



Universitat Autònoma de Barcelona

ADVERTIMENT. L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons:  http://cat.creativecommons.org/?page_id=184

ADVERTENCIA. El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons:  <http://es.creativecommons.org/blog/licencias/>

WARNING. The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license:  <https://creativecommons.org/licenses/?lang=en>



Universitat Autònoma de Barcelona

Escola d'Enginyeria

Departament d'Enginyeria Química, Biològica i Ambiental

**L'increment de la dosi gènica com a estratègia de millora de la
producció de proteïnes recombinants al llevat *Pichia pastoris*:
revertint les limitacions genètiques del sistema biològic**

Memòria per a obtenir el grau de Doctor
per la Universitat Autònoma de Barcelona,
sota la direcció del Dr. Joan Albiol i el Dr. Pau Ferrer

per

Elena Cámara Rey

Bellaterra, setembre 2016

El Dr. Joan Albiol i Sala i el Dr. Pau Ferrer i Alegre, ambdòs professors agregats del Departament d'Enginyeria Química, Biològica i Ambiental, i membres del grup de recerca d'Enginyeria de Bioprocessos i Biocatàlisi Aplicada de la Universitat Autònoma de Barcelona

CERTIFIQUEM

Que la biotecnòloga Elena Cámara Rey ha dut a terme sota la nostra direcció el treball que, amb títol: **L'increment de la dosi gènica com a estratègia de millora de la producció de proteïnes recombinants al llevat *Pichia pastoris*: revertint les limitacions genètiques del sistema biològic**, es presenta en aquesta memòria i constitueix la seva Tesi per optar al Grau de Doctor per la Universitat Autònoma de Barcelona en el programa de doctorat de Biotecnologia.

I per tal que se'n prengui coneixement i consti als efectes oportuns, signem la present a Bellaterra, 15 setembre del 2016,

Dr. Joan Albiol i Sala
(Co-tutor)

Dr. Pau Ferrer Alegre
(Co-tutor)

Elena Cámara Rey
(Autora)

RESUM

La producció de proteïnes recombinants ha esdevingut la opció predilecta per la producció de productes biofarmacèutics i enzims industrials. Com a factories cel·lulars, els llevats presenten nombroses avantatges enfront d'altres microorganismes, com la seva capacitat per realitzar modificacions post-traduccional a les proteïnes, l'ampli ventall d'eines genètiques i la facilitat per créixer a altes densitats, entre d'altres.

En aquest context, la idoneïtat del llevat metilotròfic *Pichia pastoris* en la producció de proteïnes recombinants ha destacat per sobre del llevat clàssic *Saccharomyces cerevisiae*, presentant una rellevància creixent en la producció de productes comercials recombinants. Entre d'altres factors, la presència del promotor *AOX1* (P_{AOX1}), d'expressió elevada i fortament regulat pel metanol, ha contribuït a la popularitat d'aquest sistema d'expressió. Per tal de millorar la producció de proteïna recombinant, s'han provat diverses estratègies, sent l'increment de la dosi gènica una de les més populars. Tot i que es pot obtenir una millora de la productivitat heteròloga amb aquesta metodologia, elevades dosis gèniques normalment comporten una sobrecàrrega metabòlica que deriva en productivitats baixes.

Durant el transcurs d'aquesta tesi, s'ha utilitzat la lipasa de *Rhizopus oryzae* (Rol) com a proteïna model, expressada sota el control de P_{AOX1} . Primer, es va construir una col·lecció de soques de *P. pastoris* amb un nombre creixent de còpies de *ROL*, i es va descriure per primer cop l'ús de la ddPCR (de l'anglès Droplet Digital PCR) per la determinació acurada de la dosi gènica a *P. pastoris*, demostrant una precisió més elevada que amb l'ús de la qPCR (de l'anglès quantitative real-time PCR) convencional. Segon, es va realitzar un estudi transcriptòmic mitjançant microarrays d'un grup de soques multicòpia, per tal d'entendre millor l'impacte de la dosi gènica de *ROL* a les rutes metabòliques. Concretament, les dades van revelar una forta regulació negativa de les rutes relacionades amb el metanol (entre d'altres), en consonància amb les dades fisiològiques obtingudes, que mostraven que les taxes de consum de la font de carboni i el creixement cel·lular eren dependents del nombre de còpies. A partir d'aquestes observacions, es va modificar una soca amb 4 còpies de *ROL* per tal que expressés una variant desregulada del factor de transcripció Mxr1, principal regulador de la ruta del metanol. Les dades fisiològiques i els nivells transcripcionals dels principals gens implicats en el metabolisme del metanol van confirmar la millora en la producció de Rol i l'increment en l'expressió dels gens relacionats amb el metanol i *ROL*. Finalment, es va dur a terme un estudi comparatiu a nivell transcriptòmic de la resposta de *P. pastoris* a l'expressió de tres proteïnes recombinants diferents, per tal d'estudiar l'impacte de la complexitat proteica i l'ús de sistemes d'expressió constitutius o induïbles.

En resum, el present estudi descriu per primera vegada una aproximació a nivell de sistema de l'impacte d'una dosi gènica creixent en el llevat *P. pastoris*, permetent així el disseny de noves estratègies d'enginyeria cel·lular per tal de generar soques més productores, mitjançant la biologia sintètica i l'enginyeria metabòlica.

ABSTRACT

Recombinant protein production is becoming the preferred option to synthesized biopharmaceutical and industrial enzymes. As a host system, yeast factories provide numerous advantages over other microorganisms, such as their capacity to perform post-translational protein modifications, the wide genetic available tools, and the facility to growth at high densities, among others.

In this context, the suitability of methylotrophic yeast *Pichia pastoris* for recombinant protein production has stand out over the classical yeast *Saccharomyces cerevisiae*, presenting an increasing relevance for the production of commercial recombinant products. Among several factors, the presence of the strong and tightly regulated *AOX1* promoter (P_{AOX1}) has contributed to the popularity of this expression system. To further increase the recombinant protein production, several strategies have been attempted, being the increase of the heterologous gene dosage one of the most widely reported. Although a heterologous productivity improvement can be achieved with this methodology, high gene dosages normally entail a metabolic burden in the cell that results in lower productivities.

In the course of this thesis, it has been used the lipase of *Rhizopus oryzae* (Rol) as a protein model expressed under the control of the P_{AOX1} . Firstly, we constructed a series of *P. pastoris* strains harbouring an increase gene dosage of *ROL*, and we reported for the first time the use of Droplet Digital PCR (ddPCR) for an accurate copy number determination in *P. pastoris*, allowing a higher precision than using conventional quantitative real-time PCR (qPCR). Secondly, a transcriptomic study of a set of multicopy strains was carried out using microarrays, to better understand the impact of *ROL* gene dosage on the metabolic pathways. Concretely, transcriptomic data revealed a high downregulation of methanol-related pathways (among others), in agreement with the physiological data that had been showed that C-source uptake rate and growth yields were gene dosage dependents. Based on these observations, a strain carrying 4 copies of *ROL* was remodified to express a deregulated variant of the Mxr1 transcription factor, the main regulator of methanol metabolism. Physiological data and transcriptional levels of the main genes involved in methanol metabolism confirmed the improvement of Rol production and the increasing expression levels of methanol-related genes and *ROL*. Finally, it was performed a comparative transcriptomic analysis of the response of *P. pastoris* to the expression of three different recombinant proteins, to study the impact of protein complexity and the use of constitutive or inducible expression systems.

Overall, in the present study we describe the first systems-level approach to determine the impact of increasing gene dosage in the yeast *P. pastoris*, hence allowing the design of novel cell engineering strategies to generate high-producing strains by means of the synthetic biology and metabolic engineering.

ÍNDEX

Resum	I
Abstract	II
Índex	III
1. Introducció general	1
1.1 La producció de proteïnes recombinants: un mercat en expansió	1
1.2 <i>Pichia pastoris</i> com a sistema d'expressió	2
1.3 El metabolisme del metanol de <i>P. pastoris</i>	5
1.4 El sistema P_{AOX1} per l'expressió de proteïnes recombinants	5
1.5 Referències	7
Objectius	10
2. Construcció i caracterització de soques multicòpia de <i>ROL</i> de <i>Pichia pastoris</i>	11
2.1 Introducció	12
2.2 Materials i mètodes	15
2.2.1 Soques i plasmidis	15
2.2.2 Determinació del nombre de còpies de <i>ROL</i> per ddPCR	15
2.2.3 Determinació del nombre de còpies de <i>ROL</i> per qPCR	16
2.2.4 Anàlisi per citometria de flux	17
2.2.5 Procediments analítics	17
2.2.5.1 Anàlisi de biomassa	17
2.2.5.2 Assaig d'activitat lipolítica	17
2.2.5.3 Quantificació dels metabòlits extracel·lulars	18
2.2.6 Condicions de cultiu	18
2.2.6.1 Cultius en microplaca	18
2.2.6.2 Cultius en Erlenmeyer	19
2.2.6.3 Cultius en reactor operats en continu	19
2.2.7 Anàlisi estadístic	20
2.3. Resultats i discussió	21
2.3.1 Construcció d'una col·lecció de soques amb diferent nombre de còpies de la lipasa de <i>R. oryzae</i> (<i>ROL</i>)	21
2.3.2 Comparació de la qPCR vs ddPCR per a la determinació del nombre de còpies de <i>ROL</i>	22
2.3.3 Estudi dels clons transformants en plaques de 24 pous i cultius en Erlenmeyer: correlació de la dosi gènica amb el creixement cel·lular i la producció de lipasa	24

2.3.4 Validació del comportament dels clons seleccionats en cultius operats en continu	27
2.4 Conclusions	30
2.5 Referències	31
3. L'increment de la dosi gènica dels cassets d'expressió regulats pel promotor AOX1 comporta una atenuació dels nivells transcripcionals del metabolisme del metanol	37
3.1 Introducció	38
3.2 Materials i mètodes	40
3.2.1 Construcció de soques	40
3.2.2 Cultius operats en continu	40
3.2.3 Procediments analítics	41
3.2.3.1 Determinació de la biomassa	41
3.2.3.2 Determinació de l'activitat lipolítica	41
3.2.3.3 Disrupció cel·lular	41
3.2.3.4 Determinació de la proteïna total	41
3.2.3.5 Quantificació dels metabòlits	42
3.2.3.6 Composició d'àcids grassos	42
3.2.4 Quantificació de Rol per Western Blot	42
3.2.5 Determinació dels nivells transcripcionals de <i>ROL</i> per ddPCR	43
3.2.6 Determinació del nombre de còpies de <i>ROL</i> per ddPCR	43
3.2.7 Anàlisi estadístic	44
3.2.8 Extracció del RNA i hibridació dels microarrays	44
3.2.9 Anàlisi transcriptòmic	45
3.2.9.1 Diagrames de Venn	45
3.2.9.2 Anàlisi de clústers	45
3.2.9.3 Anàlisi Principal de Components	45
3.2.9.4 Enriquiment funcional	45
3.3 Resultats i discussió	47
3.3.1 Dosis gèniques altes de <i>ROL</i> provoquen reduccions al creixement cel·lular i al consum de metanol	47
3.3.2 Els nivells baixos de <i>ROL</i> a les soques multicòpia podrien ser deguts a una transcripció ineficient	48
3.3.3 Perfil transcripcional de les soques multicòpia	51
3.3.4 Enriquiment funcional	53
3.3.5 Els gens d'utilització del metanol (<i>Mut</i>) presenten una expressió reduïda a mesura que augmenta el nombre de còpies de <i>ROL</i> , suggerint un impacte per la disponibilitat limitada de <i>Mxr1</i>	56
3.3.6 La biogènesi dels peroxisomes es regula negativament a les soques amb alta dosi gènica de <i>ROL</i>	60
3.3.7 El metabolisme dels àcids grassos es troba significativament alterat, mostrant una regulació positiva de les rutes biosintètiques i la β -oxidació a les soques productores de Rol	62

3.3.8 L'impacte de la dosi gènica de <i>ROL</i> al metabolisme dels aminoàcids, el metabolisme energètic i del glicerol, la síntesi, plegament, secreció i degradació de proteïnes és limitat o no significatiu	65
3.4 Conclusions	67
3.5 Referències	68
4. La desregulació del metabolisme del metanol reverteix les limitacions transcripcionals de les soques recombinants de <i>P. pastoris</i> amb múltiples cassets d'expressió sota el control de P_{AOX1}	75
4.1. Introducció	76
4.2. Materials i mètodes	78
4.2.1 Soques i plasmidis	78
4.2.2 Construcció del vector d'expressió pGAPHA_MXR1*	78
4.2.3 Transformació de <i>P. pastoris</i> i selecció de clons	78
4.2.4 Condicions de creixement	79
4.2.4.1 Cultius en Erlenmeyer	79
4.2.4.2 Cultius operats en continu	79
4.2.5 Procediments analítics	80
4.2.5.1 Determinació de la biomassa	80
4.2.5.2 Assaig d'activitat lipolítica	80
4.2.5.3 Quantificació dels metabòlits	80
4.2.6 Anàlisi per citometria de flux	80
4.2.7 Determinació de la dosi gènica per ddPCR	81
4.2.8 Determinació dels nivells transcripcionals per ddPCR	82
4.2.9 Anàlisi estadístic	82
4.3. Resultats i discussió	83
4.3.1 Construcció i caracterització de soques productores de Rol coexpressant Mxr1*	83
4.3.2 L'expressió de <i>MXR1</i> * incrementa els nivells de producció de Rol i reverteix l'efecte perjudicial de la sobreexpressió de <i>ROL</i> en l'assimilació del metanol a les soques multicòpia	85
4.3.3 Els cultius en continu van revelar un canvi de preferència en la font de carboni depenent de la soca	87
4.3.4 L'anàlisi transcripcional va confirmar una sobreexpressió dels gens clau de la ruta del metanol a la soca 4C_2MXR1*	88
4.4 Conclusions	93
4.5 Referències	94

5. Anàlisi transcriptòmic comparatiu de la resposta fisiològica de <i>P. pastoris</i> a l'expressió de tres proteïnes recombinants diferents	99
5.1 Introducció	100
5.2 Resultats i discussió	101
5.2.1 Caracterització fisiològica de les soques productores	101
5.2.2 Anàlisi transcripcional global	102
5.2.3 La font de carboni té un impacte major en l'expressió gènica que els nivells de producció recombinant	104
5.2.4 Anàlisi de <i>heatmaps</i>	106
5.2.5 La producció de proteïna recombinant disminueix les vies de reproducció sexual i afavoreix la biogènesi ribosomal i la síntesi d'aminoàcids a les soques P_{GAP}	107
5.2.6 Els nivells de complexitat proteica tenen un impacte mínim sobre el metabolisme global i l'estrès quan no es detecta proteïna acumulada	110
5.2.7 Els nivells elevats d'albúmina de sèrum humà podrien ser deguts a una sobreexpressió de gens relacionats amb l'inici de la transcripció i la maquinària de reparació de DNA	111
5.2.8 La limitada anotació funcional del genoma de <i>P. pastoris</i> dificulta l'anàlisi de diversos grups d'interès	113
5.2.9 La comparativa dels perfils transcriptòmics de les soques d'alta dosi gènica envers les de la baixa dosi pot proporcionar diverses dianes per futures estratègies d'enginyeria genètica	114
5.3 Referències	117
6. Conclusions generals i perspectives futures	121
7. Contribucions científiques	123
8. Annex	125

1. INTRODUCCIÓ

1.1 La producció de proteïnes recombinants: un mercat en expansió

Segons la Federació Europea de la Biotecnologia (EFB), la biotecnologia es defineix com “l’ús integratiu de la bioquímica, la microbiologia i l’enginyeria per tal d’aconseguir aplicacions tecnològiques (industrials) de les característiques dels microorganismes, dels cultius de teixit cel·lular o parts d’aquests”. És a dir, l’obtenció de bens i serveis mitjançant eines biològiques (Bull et al., 1982).

Des dels anys 80, el desenvolupament de l’expressió de proteïnes heteròlogues per part de diferents microorganismes (bactèries, llevats, cèl·lules animals, entre d’altres) ha viscut un creixement prominent, degut els avantatges que comporta enfront la producció química de compostos d’elevada complexitat (com pèptids i proteïnes) o les limitacions econòmiques de l’extracció dels productors naturals (Mattanovich et al., 2012; Pollack, 2007). Així, des de la comercialització del primer producte recombinant, la insulina humana Humulin (comercialitzada per Eli Lilly al 1982, Change et al., 1999) fins al dia d’avui, s’han comercialitzat al voltant de 400 productes biofarmacèutics produïts mitjançant tecnologia recombinant, i 1300 productes estan en vies de desenvolupament (Sánchez-García et al., 2016).

D’altra banda, la irrupció de l’ús dels enzims en la indústria enfront dels catalitzadors químics clàssics (donada la seva elevada estereoselectivitat, condicions suaus d’operació entre d’altres avantatges; Cooper, 2000) també s’ha vist reforçada per la tecnologia recombinant. Actualment, un 60% dels enzims industrials es produeixen de manera heteròloga, per tal d’evitar el cultiu d’aquells microorganismes originals tòxics o amb baix creixement a més de poder millorar les

seves limitacions (eficiència catalítica, estabilitat) mitjançant l'enginyeria de proteïnes (Adrio i Demain, 2012; Yang et al., 2014).

D'aquesta manera, el creixent interès en la tecnologia recombinant per la síntesi de productes d'alt valor afegit s'ha de veure acompanyat de millores a diferents nivells, com els microorganismes hostes, les estratègies de producció (condicions d'operació) i l'increment dels nivells d'expressió, per tal d'aconseguir productivitats més elevades i permetre una major comercialització dels processos (Palomares et al., 2004).

1.2 *Pichia pastoris* com a sistema d'expressió

En la producció de proteïna recombinant l'elecció de l'hoste és un element clau. Entre els sistemes disponibles es poden trobar bacteris (*Escherichia coli*, *Bacillus subtilis*), llevats (*Saccharomyces cerevisiae*, *Pichia pastoris*, *Hansenula polymorpha*), fongs filamentosos (*Aspergillus Niger*), cèl·lules d'insecte (*Spodoptera frugiperda*) i cèl·lules animals (cèl·lules CHO), entre d'altres (Demain i Vaishnav, 2009). Cap sistema d'expressió és òptim per la producció de totes les proteïnes recombinants, per tant la tria de l'hoste dependrà majoritàriament de les característiques estructurals i biològiques de la proteïna a expressar, i tot i així, es requerirà una comprovació empírica per tal de determinar-ho (Tate et al., 2002; Bernaudat et al., 2011; Meyer et al., 2013).

En el cas concret dels llevats, aquests es caracteritzen per la seva capacitat de realitzar modificacions post-traduccionals típiques d'organismes superiors, combinat amb la simplicitat dels organismes unicel·lulars (fàcil manipulació genètica i ràpid creixement). Donat l'ampli coneixement sobre la genètica, la fisiologia i les diferents estratègies de cultiu, *S. cerevisiae* s'ha establert durant les darreres dècades com el llevat més àmpliament utilitzat per la producció de proteïnes recombinants (Johnson i Echavarri-Erasun, 2011). No obstant, el seu metabolisme fermentatiu, juntament amb una productivitat limitada i la hiperglicosilació de les proteïnes recombinants ha fet que darrerament la tria d'altres llevats s'hagi estès per a la producció heteròloga, com és el cas de *P. pastoris* (Mattanovich et al., 2012).

Pichia pastoris (reassignada recentment al gènere *Komagataella*; Kurtzman, 2009), és un llevat no convencional metil·lotròfic. El seu interès biotecnològic es caracteritza, principalment, pels següents aspectes:

- capacitat de créixer a altes concentracions cel·lulars, arribant als 100 g/L (de pes sec) (Cereghino i Cregg, 2000)
- disponibilitat d'un sistema genètic d'expressió de fàcil manipulació (Cregg i Madden, 1988)
- possibilitat de secretar les proteïnes heteròlogues mitjançant diverses senyals de secreció (Cereghino i Cregg, 2000; Lin-Cereghino et al., 2013)
- reconeixement d'estatus d'organisme GRAS (de l'anglès *Generally Recognized As Safe*) per part de l'Administració d'Aliments i Fàrmacs (FDA) (Ciofalo et al., 2006)

Els avantatges comentats anteriorment, juntament amb la completa seqüenciació del genoma de les principals soques utilitzades a *P. pastoris* (De Schutter et al., 2009; Mattanovich et al., 2009; Küberl et al., 2011) i les darreres millores en la seva anotació (Valli et al., 2016), fan que les avantatges que podria presentar *S. cerevisiae* enfront *P. pastoris* (bàsicament, un catàleg més ampli de tècniques moleculars i una bibliografia més extensa, Darby et al., 2012) quedessin enrere, i *P. pastoris* passés a ser el llevat preferent per la producció de proteïnes recombinants.

1.3 El metabolisme del metanol a *P. pastoris*

P. pastoris és un llevat metil·lotròfic; és a dir, pot utilitzar el metanol com a única font d'energia i carboni. La ruta d'oxidació del metanol és una via compartida a tots els llevats metil·lotròfics (Hartner i Glieder, 2006), i la primera caracterització dels enzims implicats a *P. pastoris* va ser descrita per Couderc i Baratti (1980). Breument, el metanol captat per la cèl·lula s'oxida a formaldehid i peròxid d'hidrogen, mitjançant l'enzim alcohol oxidasa. Posteriorment, aquest formaldehid pot tenir dues destinacions diferents: la via deassimilatòria o la via assimilatòria. En la primera, el formaldehid és oxidat per la deshidrogenasa depenent de glutatió Fld, i subseqüentment reacciona amb la hidrolasa Fgh1 i la format deshidrogenasa Fdh1, resultant finalment en CO₂. En la ruta assimilatòria, el formaldehid es condensa amb la xilulosa-5-fosfat per tal d'acabar formant constituents cel·lulars mitjançant l'acció de la dihidroxiacetona sintasa Das1, generant dihidroxiacetona i gliceraldehid-3-fosfat (Figura 1).

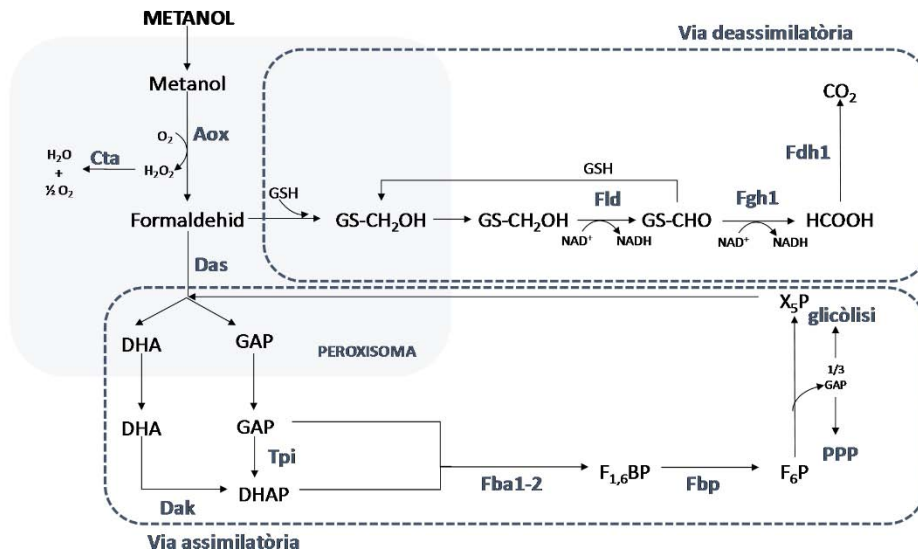


Figura 1. Ruta del metanol al llevat *P. pastoris*. Els principals enzims implicats en la metabolització del metanol es troben en blau. Aox, alcohol oxidasa; Cat, catalasa; Das, dihidroxiacetona sintasa; Dak, dihidroxiacetona quinasa; Tpi, triosafosfat isomerasa; Fba1-2, fructosa 1,6-bifosfat aldolasa; Fbp, fructosa 1,6-bifosfatasa; Fld, formaldehid deshidrogenasa; Fgh, s-formilglutaciona hidrolasa; Fdh, format deshidrogenasa; DHA, dihidroxiacetona; GAP, gliceraldehid-3-fosfat, DHAP, dihidroxiacetona fosfat; F_{1,6}BP, fructosa1,6-bifosfat; F₆P, frucosa 6-fosfat; X₅P, xilulosa 5-fosfat; PPP, ruta de les pentoses fosfat.

Donada l'alta toxicitat dels compostos que es generen en la metabolització del metanol (formaldehid, peròxid d'hidrogen), aquesta ruta té lloc als peroxisomes, uns orgànuls àmpliament conservats i de vital importància als eucariotes. Recentment, Rußmayer et al. (2015) van proposar un nou model descriptiu del metabolisme del metanol a *P. pastoris*, a partir de les senyals de localització peroxisomal de diferents gens assignats clàssicament a la ruta de les pentoses fosfat, i on s'explicaria com es regenera la xilulosa 5-fosfat utilitzada al procés. Segons aquesta hipòtesi, determinades isoformes dels enzims implicats en la ruta no oxidativa de les pentoses fosfat (la transaldolasa Tal, la ribosa 5-fosfat cetol-isomerasa Rki i la ribosa 5-fosfat epimerasa Rpe) es trobarien localitzades a l'interior del peroxisoma, i la seva expressió es veuria augmentada dràsticament en presència del metanol (a l'igual que la biogènesi dels peroxisomes i els altres enzims implicats en la ruta). La reestructuració suggerida es mostra a la figura 2, resultant en un flux net d'una molècula de gliceraldehid 3-fosfat produïda per cada 3 molècules de metanol metabolitzades.

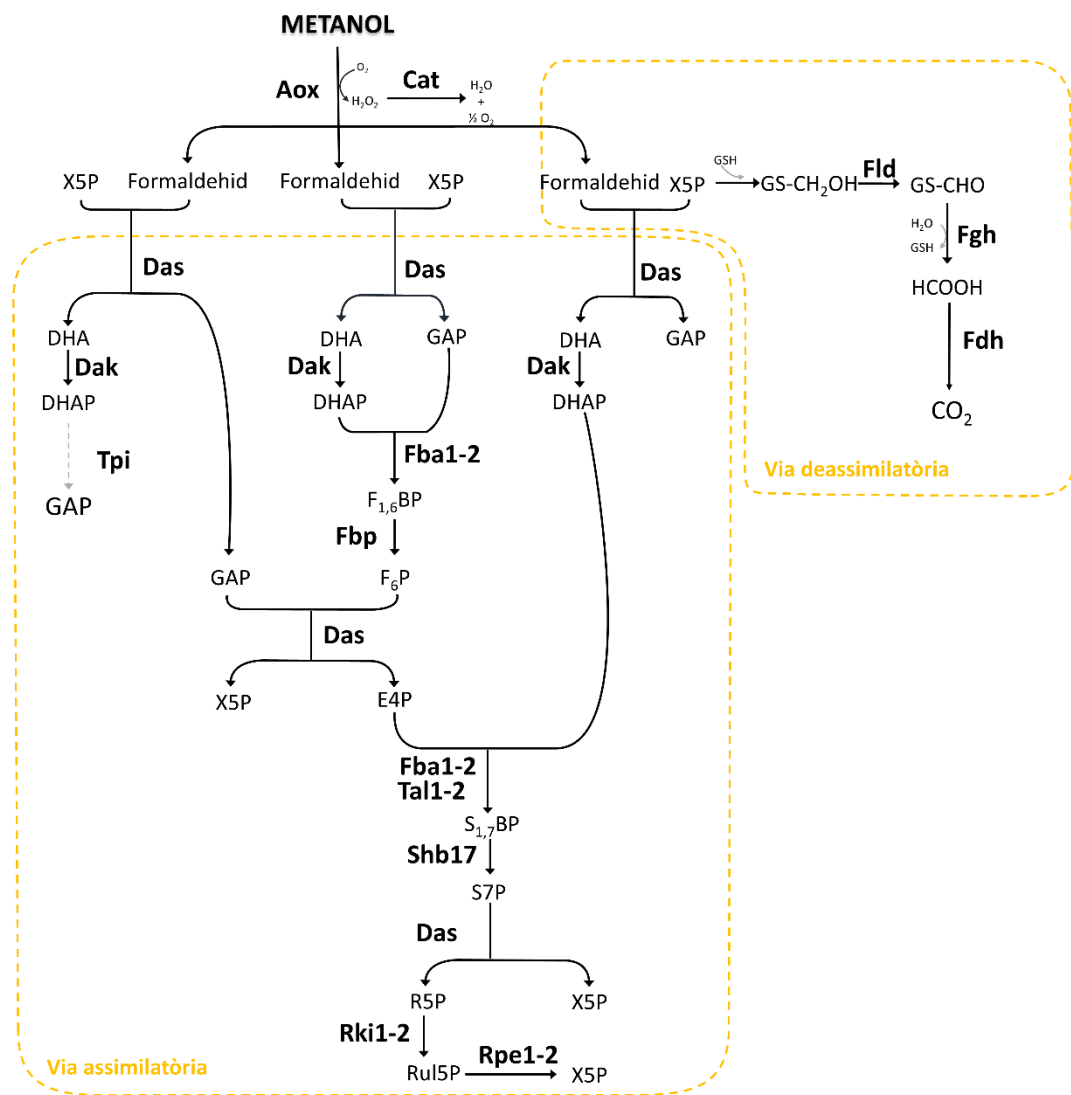


Figura 2. Ruta del metanol suggerida per Rußmayer et al. (2015). Per tal de generar una molècula de GAP, es requereixen 3 molècules de metanol. Tal1-2, transaldolasa; Shb17, sedoheptulosa 1,7-bifosfat; Rki1-2, ribosa-5-fosfat cetol-isomerasa; Rpe1-2, ribosa-5-fosfat epimerasa

1.4 El sistema P_{AOX1} per l'expressió de proteïnes recombinants

Tal i com s'ha mencionat anteriorment, l'enzim alcohol oxidasa és el responsable de la oxidació del metanol a formaldehid. Aquest enzim presenta una constant d'afinitat molt baixa pel metanol i l'oxigen, però en condicions on la única font de carboni és el metanol, els alts nivells d'expressió de l'enzim compensen la baixa afinitat (Couderc i Baratti, 1980). Aquesta característica s'ha emprat àmpliament per tal d'expressar diferents proteïnes heteròlogues sota el control del promotor *AOX1* (P_{AOX1}), resultant en el promotor més utilitzat per l'expressió de proteïnes recombinants al llevat *P. pastoris* (Vogl i Glieder, 2013).

P_{AOX1} , a banda de presentar una elevada expressió, es caracteritza per la forta regulació mitjançant diferents fonts de carboni. En presència de metanol, l'expressió d'alcohol oxidasa sota P_{AOX1} pot resultar en el 5% del mRNA total i fins un 30% de la proteïna soluble. D'altra banda, en presència de fonts de carboni repressores, com la glucosa, el glicerol o l'etanol, els nivells basals de mRNA d' $AOX1$ són mínims, mentre que l'enzim alcohol oxidasa és indetectable (Inan i Meagher).

En els darrers anys, múltiples investigacions s'han dut a terme per tal de determinar la complexa regulació de P_{AOX1} . Aquest promotor presenta una regulació a tres nivells: I) repressió, II) derepressió i III) inducció. Així, tot i l'absència de fonts de carboni repressores, el promotor necessita de la presència del metanol per tal de poder expressar-se.

1.5 REFERÈNCIES

- Adrio JL, Demain AL. 2010. Recombinant organisms for production of industrial products. *Bioeng. Bugs* 1:116–131.
- Bernaodat F, Frelet-Barrand A, Pochon N, Dementin S, Hivin P, Boutigny S, Rioux JB, Salvi D, Seigneurin-Berny D, Richaud P, Joyard J, Pignol D, Sabaty M, Desnos T, Pebay-Peyroula E, Darrouzet E, Vernet T, Rolland N. 2011. Heterologous expression of membrane proteins: Choosing the appropriate host. *PLoS One* 6.
- Bull AT, Holt G, Lilly MD. 1982. *Biotechnology: International Trends and Perspectives*. Paris: Organisation for Economic Co-Operation and Development
- Cereghino JL, Cregg JM. 2000. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol. Rev.* 24:45–66.
- Chance R, Glazer N, Wishner K. Insulin Lispro (Humalog) In: Walsh G, Murphy B, editors. *Biopharmaceuticals, an Industrial Perspective*. Kluwer: Dordrecht; 1999. pp. 149–172.
- Ciofalo V, Barton N, Kreps J, Coats I, Shanahan D. 2006. Safety evaluation of a lipase enzyme preparation, expressed in *Pichia pastoris*, intended for use in the degumming of edible vegetable oil. *Regul. Toxicol. Pharmacol.* 45:1–8.
- Cooper GM. 2000. *The Cell: A Molecular Approach*. 2nd edition
- Couderc R, Baratti J. 1980. Oxidation of Methanol by the Yeast, *Pichia pastoris*. Purification and Properties of the Alcohol Oxidase. *Agric. Biol. Chem.* 44:2279–2289.
- Cregg JM, Madden KR. 1988. Development of the methylotrophic yeast, *Pichia pastoris*, as a host system for the production of foreign proteins. *Dev. Ind. Microbiol.* 29:33–41
- Darby RA, Cartwright SP, Dilworth MV, Bill RM. 2012. Which yeast species shall I choose? *Saccharomyces cerevisiae* versus *Pichia pastoris* (review). *Methods Mol Biol* 866:11–23
- De Schutter K, Lin Y-C, Tiels P, Van Hecke A, Glinka S, Weber-Lehmann J, Rouzé P, Van de Peer Y, Callewaert N. 2009. Genome sequence of the recombinant protein production host *Pichia pastoris*. *Nat. Biotechnol.* 27:561–566.
- Demain AL, Vaishnav P. 2009. Production of recombinant proteins by microbes and higher organisms. *Biotechnol. Adv.* 27:297–306.

Hartner FS, Glieder A. 2006. Regulation of methanol utilisation pathway genes in yeasts. *Microb. Cell Fact.* 5:39.

Inan M, Meagher MM. 2001. Non-Repressing Carbon Sources for Alcohol Oxidase (AOX1) Promoter of *Pichia pastoris*. *Pap. Biochem. Eng.* 92:585–589.

Johnson et al. 2011. Yeast Biotechnology. In: *The Yeasts. A taxonomic study.* v1

Küberl A, Schneider J, Thallinger GG, Anderl I, Wibberg D, Hajek T, Jaenicke S, Brinkrolf K, Goesmann A, Szczepanowski R, Pühler A, Schwab H, Glieder A, Pichler H. 2011. High-quality genome sequence of *Pichia pastoris* CBS7435. *J. Biotechnol.* 154:312–20.

Kurtzman CP. 2009. Biotechnological strains of *Komagataella (Pichia) pastoris* are *Komagataella phaffii* as determined from multigene sequence analysis. *J. Ind. Microbiol. Biotechnol.* 36:1435–1438.

Lin-Cereghino GP, Godfrey L, de la Cruz BJ, Johnson S, Khuongsathiene S, Tolstorukov I, Yan M, Lin-Cereghino J, Veenhuis M, Subramani S, Cregg JM. 2006. Mxr1p, a key regulator of the methanol utilization pathway and peroxisomal genes in *Pichia pastoris*. *Mol. Cell. Biol.* 26:883–897.

Lin-cereghino GP, Stark C, Kim D, Chang J, Nadia S, Poerwanto H, Agari K, Moua P, Low L, Tran N, Huang A, Maria N, Oshiro K, Chang J, Chavan A, Tsai J, Lin-Cereghino J. 2013. The Effect of α -Mating Factor Secretion Signal Mutations on Recombinant Protein Expression in *Pichia pastoris* Geoff. *Gene* 519:311–317.

Mattanovich D, Graf A, Stadlmann J, Dragosits M, Redl A, Maurer M, Kleinheinz M, Sauer M, Altmann F, Gasser B. 2009. Genome, secretome and glucose transport highlight unique features of the protein production host *Pichia pastoris*. *Microb. Cell Fact.* 8:29.

Mattanovich D, Branduardi P, Dato L, Gasser B, Sauer M, Porro D. 2012. Recombinant Protein Production in Yeasts. *Methods Mol. Biol.* 824:329–358.

Meyer S, Lorenz C, Baser B, Wördehoff M, Jäger V, van den Heuvel J. 2013. Multi-Host Expression System for Recombinant Production of Challenging Proteins. *PLoS One* 8:1–9.

Palomares LA, Estrada-Mondaca S, Ramírez OT. 2004. Production of Recombinant Proteins: Challenges and Solutions. *Recomb. Gene Expr.* 267:1–39.

Pollack, P. *Fine Chemicals. The Industry and the business.* 2007

Rußmayer H, Buchetics M, Gruber C, Valli M, Grillitsch K, Modarres G, Guerrasio R, Klavins K, Neubauer S, Drexler H, Steiger M, Troyer C, Al Chalabi A, Krebiehl G, Sonntag D, Zellnig G, Daum G, Graf AB, Altmann F, Koellensperger G, Hann S, Sauer M, Mattanovich D, Gasser B. 2015. Systems-level organization of yeast methylotrophic lifestyle. *BMC Biol.* 13:80.

