constructed previously [1, 2,25]. The strains are listed in Supplementary Table S1. *Y. lipolytica HAC1_scYFP* (ECY212.3) strain was constructed on a platform of Po1f strain (*MatA*, *leu2-270*, *ura3-302*, *xpr2-322*, *axp-2*) by

co-over-expressing: (i) a secretory yellow fluorescence protein (*SCYFP*), and (ii) a native bZIP transcription factor *HAC1* (*B12716g*) (Supplementary Table S2). The genes were cloned in JMP62-URA3ex and JMP62-LEU2ex vectors (Supplementary Fig. S1), respectively, according to a previously described methodology [26]. *Y. lipolytica* strains GGY251 and GGY249 are Po1f-derivates co-over-expressing a YFP reporter in secreted (sc-YFP) and intracellular form (in-YFP), respectively [25]. For secretion, the YFP-encoding gene was transcriptionally fused with a signal sequence SP1, native for Exo-1,3-beta-glucanase (*B03564g*) [27, 28]. The reference strain, GGY231, is a prototrophic version of Po1f (not over-expressing YFP). The *scYFP*, *inYFP*, and the reference strain were constructed via the Golden Gate strategy [29–31]. All the strains were maintained in glycerol stocks at -80 °C, and revived by plating on YPD agar ([g L⁻¹]: yeast extract (BTL, Poland), 10; bactopectone (BTL), 20; glucose (POCh, Poland), 20; agar (BTL, Poland), 15) and incubation at 30 °C. Visualization of the strains' growth on agar plates and microscopic images are provided in Supplementary Fig. S2.

2.2. Chemostat cultivation

Data on growth, r(s)-Prot synthesis, and carbon utilization by *scYFP* and *inYFP* strains were obtained in the previous study [1]. Bioreactor cultivations of the *HAC1_scYFP* strain were executed according to a previously described methodology [1,25]. Briefly, pre-cultures were developed in shake flasks in 30 mL YPG₂₀ medium ([g L⁻¹]: yeast extract, 10; bactopectone, 20; glycerol (POCh, Poland), 20), at 28 °C, agitation at 250 rpm for 22 h. Continuous cultures were conducted in 0.5 L stirred tank bioreactors Multifors 2 (Infors HT, Bottmingen Basel, Switzerland), in 0.3 L of YPG₁₀₀ medium ([g L⁻¹]: yeast extract, 10; bactopectone, 20; glycerol, 100), inoculated at 10%. The conditions were monitored and controlled at the following set points: temperature, 28 °C; pH, 5.5 (correction by: 20% NaOH, 10% H₂SO₄); airflow, 2.0 vvm; pO₂ was kept at 20% by setting a cascade with stirring (100 to 1200 rpm). Foaming was minimized by the automatic addition of a defoaming agent AntiFoam 204 (Merck/Sigma-Aldrich; USA). The initial batch mode culture was continued until full consumption of glycerol. Feeding was initiated at a dilution rate [D] of 0.20 [h⁻¹]. A steady state was reached after at least 6 residence times [τ]. The cultures were conducted in biological duplicates.

2.3. Biomass concentration measurements

Yeast biomass concentration was analyzed by measuring optical density at 600 nm (OD 600 nm). Prior to the measurement the biomass was washed twice and diluted in a sterile saline solution (0.85% NaCl). OD600 measurements were done using flat-bottomed 96-well microplates (Corning; Sigma-Aldrich, USA) in Tecan Infinite M200 automatic plate reader (TECAN, Austria). All the measurements were conducted in technical duplicates of each biological replicate.

2.4. Fluorometry

Fluorescence (FL) of intracellular (in-YFP) and secreted (sc-YFP) fraction of the YFP reporter protein in the *HAC1_scYFP* cells was analyzed in flat-bottomed 96-well microplates (Corning; Sigma-Aldrich) in Tecan Infinite M200 automatic plate reader at the wavelengths settings (excitation/emission) 495/527 nm. The FL of sc-YFP was measured in 200 µl of biomass-devoid supernatant. FL of in-YFP was measured in 200 µl of pre-washed culture biomass, re-suspended, and diluted in the saline solution. FL measurements were normalized vs background fluorescence: either biomass of the control strain for in-YFP or YPG medium for sc-YFP. FL results were expressed as: (i) raw FL readout of the washed biomass in FL units [RFU – Relative Fluorescence Unit], (ii) sp_in-YFP - specific FL of the washed biomass in [RFU OD600 nm⁻¹], (iii) sc-YFP – raw FL readout of the supernatant in [RFU], (iv) sp_sc-YFP - specific FL of the supernatant in [RFU OD600 nm⁻¹]. All the

measurements were conducted in technical duplicates of each biological replicate. Additionally, the FL results were randomly verified through fluorescence microscopy observations (Zeiss, Axiovert 200).

2.5. High-performance liquid chromatography

The concentration of glycerol and typical products of *Y. lipolytica* metabolism (mannitol, erythritol, citric acid, and α -ketoglutaric acid) was measured by High-Performance Liquid Chromatography (HPLC), as described previously [25]. Briefly, the analysis was executed using Agilent Technologies 1200 series (Agilent Technologies, Santa Clara, USA), equipped with a refractive-index detector (G1362A) and a Rezex ROA-Organic Acid H+ column (Phenomenex, Torrance, USA), and 0.005 N H₂SO₄ as eluent at a flow rate of 0.6 [mL min⁻¹], at 40 °C.

2.6. Protein isolation and Protease activity assay

2.6.1. Total protein isolation and BCA assay

Total protein extraction was conducted according to the previously described protocol [32]. Briefly, cellular pellets were resuspended in ice-cold breaking buffer (0.1 M sodium phosphate buffer pH 7.4, 5 μ M DTT, 5% GLY). The cells were disrupted by glass beads (Sigma-Aldrich, St. Louis, USA) under repeated cycles (5x) of mixing at 30 strokes s⁻¹ for 30 s in a MixerMill MM400 (Retsch GmbH, Haan, Germany) and incubation on ice for 1 min. The cellular debris was then separated by centrifugation (24,652 xg, 4 °C, 10 min; Eppendorf 5430 R; Eppendorf, Hamburg, Germany). The cellular extracts were subjected to buffer exchange by ultrafiltration through a 3 kDa MWCO point cellulose membrane (Amicon Ultra; Merck, Darmstadt, Germany).

Total protein concentration in the obtained extracts was determined using a commercial BCA (Bicinchoninic Acid) Protein Assay Kit (Merck) in a microtiter plate format. A standard curve was prepared using a BSA standard solution. The reaction was conducted according to the manufacturer's protocol. Reading of the assay colorimetric was conducted in transparent flat-bottomed 96-well microplates (Corning; Sigma-Aldrich) using a Multiskan SkyHigh Plate Reader - UV/Vis Spectrophotometer at 562 nm wavelength (ThermoFisher Scientific, Waltham, USA).

The assay was conducted in two independent runs and a technical triplicate.

2.6.2. Protease activity assay with FTC-casein

Proteolytic activity in the total protein extracts was examined using a commercial Pierce™ Fluorescent Protease Assay Kit (ThermoFisher Scientific), relying on FTC-casein substrate digestion and its fluorescence. The protein extracts were subjected to the reaction in a microtiter plate format. A standard curve was prepared using a standardized solution of trypsin. Fluorescence reading was conducted in white opaque plates, as recommended by the kit's manufacturer, using a Tecan Infinite M200 automatic plate reader at the wavelength settings (excitation/emission) 485/538 nm.

The assay was conducted in two independent runs and a technical triplicate.

2.6.3. Statistical analysis

Initial data processing and their graphical representation were done using Microsoft Excel and Data Analysis plugin. The statistical importance of the differences between compared sets of data was analyzed using two-way analysis of variance (ANOVA) and Tukey's HSD multiple comparison tests using the STATISTICA software (StatSoft, Inc., Tulsa, USA). Datasets were initially tested by Kruskal-Wallis and Levene's tests to verify ANOVA's presumptions. The results were considered statistically significant at a *p*-value < 0.05. The results were expressed as mean \pm standard deviation (\pm SD) of the replicates.

2.7. RNA extraction and NGS sequencing

Steady-state-maintained yeast cells were harvested and processed according to a Bead-Beat Total RNA Mini Kit protocol (A&A Biotechnology, Gdynia, Poland) to extract total RNA. The quality of RNA was verified through agarose gel electrophoresis [33] and spectrophotometry (NanoDrop, Thermo Fisher Scientific). Library construction and NGS RNA sequencing were executed using NovaSeq 6000 (Illumina) apparatus at Genomed S.A. (Warsaw, Poland). Raw sequencing reads are available on the NCBI Sequence Read Archive under the BioProject number PRJNA869113 (see Data Availability section).

2.7.1. RNAseq data analysis

2.7.1.1. Raw data processing and data filtration. Cutadapt [34] was used for NGS adapters removal. Initial filtering was done using the *q25* quality parameter and *m15* the minimum length of the reads, followed by re-filtering to remove reads < 20 bp. Quality control was conducted using FastQC (2010; www.bioinformatics.babraham.ac.uk). Post-QC data were aligned and mapped to the *Y. lipolytica* CLIB122 reference genome (GenBank assembly identifier: GCA_000002525.2) using Hisat2 [35]. Reading pairs were counted using the HTseq program [36]. Primary Gene Ontology (GO) annotation was conducted using the reference genome template (GCF_000002525.2; manually curated). The final results were processed in R (www.rproject.org) using the DESeq2 package [37]. All Differentially Expressed Genes (DEGs) met the requirement of FDR \leq 0.05 (adjusted *p*-value) tested via the Wald test. Only DEGs meeting all those criteria were considered in further analyses. Gene name and function annotation was manually curated by cross-referencing several databases – Panther, UniProt, GRYC, and NCBI and blasting (blastp) the sequences against the database's collections. The most relevant matches were chosen to finally describe a given gene. A complete list of all DEGs is given in Supplementary Table S3. Relative gene expression level is expressed as fold change (FC) over the control prototrophic strain.

2.7.1.2. Enrichment / overrepresentation tests. Statistical overrepresentation test and Enrichment Test were conducted using PANTHER v14.0 online tool and database [38,39]. The initial GO terms annotation by DESeq2 and the reference genome was followed by manual curation using PANTHER resources and tools. The biological function was assigned to each DEG according to PANTHER's GO Ontology database DOI: 10.5281/zenodo.4437524 Released 2021- 01–01. Overrepresentation of specific biological processes, molecular functions, and cellular components in a subset of DEGs was tested using GO complete annotation data sets, Fisher's exact test, and False Discovery Rate for multiple testing. Only results with *p* < 0.05 were considered significant and displayed.

2.7.1.3. Comparative analysis through Venn diagrams. Prior to the Venn diagram analysis, DEGs predicted by DESeq2 were filtered for log₂-fold change > 1.0; only such DEGs were considered in this analysis. Similarities and dissimilarities in DEGs subsets were searched using InteractiVenn online tool; graphics were adjusted using Inkscape (https://inkscape.org/release/inkscape-1.0.2/).

2.7.1.4. Comparative analysis – heat-map generation. Heat-maps of DEGs in the studied strains were prepared using Morpheus online tool (https://software.broadinstitute.org/morpheus).

2.7.1.5. Quantitation of spliced-to-unspliced HAC1 transcript ratio. As the presence of an intron site was not described in the reference genome used, the intron existence and location within the *B12716g* locus had to be identified de novo. The detection and visualization of the potential splice sites were conducted using the SGSeq package [40] in the R

programming language. Mapping of exons and introns in the *B12716g* gene was conducted in the IGV program [41]. Coverage depth within the gene fragments (intron and exons) was used as an indicator of actual splicing events. Coverage was calculated in fragments: intron (NC_006068.1:1702623–1702651), exons (NC_006068.1:1702612–1702622, NC_006068.1:1702652–1702662). The average coverage depth was calculated by SamTools [42].

2.8. Basic bioinformatics

DNA and protein sequences were retrieved from NCBI's GenBank, Nucleotide or Protein databases (<https://www.ncbi.nlm.nih.gov/>), UniProt (<https://www.uniprot.org/>), or GRYC (<http://gryc.inra.fr/>). Sequence analysis and primer design were done in Benchling (<https://www.benchling.com/>). Multiple sequence alignments were done using BioEdit [43] or MultAlin [44].

3. Results and discussion

The principal aim of this study was to investigate in more detail the *HAC1* gene operation in *Y. lipolytica* and its effect on r(s)-Prot synthesis, as well as to verify the previous unexpected observations on its direct (co-)regulome. To this end, three *Y. lipolytica* strains over-synthesizing YFP reporter protein were investigated: (i) strain *scYFP* expressing *HAC1* at the level of control strain, despite high-level over-synthesis of the secretory reporter protein [1], (ii) strain *inYFP* over-synthesizing intracellular YFP, with elevated *HAC1* expression level by 1.23-fold over the control (assayed by RNAseq) [1], (iii) strain *HAC1_scYFP* with synthetically induced *HAC1* expression by more than 6-fold over the prototrophic control and ~30% higher synthesis of YFP over the control, seen previously in batch cultures [2]. All the strains over-synthesized YFP reporter, to enable assessment of practical consequences of the *HAC1* overexpression on rs-Prot formation. We used the same model protein in all the strains, to eliminate protein-specific effects, seen previously [25]. Transcriptomes of *inYFP* and *scYFP* strains were obtained and analyzed previously [1]. Here they are used as references, enabling a more thorough characterization of the *HAC1* transcriptome.