Tate CG, Haase J, Baker C, Boorsma M, Magnani F, Vallis Y, Williams DC. 2003. Comparison of seven different heterologous protein expression systems for the production of the serotonin transporter. *Biochim. Biophys. Acta - Biomembr.* 1610:141–153.

Yang H, Li J, Shin H-D, Du G, Liu L, Chen J. 2014. Molecular engineering of industrial enzymes: recent advances and future prospects. *Appl. Microbiol. Biotechnol.* 98:23–9.

Valli M, Totto NE, Peymann A, Gruber C, Landes N, Ekker H, Thallinger GG, Mattanovich D, Gasser B, Graf AB. 2016. Curation of the genome annotation of *Pichia pastoris* (*Komagataella phaffii*) CBS7435 from gene level to protein function. *FEMS Yeast Res.*:1–28.

Vogl T, Glieder A. 2013. Regulation of *Pichia pastoris* promoters and its consequences for protein production. *N. Biotechnol.* 30:385–404.

OBJECTIUS

L'objectiu principal d'aquest treball és l'estudi de l'impacte de l'increment del nombre de còpies d'un gen heteròleg en la fisiologia del llevat metil·lotròfic *Pichia pastoris*, utilitzant com a proteïna model la lipasa de *Rhizopus oryzae* (Rol). Tot i ser una estratègia àmpliament reportada per la millora de la producció de proteïna recombinant, sovint altes dosis gèniques acaben comportant una sobrecàrrega metabòlica a les soques, resultant en baixes productivitats de la proteïna recombinant. Per tal d'estudiar aquestes limitacions, s'han d'acomplir els següents objectius parcials:

- Construcció d'una bateria de soques amb un nombre creixent de còpies del gen *ROL*, per tal d'estudiar els efectes fisiològics de la diferent dosi gènica
- Implementació d'una tècnica de quantificació precisa del nombre de còpies que permeti la classificació acurada de les diferents soques segons la dosi gènica
- Caracterització de les soques mitjançant cultius a diferent escala de l'impacte de l'increment del nombre de còpies sobre diferents paràmetres fisiològics (creixement, producció, estrés cel·lular)
- Anàlisi transcriptòmic d'una col·lecció de soques representatives de diferents rangs de còpies per estudiar l'impacte de la dosi gènica al metabolisme general
- Disseny de noves estratègies d'enginyeria genètica a partir de les dades transcriptòmiques obtingudes per tal de revertir les limitacions de les soques
- Estudi transcriptòmic comparatiu de l'expressió de tres proteïnes diferents a *P. pastoris*, en condicions de baixa i alta dosi gènica

2. Construcció i caracterització de soques multicòpia de *ROL* de *Pichia pastoris*

Paraules clau: ddPCR, dosi gènica, *Pichia pastoris*, qPCR, lipasa *Rhizopus oryzae*, proteïna recombinant

2.1 INTRODUCCIÓ

Pichia pastoris és un hoste per la producció de proteïnes recombinants sòlidament establert, i una plataforma prometedora per a la producció de compostos d'alt valor afegit. Generalment, l'estratègia per la sobreexpressió de proteïna heteròloga en aquest hoste consisteix en la integració estable d'un vector d'expressió en el genoma mitjançant recombinació homòloga (Ahmad et al., 2014; Cregg et al., 2000). La producció de proteïnes recombinants pot ser millorada a diferents nivells, des de la transcripció o la traducció fins l'abastiment d'energia i components estructurals pel plegament i secreció de proteïnes (Bai et al., 2011; Heyland et al., 2011; Ruth et al., 2014). En aquest context, l'increment de la dosi gènica és una estratègia àmpliament utilitzada per millorar la producció de proteïnes recombinants (Aw i Polizzi, 2013). Diversos estudis descriuen com un increment del nombre de còpies resulta en una producció més elevada, tot i que un increment més enllà de l'òptim pot derivar en productivitats més baixes (Liu et al., 2014; Marx et al., 2009; Zhu et al., 2009). En el cas específic de la lipasa de *Rhizopus oryzae* (Rol), estudis previs per Jordà et al. (2012) van evidenciar una sobrecàrrega metabòlica significativa provocada per la sobreexpressió de Rol, fins i tot en aquelles soques amb una única còpia del gen *ROL*. No obstant, no es van poder realitzar anàlisis addicionals degut a la dificultat de determinar de manera precisa la dosi gènica mitjançant qPCR (de l'anglès *Quantitative real-time PCR*), concretament entre soques que diferissin en una còpia de *ROL*, però tot i així amb comportaments diferents en termes de productivitat i paràmetres de creixement (resultats no publicats del nostre grup). Negligències a l'hora de determinar el nombre de còpies poden derivar en interpretacions incorrectes dels resultats experimentals, com va ser descrit per Abad et al. (2010). Per tant, es necessita una determinació exacta de la dosi gènica en els processos de cribratge/*screening* per tal de validar la selecció de clons, així com per estudis fisiològics en condicions de procés, on l'estabilitat genètica és un factor clau. Anteriorment, el *Southern Blot* i la qPCR havien estat les tècniques més comunes a l'hora de determinar el nombre de còpies d'un gen (Abad et al., 2010; Sunga et al., 2008). Amb tot, diverses limitacions han anat en detriment d'aquestes tècniques. En el cas concret de la qPCR els problemes més destacats estan relacionats amb l'eficiència de l'assaig, la determinació del llinard i el mètode de quantificació emprat (Weaver et al., 2010). La ddPCR (de l'anglès *Droplet Digital PCR*) és basa en una tècnica microfluídica d'anàlisi de milers de gotes per mostra, permetent així un càlcul acurat de les molècules de DNA utilitzant una distribució de Poisson (Hindson et al., 2011)(Figura 1).

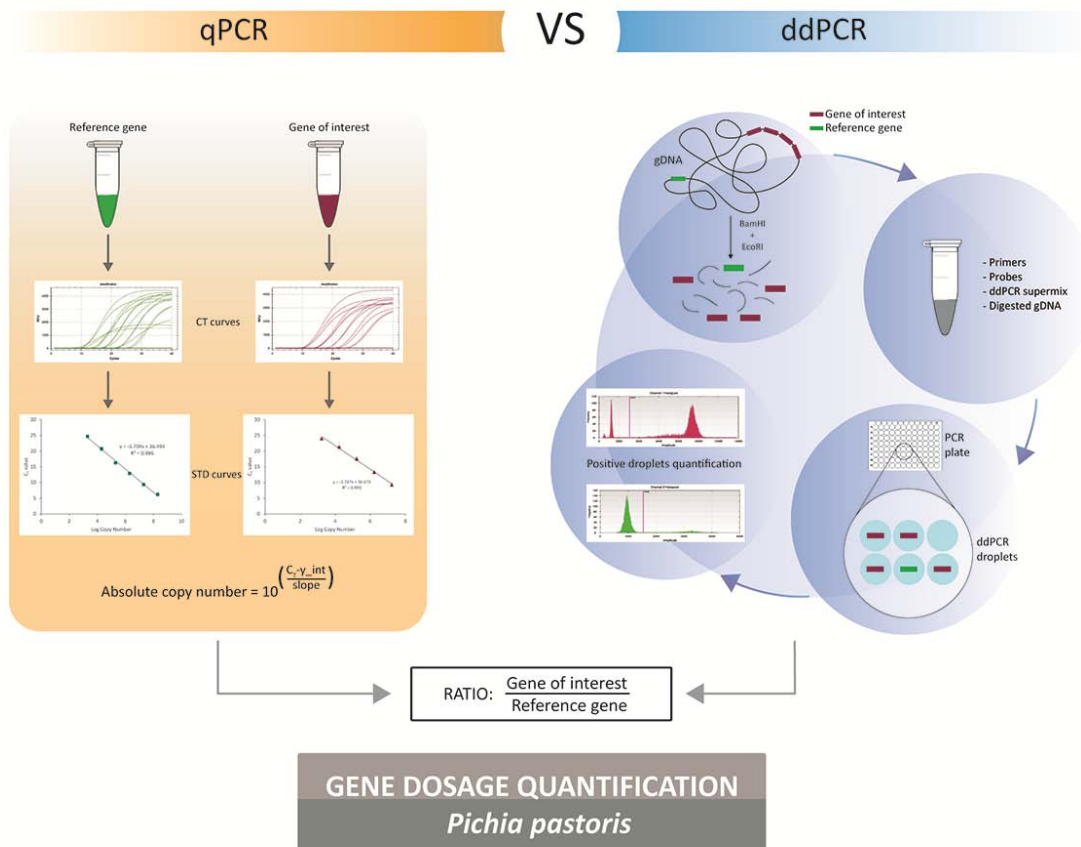


Figura 1. Comparació entre la metodologia de la qPCR (quantitative real-time PCR) i la ddPCR (Droplet Digital PCR) per la quantificació de la dosi gènica en *Pichia pastoris*. En el cas de la qPCR es requereix una corba estàndard pel gen d'interès i el gen referència, mentre que la determinació del nombre de còpies es pot realitzar mitjançant diferents mètodes de càlcul. La ddPCR es basa en la partició de la mostra en milers de gotes i la detecció de la senyal de cada gen, calculant la dosi gènica mitjançant el quocient del nombre de gotes positives pel gen d'interès i el gen de referència. CT, cicle límit; STD, corbes estàndard.

S'han descrit una àmplia varietat d'aplicacions per la ddPCR en diferents camps de recerca, confirmant el gran potencial d'aquesta tècnica (Cao et al., 2015; Corbisier et al., 2015; Floren et al., 2015). En conseqüència, durant els últims anys s'han publicat nombrosos estudis comparant ambdós mètodes (Hayden et al., 2013; Hindson et al., 2013; Te et al., 2015; Zhang et al., 2015). Darrerament la ddPCR s'ha postulat com una alternativa rellevant a la determinació de la dosi gènica. Malgrat això, l'aplicació de la ddPCR en llevats està actualment – al nostre parer – en vies de desenvolupament, ja que només s'ha utilitzat per determinació de dosi gènica en *Saccharomyces cerevisiae* una sola ocasió (Bozdag i Greig, 2014). De fet, fins l'actualitat, no s'han reportat aplicacions per *P. pastoris*. L'establiment de metodologies basades en ddPCR per aquest llevat podria proporcionar una eina potent i més fiable per accelerar properes millores

en *P. pastoris* com a plataforma de factoria cel·lular. En aquest capítol es descriu per primer cop l'ús de la ddPCR per la caracterització clonal a *P. pastoris*. En particular, aquesta tècnica va permetre determinar la dosi gènica de *ROL*, especialment en aquelles soques amb un interval entre 1 – 5 còpies integrades al genoma, facilitant d'aquesta manera la discriminació entre els efectes de la dosi gènica i els causats per la variació clonal (Clare et al., 1998; Marx et al., 2009; Viader-Salvado et al., 2006). Així mateix, la classificació precisa dels clons en funció del nombre de còpies va permetre determinar la dosi gènica òptima de *ROL* per la producció de lipasa, així com establir una sèrie de soques ben definida albergant fins a 15 còpies de *ROL*, aptes per posteriors anàlisis fisiològics sobre l'efecte de la dosi gènica en la fisiologia cel·lular.

2.2 MATERIALS I MÈTODES

2.2.1 Soques i plasmidis

En aquest estudi es va construir una sèrie de soques de *P. pastoris* derivades de X-33 (Invitrogen, Carlsbad, CA, EUA), cadascuna amb un nombre diferent de còpies del gen de la lipasa de *Rhizopus oryzae* (*ROL*) clonat en el vector pPICZ α A (Invitrogen). La soca sense transformar X-33 es va utilitzar com a control negatiu (C-). Dues soques amb una (C+1) i dues còpies de *ROL* (C+2), prèviament obtingudes per Minning et al. (2001) i Cos et al. (2005), respectivament, es van utilitzar com a controls positius. Ambdues soques es van seqüenciar pel Centre for Research in Agricultural Genomics (Bellaterra, Cerdanyola del Vallès, Espanya) per tal de verificar el contingut de *ROL* utilitzant la seqüenciació Ion Torrent (Ion Personal Genome Machine, Life Technologies, Carlsbad, CA, EUA). Es van obtenir diferents vectors amb una, dues o quatre còpies de *ROL* mitjançant la construcció de multimers *in vitro* i es van introduir a *P. pastoris* X-33 per tal de generar soques multicòpia. Totes les soques es van caracteritzar posteriorment en termes de dosi gènica de *ROL* per qPCR i ddPCR.

2.2.2 Determinació del nombre de còpies de *ROL* per ddPCR

Per tal de realitzar la ddPCR, es va aïllar el DNA genòmic utilitzant un kit comercial de purificació (Wizard® Genomic DNA Purification Kit, Promega). Per assegurar una quantificació acurada i evitar un impacte negatiu en el volum de les gotes, el DNA genòmic dels clons seleccionats es va digerir amb els enzims de restricció BamHI i EcoRI (Thermo Scientific™, Waltham, MA, EUA). Aquests enzims es van seleccionar amb l'eina informàtica desenvolupada per Bikandi et al. (2004), amb l'objectiu d'obtenir fragments menors a les 5 kb i la separació dels amplicons de *ROL* evitant la seva fragmentació. Considerant 6000 còpies de genoma per μ L com l'òptim del rang dinàmic del sistema QX100 (Droplet Digital™ PCR Applications guide, Bio-Rad), al mix de reacció es van afegir 0.4 ng de DNA genòmic digerit de cada mostra, equivalent a 40,000 còpies de genoma haploid de *P. pastoris*. Per l'amplificació de DNA, es van dissenyar un parell d'oligonucleòtids de *ROL* i una sonda específica marcada a l'extrem 5' amb amidita de fluoresceïna (6-FAM). Amb l'objectiu de normalitzar els resultats, es va utilitzar el gen de referència β -actina (*ACT1*) com a control endogen, i es van dissenyar uns oligonucleòtids específics, juntament amb una sonda marcada a l'extrem 5' amb hexaclaro-6-carboxi-fluoresceïna (HEX). Ambdues sondes es van marcar a l'extrem 3' amb la molècula *Black Hole Quencher-1* (BHQ-1). Tots els oligonucleòtids i sondes van ser sintetitzats per Biomers.net GmbH

(Ulm, Alemanya). Les seqüències dels primers es mostren a l'annex. Cada reacció de ddPCR es va dur a terme en un volum final de 20 µL, amb 12.5 µL de ddPCR™ Supermix for Probes (Bio-Rad), 300 nM de cada oligonucleòtid, 200 nM de sonda, 0.4 ng de DNA genòmic digerit i la quantitat requerida d'aigua *Dnase/Rnase-free* (Thermo Fisher Scientific, Waltham, MA, EUA). Les gotes amb les barres de reacció de PCR es van formar amb el dispositiu Droplet Generator (Bio-Rad) i subseqüentment es van transferir a una placa de 96 pous. Les reaccions es van incubar a 95°C durant 10 min per tal d'activar la polimerasa *hot-start*, i posteriorment es van sotmetre a un protocol de dos passos, incloent la desnaturalització (94°C, 30 s) i l'alineament/extensió (60°C, 1 min) durant 40 cicles. La detecció de gotes es realitza amb el QX100 Droplet Digital PCR System i el software QuantaSoft v. 1.5.38 (Bio-Rad). El nombre de còpies es va determinar a cada mostra per triplicat, calculant la ràtio entre les gotes positives de *ROL* i *ACT1*.

2.2.3 Determinació del nombre de còpies de *ROL* per qPCR

La qPCR es va dur a terme en reaccions de 20 µL utilitzant plaques de 96 pous iQ *semi-skirted* i SsoFast™ EvaGreen® Supermix (ambdós de Bio-Rad). Cada reacció de PCR contenia 5 µL de EvaGreen, 300 nM de cada oligonucleòtid, 5 ng de DNA genòmic i la quantitat requerida d'aigua *Dnase/Rnase-free*. Les mostres i els estàndars es van mesurar per triplicat en el dispositiu CFX384 Touch™ Real-Time PCR Detection System (Bio-Rad). Es va afegir un control amb aigua (sense DNA motlle) per cada parell d'oligonucleòtids per tal d'evitar la detecció d'oligonucleòtids units de manera inespecífica, i el gen *ACT1* es va seleccionar com a referència. Les reaccions es van incubar a 98°C durant dos minuts per tal d'activar la polimerasa Taq, i posteriorment es va realitzar un protocol de dos passos, consistent en un pas de desnaturalització (98°C, 5 s) i un pas d'alineament/extensió (60°C, 15 s) durant 40 cicles. Amb l'objectiu de verificar l'especificitat de les dianes i els gens de referència, es va analitzar la corba de fusió amb un rang de temperatura de 65°C fins 95°C, amb un increment de 0.5°C/cicle (interval de 10 s). Donat que les eficiències d'amplificació de les dianes i els gens referència eren similars, es va seleccionar el mètode de càlcul de quantificació absoluta (Lee et al., 2008) per la determinació del nombre de còpies de *ROL*. Totes les seqüències dels oligonucleòtids i les sondes es troben a l'annex.

2.2.4 Anàlisi per citometria de flux

Per evitar agregats cel·lulars, prèviament a l'anàlisi amb el citòmetre totes les mostres es van sonicar durant 3 cicles de 4 s a 50 W amb el processador d'ultrasons VC-5 (Vibracell, Sonics&Materials, Newtown, CA, EUA). Pel comptatge cel·lular, les mostres es van diluir amb PBS filtrat (0.22 µm, MCE membrane, Millex-GS, Millipore, Jaffrey, NH, EUA) a una concentració final de 50-500 cèl·lules/µL. Per la tinció de viabilitat, es va diluir 1 mL de cultiu cel·lular amb PBS filtrat a una concentració final de 10⁶ cèl·lules/µL, i posteriorment es va addicionar 1 mL de iodur de propidi (PI, de l'anglès Propidium Iodide, 1 mg/mL, Sigma-Aldrich, St Louis, MO, EUA) i les mostres es van analitzar al cap de 5 min. Com a control positiu per les cèl·lules no viables, es va incubar 1 mL de cultiu durant 3 min a 80°C. Amb l'objectiu de mesurar la presència d'espècies reactives d'oxigen (ROS, de l'anglès *Reactive Oxygen Species*) es van afegir 0.2 mL de dihidroetidi (DHE; 12.5 µg/mL, Sigma-Aldrich) a 1 mL de cultiu amb una concentració de 10⁶ cèl·lules/mL. Després d'una incubació de 30 min a 30°C amb agitació constant, les cèl·lules es van centrifugar a 13,000 rpm durant 3 min (Espresso Microcentrifuge, Thermo Electron Corporation, Waltham, MA, EUA) i es van resuspendre en 1 mL de PBS. Com a control positiu d'estrès cel·lular es va utilitzar un cultiu crescut en un 20% de metanol durant 30 min (Costa et al., 1997). Les mostres es van processar amb el citòmetre Guava EasyCyte Mini (Millipore), amb un làser d'argó de 488 nm. Per cada mostra es van mesurar 5,000 cèl·lules amb un cabal de 0.59 µL/s. L'emissió del PI i del DHE es va detectar amb el filtre de 680 nm.

2.2.5 Procediments analítics

2.2.5.1 Anàlisi de biomassa

La biomassa es va monitoritzar mesurant la densitat òptima a 600 nm (DO₆₀₀) en mostres per triplicat. La concentració de biomassa als cultius operats en continu es va determinar mitjançant el pes sec (DCW, de l'anglès Dry Cell Weight) utilitzant el mètode descrit prèviament per Jordà et al. (2012). Les determinacions es van realitzar per triplicat i les desviacions estàndard relatives (RSD, de l'anglès *Relative Standard Deviation*) van ser del 2% aproximadament.

2.2.5.2 Assaig d'activitat lipolítica

L'anàlisi d'activitat lipolítica es va dur a terme en triplicats seguint el protocol descrit a Resina et al. (2004).

2.2.5.3 Quantificació dels metabòlits extracel·lulars

Per l'anàlisi dels metabòlits extracel·lulars es van centrifugar 2 mL de sobrenedant de cultiu a 12,000 rpm durant 1 min en la microcentrífuga Minispin (Eppendorf, Hamburg, Alemanya) per tal d'eliminar les cèl·lules, i posteriorment es va filtrar (0.45 µm, tipus HAWP, Millipore). El glicerol, metanol i d'altres compostos extracel·lulars d'interès potencial dels cultius operats en continu es van analitzar per duplicat per HPLC (Dionex Ultimate 3000, Dionex Corporation, Sunnyvale, CA, EUA) utilitzant una columna d'intercanvi iònic (ICSep ICE-COREGEL 87 H3, Transgenomic Inc., Omaha, NE, EUA). La fase mòbil estava formada per àcid sulfúric 8 mM. El volum d'injecció va ser de 20 µL i el cromatograma es va quantificar amb el software Chromeleon v. 6.8 (Dionex).

2.2.6 Condicions de cultiu

2.2.6.1 Cultius en microplaca

Pels cultius en microplaca de 24 pous (Whatman, GE Healthcare Life Sciences, Buckinghamshire, Regne Unit) es va realitzar un precultiu dels clons en 5 mL de medi YPD (de l'anglès *Yeast extract Peptone Dextrose*) (1% (w/v) extracte de llevat, 2% (w/v) peptona, 2% (w/v) dextrosa) en tubs Falcon de 50 mL. Els cultius es van inocular amb colònies provinents de plaques i es van incubar tota la nit a 30°C i 150 rpm en l'agitador orbital Multitron II (Infors, Einsbach, Alemanya). Després de 14-16 h, es va mesurar la densitat òptica i les cèl·lules es van centrifugar i resuspendre en 5 mL de medi BMG (de l'anglès *Buffered Minimal Glycerol*) (100mM fosfat potàssic (pH 6.0), 1.34% (w/v) base de llevat nitrogenada sense aminoàcids, 4·10⁻⁵% (w/v) biotina, 1% (v/v) glicerol) fins una DO₆₀₀ de 0.1, i es va continuar la incubació. Després de 16-18 h, es van centrifugar les cèl·lules a temperatura ambient i es van inocular les microplaques a una DO₆₀₀ d'1, per triplicat per cada soca, amb 2 mL de medi BMM (de l'anglès *Buffered Minimal Methanol*) (100mM fosfat potàssic (pH 6.0), 1.34% (w/v) base de llevat nitrogenada sense aminoàcids, 4·10⁻⁵% (w/v) biotina, 0.5% (v/v) metanol) a cada pou. Les plaques inoculades es van segellar amb les membranes BREATHseal™ (Greiner, Monroe, NC, EUA), estèrils i permeables a gas, per tal de mantenir una aportació òptima d'oxigen als cultius. Les cèl·lules es van incubar durant 72 h a 25°C, amb un 80% d'humitat i 200 rpm d'agitació. Es va addicionar un pols de metanol pur a una concentració final de 0.5% v/v als cultius cada 24 h per tal de compensar el consum i evaporació del metanol. Tots els medis (menys la soca control X-33) es van suplementar amb Zeocina (100 µg/mL).

2.2.6.2 Cultius en Erlenmeyer

Es van realitzar triplicats dels cultius en Erlenmeyer per cada clon de la següent manera: es van inocular amb criovials de la soca seleccionada, a una DO_{600} inicial de 0.1, Erlenmeyers amb deflectors d'1 L amb 100 mL de medi BMG, i es van incubar a 30°C i 150 rpm. Després de 18-20 h, les cèl·lules es van centrifugar i es van utilitzar per inocular Erlenmeyers amb deflectors de 0.5 L amb 50 mL de medi BMM a una DO_{600} d'1. Els cultius es incubar durant 70 h a 25°C, amb un 80% d'humitat i 150 rpm d'agitació (incubador Multitron II). Cada 24 h es va afegir un pols de metanol pur a una concentració final de 0.5% v/v. Tots els medis (menys la soca control X-33) es van suplementar amb Zeocina (100 µg/mL).

2.2.6.3 Cultius en reactor operats en continu

Els cultius operats en continu es van dur a terme a un volum de treball d'1 L en un bioreactor d'1.5 L (Biostat B Plus, Sartorius AG, Göttingen, Alemanya). Com a precultiu, es va inocular un Erlenmeyer d'1 L i 150 mL de medi YPD-Zeocina amb un criovial de la soca seleccionada a una DO_{600} inicial d'1, i es va incubar durant aproximadament 24 h a 30°C i 150 rpm, per posteriorment inocular un volum d'1 L de medi *batch*. Al final de la fase *batch* (indicat per una caiguda sobtada dels nivells de CO₂ en el reactor), el cultiu en continu es va iniciar a una taxa de dilució (D) de 0.1 h⁻¹ amb un medi de creixement definit i una barreja glicerol/metanol (60/40%, w/w) com a font de carboni. La temperatura del cultiu es va mantenir a 25°C i el pH es va controlar a 5 amb amoni 10%. La taxa d'aeració d'1 vvm i l'agitació a 500 rpm van mantenir els nivells d'oxigen dissolt en un mínim de saturació de l'aire del 10%. Es va aplicar al sistema una sobrepressió de 0.2 bars per tal de facilitar la presa de mostra. Els gasos de sortida del reactor es van fer circular a través d'un condensador refrigerat a 4°C mitjançant un criòstat (FrigoMix 1000, B. Braun Biotech International, Melsungen, Alemanya), i seguidament assecat per dues columnes de sílice gel, per ser finalment analitzat amb l'espectròmetre de masses Omnistar™ GSD 300 02 (Balzers Instruments, Balzers, Liechtenstein). L'espectròmetre de masses es va calibrar utilitzant gasos de calibratge amb 5 i 21% d'O₂ i 3, 5 i 7% de CO₂. Les mostres de cultiu es van prendre al cap d'un mínim de cinc temps de residència, un cop assolit l'estat estacionari (Jordà et al., 2013). La composició del medi *batch* va ser la següent: 9.98 g/L glicerol, 0.45 g/L àcid cítric, 3.15 g/L (NH₄)₂HPO₄, 5 mg/L CaCl₂, 0.23 g/L KCl, 0.15 g/L MgSO₄·7H₂O, 0.5 mL biotina (0.2 g/L; Sigma-Aldrich), 1.15 mL sals traça PTM1. El medi del continu contenia: 7.5 g/L glicerol, 5 g/L metanol, 2.3 g/L àcid cítric·H₂O, 1.09 g/L (NH₄)₂HPO₄, 2.5 mg/L CaCl₂·2H₂O, 0.43 g/L KCl, 0.16 g/L MgSO₄·7H₂O, 0.25 mL biotina (0.2 g/L) i 0.4 mL de sals traça PTM1. El pH es va ajustar a 5.0 en

ambdós medis mitjançant HCl 25%. La composició de les sals traça PTM1 és la mateixa descrita a Baumann et al. (2008).

2.2.7 Anàlisi estadístic

Es van preparar dues i tres plaques de reaccions per la qPCR i la ddPCR, respectivament, amb tres rèpliques tècniques per cada cas (és a dir, sis reaccions per mostra per la qPCR, i nou reaccions per mostra en el cas de la ddPCR). Els cultius en microplaca es van realitzar en triplicats tècnics per cada soca, mentre que els cultius en Erlenmeyer es van dur a terme per triplicat provinents de precultius independents. Les dades obtingudes dels cultius en continu es van verificar utilitzant procediments estàndard de consistència i reconciliació (Wang i Stephanopoulos, 1983; van der Heijden et al., 1994; Verheijen, 2010), amb la restricció que es satisfessin les relacions de conservació dels elements. Per tots els cultius realitzats el test de consistència va acomplir un nivell de confiança del 95%, i en conseqüència, no es van detectar crassos errors de mesura. Els resultats s'expressen en termes de mitjana \pm SD (de l'anglès *standard deviation*). L'anàlisi estadístic dels resultats es va realitzar utilitzant el test ANOVA, seguit d'una prova *t* de Student (mostres no aparellades), amb el software Excel de Microsoft. Es va considerar com a significatiu un *P-val* menor al 0.05.

2.3 RESULTATS I DISCUSSIÓ

2.3.1 Construcció d'una col·lecció de soques amb diferent nombre de còpies de la lipasa de *R. oryzae* (ROL)

Per a generar una sèrie de clons amb diferent dosi gènica de *ROL*, es varen construir de manera esglaonada vectors derivats del pPICZαA amb un nombre diferent de cassets gènics (Figura 2). En primer lloc, el casset d'expressió del gen *ROL* (2446 pb) (constituït per la regió

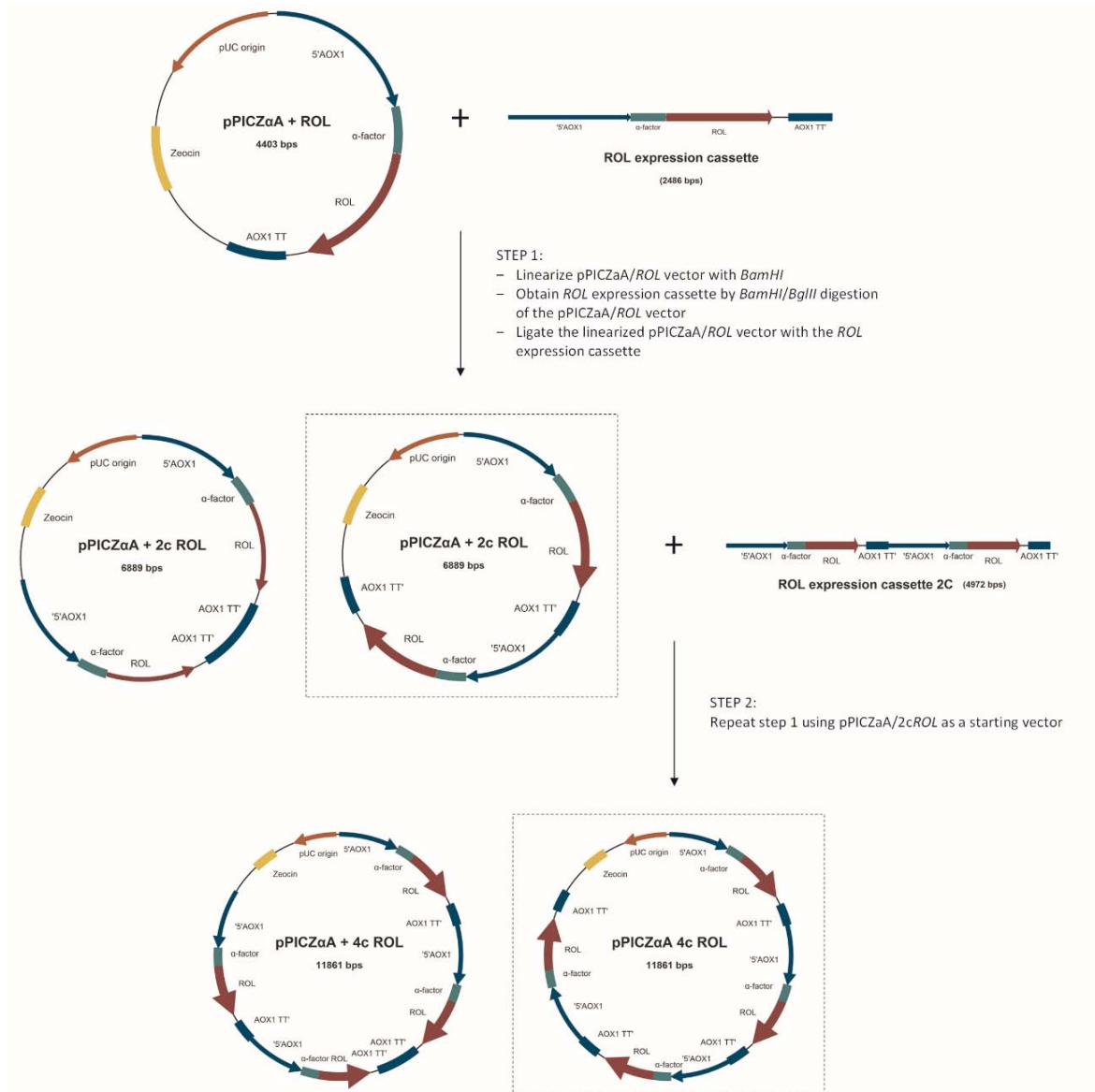


Figura 2. Construcció dels plasmidus multimèrics de *ROL*. Diagrama de flux de les etapes per l'assemblatge dels cassets d'expressió de *ROL* al vector monocòpia pPICZαA. Origen pUC, origen de replicació; 5'AOX1, promotor *AOX1* de *P. pastoris*; α-factor, senyal de secreció α-mating factor de *S. cerevisiae*; *AOX1* TT, regió terminador de la transcripció; Zeocin, gen de resistència a l'antibiòtic Zeocina; *ROL*, gen de la lipasa de *Rhizopus oryzae*.

promotora 5'AOX1 de l'alcohol oxidasa, la seqüència codificant per la senyal de secreció *α -mating factor*, la seqüència codificant per la forma madura de la Rol i la regió terminador de la transcripció d'AOX1) es va escindir del vector monocòpia preexistent pPICZ α A/ROL mitjançant els enzims de restricció BglII i BamHI. Seguidament, el casset d'expressió i el vector monocòpia (prèviament linealitzat amb BamHI) es van lligar per obtenir un vector amb dues còpies de ROL. Aquest pas es va repetir per obtenir un vector de quatre còpies. Amb l'objectiu d'evitar interferències transcripcionals entre els gens disposats en tàndem, només es van seleccionar cassets d'expressió amb la mateixa orientació (Gullerova i Proudfoot, 2010; Hobson et al., 2012). Malauradament, no va ser possible obtenir vectors amb una dosi gènica més elevada. Aquest fet podria ser degut a recombinacions en la soca hoste o la grandària del vector (Vassileva et al., 2001). El nombre de còpies del casset d'expressió de ROL dels constructes obtinguts es va verificar posteriorment mitjançant una restricció enzimàtica (dades no mostrades). Posteriorment es van realitzar transformacions independents de *P. pastoris* amb cada vector, obtenint un total de 70 clons. El conjunt de clons es va analitzar per PCR per confirmar la presència de ROL (dades no mostrades), resultant en 35 clons positius.

2.3.2 Comparació de la qPCR vs ddPCR per la determinació del nombre de còpies de ROL

Es va realitzar una anàlisi comparativa entre la tècnica de la qPCR i la ddPCR per trobar el millor mètode per determinar la dosi gènica de ROL. Primerament, tots els clons positius es van analitzar amb la qPCR. A fi d'establir una comparació, es va seleccionar un subgrup de 13 clons (representatius de diferents rangs de còpies) per la posterior quantificació per ddPCR. De manera global, la tècnica de la ddPCR va presentar variacions més baixes que la qPCR en gairebé totes les mostres (Taula II). Respecte les mostres de referència, la precisió de la ddPCR va ser més elevada comparada amb la qPCR, la qual va presentar unes desviacions estàndard relatives (RSD, de l'anglès *relative standard deviation*) al voltant del 50% en les soques control 1-còpia i 2-còpies. En el cas de les soques de rang baix (1 – 5 còpies) la precisió de la ddPCR va ser també millor, amb valors de RSD entre 6 – 11%, mentre que el rang de RSD de la qPCR va ser del 25 - 50%. Cal destacar que la soca #24, prèviament caracteritzada com a 2-còpies mitjançant la qPCR, es va reassignar com una soca monocòpia segons la tècnica de la ddPCR.

Taula I. Determinació del nombre de còpies de *ROL* mitjançant qPCR i ddPCR. Dos i tres plaques es van preparar per la qPCR i la ddPCR respectivament, incloent 3 rèpliques tècniques a cada mostra. Els resultats es mostren com la mitjana \pm SD. C-: control negatiu, soca X-33 sense vector; C+1, control positiu, 1 còpia de *ROL*, soca seqüenciada; C+2, control positiu, 2 còpies de *ROL*, soca seqüenciada; RSD, desviació estàndard relativa. (*) El nombre de còpies dels cassets d'expressió de *ROL* va ser posteriorment amplificat mitjançant el mètode d'enriquiment per Zeocina (Sunga et al., 2008).

Soca	Nombre de cassets de <i>ROL</i> en el vector de transformació	qPCR		ddPCR	
		Nombre de còpies	% RSD	Nombre de còpies	%RSD
C-	-	0.0	75	0.2	100
C+1	*	0.4	50	0.7	14
C+2	*	1.0	45	2.1	10
30p	2	9.6	1	8.7	19
#24	*	2.3	6	0.7	9
31p	2	1.1	41	1.7	6
33p	2	1.3	49	1.9	11
37p	4	1.9	32	3.9	8
38p	4	1.3	25	4.1	10
52p	2	32.1	10	14.6	13
53p	2	22.6	12	11.3	8
55p	2	5.2	1	1.8	8
61p	2	29.6	5	6.9	10
65p	4	1.6	31	1.9	9

Aquest nou índex podria explicar el fet que la soca 1-còpia referència i la soca #24 no mostressin diferències significatives en termes de fluxos metabòlics (Jordà et al., 2012). Per les soques de rang alt (> 5 còpies) els RSD d'ambdues tècniques van ser similars. Tanmateix, l'estimació del nombre de còpies va ser significativament més elevat utilitzant el mètode de la qPCR. Com a exemple, tot i que en ambdós mètodes la soca amb un nombre de còpies més elevat va ser la 52p (amb un RSD aproximadament del 10%), la tècnica de la qPCR va determinar que el nombre de còpies era 32, mentre que la ddPCR va establir que era de 15 còpies. Això no obstant, per tal d'obtenir 32 còpies de *ROL* s'haurien necessitat 16 integracions del casset, un esdeveniment improbable considerant que l'estratègia de transformació no implicava cap metodologia per prioritzar les soques d'alt nombre de còpies (> 7 còpies), tal com un procés d'enriquiment post-transformació mitjançant Zeocina (Sunga et al., 2008) o la integració al locus del DNA ribosomal (Marx et al., 2009). En conjunt, la ddPCR sembla un mètode més precís per la caracterització de col·leccions de clons, tot i que es requeriria una optimització de l'anàlisi de les soques de rang

alt per millorar la precisió i validació de la tècnica. A fi d'avaluar posteriorment la fiabilitat i la solidesa de la ddPCR, es va determinar el nombre de còpies de *ROL* d'un total de 30 clons mitjançant aquesta tècnica, resultant en dosis gèniques entre 1 – 15 còpies (Taula III).

Taula II. Quantificació del nombre de còpies de *ROL* mitjançant el mètode de la ddPCR. Es van establir tres rangs, depenent del nombre de còpies de *ROL*: baix (1 – 2 còpies) (●), intermedi (3 – 6 còpies) (■) i alt (7 – 15 còpies) (■). Les soques seleccionades per posteriors cultius operats en continu s'han marcat amb *. Els resultats s'expressen com la mitjana ± SD de tres experiments independents. SD, desviació estàndard

Soca	Reassignació	Nombre de cassets de <i>ROL</i> en el vector de transformació	Nombre de còpies ± SD	Soca	Reassignació	Nombre de cassets de <i>ROL</i> en el vector de transformació	Nombre de còpies ± SD
C-*	0C	x1	0.2 ± 0.1	67p		x1	4.2 ± 0.3
C+*	1C	x4	0.2 ± 0.1	59p		x2	6.1 ± 0.2
68p		x1	1.3 ± 0.1	66p		x4	6.4 ± 0.2
55p		x2	1.8 ± 0.2	61p		x2	6.9 ± 0.7
31p		x2	1.7 ± 0.1	19p		x1	7.2 ± 2.2
58p		x2	1.7 ± 0.2	25p		x2	7.4 ± 0.1
63p		x2	1.9 ± 0.1	51p		x2	7.5 ± 0.2
33p*	2C	x2	1.9 ± 0.2	14p		x4	7.5 ± 3.2
65p		x4	1.9 ± 0.2	64p*	8C	x4	7.9 ± 0.1
29p		x2	2.1 ± 0.2	54p		x2	10.2 ± 0.6
56p		x2	2.1 ± 0.2	69p		x1	10.5 ± 1.6
32p		x2	2.6 ± 1.4	30p		x2	8.7 ± 1.7
60p		x2	2.7 ± 0.1	15p		x4	10.7 ± 6.7
70p		x1	3.7 ± 0.1	53p		x2	11.3 ± 0.9
37p		x4	3.9 ± 0.3	57p		x2	11.7 ± 0.9
38p*	4C	x4	4.1 ± 0.4	52p*	15C	x2	14.6 ± 2.0

2.3.3 Estudi dels clons transformants en plaques de 24 pous i cultius en Erlenmeyer: correlació de la dosi gènica amb el creixement cel·lular i la producció de lipasa

La col·lecció de 30 clons transformants caracteritzats per ddPCR es va analitzar a petita escala mitjançant cultius en plaques de 24 pous (Figura 3). Després de 70 h es va mesurar la biomassa i l'activitat lipàsica per tal d'estudiar l'efecte de la dosi gènica de *ROL*. Amb la finalitat d'establir una classificació de les soques, es van establir tres subgrups depenent del nombre de còpies de

ROL: baix (1 – 2 còpies), intermig (3 – 6 còpies), i alt (7 – 15 còpies). Els resultats d'aquest cribratge preliminar es mostren a la Figura 3. Degut a les limitacions de tècniques tals com la densitat òptica o la determinació del pes sec en cultius convencionals a petita escala (Quirós et al., 2007), els nivells de biomassa s'expressen en unitats de cèl·lules/mL, permetent un anàlisi més acurat. Malgrat no ser estadísticament significatiu, el creixement cel·lular presentà de manera global una correlació negativa amb el nombre de còpies de *ROL*. És a dir, els contatge cel·lular era més elevat per la soca control i decreixia gradualment a mesura que s'incrementava el nombre de còpies de *ROL* a la soca. Cal destacar que alguns clons amb el mateix nombre de còpies presentaven diferències importants en el comptatge cel·lular, reflex d'una destacable variació clonal. Aquesta variació ja havia estat observada prèviament en la construcció de soques multicòpia (Marx et al., 2009). Respecte a l'activitat lipàsica específica (UA/cèl·lula), els nivells més alts van ser obtinguts pels clons de rang baix, amb la soca 33p (amb 2 còpies de *ROL*) produïnt 24 vegades més que la soca C+1, com es mostra a la Figura 3. Les soques de rang intermig van presentar entre elles una notable heterogeneïtat. És el cas de les soques 32p i 67p (amb 3 – 4 còpies de *ROL*) que van mostrar nivells de producció significatius, mentre que no es va poder detectar activitat lipàsica als clons amb 6 – 7 còpies. Convé destacar que es va observar una diferència significativa en la producció lipàsica entre les soques de rang baix i alt. Estudis previs ja havien descrit com la producció de proteïna recombinant normalment pot causar estrés oxidatiu a les soques, degut al plegament de proteïnes al reticle endoplasmàtic (Gasser et al., 2008) i la generació de ROS (de l'anglès *Reactive Oxygen Species*) durant la oxidació del metanol en *P. pastoris* (Vanz et al., 2012), afectant així al creixement cel·lular i a la producció de proteïna. Per aquest motiu, els nivells de ROS i la viabilitat cel·lular es van analitzar durant la fase d'inducció. En relació a l'estrés oxidatiu, es va observar que la soca més productora (33p, 2 còpies) mostrava nivells de ROS per sota del 20% (expressat com % cèl·lules amb ROS tenyides), mentre que els nivells més alts es van detectar per les soques d'alt rang 64p (8 còpies) i 52p (15 còpies), acompanyat d'un nombre de cèl·lules no viables més elevat.

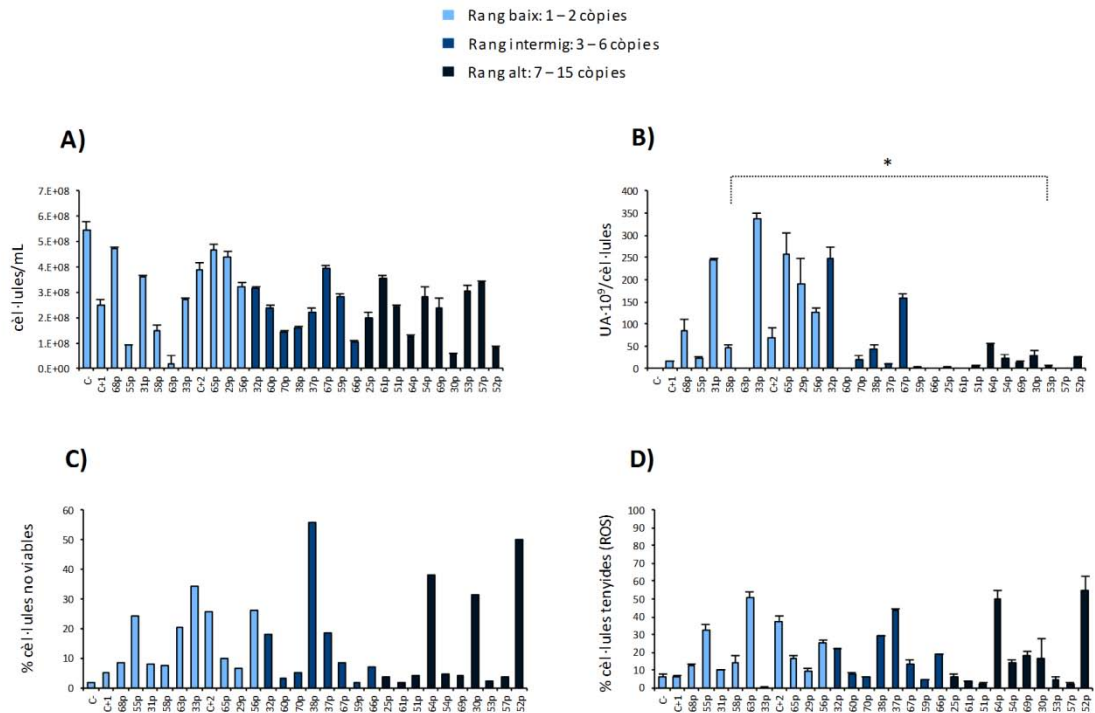


Figura 3. Anàlisi dels clons transformats en plaques de 24 pous. A) Creixement cel·lular B) Producció de Rol expressada en termes d'activitat lipàsica C) Viabilitat cel·lular D) Detecció de ROS. La línia sòlida indica la comparació entre el rang baix i l'intermig, la línia de guions senyala la comparació entre el rang mig i l'alt i la línia de punts mostra la comparativa entre el rang baix i l'alt. Les barres d'errors indiquen la desviació estàndard de les rèpliques. * $P < 0.05$, ** $P < 0.1$. UA: unitats d'activitat

A continuació es va seleccionar un subconjunt de 14 clons, representatius de cada rang de còpies en termes de creixement (és a dir, amb un creixement cel·lular pròxim al promig de cada subgrup) per l'anàlisi en cultius en Erlenmeyer. Aquesta sèrie de cultius va confirmar la tendència observada en les plaques de 24 pous, amb una disminució significativa en el comptatge cel·lular a mesura que la dosi de *ROL* augmentava, i un màxim en la activitat lipàsica específica per la soca 33p (2 còpies) (Figura 4A). Concretament, es va mesurar una caiguda en les soques amb un baix nombre de còpies respecte la soca C-. Aquest impacte negatiu de l'expressió de Rol en el creixement cel·lular va ser més notori en els clons de rang intermig i alt. Respecte l'activitat lipàsica (Figura 4B), totes les soques multicòpia (amb excepció del clon 53p) van manifestar productivitats més altes que la soca C+1 (monocòpia), amb un òptim aparent en les soques de dues còpies, seguit d'una caiguda en la productivitat a les soques amb una dosi gènica superior. Els nivells de ROS van ser similars als observats a les plaques de 24 pous (Figura 4D), amb excepció dels controls positius. Finalment, es va detectar una disminució de les

cèl·lules no viables a la soca 38p, la qual havia presentat nivells de viabilitat més baixos que el promig als cultius de 24 pous (Figura 4C).

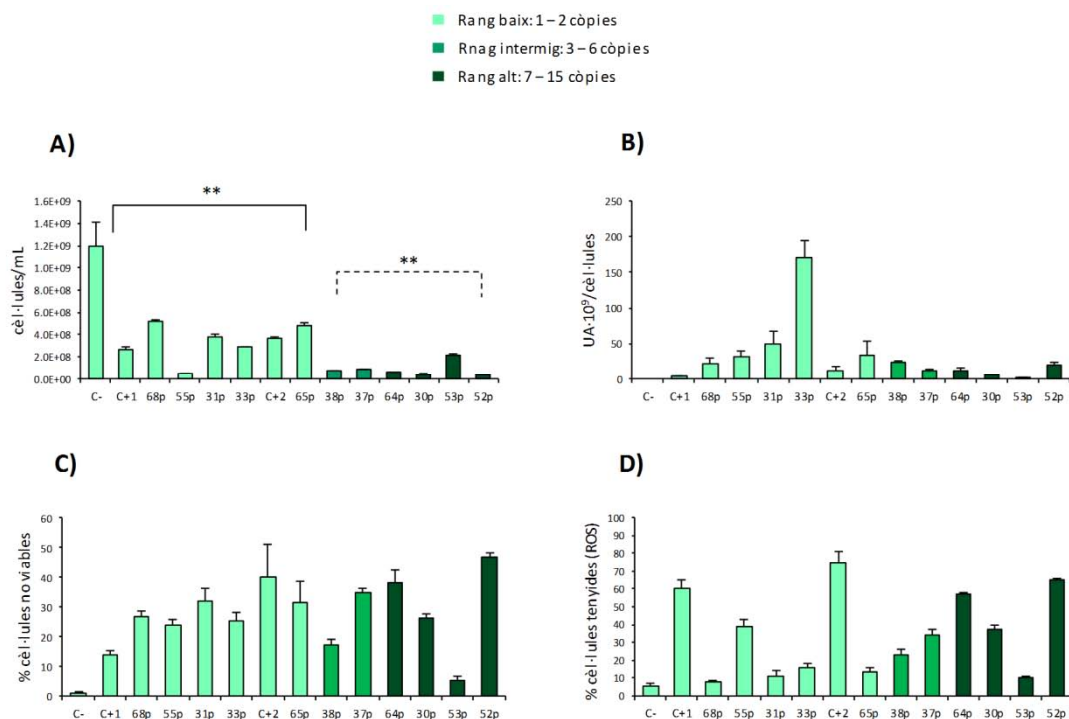


Figura 4. Anàlisi dels clons transformats en cultius en Erlenmeyer. A) Creixement cel·lular B) Producció de Rol expressada en termes d'activitat lipàsica C) Viabilitat cel·lular D) Detecció de ROS. La línia sòlida indica la comparació entre el rang baix i l'intermedi, la línia de guions senyala la comparació entre el rang mig i l'alt i la línia de punts mostra la comparativa entre el rang baix i l'alt. Les barres d'errors indiquen la desviació estàndard de les rèpliques. * $P < 0.05$, ** $P < 0.1$. UA: unitats d'activitat

2.3.4 Validació del comportament dels clons seleccionats en cultius operats en continu

Amb l'objectiu de verificar els resultats obtinguts en els cultius a petita escala, es van seleccionar 6 clons representatius de cada rang de còpies en termes de creixement cel·lular i producció de Rol per ser cultivats en un reactor operat en continu. De manera prèvia, es van establir les condicions d'operació per tal d'augmentar la productivitat volumètrica i específica del procés. Tal i com s'ha reportat àmpliament en la bibliografia, l'ús de mescles de glicerol i metanol com a font de carboni pot reduir la sobrecàrrega metabòlica derivada de la producció de proteïna recombinant a *P. pastoris* (McGrew et al., 1997; Anjou i Daugulis, 2000; Zhang et al., 2003; Jungo et al., 2007). Al treball de Jordà et al. (2014), on es va feia créixer la soca 1C en

cultius en continu amb una barreja de glicerol/metanol en diferents proporcions com a font de carboni i amb dues taxes de dilució diferent (0.05 i 0.16 h⁻¹).

La sèrie de cultius va confirmar un destacable impacte negatiu de l'increment de la dosi gènica de *Rol* en el creixement cel·lular, tal i com s'havia descrit prèviament per Cos et al. (2005) i Serrano et al. (2001). Concretament, els nivells de biomassa en la soca monocòpia i les soques multicòpia van ser un 20% i 50% més baixos que el control negatiu, respectivament (Figura 5A). D'acord amb el que s'havia observat als cultius a menor escala, els nivells de lipasa no es correlacionaven de manera lineal amb el nombre de còpies de *ROL*. La màxima producció de lipasa es va obtenir mitjançant la soca amb 2 còpies, mentre que dosis gèniques més elevades van resultar en nivells d'activitat lipàsica dos vegades menor en les soques 8C i 15C comparades amb la 2C (Figura 5B). Aquests resultats coincideixen amb diversos estudis on l'òptim de producció es trobava amb una dosi gènica baixa, i subseqüents increments en el nombre de còpies provocaven productivitats més baixes (Liu et al., 2014; Sha et al., 2013; Zhu et al., 2009). Per contra, a diferència dels cultius en 24 pous i Erlenmeyers, els nivells de ROS i de cèl·lules no viables van estar per sota dels nivells de detecció durant els cultius en continu (Figura 5C i 5D).

Aquesta observació coincideix amb els estudis previs realitzats per Hesketh et al. (2013), on la resposta a estrès en les cèl·lules recombinants de *P. pastoris* es va detectar a nivells transcripcionals poc després del canvi a condicions inductores de proteïna recombinant, és a dir, condicions transitòries. No obstant, aquesta resposta transcripcional va dissipar-se un cop les cèl·lules es van adaptar a les noves condicions (estat estacionari). Respecte la taxa de consum de la font de C, aquesta també es va veure afectada per la dosi gènica de *ROL* (Taula III), tal i com s'havia reportat per Cos et al. (2005) en cultius en discontinu alimentat utilitzant metanol com a única font de carboni, i per Jordà et al. (2012) en cultius en continu creixent amb una barreja glucosa:glicerol.

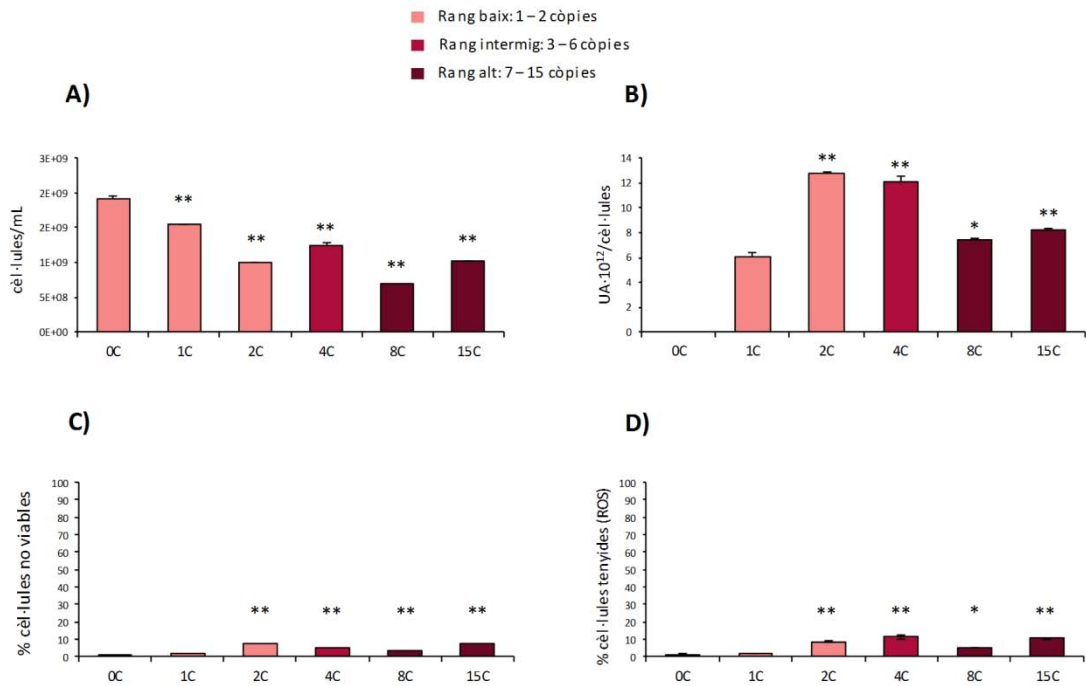


Figura 5. Comportament dels clons seleccionats, representatius de cada rang, crescuts en reactor operant en continu. A) Creixement cel·lular B) Producció de Rol C) Viabilitat cel·lular D) Detecció de ROS. Les barres d'error indiquen la desviació estàndard. * $P < 0.05$, ** $P < 0.1$ (En A,C i D, soques productores de Rol comparades amb la soca control no productora; en B, soques multicòpia comparades amb la soca 1C). UA: unitats d'activitat lipàsica

En concret, la taxa de consum específic de glicerol es va incrementar de 1.09 a la soca no productora fins 2.24 mmol/g·h a la soca 2C, tot i que aquest augment va ser menor a les altres soques amb més dosi gènica (aproximadament 1.59 mmol/g·h). D'altra banda, el consum específic de metanol va disminuir notòriament a les soques multicòpies, de 2.36 fins 1.18 mmol/g·h al control negatiu i la soca amb dosi gènica més elevada (15C), respectivament. Curiosament, aquest valor va ser encara més baix a la soca de 2 còpies (0.52 mmol/g·h). De manera conseqüent, es van detectar nivells significatius de metanol al medi de cultiu de les soques multicòpia, assolint 4.4 g/L a la soca 2C (Taula III), confirmant la reduïda capacitat assimilatòria del metanol per part d'aquestes soques. De manera interessant, Jordà et al. (2012) va descriure un increment en el consum específic de metanol entre la soca no productora i la soca monocòpia en cultius en continu amb una mescla glucosa:metanol com a font de carboni. Aquests resultats apuntarien a una adaptació diferent davant un increment en la carrega metabòlica dependent de la font de carboni utilitzada.

Taula III. Paràmetres de creixement de *P. pastoris* en cultius operats en continu, amb una mescla glicerol:metanol com a font de carboni

	q_s , glicerol (mmol g ⁻¹ DCW h ⁻¹)	q_s , metanol (mmol g ⁻¹ DCW h ⁻¹)	$Y_{x/s}$ (g DCW g ⁻¹) ^a	Activitat lipàsica (UA 10 ¹² cèl·lula ⁻¹)	Metanol residual (g/L)
0C	-1.09 ± 0.07	-2.36 ± 0.07	0.50 ± 0.09	n.d.	-
1C	-1.44 ± 0.22	-2.15 ± 0.65	0.40 ± 0.08	6.09 ± 0.33	-
2C	-2.24 ± 0.09	-0.52 ± 0.07	0.48 ± 0.11	12.73 ± 0.15	4.43 ± 0.00
4C	-1.49 ± 0.07	-1.24 ± 0.07	0.45 ± 0.10	12.09 ± 0.42	2.87 ± 0.05
8C	-1.70 ± 0.07	-0.63 ± 0.07	0.36 ± 0.09	7.40 ± 0.14	3.79 ± 0.05
15C	-1.57 ± 0.07	-1.18 ± 0.07	0.42 ± 0.10	8.19 ± 0.17	3.09 ± 0.02

q_s , taxa de consum específic de substrat reconciliat; $Y_{x/s}$, rendiment biomassa/substrat reconciliat; DCW, pes sec (de l'anglès *Dry Cell Weight*); n.d., no detectable. L'índex de consistència h va ser menor que 5 en tots els casos (nivell de confiança del 95%). Els resultats es mostren com la mitjana ± SD basada en mesures per triplicat. ^a Rendiment experimental abans de la reconciliació.

Cal remarcar que la classificació establerta segons la qPCR no permetria discriminar entre les soques de rang baix que mostraven comportaments fisiològics i productivitats de lipasa diferents, mentre que el rànquing basat en la tècnica de la ddPCR permet observar l'efecte fenotípic resultant de l'increment d'una sola còpia de *ROL* (Figura 6).

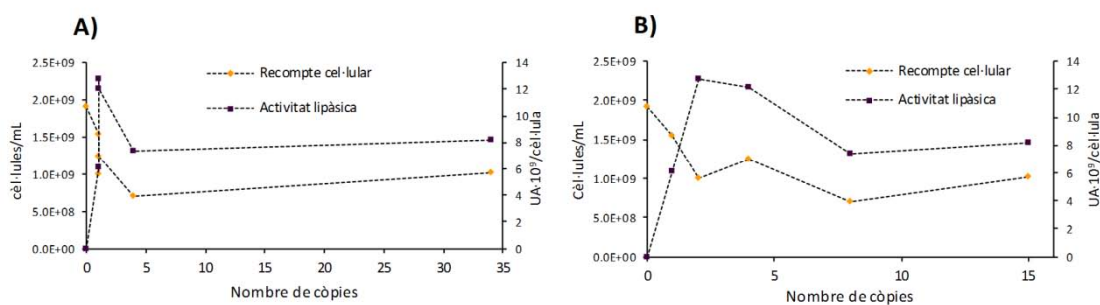


Figura 6. Creixement cel·lular (◆) i activitat lipàsica (■) en el cultiu operats en continu enfront el nombre de còpies de *ROL*. A) Nombre de còpies determinat per qPCR B) Nombre de còpies determinat per ddPCR.

2.4 CONCLUSIONS

En aquest estudi s'ha desenvolupat un mètode basat en la ddPCR per la determinació de la dosi gènica a *P. pastoris*, i s'ha utilitzat de manera satisfactòria per a caracteritzar una sèrie de soques amb diferent nombre de còpies del gen de la lipasa de *Rhizopus oryzae* (*ROL*). La ddPCR ens va permetre ordenar els transformants segons el contingut de còpies gèniques (d'1 fins a 15 còpies) amb una precisió major que el mètode convencional de la qPCR, facilitant així la subsegüent selecció i caracterització clonal. De manera interessant, va ser possible distingir entre els transformants de dosi gènica baixa que diferien només d'una còpia, mentre que amb la qPCR no es podien classificar de manera clara. Aquest fet va facilitar la identificació de la dosi gènica òptima per *ROL* en termes de productivitat, alhora que es va obtenir una reclassificació fiable de soques històriques utilitzades en estudis anteriors. De manera addicional, es va caracteritzar una sèrie de soques amb un rang entre 1 i 15 còpies de *ROL* en condicions definides i controlades de cultiu, establint així la relació entre la dosi gènica i paràmetres fisiològics com els rendiments de biomassa i producte, i les taxes de conversió de substrat. D'una altra banda, l'anàlisi amb citometria de flux dels cultius a petita escala va suggerir que les dosis gèniques altes provoquen nivells elevats de ROS i una viabilitat cel·lular compromesa després de la inducció amb metanol, és a dir, a les etapes inicials de la fase d'inducció. La caracterització futura d'aquesta sèrie de soques mitjançant el perfil transcripcional i altres plataformes "òmiques" permetria incrementar la qualitat de la informació sobre la resposta fisiològica de *P. pastoris* a la producció de proteïnes recombinants.

2.5 REFERÈNCIES

- Abad S, Kitz K, Hörmann A, Schreiner U, Hartner FS, Glieder A. 2010. Real-time PCR-based determination of gene copy numbers in *Pichia pastoris*. *Biotechnol J* 5:413–20.
- Ahmad M, Hirz M, Pichler H, Schwab H. 2014. Protein expression in *Pichia pastoris*: Recent achievements and perspectives for heterologous protein production. *Appl Microbiol Biotechnol* 98:5301–5317.
- Aw R, Polizzi KM. 2013. Can too many copies spoil the broth? *Microb Cell Fact* 12:128.
- Bai J, Swartz DJ, Protasevich II, Brouillette CG, Harrell PM, Hildebrandt E, Gasser B, Mattanovich D, Ward A, Chang G, Urbatsch IL. 2011. A gene optimization strategy that enhances production of fully functional P-Glycoprotein in *Pichia pastoris*. *PLoS One* 6:e22577
- Baumann K, Maurer M, Dragosits M, Cos O, Ferrer P, Mattanovich D. 2008. Hypoxic fed-batch cultivation of *Pichia pastoris* increases specific and volumetric productivity of recombinant proteins. *Biotechnol Bioeng* 100:177–83.
- Bikandi J, Millán RS, Rementeria A, Garaizar J. 2004. In silico analysis of complete bacterial genomes: PCR, AFLP-PCR and endonuclease restriction. *Bioinformatics* 20:798–799.
- Bozdag GO, Greig D. 2014. The genetics of a putative social trait in natural populations of yeast. *Mol Ecol* 23:5061–71.
- Cao Y, Raith MR, Griffith JF. 2015. Droplet digital PCR for simultaneous quantification of general and human-associated fecal indicators for water quality assessment. *Water Res* 70:337–349.
- Clare J, Sreekrishna K, Romanos M. 1998. Expression of Tetanus Toxin Fragment C. In: Higgins, DR; Cregg JM, editor. *Pichia* protocols. Totowa: Humana Press. P 193-208
- Corbisier P, Pinheiro L, Mazoua S, Kortekaas A-M, Chung PYJ, Gerganova T, Roebben G, Emons H, Emslie K. 2015. DNA copy number concentration measured by digital and droplet digital quantitative PCR using certified reference materials. *Anal Bioanal Chem* 407:1831–1840.
- Cos O, Serrano A, Montesinos JL, Ferrer P, Cregg JM, Valero F. 2005. Combined effect of the methanol utilization (Mut) phenotype and gene dosage on recombinant protein production in *Pichia pastoris* fed-batch cultures. *J Biotechnol* 116:321–35.

Costa V, Amorim MA, Reis E, Quintanilha A, Moradas-Ferreira P. 1997. Mitochondrial superoxide dismutase is essential for ethanol tolerance of *Saccharomyces cerevisiae* in the post-diauxic phase. *Microbiology* 143:1649–1656.

Cregg JM, Cereghino JL, Shi J, Higgins DR. 2000. Recombinant Protein Expression in *Pichia pastoris*. *Mol Biotechnol* 16:23–52.

Floren C, Wiedemann I, Brenig B, Schütz E, Beck J. 2015. Species identification and quantification in meat and meat products using droplet digital PCR (ddPCR). *Food Chem* 173:1054–1058.

Gasser B, Saloheimo M, Rinas U, Dragosits M, Rodríguez-Carmona E, Baumann K, Giuliani M, Parrilli E, Branduardi P, Lang C, Porro D, Ferrer P, Tutino ML, Mattanovich D, Villaverde A. 2008. Protein folding and conformational stress in microbial cells producing recombinant proteins: a host comparative overview. *Microb Cell Fact* 7:11.

Gullerova M, Proudfoot NJ. 2010. Transcriptional interference and gene orientation in yeast: Noncoding RNA connections. *Cold Spring Harb Symp Quant Biol* 75:299–311.

Hayden RT, Gu Z, Ingersoll J, Abdul-Ali D, Shi L, Pounds S, Caliendo AM. 2013. Comparison of droplet digital PCR to real-time PCR for quantitative detection of cytomegalovirus. *J Clin Microbiol* 51:540–546.

Hesketh AR, Castrillo JI, Sawyer T, Archer DB, Oliver SG. 2013. Investigating the physiological response of *Pichia (Komagataella) pastoris* GS115 to the heterologous expression of misfolded proteins using chemostat cultures. *Appl Microbiol Biotechnol* 97:9747–9762.

Heyland J, Fu J, Blank LM, Schmid A. 2011. Carbon metabolism limits recombinant protein production in *Pichia pastoris*. *Biotechnol Bioeng* 108:1942–53.

Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, Bright IJ, Lucero MY, Hiddessen AL, Legler TC, Kitano TK, Hodel MR, Petersen JF, Wyatt PW, Steenblock ER, Shah PH, Bousse LJ, Troup CB, Mellen JC, Wittmann DK, Erndt NG, Cauley TH, Koehler RT, So AP, Dube S, Rose KA, Montesclaros L, Wang S, Stumbo DP, Hodges SP, Romine S, Milanovich FP, White HE, Regan JF, Karlin-Neumann GA, Hindson CM, Saxonov S, Colston BW. 2011. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal Chem* 83:8604–10.

- Hindson CM, Chevillet JR, Briggs HA, Gallichotte EN, Ruf IK, Hindson BJ, Vessella RL, Tewari M. 2013. Absolute quantification by droplet digital PCR versus analog real-time PCR. *Nat Methods* 10:1003–5
- Hobson DJ, Wei W, Steinmetz LM, Svejstrup JQ. 2012. RNA Polymerase II Collision Interrupts Convergent Transcription. *Mol Cell* 48:365–374.
- Jordà J, Jouhten P, Cámara E, Maaheimo H, Albiol J, Ferrer P. 2012. Metabolic flux profiling of recombinant protein secreting *Pichia pastoris* growing on glucose:methanol mixtures. *Microb Cell Fact* 11:57.
- Jordà J, Suarez C, Carnicer M, ten Pierick A, Heijnen JJ, van Gulik W, Ferrer P, Albiol J, Wahl A. 2013. Glucose-methanol co-utilization in *Pichia pastoris* studied by metabolomics and instationary ¹³C flux analysis. *BMC Syst Biol* 7:17.
- Lee C, Lee S, Shin SG, Hwang S. 2008. Real-time PCR determination of rRNA gene copy number: Absolute and relative quantification assays with *Escherichia coli*. *Appl Microbiol Biotechnol* 78:371–376.
- Liu H, Qin Y, Huang Y, Chen Y, Cong P, He Z. 2014. Direct Evaluation of the Effect of Gene Dosage on Secretion of Protein from Yeast *Pichia pastoris* by Expressing EGFP. *J Microbiol Biotechnol* 24:144–151.
- Marx H, Mecklenbräuker A, Gasser B, Sauer M, Mattanovich D. 2009. Directed gene copy number amplification in *Pichia pastoris* by vector integration into the ribosomal DNA locus. *FEMS Yeast Res* 9:1260–70.
- Minning S, Serrano a, Ferrer P, Solá C, Schmid RD, Valero F. 2001. Optimization of the high-level production of *Rhizopus oryzae* lipase in *Pichia pastoris*. *J Biotechnol* 86:59–70.
- Quirós C, Herrero M, García LA., Díaz M. 2007. Application of flow cytometry to segregated kinetic modeling based on the physiological states of microorganisms. *Appl Environ Microbiol* 73:3993–4000.
- Resina D, Serrano A, Valero F, Ferrer P. 2004. Expression of a *Rhizopus oryzae* lipase in *Pichia pastoris* under control of the nitrogen source-regulated formaldehyde dehydrogenase promoter. *J Biotechnol* 109:103–113.

Ruth C, Buchetics M, Vidimce V, Kotz D, Naschberger S, Mattanovich D, Pichler H, Gasser B. 2014. *Pichia pastoris* Aft1 - a novel transcription factor, enhancing recombinant protein secretion. *Microb Cell Fact* 13:1–15.

Serrano A, Cereghino GU, Ferrer P, Cregg JM, Valero F. 2001. Overexpression of a *Rhizopus oryzae* lipase in *Pichia pastoris* strains containing multiple copies of the target gene. In: Merten OW, Mattanovich D, Lang C, Larsson G, Neubauer P, Porro D, Postma P, Teixeira de Mattos J, Cole JA, editors. *Recombinant Protein Production with Prokaryotic and Eukaryotic Cells. A Comparative View on Host Physiology*. Dordrecht: Springer Netherlands. p 259-267

Sha C, Yu XW, Lin NX, Zhang M, Xu Y. 2013. Enhancement of lipase r27RCL production in *Pichia pastoris* by regulating gene dosage and co-expression with chaperone protein disulfide isomerase. *Enzyme Microb Technol* 53:438–443.

Sunga AJ, Tolstorukov I, Cregg JM. 2008. Posttransformational vector amplification in the yeast *Pichia pastoris*. *FEMS Yeast Res* 8:870–6.

Te SH, Chen EY, Gin KY-H. 2015. Multiplex assay for two bloom-forming cyanobacteria, *Cylindrospermopsis* and *Microcystis* – A comparison between qPCR and ddPCR. *Appl Environ Microbiol* 81:5203-5211

Van der Heijden RTJM, Romein B, Heijnen JJ, Hellinga C, Luyben KCAM. 1994. Linear constraint relations in biochemical reaction systems: II. Diagnosis and estimation of gross errors. *Biotechnol Bioeng* 43:11–20.

Vanz AL, Lünsdorf H, Adnan A, Nimtz M, Gurramkonda C, Khanna N, Rinas U. 2012. Physiological response of *Pichia pastoris* GS115 to methanol-induced high level production of the Hepatitis B surface antigen: catabolic adaptation, stress responses, and autophagic processes. *Microb Cell Fact* 11:103.

Vassileva A, Chugh DA, Swaminathan S, Khanna N. 2001. Effect of copy number on the expression levels of hepatitis B surface antigen in the methylotrophic yeast *Pichia pastoris*. *Protein Expr Purif* 21:71–80.

Verheijen, PJT. 2010. Data Reconciliation and Error Detection. In: Smolke C, editor. *The Metabolic Pathway Engineering Handbook*. Boca Raton: CRC Press. p 8.1-8.13

Viader-Salvadó JM, Cab-Barrera EL, Galán-Wong LJ, Guerrero-Olazarán M. 2006. Genotyping of recombinant *Pichia pastoris* strains. *Cell Mol Biol Lett* 11:348–359.

Wang NS, Stephanopoulos G. 1983. Application of macroscopic balances to the identification of gross measurement errors. *Biotechnol Bioeng* 25:2177–2208.

Weaver S, Dube S, Mir A, Qin J, Sun G, Ramakrishnan R, Jones RC, Livak KJ. 2010. Taking qPCR to a higher level: Analysis of CNV reveals the power of high throughput qPCR to enhance quantitative resolution. *Methods* 50:271–276.