Noteworthy, the up-regulation of *HAC1* expression observed in the *inYFP* strain [1] cannot result from a canonical UPR induction, since in-YFP protein is not targeted for secretion and hence does not enter ER compartment. On the other hand, the *HAC1* up-regulation in the *inYFP* strain was not an isolated phenomenon caused by in-YFP over-synthesis but was indeed accompanied by a number of other gene deregulations seen also in the other strain, over-synthesizing secretory glucoamylase TIG [1], facing actual UPR. Their transcriptomes shared a high level of similarity, suggesting a similar biological response to the r(s)-Prot over-synthesis. It is possible that this *HAC1* upregulation in the *inYFP* strain results from an indirect action of “cytosolic UPR” stress [13] governed by Hsf1, or some other cross-talk mechanism. Such a tight cross-talk between “cytosolic-” and ER-UPR is well described in plant species [45], for which the implication of bZIP60 and bZIP28/17 transcription factors (40% similarity of the former to Hac1, based on blastp) was studied in detail. The here elicited transcriptome of the *HAC1_scYFP* strain shares similarities with both *S. cerevisiae*'s transcriptomes: induced by the “ER-specific responder” Hac1 [46], and by the “cytosol-specific responder” Hsf1 [13]. For example, genes involved in fatty acid metabolism (*MGA2/D24101g* FC: 1.18), heme biosynthesis (*HEM12/C01716g* FC: 0.61, *HEM13/E29975g* FC: 0.78, *HEM15/F19470g* FC: 0.33), cell wall biogenesis (*EXG2/F05390g* FC: 1.29, *PKC1/F09746g* FC: 0.85), or phospholipid biosynthesis (*OPI3/E12441g* FC: 1.27) share the same expression pattern with *S. cerevisiae*'s *HAC1*-induced regulome [46]. On the other hand, the other genes deregulated in the *HAC1_scYFP* strain, playing role in ubiquitination and proteolysis (*UBI4/F18403g* FC: 1.27, *PIB1/D15510g* FC: 0.78), carbohydrate metabolism (*UGP1/A02310g* FC: 1.16), signal

transduction (*YDR247W/B08558g* FC: 0.83), energy generation (*PGK1/D12400g* FC: 1.19), or chaperones (*SSA8 (SSA4)/D22352g* FC: 2.17, *HSP12/D20526g* FC: 1.41), were previously identified as *HSF1*-specific [13]. Such a strong overlap of here observed transcriptome of the *HAC1_scYFP* strain indicates that indeed the transcriptome of the *inYFP* strain could develop based on the accumulation of the in-YFP in the cytoplasm, which prompted *HSF1* up-regulation and concomitant *HAC1* induction (by an unknown mechanism) followed by activation of its downstream specific targets. What supports this hypothesis is that the canonical UPR transcriptome of the TIG strain shares the majority of the abovementioned marker genes with *HAC1*-specific transcriptome, and only one gene with the *HSF1*-specific transcriptome from *S. cerevisiae* (assayed previously [1], analyzed here).

3.1. *HAC1* prevents excessive accumulation of r(s)-Prots

The strains analyzed in this study were cultured in continuous mode, and at the moment of transcriptome profiling, they were maintained in steady-state (for *HAC1_scYFP*: Supplementary Fig. S3; for *inYFP* and *scYFP* strains: [1]). The reporter YFP protein was synthesized at the levels presented in Table 1. Interestingly, co-over-expression of *HAC1* caused a nearly 7-fold drop in the retained fraction of YFP but promoted its secretion, suggesting that the multidirectional activity of Hac1 was “used” for both decreasing YFP synthesis and enhancing its secretion. The level of extracellular FL generated by *HAC1_scYFP* was nearly 2.5-fold higher over the *scYFP* strain in terms of total amounts [kRFU L^{-1}]. While the overall amounts of rs-Prots upon *HAC1* co-overexpression were lower than in the more efficient strain, *scYFP*, the latter encountered an undesired phenomenon of high YFP retention within the cell.

3.2. *HAC1*'s unconventional splicing

While high-level over-expression of *HAC1* in the *HAC1_scYFP* strain was confirmed earlier [2], here we quantified the fraction of transcripts that were spliced to generate an active form of Hac1 [4,18–20,47]. First, it was necessary to identify and map the putative splicing site *de novo*, as it turned out that the *HAC1* gene from the reference genome (GCF_000002525.2) lacked this feature; even though the sequence of the intron was previously described [19,20]. The splicing sites were identified *de novo* based on the transcripts (cDNA) sequencing results and their coverage. The 29 bp fragment with the decreased coverage level visualized in Supplementary Fig. S4, corresponds to the previously described intron ([19,20]; *cagtgatgactgtcgcaactactgaccag*). This result provides evidence from the RNA level, that the previously identified “most likely” splicing sites [20] are operable.

Despite the high over-representation of *HAC1* transcripts (Fig. 1.A), the counts of a spliced form were comparable to the prototrophic control level, and the UPR-less *scYFP* strain (Fig. 1.B). It is commonly accepted that for effective operation, *HAC1* transcript requires unconventional splicing, to remove the translation-inhibiting short intron sequence [4,18,47]. The results presented in (Fig. 1.A.B) suggest that it is another mechanism, rather than accumulation of the substrate (i.e. *HAC1* mRNA), that promotes the formation of an effective = spliced *HAC1* transcript in *Y. lipolytica*. As clearly demonstrated previously, accumulation of *XBPI* mRNA (a metazoan homolog of yeast *HAC1* [48]; for sequence analysis see Supplementary Fig. S5) was a prerequisite for the following Ire1-mediated splicing reaction, to overcome the degradation of protein and to produce significant amounts of the spliced form of *XBPI* [49]. Based on the here observed patterns of the intron-less transcript accumulation (Fig. 1.B) and distribution of the YFP protein (Table 1), it can be concluded that the high accumulation of YFP inside the cell (the highest in *scYFP* strain) does not promote the rate of the *HAC1* splicing (the lowest in the *scYFP* and *HAC1_scYFP* strains). The “easy-to-fold” YFP did not elicit the necessary signal of the hydrophobic domain exposure, even when accumulated at a high level.

Table 1
Amounts of YFP synthesized by *Y. lipolytica* strains under steady-state conditions.

Fraction	<i>HAC1</i> expression level by RNAseq [fold over control]	biomass in-YFP [RRFU L ⁻¹]	supernatant sc-YFP [RRFU L ⁻¹]	biomass sp.in-YFP [RRFU OD ₆₀₀ ⁻¹]	supernatant sp.sc-YFP [RRFU OD ₆₀₀ ⁻¹]	Reference
<i>inYFP</i>	1.23	130 809.60 ± 5 201.20		42 726.32 ± 3635.71		
<i>scYFP</i>	1.0	812 108.30 ± 5 715.62	24 785.28 ± 187.33	257 051.30 ± 23 576.03	7 899.84 ± 1 260.26	[1]
<i>HAC1_scYFP</i>	21.0	118 922.90 ± 3 224.39	60 123.13 ± 847.64	18 736.20 ± 923.66	9 415.35 ± 843.21	This study

The previous detailed biochemical studies [22] evidenced the key role of the Sls1 nucleotide exchange factor in regulating the interaction between Ire1 and Kar2 in *Y. lipolytica*; that ultimately leads to *HAC1* splicing. As evidenced previously at the protein level, the intracellular levels of Sls1 are induced by the accumulation of unfolded proteins [23, 24,50]. In the present study, the *SLS1* transcript levels in the *HAC1_scYFP* and *inYFP* strains were not significantly different from the control, but in the UPR-less strain (*scYFP*) its expression was decreased (see Fig. 2), which is consistent with the results obtained on the protein level [22]. On the other hand, previous studies showed that *SLS1* co-overexpression caused a significant enhancement in the rs-Prot synthesis under permissive temperature [2]. The level of improvement was comparable to the one triggered by co-overexpression of *PDII*, *CNE1*, and *HAC1* [2]. The key question arose from results depicted in Fig. 1.A.B is about the mechanism underlying the observed improvement in the rs-YFP secretion, operating despite the comparable amounts of the intron-less, effective *HAC1* transcript in the *HAC1_scYFP* and the control strains. One possible explanation is that, as evidenced for *XBPI* [49], both forms – spliced and unspliced *HAC1* can modulate cellular UPR activity in yeast via direct binding to the complementary cis-elements; but the spliced variant is much more efficient. At this moment, it is only speculation that requires further in-depth studies.

3.3. Genes highly responsive to *HAC1* co-over-expression and *r(s)*-Prot over-synthesis – dissection of the effects

Co-over-expression of *SCYFP* and *HAC1* genes contributed to significant up-/ down-regulation of 316 / 237 genes over the prototrophic control strain (Supplementary Fig. S6; Supplementary Table S3). Amongst the most up-regulated DEGs (fold change > 2) we identified *i. a. F15411g* encoding Zn-Fe transporter (up-regulated by > 18-fold), other genes involved in transmembrane transportation of Fe (*TAF1 D05401g*, *D26081g*, *FRE2 B10846g*), phosphate (*PHO89 E23859g*), MFS-type transporters (*C00825g*), Zn-regulated transporter (*F21659g*) and other membrane proteins putatively involved in transmembrane transportation (*A07623g*, *E12881g*). In addition, two ribonucleases (*RNH1 C09658g*, *E12331g*) and reverse transcriptase domain-bearing protein (*C19888g*), were identified amongst the most up-regulated genes upon *HAC1+SCYFP* co-over-expression, highlighting intensive RNA degradation. Strikingly, all these genes were also found to be up-regulated in the *inYFP* strain with natively induced *HAC1* expression (Fig. 3.A). However, *F15411g*, *TAF1*, *D26081g*, *FRE2*, *PHO89*, *C00825g*, and *E12881g* were all previously identified in the “up-regulome” of the high-level synthesis strains (*scYFP* and *SoA*), having a basic level of *HAC1* expression [1], but decreased *SLS1* expression. On the other hand, transporters *F21659g*, *A07623g*, as well as RNA metabolism-involved genes *RNH1*, *E12331g*, and *C19888g* were up-regulated solely when *HAC1* expression level was enhanced / *SLS1* transcription was not inhibited (Fig. 1.A; Fig. 2).

Amongst the genes up-regulated in the *HAC1_scYFP* strain, but to a lesser extent, we found aquaporin (*F01210g*), biofilm regulator (*E31757g*), phosphatase *PHO2* (*C19866g*), DeoR-type transcription factor (*E19030g*), Dauer up regulated-related putative ER-resident thioltransferase (*F22187g*), alternative oxidase *AOX1* (*D09933g*) and catalase *CTT1* (*E34749g*) (Fig. 3.A). All these genes were previously found up-regulated in response to high-level over-synthesis of rs-Prot in at least one of the analyzed strains (*scYFP* or *SoA*) [1]; so their up-regulation in the *HAC1_scYFP* strain results rather from *sc-YFP* over-synthesis than *HAC1* up-regulation (to be discussed hereafter).

Within the most down-regulated genes in the *HAC1_scYFP* strain, we identified a transcription factor *D01353g* of an unknown specific role, followed by reverse-transcriptase *A07238g* (showing the opposite pattern of expression to *C19888g* having the same function), and *URC5* involved in uracil catabolism (*E35156g*) (Supplementary Table S3). The same transcription factor *D01353g* was found amongst the most down-regulated DEGs in both the *Y. lipolytica* strains over-synthesizing rs-Prot

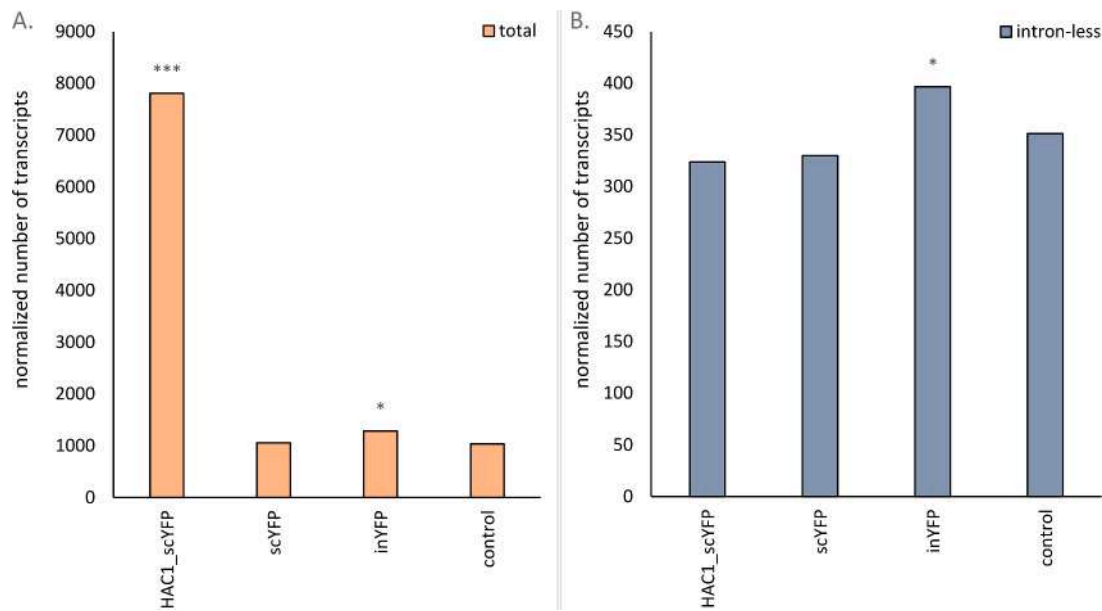


Fig. 1. Quantitation of *HAC1* transcripts total number and number of transcripts devoid of an intron in the analyzed strains' transcriptomes. The *HAC1* transcripts sequenced and quantified by RNAseq analysis were mapped for the occurrence of exons and introns. Total denotes a normalized number of all transcripts based on exon counts. Intron-less denotes a normalized number of transcripts that were devoid of an intron. *, *** signifies statistical significance at $p < 0.05$ and 0.001 .

(sc-YFP and SoA), studied previously [1]. A deeper search for the *D01353g* function revealed that it bears some similarities with the repressor of RNA polymerase II. So it seems that its down-regulation releases RNA polymerase II from its repressive activity. Inhibition of uracil catabolism was observed in both the *HAC1*-over-expressing strains and in the *scYFP* strain (Fig. 3.B), as highlighted by the

down-regulation of multiple DEGs in the latter [1].