Zhang B, Xu C, Shao Y, Wang H, Wu Y, Song Y, Li X, Zhang Z, Wang W, Li L, Cai C. 2015. Comparison of droplet digital PCR and conventional quantitative PCR for measuring EGFR gene mutation. *Exp Ther Med* 9:1383–1388

Zhu T, Guo M, Tang Z, Zhang M, Zhuang Y, Chu J, Zhang S. 2009. Efficient generation of multi-copy strains for optimizing secretory expression of porcine insulin precursor in yeast *Pichia pastoris*. *J Appl Microbiol* 107:954–63.

Zhu T, Guo M, Zhuang Y, Chu J, Zhang S. 2011. Understanding the effect of foreign gene dosage on the physiology of *Pichia pastoris* by transcriptional analysis of key genes. *Appl Microbiol Biotechnol* 89:1127–35.

3. L'increment de la dosi gènica dels cassets d'expressió regulats pel promotor *AOX1* comporta una atenuació dels nivells transcripcionals del metabolisme del metanol

Paraules clau: *Pichia pastoris*, producció proteïnes recombinants, dosi gènica, *ROL*, lipasa, metabolisme del metanol, regulació *AOX1*, *Mxr1*, àcids grassos, acetil-coA, biogènesi dels peroxisomes, transcriptòmica.

3.1 INTRODUCCIÓ

El llevat metilotròfic *Pichia (Komagataella) pastoris* és un conegut hoste per la producció de proteïnes, amb una creixent rellevància en la producció de productes recombinants amb interès comercial (Ahmad et al., 2014). Entre d'altres factors, la presència del promotor *AOX1* (P_{AOX1}), d'expressió forta i regulable per metanol, ha contribuït a la popularitat d'aquest sistema d'expressió. En presència del metanol, P_{AOX1} és fortament induït, donant lloc a elevades quantitats de l'enzim alcohol oxidasa (Couderc i Baratti, 1980; Vanz et al., 2012; Rußmayer et al., 2015). Conseqüentment, P_{AOX1} és el promotor més usat per la producció de proteïnes recombinants en *Pichia pastoris* (Cereghino et al., 2000; Vogl i Glieder, 2012). Per tal d'incrementar la producció de proteïna recombinant en els sistemes P_{AOX1} , s'han intentat diferents estratègies, tals com la optimització de les condicions de cultiu i estratègies d'aliment (Cos et al., 2006; Looser et al., 2015), la coexpressió de factors de plegament i/o secreció (Inan et al., 2006; Guerfal et al., 2010; Ruth et al., 2014; Yang et al., 2013) o la optimització de les propietats de la seqüència de nucleòtids forana (ús de codó) (Yadava et al., 2003; Bai et al., 2011; Wang et al., 2014). En aquest context, la inserció de múltiples còpies del gen recombinant s'ha emprat amb èxit (Vassileva et al., 2001; Cos et al., 2005; Zhu et al., 2009; Cámara et al., 2016). No obstant, aquests estudis van descriure que més enllà d'un determinat llindar, un subseqüent increment en la dosi gènica sovint resultava en productivitats més baixes i un sever impacte negatiu en el creixement cel·lular, una reducció en el consum de metanol i una disminució en la viabilitat cel·lular; és a dir, una sobrecàrrega metabòlica. Fins el dia d'avui, la majoria d'estudis pretenien entendre la resposta fisiològica i l'adaptació relacionada amb la producció de proteïnes recombinants en les soques multicòpia, centrant-se en les respostes d'estrès provocades pels alts nivells de proteïna sintetitzada. En el cas de les proteïnes secretades, s'han reportat l'estrès desencadenat per l'acumulació de proteïna recombinant en la via secretora (és a dir, l'UPR, de l'anglès Unfolded Protein Response), l'estrès oxidatiu en el reticle endoplasmàtic (RE) i la degradació de proteïnes al RE (via ERAD, de l'anglès *Endoplasmatic Reticulum Associated protein Degradation*) (Hohenblum et al., 2004; Liu et al., 2014; Azoun et al., 2016). Tanmateix, hi ha una evidència creixent que l'increment en la dosi gènica comporta una sobrecàrrega metabòlica en el metabolisme del carboni central, indicant una limitació en l'abastiment de carboni i recursos energètics (Jordà et al., 2012; Zhu et al., 2011). Tot i que el nombre d'aproximacions a nivell de sistema per tal d'entendre la fisiologia de *P. pastoris* en la producció de proteïnes recombinants va en augment (Graf et al., 2008; Baumann et al., 2011; Hesketh et al., 2013; Liang et al., 2012; Vanz et al., 2012; Ferrer i Albiol, 2014; Rebnegger et al., 2014), l'anàlisi a nivell genòmic per tal de conèixer l'impacte de la dosi gènica ha estat llargament deixat

de banda, especialment pels sistemes basats en P_{AOX1} . En un estudi previ (Cámara et al., 2016), es va construir i caracteritzar una col·lecció de soques amb diferent nombre de còpies de la lipasa de *Rhizopus oryzae* (*ROL*) sota el control de P_{AOX1} , mostrant una disminució en els rendiments de producte i biomassa, així com una alteració en el consum de substrat en les soques multicòpies en comparació amb la soca monocòpia.

En aquest estudi s'ha realitzat un estudi transcriptòmic a nivell global d'una sèrie de soques multicòpia pel gen *ROL* per obtenir un coneixement més ampli de l'impacte de la dosi gènica en la xarxa metabòlica de *P. pastoris* en presència de metanol. Amb l'objectiu de minimitzar els efectes d'altres variables (per exemple, la taxa específica de creixement) aquests experiments es van realitzar en cultius operats en continu sota condicions d'estat estacionari, utilitzant com a font de carboni una barreja de glicerol i metanol, per tal de permetre l'expressió de *ROL*. D'altra banda, es va monitoritzar la producció de Rol a nivell transcripcional (mitjançant la ddPCR) i traduccional (mesura de l'activitat enzimàtica intra/extracel·lular i Western Blot).

3.2 MATERIALS I MÈTODES

3.2.1 Construcció de soques

En aquest estudi es van utilitzar una col·lecció de soques derivades de *P. pastoris* X-33 (Invitrogen, Carlsbad, CA, US), amb un increment en el nombre de còpies del gen de la lipasa de *Rhizopus oryzae* (*ROL*) clonades en el vector pPICZαA (Invitrogen, Carlsbad, CA, US). La metodologia utilitzada per la construcció i caracterització de les soques es va reportar prèviament per Cámara et al. (2016). Breument, es van obtenir diferents vectors pPICZαA amb una, dues o quatre còpies de *ROL* mitjançant la construcció de tàndems *in vitro*, i seguidament es van introduir a la soca X-33 per electroporació per tal de generar les soques multicòpia. Es va seleccionar un conjunt representatiu de soques amb 2, 4, 8 i 15 còpies de *ROL* (denominades 2C, 4C, 8C i 15C, respectivament) i es van caracteritzar posteriorment a nivell transcripcional. La soca X-33 sense transformar (0C) es va utilitzar com a control negatiu, mentre que la soca amb una còpia de *ROL* (1C), obtinguda prèviament per Minning et al. (2001) es va seleccionar com a soca monocòpia.

3.2.2 Cultius operats en continu

Per cada soca es van realitzar triplicats dels cultius operats en continu en reactors d'1 L (DASGIP® Parallel Bioreactor Systems, Eppendorf, Hamburg, Alemanya), amb un volum de treball de 0.35 L. Com a pre-cultiu, es va inocular amb un criovial de la soca de *P. pastoris* seleccionada un Erlenmeyer de 0.5 L amb 50 mL de medi YPD-Zeo (per litre: 10 g extracte de llevat, 20 g peptona, 20 g glucosa, 100 mg Zeocina) a una densitat òptica a 600 nm (DO_{600}) de 0.3, i es va incubar durant aproximadament 24 h a 25°C i 180 rpm en un incubador Infors Orbitron (Infors AG, Bottmingen, Suïssa), per tal d'inocular 0.4 L de medi batch. Un cop assolit el final de la fase batch del cultiu (amb una DO_{600} de 50) es va iniciar la fase en continu, amb una taxa de dilució (D) de 0.1 h^{-1} , amb un cabal d'alimentació de medi definit (controlat per balances) compost per 25 g L^{-1} d'una mescla de glicerol/metanol (60% glicerol / 40% metanol, w/w) coma font de carboni. La composició dels medis per la fase batch i la fase en continu és la mateixa descrita per Cámara et al. (2016). L'antiescumant (Glanapon 2000 Konc, Bussetti, Viena, Àustria) es va addicionar de manera constant a una taxa de $13 \text{ } \mu\text{L h}^{-1}$. La temperatura del cultiu es va mantenir constant a 25°C, el pH es va controlar a 5 mitjançant amoni 10% i la concentració d'oxigen dissolt es va mantenir al 20% controlant la velocitat de l'agitador fins a un màxim de 1250 rpm i un cabal

constant d'aire de 21 L h⁻¹. Les mostres es van prendre després de cinc temps de residència, un cop assolit l'estat estacionari.

3.2.3 Procediments analítics

3.2.3.1 Determinació de la biomassa

Els nivells de biomassa durant la fase en continu es van monitoritzar mesurant la DO₆₀₀. La concentració de biomassa també es va determinar per cada soca en triplicat mitjançant el pes sec (DCW, de l'anglès *Dry Cell Weight*). Es van centrifugar 5 mL del brou de cultiu a 4100 rpm durant 5 min, i seguidament es van resuspendre en 5 mL d'aigua desionitzada. Les cèl·lules es van rentar dues vegades i es van assecar en vasos de precipitats (prèviament pesats) a 105 °C i 24 h.

3.2.3.2 Determinació de l'activitat lipolítica

L'assaig d'activitat lipolítica es va realitzar per triplicat per cada experiment independent, tal i com es descriu a Resina et al. (2004).

3.2.3.3 Disrupció cel·lular

L'activitat lipàsica intracel·lular es va mesurar després de disruptar les cèl·lules mecànicament, utilitzant el mètode descrit per García-Ortega et al. (2016). Es va obtenir un extracte cel·lular soluble i insoluble, corresponent al lisat cel·lular clarificat i la fracció associada a la membrana, respectivament. L'eficiència en la disrupció cel·lular es va calcular amb la següent equació:

$$E (\%) = 100 \cdot \frac{x_i - x_f}{x_i}$$

on E (%) és el grau de disrupció; x_i , el nombre de cèl·lules abans de la disrupció; x_f , el nombre de cèl·lules no disruptades després de la disrupció. Les mesures es van realitzar de mostres de tres experiments independents per cada soca.

3.2.3.4 Determinació de la proteïna total

La quantificació del contingut de proteïna total del sobrenedant i les fraccions del lisat es va determinar amb el kit Pierce BCA Protein Assay, seguint les instruccions del fabricant. Com a estàndard de proteïna per la corba de calibratge es va utilitzar albúmina (BSA, de l'anglès *Bovine Serum Albumin*). L'absorbància es va mesurar en plaques de 96 pous, utilitzant l'espectrofotòmetre Multiskan™ FC (ThermoFisher Scientific, Waltham, MA, US). Tots els

reactius es van adquirir a ThermoFisher Scientific (Waltham, MA, US). Les determinacions es van realitzar per duplicat per cada experiment independent.

3.2.3.5 Quantificació dels metabòlits

El glicerol, el metanol, i altres compostos extracel·lulars potencials en els cultius operats en continu es van analitzar tal i com es descriu a Cámara et al. (2016). Es van dur a terme duplicats per cada experiment independent.

3.2.3.6 Composició d'àcids grassos

La quantificació del contingut total d'àcids grassos en les mostres dels cultius en continu es van realitzar per BIOCRATES Life Science AG (Innsbruck, Àustria). Breument, es van tractar mostres de cèl·lules liofilitzades de cada cultiu en continu amb una solució de metanol-HCl a temperatures elevades durant un temps prolongat, per tal de convertir completament els àcids grassos lliures i associats a membranes en els corresponents metil èsters, i representar així el contingut total d'àcids grassos. El contingut individual d'àcids grassos a les mostres es va determinar mitjançant el metil èster derivat corresponent (FAME's, de l'anglès *Fatty Acid Methyl Ester*), utilitzant una cromatografia de gasos acoblada a una detecció per espectrometria de masses (Agilent 7890 GC / 5975 MSD). Amb l'objectiu de calcular les concentracions corresponents, es van utilitzar corbes de calibratge d'estàndards externs.

3.2.4 Quantificació de Rol per Western Blot

Per tal de determinar el contingut de lipasa de *R. oryzae* (Rol), el sobrenedant dels cultius en continu, i les fraccions soluble i insoluble dels lisats es van analitzar mitjançant Western Blot (WB). A fi d'obtenir resultats fiables a nivell quantitatiu, es van seguir les directrius recomanades per Taylor et al (2013). Les mostres es van diluir 3:1 en tampó Laemmli i es van incubar a 95 °C durant 10 min, seguit de 10 min en gel. Es van carregar les mostres en un gel 12% SDS-PAGE (de l'anglès *Dodecyl Sulphate–PolyAcrylamide Gel Electrophoresis*), en una cubeta Mini-PROTEAN II utilitzant els gels Mini-PROTEAN® TGX Stain-Free™ Protein Gels, seguint els protocols recomanats pel fabricant. Com a marcadors de pes molecular es van utilitzar els estàndards Precision Plus Protein™ Unstained i All Blue Standards, per la visualització del gel SDS-PAGE i la membrana WB, respectivament. La visualització del gel es va realitzar després d'una activació amb UV amb el sistema Gel Doc EZ. Per tal de dur a terme els Western Blots després del gel SDS-PAGE, les proteïnes es van transferir a una membrana de nitrocel·lulosa (10 min, 2.5 A, 25 V) amb l'aparell Trans-Blot® Turbo™ Transfer System. La membrana es va bloquejar durant 1 h amb

una solució 5% de llet desnatada en pols, 1 mL de Tween 20, seguit d'una incubació amb sèrum de ratolí anti-Rol 1:10000 (obtingut del Servei de Cultius Cel·lulars, Producció d'Anticossos i Citometria, Universitat Autònoma de Barcelona, Bellaterra, Espanya) durant 1 h. Després dels rentats, la membrana es va incubar amb una solució 1:1000 de IgG policlonal anti-ratolí conjugat amb HRP (Sigma Aldrich, St Louis, MO, EUA) durant 1 h. La detecció es va realitzar incubant la membrana durant cinc minuts amb el reactiu Clarity™ Western ECL Substrate. L'anàlisi de les imatges i la corresponent quantificació es va dur a terme amb el dispositiu Molecular Imager® ChemiDoc™ XRS System i el software ImageLab v.5.2. Les quantificacions per cada soca es van realitzar de mostres provinents de tres experiments independents. Si no és especificat, tots els reactius es van adquirir a Bio-Rad (Hercules, CA, EUA).

3.2.5 Determinació dels nivell transcripcionals de *ROL* per ddPCR

Els nivells d'expressió de *ROL* es van quantificar mitjançant la ddPCR (de l'anglès *Digital Droplet PCR*) en mostres de RNA total extretes pels microarrays (veure secció a continuació). Per tal de sintetitzar el cDNA es va utilitzar el iScript™ cDNA Synthesis kit (Bio-Rad, Hercules, CA, US) seguint les indicacions del fabricant. Cada reacció de PCR es va preparar en un volum final de 20 µL, amb 10 µL de QX200™ ddPCR™ EvaGreen Supermix, 200 nM de l'oligonucleòtid *forward*, 100 nM of de l'oligonucleòtid *reverse*, 0.4 ng de cDNA i la quantitat necessària d'aigua *Dnase/Rnase-free*. Les gotes es van formar utilitzant el Droplet Generator QX200™ i posteriorment es van transferir a una placa de 96 pous. Les reaccions es van incubar a 95°C durant 10 min i subsegüentment es van sotmetre a un primer pas de desnaturalització (94°C, 30 s), seguit d'un pas d'alineament/extensió (60.2 °C, 1 min, pels oligonucleòtids de *ROL*; 56.5°C, 1 min, pels oligonucleòtids d'*ACT1*) durant un total de 40 cicles. La detecció de les gotes es va completar mitjançant el dispositiu QX100™ Droplet Digital Reader i el software QuantaSoft v. 1.7.4.0917. Amb la finalitat de normalitzar les dades obtingudes, es va seleccionar el gen *house-keeping* β -actin (*ACT1*). D'aquesta manera, l'expressió de *ROL* es va calcular per cada mostra en duplicat, normalitzant els resultats calculant la ràtio entre les gotes positives de *ROL* i les d'*ACT1*. Els reactius per la ddPCR es van obtenir de Bio-Rad (Hercules, CA, EUA), mentre que els oligonucleòtids es van ser sintetitzats per Sigma Aldrich (St Louis, MO, EUA). La seqüència dels oligonucleòtids *ROL* i *ACT1* es mostra a l'annex.

3.2.6 Determinació del nombre de còpies de *ROL* per ddPCR

Per tal de verificar l'estabilitat de la dosi gènica de *ROL*, es van prendre mostres de biomassa dels cultius operats en continu després de cinc temps de residència, un cop s'havia assolit l'estat estacionari. La dosi gènica es va analitzar per ddPCR utilitzant el mètode descrit prèviament per Cámara et al. (2016).

3.2.7 Anàlisi estadístic

Per cada soca es van realitzar tres cultius en continu (rèpliques biològiques), des de precultius independents. Les dades fisiològiques de creixement es van verificar utilitzant procediments estàndards de consistència i reconciliació, tal i com es va descriure prèviament a Cámara et al. (2016). Les dades s'expressen com la mitjana \pm desviació estàndard (SD, de l'anglès Standard Deviation). L'anàlisi estadístic dels resultats es va realitzar mitjançant un test ANOVA seguit d'una prova *t* de Student (mostres no aparellades), amb el software Excel de Microsoft. Es va considerar estadísticament significatiu un valor *p* (*p-val*) inferior a 0.05.

3.2.8 Extracció del RNA i hibridació dels microarrays

Per l'anàlisi transcriptòmic es va extreure el RNA total de tres rèpliques biològiques per cada soca (provinents de cultius en continu independents). A les mostres dels cultius en continu es va afegir immediatament una solució fixadora freda (5% v/v de fenol en etanol absolut) en una ràtio de 2:1, i es van centrifugar a 10,000 x g durant 1 min. Els pèl·lets es van guardar a -80°C fins el processament pertinent. L'aïllament del RNA es va dur a terme utilitzant el reactiu TRI, d'acord amb les indicacions del fabricant (Ambion, CA, EUA). La integritat del RNA i la seva concentració es van analitzar amb els xips RNA Nano (Agilent, Santa Clara, CA, EUA) i el Nanodrop (Thermo Scientific™, Waltham, MA, EUA), respectivament. El marcatge, la hibridació dels microarrays específics de *P. pastoris* (AMAD-ID 034821, 8x15K custom arrays, Agilent) i l'escanejat es van realitzar segons el protocol Agilent's Two-Colour Microarray-Based Gene Expression Analysis (Low Input Quick Amp Labelling Kit). Totes les mostres es van marcar de forma *dye-swap*, i es van hibridar contra un cDNA de referència, el qual provenia d'una mescla de cèl·lules crescudes en diferents condicions de cultiu (descrit en Graf et al, 2008). La normalització i l'anàlisi estadístic de les dades obtingudes dels microarrays va consistir en rectificar la desviació del color utilitzant el mètode LOESS (de l'anglès *locally weighted MA-*

scatterplot smoothing) seguit d'una normalització amb el mètode *Aquantile*. El *background* dels resultats dels microarrays no es va corregir. Els *p-val* obtinguts es van ajustar mitjançant el mètode de Benjamini i Yekutieli (Reiner et al., 2003). Per cada soca, es va calcular el Log_2 *fold change* (FC) envers la soca control no productora (OC), utilitzant el paquet Limma del software R (Smith GK, 2005). Per tal de considerar un gen expressat de manera diferencial, es va establir un límit de $-0.58 > \text{Log}_2\text{FC} > 0.58$ (equivalent a un *fold change* de 1.5) i un *p-val* < 0.05 . Les dades dels microarrays estan disponibles en la base de dades ArrayExpress (<http://www.ebi.ac.uk/arrayexpress>) amb el número d'accés E-MTAB-4582.

3.2.9 Anàlisi transcriptòmic

2.2.9.1 Diagrames de Venn

Els diagrames de Venn es van crear utilitzant l'aplicació web <http://bioinformatics.psb.ugent.be/webtools/Venn/>

3.2.9.2 Anàlisi de clústers

L'anàlisi de clústers de les dades transcriptòmiques es va realitzar per tal d'agrupar els gens regulats amb tendències d'expressió similars. L'anàlisi es va dur a terme utilitzant els valors de log_2FC amb l'algoritme *k-means* (MacQueen, 1967) i aplicant una distància Euclidiana, implementat al software Genesis v.1.7.7 (Sturn et al., 2002).

3.2.9.3 Anàlisi Principal de Components

L'anàlisi principal de components (PCA, de l'anglès *Principal Component Analysis*) es va utilitzar per identificar patrons amagats al conjunt de dades i identificar com es correlacionaven entre elles les variables (Yeung KY., Ruzzo WL., 2001). L'anàlisi es va realitzar amb el plug-in d'Excel XLSTAT.

3.2.9.4 Enriquiment funcional

Les xarxes funcionals dels termes enriquits (utilitzant les bases de dades KEGG i GO) en comparació amb la soca control es van visualitzar amb el plug-in ClueGO (v. 2.1.7) (Bindea et al., 2009) del software Cytoscape (v. 3.2.1). Els arxius d'ontologia i anotació funcional per l'anàlisi amb la base de dades GO es van descarregar a data del 24.09.2015, mentre que per la base de dades KEGG es van fer servir les publicades a data del 25.09.2015. L'interval en l'esquema GO va ser del 8-15, i el nombre mínim de gens per agrupar es va establir en dos. Amb l'objectiu d'identificar els termes GO i KEGG sobrerrepresentats es va utilitzar el test hipergeomètric doble

amb un *p-val* de 0.05, i el mètode Benjamini-Hochberg (amb un paràmetre kappa de 0.4) es va aplicar per corregir la taxa de fals descobriment (FDR, de l'anglès *False Discovery Rate*). Les representacions en forma de roda de les principals categories funcionals es van representar amb el software Excel de Microsoft. En aquestes representacions, l'interval de l'esquema GO va ser de 7-10, amb un mínim de tres gens per grup i un mínim del 5% de representació de tots els gens associats al terme GO corresponent.

3.3 RESULTATS I DISCUSSIÓ

3.3.1 Dosis gèniques altes de *ROL* provoquen reduccions al creixement cel·lular i al consum de metanol

Un conjunt de 6 soques de *P. pastoris* amb un nombre creixent de còpies de *ROL*, prèviament seleccionades d'una col·lecció de transformants (Cámara et al., 2016) es va fer créixer en glicerol-metanol en cultius operats en continu a una taxa de dilució de 0.1 h⁻¹. En aquestes condicions, els cultius eren limitats per glicerol, permetent així la derepressió i la inducció per metanol dels gens implicats en el metabolisme del metanol. Comparant diferents paràmetres relacionats amb el creixement (Taula I), es va poder observar un clar impacte fisiològic a mesura que la dosi gènica de *ROL* augmenta, en concordança amb els resultats obtinguts prèviament (Cámara et al., 2016). Concretament, les concentracions de biomassa van presentar una correlació negativa amb la dosi gènica de *ROL*, amb una reducció del 30-35% en les soques multicòpia en comparació a la soca control no productora, també reflectit en una disminució significativa (5-15%) dels rendiments de biomassa. En termes de consum de la font de C, la taxa de consum de glicerol va ser 1.5 vegades superior en aquestes soques en comparació amb la soca control 0C, mentre que el consum específic de metanol va caure bruscament de 2.01 mmol g⁻¹ h⁻¹ (soca referència) fins 1.08 mmol g⁻¹ h⁻¹ en la soca 15C (i amb taxes encara més baixes a les soques 4C i 8C), acompanyat de quantitats significatives de metanol residual en el medi de cultiu (Taula I i Annex).

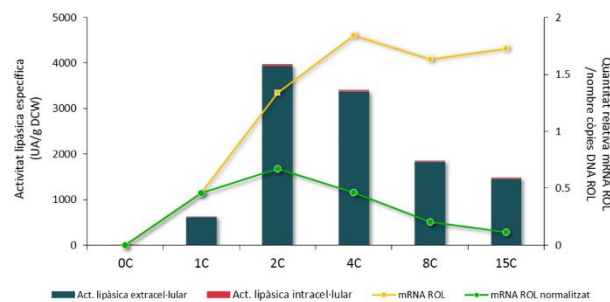
Taula I Paràmetres macroscòpics de creixement reconciliats de les soques productores de Rol i la soca referència, en cultius operats en continu. Taxes de consum de substrat i producció de biomassa per cada soca. L'índex de consistència h era inferior a 5 a tots els casos (95% de nivell de confiança). Els resultats es mostren com la mitjana ± sd basat en tres experiments independents per cada soca. q_s: taxa reconciliada de consum específic de substrat; Y_{x/s}: rendiment reconciliat biomassa/substrat.

	q _s , glicerol (mmol g ⁻¹ DCW h ⁻¹)	q _s , metanol (mmol g ⁻¹ DCW h ⁻¹)	DCW (g ⁻¹ L ⁻¹)	Y _{x/s} (g ⁻¹ DCW C-mol ⁻¹)	Metanol residual (g/L)	Glicerol residual (g/L)
0C	-1.12 ± 0.01	-2.01 ± 0.11	12.82 ± 0.16	16.84 ± 0.53	0.11 ± 0.10	0.04 ± 0.02
1C	-1.20 ± 0.03	-2.22 ± 0.10	12.03 ± 0.11	15.97 ± 0.18	0.10 ± 0.09	0.03 ± 0.01
2C	-1.51 ± 0.03	-1.59 ± 0.14	8.99 ± 0.15	14.35 ± 0.08	4.39 ± 0.28	0.04 ± 0.01
4C	-1.71 ± 0.09	-0.62 ± 0.21	8.66 ± 0.41	16.18 ± 1.04	7.07 ± 0.19	0.03 ± 0.03
8C	-1.75 ± 0.05	-0.67 ± 0.12	8.26 ± 0.10	15.37 ± 0.14	7.52 ± 0.23	0.03 ± 0.10
15C	-1.73 ± 0.12	-1.08 ± 0.24	8.55 ± 0.34	14.92 ± 0.13	5.90 ± 0.92	0.02 ± 0.02

3.3.2 Els nivells baixos de *ROL* a les soques multicòpia podrien ser deguts a una transcripció ineficient

La quantificació de la producció de lipasa en *P. pastoris* es va realitzar a diferents nivells. Primer, es va mesurar la lipasa secretada mitjançant l'activitat lipolítica als sobrenedants dels cultius. El perfil d'activitat lipàsica específica (UA/g DCW) obtingut per la col·lecció de soques (Figura 1A) coincidí amb els resultats prèviament reportats (Cámara et al., 2016).

A)



B)

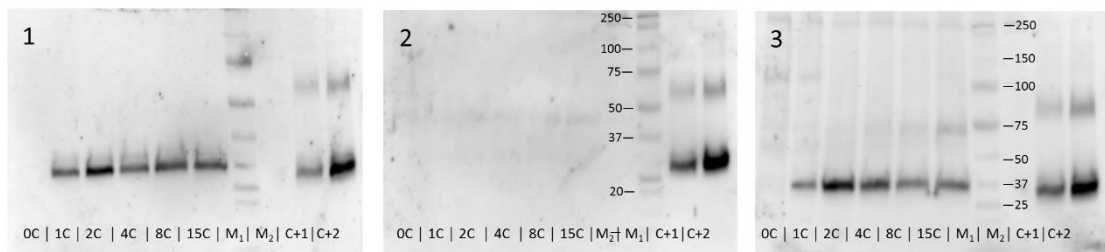


Figura 1. Anàlisi de la producció de Rol a diferents nivells. A) En barres, activitat lipàsica extracel·lular i intracel·lular de les soques recombínants. El gràfic de línies indica els nivells d'expressió de mRNA. B) Anàlisi per Western Blot de tres mostres representatives de diferents fraccions cel·lulars al final dels cultius en continu. A la membrana 1 les mostres 2C i 4C estaven diluïdes en proporció 1:2 per tal d'evitar una saturació del senyal. Membrana 1. Sobrenedant del cultiu cel·lular. Membrana 2. Fracció soluble dels lisats. Membrana 3. Fracció insoluble dels lisats. M1: marcador visible; M2: marcador no tenyit; C+1, C+2: controls positius, 9 ng i 18 ng de Rol purificada, respectivament.

Breument, els nivells òptims de producció es van obtenir en la soca 2C, mentre que increments superiors en la dosi gènica comportaven activitats lipàsiques específiques inferiors. Concretament, els nivells d'activitat aconseguits per la soca 2C eren 6 vegades superior comparats amb la soca monocòpia, mentre que les soques amb un nombre superior de còpies de *ROL* (8C i 15C) exhibien millores més modestes comparant amb 1C (3 i 2 vegades més, respectivament). Cal destacar que l'activitat lipàsica intracel·lular només va representar un 2%

del total d'activitat produïda per les soques, independentment de la dosi gènica de *ROL*. En segon lloc, per tal de determinar si els perfils d'activitat lipàsica intra i extracel·lular es correlacionaven amb el contingut de Rol (és a dir, la massa de Rol), es van analitzar per Western Blot les fraccions solubles i insolubles dels lisats cel·lulars, juntament amb els sobrenedants dels cultius. No es va detectar acumulació de lipasa en la fracció soluble dels lisats, tal i com es mostra a la Figura 1B. D'una altra banda, es van detectar nivells significatius de lipasa acumulada en la fracció insoluble (la fracció associada a les membranes, la qual conté els orgànuls de la via secretora) (Taula II). No obstant, sembla que no hi ha diferències entre les soques de baix i alt rang de còpies de *ROL*. En totes les soques testades el contingut de lipasa detectat per WB representava un 30-40% del total màssic de Rol (contingut extracel·lular + intracel·lular).

Taula II. Quantificació del contingut de Rol. Les dades es mostren com la mitjana \pm sd, basat en mesures per triplicat (una per cada cultiu en continu realitzat). *P < 0.05 (en comparació a la soca referència en el cas del contingut de proteïna, comparat enfront la soca monocòpia en el resta de casos). SN, sobrenedant; FI, fracció insoluble; FS, fracció soluble.

	Contingut proteïna SN (mg proteïna/g DCW)	Contingut Rol SN (μ g Rol/g DCW)	Contingut Rol FI (μ g Rol/g DCW)	Activitat lipàsica SN (UA/ μ g Rol)	Activitat lipàsica FS (UA/mg proteïna)
0C	1.74 \pm 0.41	-	-	-	-
1C	1.70 \pm 0.03	24.16 \pm 13.13	14.93 \pm 3.65	29.08 \pm 10.18	0.09 \pm 0.03
2C	2.56 \pm 0.19 *	175.20 \pm 18.90 *	76.01 \pm 10.83 *	22.64 \pm 3.91	0.29 \pm 0.05 *
4C	2.69 \pm 0.08 *	146.86 \pm 29.38 *	63.33 \pm 17.05 *	23.62 \pm 5.90	0.30 \pm 0.07 *
8C	2.24 \pm 0.11 *	52.54 \pm 7.66 *	35.29 \pm 9.69 *	28.93 \pm 2.97	0.23 \pm 0.05 *
15C	2.29 \pm 0.08 *	97.00 \pm 3.03 *	58.55 \pm 13.39 *	18.35 \pm 2.64	0.27 \pm 0.03 *

Tot i que l'activitat proteàsica extracel·lular no es va mesurar als sobrenedants, les condicions operacionals (pH, temperatura, font de nitrogen) havien estat seleccionades per tal de minimitzar la lisi cel·lular i la secreció de proteases, tal i com s'havia descrit prèviament (Gao i Shi, 2013). Altrament, no es van detectar bandes de degradació en els WB de les fraccions intracel·lulars (Figura 1B), reforçant la hipòtesi de que la degradació preteolítica (per exemple, a través de la via ERAD) no és el principal coll d'ampolla en les soques amb alt nombre de còpies de *ROL*. Encara més, la relació constant entre el contingut extracel·lular de Rol i l'activitat lipolítica a totes les soques indica que no hi ha inactivació de la lipasa. De manera consistent, els nivells específics d'activitat lipàsica (UA Rol/mg proteïna total secretada) al sobrenedant es

correlaciona amb les quantitats relatives de Rol (mg Rol secretada/mg proteïna total secretada), sent del 0.5%, 7% i 3% per la soca monocòpia, la soca 2C i les soques d'alt rang de còpies, respectivament. En conjunt, aquestes observacions indiquen que la retenció intracel·lular de Rol no és la principal causa dels nivells inferiors de Rol extracel·lular en les soques amb alta dosi gènica (4, 8 i 15 còpies). Això contrasta amb altres resultats prèviament reportats, on les soques amb alta dosi gènica acumulaven quantitats superiors de proteïna recombinant a l'interior de la cèl·lula comparant amb les soques amb baix nombre de còpies (Ben Azoun et al., 2016; Yang et al., 2016). De manera similar, estudis previs amb *P. pastoris* expressant *ROL* sota el control del promotor *FDL1* mostraren que una sobreexpressió de *ROL* en cultius operats en fed-batch resultaven en una acumulació significativa de Rol (fins a 700 UA Rol/mg proteïna total) en la fracció soluble dels lisats cel·lulars després de 55 h d'inducció (Resina et al., 2005), provocant així una inducció de la UPR (Resina et al., 2007; Resina et al., 2009). Tanmateix, aquests valors eren tres ordres de magnitud superior als del present estudi, és a dir, en condicions d'estat estacionari en cultius operats en continu. Per això, és plausible que en aquestes condicions els valors de estrès relacionat amb el RE siguin negligibles o molt baixos.

Amb relació als nivells transcripcionals de *ROL*, es va observar un increment de 4 vegades entre les soques 1C i la 4C (Figura 1A), mentre que subseqüents increments en la dosi gènica (soques 8C i 15C) no va resultar en nivells superiors de mRNA de *ROL*. Aquests resultats coincideixen amb diversos casos reportats prèviament, on s'observava una relació amb forma de *plateau* (Liu et al., 2014; Sha et al., 2013; Huang et al., 2014; Yang et al., 2016a). No obstant, aquest patró no sembla que sigui una norma general pels sistemes P_{AOX1} , ja que d'altres estudis descriuen una correlació positiva entre el nombre de còpies i els nivells de mRNA quan s'incrementa la dosi gènica (Yang et al., 2016b; Ben Azoun et al., 2016). Sorprenentment, la normalització a cada soca dels nivells de mRNA amb el nombre de còpies de *ROL* va resultar en una tendència anàloga al perfil d'activitat lipàsica extracel·lular (Figura 1A). Aquest fet suggereix una limitació transcripcional en l'expressió de *ROL* a mesura que s'augmenta el nombre de còpies, i no tant una limitació en el procés de plegament i/o secreció. Cal destacar que es va verificar la dosi gènica al final de tots els cultius operats en continu (després de 5 temps de residència, equivalent a 7 generacions cel·lulars), i no es va detectar cap pèrdua de còpies de *ROL* en les soques (dades no mostrades), descartant així que els nivells d'expressió més baixos detectats en les soques d'alt rang de còpies fossin deguts a una inestabilitat gènica (pèrdua dels cassets d'expressió) al llarg de les generacions en els cultius en continu.

3.3.3 Perfil transcripcional de les soques multicòpia

Els nivells d'expressió de 707 gens diferents va canviar ($\text{Log}_2\text{FC} \pm 0.58$, $p\text{-val} < 0.05$) com a mínim en una de les soques productores de *ROL*, quan aquestes es van comparar amb la soca no productora. El nombre total de gens regulats va incrementar amb la dosi gènica de *ROL* de manera general per totes les soques, excepte per la 15C, que va mostrar un nombre de gens regulats inferior amb relació a la soca 8C (Figura 2A, requadre superior). Així mateix, la fracció de gens amb expressió disminuïda (comparat amb la soca no productora) va augmentar en les soques d'alt rang de còpies, d'un 34% per 1C fins un 54% en la soca 15C. Quatre dels gens més regulats a les soques 2C i 4C (les soques més productores) eren idèntics (Figura 2A, requadre inferior). *ACS2* (codificant per la sintasa acetil-coA, que actua com a font d'acetil-coA al nucli per l'acetilació d'histones) va ser el gen amb expressió més augmentada en aquestes soques, amb un Log_2FC superior en les soques multicòpia. *FRE2* i *FRE3*, codificants per reductases de ferro i coure que faciliten la captació de metall per la cèl·lula, també van presentar una expressió augmentada en totes les soques, d'igual manera que l'aldehid deshidrogenasa *ALD6-2*. De manera sorprenent, a totes les soques multicòpia el gen amb expressió més disminuïda va ser el de l'alcohol oxidasa *AOX1*. També es va observar una regulació negativa als transportadors d'acetat (codificats pels gens *PAS_chr1-1_0417* i *PAS_chr1-1_0418*) a totes les soques multicòpia. Curiosament, el gen anotat com la floculina *FLO11* (una glicoproteïna de la superfície cel·lular requerida per la formació de pseudohifes i la floculació de *Saccharomyces cerevisiae*) va presentar un fort augment en la seva expressió a la soca 1C, mentre que en les soques multicòpia va presentar una regulació negativa, particularment en les soques 4C, 8C i 15C. No obstant, no es van observar canvis morfològics sota el microscopi a cap de les soques (dades no mostrades).

Per tal d'avaluar l'impacte de la dosi gènica de *ROL* en el perfil d'expressió gènica, es van analitzar amb diagrames de Venn la soca monocòpia (1C), la millor productora (2C) i la soca amb un nombre més alt de còpies (15C). La Figura 2B mostra com només una petita fracció de gens eren regulats a totes les soques. Respecte els gens amb expressió augmentada (comparant amb la soca 0C), un 14% dels 265 eren comuns per totes tres soques, comprenent principalment la síntesi d'àcids grassos i la β -oxidació. D'una altra banda, un 7% dels gens van presentar una expressió disminuïda a totes les soques, relacionats amb el metabolisme de l'acilglicerol, les feromones i la senyalització cel·lular. Mentre que un 30% dels gens regulats eren comuns per les soques 2C i 15C, el grup més prominent estava format per gens regulats (tant positiva com negativament) únicament a la soca 15C.

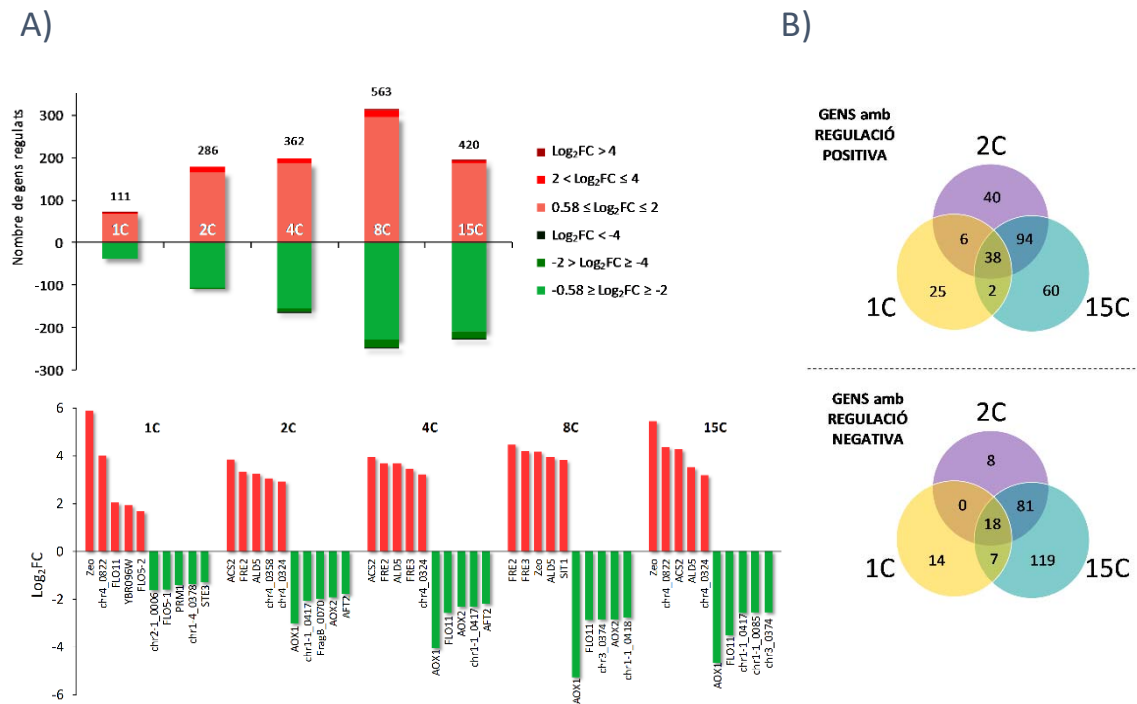


Figura 2. Anàlisi transcriptòmic global dels gens regulats a les soques productores de ROL. Es consideren gens regulats aquells amb un \log_2FC de 0.58 comparat amb la soca control 0C i un $p\text{-val} \leq 0.05$. A) Panel superior: nombre de gens regulats a cada soca. Els valors negatius indiquen els gens d'expressió disminuïda (color verd), mentre que els valors positius designen gens amb expressió augmentada (color vermell). El número sobre les barres denota el nombre total de gens regulats a cada soca. Panel inferior: classificació dels gens més regulats. B) Diagrames de Venn on es mostra el nombre de gens específics o comuns entre les soques 1C (soca productora control), 2C (nivells més alts de producció) i 15C (nombre més elevat de còpies).

Amb l'objectiu d'analitzar més detalladament la variació de l'expressió gènica a totes les soques es va realitzar un Anàlisi Principal de Components (PCA; Figura 3A i 3B). El primer i el segon component (PCA1 i PCA2) explicaven un total del 93% de la variabilitat, agrupant totes les soques multicòpia i separant-les de la soca monocòpia. La representació de PCA2 enfront PCA3 (Figura 2D) va permetre reordenar les soques multicòpia, mostrant un comportament diferent de 8C (amb el nombre més elevat de gens regulats) comparat amb les altres soques.

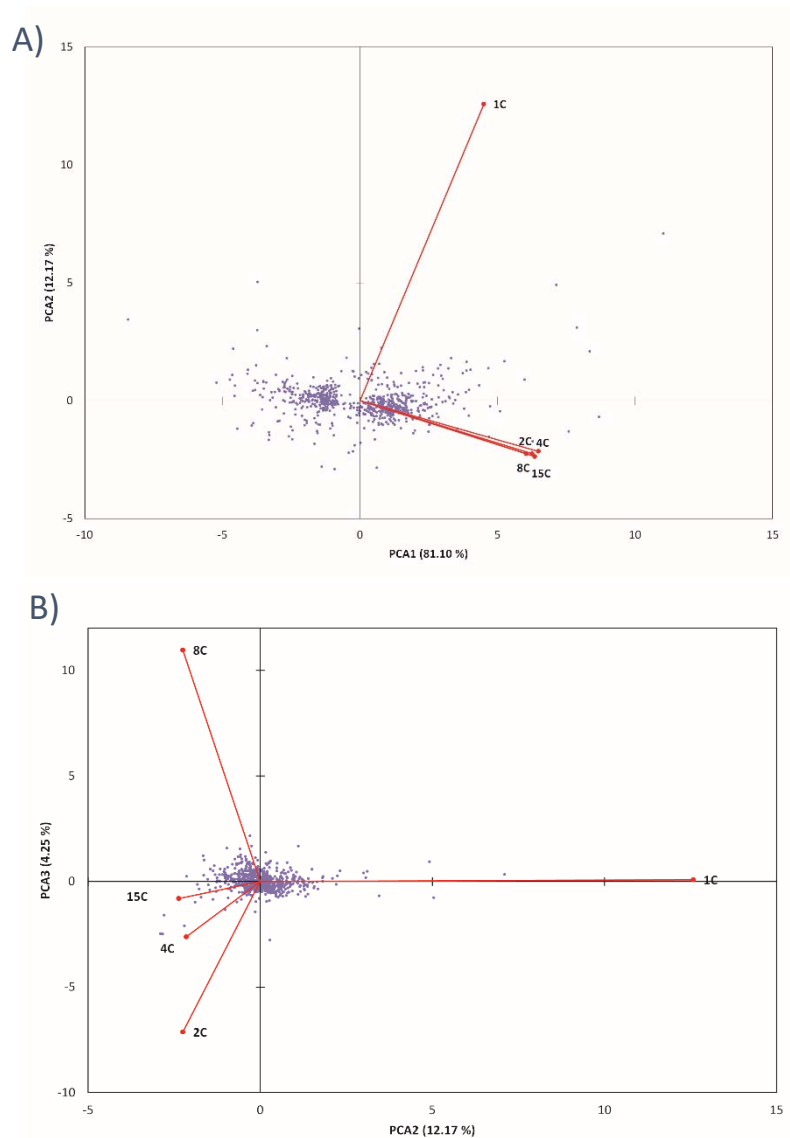
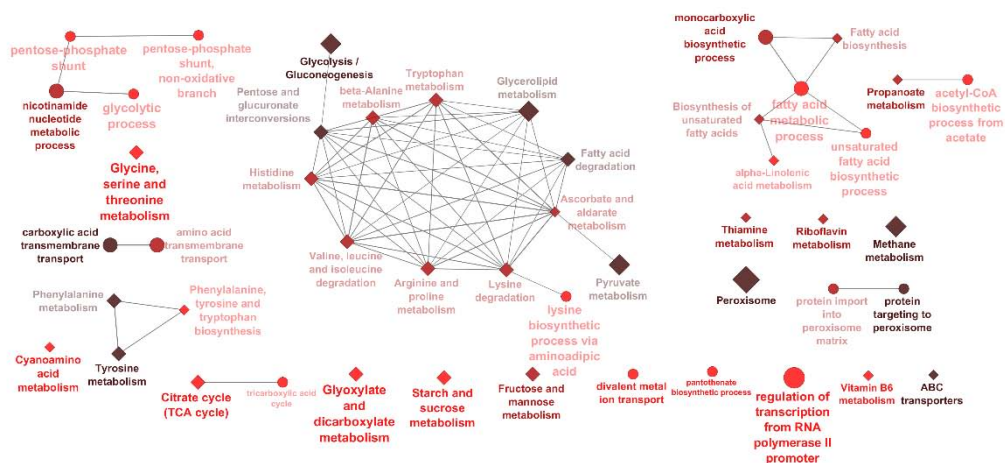


Figura 3. Anàlisi Principal de Components (PCA) de les dades d'expressió. Els gens es representen com a punts (observacions), mentre que les línies vermelles indiquen la projecció de la tendència de cada soja. A) Primer (PCA1) i segon (PCA2) components principals B) Representació del segon i el tercer (PCA3) component.

3.3.4 Enriquiment funcional

Es va realitzar un anàlisi per enriquiment funcional dels 707 gens regulats per tal d'identificar els principals processos biològics implicats i reconèixer la relació entre ells. Es va utilitzar el software ClueGO amb les bases de dades Gene Ontology (GO) i la Kyoto Encyclopaedia of Genes and Genomes (KEGG) (Figura 4A).

A)



B)

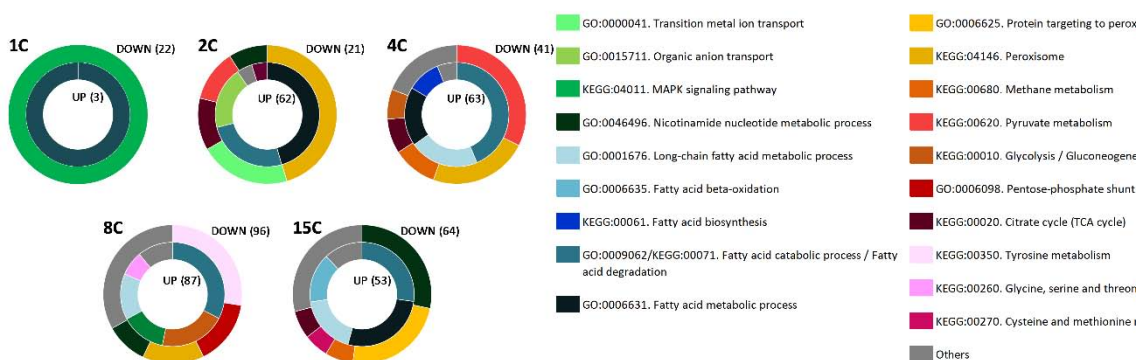


Figura 4. Anàlisi funcional dels gens regulats. A) Representació en xarxa de les categories funcionals enriquides entre les diferents soques productores de *ROL*, generat amb el software ClueGO. La mida dels nòduls reflecteix el nombre de gens associats a cada terme, mentre que el gradient de color indica el nivell de significació (amb correcció de Bonferroni-Hochberg). Es van utilitzar les bases de dades GO i KEGG, representades com a cercles i rombes, respectivament. Les xarxes individuals es troben agrupades segons les vies metabòliques. B) Representacions en forma de roda de les principals categories funcionals de cada soca. El cercle exterior i interior fa referència als gens amb expressió negativa i positiva, respectivament. Els números entre parèntesi indiquen el nombre total de gens anotats.

Els processos estadísticament sobrerrepresentats ($p\text{-val} < 0.05$, amb correcció Benjamini-Hochberg) es van agrupar en 6 categories: I) transport de metalls/poliamines (39 gens) II) metabolisme del metanol/peroxisomes (56 gens) III) metabolisme dels aminoàcids (74 gens) IV) metabolisme dels àcids grassos (50 gens) V) regulació de la transcripció (39 gens) i VI) metabolisme del carboni central/ reaccions redox (82 gens). D'altra banda també es van

analitzar a totes les soques els perfils d'expressió dels gens implicats en la UPR. A la figura 4B es mostra un resum dels processos regulats a cada soca. Tal i com es pot observar a la figura, entre els gens amb expressió augmentada es trobava enriquit a totes les soques el procés catabòlic dels àcids grassos. No obstant, només les soques multicòpia (2C, 4C, 8C i 15C) van presentar una àmplia varietat de processos regulats on es trobaven implicats els àcids grassos. La principal categoria de gens amb expressió reduïda a totes les soques multicòpia estava relacionada amb els peroxisomes. Així mateix, la regulació negativa de la via del metanol estava sobrerrepresentada a les soques més productores (2C i 4C). És més, només en aquestes soques es va detectar una expressió reduïda dels gens implicats en el cicle dels àcids tricarboxílics (TCA, de l'anglès *Tricarboxylic Acid Cycle*) i el metabolisme del piruvat. Finalment, es va trobar que el procés metabòlic de les nicotinamides estava regulat a totes les soques multicòpia, especialment en el cas de la soca 15C.

Els gens regulats es van sotmetre a un procés d'agrupament (*clustering*), utilitzant l'algoritme k-means, per tal de classificar-los en 12 grups diferents en funció dels perfils d'expressió (Figura 5). 7 clústers van mostrar una evident correlació negativa a mesura que augmentava el nombre de còpies de *ROL* a les soques (clústers 1, 2, 3, 5, 7, 8 i 12). Entre aquests grups, els clústers 1, 2 i 3 comprenien els gens més regulats (prèviament representats a la Figura 2A, requadre inferior). El clúster 3 estava significativament enriquit en termes relacionats amb els peroxisomes, el metabolisme de les piridines i de la tiamina, el transport del ferro i la glicòlisi. Els clústers 8 i 12, també amb una tendència marcadament negativa, contenia, entre d'altres, gens relacionats amb els peroxisomes, el metabolisme del metanol i el cicle del glioxilat. Amb relació als clústers amb regulació positiva (és a dir, una correlació positiva entre els nivells transcripcionals i la dosi gènica de *ROL*; clústers 9 i 11), tots els processos biològics enriquits al clúster 9 estaven relacionats amb el metabolisme dels àcids grassos. Al clúster 11 els termes enriquits implicaven la glicòlisi, el metabolisme dels àcids grassos i el transport d'anions, entre d'altres.

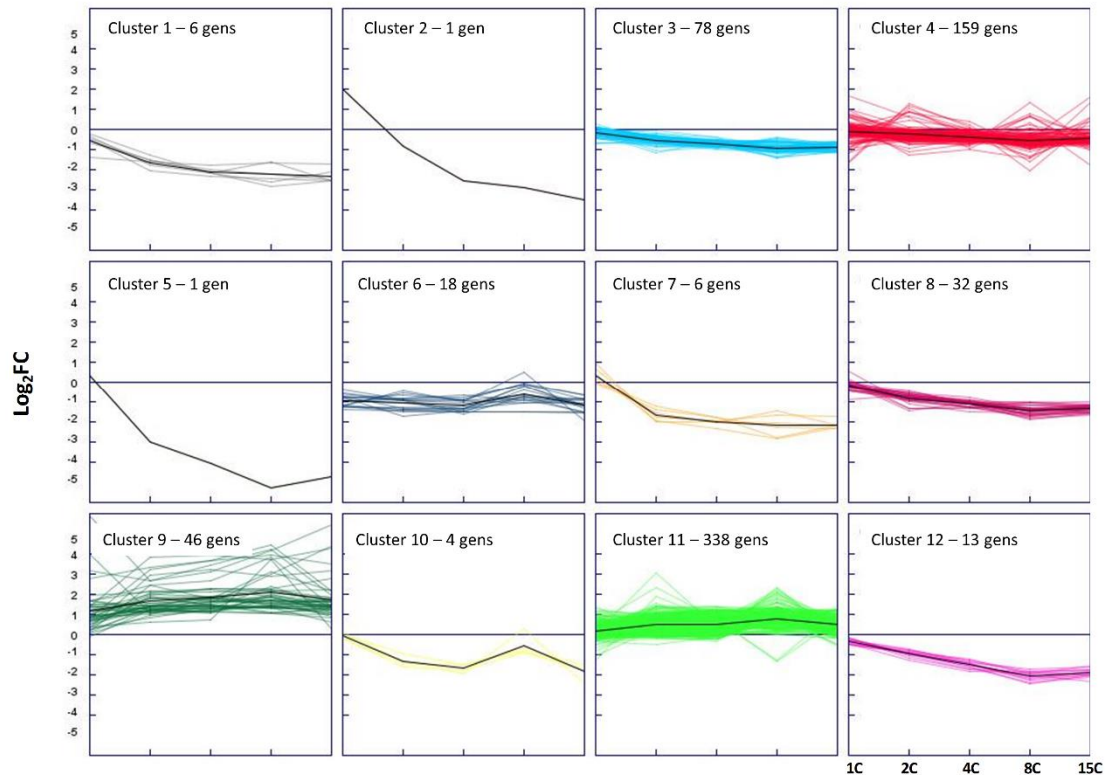


Figura 5. Clustering dels gens regulats a les soques productores. Els perfils d'expressió es van agrupar en 12 categories utilitzant l'algoritme *k-means* per tal de dividir la població segons les tendències principals d'expressió. L'expressió mitjana de cada gràfic es representa amb una línia negra.

3.3.5 Els gens d'utilització del metanol (Mut) presenten una expressió reduïda a mesura que incrementa el nombre de còpies de *ROL*, suggerint un impacte per la disponibilitat limitada de *Mxr1*

P. pastoris, com a llevat metilotròfic, es capaç de créixer amb metanol com a única font de carboni. En aquest estudi, on les cèl·lules creixien en cultius operats en continu amb una barreja de glicerol:metanol com alimentació i on s'assegurava la coassimilació de ambdues fonts de carboni, es va observar una clara regulació global en la ruta del metanol (Figura 6). En primer lloc, l'oxidació de metanol a formaldehid, principalment dirigit per l'enzim alcohol oxidasa (codificat pel gen *AOX1*), va presentar una expressió més baixa de 8 fins a 30 vegades, per la soca 2C i les soques d'alt rang de còpies, respectivament, comparant amb la soca control no productora. Només en el cas de la soca 1C, *AOX1* semblava lleugerament regulat de manera positiva. *AOX2* també va presentar una disminució de 5 vegades en la seva expressió a totes les soques multicòpia. Segon, tots els gens implicats en la via deassimilatòria (oxidació directa del

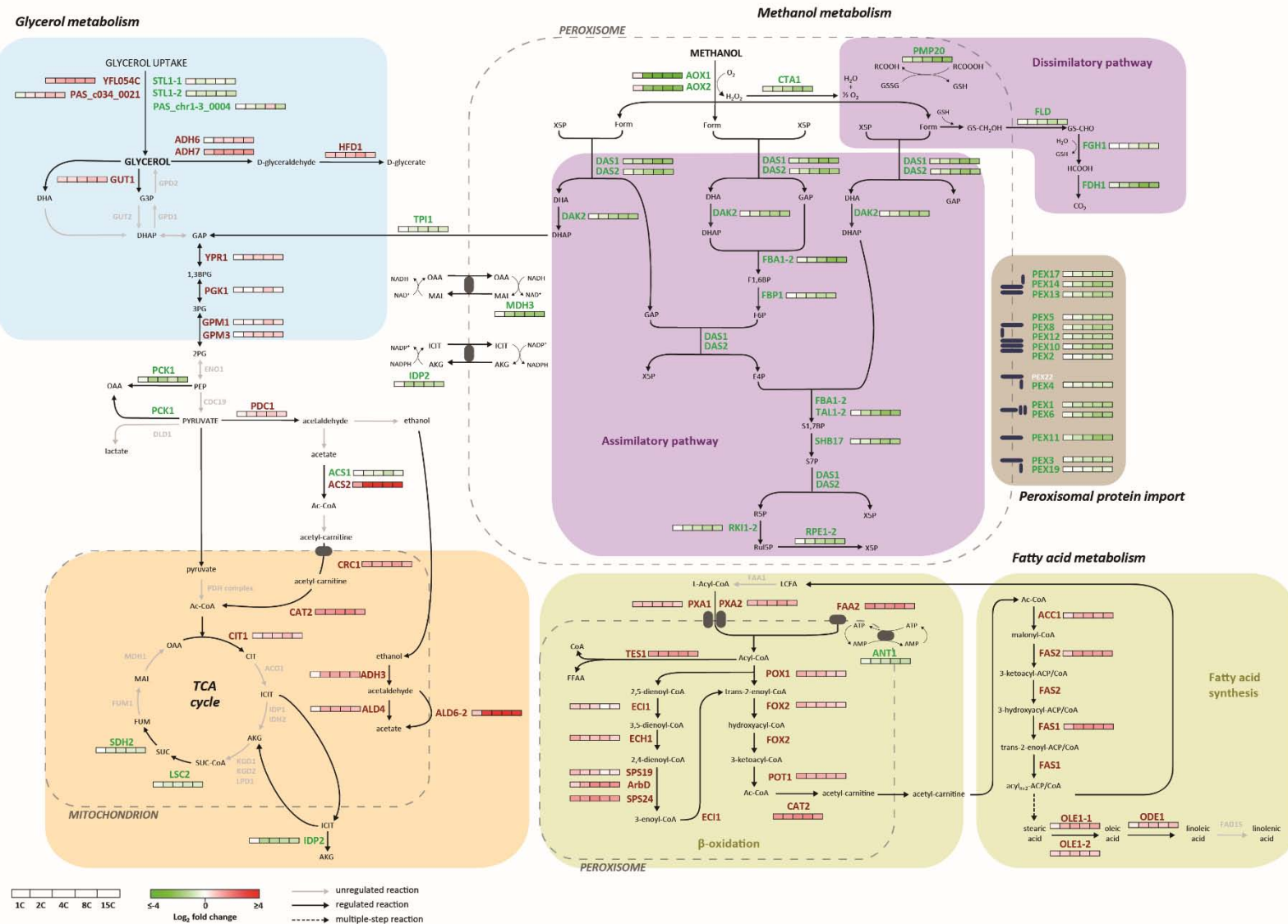


Figura 6. Principals rutes metabòliques afectades per la dosi gènica de *ROL* en cultius de *P. pastoris* operats en continu amb una barreja de glicerol:metanol com a font de carboni. Es troben representades la ruta de la captació de glicerol (blau), el cicle dels àcids tricarbòxics (TCA) (taronja), el metabolisme dels àcids grassos (síntesi i β-oxidació) (marró), l'assimilació del metanol (lila) i l'importació de proteïnes peroxisomals (gris). La regulació gènica de cada pas de la ruta es troba indicat per un gradient de colors. Els gens no regulats es mostren en color gris.

metanol a CO_2) van presentar una regulació transcripcional negativa (gens *FLD*, *FGH1* i *FDH1*). Finalment, tots els gens de la via assimilatòria del metanol, on es generen constituents cel·lulars des del formaldehid, també presentaven una expressió disminuïda (*DAS1*, *DAS2*, *TPI1*, *DAK2*, *FBA1-2* i *FBP1*). De manera interessant, mentre que els gens codificants per la branca no oxidativa de les pentoses fosfat (*TAL1-1*, *RKI1-1* i *RPE1-1*) no estaven regulats, es va observar una regulació negativa en les isoformes peroxisomals d'aquests gens (*TAL1-2*, *RKI1-2*, *RPE1-2* i *SHB17*) en les soques multicòpia, d'igual manera que els gens implicats en la ruta del metanol. Recentment, Rußmayer et al. (2015) van reportar la presència d'un senyal peroxisomal (PTS1, de l'anglès *Peroxisomal Targeting Signal*) en aquests enzims, definint així el reordenament de les reaccions de manera que tenien lloc en un cicle xilulosa-monofosfat al peroxisoma. En síntesi, la xilulosa-5-fosfat és regenerada utilitzant un conjunt d'enzims especialitzats localitzats al peroxisoma, els quals es sintetitzen només en presència de metanol. A més, la regulació negativa de la ruta assimilatòria del metanol (depenent del nombre de còpies de *ROL*) explicaria la disminució en els nivells de biomassa i la reducció del consum de metanol a les soques multicòpia. En efecte, els nivells transcripcionals d'*AOX1* es correlacionaven amb el perfil del consum de metanol observat en aquestes soques (Figura 7).

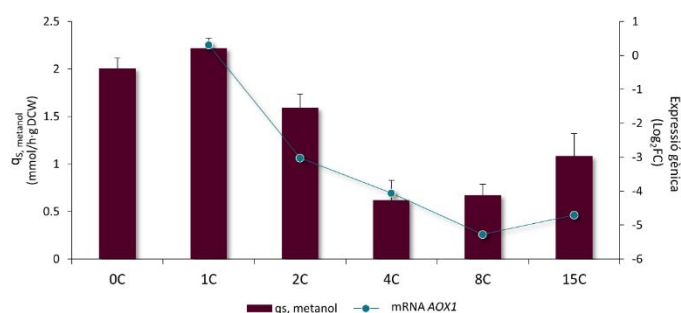


Figura 7. La taxa de consum de metanol està relacionada amb l'expressió de l'alcohol oxidasa. Representat en barres, la taxa de consum de metanol a les soques multicòpia crescudes en cultius operats en continu. El gràfic de línia representa els nivells transcripcionals d'*AOX1* mesurats als microarrays, en comparació a la soca referència 0C. Les barres d'error representen la desviació estàndard entre tres experiments independents.

A mesura que el nombre de còpies incrementava, la capacitat d'assimilació del metanol disminuïa. Això va resultar en nivells significatius de metanol acumulat al medi (fins a 7 g/L, Taula II) en els cultius de les soques multicòpia. Coherentment, tots els gens de les riboflavines implicats en la síntesi del coenzim de la flavina adenina dinucleòtid (FAD, cofactor essencial per

l'assemblatge de l'alcohol oxidasa, i possiblement regulat per *AOX1*, segons Abbas et al., 2011) estaven regulats negativament (*RIB1*, *RIB3*, *RIB4*, *RIB5*, *FAD1*) a les soques multicòpia. Així mateix, tal i com indica Rußmayer et al. (2015), la tiamina pirofosfat (TPP) és el cofactor de les transcetolases *Das1* i *Das2*. De manera consistent, diversos gens implicats en la síntesi de la tiamina i de la TPP presentaven una expressió disminuïda depenent de la dosi gènica de *ROL* (*THI13*, *THI14*, *THI20*, *THI21*, *THI73* i *THI80*).

El metabolisme del metanol a *P. pastoris* es regula principalment per *Mxr1*, un factor de transcripció el qual s'expressa de manera constitutiva a nivells baixos i que és crucial per la inducció de P_{AOX1} i P_{DAS} (Kranthi et al., 2010; Lin-Cereghino et al., 2008; Vogl i Glieder, 2013). Sembla plausible que la inserció múltiple de vectors d'expressió (resultant en un increment del nombre de seqüències de P_{AOX1}) derivi en una limitació de molècules lliures de *Mxr1* a les soques multicòpia, afectant així negativament els nivells transcripcionals dels gens induïts pel metanol, inclòs el gen *ROL*. Tal i com es proposa a la Figura 8A, els nivells baixos d'expressió de *Mxr1* serien suficients per tal d'assolir l'activació total de P_{AOX1} i P_{DAS} en la soca monocòpia i conseqüentment la correcta assimilació del metanol com a font de carboni. En canvi en el cas de la soca 4C (Figura 8B), la distribució de *Mxr1* entre els múltiples llocs d'unió no seria suficient per tal d'aconseguir una expressió regular de *DAS* i *AOX1*, resultant en una reducció en l'assimilació del metanol. Tot i això, aquesta hipòtesi no explicaria el fet que mentre que els nivells de mRNA de *ROL* són constants per les soques multicòpia, s'observi una baixada en els nivells extracel·lulars d'activitat lipàsica (Figura 1A). Així mateix, la baixa activitat lipàsica detectada intracel·lularment podria indicar una aturada traduccional i no una activació de l'UPR com la raó per la qual s'observa una disminució en la producció de *Rol* a les soques 4C, 8C i 15C, tal i com va reportar recentment Edward-Jones et al. (2015) per soques multicòpia expressant tripsinogen humà. Breument, aquests autors van observar que en cultius de *P. pastoris* operats en fed-batch la soca multicòpia no traduïa la totalitat dels transcrits del tripsinogen, i en conseqüència, es detectava una reducció en el tràfic a través del RE en comparació amb la soca monocòpia. Aquest fet es pot relacionar amb les nostres observacions que no només els nivells extracel·lulars, sinó també els nivells intracel·lulars (fracció insoluble) de *Rol* eren més baixos a les soques 4C i 8C. De manera addicional, a Edward-Jones et al. (2015) aquest fenomen es va veure acompanyat d'una adaptació lenta al metabolisme del metanol, consistent amb la nostra hipòtesi d'una limitació de *Mxr1* lliures a les soques multicòpia. Com en el nostre cas, no es van trobar evidències d'activació de l'UPR, apuntant cap a una elevada demanda metabòlica per la síntesi de proteïna recombinant com la principal causa d'aquesta aturada traduccional.

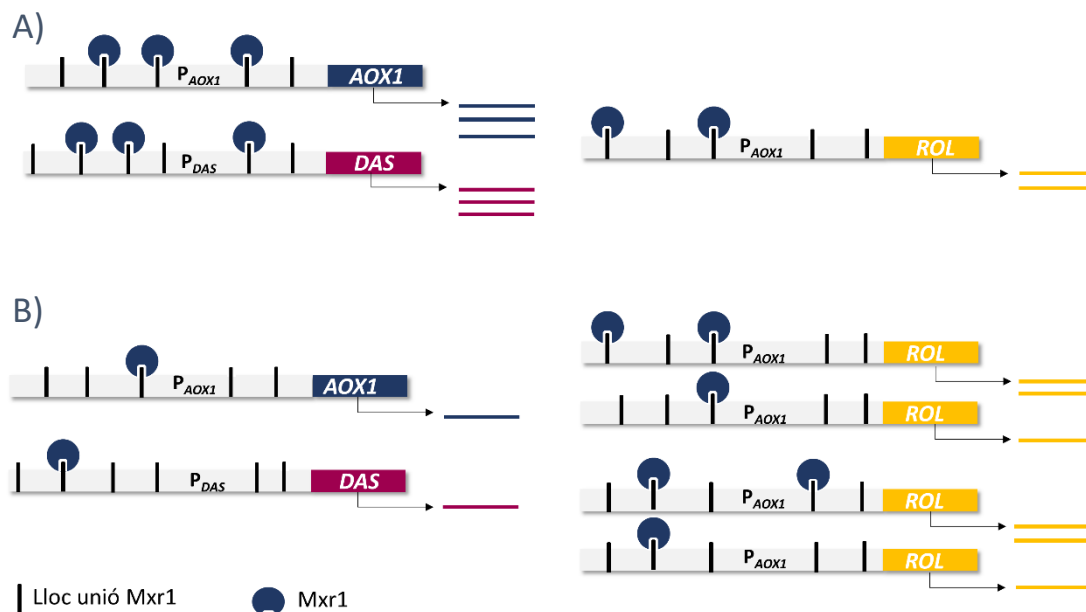


Figura 8. Proposta del mecanisme de regulació dels gens Mut mitjançant el factor de transcripció Mxr1 en presència de metanol. A) Regulació dels gens Mut a la soca 1C. B) Regulació dels gens Mut a la soca 4C. Degut als nombrosos llocs d'unió de Mxr1 als promotors dels gens implicats en el metabolisme del metanol (com l'alcohol oxidasa 1 i la dihidroxiacetona sintasa) i els nivells relativament baixos de Mxr1, és possible que es doni una limitació en la regulació dels gens Mut a les soques multicòpia. Mentre que a la soca 1C el nombre de molècules de Mxr1 és suficient per una expressió correcta dels gens Mut, l'increment dels llocs d'unió a la soca 4C resulta en una probabilitat menor d'unió de Mxr1 al promotor d'aquests gens, afectant així de manera negativa a la capacitat assimilatòria del metanol de la soca.

3.3.6 La biogènesi dels peroxisomes es regula negativament a les soques amb alta dosi gènica de *ROL*

Els peroxisomes són orgànuls crucials a *P. pastoris*, duent a terme tres funcions principals: I) acollir els enzims necessaris per la metabolització del metanol a través de la ruta assimilatòria, II) proveir un compartiment on confinar i eliminar el peròxid d'hidrogen produït per l'alcohol oxidasa, i III) albergar la β -oxidació dels àcids grassos (Lin-Cereghino et al., 2006; Yurimoto et al., 2011). En aquest estudi, es va observar una completa regulació negativa del metabolisme del metanol. En consonància, també es va observar una regulació global dels gens de les peroxines, codificants per proteïnes implicades en la biogènesi dels peroxisomes. D'una banda, els gens *PEX1*, *PEX3*, *PEX6*, *PEX11C*, *PEX19*, *PEX25* i *PEX30*, relacionats amb la biogènesi de la membrana peroxisomal (provinent del RE o per fissió; Smith i Aitchison, 2013) presentaven una regulació disminuïda (en comparació a la soca 0C), especialment les soques d'alt rang de còpies 8C i 15C.

D'altra banda, *PEX2*, *PEX4*, *PEX5*, *PEX8*, *PEX10*, *PEX12*, *PEX13*, *PEX14* i *PEX17*, els quals estan relacionats funcionalment amb l'import de proteïnes a la matriu peroxisomal, també estaven regulats negativament (Figura 6). Les proteïnes de la matriu peroxisomal contenen senyals específics (PTS1 o PTS2) que es reconeixen al citoplasma de manera post-traducciona per els receptors Pex5 i Pex7, respectivament (Hettema et al., 2014). De manera interessant, *PEX5* presentava una regulació negativa, d'acord amb la regulació negativa dels gens clau de la ruta del metanol (*AOX1*, *CTA1*, *DAS1* i *DAS2*), tots ells codificants per proteïnes PTS1 (Kiel, 2006). Contràriament, no es va detectar cap regulació en el gen *PEX7*, en concordança amb les observacions prèvies d'Elgersma et al. (1998) i Rußmayer et al. (2015), reforçant la idea que el metabolisme del metanol no és dependent de la via d'importació per PTS2. Sorprenentment, *ATG8* i *ATG18*, implicats en la pexofàgia (és a dir, l'autofàgia dels peroxisomes) (Tamura et al., 2013) van presentar una regulació negativa en comparació a la soca no productora. Segons Vanz et al. (2012), als cultius en metanol hi ha un reciclatge constitutiu dels peroxisomes mitjançant la pexofàgia per tal d'eliminar els peroxisomes malmesos per les espècies reactives de l'oxigen (ROS, de l'anglès *Reactive Oxygen Species*). D'acord amb aquestes observacions, els nivells de ROS detectats en aquesta sèrie de soques crescudes en cultius operats en continu en condicions anàlogues van ser molt baixos (Cámara et al., 2016). De manera similar, Hesketh et al. (2013) van descriure l'acimatació de soques productores creixent en cultius en continu amb medi sorbitol-metanol, on només es va detectar una resposta oxidativa transitòria després de començar la fase d'inducció (en el canvi de condicions no induïbles (medi sorbitol) a induïbles (sorbitol+metanol) en cultius operats en continu), però no un cop es va assolir l'estat estacionari. En conjunt, els nivells baixos de ROS, la manca de regulació dels marcadors d'estrès oxidatiu i la regulació negativa del gen *CTA1* (codificant per l'enzim catalasa, i responsable de la degradació del peròxid d'hidrogen als peroxisomes) suggereix que l'estrès oxidatiu té un rol menor en l'adaptació fisiològica a altes dosis gèniques de *ROL* en condicions d'estat estacionari. No obstant, no es pot descartar que els nivells de ROS puguin augmentar transitòriament a l'inici de la inducció per metanol, tal i com es despenia de l'anàlisi previ d'aquestes soques en cultius en Erlenmeyer (Cámara et al., 2016). Així mateix, el gen *HEM1*, que codifica pel primer pas de la síntesi del grup hemo (cofactor de la catalasa), i *PET18*, codificant per una proteïna de tipus hemo (que catalitza la degradació dels grups hemo) presentaven una regulació disminuïda dependent del nombre de còpies de *ROL*. Segons Rußmayer et al. (2015), el metabolisme del ferro es troba estretament relacionat amb la detoxificació duta a terme per la catalasa i el metabolisme del metanol. En el nostre cas, diversos gens codificants per permeases del ferro, entre d'altres gens relacionats amb el metabolisme del ferro, estaven regulats positivament en

les soques productores (*FET3*, *FET4-2*, *FRE2*, *FRE3*, *FRE6*, *FTR1*, *SIT1-2*) en contrast amb la regulació negativa global dels gens Mut.