Transcriptomes of *HAC1_scYFP* and *inYFP* strains shared down-regulation of RNA endonuclease (*A12683g*) that specifically degrades the RNA of RNA-DNA hybrids (Fig. 3.B), Fe-O2 oxidoreductase (*D07282g*), and nicotinic acid plasma membrane permease (*F21109g*). These genes were significantly down-regulated solely in the strains with

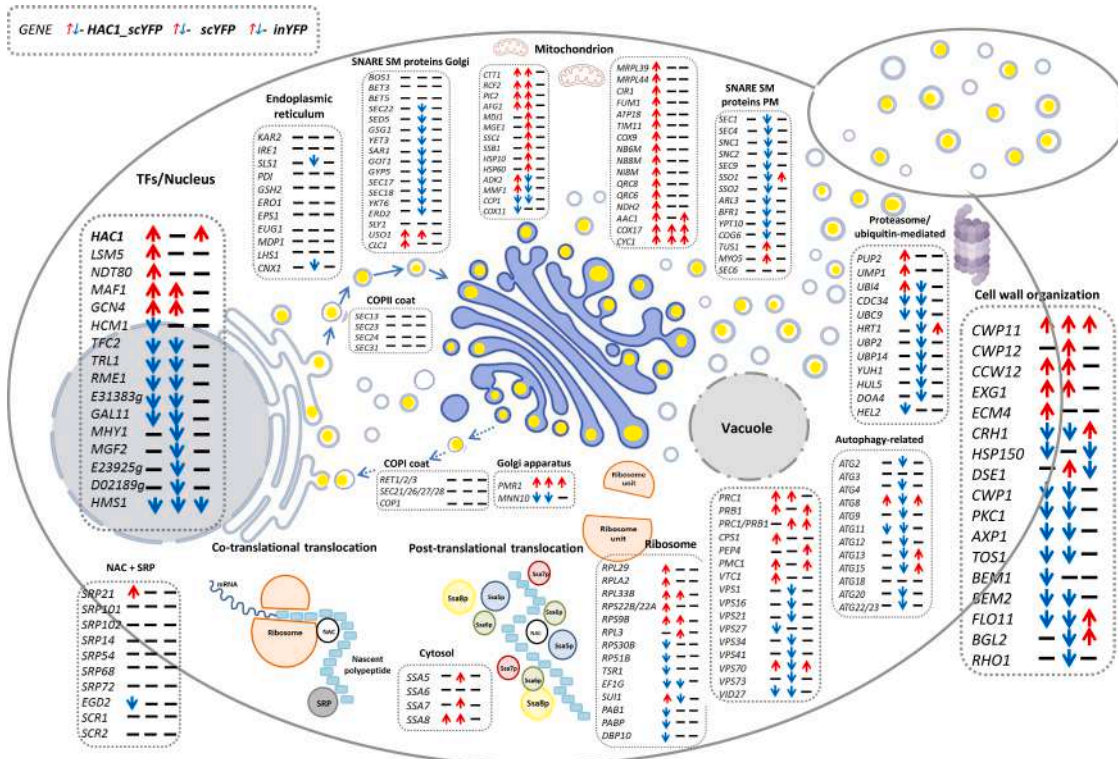


Fig. 2. Overview of expression patterns of selected genes of relevance to r(s)-Prot synthesis and secretion. Red arrows (\uparrow) significant up-regulation over the control strain; blue arrows (\downarrow) down-regulation over the control strain; black dash (-) no significant change in expression level over the control strain. Legend indicates the order of strains: left column: *HAC1_scYFP*; middle column: *scYFP*; right column: *inYFP*. Numerical values of estimated Fold Change (FC) and a full list of genes depicted in this Fig., with their specific functions, is given in Supplementary Table S4.

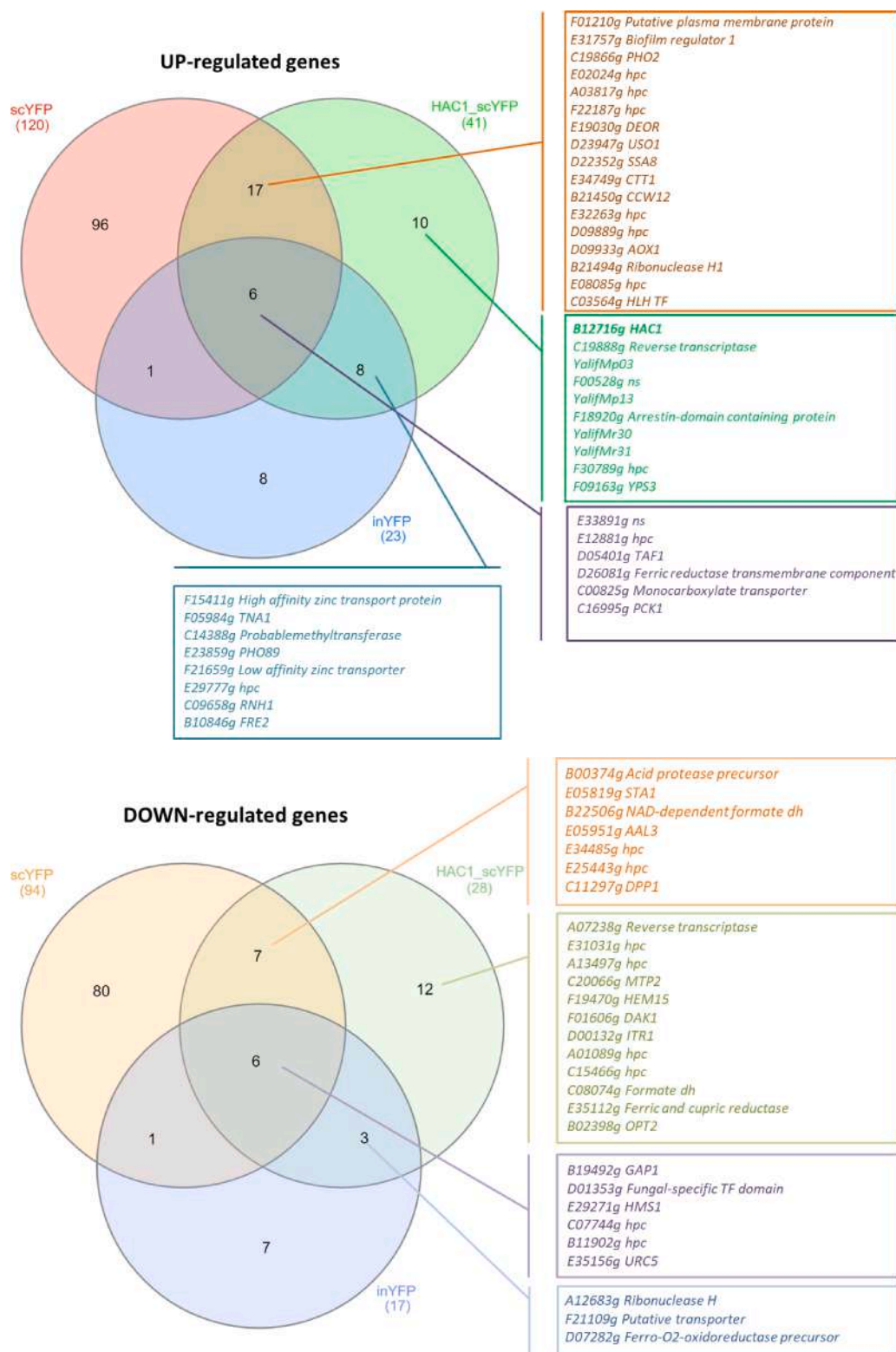


Fig. 3. Venn Diagrams analysis illustrating shared and dissimilar gene sets for individual strains. The numbers indicate the number of shared or strain-specific genes, up-/downregulated in *HAC1_scYFP*, *scYFP*, and *inYFP* strains; (A) up-regulated DEGs or (B) down-regulated DEGs. Venn diagrams analysis was done with the DEGs lists restricted to \log_2 -fold change >1 and p -value < 0.05. YALI_code of the gene and a short description is given for selected genes. Abbreviations: TF – Transcription Factor, dh – dehydrogenase, hpc – hypothetical protein conserved in the *Yarrowia* clade, ns – no similarity.

an elevated level of *HAC1* expression (Fig. 1.A). On the other hand, there was a group of genes that were commonly down-regulated in the strains over-synthesizing sc-/in-YFP, like *AAL3* (*E05951g*) activating acyl chains, *CUP9*-related transcription factor regulating RNA polymerase II

(*E29271g*), but also *GAP1* general amino acid permeases (*B19492g*), and flocculin *STA1* (*E05819g*) (Fig. 3.B; not all the genes are shown in Venn diagrams, as they did not always meet 2-fold change criterion for all the strains).

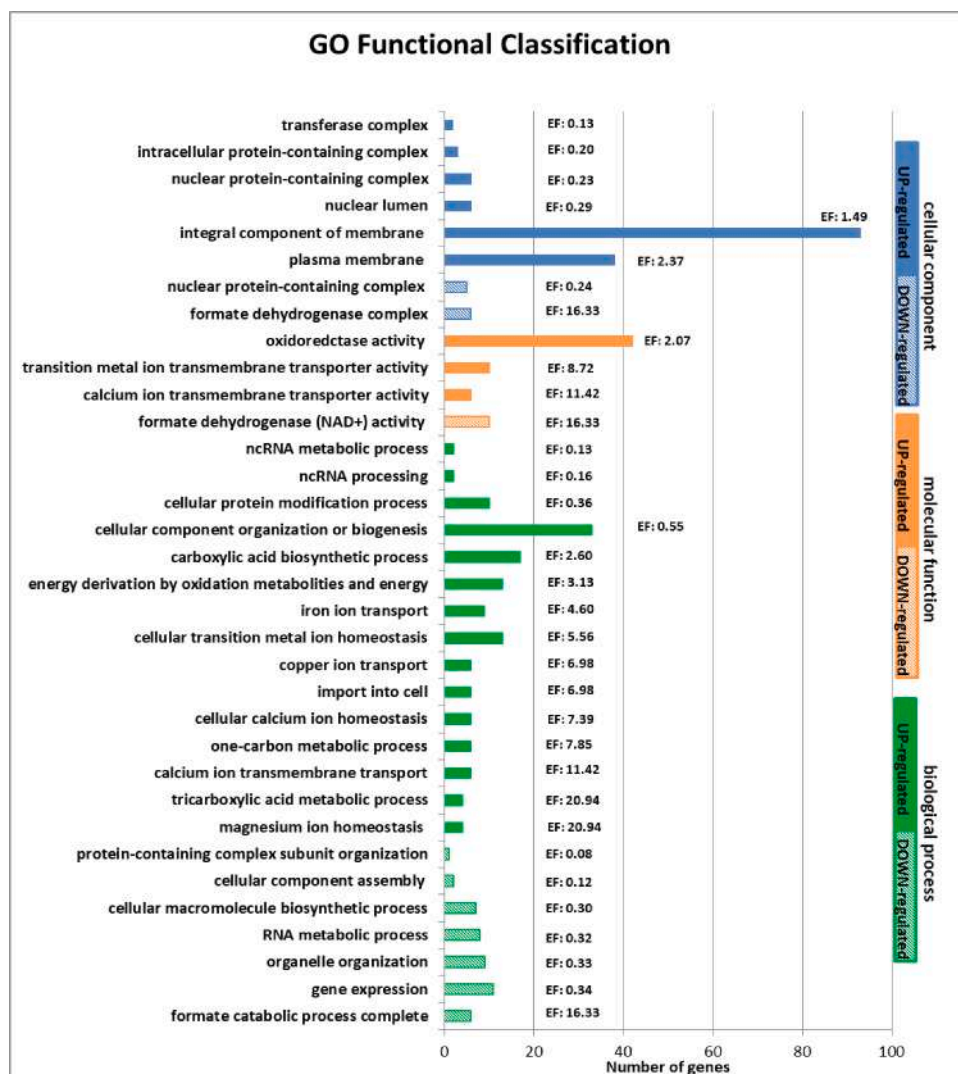


Fig. 4. Functional classification and over-representation test of DEGs in *HAC1-scYFP* strain transcriptome. GO term statistical over-representation test was conducted using the Panther tool on complete DEGs lists. The vertical coordinates are the significantly enriched GO terms, and the horizontal coordinates are the numbers of the up-regulated (full) and down-regulated (dashes) genes in these GO terms. Blue - cellular component; orange - molecular function; green - biological process. Only GO terms enriched at a significance level of $p < 0.05$ are shown. EF – Enrichment fold.

From the above it is clear that the transcriptome of the *HAC1_scYFP* strain bears several unique biological features, but also shares the others with one of the compared strains – *inYFP* or *scYFP* (Figs. 2, 3.A.B). It implies that in the *HAC1_scYFP* strain, the deregulation in the gene expression could equally well result from *scYFP* over-synthesis, the metabolic burden caused by the co-over-expression of the two genes, or over-representation of *HAC1* transcript (and notably - concomitant constant level of *SLS1*, which is an important finding of this study).

For example, several membrane transporters, genes involved in uracil catabolism, and RNA degradation were deregulated solely upon *HAC1* over-representation (*HAC1_scYFP* and *inYFP*). The former observation well corresponds with the results of functional classification followed by an overrepresentation test for that strain (Fig. 4) showing that the cellular component “plasma membrane” and molecular function “transporter” were the most affected (vs the control strain) upon *HAC1* and *SCYFP* co-over-expression. Inhibition of uracil catabolism concomitant with up-regulation of ribonucleases, which decompose RNA to single nucleotides, suggests that RNA components are not degraded but intensively re-used. These biological processes could be linked with the high disproportion between *HAC1* transcript in spliced and un-spliced form (Fig. 1.A.B) seen in the *HAC1_scYFP* strain; and to a lesser extent, the *HAC1* mRNA turnover in the *inYFP* strain.

Genes involved in oxidative stress response (Fig. 3.A.B), cell cycle progression (Fig. 5), and several RNA polymerase II and III repressors (Figs. 2, 3.A.B, 4) were deregulated in response to a rs-Prot over-

synthesis, rather than to the specific action of operable Hac1. Likewise, based on DEGs profiles illustrated in Fig. 2, it seems that nucleus-localized genes (mainly transcription factors) were more responsive to the rs-Prot over-synthesis rather than to the increased occupancy of the *HAC1* transcript; inferred based on the high similarity of *HAC1_scYFP* and *scYFP* strains profiles.

Finally, the global transcriptomes comparison suggests that the unique up-regulation of multiple genes localized to mitochondria and involved in cellular energetics in the *HAC1_scYFP* is a consequence of the co-over-expression of two heterologous genes, and the high-energy-consuming secretion process (Table 1). In this regard, the high-level synthesis and secretion strain, *scYFP*, shows partial similarity.

Still, we are aware that the abovementioned correlations between unique DEGs indicated above and generated phenotypes may equally well indicate a direct functional interaction, or be a consequence of some other not identified phenomena.

3.4. (Co)-Regulome of *HAC1*

Further in relation to the specificity of *HAC1* operation in *Y. lipolytica*, of key importance for us was to investigate the “direct regulome” of *HAC1*, including ER-resident foldases and chaperones (e.g. *KAR2*, *PD11*, etc.), since we previously observed that they are insensitive to changes in *HAC1* expression level in *Y. lipolytica* [1,2]. That observation stayed in contrast to what has been earlier seen in *S. cerevisiae* and

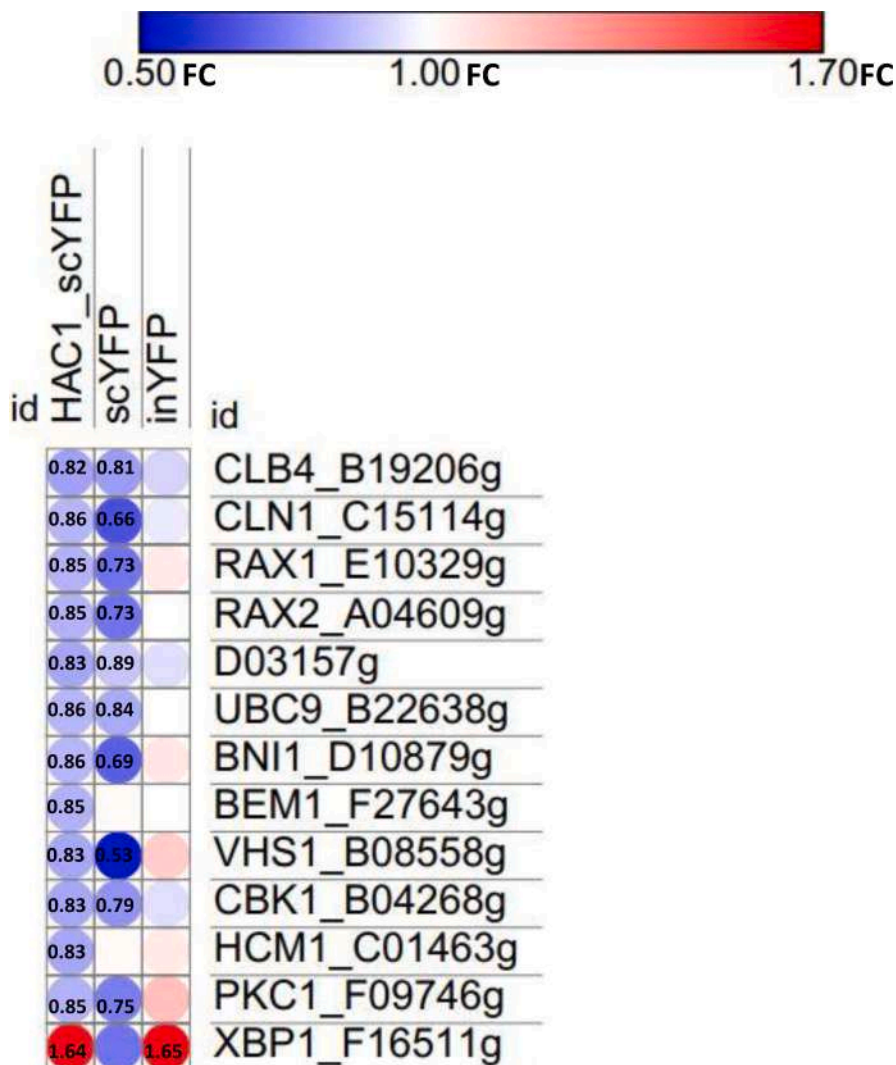


Fig. 5. Expression patterns of DEGs involved in cell cycle progression in HAC1_scYFP, scYFP, and inYFP strains. YALI_code and a short description of a given gene are shown. Assigned names and functions were manually curated for all the ambiguous hits by blasting the sequence against UniProt and NCBI non-redundant protein database, and inferring the most accurate match. Red (up-regulation), blue (down-regulation). Numbers indicate estimated Fold Change (FC). A full list of genes depicted in this Fig., with their specific functions, is given in Supplementary Table S5.

P. pastoris [5,11,18,21], or even for *Y. lipolytica* [19] (assayed by RT-qPCR). The current transcriptomics analysis indicates that none of the ER foldases nor chaperones exhibited any changes in transcription level upon *HAC1* synthetic over-expression or its native induction in the *inYFP* strain (Fig. 2; Supplementary Table S4). Their expression level was not affected either when *HAC1* was synthetically co-overexpressed (assayed here by RNAseq or previously by RT-qPCR [2]), or when the spliced *HAC1* variant was in abundance in the *inYFP* strain (assayed previously by RNAseq [1]). This observation stays in agreement, with our previous observation, and supports the statement that their transcription is not correlated with the *HAC1* expression level in *Y. lipolytica*.

Amongst the genes known to be implicated in *HAC1*-driven UPR in yeast is *XBP1*, which in the present study showed an expression pattern highly corresponding to *HAC1* (Fig. 5). As described for *S. cerevisiae*, *XBP1* encodes a transcriptional repressor binding to promoter sequences of cyclin genes, *CYS3*, *SMF2*, and *CLN1* hence is also claimed to be involved in the G1/S transition of the mitotic cell cycle [51,52]. Homonimic but not homologous *XBP1* (X-box binding protein 1; see alignments in Supplementary Fig. S5) is a widely studied metazoan transcriptional repressor, undergoing unconventional mRNA splicing under conditions of ER stress [48,49,53]. Noteworthy, metazoan *XBP1* is a homolog of yeast *HAC1*, while yeast *XBP1* is a separate gene, operable in yeast (at least in *S. cerevisiae* and *Y. lipolytica*; please see Supplementary Fig. S5). Both transcription factors Hac1 and Xbp1 operate in the yeast cells, and both undergo unconventional splicing by Ire1. Upon

ER stress, *XBP1* mRNA is spliced by Ire1 (as *HAC1*, discussed above), thereby generating functional *XBP1* transcript, which is translocated to the nucleus to initiate transcriptional programs that regulate a subset of UPR- and non-UPR-associated genes. In yeast, Xbp1 is known to repress 15% of all genes as cells transition to quiescence (G1 arrest). It is specifically interesting since *XBP1* was also claimed to be implicated in UPR. It is also known that *XBP1* expression is induced by different stress conditions. A protein encoded by the *yXBP1* gene (*F16511g*) shows significant similarity to the *S. cerevisiae*'s homolog (YIL101C1; Supplementary Fig. S5). Based on its unique up-regulation in *HAC1_scYFP* and *inYFP* it can be speculated that its expression is associated with the *HAC1* transcription / (cytosolic) UPR induction in *Y. lipolytica*. However, this statement requires further investigation.

Intriguingly, amongst the genes significantly down-regulated in the scYFP-over-synthesizing strains we identified a tRNA ligase *TRL1* involved in tRNA biogenesis (*D14916g*). Strikingly, one of the *TRL1*'s functions is the ligation of exons after unconventional splicing of *HAC1/XBP1* transcript to initiate UPR [54]. Aside from *HAC1* and *IRE1*, *TRL1* (*RLG1* in *S. cerevisiae*) is considered the key element of UPR response in *S. cerevisiae* [49]. *TRL1*'s down-regulation was observed in all the *Y. lipolytica* strains co-overexpressing any secretory r-Prot (by 14%, 33%, 14%, and 24% for *HAC1_scYFP*, *scYFP*, *TIG*, and *SoA* strain, respectively [1]), but not for *inYFP*, which is the only case for which a significant increase in spliced *HAC1* transcript counts was observed (Fig. 1.B). While it does not explain how co-over-expression of *HAC1*

contributed to enhanced rs-Prot secretion, it suggests that Trl1 ligase is indeed implicated in unconventional splicing of *HAC1/XBP1* in *Y. lipolytica*, and is the Rlg1's homolog.

3.5. Biological processes affected by *HAC1* and *SCYFP* co-over-expression

Biological processes are most frequently modulated by the cumulative action of many minor modifications of multiple genes. The most efficient way of finding and characterizing such global changes is functional classification followed by an overrepresentation test, that yields numerical value (enrichment fold, EF) indicating whether the group of DEGs is enriched in a specific biological process, molecular function, or localized to a specific cellular compartment. Therefore, all DEGs identified in the *HAC1_scYFP* strain (also those showing expression fold change < 2, but significantly deregulated) were searched for enrichment in specific biological functions. The results of this analysis are illustrated in Fig. 4. Functional categorization showed that the most affected cellular component was the plasma membrane and its integral components, represented by more than 100 up-regulated DEGs. It could be expected since multiple highly-deregulated DEGs were represented by trans-membrane transporters of various types. No such clear indication could be inferred from the enrichment test on down-regulated cellular components, except from the formate dehydrogenase complex, which also constituted an enriched molecular function and a biological process identified amongst the down-regulated DEGs (see Fig. 4). Oxidoreductase activity and ion transporters significantly up-regulated molecular functions in the *HAC1_scYFP* strain. Up-regulation of ion transporter function well corresponds with enrichment in plasma membrane cellular component. Likewise, many different biological processes related to ions transportation and import into the cell were identified as enriched biological processes up-regulated upon co-over-expression of *HAC1* and *SCYFP*. Energy derivation, tricarboxylic acid cycle (TCA), and one-carbon metabolic processes were also identified as significantly up-regulated biological processes (Fig. 4, Supplementary Table S3). Within the latter group, we found many genes involved in amino acid synthesis and conversion (see Supplementary Table S3). Enrichment in energy derivation and TCA well corresponds with concerted up-regulation of mitochondria-localized genes involved in cellular energetics, which was unique for the *HAC1_scYFP* strain (see Fig. 2; Supplementary Table S4). Intensive rebuilding of cellular structures was highlighted by over 30 up-regulated DEGs assigned to cellular components organization and biogenesis, and over 10 DEGs assigned to cellular components assembly that were down-regulated (Fig. 4). Careful analysis of the DEGs assigned to these categories showed that cellular components organization and biogenesis were specifically represented by genes localized to mitochondria, involved in vesicle transportation and chromatin re-structuring (specific DEGs can be found in Supplementary Table S3).

Changes to the protein synthesis process were hallmarked by the deregulation of several ribosomal subunit proteins, and eukaryotic translation initiation factor *SUI1* (*C07524g*; up-regulated by 23%; Fig. 2; Supplementary Table S4). On the contrary, RNA metabolism and gene expression, as well as macromolecule biosynthesis were all down-regulated biological processes upon *HAC1* and *SCYFP* co-over-expression. Specific DEGs classified to these down-regulated categories were ribosomal proteins *RP51B* (*F14465g*; down-regulated by 18%), *RPS30B* (*F14663g*; down-regulated by 29%), pre-rRNA processing protein *TSR1* (*B08756g*; down-regulated by 13%), elongation factor EF-1-gamma (*C16049g*; down-regulated by 20%), as well as tRNA ligase *TRL1* involved in tRNA biogenesis (*D14916g*; down-regulated by 14%).

3.6. Biological processes of specific interest for rs-Prots production

As inferred from the previous [1] and current transcriptomics data, as well as the amounts of YFP reporter protein distribution (Table 1),

three specific biological processes were postulated as Hac1-driven mechanisms for optimization of rs-Prots synthesis in *Y. lipolytica*: (i) rate of r(s)-Prot synthesis, (ii) degradation of proteins, and (iii) vesicular transportation across the secretory pathway. Having transcriptomics data, we were able to specifically track these processes in the analyzed strains at the gene expression level. Transcriptomics data for the genes specifically implicated in these biological processes are globally presented in Fig. 2.

3.6.1. Gene expression regulation

A wide set of genes deregulated upon *HAC1* co-over-expression with *SCYFP* was localized to the nucleus (Fig. 2; Supplementary Table S4). Many of the identified DEGs were transcription factors of different scopes. For example, *MAF1* (*F10541g*; up-regulated by 82%), and *TFC2* (*F05104g*; down-regulated by 25%) exert their regulatory action towards RNA polymerase III. Maf1 is a negative regulator of RNA polymerase III, involved in tRNA processing and stability; while Tfc2 is a positive regulator of rRNA expression. The down-regulated transcription factor *E31383g* is involved in the positive regulation of transcription by RNA polymerase II; while transcription factor *GCN4* (*E27742g*), up-regulated in *HAC1_scYFP* and *scYFP* strains (Fig. 2; Supplementary Table S4), is a transcription factor that is responsible for the activation of more than 30 genes required for amino acid and purine biosynthesis in response to their deficiency. The expression pattern of transcription factors *MAF1*, *TFC2*, and *E31383g* across the analyzed strains suggests that transcription by the RNA polymerase III and II is limited upon high-level rs-Prot synthesis and secretion. This observation well corresponds with the down-regulation of the "gene expression" functional category (Fig. 4). Up-regulation of transcription factor *GCN4* in *HAC1_scYFP* and *scYFP* strains (Fig. 2; Supplementary Table S4) highlights the occurrence of nitrogen starvation in the *scYFP* over-synthesizing strains. Upon induction of UPR, *GCN4* was up-regulated in *P. pastoris* but down-regulated in *S. cerevisiae* [11]. Gcn4 was shown to be activated upon ER stress and to play an essential role in the UPR together with or downstream of Hac1 [55]. Our data show that *GCN4* expression level is not directly linked to *HAC1* expression in *Y. lipolytica*. It is more associated with amino acid starvation under intensive over-synthesis of rs-Prot. Co-over-expression of *HAC1* and *SCYFP* caused significant changes to the ribosome operation (Fig. 2; Supplementary Table S4). In the face of indicated above attenuation of transcription by RNA polymerases III and II, the down-regulation of multiple translation-associated genes is more understandable (Fig. 2, Supplementary Table S4). Notably, such concerted changes to ribosome-related genes were observed neither in highly-synthesizing *scYFP* nor in spliced *HAC1*-up-regulated *inYFP* strain, so it is plausible that a combination of both genes' co-over-expression contributed to this effect. Altogether these data indicate that upon high-level rs-Prot synthesis gene expression is attenuated and the cells encounter nitrogen starvation.

3.6.2. Protein degradation via different pathways

The second proposed mechanism by which Hac1 executes its protective role in *Y. lipolytica* cells is targeting excessively accumulated, incorrectly folded r(s)-Prots to degradation (Fig. 2). We observed that indeed this biological process was up-regulated in *HAC1_scYFP* strain, even though the model protein was "easy-to-fold" (based on previous comparisons [1,25]). Specifically, the main vacuolar protease *PRB1* (*B16500g*; up-regulation by 22%), vacuolar Ca²⁺ ATPase *PMC1* (*D04873g*; up-regulation by 37%), vacuolar protein sorting factor *VPS70* (*B05258g*; up-regulation by 31%) and essential for autophagy *ATG8* gene (*E02662g*; up-regulation by 22%) were all up-regulated. This observation was contrasted by global down-regulation of the genes involved in any of the protein degradation pathways (proteasome, autophagy / vacuolar) in the *scYFP* strain and the other highly over-synthesizing strain SoA [1], with no change or up-regulation in the *HAC1*-up-regulation strains (*HAC1_scYFP*, *inYFP* and *TIG*), (Fig. 2; Supplementary Table S4) [1]. Amongst the genes specifically up-regulated

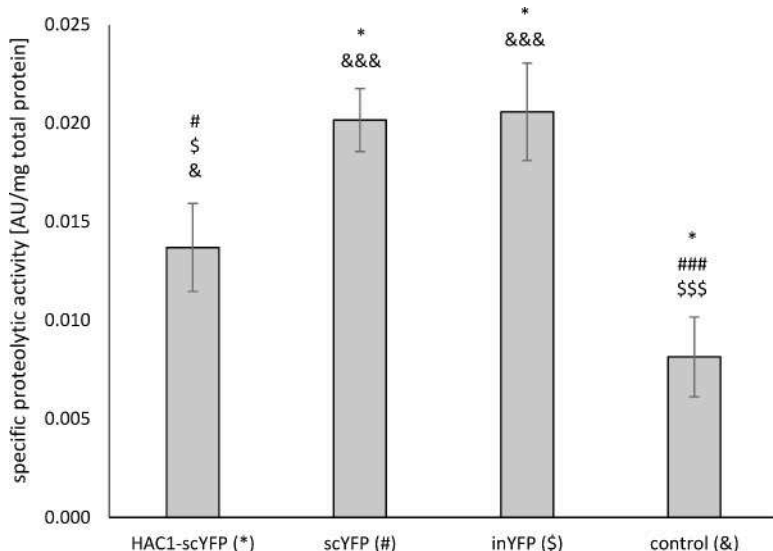


Fig. 6. Total intracellular proteolytic activity of the analyzed strains. Proteolytic activity was determined in total protein extracts. Y axis: specific proteolytic activity equivalent to trypsin activity in standardized solution per a unit of total protein concentration expressed in [AU/mg total protein]. Each strain is marked with a symbol: *, #, \$, and &, for *HAC1_scYFP*, *scYFP*, *inYFP*, and control, respectively. Top labels indicate the statistical significance of the differences at the level of $p < 0.05$ and 0.001 .

in *HAC1*-induced strains was *E02662g* encoding the crucial for autophagy initiation - Atg8. *ATG8* is the key gene involved in autophagy, directing the vacuolar degradation of autophagosomes. Interestingly, it is also known to be involved in the ER-specific autophagy process and is essential for the survival of cells subjected to severe ER stress. Hence, based on such a specific expression pattern, conserved across so many genes, we hypothesized that the *HAC1* expression level is directly linked to the vacuolar sorting and degradation.