3.3.7 El metabolisme dels àcids grassos es troba significativament alterat, mostrant una regulació positiva de les rutes biosintètiques i la β -oxidació a les soques productores de Rol

Als llevats, la β -oxidació (la principal ruta de degradació dels àcids grassos) es dona lloc de manera exclusiva als peroxisomes. L'acetil-CoA format pot ser redirigit cap el cicle del glicoxilat o transportat a la mitocondria a través de la carnitina, sent així una font d'energia i una font anapleròtica de blocs estructurals. Breument, els àcids grassos de cadena mitjana (MCFA, de l'anglès *Medium Chain Fatty Acids*) són transportats dins els peroxisomes per difusió, i són activats per l'acil-CoA sintasa depenent *Faa2*. En el cas dels àcids grassos de cadena llarga (LCFA, de l'anglès *Long Chain Fatty Acids*), es requereix l'heterodímer format per *Pxa1-Pxa2* per tal d'introduir-los dins el peroxisoma. Tal i com va apuntar Van Roermund et al. (2012), *Pxa1-Pxa2* també accepta LCFA activats, els quals són posteriorment reesterificats per l'acil-CoA sintasa *Faa2* o *Fat1*. El procés de la β -oxidació s'inicia amb l'oxidació dels substrats de l'acil-CoA per *Pox1*. Les següents reaccions són catalitzades per l'enzim multifuncional *Fox2*, generant hidroxiacil-CoA i 3-cetoacil-CoA. La reacció final, duta a terme per la tiolasa 3-cetoacil-CoA *Pot1*, resulta en una molècula d'acetil-CoA i una cadena d'acil-CoA 2 carbonis més curta, que pot tornar a entrar al cicle de la β -oxidació. A totes les soques productores es va observar una regulació global positiva a tots els passos de la β -oxidació (gens *FAA2*, *PXA1*, *PXA2*, *FAT1*, *POX1*, *FOX2* i *POT1*), tal i com es mostra a la Figura 6. Fins i tot *TES1*, codificant per la tioesterasa acil-CoA implicada en el manteniment del conjunt de CoA lliures dins el peroxisoma (Maeda et al., 2006), va presentar una regulació significativament positiva a totes les soques. Els gens codificants per la β -oxidació dels àcids grassos insaturats amb dobles enllaços cis i trans a posicions senars, o enllaços dobles cis a posicions parells (àcid oleic) també van presentar una regulació augmentada, però en menor magnitud. A diferència de la ruta del metanol, aquesta regulació global dels gens relacionats amb el catabolisme dels àcids grassos no semblava depenent del nombre de còpies de *ROL*; és a dir, afectava de manera similar a totes les soques productores.

L'acetil-CoA generat a la β -oxidació té dues maneres de ser exportat dels peroxisomes: Primer, a través de la conjugació amb la carnitina, mediat per la carnitina acetiltransferasa *Cat2*, i subsegüentment transportat a la mitocondria, on s'uneix a la proteïna transportadora *Crc1*. El

segon possible destí de l'acetil-CoA és el cicle del glioxilat, on es genera succinat com a producte final (Van Roermund et al., 2003). La dèbil regulació dels gens codificants pels enzims implicats en el cicle del glioxilat, juntament amb la regulació positiva de *CAT2* i *CRC1* a totes les soques productores podria indicar que l'acetil-CoA és principalment redirigit cap a la mitocondria, a expenses del cicle del glioxilat. Donat que els peroxisomes són orgànuls impermeables als dinucleòtids (per exemple, FAD, NAD⁺, NADPH) i l'acetil-CoA (Hiltunen et al., 2003), aquests depenen de diversos transportadors per tal de mantenir el conjunt de cofactors disponibles. En el cas dels enzims que contenen FAD (acil-CoA oxidasa, alcohol oxidasa), aquests són ensamblats amb el cofactor al citoplasma i posteriorment són reoxidats al peroxisoma, generant H₂O₂ (Rottensteiner i Theodoulou, 2006). Contràriament, el NAD⁺ es regenera a través de la llançadora malat/aspartat, via la malat deshidrogenasa peroxisomal Mdh3, mentre que la reducció de NADP⁺ pot ser realitzada per la llançadora 2-cetoglutarat/isocitrat, tal i com s'ha proposat per *S. cerevisiae* (Van Roermund et al., 2003). En ambdós casos, els gens corresponents van presentar una expressió disminuïda comparat amb la soca control no productora (Figura 6). De manera similar, el transportador del nucleòtid adenina codificat per *ANT1* (i acoblat a l'activació dels MCFA pel gen *FAA2*) va presentar una regulació negativa, juntament amb la catalasa *CTA1*.

Curiosament, la biosíntesi dels àcids grassos també es va regular de manera positiva amb l'increment del nombre de còpies de *ROL*. Tant el gen de l'acetil-CoA carboxilasa (*ACC1*) com el del complex de la sintasa dels àcids grassos (codificat per *FAS1* i *FAS2*) van presentar una expressió incrementada, particularment en aquelles soques amb més d'una còpia de *ROL*. A més, tres dels quatre gens codificants per les desaturases dels àcids grassos descrites a *P. pastoris* (Yu et al., 2012) van mostrar una regulació positiva (*OLE1-1*, *OLE1-2*, *ODE1*). La composició total d'àcids grassos de les soques es va analitzar posteriorment per tal d'estudiar l'efecte de la regulació de les desaturases en el perfil dels àcids grassos (Taula III). Els resultats van confirmar que els àcids grassos més abundants a *P. pastoris* eren l'àcid oleic, l'àcid linoleic, l'àcid palmític, l'àcid linolènic i l'àcid esteàric, tal i com s'havia reportat anteriorment (Grillitsch et al., 2014; Yu et al., 2012). De manera interessant, sembla que a les soques productores l'àcid gras més abundant era el linoleic, a diferència de la soca referència OC, on el principal àcid gras era l'oleic. L'increment a les soques productores de l'expressió del gen *ODE1*, codificant per la desaturasa Fad12 (la qual catalitza la síntesi d'àcid linoleic a partir de oleic com a substrat), podria explicar aquest canvi en el perfil dels àcids grassos. Així mateix, el contingut total d'àcids grassos de cada soca es va correlacionar de manera negativa amb el nombre de còpies de *ROL*, especialment en les soques multicòpia amb una elevada dosi gènica de *ROL*.

Taula III Composició d'àcids grassos totals del les soques productores de Rol i la soca de referència. El contingut d'àcids grassos es va quantificar per cromatografia de gasos acoblada a una espectrometria de masses. Només es representen els 5 àcids grassos majoritaris. Els resultats es mostren com la mitjana \pm sd, basat en tres experiments independents per cada soca. * indica $P < 0.05$ comparat amb la soca 0C.

	% oleic (C18:1)	% linoleic (C18:2n-6)	% palmític (C16:0)	% α -linolènic (C18:3n-3)	% esteàric (C18:0)	TOTAL (μ g àc. grassos/mg DCW)
0C	35.3 \pm 3.7	26.9 \pm 2.6	13.6 \pm 1.3	10.8 \pm 0.1	2.1 \pm 0.1	45.9 \pm 4.2
1C	28.2 \pm 1.9 *	31.6 \pm 2.1 *	11.9 \pm 0.9 *	12.9 \pm 0.1 *	1.6 \pm 0.4	40.1 \pm 2.9
2C	28.4 \pm 0.7 *	31.4 \pm 1.8 *	11.7 \pm 0.5 *	10.4 \pm 0.1	2.5 \pm 0.5	40.2 \pm 1.4
4C	28.2 \pm 3.1 *	31.6 \pm 2.1 *	12.0 \pm 1.2 *	11.1 \pm 0.1	2.1 \pm 0.3	35.6 \pm 2.7 *
8C	26.2 \pm 1.5 *	32.0 \pm 2.3 *	13.6 \pm 1.1	13.5 \pm 0.2 *	2.8 \pm 0.3 *	35.9 \pm 2.5 *
15C	26.5 \pm 0.7 *	30.8 \pm 0.2 *	14.0 \pm 0.7	12.0 \pm 0.5	2.7 \pm 0.2 *	32.5 \pm 5.2 *

L'increment dels nivells transcripcionals dels gens codificants per la biosíntesi dels àcids grassos així com la β -oxidació podria reflectir un increment en el recanvi dels àcids grassos, tal i com havia reportat prèviament Marchesini et al. (2003) a *S. cerevisiae*. Aquesta evidència es veuria reforçada pel fet que el gen codificant pel contribuent principal a la síntesi dels triacilglicèrids, la acetiltransferasa Dga1, presenta una regulació negativa i cap altre gen implicat en el metabolisme dels lípids (esfingolípids, fosfolípids) es va detectar regulat. Consegüentment, sembla que a les soques productores s'estigués propiciant un cicle fútil on estaria implicat l'acetil-CoA. Tal i com s'ha mencionat anteriorment, es va observar una elevada regulació positiva del gen *ACS2* a les soques multicòpia. Aquesta acetil-CoA sintasa és la principal responsable de la síntesi de l'acetil-CoA d'origen nucleic en els llevats, i tal i com va reportar Eisenberg et al. (2004), una hiperactivació d'aquest enzim pot provocar l'acetilació d'histones i la subsegüent repressió dels gens de l'autofàgia, en consonància amb la regulació negativa dels gens de la pexofàgia *ATG8* i *ATG18* mencionats anteriorment. D'una altra banda, la regulació de l'acetil-CoA també es dona per la síntesi *de novo* dels àcids grassos (Galdieri i Vacuna, 2012), que competeixen amb l'acetilació de proteïnes pel mateix grup de molècules d'acetil-CoA. En conjunt, aquests resultats apunten a un complex mecanisme de regulació a nivell transcripcional i traduccional. De manera alternativa, la forta regulació del metabolisme dels àcids grassos podria reflectir una limitació en la disponibilitat d'acetil-CoA com a resultat d'un increment en la demanda anabòlica de blocs estructurals (per exemple, per la síntesi de membrana lipídica, requerida per tal incrementar la capacitat secretòria), alhora que per alimentar el TCA (és a dir,

per producció d'energia), tal i com ha suggerit recentment Klein et al. (2014). No obstant, tal i com s'ha mencionat anteriorment, no es va detectar cap regulació de les vies biosintètiques dels lípids de membrana a les soques productores de Rol.

3.3.8 L'impacte de la dosi gènica de *ROL* al metabolisme dels aminoàcids, el metabolisme energètic i del glicerol, la síntesi, plegament, secreció i degradació de proteïnes és limitat o no significatiu

Tot i que s'esperava un impacte en el metabolisme dels aminoàcids degut als requeriments per la síntesi de proteïna recombinant (Jordà et al., 2015), no es va detectar cap regulació global en les diferents rutes biosintètiques dels aminoàcids ni les seves permeases, excepte per la lisina, amb una lleugera regulació negativa a les soques multicòpia en algunes de les etapes de síntesi d'aquest aminoàcid (*LYS1*, *LYS2*, *LYS4*, *LYS20*, *MKS1*). Curiosament, *LYP1* i *PAS_chr2-1_0649*, codificants per permeases de la lisina, van presentar un augment de l'expressió en comparació amb la soca de referència OC. De manera consistent, no es va detectar una regulació significativa dels gens associats als ribosomes. Tot i que es van observar taxes de consum de glicerol més elevades a mesura que augmentava el nombre de còpies de *ROL*, només dues possibles permeases del glicerol (gens *YFL054C* i *PAS_c034_0021*) van mostrar una regulació positiva a les soques multicòpia. De manera interessant, tot i que estudis previs havien revelat un increment en l'activitat metabòlica del TCA per tal d'aconseguir més cofactors energètics (NADH, ATP) degut a la producció de proteïna recombinant (Jordà et al., 2012; Jordà et al., 2014), només *CAT2* (essencial pel transport mediat per la carnitina acetiltransferasa de l'acetil-CoA citosòlic cap a la mitocòndria) i la citrat sintasa (codificat per *CIT1*) van presentar una lleugera regulació positiva a les soques multicòpia.

Tal i com s'ha comentat anteriorment (Taula III) aproximadament un 30% del total de Rol es va trobar a la fracció cel·lular insoluble de les soques productores. Amb l'objectiu d'estudiar la relació entre la retenció de lipasa i l'UPR, es van estudiar els gens implicats en la secreció i plegament de proteïnes. Tal i com es mostra a la Figura 9, un total de 17 gens es van trobar regulats en almenys una soca en comparació amb la soca no productora OC.

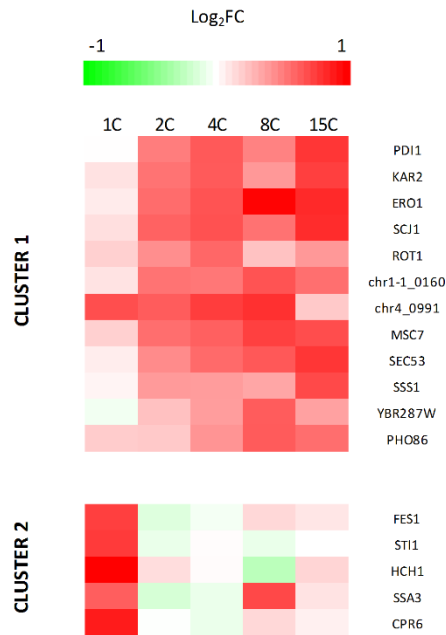


Figura 9. Regulació dels gens implicats en plegament i secreció. Es va realitzar un anàlisi de *clustering* utilitzant l'algoritme *k-means* (k=2). El clúster 1 comprenia principalment proteïnes residents del RE, mentre que el clúster 2 contenia xaperones citosòliques.

La regulació d'aquests gens no va ser intensa, amb una diferència màxima d'expressió del doble. Després de realitzar un anàlisi de clústers es van obtenir dos grups principals. D'una banda, un dels grups exhibia una regulació positiva a les soques multicòpia, mentre que no es detectava regulació a la soca monocòpia (amb l'excepció del gen PAS_chr4_0991). L'anàlisi funcional d'aquest grup va revelar que aquest grup de gens eren residents del RE, principalment xaperones i isomerases. D'una altra banda, el segon grup comprenia gens regulats únicament a la soca 1C (excepte SSA3). En aquest cas, tots els gens codificaven proteïnes del tipus xoc tèrmic citosòliques, implicades en el plegament proteic. De manera consistent, no es va observar cap regulació del factor de transcripció *HAC1* (principal regulador de l'UPR), confirmant la hipòtesi que aquesta resposta cel·lular no s'havia activat en les soques productores en condicions d'estat estacionari. De manera addicional, no es va trobar una regulació global per les proteases vacuolars o els gens relacionats amb la via ERAD, indicant així que l'estrès al RE no es trobava activat a les soques multicòpia.

3.4 CONCLUSIONS

L'increment del nombre de còpies d'un gen heteròleg és una estratègia àmpliament establerta per tal de millorar la producció de proteïnes recombinants a *P. pastoris*. No obstant, aquesta estratègia sovint presenta limitacions quan la dosi gènica s'incrementa més enllà de l'òptim, provocant un important impacte fisiològic en cèl·lules productores.

Aquest estudi transcriptòmic ajuda a entendre les adaptacions fisiològiques per part dels mecanismes cel·lulars en resposta a l'increment de la dosi gènica en sistemes on s'utilitza el promotor d'*AOX1*, induïble per metanol. Destacablement, l'expressió del sistema sembla estar fortament limitada a nivell transcripcional, per sobre d'altres limitacions com podrien ser la disponibilitat de recursos energètics i constituents estructurals, o el plegament i la secreció de la proteïna recombinant. Concretament, l'atenuació transcripcional a les soques multicòpia dels gens implicats en la utilització del metanol, incloent el P_{AOX1} utilitzat per dirigir l'expressió del gen heteròleg, va resultar en una forta regulació negativa del metabolisme del metanol, la biogènesi de peroxisomes i el metabolisme dels àcids grassos. Aquest fet va impactar posteriorment als paràmetres de creixement i les productivitats de les soques multicòpia creixent en una barreja de glicerol:metanol (elevat consum del glicerol enfront una disminució de les taxes de consum del metanol).

Aquest treball proporciona un conjunt de dades que permet entendre les limitacions dels sistemes basats en P_{AOX1} quan es desenvolupen soques amb un increment de còpies del gen heteròleg, i tindrà un gran impacte en el disseny de noves estratègies d'enginyeria cel·lular per tal de generar soques més productores.

3.5 REFERÈNCIES

- Abbas CA, Sibirny AA. 2011. Genetic Control of Biosynthesis and Transport of Riboflavin and Flavin Nucleotides and Construction of Robust Biotechnological Producers. *Microbiol. Mol. Biol. Rev.* 75:321–360.
- Ahmad M, Hirz M, Pichler H, Schwab H. 2014. Protein expression in *Pichia pastoris*: Recent achievements and perspectives for heterologous protein production. *Appl. Microbiol. Biotechnol.* 98:5301–5317.
- Bai J, Swartz DJ, Protasevich II, Brouillette CG, Harrell PM, Hildebrandt E, Gasser B, Mattanovich D, Ward A, Chang G, Urbatsch IL. 2011. A gene optimization strategy that enhances production of fully functional P-Glycoprotein in *Pichia pastoris*. *PLoS One* 6:e22577
- Baumann K, Dato L, Graf AB, Frascotti G, Dragosits M, Porro D, Mattanovich D, Ferrer P, Branduardi P. 2011. The impact of oxygen on the transcriptome of recombinant *S. cerevisiae* and *P. pastoris* - a comparative analysis. *BMC Genomics* 12:218.
- Ben Azoun S, Belhaj AE, Göngrich R, Gasser B, Kallel H. 2016. Molecular optimization of rabies virus glycoprotein expression in *Pichia pastoris*. *Microb. Biotechnol.* 9:355–368
- Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pagès F, Trajanoski Z, Galon J. 2009. ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 25:1091–1093.
- Cámara E, Albiol J, Ferrer P. 2016. Droplet digital PCR-aided screening and characterization of *Pichia pastoris* multiple gene copy strains. *Biotechnol. Bioeng.* 113:1542–1551.
- Cereghino JL, Cregg JM. 2000. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol. Rev.* 24:45–66.
- Cos O, Ramón R, Montesinos JL, Valero F. 2006. Operational strategies, monitoring and control of heterologous protein production in the methylotrophic yeast *Pichia pastoris* under different promoters: a review. *Microb. Cell Fact.* 5:17.
- Cos O, Serrano A, Montesinos JL, Ferrer P, Cregg JM, Valero F. 2005. Combined effect of the methanol utilization (Mut) phenotype and gene dosage on recombinant protein production in *Pichia pastoris* fed-batch cultures. *J. Biotechnol.* 116:321–35.

Couderc R, Baratti J. 1980. Oxidation of Methanol by the Yeast, *Pichia pastoris*. Purification and Properties of the Alcohol Oxidase. *Agric. Biol. Chem.* 44:2279–2289.

Edwards-Jones B, Aw R, Barton GR, Tredwell GD, Bundy JG, Leak DJ. 2015. Translational Arrest Due to Cytoplasmic Redox Stress Delays Adaptation to Growth on Methanol and Heterologous Protein Expression in a Typical Fed-Batch Culture of *Pichia pastoris*. *PLoS One* 10:e0119637.

Eisenberg T, Schroeder S, Andryushkova A, Pendl T, Küttner V, Bhukel A, Mariño G, Pietrocola F, Harger A, Zimmermann A, Moustafa T, Sprenger A, Jany E, Büttner S, Carmona-Gutierrez D, Ruckstuhl C, Ring J, Reichelt W, Schimmel K, Leeb T, Moser C, Schatz S, Kamolz LP, Magnes C, Sinner F, Sedej S, Fröhlich KU, Juhasz G, Pieber TR, Dengjel J, Sigrist SJ, Kroemer G, Madeo F. 2014. Nucleocytoplasmic depletion of the energy metabolite acetyl-coenzyme A stimulates autophagy and prolongs lifespan. *Cell Metab.* 19:431–444.

Elgersma Y, Elgersma-Hooisma M, Wenzel T, McCaffery JM, Farquhar MG, Subramani S. 1998. A mobile PTS2 receptor for peroxisomal protein import in *Pichia pastoris*. *J. Cell Biol.* 140:807–20.

Ferrer, P, Albiol, J. 2014. ¹³C-based metabolic flux analysis of recombinant *Pichia pastoris*. *Meth. Mol. Biol.* 1191:291-313

Gao M, Shi Z. 2013. Process control and optimization for heterologous protein production by methylotrophic *Pichia pastoris*. *Chinese J. Chem. Eng.* 21:216–226.

Galdieri L, Vancura A. 2012. Acetyl-CoA carboxylase regulates global histone acetylation. *J. Biol. Chem.* 287:23865–23876.

Garcia-Ortega X, Adelantado N, Ferrer P, Montesinos JL, Valero F. 2016. A step forward to improve recombinant protein production in *Pichia pastoris*: from specific growth rate effect on protein secretion to carbon-starving conditions as advanced strategy. *Process Biochem.* 51:681-691

Guerfal M, Ryckaert S, Jacobs PP, Ameloot P, Van Craenenbroeck K, Derycke R, Callewaert N. 2010. 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.* 9:49.

Graf A, Gasser B, Dragosits M, Sauer M, Leparac GG, Tüchler T, Kreil DP, Mattanovich D. 2008. Novel insights into the unfolded protein response using *Pichia pastoris* specific DNA microarrays. *BMC Genomics* 9:390.

- Grillitsch K, Tarazona P, Klug L, Wriessnegger T, Zellnig G, Leitner E, Feussner I, Daum G. 2014. Isolation and characterization of the plasma membrane from the yeast *Pichia pastoris*. *Biochim. Biophys. Acta* 1838:1889–97.
- Hesketh AR, Castrillo JI, Sawyer T, Archer DB, Oliver SG. 2013. Investigating the physiological response of *Pichia (Komagataella) pastoris* GS115 to the heterologous expression of misfolded proteins using chemostat cultures. *Appl. Microbiol. Biotechnol.* 97:9747–9762.
- Hettema EH, Erdmann R, van der Klei I, Veenhuis M. 2014. Evolving models for peroxisome biogenesis. *Curr. Opin. Cell Biol.* 29:25–30.
- Hiltunen JK, Mursula AM, Rottensteiner H, Wierenga RK, Kastaniotis AJ, Gurvitz A. 2003. The biochemistry of peroxisomal β -oxidation in the yeast *Saccharomyces cerevisiae*. *FEMS Microbiol. Rev.* 27:35–64.
- Hohenblum H, Gasser B, Maurer M, Borth N, Mattanovich D. 2004. Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant *Pichia pastoris*. *Biotechnol. Bioeng.* 85:367–75.
- Huang J, Xia J, Yang Z, Guan F, Cui D, Guan G, Jiang W, Li Y. 2014. Improved production of a recombinant *Rhizomucor miehei* lipase expressed in *Pichia pastoris* and its application for conversion of microalgae oil to biodiesel. *Biotechnol. Biofuels* 7:111.
- Inan M, Aryasomayajula D, Sinha J, Meagher MM. 2006. Enhancement of protein secretion in *Pichia pastoris* by overexpression of protein disulfide isomerase. *Biotechnol. Bioeng.* 93:771–778.
- Jordà J, Jouhten P, Cámara E, Maaheimo H, Albiol J, Ferrer P. 2012. Metabolic flux profiling of recombinant protein secreting *Pichia pastoris* growing on glucose:methanol mixtures. *Microb. Cell Fact.* 11:57.
- Jordà J, Rojas HC, Carnicer M, Wahl A, Ferrer P, Albiol J. 2014. Quantitative Metabolomics and Instationary 13 C-Metabolic Flux Analysis Reveals Impact of Recombinant Protein Production on Trehalose and Energy Metabolism in *Pichia pastoris*. *Metabolites* 4:281–299.
- Kiel JAKW, Veenhuis M, van der Klei IJ. 2006. PEX genes in fungal genomes: Common, rare or redundant. *Traffic* 7:1291–1303.

- Klein T, Lange S, Wilhelm N, Bureik M, Yang TH, Heinzle E, Schneider K. 2014. Overcoming the metabolic burden of protein secretion in *Schizosaccharomyces pombe* - A quantitative approach using ¹³C-based metabolic flux analysis. *Metab. Eng.* 21:34–45.
- Kranthi BV, Kumar HR, Rangarajan PN. 2010. Identification of Mxr1p-binding sites in the promoters of genes encoding dihydroxyacetone synthase and peroxin 8 of the methylotrophic yeast *Pichia pastoris*. *Yeast* 27:705–711.
- Liang S, Wang B, Pan L, Ye Y, He M, Han S, Zheng S, Wang X, Lin Y. 2012. Comprehensive structural annotation of *Pichia pastoris* transcriptome and the response to various carbon sources using deep paired-end RNA sequencing. *BMC Genomics* 13:738.
- Lin-Cereghino GP, Godfrey L, de la Cruz BJ, Johnson S, Khuongsathiene S, Tolstorukov I, Yan M, Lin-Cereghino J, Veenhuis M, Subramani S, Cregg JM. 2006. Mxr1p, a key regulator of the methanol utilization pathway and peroxisomal genes in *Pichia pastoris*. *Mol. Cell. Biol.* 26:883–897.
- Liu H, Qin Y, Huang Y, Chen Y, Cong P, He Z. 2014. Direct Evaluation of the Effect of Gene Dosage on Secretion of Protein from Yeast *Pichia pastoris* by Expressing EGFP. *J. Microbiol. Biotechnol.* 24:144–151.
- Looser V, Brühlmann B, Bumbak F, Stenger C, Costa M, Camattari a., Fotiadis D, Kovar K. 2015. Cultivation strategies to enhance productivity of *Pichia pastoris*: A review. *Biotechnol. Adv.* 33:1177-1193
- MacQueen JB. 1967. Some methods for classification and analysis of multivariate observations. 5th Berkeley Symp. Math. Stat. Probab. 1967 1:281–297.
- Maeda I, Delessert S, Hasegawa S, Seto Y, Zuber S, Poirier Y. 2006. The peroxisomal acyl-CoA thioesterase Pte1p from *Saccharomyces cerevisiae* is required for efficient degradation of short straight chain and branched chain fatty acids. *J. Biol. Chem.* 281:11729–11735.
- Marchesini S, Poirier Y. 2003. Futile cycling of intermediates of fatty acid biosynthesis toward peroxisomal β -oxidation in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 278:32596–32601.
- Minning S, Serrano A, Ferrer P, Solà C, Schmid RD, Valero F. 2001. Optimization of the high-level production of *Rhizopus oryzae* lipase in *Pichia pastoris*. *J. Biotechnol.* 86:59–70.

- Pfeffer M, Maurer M, Köllensperger G, Hann S, Graf AB, Mattanovich D. 2011. Modeling and measuring intracellular fluxes of secreted recombinant protein in *Pichia pastoris* with a novel 34S labeling procedure. *Microb. Cell Fact.* 10:47.
- 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.
- Rebnegger C, Graf AB, Valli M, Steiger MG, Gasser B, Maurer M, Mattanovich D. 2014. In *Pichia pastoris*, growth rate regulates protein synthesis and secretion, mating and stress response. *Biotechnol. J.* 9:511–525.
- Reiner A, Yekutieli D, Benjamini Y. 2003. Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics* 19:368–375.
- Resina D, Cos O, Ferrer P, Valero F. 2005. Developing high cell density fed-batch cultivation strategies for heterologous protein production in *Pichia pastoris* using the nitrogen source-regulated FLD1 promoter. *Biotechnol. Bioeng.* 91:760–767.
- Resina D, Bollók M, Khatri NK, Valero F, Neubauer P, Ferrer P. 2007. Transcriptional response of *P. pastoris* in fed-batch cultivations to *Rhizopus oryzae* lipase production reveals UPR induction. *Microb. Cell Fact.* 6:21.
- Resina D, Maurer M, Cos O, Arnau C, Carnicer M, Marx H, Gasser B, Valero F, Mattanovich D, Ferrer P. 2009. Engineering of bottlenecks in *Rhizopus oryzae* lipase production in *Pichia pastoris* using the nitrogen source-regulated FLD1 promoter. *N. Biotechnol.* 25:396–403.
- Rottensteiner H, Theodoulou FL. 2006. The ins and outs of peroxisomes: Coordination of membrane transport and peroxisomal metabolism. *Biochim. Biophys. Acta - Mol. Cell Res.* 1763:1527–1540.
- Ruth C, Buchetics M, Vidimce V, Kotz D, Naschberger S, Mattanovich D, Pichler H, Gasser B. 2014. *Pichia pastoris* Aft1 - a novel transcription factor, enhancing recombinant protein secretion. *Microb. Cell Fact.* 13:120.
- Rußmayer H, Buchetics M, Gruber C, Valli M, Grillitsch K, Modarres G, Guerrasio R, Klavins K, Neubauer S, Drexler H, Steiger M, Troyer C, Al Chalabi A, Krebiehl G, Sonntag D, Zellnig G, Daum G, Graf AB, Altmann F, Köllensperger G, Hann S, Sauer M, Mattanovich D, Gasser B. 2015. Systems-level organization of yeast methylotrophic lifestyle. *BMC Biol.* 13:80

- Sha C, Yu XW, Lin NX, Zhang M, Xu Y. 2013. Enhancement of lipase r27RCL production in *Pichia pastoris* by regulating gene dosage and co-expression with chaperone protein disulfide isomerase. *Enzyme Microb. Technol.* 53:438–443
- Smyth, G. K. (2005). Limma: linear models for microarray data. In: *Bioinformatics and Computational Biology Solutions using R and Bioconductor*, R. Gentleman, V. Carey, S. Dudoit, R. Irizarry, W. Huber (eds.), Springer, New York, pages 397–420.
- Smith JJ, Aitchison JD. 2013. Peroxisomes take shape. *Nat. Rev. Mol. Cell Biol.* 14:803–817.
- Sturn A, Quackenbush J, Trajanoski Z. 2002. Genesis: cluster analysis of microarray data. *Bioinformatics* 18:207–8
- Tamura N, Oku M, Ito M, Noda NN, Inagaki F, Sakai Y. 2013. Atg18 phosphoregulation controls organellar dynamics by modulating its phosphoinositide-binding activity. *J. Cell Biol.* 202:685–698.
- Taylor SC, Berkelman T, Yadav G, Hammond M. 2013. A defined methodology for reliable quantification of Western Blot data. *Mol. Biotechnol.* 55:217–226.
- Van Roermund CWT, Ijlst L, Majczak W, Waterham HR, Folkerts H, Wanders RJ a, Hellingwerf KJ. 2012. Peroxisomal fatty acid uptake mechanism in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 287:20144–20153.
- Van Roermund CWT, Waterham HR, Ijlst L, Wanders RJ a. 2003. Fatty acid metabolism in *Saccharomyces cerevisiae*. *Cell. Mol. Life Sci.* 60:1838–1851
- Vanz AL, Lünsdorf H, Adnan A, Nimtz M, Gurramkonda C, Khanna N, Rinas U. 2012. Physiological response of *Pichia pastoris* GS115 to methanol-induced high level production of the Hepatitis B surface antigen: catabolic adaptation, stress responses, and autophagic processes. *Microb. Cell Fact.* 11:103.
- Vassileva A, Chugh D A, Swaminathan S, Khanna N. 2001. Effect of copy number on the expression levels of hepatitis B surface antigen in the methylotrophic yeast *Pichia pastoris*. *Protein Expr. Purif.* 21:71–80.
- Vogl T, Glieder A. 2013. Regulation of *Pichia pastoris* promoters and its consequences for protein production. *N. Biotechnol.* 30:385–404.

- Yadava A, Ockenhouse CF. 2003. Effect of Codon Optimization on Expression Levels of a Functionally Folded Malaria Vaccine Candidate in Prokaryotic and Eukaryotic Expression Systems . *Infect. Immun.* 71:4961–4969.
- Yang H, Zhai C, Yu X, Li Z, Tang W, Liu Y, Ma X, Zhong X, Li G, Wu D, Ma L. 2016a. High-level expression of Proteinase K from *Tritirachium album Limber* in *Pichia pastoris* using multi-copy expression strains. *Protein Expr. Purif.* 122:38–44.
- Yang J, Lu Z, Chen J, Chu P, Cheng Q, Liu J, Ming F, Huang C, Xiao A, Cai H, Zhang L. 2016b. Effect of cooperation of chaperones and gene dosage on the expression of porcine PGLYRP-1 in *Pichia pastoris*. *Appl. Microbiol. Biotechnol.* 100: 5453-5465
- Yang S, Kuang Y, Li H, Liu Y, Hui X, Li P, Jiang Z, Zhou Y, Wang Y, Xu A, Li S, Liu P, Wu D. 2013. Enhanced Production of Recombinant Secretory Proteins in *Pichia pastoris* by Optimizing Kex2 P1' site. *PLoS One* 8:1–11.
- Yeung KY, Ruzzo WL. 2001. Principal component analysis for clustering gene expression data. *Bioinformatics* 17:763–774.
- Yu AQ, Zhu JC, Zhang B, Xing LJ, Li MC. 2012. Knockout of fatty acid desaturase genes in *Pichia pastoris* GS115 and its effect on the fatty acid biosynthesis and physiological consequences. *Arch. Microbiol.* 194:1023–1032.
- Yurimoto H, Oku M, Sakai Y. 2011. Yeast methylotrophy: Metabolism, gene regulation and peroxisome homeostasis. *Int. J. Microbiol* 2011:101298
- Wang JR, Li YY, Liu DN, Liu JS, Li P, Chen LZ, Xu SD. 2015. Codon Optimization Significantly Improves the Expression Level of α -Amylase Gene from *Bacillus licheniformis* in *Pichia pastoris*. *Biomed Res. Int.* 2015:1–9.
- Zhu T, Guo M, Tang Z, Zhang M, Zhuang Y, Chu J, Zhang S. 2009. Efficient generation of multi-copy strains for optimizing secretory expression of porcine insulin precursor in yeast *Pichia pastoris*. *J. Appl. Microbiol.* 107:954–63.
- Zhu T, Guo M, Zhuang Y, Chu J, Zhang S. 2011. Understanding the effect of foreign gene dosage on the physiology of *Pichia pastoris* by transcriptional analysis of key genes. *Appl. Microbiol. Biotechnol.* 89:1127–35.

4. La desregulació del metabolisme del metanol reverteix les limitacions transcripcionals de les soques recombinants de *P. pastoris* amb múltiples cassets d'expressió sota el control de P_{AOX1}

Paraules clau: *Pichia pastoris*, Mxr1, regulació AOX1, promotor AOX1, metabolisme del metanol, producció de proteïna recombinant, ROL, factor de transcripció, proteïnes 14-3-3, lipasa

4.1 INTRODUCCIÓ

El llevat metilotròfic *Pichia pastoris* (*Komagataella*) és un hoste àmpliament utilitzat en la producció de proteïnes recombinants, i un dels sistemes preferits per la producció d'un ampli ventall de proteïnes, des de biofarmacèutics i enzims industrials fins proteïnes de membranes i nanoestructures complexes (Corchero et al., 2013; Bill, 2015; Byrne, 2015; Rabert et al., 2013; Spohner et al., 2015). Les característiques més importants i les metodologies implementades per l'expressió de proteïnes en *P. pastoris* han estat revisades extensament (Ahmad et al., 2014; Cregg et al., 2000; Gasser et al., 2013; Vogl et al., 2013). Un dels factors més determinants per la tria d'aquest hoste és la presència del promotor de l'alcohol oxidasa 1 (P_{AOX1}), un promotor regulable d'expressió forta, convertint-lo així en l'opció més freqüent per dirigir l'expressió de gens forans a *P. pastoris* (Vogl i Glieder, 2013). La repressió per catabòlit de P_{AOX1} per part de la glucosa i el glicerol, entre d'altres fonts de carboni, es troba descrita des de fa dècades (Tschopp et al., 1987), i conseqüentment, al llarg del temps s'han anat optimitzant de manera extensiva les estratègies operacionals (Looser et al., 2014). D'altra banda, s'han realitzat diversos estudis de fluxos metabòlics on es detalla l'impacte de la producció de proteïna recombinant en el metabolisme de *P. pastoris* creixent en substrats mixtes de metanol i altres fonts de carboni (Isidro et al., 2016; Jordà et al., 2012; Jordà et al., 2014a; Jordà et al., 2014b; Ferrer i Albiol, 2014). No obstant, les bases moleculars de la regulació transcripcional de P_{AOX1} s'han desvelat els darrers 10 anys. Primerament, Lin-Cereghino et al. (2006) van identificar un factor de transcripció (TF, de l'anglès *Transcription Factor*) amb un domini d'unió al DNA del tipus dit de zinc, anomenat Mxr1 (de l'anglès *Methanol Expression Regulator 1*), homòleg al TF Adr1 de *Saccharomyces cerevisiae*. Les mutacions en aquest gen van resultar en soques incapaces de créixer en metanol i oleat i d'induir la transcripció de l'alcohol oxidasa (*AOX1*), la dihidroxiacetona sintasa (*DAS*), la peroxina 8 (*PEX8*) i la peroxina 14 (*PEX14*). Així mateix, posteriors investigacions per Kranthi et al. (2009 i 2010) van revelar 5, 6 i 7 llocs d'unió al Mxr1 en els promotors de *PEX8*, *AOX1* i *DAS*, respectivament. Posteriorment, es va caracteritzar una família de proteïnes 14-3-3 (reguladores d'innombrables processos biològics als eucariotes, Fu et al., 2000) per part de Parua et al. (2012) a *P. pastoris*. Aquests autors van descriure a Mxr1 un motiu d'unió a les proteïnes 14-3-3, desvelant que la interacció entre ambdues proteïnes era deguda a la fosforilació de la Ser215 a Mxr1 en presència de fonts de carboni repressores, comportant així la repressió dels gens dependents de Mxr1. Més tard es van descriure d'altres factors de transcripció responsables de la regulació del metanol: Mit1 (de l'anglès *methanol-induced transcription factor*), Rop (de l'anglès *repressor of phosphoenolpyruvate carboxykinase*) i Trm1 (Kumar i Rangarajan, 2012; Sahu et al., 2014; Wang et al., 2016). Breument, Wang et al.