To verify this statement, we examined the total intracellular proteolytic activity of the strains studied here. Based on the assay results (Fig. 6), any over-synthesis of r(s)-Prot led to increased total proteolytic activity. On the other hand, co-over-expression of *HAC1* contributed to a significant decrease of proteolysis vs the other two YFP-over-synthesizing strains, but still it was higher in *HAC1_scYFP* strain than in the control. The biological sense of the enhancement in total proteolytic activity upon over-synthesis of r(s)-Prot seems clear and is most probably related to globally enhanced protein turnover. However, such an explanation does not correspond to the overall profile of transcriptomics data, from which it is clear that proteolysis is inhibited in the *scYFP* strain. Careful analysis of the transcriptomics data revealed that many of the down-regulated genes assigned to the functional category “proteolysis” are vacuolar sorting factors, ubiquitination-mediating factors, and autophagy-related factors, which do not possess proteolytic activity *per se*, which was assayed in the test. In other words, it was the process of the protein targeting for degradation, that was decreased in the *scYFP* strain. Specific proteolytic activity was the highest in the strains facing enhanced YFP retention, i.e. *scYFP* and *inYFP* (Table 1, Supplementary Fig. S2). On the other hand, co-over-expression of *HAC1* contributed to a significant decrease of proteolysis vs these two strains, which still was higher than in the control. The decreased proteolytic activity in the *HAC1_scYFP* strain could partially explain the higher amounts of secreted YFP (Table 1).

3.6.3. Vesicular transportation across the secretory pathway

The data presented in Table 1 suggest that one of the reasons for increased amounts of the secreted YFP reporter in the *HAC1_scYFP* strain could be intensified transportation through the secretory pathway. From amongst the genes involved in anterograde and retrograde COP-mediated transport, intra- and post-Golgi trafficking, all except two (*USO1*, *D23947g*, essential for vesicle-mediated ER to Golgi transport; up-regulated by 2.58-fold; *CLC1*, *E33253g*, vesicle coat protein, clathrin light chain; up-regulated by 18%) remained at the control expression level. At the same time, many genes involved in vesicular transportation were down-regulated in the *scYFP* strain (Fig. 2; Supplementary

Table S4), which over-synthesized the rs-Prot at a very high level and secreted it. Looking at the data illustrated in Fig. 2, it was not possible to indicate specific DEGs that could account for such an outcome. The surprising down-regulation of vesicular transportation genes under high-level synthesis and secretion in the *scYFP* strain is not well understood at this moment.

3.6.4. Cell cycle progression

Based on the conducted enrichment test (not shown), we identified another functional category that was enriched amongst DEGs. Namely, a set of genes involved in cell cycle progression was significantly down-regulated in response to the co-over-expression of *HAC1* and *SCYFP* (Fig. 5; Supplementary Table S5). A similar, yet not identical, expression pattern was observed for the *scYFP* strain. The expression pattern of the genes involved in cell-cycle progression was characterized by concerted down-regulation of several serine/threonine-protein kinases involved in cell division and morphogenesis (*VHS1*, *B08558g*; *CBK1*, *B04268g*; *PKC1*, *F09746g*, G2/M-specific), two cyclins *CLB4* (*B19206g*, G2/M-specific) and *CLN1* (*C15114g*, G1/S-specific), as well as *UBC9* gene (*B22638g*) specifically involved in the degradation of cyclins of S- and M-phases. Furthermore, we observed significant down-regulation of an *HCMI* transcription factor involved in cell cycle-specific transcription (*C01463g*), proteins *RAX1* (*E10329g*) and *RAX2* (*A04609g*), *BNI1* (*D10879g*) and *BEM1* (*F27643g*), all implicated in biological processes like bud site selection, cytokinesis, maintenance of cell polarity; as well as GTP-binding protein *D03157g* involved in the mitotic spindle and division septum assembly. Strikingly, in line with the down-regulation of cell cycle progression, we noted significant up-regulation of the *XBPI* repressor gene (*F16511g*) in both *HAC1_scYFP* and *inYFP* strains, contrasted by its down-regulation in the *scYFP* strain (Fig. 5); its involvement in cell cycle regulation but also in ER stress response was confirmed for *S. cerevisiae* [51–53]. Considering previously [1] and currently observed expression patterns of *HAC1* and *XBPI* across all the strains analyzed (*SoA*, *scYFP*, *HAC1_scYFP*, *inYFP*, and *TIG*) it can be concluded that the expression of the two genes is correlated. Based on the “guilt-by-association” rule, it suggests their functional interaction of yet unknown mechanisms.

4. Conclusions

In conclusion, this study provides insight into several new elements of the secretory pathway in *Y. lipolytica*. A practical outcome of this research is demonstrating that *HAC1* co-overexpression in *Y. lipolytica* leads to enhanced secretion of the correctly folded target protein and its

limited retention. The identified associated molecular mechanisms were massive changes in mitochondria and ribosomes operation, cell cycle arrest, attenuation of gene expression by RNA polymerase III and II, as well as modulation of proteolysis and RNA metabolism; but whether the *HAC1* co-over-expression/induction was the actual causative agent for these changes, was not always clear. On the other hand, it was settled that the expression level of the “conventional” *HAC1* targets, like *KAR2*, *PDII*, and other ER-localized foldases, is not affected by increased expression of *HAC1* in *Y. lipolytica*. Moreover, our transcriptomics data suggest the implication of the Xbp1 transcription factor (*F16511g*) and the Trl1 ligase (*D14916g*) in the *HAC1*-driven cellular response and unconventional splicing of *HAC1* mRNA in *Y. lipolytica*, respectively.

Funding

This study was partly financially supported by the Ministry of Science and Higher Education (currently, Ministry of Education and Science) project number: DI 2017 000947, in the part concerning chemostat cultivations, samples processing, and part of RNAseq and bioinformatics. The research was also partly funded by National Science Centre, Poland, grant number 2021/41/B/NZ9/00086; in the part concerning the investigation of the *HAC1* splicing based on transcripts sequencing, and the impact on *HAC1* overexpression on the intracellular proteolysis.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

CRedit authorship contribution statement

Paulina Korpys-Woźniak: Investigation, Methodology, Visualization, Data curation, Writing – original draft, Writing – review & editing.
Ewelina Celińska: Conceptualization, Supervision, Formal analysis, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data accompanying this research are presented directly in the manuscript, and supplementary materials or are available in Sequence Read Archive, NCBI database under BioProject number: PRJNA869113.

Acknowledgments

Graphical Abstract was prepared with BioRender.com.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.btre.2023.e00801](https://doi.org/10.1016/j.btre.2023.e00801).

References

- [1] P. Korpys-Woźniak, E. Celińska, Global transcriptome profiling reveals genes responding to overproduction of a small secretory, a high cysteine- and a high glycosylation-bearing protein in *Yarrowia lipolytica*, *Biotechnol. Rep.* 31 (2021) e00646, <https://doi.org/10.1016/j.btre.2021.e00646>.

- [2] P. Korpys-Woźniak, P. Kubiak, E. Celińska, Secretary helpers for enhanced production of heterologous proteins in *Yarrowia lipolytica*, *Biotechnol. Rep.* 32 (2021) e00669, <https://doi.org/10.1016/j.btre.2021.e00669>.
- [3] K. Mori, T. Kawahara, H. Yoshida, H. Yanagi, T. Yura, Signalling from endoplasmic reticulum to nucleus: transcription factor with a basic-leucine zipper motif is required for the unfolded protein-response pathway, *Genes Cells* 1 (1996) 803–817, <https://doi.org/10.1046/j.1365-2443.1996.d01-274.x>.
- [4] M. Guerfal, S. Ryckaert, P.P. Jacobs, P. Ameloot, K. Van Craenenbroeck, R. Derycke, N. Callewaert, The *HAC1* gene from *Pichia pastoris*: characterization and effect of its overexpression on the production of secreted, surface displayed and membrane proteins, *Microb. Cell Fact* 9 (2010) 1–12, <https://doi.org/10.1186/1475-2859-9-49>.
- [5] A.R. Heskeith, J.I. Castrillo, T. Sawyer, D.B. Archer, S.G. Oliver, Investigating the physiological response of *Pichia (Komagataella) pastoris* GS115 to the heterologous expression of misfolded proteins using chemostat cultures, *Appl. Microbiol. Biotechnol.* 97 (2013) 9747–9762, <https://doi.org/10.1007/s00253-013-5186-1>.
- [6] D. Mattanovich, B. Gasser, H. Hohenblum, M. Sauer, Stress in recombinant protein producing yeasts, *J. Biotechnol.* 113 (2004) 121–135, <https://doi.org/10.1016/j.jbiotec.2004.04.035>.
- [7] A. Graf, B. Gasser, M. Dragosits, M. Sauer, G.G. Leparc, T. Tüchler, D.P. Kreil, D. Mattanovich, Novel insights into the unfolded protein response using *Pichia pastoris* specific DNA microarrays, *Biomed. Chromatogr.* 9 (2008) 1–13, <https://doi.org/10.1186/1471-2164-9-390>.
- [8] K.J. Kauffman, E.M. Pridgen, F.J. Doyle, P.S. Dhurjati, A.S. Robinson, Decreased protein expression and intermittent recoveries in BiP levels result from cellular stress during heterologous protein expression in *Saccharomyces cerevisiae*, *Biotechnol. Prog.* 18 (2002) 942–950, <https://doi.org/10.1021/bp025518g>.
- [9] K.E.J. Tyo, Z. Liu, D. Petranovic, J. Nielsen, Imbalance of heterologous protein folding and disulfide bond formation rates yields runaway oxidative stress, *BMC Biol.* 10 (2012), <https://doi.org/10.1186/1741-7007-10-16>.
- [10] M. Hernández-Elvira, F. Torres-Quiroz, A. Escamilla-Ayala, E. Domínguez-Martin, R. Escalante, L. Kawasaki, L. Ongay-Larios, R. Coria, The unfolded protein response pathway in the yeast *Kluyveromyces lactis*. A comparative view among yeast species, *Cells* 7 (2018), <https://doi.org/10.3390/cells7080106>.
- [11] B. Gasser, M. Maurer, J. Rautio, M. Sauer, A. Bhattacharyya, M. Saloheimo, M. Penttilä, D. Mattanovich, Monitoring of transcriptional regulation in *Pichia pastoris* under protein production conditions, *Biomed. Chromatogr.* 8 (2007) 1–18, <https://doi.org/10.1186/1471-2164-8-179>.
- [12] J. Liu, Q. Han, Q. Cheng, Y. Chen, R. Wang, X. Li, Y. Liu, D. Yan, Efficient expression of human lysozyme through the increased gene dosage and co-expression of transcription factor Hac1p in *Pichia pastoris*, *Curr. Microbiol.* 77 (2020) 846–854, <https://doi.org/10.1007/s00284-019-01872-9>.
- [13] J. Hou, H. Tang, Z. Liu, T. Österlund, J. Nielsen, D. Petranovic, Management of the endoplasmic reticulum stress by activation of the heat shock response in yeast, *FEMS Yeast Res* 14 (2014) 481–494, <https://doi.org/10.1111/1567-1364.12125>.
- [14] B. Gasser, M. Sauer, M. Maurer, G. Stadlmayr, D. Mattanovich, Transcriptomics-based identification of novel factors enhancing heterologous protein secretion in yeasts, *Appl Environ Microbiol* 73 (2007) 6499–6507, <https://doi.org/10.1128/AEM.01196-07>.
- [15] X. Zhao, W. Xie, Y. Lin, X. Lin, S. Zheng, S. Han, Combined strategies for improving the heterologous expression of an alkaline lipase from *Acinetobacter radioresistens* CMC-1 in *Pichia pastoris*, *Process Biochem.* 48 (2013) 1317–1323, <https://doi.org/10.1016/j.procbio.2013.06.021>.
- [16] H. Raschmanová, A. Weninger, Z. Knejzlík, K. Melzoch, K. Kovar, Engineering of the unfolded protein response pathway in *Pichia pastoris*: enhancing production of secreted recombinant proteins, *Appl. Microbiol. Biotechnol.* 105 (2021) 4397–4414, <https://doi.org/10.1007/s00253-021-11336-5>.
- [17] R. Gordán, K.F. Murphy, R.P. McCord, C. Zhu, A. Vedenko, M.L. Bulyk, Curated collection of yeast transcription factor DNA binding specificity data reveals novel structural and gene regulatory insights, *Genome Biol.* 12 (2011), <https://doi.org/10.1186/gb-2011-12-12-r125>.
- [18] G. Whyteside, R.M. Nor, M.J.C. Alcocer, D.B. Archer, Activation of the unfolded protein response in *Pichia pastoris* requires splicing of a *HAC1* mRNA intron and retention of the C-terminal tail of Hac1p, *FEBS Lett.* 585 (2011) 1037–1041, <https://doi.org/10.1016/j.febslet.2011.02.036>.
- [19] N. Callewaert, W. Vervecken, K. De Pourq, S. Geysens, M. Guerfal, *Yarrowia lipolytica* and *Pichia pastoris* Hac1 nucleic acids, *US* 8026083 B2, 2011.
- [20] M.H. Oh, S.A. Cheon, H.A. Kang, J.Y. Kim, Functional characterization of the unconventional splicing of *Yarrowia lipolytica* *HAC1* mRNA induced by unfolded protein response, *Yeast* 27 (2010) 443–452, <https://doi.org/10.1002/yea.1762>.
- [21] Z. Liu, J. Hou, J.L. Martínez, D. Petranovic, J. Nielsen, Correlation of cell growth and heterologous protein production by *Saccharomyces cerevisiae*, *Appl. Microbiol. Biotechnol.* 97 (2013) 8955–8962, <https://doi.org/10.1007/s00253-013-4715-2>.
- [22] A. Babour, M. Kabani, A. Boisramé, J.M. Beckerich, Characterization of Ire1 in the yeast *Yarrowia lipolytica* reveals an important role for the Sls1 nucleotide exchange factor in unfolded protein response regulation, *Curr. Genet.* 53 (2008) 337–346, <https://doi.org/10.1007/s00294-008-0190-1>.
- [23] A. Boisramé, J.M. Beckerich, C. Gaillardin, Sls1p, an endoplasmic reticulum component, is involved in the protein translocation process in the yeast *Yarrowia lipolytica*, *J. Biol. Chem.* 271 (1996) 11668–11675, <https://doi.org/10.1074/jbc.271.20.11668>.
- [24] A. Boisramé, M. Kabani, J.M. Beckerich, E. Hartmann, C. Gaillardin, Interaction of Kar2p and Sls1p is required for efficient co-translational translocation of secreted proteins in the yeast *Yarrowia lipolytica*, *J. Biol. Chem.* 273 (1998) 30903–30908, <https://doi.org/10.1074/jbc.273.47.30903>.

- [25] P. Korpys-Woźniak, P. Kubiak, W. Białas, E. Celińska, Impact of overproduced heterologous protein characteristics on physiological response in *Yarrowia lipolytica* steady-state-maintained continuous cultures, *Appl. Microbiol. Biotechnol.* 104 (2020) 9785–9800, <https://doi.org/10.1007/s00253-020-10937-w>.
- [26] J.M. Nicaud, C. Madzak, P. Van Den Broek, C. Gysler, P. Duboc, P. Niederberger, C. Gaillardin, Protein expression and secretion in the yeast *Yarrowia lipolytica*, *FEMS Yeast Res.* 2 (2002) 371–379, [https://doi.org/10.1016/S1567-1356\(02\)00082-X](https://doi.org/10.1016/S1567-1356(02)00082-X).
- [27] E. Celińska, M. Borkowska, W. Białas, P. Korpys, J.M. Nicaud, Robust signal peptides for protein secretion in *Yarrowia lipolytica*: identification and characterization of novel secretory tags, *Appl. Microbiol. Biotechnol.* 102 (2018) 5221–5233, <https://doi.org/10.1007/s00253-018-8966-9>.
- [28] E. Celińska, M. Borkowska, P. Korpys-Woźniak, M. Kubiak, J. Nicaud, P. Kubiak, M. Gorczyca, W. Białas, Optimization of *Yarrowia lipolytica*-based consolidated biocatalyst through synthetic biology approach: transcription units and signal peptides shuffling, *Appl. Microbiol. Biotechnol.* 104 (2020) 5845–5859, <https://doi.org/10.1007/s00253-020-10644-6>.
- [29] E. Celińska, R. Ledesma-Amaro, M. Larroude, T. Rossignol, C. Pauthenier, J. M. Nicaud, Golden Gate Assembly system dedicated to complex pathway manipulation in *Yarrowia lipolytica*, *Microb Biotechnol.* 10 (2017) 450–455, <https://doi.org/10.1111/1751-7915.12605>.
- [30] P. Soudier, M. Larroude, E. Celińska, T. Rossignol, J.M.J.-M. Nicaud, Selection of heterologous protein-producing strains in *Yarrowia lipolytica*, in: B. Gasser, D. Mattanovich (Eds.), *Methods in Molecular Biology*, Springer Science+Business Media, LLC, part of Springer Nature, 2019, pp. 153–168, https://doi.org/10.1007/978-1-4939-9024-5_6.
- [31] P. Korpys-Woźniak, M. Kubiak, M. Borkowska, E. Celińska, Construction and Assembly of Standardized Biobricks for Synthetic Pathways Engineering in Yeasts, in: F. Darvishi Harzevili (Ed.), *Synthetic Biology of Yeasts: Tools and Applications*, Springer International Publishing, Cham, 2022, pp. 27–62, https://doi.org/10.1007/978-3-030-89680-5_2.
- [32] M. Kubiak-Szymendera, B. Skupień-Rabian, U. Jankowska, E. Celińska, Hyperosmolarity adversely impacts recombinant protein synthesis by *Yarrowia lipolytica* – molecular background revealed by quantitative proteomics, *Appl. Microbiol. Biotechnol.* 106 (2021) 349–367, <https://doi.org/10.1007/s00253-021-11731-y>.
- [33] J. Sambrook, D. Russell, *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory Press, New York, 2001.
- [34] M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads, *EMBnet J.* 17 (2011) 10–12, <https://doi.org/10.14806/ej.17.1.200>.
- [35] D. Kim, J. Paggi, C. Park, C. Bennett, S. Salzberg, Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype, *Nat. Biotechnol.* 37 (2019) 907–915, <https://doi.org/10.1038/s41587-019-0201-4>.
- [36] S. Anders, P.T. Pyl, W. Huber, HTSeq-A Python framework to work with high-throughput sequencing data, *Bioinformatics* 31 (2015) 166–169, <https://doi.org/10.1093/bioinformatics/btu638>.
- [37] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, *Genome Biol.* 15 (2014) 1–21, <https://doi.org/10.1186/s13059-014-0550-8>.
- [38] H. Mi, A. Muruganujan, X. Huang, D. Ebert, C. Mills, X. Guo, P. Thomas, Protocol update for large-scale genome and gene function analysis with PANTHER Classification System (v.14.0), *Nat. Protoc.* 14 (2019) 703–721, <https://doi.org/10.1038/s41596-019-0128-8>.
- [39] H. Mi, D. Ebert, A. Muruganujan, C. Mills, L.-P. Albou, T. Mushayama, P. D. Thomas, PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API, *Nucleic Acids Res.* 49 (2021) D394–D403, <https://doi.org/10.1093/nar/gkaa1106>.
- [40] L.D. Goldstein, Y. Cao, G. Pau, M. Lawrence, T.D. Wu, S. Seshagiri, R. Gentleman, Prediction and quantification of splice events from RNA-Seq Data, *PLoS ONE* 11 (2016), e0156132, <https://doi.org/10.1371/journal.pone.0156132>.
- [41] J.T. Robinson, H. Thorvaldsdóttir, W. Winckler, M. Guttman, E.S. Lander, G. Getz, J.P. Mesirov, Integrative genomics viewer, *Nat. Biotechnol.* 29 (2011) 24–26, <https://doi.org/10.1038/nbt.1754>.
- [42] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The sequence alignment/map format and SAMtools, *Bioinformatics* 25 (2009) 2078–2079, <https://doi.org/10.1093/bioinformatics/btp352>.
- [43] T. Hall, Bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, *Nucleic Acids Symp. Ser.* 41 (1999) 95–98.
- [44] F. Corpet, Multiple sequence alignment with hierarchical clustering, *Nucleic Acids Res.* 16 (1988) 10881–10890, <https://doi.org/10.1093/nar/16.22.10881>.
- [45] K. Löchli, E. Torbica, M. Haile-Weldeslasie, D. Baku, A. Aziz, D. Bublak, S. Fragkostefanakis, Crosstalk between endoplasmic reticulum and cytosolic unfolded protein response in tomato, *Cell Stress Chaperones* (2022), <https://doi.org/10.1007/s12192-022-01316-7>.
- [46] K.J. Travers, C.K. Patil, L. Wodicka, D.J. Lockhart, J.S. Weissman, P. Walter, Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation, 2000. <https://www.cell.com/cgi/content/full/101/3/249/DC1>.
- [47] J.S. Cox, P. Walter, A novel mechanism for regulating activity of a transcription factor that controls the unfolded protein response, *Cell* 87 (1996) 391–404.
- [48] H. Yoshida, K. Haze, H. Yanagi, T. Yura, K. Mori, Identification of the cis-acting endoplasmic reticulum stress response element responsible for transcriptional induction of mammalian glucose-regulated proteins: involvement of basic leucine zipper transcription factors, *J. Biol. Chem.* 273 (1998) 33741–33749, <https://doi.org/10.1074/jbc.273.50.33741>.
- [49] H. Yoshida, T. Matsui, A. Yamamoto, T. Okada, K. Mori, XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor, *Cell* 107 (2001) 881–891.
- [50] M. Kabani, J.-M. Beckerich, C. Gaillardin, Sls1p stimulates Sec63p-mediated activation of Kar2p in a conformation-dependent manner in the yeast endoplasmic reticulum, *Mol. Cell Biol.* 20 (2000) 6923–6934, <https://doi.org/10.1128/mcb.20.18.6923-6934.2000>.
- [51] B. Mai, L. Breeden, Xbp1, a stress-induced transcriptional repressor of the *Saccharomyces cerevisiae* Swi4/Mbp1 family, *Mol. Cell Biol.* 17 (1997) 6491–6501, <https://doi.org/10.1128/mcb.17.11.6491>.
- [52] B. Mai, L. Breeden, CLN1 and its repression by Xbp1 are important for efficient sporulation in budding yeast, *Mol. Cell Biol.* 20 (2000) 478–487, <https://doi.org/10.1128/mcb.20.2.478-487.2000>.
- [53] T. Iwawaki, M. Tokuda, Function of yeast and amphioxus tRNA ligase in IRE1 α -dependent XBP1 mRNA splicing, *Biochem. Biophys. Res. Commun.* 413 (2011) 527–531, <https://doi.org/10.1016/j.bbrc.2011.08.129>.
- [54] K.B. Hooks, S. Griffiths-Jones, Conserved RNA structures in the non-canonical Hac1/Xbp1 intron, *RNA Biol.* 8 (2011) 552–556, <https://doi.org/10.4161/rna.8.4.15396>.
- [55] C.K. Patil, H. Li, P. Walter, Gcn4p and novel upstream activating sequences regulate targets of the unfolded protein response, *PLoS Biol.* 2 (2004), <https://doi.org/10.1371/journal.pbio.0020246>.

SUPPLEMENTARY

Molecular background of *HAC1*-driven improvement in the secretion of recombinant protein in *Yarrowia lipolytica* based on comparative transcriptomics

List of Supplementary Materials

Supplementary Figure S1. Schematic representation of cloning vectors used in this study.

Supplementary Figure S2. Growth and morphology of *Y. lipolytica* strains used in this study.

Supplementary Figure S3. Kinetics of Po1f-derived *HAC1_scYFP* *Y. lipolytica* strain culture.

Supplementary Figure S4. Visualization of *HAC1* sequence mapping to the reference genome using IGV software.

Supplementary Figure S5. Amino acid sequences and multiple sequence alignments of *Hac1* and *Xbp1* from *S. cerevisiae* and *Y. lipolytica*.

Supplementary Figure S6. Volcano plots showing the dependence of the logarithmic changes of expression fold [\log_2 _fold change] values from the logarithmic values of mean expression [\log_2 _mean_expression] of the *HAC1_scYFP* strain vs control.

Supplementary Table S1. Strains used in this study.

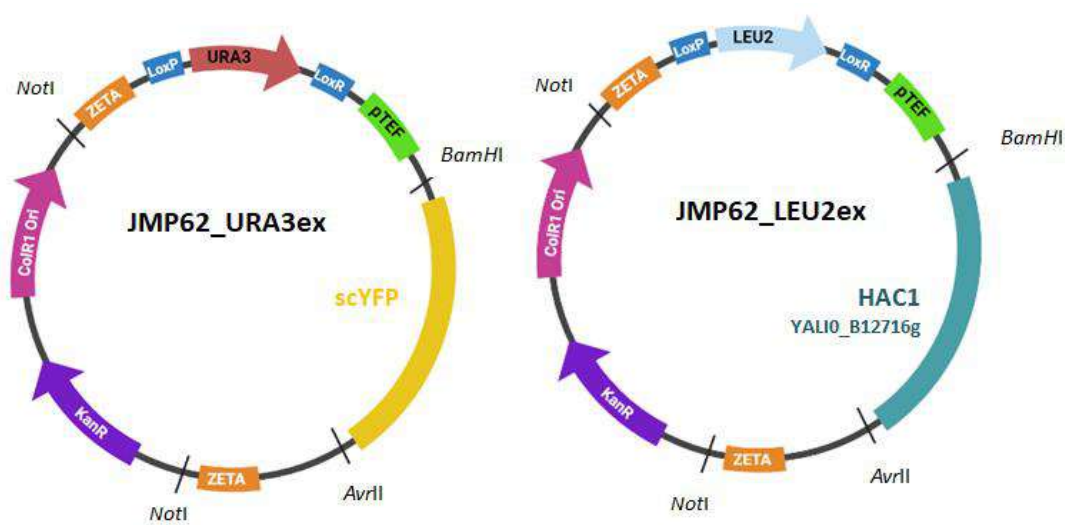
Supplementary Table S2. Nucleotide sequences of the genes cloned in this study (*HAC1* and *SP1_YFP*)

Supplementary Table S3. Full list of DEGs identified in *HAC1_scYFP* strain over the control. Any DEGs listed here must meet the criteria indicated in section **2.7.1.1** – online Excel file

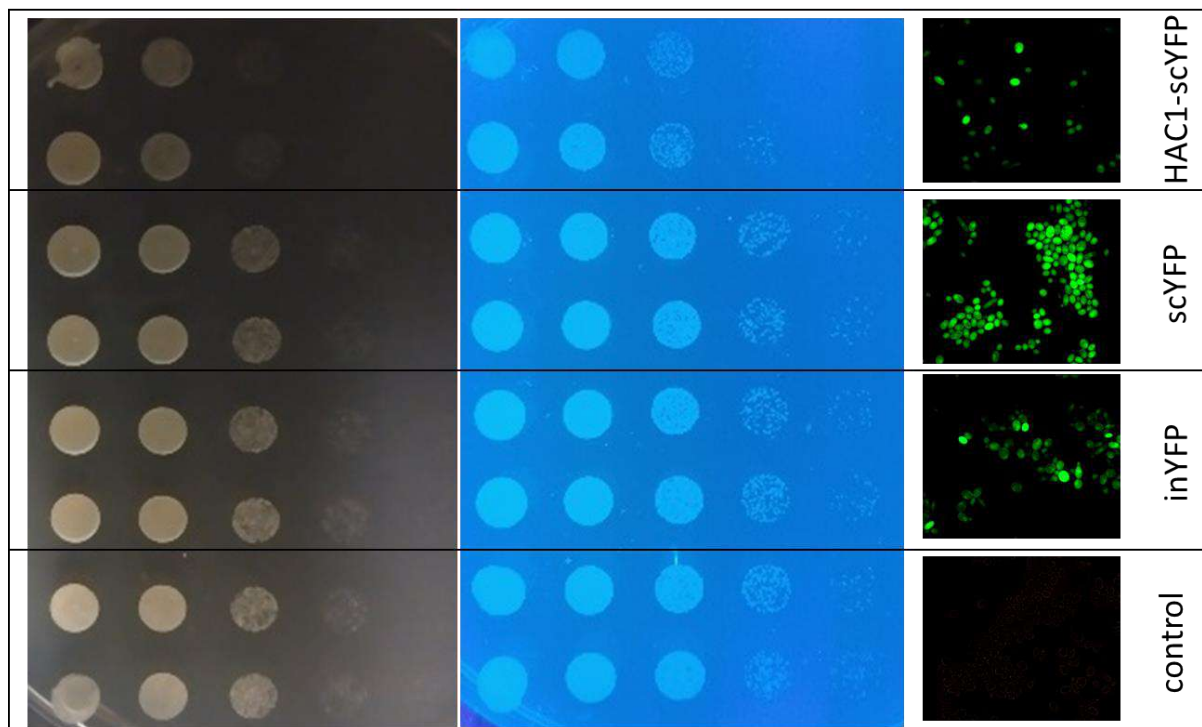
Supplementary Table S4. The list of genes displayed in **Figure 2** with fold change (FC) values – online Excel file

Supplementary Table S5. The list of genes displayed in **Figure 5** – online Excel file.