(2016) van proposar una transmissió en cascada del senyal d'inducció per metanol entre Mxr1, Mit1 i Trm1, sent Mxr1 el responsable de la derepressió de P_{AOX1} , mentre que el metanol induiria Trm1 i seguidament l'expressió de Mit1. Finalment, Kumar i Rangarajan (2012) van demostrar que Rop és un repressor del metabolisme del metanol, competint amb Mxr1 pels mateixos llocs d'unió a P_{AOX1} , resultant en una disminució dels nivells transcripcionals d' $AOX1$.

En aquest context, un estudi recent del nostre grup de recerca (Cámara et al., 2016b) ha reportat una limitació transcripcional en la capacitat assimilatòria del metanol en diferents soques de *P. pastoris* expressant més d'una còpia del gen codificant per la lipasa de *Rhizopus oryzae* (Rol) sota el control del promotor P_{AOX1} .

Degut l'expressió constitutiva de Mxr1 a nivells baixos (Lin-Cereghino et al., 2006), els nostres resultats van suggerir que la presència de múltiples còpies (o regions) de P_{AOX1} resultava en un nombre insuficient de molècules de Mxr1 per tal d'induir l'expressió dels gens relacionats amb el metabolisme del metanol, incloent la proteïna recombinant Rol. D'acord amb aquesta hipòtesi, Takagi et al. (2009) van reportar prèviament un increment en la producció de proteïna recombinant degut a la sobreexpressió de Mxr1, tot i que no es van descriure paràmetres fisiològics a la patent.

En aquest estudi es descriu la coexpressió d'una variant desregulada de Mxr1 (Parua et al., 2012), anomenada Mxr1*, en una sèrie de soques de *P. pastoris* amb una i quatre còpies de *ROL* (anomenades 1C i 4C, respectivament, Cámara et al., 2016a). Mxr1* conté una mutació a la posició Ser215 que prevén completament la unió de les proteïnes 14-3-3 a aquest TF, assegurant així l'acció constant de Mxr1. L'efecte de l'expressió de Mxr1* a les soques 1C i 4C modificades es va testar en cultius en Erlenmeyer. A més, es va realitzar un anàlisi transcripcional de 8 gens clau implicats en diferents processos biològics rellevants (metabolisme del metanol, plegament i biogènesi dels peroxisomes) entre la soca control 4C i una soca 4C amb dues còpies addicionals de Mxr1*, en cultius operats en continu i condicions de creixement controlades.

4.2 MATERIALS I MÈTODES

4.2.1 Soques i plasmidis

Les soques de *P. pastoris* amb una i quatre còpies de la lipasa de *R. oryzae* (anomenades 1C i 4C, respectivament) es van construir prèviament per Cámara et al. (2016a), i juntament amb la soca no productora X-33 (anomenada 0C) es van designar com a soques referència. Aquestes dues soques es van utilitzar per a la transformació amb el vector d'expressió. El vector pGAPHA, derivat del plasmidi pGAPZ α en el qual la resistència a Zeocina s'havia substituït per una resistència a Higromicina (construït prèviament al nostre laboratori; Adelantado, 2016), es va utilitzar en aquest estudi com a vector d'expressió de Mxr1*.

4.2.2 Construcció del vector d'expressió pGAPHA_MXR1*

La variant mutant del gen MXR1 (MXR1*), amb la posició Ser215 canviada per una alanina, es va amplificar per PCR des del plasmidi pG213 (cortesia d'E.T. Young) utilitzant els oligonucleòtids MXR1*_PmlI i MXR1*_KpnI (l·listats a l'annex) i la polimerasa Q5® Hot Start High-Fidelity (New England Biolabs, Ipswich, MA, EUA). Després de l'amplificació, el producte de PCR es va seqüenciar utilitzant els oligonucleòtids MXR1*_seq_1-5 (annex). Posteriorment, MXR1* es va clonar en el vector pGAPHA entre els llocs PmlI i KpnI, sota el control del promotor del gen de la gliceraldehid-3-fosfat deshidrogenasa (P_{GAP}). El plasmidi resultant (Figura 1) es va seqüenciar utilitzant els oligonucleòtids pGAP_fw, MXR1*_int_rev i MXR1*_seq_5-6 (annex) per tal de confirmar la inserció en el marc de lectura correcte. La seqüenciació dels plasmidis i els productes de PCR es van dur a terme pel Servei de Genòmica i Bioinformàtica (Bellaterra, Cerdanyola del Vallès, Espanya).

4.2.3 Transformació de *P. pastoris* i selecció de clons

Les cèl·lules competents es van preparar seguint el protocol descrit per Cregg (2007). Es van introduir per electroporació 100 ng de pGAPHA linealitzar per XbaI, mitjançant el dispositiu Gene Pulser Xcell™ Electroporation System (Bio-Rad, Hercules, CA, EUA). Els ajustaments van ser 1500 V, 25 μ F, i 200 Ω . Seguidament es van seleccionar els transformants en plaques YPD (de l'anglès *Yeast Extract Peptone Dextrose*) ((1% (w/v) extracte de llevat, 2% (w/v) peptona, 2% (w/v) dextrosa)) amb 250 μ g/mL d'Higromicina (Invivogen, San Diego, CA, EUA). Els clons positius es van confirmar per PCR (prèvia extracció genòmica (Wizard® Genomic DNA Purification Kit, Promega, Madison, WI, EUA)), mitjançant l'amplificació de l'insert MXR1* utilitzant els oligonucleòtids pGAP i MXR1*_int (annex).

4.2.4 Condicions de creixement

4.2.4.1 *Cultius en Erlenmeyer*

Els cultius de *P. pastoris* en Erlenmeyer es van realitzar per triplicar de la següent manera: 50 mL de medi fresc BMG (de l'anglès *Buffered Minimal Glycerol*) (100mM fosfat potàssic (pH 6.0), 1.34% (w/v) base de llevat nitrogenada sense aminoàcids, $4 \cdot 10^{-5}$ % (w/v) biotina, 1% (v/v) glicerol) en Erlenmeyers de 0.5 L amb deflectors es van inocular amb el cultiu d'una nit en YPD de la soca seleccionada a una densitat òptica a 600 nm (DO_{600}) de 0.1, i es van incubar a 30°C i 150 rpm en un agitador orbital Multitron II (Infors, Einsbach, Alemanya). Després de 20 h, el cultiu (amb una DO_{600} aproximada de 2-6) es va centrifugar, i es va utilitzar per inocular Erlenmeyers de 0.25 L, amb deflectors i taps amb filtre (Duran Group, Mainz, Alemanya), amb 35 mL de medi BMM (de l'anglès *Buffered Minimal Methanol*) (100mM fosfat potàssic (pH 6.0), 1.34% (w/v) base nitrogenada de llevat sense aminoàcids, $4 \cdot 10^{-5}$ % (w/v) biotina, 0.5% (v/v) metanol) a una DO_{600} de 1. Cada 24 h es va afegir a cada cultiu un pols de metanol pur a una concentració final de 0.5% v/v. Tots els medis es van suplementar amb Higromicina (150 µg/mL).

4.2.4.2 *Cultius operats en continu*

Els cultius operats en continu de les soques seleccionades es van dur a terme en un reactor de 1.5 L (Biostat B Plus, Sartorius AG, Göttingen, Alemanya), en un volum de treball d'1 L. Com a precultiu, es van inocular 150 mL de medi YPD-Zeocina o YPD-Higromicina (per la soca referència i per la soca MXR1*, respectivament) amb 1 mL de criovial de la soca seleccionada en un Erlenmeyer d'1 L, es va incubar aproximadament 24 h a 30°C i 150 rpm en un agitador orbital Multitron II, i posteriorment es va utilitzar per tal d'inocular un volum d'1 L de medi *batch*. Al final de la fase *batch*, les soques es van fer créixer en condicions de carboni limitant a una taxa de dilució (D) de 0.1 h^{-1} , amb un medi de creixement definit i una barreja de glicerol/metanol (60%/40%, w/w) com a font de carboni. Pressió, pH, agitació, aeració i temperatura es van controlar a 1.2 bars, pH 5 (ajustat amb 10% NH_3 v/v), 700 rpm, 1 vvm i 25°C, respectivament. Els gasos de sortida del reactor es van fer passar per un condensador refrigerat a 4°C mitjançant un criòstat Ecoline Staredition RE 106 (Lauda Dr. R. Wobser GmbH&Co. KG, Lauda-Königshofen, Alemanya), seguit d'un assecat mitjançant dues columnes de sílice gel i subseqüentment analitzat amb l'espectròmetre Omnistar™ GSD 300 02 (Balzers Instruments, Balzers, Liechtenstein). Les mostres es van prendre al cap d'un mínim de cinc temps de residència, un cop es va assolir l'estat estacionari (Jordà et al., 2013).

El medi *batch* contenia: : 19.95 g/L glicerol, 0.9 g/L àcid cítric, 6.3 g/L $(\text{NH}_4)_2\text{HPO}_4$, 0.01 g/L CaCl_2 , 0.45 g/L KCl, 0.25 g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 1 mL/L biotina (0.2 g/L; Sigma, Munich, Alemanya) i 2.3 mL/L de solució sals traça PTM1.

El medi de la fase en continu contenia: 15 g/L glicerol, 10 g/L metanol, 0.42 g/L àcid cítric, 2.18 g/L $(\text{NH}_4)_2\text{HPO}_4$, 5.5 mg/L $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 0.85 g/L KCl, 0.32 g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.5 mL biotina (0.2 g/L) i 0.8 mL de solució sals traça PTM1. El pH es va ajustar a 5.0 en ambdós medis mitjançant HCl 25%. La composició de les sals traça PTM1 és la mateixa descrita a Baumann et al. (2008).

4.2.5 Procediments analítics

4.2.5.1 Determinació de la biomassa

La biomassa cel·lular es va monitorar mesurant la DO_{600} . Per la determinació del pes sec, es va utilitzar el mètode prèviament reportat per Jordà et al. (2012). Totes les mesures es van realitzar per triplicat.

4.2.5.2 Assaig d'activitat lipolítica

La determinació de l'activitat lipolítica es va realitzar per triplicat tal i com va descriure prèviament Resina et al. (2004) amb l'espectrofotòmetre Specord 200 Plus (Analytik Jena AG, Jena, Alemanya).

4.2.5.3 Quantificació de metabòlits

El glicerol, el metanol i d'altres compostos extracel·lulars d'interès en els cultius en Erlenmeyer i en els cultius en continu es van analitzar tal i com es descriu a Cámara et al. (2016a). Les anàlisis es van realitzar per duplicat per cada experiment independent.

4.2.6 Anàlisi per citometria de flux

El comptatge i la viabilitat cel·lular es van mesurar utilitzant els mètodes descrits a Cámara et al. (2016a). Per tal de determinar la mida cel·lular, es va fer servir el kit Flow Cytometry Size Calibration (Thermo Fisher Scientific, Waltham, MA, EUA). Breument, diferents suspensions de microesferes amb un diàmetre de 1 μm , 2 μm , 4 μm , 6 μm , 10 μm i 15 μm es van sonicar durant 5 s a 50 W amb el processador d'ultrasons VC-5 (Vibracell, Sonics & Materials, Newtown, CT, EUA) abans de passar les mostres pel citòmetre i procedir amb l'anàlisi. Tot seguit, es van establir diferents marcadors visuals al software per tal de definir la mida de cada interval a les mostres processades. Els experiments es van dur a terme amb el citòmetre Guava EasyCyte Mini

(Millipore, Jaffrey, NH, EUA), equipat amb un làser d'argó de 488 nm. Per cada anàlisi es van mesurar 5,000 cèl·lules a un cabal de 0.59 mL/s. Els assajos de viabilitat es van realitzar per duplicat, mentre que el comptatge i la mida cel·lular es van determinar per triplicat per cada cultiu en Erlenmeyer o cultiu en continu.

4.2.7 Determinació de la dosi gènica per ddPCR

La quantificació de la dosi gènica de *ROL* i *MXR1* es va realitzar amb la ddPCR, utilitzant com a gen referència la β -actina (*ACT1*) com a control endogen per normalitzar els resultats. Per l'amplificació del DNA, es van utilitzar els oligonucleòtids *MXR1*, *ACT1* i *ROL* (annex) (Cámara et al., 2016a). Breument, el DNA genòmic es va extreure utilitzant un kit comercial de purificació (Wizard® Genomic DNA Purification Kit, Promega, Madison, WI, EUA) i subseqüentment va ser digerit amb BamHI i EcoRI (Thermo Scientific™, Waltham, MA). Cada reacció de ddPCR es va dur a terme en un volum final de 20 μ L, amb 10 μ L QX200™ ddPCR™ EvaGreen Supermix, 200 nM de l'oligonucleòtid *forward*, 100 nM de l'oligonucleòtid *reverse*, 0.8 ng de DNA genòmic digerit i la quantitat requerida d'aigua *DNase/RNase-free*. Les condicions de la ddPCR van ser les següents: preincubació a 95°C durant 10 min i posterior pas de desnaturalització (94°C, 30 s), seguit d'un pas de alineament/extensió (56.5°C per *ACT1* i *MXR1*, 60.2°C per *ROL*; 1 min) durant 40 cicles. La formació de gotes es va realitzar amb el dispositiu Droplet Generator QX200™, mentre que la detecció es va dur a terme amb el QX100™ Droplet Digital Reader i el software Quantasoft v. 1.7.4.0917. El nombre de còpies de cada gen es va calcular per cada mostra per duplicat, calculant la ràtio entre les gotes positives de les reaccions de PCR de *MXR1* o *ROL* i *ACT1*. Per tal de verificar l'estabilitat de la dosi gènica de *MXR1* i *ROL* al final dels cultius en continu, es va prendre mostres de biomassa al cap de cinc temps de residència, un cop assolit l'estat estacionari. Tots els reactius es van adquirir a Bio-Rad (Hercules, CA, EUA), mentre que els oligonucleòtids van ser sintetitzats per Sigma-Aldrich (St Louis, MO, EUA).

4.2.8 Determinació dels nivells transcripcionals per ddPCR

Per la quantificació transcripcional es van centrifugar 5 mL de cèl·lules al final dels cultius en Erlenmeyer, prèviament tractats amb fenol (5%, v/v) i posteriorment es van emmagatzemar a -80°C. L'extracció de RNA es va realitzar amb el kit Rneasy Mini (Qiagen, Hilden, Alemanya) mentre que el cDNA es va sintetitzar amb el kit iScript™ cDNA Synthesis (Bio-Rad), seguint les instruccions del fabricant. La qualitat del RNA es va determinar mesurant la ràtio d'absorbància a 260 i 280 nm, utilitzant el Nanodrop 1000 (Thermo Fisher Scientific). Les reaccions de ddPCR contenien 10 µL de QX200™ ddPCR™ EvaGreen Supermix, 200 nM de l'oligonucleòtid *forward*, 100 nM de l'oligonucleòtid *reverse*, 0.4 ng de cDNA i la quantitat requerida d'aigua *DNase/RNase-free* fins a un volum final de 20 µL. Les reaccions es van incubar a 95°C durant 10 min, seguit d'un pas de desnaturalització (94°C, 30 s) i un pas d'alineament/extensió (60.2°C per *KAR2* i *ROL*; 56.5°C per la resta; 1 min) durant 40 cicles. Les gotes positives de cada reacció de ddPCR es van normalitzar en relació a la senyal d'*ACT1*. Les seqüències dels oligonucleòtids es mostren a l'annex.

4.2.9 Anàlisi estadístic

Els resultats s'expressen com a mitjana \pm SD. L'anàlisi estadístic dels resultats es va realitzar la prova *t* de Student (mostres no aparellades) utilitzant el software Excel de Microsoft. Es va considerar estadísticament significatiu un valor de *P*-val inferior a 0.05.

4.3 RESULTATS I DISCUSSIÓ

4.3.1 Construcció i caracterització de soques productores de Rol coexpressant *MXR1**

Amb l'objectiu d'incrementar els nivells de Mxr1 a les soques productores de Rol, alhora que assegurar una activació constant dels promotors dependents de Mxr1 en condicions d'inducció, es van introduir múltiples còpies de *MXR1** (codificant per una variant de Mxr1 amb una mutació a la Ser215) a les soques referència 1C i 4C. Donat que estudis previs havien reportat l'efecte letal de la sobreexpressió de Mxr1 sota el control de promotors induïbles (P_{AOX1} i P_{PEX8} , Lin-Cereghino et al., 2006), però cert creixement amb promotors constitutius (Takagi et al., 2012; Sahu i Rangarajan, 2015), es va seleccionar P_{GAP} com a promotor per dirigir l'expressió de *MXR1**. El vector d'expressió pGAPHA_ *MXR1** (Figura 1A) es va transformar en les soques productores de Rol 1C i 4C, resultant en 5 i 11 clons aïllats en plaques d'agar amb Higromicina per cada soca, respectivament. Tots els transformants per la soca 1C i 3 d'11 per la soca 4C van amplificar pel producte de PCR *MXR1** (Figura 1B). Curiosament, només 3 dels clons 1C_ *MXR1** (1C₁, 1C₂ i 1C₃) van créixer en medi líquid. La posterior quantificació per ddPCR va determinar la dosi gènica total de *MXR1** (*MXR1*+*MXR1**) per cada clon (Figura 1C), obtenint així una població amb una còpia (1C_ *MXR1**a i 1C_ *MXR1**b provinents de la soca 1C; 4C_ *MXR1**a i 4C_ *MXR1**b, derivats de la soca 4C) o dues còpies de *MXR1** (1C_2*MXR1** i 4C_2*MXR1**), demostrant que P_{GAP} era una opció adient per l'expressió de *MXR1**. El fet que només els clons amb una dosi gènica baixa de *MXR1** fossin capaços de créixer en medi líquid podria ser degut a l'efecte nociu d'altres dosis d'aquest gen, suggerint que els clons positius que només creixien en placa podrien portar més de dues còpies de *MXR1**.

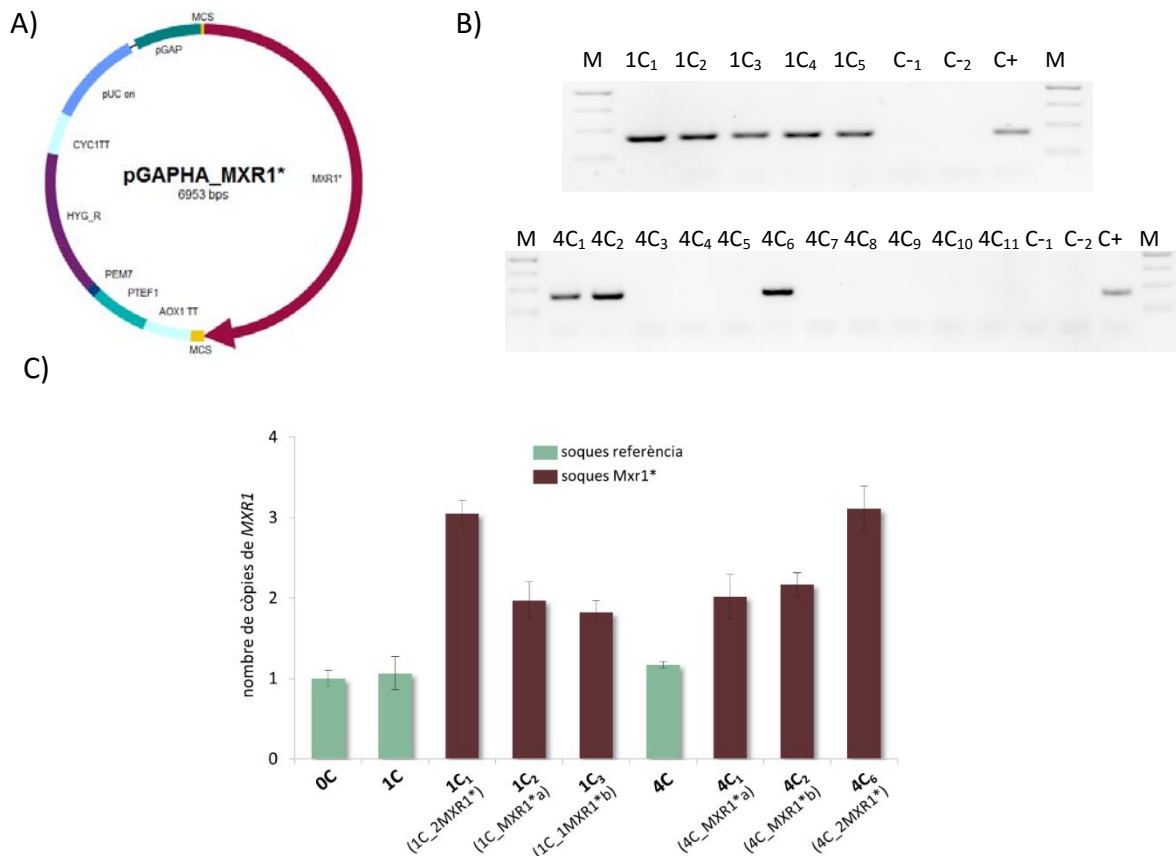


Figura 1. Construcció de la sèrie de soques expressant *MXR1.** A) Mapa del plasmidi d'expressió pGAPHA_MXR1*. pUC, origen de replicació; pGAP, promotor *GAP* de *P. pastoris*; MCS, lloc múltiple de clonatge; *MXR1**, variant del gen *MXR1* amb la posició Ser215 mutada; AOX1 TT, terminador de la transcripció d'*AOX1*; pTEF1, promotor *TEF1* de *P. pastoris*, Hyg_R, gen de la resistència a Higromicina B) Verificació del producte de PCR de *MXR1**. C-1, soca salvatge OC; C-2, aigua; C+, fragment de *MXR1**, ; M, marcador 1 kb DNA (NEB, Ipswich, MA, EUA) C) Determinació del nombre de còpies total de *MXR1* (*MXR1*+*MXR1**) mitjançant ddPCR. Donat que al genoma de *P. pastoris* només hi ha una còpia del gen *MXR1* (De Schutter et al., 2009; Mattanovich et al., 2009), es va considerar que els clons amb dues o tres còpies de *MXR1* tenien en conseqüència una o dues còpies de *MXR1**, respectivament. En negreta, la nomenclatura utilitzada en la verificació de *MXR1** per PCR. Entre parèntesis, la nomenclatura final basada en els resultats de la quantificació de *MXR1** per ddPCR.

4.3.2 L'expressió de *MXR1 incrementa els nivells de producció de Rol i reverteix l'efecte perjudicial de la sobreexpressió de *ROL* en l'assimilació del metanol a les soques multicòpia**

Segons estudis previs (Cámara et al., 2016a; Cámara et al., 2016b), les soques de *P. pastoris* amb més d'una còpia del casset d'expressió *ROL* mostren una disminució en el creixement i la capacitat d'assimilació del metanol. Per això, primerament es va dur a terme una sèrie de cultius en Erlenmeyer per tal d'analitzar l'efecte de la coexpressió de *MXR1** en el creixement, el consum de metanol i la producció de Rol després de 70 h de cultiu. Tal i com s'esperava, els nivells de biomassa de la soca 4C van ser 3 vegades menor que la soca productora 1C (Figura 2A). A més, es van detectar nivells significatius de metanol residual al final del cultiu de la soca 4C (Figura 2D), mentre que no es va observar acumulació en els cultius de 1C. En el cas de les soques coexpressant *MXR1**, una de les soques 1C amb una còpia de *MXR1** (1C_*MXR1**a) va mostrar una disminució significativa en el seu creixement comparat amb la soca 1C, tot i que no es va detectar metanol residual. Contràriament, dues de les tres soques 4C_*MXR1** van incrementar els nivells de biomassa dues vegades en comparació a la soca referència 4C. De manera consistent, mentre que a la soca 4C i 4C_*MXR1**a (ambdues amb nivells similars de creixement) es van mesurar 5-6 g/L de metanol residual, els nivells més elevats de biomassa de les soques 4C_*MXR1**b i 4C_2*MXR1** es van veure acompanyats de nivells significativament més baixos de metanol al medi. En termes de viabilitat cel·lular, gairebé totes les cèl·lules en el cultiu de la soca no productora 0C eren viables (97%, dades no mostrades), mentre que es va detectar una baixada del 20% en la viabilitat de les soques 1C i 4C. De manera similar, totes les soques *MXR1** van mostrar un 10-20% de cèl·lules no viables, excepte la soca 4C_*MXR1**a, que va presentar un 40% de cèl·lules no viables.

En relació a la mida cel·lular, mentre que es va observar un lleuger (però significatiu) canvi global en la mida cel·lular de les soques 0C i 1C, es va mesurar un increment notable en les fraccions de mida més gran (6 μm i $\geq 10 \mu\text{m}$) en comparació amb les altres soques referència (Figura 2B). De manera interessant, també es van observar diferències significatives en la distribució de la mida cel·lular entre les poblacions de les soques 4C_*MXR1**b i 4C_2*MXR1** en comparació a la soca control 4C (Figura 2B). Concretament, es va detectar un increment del 20% en la població de 4 μm en les soques 4C-*MXR1**, mentre que la fracció 6 μm es va reduir un 15%, resultant en una distribució de la mida similar a la de la soca control 1C, indicant així una possible reversió dels efectes nocius de la sobreexpressió de *ROL* a la soca 4C mitjançant l'expressió de *MXR1**.

Tal i com havíem reportat anteriorment (Cámara et al., 2016b), l'increment de la dosi gènica de *ROL* comporta nivells superiors de productivitat a la soca 4C en comparació a 1C. De manera interessant, a totes les soques 1C_MXR1* la producció de Rol va ser dos i tres vegades superior a la soca control 1C (Figura 2C).

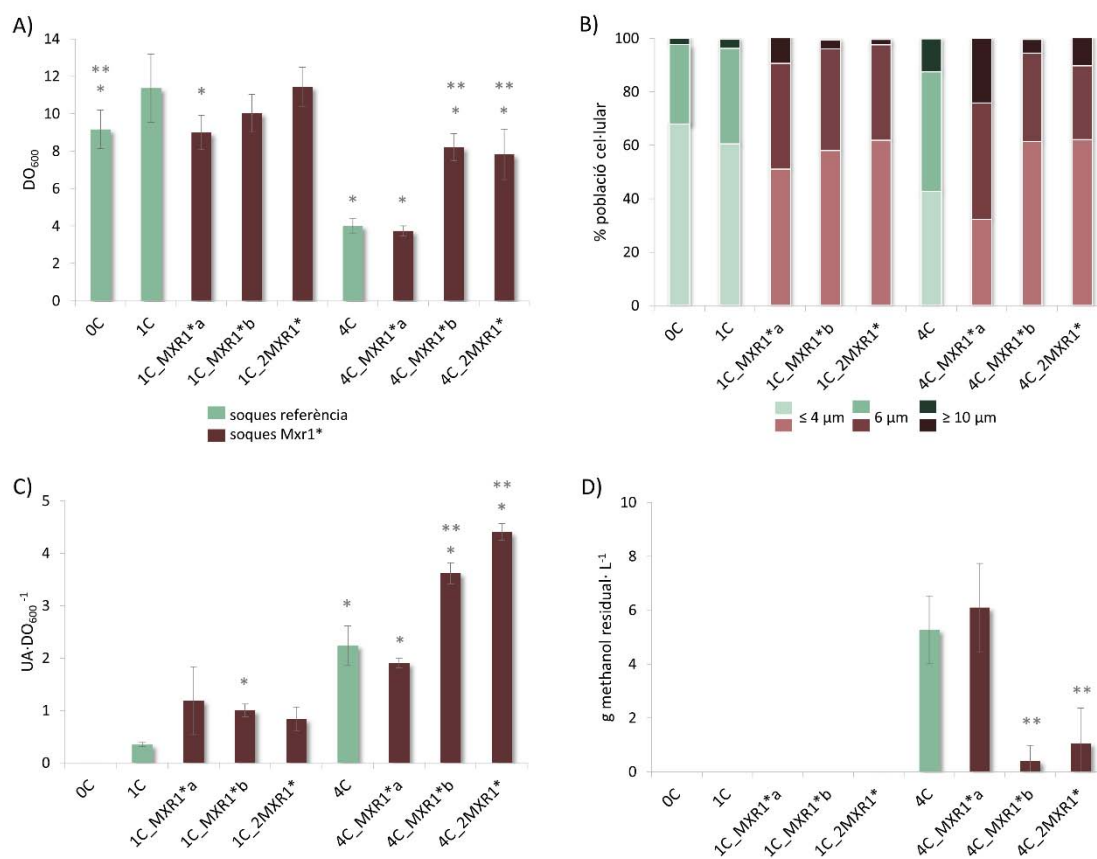


Figura 2. Cultius en Erlenmeyer dels clons transformants. A) Nivells de biomassa mesurats per DO₆₀₀ B) Mida de la població cel·lular mesurada per citometria de flux C) Activitat lipàsica específica als cultius D) Metanol residual al final dels cultius quantificat per HPLC. Totes les mostres es van prendre després de 70 h de cultiu. En blau, 0C, 1C i 4C (soques referència). En granat, soques expressant *MXR1**. 1C_MXR1*a, 1C_MXR1*b i 1C_2MXR1*, soques amb una i dues còpies de *MXR1**, resultants de la transformació de 1C; 4C_MXR1*a, 4C_MXR1*b i 4C_2MXR1*, soques amb una i dues còpies de *MXR1** de la transformació de 4C. UA, unitats d'activitat lipàsica. **P* < 0.05, comparat amb la soca control 1C; ***P* < 0.05, comparat amb la soca control 4C.

En relació a les soques 4C modificades, 4C_MXR1*b i 4C_2MXR1* van presentar una activitat lipàsica total dues vegades superior a la soca 4C (és a dir, un increment de 6 vegades en relació a la soca control 1C). Tal i com es mostra a la Figura 2D, aquest increment en la producció de lipasa va

anar associat a una caiguda sobtada dels nivells de metanol residual al medi, reflectint una taxa de consum de metanol similar a les soques 1C. Encara més, les soques amb una assimilació del metanol més elevada (0C i 1C) presentaven una mida mitjana de la població inferior a aquelles amb taxes de consum més baixes (4C i 4C_MXR1*a). Aquests resultats podrien explicar la mida similar de les soques 1C, 4C_MXR1*b i 4C_2MXR1, com a resultat d'una redistribució dels fluxos metabòlics que comportaria una assimilació similar del metanol. Curiosament, tot i que es van observar millores significatives en els rendiments de biomassa i producte (juntament amb les taxes de consum de metanol) a les soques 4C_MXR1*b i 4C_2MXR1*, no es van detectar canvis a la soca 4C_MXR1*a (amb una còpia de *MXR1**) en comparació a la soca 4C.

4.3.3 Els cultius en continu van revelar un canvi de preferència en la font de carboni dependent de la soca

Per tal de reproduir i validar els resultats obtinguts en els cultius en Erlenmeyer, es va fer créixer en cultius en continu les soques 4C i 4C_2MXR1*, amb una D de 0.09 h⁻¹ utilitzant una barreja de glicerol metanol (60/40% v/v) com a aliment (és a dir, en condicions anàlogues als nostres estudis previs, Cámara et al., 2016a; Cámara et al., 2016b). En termes d'activitat lipàsica extracel·lular, es va arribar a un increment del 30% en la productivitat volumètrica de la soca 4C_2MXR1* en comparació amb 4C, corresponent a un lleuger augment del 10% en termes de productivitat lipàsica específica (Taula I). Així mateix, els cultius van confirmar l'impacte fisiològic de la sobreexpressió de *MXR1** en 4C. Concretament, els valors de biomassa van ser un 30% superiors a 4C_2MXR1* que a la soca 4C, també reflectit en un increment del 15% en el rendiment biomassa/substrat (Taula I). En les condicions de cultiu establides, els continus de les soques referència 0C i 1C es troben en condicions de carboni limitant, però en el cas de les soques multicòpia (4C) s'observa metanol residual al reactor. En aquest experiment es va observar un canvi substancial entre la soca *MXR1** i la soca control en referència al consum específic de la font de carboni. El consum específic de glicerol va disminuir un 30%, de 19.59 a 4C fins 15.22 mmol g⁻¹ DCW h⁻¹ a la soca 4C_2MXR1*, ja que tot i que no es va detectar glicerol acumulat al medi, l'increment dels nivells de biomassa van modificar la taxa d'assimilació. En canvi, el consum específic de metanol va apuntar a un augment a la soca expressant *MXR1**, de 14.41 fins 17.06 mmol g⁻¹ DCW h⁻¹, comparant 4C amb 4C_2MXR1*, respectivament. Conseqüentment, els nivells detectats de metanol residual van ser un 40% menors a 4C_2MXR1*. No obstant això, encara es va mesurar certa quantitat de metanol al reactor (no observat en condicions anàlogues a les soques referència 0C i 1C), fet que podria indicar la

possibilitat de millores posteriors en el consum de metanol i apuntant a la necessitat de fer un anàlisi més detallat d'aquesta ruta metabòlica. De manera interessant, el canvi de preferència per la font de carboni a les soques podria apuntar a una disminució en les limitacions del metabolisme del metanol, que podrien veure's reflectides a nivell transcripcional.