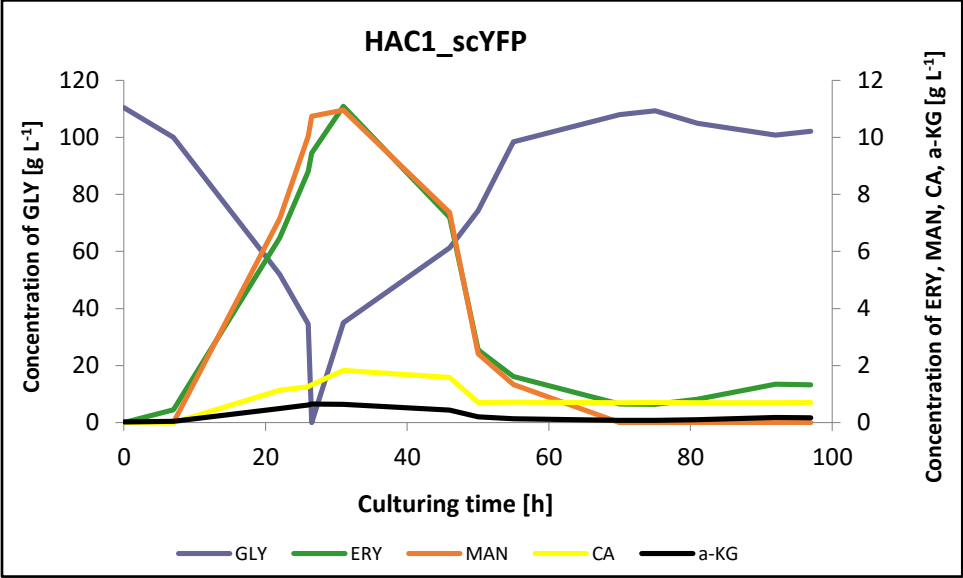
Supplementary Figure S1. Schematic representation of cloning vectors used in this study [1]. Two vectors from the JMP62 series were co-transformed into the *Y. lipolytica* recipient strain. Elements of the vectors were indicated in colors: zeta region (orange), *LoxP* and *LoxR* (dark blue), constitutive *pTEF* promoter (green), genes of interest: *scYFP* (A. yellow) or *HAC1* (B. blue) equipped in the *Bam*HI-*Avr*II fragments, the *URA3ex* (A. red) or *LEU2ex* (B. light blue) excisable auxotrophic markers. Bacterial *ori* of replication (pink) and kanamycin resistance gene *KanR* (purple), contained in the “bacterial” part of the vectors, were discarded before the yeast cell transformation by *Not*I restriction digestion.



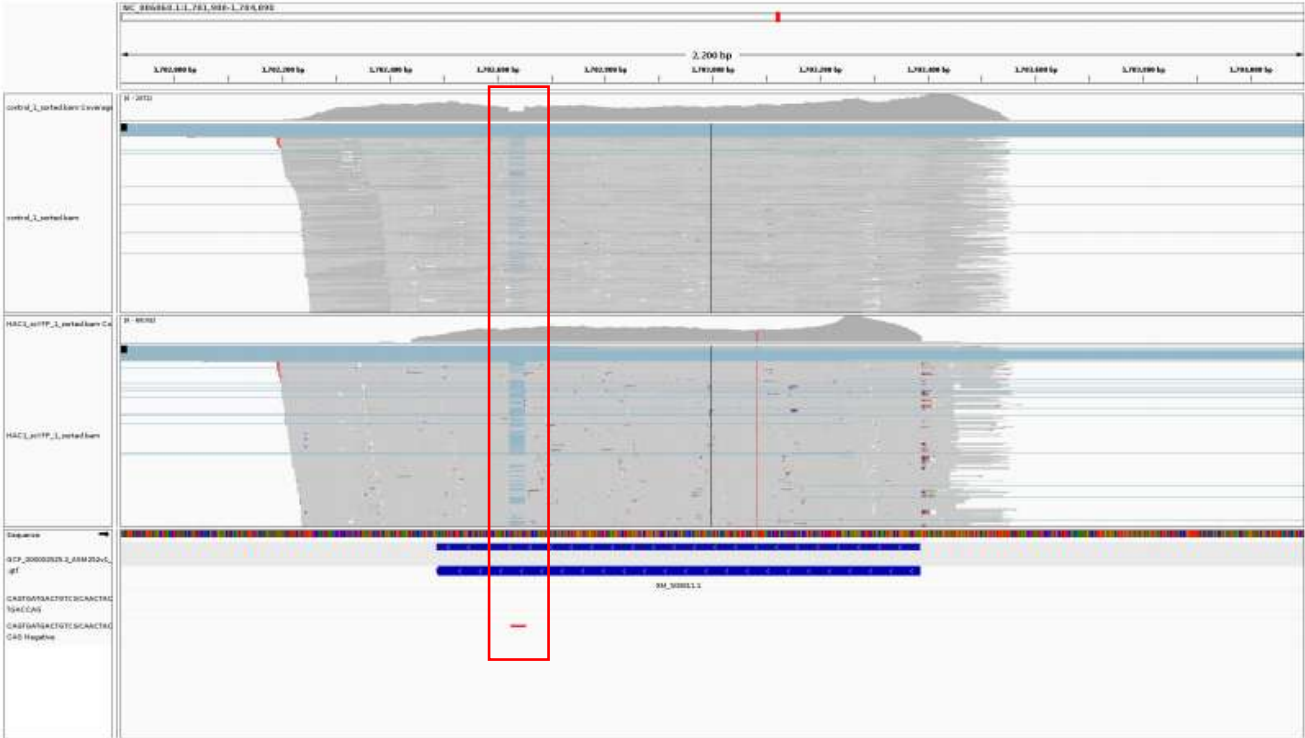
Supplementary Figure S2. Growth and morphology of *Y. lipolytica* strains used in this study. Strains were cultured over 22 h in liquid YPG medium, decimally diluted, and spotted onto a YPD agar plate. After 48 h, the images were taken in white, visual light and under UV irradiation. Microscopic images of washed biomass were taken under magnification 1000x, using an Axiovert 200M light fluorescent microscope (ZEISS, Oberkochen, Germany) using a filter set 09, as described previously [2,3]. The images were processed and archived using AxioVision software.



Supplementary Figure S3. Kinetics of Po1f-derived *HAC1_scYFP* *Y. lipolytica* strain culture, including the batch and continuous stages; X-axis: Culturing time (h); Y axis: glycerol ([g L⁻¹]; GLY), erythritol [g L⁻¹]; ERY), mannitol ([g L⁻¹]; MAN), citric acid ([g L⁻¹]; CA) and α-ketoglutaric acid ([g L⁻¹]; a-KG) concentrations.



Supplementary Figure S4. Visualization of HAC1 sequence mapping to the reference genome using IGV software. The blueish block within the reads alignment (also seen as the red bar below the alignments) indicates the intron position, which also corresponds to lower coverage of this element in the sequencing. The predicted splicing site was framed in red.



Supplementary Figure S5. Amino acid sequences and multiple sequence alignments of Hac1 and Xbp1 from *S. cerevisiae* and *Y. lipolytica*; and Xbp1 from *Homo sapiens*. Sequences were retrieved from GRYC and UniProt database. Alignment was conducted using MultAlin [4].

A. Amino acid sequences

```
>YALI0F16511g1_1_Xbp1_Yl
MTTQEPGFLATSSVSPANREFFVYAVNWPIYTIIPSCLIIPNDQKELVPSPEISIKKRYAT
SVDERNYLTVEEYQINNQWI IWDYHTGYVHLTGLWKAIGNSKADIVKLIIDNSPDLEAVIR
RVRGGYLYKIQGTWVPYDIARALASRTCYFIRFALIPFLGQDFPGTCLKPHEPGFGYLMQH
FNSTKKRRKSNAKAQQOGLAHATSPTQHSHTSPSHSPHSHMIHSQSVPHMAQHMSLQQ
HVAHSQQQQQHQHMHMHQHHMVHSHVSHPHHIQPMQSPRYPIQPAQQONQAQQSRLS
QSAPAHVSYAGSHSNVTPVPI SPAPAASASNPYYHVTPPSPHSQAVLLPRISSMPSPQQL
PHATRKYSNSKTRPILIRPNSDDCAVKTEGDEEMLSQQHTPPYTTYVSRRLPSLGGIAKPG
NAVRPRRKSRLSYDESQTSLLSAAALRSSSSVSGSAVTCSSSESEQDDDEYYRNFARRQS
ISEYNPANPPKQESPEEIMEVLQAIRSLQQLSTGGGKPGKVKMDINSVLS*
```

```
>YIL101C1_1_Xbp1_Sc
MKYPAFSINSDTVHLTDNPLDDYQRLYLVSVLDRDSPASFSAGLNIRKVNKYSSIAAQF
THPNFII SARDAGNGEEAAAQNVLNCFEYQFPNLQTIQSLVHEQTLLSQLASSATPHSAL
HLHDKNILMGKII LPSRSNKTPVSASPTKQEKALSTASRENATSSLTKNQFKLTKMDH
NLINDKLINPNNCVIWSDSGYVFMTGIWRLYQDVMKGLINLPRGDSVSTSQQQFFCKAE
FEKILSFCFYNHSSFTSEESSVLLSSSTSSPKRRTSTGSTFLDANASSSSTSSTQANN
YIDFHWNNIKPELRDLICQSYKDFLINELGPDQIDLPLNLPANFTKRIRGGYIKIQGTWL
PMEISRLLCLRFCCPIRYFLVPIFGPDFPKDCESWYLAHQNVTFASSTTGAGAATAATAA
ANTSTNFTSTAVARPRQKPRPRPRQRSTSMSSKAQKLVIEDALPSFDSFVENLGLSSND
KNFIKNSKRQKSSTYTSQTSSPIGPRDPTVQILSNLASFYNTHGHRYSYPGNIYIPQQR
YSLPPPQNLSSPQRQLNYTYDHIHPVPSQYQSPRHYNVPSSPIAPAPPTFPQPYGDDHYH
FLKYASEVYKQONQORPAHNTNTNMDTSFSPRANNSLNNFKFKTNSKQ*
```

```
>P17861|XBP1_HUMAN
MVVVAAAPNPADGTPKVLVLLSGQPASAAGAPAGQALPLMVPAQRGASPEAASGGLPQARK
RQRLTHLSPEEKALRRKLNKRVAAQTARDRKKARMSELEQQVVDLEENQKLLLENQLLR
EKTHGLVVENQELRQRLGMDALVAEEEEAEAKGNEVRPVAGSAESAALRLRAPLQQVQAQL
SPLQNI SPWILAVLTLQIQSLISCWAFWTTWTQSCSSNALPQSLPAWRSSQRSTQKDPVP
YQPPFLCQWGRHQPSWKPLMN*
```

```
>YALI0B12716g2_1_Hac1_Yl
MSIKREESFTPTPEDLGSPLTADSPGSPESGDKRKDLTLPLPAGALPPRKRAKTENEKE
QRRIERIMRNRQAAHASREKKRRHLEDLEKKCELSSENNDLHHQVTESKKTNMHLMEQH
YSLVAKLQQLSSLVNMKSSGALAGVDVPDMSDVSMAPKLEMPAAPSQPMGLASAPTLF
NHDNETVVPDSPIVKTEEV DSTNFFLHTESSPPELAESTGSGSPSSTLSCDETDYLVDR
ARHPAVMTVATTDQQRHKISFSSRTSPLTTS LDCMDCRMTSPCLKTTSSLPSTLLLI*
```

```
>YFL031W1_1_Hac1_Sc
MEMTDFELTSNSQSNLAIPTNFKSTLPPRKRAKTKEEKEQRRIERILRNRRAAHQSRK
RLHLQYLERKCSLLENLLNSVNLEKLADHEDALTC SHDAFVASLDEYRDFQSTRGASLDT
RASSHSSSDTFTPSPLNCTMEPATLSPKSMRDSASDQETSWELQMFKTENVPESTTLPAV
DNNNLFDAVASPLADPLCDDIAGNSLPFDNSIDL DNWRNPEAQSGLNSFELNDFFIT*
```

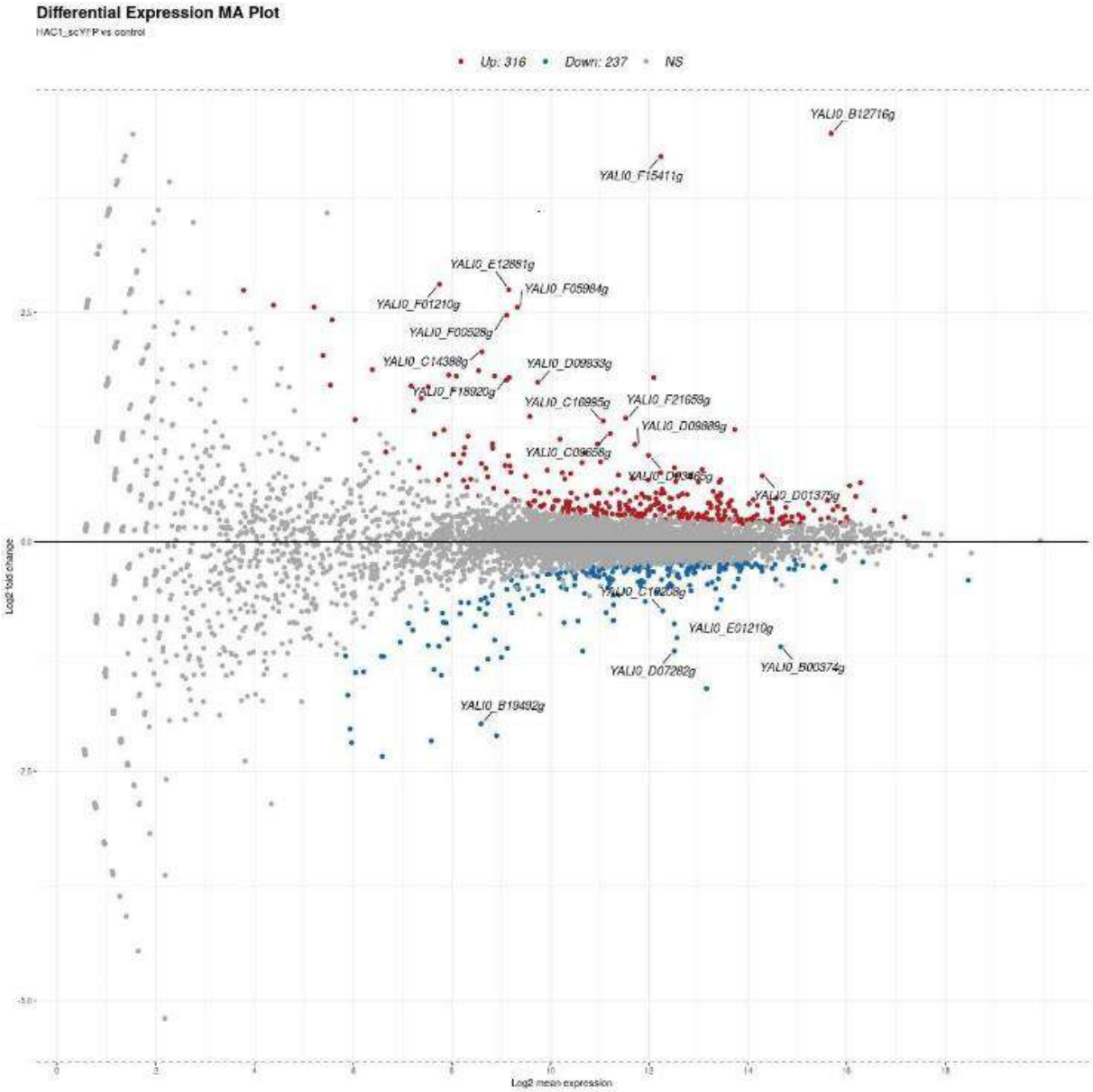

B.IV) Xbp1 from *Y. lipolytica*, *S. cerevisiae*, and *Homo sapiens*

	1	10	20	30	40	50	60	70	80	90	100	110	120	130
YALIOF16511g1_1_Xbp1	MTTQEPGLATSSVSPANREFPYAV-NHPYITIPSCLIIPN--DQKELVPSPEISIKKRYATSDVERNYLT-													
YIL101C1_1_Xbp1_Sc	MKYPAFSINDSVDLTONPLDDYQRLYLVSVDLRDSDPPASFSAGLNIRKVNKYKSSIAQFTHPFLISARDAGNGEERAQNVNCFEYQFNLQTTQSLVHEQTLLSQLASATPHSLALHDKNLLMG													
P178611XBP1_HUMANp.f.....e.....v.n...y..p...i...l...s.....a.....n.l..													
YALIOF16511g1_1_Xbp1	-VVEYQINNQIIMDYHGYVHLTGLKATIGNSKADIVK-----LIDNS--PDLFAVIRRVGGYLKIQGTWVPPY-DIARAL-----ASRTCYFIRFALIPFGQDFPGTCLKPHEPG													
YIL101C1_1_Xbp1_Sc	KIILPSSRNKTPVSRPTKQEKKALSTASRENATSLTKNQKFLTKMDNLLINDKLNPNNCYIHSADSGYVFTGIDRLYQVHKGLINLPRGDSVSTQQQFCKREFEKLSFCFYNHSSFTSEES													
P178611XBP1_HUMANn.....t.....n.....k.....li.....vi.....gy...g.w..y.d...l...s...f.....f.....e..													
YALIOF16511g1_1_Xbp1	FYGLQAHFNSTKKRKSNARQQGLAHATSPTQHSPTSPSHSPHSHHSQSVPHHAQHHS-LQHVYHSQQQQQHAAHQHQHHYSHSHVSHPHHIQPHQSPRYYPQPAQQQQAQSQSLSQS													
YIL101C1_1_Xbp1_Sc	SSVLLSSSTSPKRRSTGTSTGLD--ANSSSSTSTQANNYIDFHMMNKPELRDLICQSYKDFLINELGPDQIDLPLNLPANFTKR--IRGGYIKIQGTALPHEISRLCLLRFCEPIRYFLVPIFGPD													
P178611XBP1_HUMAN	...l...s...r.....a.a.s...s.....h...i...q...l...q.....n.....pm...r.....v.....													
YALIOF16511g1_1_Xbp1	APAHV-SYAGSHNVTVPVISPAPASASNPYYHYTPSPHSQAVLLPRISMSPPQLPHATRYKSNKTRPILIRPNSD---DCAVKTEGDEEHLQQHTPPYTYVSRPRLGGIAKPGNARVPRR													
YIL101C1_1_Xbp1_Sc	FPKDCESMHLAQNYVTFASSTTGAGATAATAANTSTNFTSTAVARQKPRPRQRQSTSMHSKQKLVIEDALPSDFSFVEMGLSSNDKQNFIKKSKRQKSSYTSQTSSTPQIPROPTVQILSNL													
P178611XBP1_HUMAN	NPDGTPKYLKLLSGQPASAGAPAGQALPLMYPARQASPEASGGLPQAKKQRLTHLSPEEKALRRKLNRYAQTARDKKAHSELEQQVVDLEENQKLLLENQLREKTH--GLVVENQELRQRL													
Consensus	.Pad..s..l.hsnvt.....agaa...a.t..sp.s.av.lPr...rprp.qls.....k.....a.p.D.....#.....#.....ty.sr..s.G..p...r.r.l													
YALIOF16511g1_1_Xbp1	KSRLSYDESQTSLLSARLSSSSVGSAGVTCSSSESEQDDDEY--YRNFARRQSTSEYVNPANPKQSPPEEIMEVL--QAIRSLQQLSTGGGKPGKPKYMDINSVLS													
YIL101C1_1_Xbp1_Sc	ASFYNTGHRYSPGNITPQRYSLPPNQLSSPQRLLHYTDHHPVPSQYSPRHYVPSPTAPAPPTFPQPYGDHYHFLKYASEYKQDNQPAHNTNTNMDTFSPPRANNSLNFKFTNSKQ													
P178611XBP1_HUMAN	GMDALVAEEAEAKGNEVYPVAGSAESALRLRAPLQVQALQSLQINSPHILAVLTLQIQSLISCAH--FATTTATQSCSSNALPQSLPAMRSSDSTQKOPVYPQPFPCQGRHOPSHKPLHN													
Consensus	.s.....e...s..gn..rp...s.sa...lssp..Q.#..y...n.....qs...y#..s.p...ap.f.....l...S.....fr.....n.....f.....k...n...													

B.V) Xbp1 from *H. sapiens* and Hac1 from *Y. lipolytica* and *S. cerevisiae*

	1	10	20	30	40	50	60	70	80	90	100	110	120	130
P178611XBP1_HUMAN	MNVVRAAPNPDGTPKYLKLLSGQPASAGAPAGQALPLMYPARQASPEASGGLPQAKKQRLTHLSPEEKALRRKLNRYAQTARDKKAHSELEQQVVDLEENQKLLLENQLREKTHGLVVEN													
YALIOB12716g2_1_Hac1	MSIKREESFTPTPEDL---GSPLTADSPGSPESGDKR--KKDLTLPLP--AGALPP--RKRRAK--TENEKEQRRTERIRNRQAHAHSEKRRHLEDLEKCELSSENNDLHHQVLESKKTMMHLMQEH													
YFL031M1_1_Hac1_Sc	MEHTDFELTNS--QSNLAIPNTNFSTLPP--RKRRAK--TENEKEQRRTERIRNRQAHAHSEKRRHLEDLEKCELSSENNDLHHQVLESKKTMMHLMQEH													
Consensustp..l...g.p..a.....#.....la.p.....g.lfp..RKRak..T...eke#rrieRi#rNR..AAH.s#kKr.h#.LE.kcs.Le.en#.l.l#.....e...l...h													
P178611XBP1_HUMAN	QELRQRLGMDALV---AEEEREAKGNEVYPVAGSAESALRLRAPLQVQALQSLQINSPHILAVLTLQIQSLISCAH--FATTTATQSCSSNALPQSLPAMRSSDSTQKOPVYPQPFPCQGRHOPSHKPLHN													
YALIOB12716g2_1_Hac1	YSLVARKLQQLSSLVNMAKSSGALAGYDVPDMSDYSHAPKLEHPTARPSPOMGLASAPTLFNHNDNETVYVPSPIVKTVEVDSTNLLHTESSPPELAESTGSGSPSTLSCDETQYLDVDRARHAPVNTVA													
YFL031M1_1_Hac1_Sc	DAFVRLDEYRDF---QSTRGASLDRASSHSSSDFTFSPNLCTNEPATLSPKSNRDSASDQETSMEIQMFKTENVPSTLLPVDNNDLFDAY---ASPLADPLCDDIAGNSLPFONSIDLNDM													
Consensus	..lva.l.....a.s...ag.#vr..s.ss...l..p.....pa.lsp.....#..q..kte.v..tt.l...sn.lp#.....sp.....scd.t...p...p.....													
P178611XBP1_HUMAN	RHQPSHAKPLHN													
YALIOB12716g2_1_Hac1	TTDQQRKIKISFSSRTSPLTSLDCHDCRHTSPCLKTSTSLPSTLLLI													
YFL031M1_1_Hac1_Sc	RNPERQSGLSFELNDFFITS													
Consensus	r.....l.sf.....t.....													

Supplementary Figure S6. Volcano plots showing the dependence of the logarithmic changes of expression fold [\log_2 fold change] values from the logarithmic values of mean expression [\log_2 mean expression] of the *HAC1_scYFP* strain vs control. Statistically significant down-regulated genes are marked as blue dots and up-regulated genes are marked as red. The chart was prepared using the EnhancedVolcano function in the EnhancedVolcano package (R program).



Supplementary Table S1. Strains used in this study.

Strain	Abbreviation	Short description	References
GGY231	Control	Reference, no synthesis of heterologous protein	[5]
GGY249	<i>inYFP</i>	Over-expression of <i>YFP</i> ; limited synthesis of in-YFP; subjected to native <i>HAC1</i> up-regulation and UPR	
GGY251	<i>scYFP</i>	Over-expression of <i>SCYFP</i> ; high-level synthesis of sc-YFP; not subjected to UPR	
ECY212	<i>HAC1_scYFP</i>	Co-over-expression of <i>HAC1</i> and <i>SCYFP</i> ; high-level synthesis of sc-YFP; subjected to UPR	[6]

Supplementary Table S2. Nucleotide sequences of the genes cloned in this study (*HAC1* and *SP1_YFP*)

Gene	Nucleotide sequence 5' → 3'
>HAC1 (YALIO_B12716g)	<p>ATGTCTATCAAGCGAGAAGAGTCCTTTACTCCCACCCCGAGGACCTGGGATCTCCCCTG ACAGCTGATTCTCCTGGCTCTCCCGAGTCTGGAGACAAGCGAAAGAAGGATCTCACTCTG CCCCTTCCTGCTGGTGTCTTCCCCCTCGAAAGAGAGCTAAGACAGAGAACGAAAAGGAG CAGAGACGCATCGAGCGGATCATGCGAAACCGGCAGGCGGCACATGCGTCTCGAGAGAAG AAGCGACGACATTTGGAGGACCTGGAGAAGAAGTGCTCGGAGTTGTCGTCCGAAAACAAC GATCTACACCACCAGGTGACTGAGTCCAAGAAGACCAACATGCACCTCATGGAACAACAC TACTCGCTGGTGGCCAAGCTGCAGCAGCTCTCGTCGCTCGTCAACATGGCCAAGTCTTCC GGAGCTTTGGCCGGCGTTGATGTCCCCGACATGAGCGATGTGTCTATGGCCCCAAGTTG GAGATGCCACCGCGGCTCCTTCCCAGCCCATGGGTCTCGCCAGCGCGCCACCCTCTTC AACCACGATAATGAGACCGTCGTCCCCGACTCTCCTATTGTGAAGACCGAGGAAGTCGAC TCTACAAACTTTCTCCTCCACACGGAGTCTCCTCCCCCCCCGAACTAGCTGAGAGCACT GGCTCAGGCTCGCCATCGTCTGACTCTGCCTGCGACGAAACTGATTATCTTGTGGACCGG GCGCGTCATCCAGCAGTGATGACTGTCGCAACTACTGACCAGCAGCGTCGGCACAAGATT TCATTTTCATCAAGGACGAGCCCGTTGACGACGAGCTTGGACTGCATGGACTGTCTGGATG ACTTACCCTGTTTGAAGACAACAAGCAGCCTGCCAGCAGCACTTTATTGCTGATCTAG CTCACTATGAGTCGAGCGTCAGCAACCTGTTTGGGGATTGGAGTGA</p>
>SP1_YFP	<p>ATGAAGTTCACATTTGCTGCCGTTACCGCCGCGCTGGCCTCGTCCGCCATGGCC GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGGCCATCCTGGTTCGAGCTGGAC GGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTAC GGCAAGCTGACCCTGAAGTTCATCTGCACCACCGCAAGCTGCCCGTGCCCTGGCCCACC CTCGTGACCACCTTCGGCTACGGCCTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAG CAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTC TTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTG GTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCAC AAGCTGGAGTACAACACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAAC GGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCC GACCACTACCAGCAGAACACCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCAC TACCTGAGCTACCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTC CTGCTGGAGTTCGTGACCGCCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAA</p>