Taula IV Paràmetres microscòpics de creixement i producció de lipasa a les soques 4C i 4C_2MXR1* en cultius operats en continu. DCW, pes sec; q_s , taxa de consum específic de substrat; $Y_{x/s}$, rendiment biomassa/substrat. Els resultats es mostren com a mitjana \pm sd basat en mesures per triplicat

	DCW (g L ⁻¹)	Biomassa (mmol g ⁻¹ DCW h ⁻¹)	Y_{x/s} (g C-mol ⁻¹)	q_s, glicerol (mmol g ⁻¹ DCW h ⁻¹)	q_s, metanol (mmol g ⁻¹ DCW h ⁻¹)	Activitat lipàsica (UA mL ⁻¹)	Activitat lipàsica (UA g ⁻¹ DCW)
4C	8.64 \pm 0.24	33.71 \pm 1.69	13.66 \pm 0.38	-19.59 \pm 1.96	-14.41 \pm 1.44	40.13 \pm 0.02	4644.33 \pm 129.04
4C_2MXR1*	11.12 \pm 0.11	33.71 \pm 1.69	15.94 \pm 0.16	-15.22 \pm 1.52	-17.06 \pm 1.71	57.57 \pm 1.23	5177.23 \pm 121.89

4.3.4 L'anàlisi transcripcional va confirmar una sobreexpressió dels gens clau de la ruta del metanol a la soca 4C_2MXR1*

Per tal d'avaluar si els canvis fisiològics observats a la soca 4C_2MXR1* en comparació a la soca referència 4C (és a dir, nivells superiors de biomassa, increment en el consum de metanol, millora en la producció de Rol) es corresponien a una alteració (o desregulació) transcripcional de la ruta metabòlica del metanol, es van analitzar diversos gens clau d'aquesta ruta (Figura 3). Concretament, els nivells d'*AOX1*, el gen codificant per l'alcohol oxidasa I (responsable de l'oxidació de metanol a formaldehid i peròxid d'hidrogen, primer pas en la ruta d'assimilació del metanol) van ser 2.3 vegades superiors a la soca 4C_2MXR1* que a 4C. De manera similar, *DAS1* (codificant per la transcetolasa dihidroxiacetona sintasa I, la qual catalitza la oxidació de formaldehid a dihidroxiacetona i gliceraldehid-3-fosfat a la branca assimilatòria) i *FDH1* (codificant per l'enzim format deshidrogenasa, catalitzador del darrer pas de la via dissimilatòria, on l'àcid fòrmic es converteix a CO₂ amb la producció de NADH⁺) van augmentar la seva expressió 3 i 1.9 vegades a la soca 4C_2MXR1*, en concordança amb els resultats observats als cultius en Erlenmeyer i en reactor, on aquesta soca mostrava un increment en la seva capacitat assimilatòria. En relació als nivells transcripcionals de *ROL*, es va mesurar un augment de 5 vegades entre les dues soques. Curiosament, aquest increment en el mRNA no es va veure reflectit en un increment similar de la producció lipàsica a la soca *MXR1**, on la productivitat específica només va ser un 10% major que

a la soca referència (Taula I), suggerint a una limitació traduccional i/o posttraduccional, tal i com s'havia proposat anteriorment (Resina et al., 2007; Resina et al., 2009). Com a marcador típic de la UPR (de l'anglès *Unfolded Protein Response*), es va triar *KAR2* (codificant per una xaperona del reticle endoplasmàtic), i els nivells de mRNA a la soca 4C_MXR1* van mostrar que una sobreexpressió del 50% en comparació a 4C. No obstant, es necessitarien anàlisis addicionals més enllà de l'objectiu d'aquest estudi per tal de confirmar si la maquinària de plegament és el pas limitant en la producció d'una lipasa funcional com a resultat d'un increment als nivells de mRNA de *ROL* a la soca 4C_2MXR1*, ja que l'anàlisi transcriptòmic anteriorment realitzat a soques multicòpia de *ROL* va revelar un impacte baix en les vies de secreció i plegament (Cámara et al., 2016b). D'altra banda, donat que l'anàlisi transcriptòmic previ de la soca 4C havia revelat un fort impacte en la via de regeneració de les pentoses fosfat i la biogènesi dels peroxisomes per la sobreexpressió de *ROL* (Cámara et al., 2016b) els gens codificants per l'aldolasa fructosa 1,6-bifosfat (*FBA1-2*) i la peroxina 6 (*PEX6*) es van seleccionar com a marcadors d'aquestes rutes. En el cas de *FBA1-2* es va detectar un increment del 40% en relació a la soca 4C. Juntament amb la implicació de *DAS1* en aquesta ruta, suggerit recentment per Rußmayer et al. (2015), aquestes dades podrien evidenciar una regulació metabòlica més enllà de la ruta metabòlica d'assimilació directa del metanol. Respecte *PEX6* (codificant per una ATPasa implicada en l'import de proteïnes peroxisomals i crucial per la biogènesi dels peroxisomes, Saffian et al., 2012), no es va detectar una regulació significativa entre les dues soques. Tanmateix, seria interessant analitzar si es donen canvis en l'expressió de *PEX8* i *PEX14*, també codificants per proteïnes implicades la maquinària d'importació i directament regulades per *MXR1*, per tal de descartar completament la no regulació de la biogènesi peroxisomal.

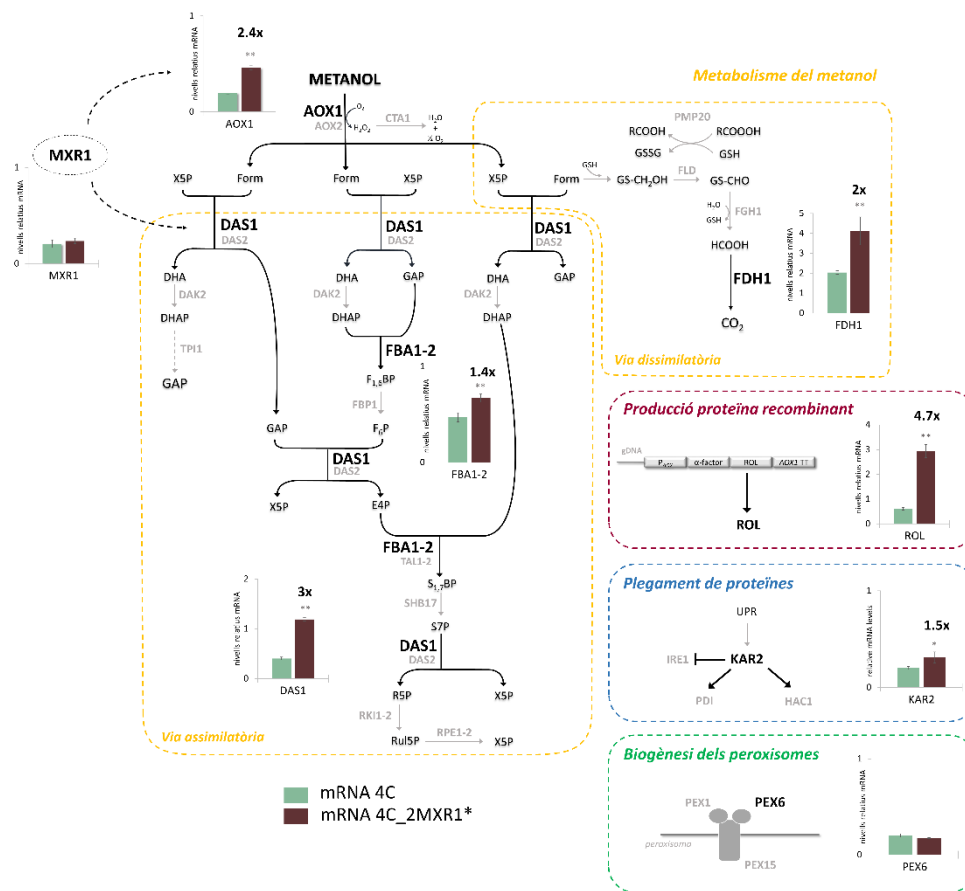


Figura 3. Nivells transcripcionals dels gens seleccionats. Comparativa dels nivells de mRNA de 4C (blau) i la soca 4C_2MXR1* (granat). Els nivells relatius de mRNA es van calcular en comparació amb els nivells d'expressió d'*ACT1* a cada soca. Les línies negres representen les rutes directament controlades pels gens analitzats. DHA, dihidroxiacetona; E4P, eritrosa-4-fosfat; HCOOH, àcid fòrmic; Form, formaldehid; F6P, fructosa-6-fosfat; GAP, gliceraldehid-3-fosfat; S1,7BP, sedoheptulosa-1,7-bifosfat; S7P, sedoheptulosa-7-fosfat; GS-CH₂OH, S-(hidroximetil)glutaciona; X5P, xilulosa-5-fosfat; R5P, ribosa-5-fosfat; Ru15P, ribulosa-5-fosfat. * $P < 0.05$, ** $P < 0.01$, en comparació a la soca control 4C

Sorprenentment, no es van detectar canvis en els nivells de mRNA de *MXR1* entre les soques 4C i 4C_2MXR1*. La autoregulació dels factors de transcripció s'ha revisat de manera extensa en procarïotes i eucariotes (MacPherson et al., 2006; Stekel i Jenkins, 2008), i fins i tot s'ha descrit per Trm1 (Wang et al., 2016), un TF també implicat en la regulació del metabolisme del metanol. La nostra hipòtesi apuntaria a una autoregulació del tipus *feedback loop* per part de Mxr1 per tal de controlar els nivells de molècules disponibles en aquestes condicions de cultiu. Tot i això, la detecció d'una regulació significativa dels gens implicats en el metabolisme del metanol es podrien explicar

pel fet de que tot i que l'eficiència transcripcional per còpia (és a dir, la ràtio entre els nivells de mRNA i el nombre de còpies d'un gen determinat) de *MXR1* va disminuir, la mutació introduïda en una subpoblació de molècules de Mxr1 (per exemple, assegurant una derepressió total de P_{AOX1}) va ser suficient per desencadenar la sobreexpressió dels gens monitoritzats. Tot i que Parua et al. (2012) va demostrar que no hi havia una inhibició dels gens Mut depenent de les proteïnes 14-3-3 en els cultius de *P. pastoris* amb glicerol o glucosa com a font de carboni, es va observar una fosforilació basal de Mxr1 (i.e. promovent la unió de les proteïnes 14-3-3) en les anàlisis publicades, fins i tot en condicions d'inducció total (només amb metanol com a font de carboni). Degut als nivells baixos de Mxr1, aquesta fosforilació basal podria afectar severament els nivells transcripcionals dels gens dependents del metanol a les soques multicòpia. D'aquesta manera, la nostra hipòtesi apuntaria a que la mutació en la posició Ser215 evitaria la unió de les proteïnes 14-3-3 a Mxr1*, provocant així els nivells transcripcionals més elevats observats a les soques *MXR1**, i en conseqüència resultant en una millor assimilació del metanol. Tot plegat explicaria la disminució dels nivells de metanol residual observats a la soca 4C_2MXR1*, tot i que degut a la impossibilitat de sobreexpressar *MXR1**, la taxa d'assimilació del metanol encara seria més baixa que la de la soca control 1C.

En conseqüència, els resultats descrits ens permetrien confirmar el paper principal de Mxr1 en la producció de proteïnes recombinants dirigides per P_{AOX1} . Donada la baixa expressió constitutiva de *MXR1* (Lin-Cereghino et al., 2008; Vogl i Glieder, 2013) un nombre limitat de molècules tindrien múltiples llocs d'unió en diversos gens (*AOX1*, *DAS*, *PEX8*, *ACS1*) (Kranthi et al., 2009; Kranthi et al., 2010; Sahu i Rangarajan, 2015) a les soques productores. En el cas de la soca monocòpia, aquest conjunt de molècules de Mxr1 seria suficient per tal d'aconseguir una assimilació apropiada del metanol, juntament amb una expressió apropiada de la proteïna recombinant (Figura 4A). D'altra banda, en el cas de les soques multicòpia (és a dir, amb un nombre més elevat de seqüència P_{AOX1}) el nombre de molècules disponibles de Mxr1 seria insuficient per tal de fer front a la regulació de tots els gens dependents de Mxr1 i per tant afectaria als nivells transcripcionals corresponents, inclòs el gen heteròleg (Figura 4B). Consegüentment, aquesta conjectura explicaria les tendències de forma *plateau* observades entre les correlacions del nombre de còpies d'un gen i els nivells de mRNA de la proteïna recombinant corresponent observades en d'altres estudis quan s'utilitza P_{AOX1} per l'expressió heteròloga (Liu et al., 2014; Sha et al., 2013; Huang et al., 2014; Yang et al., 2016). D'altra banda, els mecanismes desconeguts de la repressió de P_{AOX1} per part del glicerol permeten mantenir un sistema induïble per la producció de proteïnes recombinants, però amb l'avantatge de l'expressió de *MXR1**. Addicionalment, tal i com va descriure Wang et al. (2016), Mit i Trm1

cooperen amb Mxr1 per la regulació del metabolisme del metanol. Tenint en compte que Trm1 també s'expressa de manera constitutiva a nivells baixos (de manera similar a Mxr1), els efectes de la sobreexpressió de Mxr1* es podrien millorar de manera addicional intentant augmentar els nivells d'expressió de Trm1.

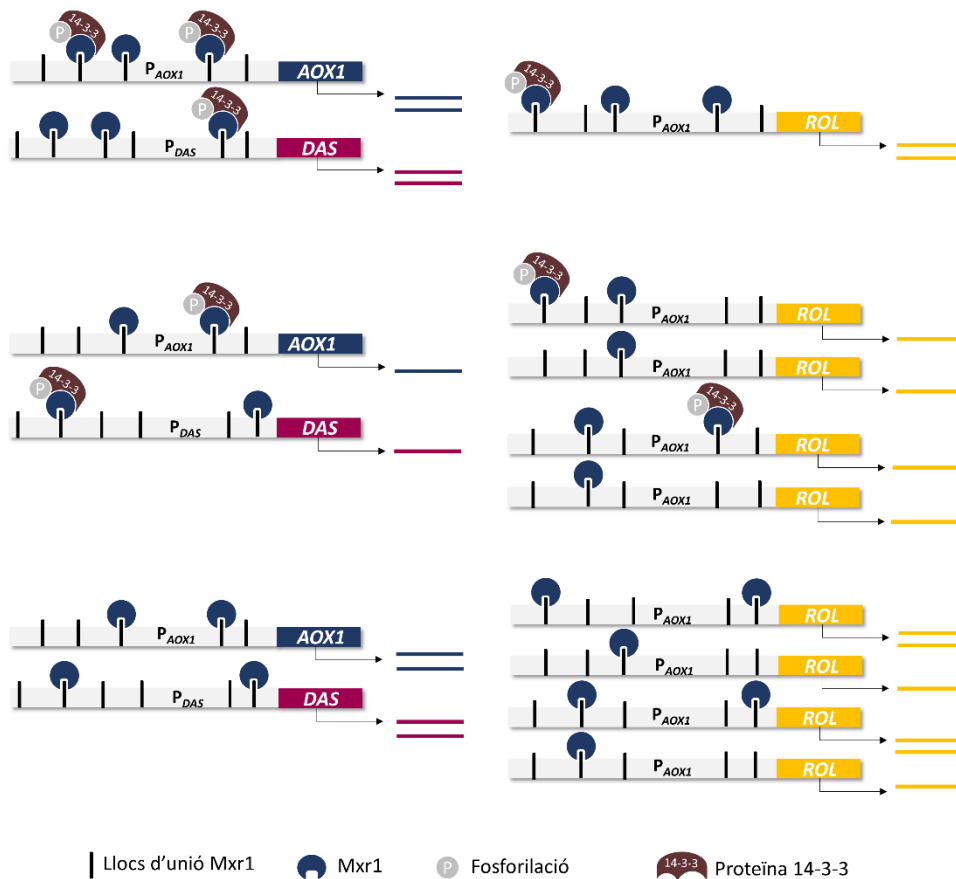


Figura 4. Model de la regulació de P_{AOX} en la producció de proteïna recombinant mediat per Mxr1. Proposem un model on la presència de múltiples llocs d'unió del TF contribueix a nivells transcripcionals més elevats del gen a mesura que s'incrementa el nombre de TF units, degut a una estimulació sinèrgica de l'expressió gènica. A) Producció de la proteïna recombinant a la soca monocòpia. El nombre de molècules disponibles de Mxr1 és limitat, i es distribueixen entre els múltiples llocs d'unió a P_{AOX1} i P_{DAS} , entre d'altres promotors regulats per Mxr1. B) Producció de la proteïna recombinant a la soca 4C. A mesura que el nombre de llocs d'unió de Mxr1 augmenta, els nivells de transcripció dels gens endògens controlats per Mxr1 disminueix. C) Producció de la proteïna recombinant a la soca 4C expressant Mxr1*. L'intercanvi a la posició Ser215 per una alanina prevén la unió de les proteïnes 14-3-3. Consegüentment, els nivells transcripcionals obtinguts pels gens relacionats amb el metanol i ROL són més elevats. Nombre de llocs d'unió a Mxr1 segons Kranthi et al. (2009) i Kranthi et al. (2010).

4.4 CONCLUSIONS

Per tal d'aconseguir millors hostes per la producció de proteïna recombinant, és necessari conèixer els mecanismes de regulació que dirigeixen l'expressió dels gens heteròlegs. Recentment, el coneixement de les bases moleculars dels sistemes basats en P_{AOX1} s'ha incrementat, però pocs s'han centrat en la seva optimització per la producció heteròloga.

En aquest estudi descrivim l'impacte beneficiós d'expressar $MXR1^*$, una versió desregulada de $MXR1$, en una soca de *P. pastoris* amb quatre còpies de ROL . Remarcablement, es va observar una millora en el creixement, la producció de Rol i l'assimilació de metanol en cultius en Erlenmeyer i en continu quan es comparava la soca $MXR1^*$ i la soca multicòpia referència. D'altra banda, l'anàlisi transcripcional dels gens relacionats amb el metabolisme del metanol va revelar una regulació global positiva, corroborant els canvis fisiològics observats a la soca. Finalment, s'ha proposat un model de regulació per P_{AOX1} mediat per Mxr1 a les soques multicòpia, proporcionant una nova aproximació per tal d'entendre els colls d'ampolla transcripcionals observats en els models basats en metanol.

Modificacions futures d'aquestes soques, sobreexpresant $MXR1^*$ i $TRM1$, haurien de comportar millores en la producció heteròloga dirigida per P_{AOX1} . Així mateix, l'anàlisi transcriptòmic d'aquestes noves soques podria aclarir l'impacte global d'aquestes modificacions a *P. pastoris*.

4.5 REFERÈNCIES

- Adelantado, N. 2016. Lipidomics studies of recombinant *Pichia pastoris* for improved recombinant protein secretion through cell engineering. Tesis doctoral, Universitat Autònoma de Barcelona. Obtingut de <http://www.tesisenred.net/handle/10803/384229>
- Ahmad M, Hirz M, Pichler H, Schwab H. 2014. Protein expression in *Pichia pastoris*: Recent achievements and perspectives for heterologous protein production. *Appl. Microbiol. Biotechnol.* 98:5301–5317.
- Bill RM. 2015. Recombinant protein subunit vaccine synthesis in microbes: A role for yeast? *J. Pharm. Pharmacol.* 67:319–328.
- Byrne B. 2015. *Pichia pastoris* as an expression host for membrane protein structural biology. *Curr. Opin. Struct. Biol.* 32:9–17.
- Cámara E, Albiol J, Ferrer P. 2016a. Droplet digital PCR-aided screening and characterization of *Pichia pastoris* multiple gene copy strains. *Biotechnol. Bioeng.* 113:1542–1551.
- Cámara E, Landes N, Albiol J, Gasser B, Mattanovich D, Ferrer P. Increased dosage of AOX1 promoter-regulated expression cassettes leads to transcription attenuation of the methanol metabolism in *Pichia pastoris*. Manuscript under submission of Scientific Reports.
- Corchero JL, Gasser B, Resina D, Smith W, Parrilli E, Vázquez F, Abasolo I, Giuliani M, Jäntti J, Ferrer P, Saloheimo M, Mattanovich D, Schwartz S, Tutino ML, Villaverde A. 2013. Unconventional microbial systems for the cost-efficient production of high-quality protein therapeutics. *Biotechnol. Adv.* 31:140–153.
- Cregg J M. 2007. DNA-mediated transformation. *Methods Mol. Biol.* Vol. 389 27-42 p.
- Cregg JM, Cereghino JL, Shi J, Higgins DR. 2000. Recombinant Protein Expression in *Pichia pastoris*. *Mol. Biotechnol.* 16:23–52.
- De Schutter K, Lin Y-C, Tiels P, Van Hecke A, Glinka S, Weber-Lehmann J, Rouzé P, Van de Peer Y, Callewaert N. 2009. Genome sequence of the recombinant protein production host *Pichia pastoris*. *Nat. Biotechnol.* 27:561–566.
- Ferrer P, Albiol J. 2014. ¹³C-based metabolic flux analysis of recombinant *Pichia Pastoris*. *Methods in Molecular Biology. Methods Mol Biol.* 1191:291–313.

Fu H, Subramanian RR, Masters SC. 2000. 14-3-3 proteins: Structure, Function, and Regulation. *Annu. Rev. Pharmacol. Toxicol.* 40:617–647.

Gasser B, Prielhofer R, Marx H, Maurer M, Nocon J, Steiger M, Puxbaum V, Sauer M, Mattanovich D. 2013. *Pichia pastoris*: protein production host and model organism for biomedical research. *Future Microbiol.* 8:191–208.

Huang J, Xia J, Yang Z, Guan F, Cui D, Guan G, Jiang W, Li Y. 2014. Improved production of a recombinant *Rhizomucor miehei* lipase expressed in *Pichia pastoris* and its application for conversion of microalgae oil to biodiesel. *Biotechnol. Biofuels* 7:111.

Isidro IA, Portela RM, Clemente JJ, Cunha AE, Oliveira R. 2016. Hybrid metabolic flux analysis and recombinant protein prediction in *Pichia pastoris* X-33 cultures expressing a single-chain antibody fragment. *Bioprocess Biosyst. Eng.*:1–13.

Jordà J, De Jesus SS, Peltier S, Ferrer P, Albiol J. 2014a. Metabolic flux analysis of recombinant *Pichia pastoris* growing on different glycerol/methanol mixtures by iterative fitting of NMR-derived ¹³C-labelling data from proteinogenic amino acids. *N. Biotechnol.* 31:120–132.

Jordà J, Rojas HC, Carnicer M, Wahl A, Ferrer P, Albiol J. 2014b. Quantitative Metabolomics and Instationary ¹³C-Metabolic Flux Analysis Reveals Impact of Recombinant Protein Production on Trehalose and Energy Metabolism in *Pichia pastoris*. *Metabolites* 4:281–299.

Jordà J, Jouhten P, Cámara E, Maaheimo H, Albiol J, Ferrer P. 2012. Metabolic flux profiling of recombinant protein secreting *Pichia pastoris* growing on glucose:methanol mixtures. *Microb. Cell Fact.* 11:57.

Kranthi BV, Kumar HR, Rangarajan PN. 2010. Identification of Mxr1p-binding sites in the promoters of genes encoding dihydroxyacetone synthase and peroxin 8 of the methylotrophic yeast *Pichia pastoris*. *Yeast* 27:705–711.

Kranthi BV, Kumar R, Kumar NV, Rao DN, Rangarajan PN. 2009. Identification of key DNA elements involved in promoter recognition by Mxr1p, a master regulator of methanol utilization pathway in *Pichia pastoris*. *Biochim. Biophys. Acta - Gene Regul. Mech.* 1789:460–468.

- Kumar NV, Rangarajan PN. 2012. The zinc finger proteins Mxr1p and repressor of phosphoenolpyruvate carboxykinase (ROP) have the same DNA binding specificity but regulate methanol metabolism antagonistically in *Pichia pastoris*. *J. Biol. Chem.* 287:34465–34473.
- Lin-Cereghino GP, Godfrey L, de la Cruz BJ, Johnson S, Khuongsathiene S, Tolstorukov I, Yan M, Lin-Cereghino J, Veenhuis M, Subramani S, Cregg JM. 2006. Mxr1p, a key regulator of the methanol utilization pathway and peroxisomal genes in *Pichia pastoris*. *Mol. Cell. Biol.* 26:883–897.
- Liu H, Qin Y, Huang Y, Chen Y, Cong P, He Z. 2014. Direct Evaluation of the Effect of Gene Dosage on Secretion of Protein from Yeast *Pichia pastoris* by Expressing EGFP. *J. Microbiol. Biotechnol.* 24:144–151.
- Looser V, Brühlmann 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
- MacPherson S, Laroche M, Turcotte B. 2006. A fungal family of transcriptional regulators: the zinc cluster proteins. *Microbiol. Mol. Biol. Rev.* 70:583–604.
- Mattanovich D, Graf A, Stadlmann J, Dragosits M, Redl A, Maurer M, Kleinheinz M, Sauer M, Altmann F, Gasser B. 2009. Genome, secretome and glucose transport highlight unique features of the protein production host *Pichia pastoris*. *Microb. Cell Fact.* 8:29.
- Parua PK, Ryan PM, Trang K, Young ET. 2012. *Pichia pastoris* 14-3-3 regulates transcriptional activity of the methanol inducible transcription factor Mxr1 by direct interaction. *Mol. Microbiol.* 85:282–298.
- Rabert C, Weinacker D, Pessoa A, Farías JG. 2013. Recombinant proteins for industrial uses: Utilization of *Pichia pastoris* expression system. *Brazilian J. Microbiol.* 44:351–356.
- Resina D, Bollók M, Khatri NK, Valero F, Neubauer P, Ferrer P. 2007. Transcriptional response of *P. pastoris* in fed-batch cultivations to *Rhizopus oryzae* lipase production reveals UPR induction. *Microb. Cell Fact.* 6:21.
- Resina D, Maurer M, Cos O, Arnau C, Carnicer M, Marx H, Gasser B, Valero F, Mattanovich D, Ferrer P. 2009. Engineering of bottlenecks in *Rhizopus oryzae* lipase production in *Pichia pastoris* using the nitrogen source-regulated FLD1 promoter. *N. Biotechnol.* 25:396–403.

- Resina D, Serrano A, Valero F, Ferrer P. 2004. Expression of a *Rhizopus oryzae* lipase in *Pichia pastoris* under control of the nitrogen source-regulated formaldehyde dehydrogenase promoter. *J. Biotechnol.* 109:103–113.
- Rußmayer H, Buchetics M, Gruber C, Valli M, Grillitsch K, Modarres G, Guerrasio R, Klavins K, Neubauer S, Drexler H, Steiger M, Troyer C, Al Chalabi A, Krebiehl G, Sonntag D, Zellnig G, Daum G, Graf AB, Altmann F, Koellensperger G, Hann S, Sauer M, Mattanovich D, Gasser B. 2015. Systems-level organization of yeast methylotrophic lifestyle. *BMC Biol.* 13:80.
- Saffian D, Grimm I, Girzalsky W, Erdmann R. 2012. ATP-dependent assembly of the heteromeric Pex1p-Pex6p-complex of the peroxisomal matrix protein import machinery. *J. Struct. Biol.* 179:126–132.
- Sahu U, Krishna Rao K, Rangarajan PN. 2014. Trm1p, a Zn(II)2Cys6-type transcription factor, is essential for the transcriptional activation of genes of methanol utilization pathway, in *Pichia pastoris*. *Biochem. Biophys. Res. Commun.* 451:158–164.
- Sha C, Yu XW, Lin NX, Zhang M, Xu Y. 2013. Enhancement of lipase r27RCL production in *Pichia pastoris* by regulating gene dosage and co-expression with chaperone protein disulfide isomerase. *Enzyme Microb. Technol.* 53:438–443
- Spohner SC, Müller H, Quitmann H, Czermak P. 2015. Expression of enzymes for the usage in food and feed industry with *Pichia pastoris*. *J. Biotechnol.* 202:118–134.
- Stekel DJ, Jenkins DJ. 2008. Strong negative self regulation of prokaryotic transcription factors increases the intrinsic noise of protein expression. *BMC Syst. Biol.* 2:6.
- Takagi S, Tsutsumi N, Terui Y, Kong XY. 2009. Method for methanol independent induction from methanol inducible promoters in *Pichia*. Patent EUA 8,236,528 B2, 7 agost, 2012.
- Tschopp JF, Brust PF, Cregg JM, Stillman CA, Gingeras TR. 1987. Expression of the lacZ gene from two methanol-regulated promoters in *Pichia pastoris*. *Nucleic Acids Res.* 15:3859–3876.
- Vogl T, Glieder A. 2013. Regulation of *Pichia pastoris* promoters and its consequences for protein production. *N. Biotechnol.* 30:385–404.
- Vogl T, Hartner FS, Glieder A. 2013. New opportunities by synthetic biology for biopharmaceutical production in *Pichia pastoris*. *Curr. Opin. Biotechnol.* 24:1094–1101.

Wang X, Wang Q, Wang J, Bai P, Shi L, Shen W, Zhou M, Zhou X, Zhang Y, Cai M. 2016. Mit1 Transcription Factor Mediates Methanol Signaling and Regulates Alcohol Oxidase 1 Promoter in *Pichia pastoris*. *J. Biol. Chem.*: 291:6245-61

Yang H, Zhai C, Yu X, Li Z, Tang W, Liu Y, Ma X, Zhong X, Li G, Wu D, Ma L. 2016. High-level expression of Proteinase K from *Tritirachium album Limber* in *Pichia pastoris* using multi-copy expression strains. *Protein Expr. Purif.* 122:38–44.

5. Anàlisi transcriptòmic comparatiu de la resposta fisiològica de *P. pastoris* a l'expressió de tres proteïnes recombinants diferents

5.1 INTRODUCCIÓ

A les darreres dècades, la irrupció de les tecnologies òmiques ha representat un pas endavant en el coneixement de la base genètica, molecular i metabòlica d'innombrables microorganismes (Westerhoff i Palsson, 2004). Així mateix, aquestes noves aproximacions (que consideren un sistema complex com un tot) utilitzen les dades disponibles per tal de generar noves hipòtesis que poden ser posteriorment verificades, a diferència del reduccionisme de la biologia molecular clàssica (Kell i Oliver, 2004). En aquest context, la transcriptòmica és una eina potent que pot aportar elevades quantitats de dades d'expressió gènica, i les millores en les plataformes de RNA-seq i la tecnologia dels microarrays s'han incrementat de manera dràstica als darrers anys (Wang et al., 2010; Dong i Chen, 2013; Ventimiglia i Petralia, 2013; Spies i Ciaudo, 2015). En el cas de *P. pastoris*, tot i que ja s'havien publicat alguns estudis transcriptòmics (Sauer et al., 2004; Gasser et al., 2007), la publicació de la seqüenciació i anotació del genoma de *P. pastoris* (De Schutter et al., 2009; Mattanovich et al., 2009) va ser el punt d'inflexió per tal d'estendre l'ús de les tecnologies òmiques en aquest sistema d'expressió. Fins el dia d'avui, s'han publicat nombrosos estudis centrats en l'efecte de diferents estrès ambientals (Dragosits et al., 2010), la resposta a proteïnes mal plegades (Hesketh et al., 2013), l'efecte de la disponibilitat d'oxigen (Baumann et al., 2010), l'impacte de la taxa de creixement (Rebnegger et al., 2014) o la dosi gènica (Cámara et al., 2016b), la resposta a diferents fonts de carboni (Liang et al., 2012; Prielhofer et al., 2015; Rußmayer et al., 2015) o anàlisis comparatius entre diferents soques o microorganismes (Baumann et al., 2011; Love et al., 2016). No obstant, mentre que diversos estudis comparatius sobre l'efecte de l'expressió de proteïna recombinant s'havien publicat per diversos llevats (Liu et al., 2014; Mojzita et al., 2014; Sims et al., 2005), encara no s'havia publicat cap centrat en *P. pastoris*.

En el present estudi, es comparen diferents conjunts de dades transcriptòmics obtingudes dels cultius operats en continu d'un grup de soques de *P. pastoris* produint tres proteïnes model secretades: l'albumina del sèrum humà (Hsa) i la cadena simple 3d6 de l'anticòs Fv-Fc anti-VIH-1 (3d6), ambdues expressades sota el control del promotor de la gliceraldehid-3-fosfat deshidrogenasa (P_{GAP}), i una lipasa de *Rhizopus oryzae* (Rol), expressada sota el control del promotor de l'alcohol oxidasa 1 (P_{AOX1}). L'estudi comparatiu es va estendre a soques amb múltiples còpies dels cassets d'expressió de cada proteïna model. Aquestes proteïnes representen diferents nivells de complexitat estructural, per tal de poder valorar la resposta fisiològica a la complexitat proteica i els corresponents nivells d'expressió.

5.2 RESULTATS I DISCUSSIÓ

5.2.1 Caracterització fisiològica de les soques productores

Les soques de *P. pastoris* expressant diferents proteïnes recombinants amb una dosi gènica diferent es van fer créixer en cultius operats en continu (amb glucosa com a font de carboni en el cas de les soques productores de Hsa i 3d6, i amb una barreja de metanol:glicerol a les soques productores de Rol). El nombre de còpies, els nivells de biomassa, els rendiments de producció i les taxes de consum es mostren a la Taula I.

En el cas de les soques Hsa i 3d6, la ràtio entre les soques amb un nombre baix i alt de cassets d'expressió era de 3, mentre que la ràtio per les soques Rol era de 4. En relació a la producció de proteïna recombinant, la soca més productora va ser Hsa_H (productora d'albumina amb alta dosi gènica), seguida de la Rol_H (productora de lipasa amb alta dosi gènica). Tot i que l'albumina de sèrum humà es considera una proteïna estructuralment més complexa que la lipasa degut a les seves modificacions post-traduccionals i la mida molecular, s'ha descrit àmpliament la gran facilitat per produir alts nivells d'Hsa (Kobayashi et al., 2000; Ohya et al., 2005), mentre que en el cas de soques sobreexpressant *ROL* s'han reportat limitacions transcripcionals i traduccionals (Cámara et al., 2016b). La soca amb una productivitat més baixa va ser 3d6_H (soca productora de 3d6 amb alta dosi gènica), en consonància amb l'elevada complexitat estructural de l'anticòs 3d6. La ràtio de productivitat entre les soques d'alta i baixa dosi gènica va ser de 6 per les soques productores d'Hsa i Rol, mentre que per les soques 3d6 va ser de 3. Remarcablement, tot i que es va observar un fort impacte en el creixement de la soca Rol_H (Taula I), no es va observar cap canvi significatiu entre les soques de baixa o alta dosi gènica crescudes en glucosa. De manera consistent, no es van mesurar canvis en les taxes de consum específic de glucosa, mentre que es va observar un augment significatiu en el consum específic de glicerol a les soques multicòpia de Rol, acompanyat d'una caiguda en l'assimilació del metanol (Cámara et al., 2016a i 2016b).

Taula II. Paràmetres fisiològics de les diferents soques productores. Hsa, soques productores de l'albumina de sèrum humà; 3d6, soques productores de la cadena simple 3d6 de l'anticòs Fv-Fc anti-VIH-1; Rol, soques productores d'una lipasa de *Rhizopus oryzae*; H/L, baixa o alta dosi gènica, respectivament.

Soca	Nº de còpies	Biomassa (g L ⁻¹)	Q _p (µg g ⁻¹ h ⁻¹)	Q _{s, glucosa} (mmol g ⁻¹ h ⁻¹)	Q _{s, glicerol} (mmol g ⁻¹ h ⁻¹)	Q _{s, metanol} (mmol g ⁻¹ h ⁻¹)
No productora	-	25.85 ± 0.19	-	0.81 ± 0.01	-	-
Hsa_L	4	26.60 ± 0.79	8.4 ± 0.5	0.78 ± 0.02	-	-
Hsa_H	11	26.69 ± 0.07	49.4 ± 2.9	0.78 ± 0.00	-	-
3d6_L	4	27.16 ± 1.33	0.9 ± 0.2	0.77 ± 0.04	-	-
3d6_H	10	26.84 ± 0.39	2.8 ± 0.3	0.78 ± 0.01	-	-
No productora	-	12.82 ± 0.16	-	-	1.12 ± 0.01	2.01 ± 0.11
Rol_L	1	12.03 ± 0.11	2.4 ± 1.3	-	1.20 ± 0.03	2.22 ± 0.10
Rol_H	4	8.66 ± 0.41	14.7 ± 2.9	-	1.71 ± 0.09	0.62 ± 0.21

5.2.2 Anàlisi transcripcional global

Per tal d'identificar els gens expressats de manera diferencial, els gens regulats es van seleccionar seguint un criteri de $-0.58 > \text{Log}_2\text{FC} > 0.58$ (equivalent a un *fold change* de 1.5) i un *p-val* < 0.05 en comparació a la soca referència utilitzada. Els nivells d'expressió de 958 gens diferents van canviar de manera significativa en alguna de les soques en comparació a la soca no productora. En el cas de les soques productores d'Hsa, el nombre de gens sobreexpressats va ser igual per les soques amb baixa i alta dosi gènica, però el nombre de gens amb expressió negativa va ser major per la soca Hsa_H (Taula II), mentre que el nombre de gens regulats va ser tres vegades més alt per les soques Rol_H que per la soca Rol amb baixa dosi gènica. En ambdòs casos, la ràtio de gens amb regulació positiva/negativa va ser gairebé igual per la soca productora d'Hsa i Rol. Respecte les soques 3d6, 3d6_L va presentar una regulació mínima, mentre que el nombre de gens regulats a la soca 3d6_H va ser significativament superior i similar a les altres soques productores amb alta dosi gènica. Aquest fet podria estar relacionat a la baixa qualitat dels microarrays de la soca 3d6_L, resultant en gens amb *p-val* elevats que no passaven el criteri establert.

Taula III. Resum dels gens regulats respecte la soca no productora.

	Hsa_L	Hsa_H	3d6_L	3d6_H	Rol_L	Rol_H
REGULACIÓ +	168	168	15	79	70	197
REGULACIÓ -	97	140	4	156	39	162
% total	4.95	5.75	0.35	4.39	2.04	6.71
% regulació +	63.4	54.5	78.9	33.6	64.2	54.9
% regulació -	36.6	45.5	21.1	66.4	35.8	45.1

A la figura 1A es pot observar la distribució dels nivells d'expressió gènica. Pel que fa a la població sobreexpressada, els nivells d'expressió de les soques Hsa i 3d6 estaven limitats a un Log2FC màxim de 2, mentre que les soques productores de Rol presentaven nivells més elevats, amb un Log2FC de 3 en diversos gens. L'anàlisi dels gens més regulats a cada soca (Figura 1B) va revelar una diversitat notable. Malauradament, només 34 dels 60 gens estaven anotats funcionalment. *MUP1*, *LYP1* i *HIP1*, codificants per permeases d'aminoàcids, es trobaven entre els gens més regulats a les soques Hsa i 3d6, juntament amb diversos gens relacionats amb els ribosomes i el metabolisme del RNA. De manera interessant, diferents gens codificants per proteïnes reguladores per feromones (*KAR4*, *PRM1*, *STE3*) i proteïnes relacionades amb la floculació (codificades per *FLO100*, *FLO10*, *FLO5-1* i *FLO11*) es trobaven regulats de manera negativa a totes les soques productores.

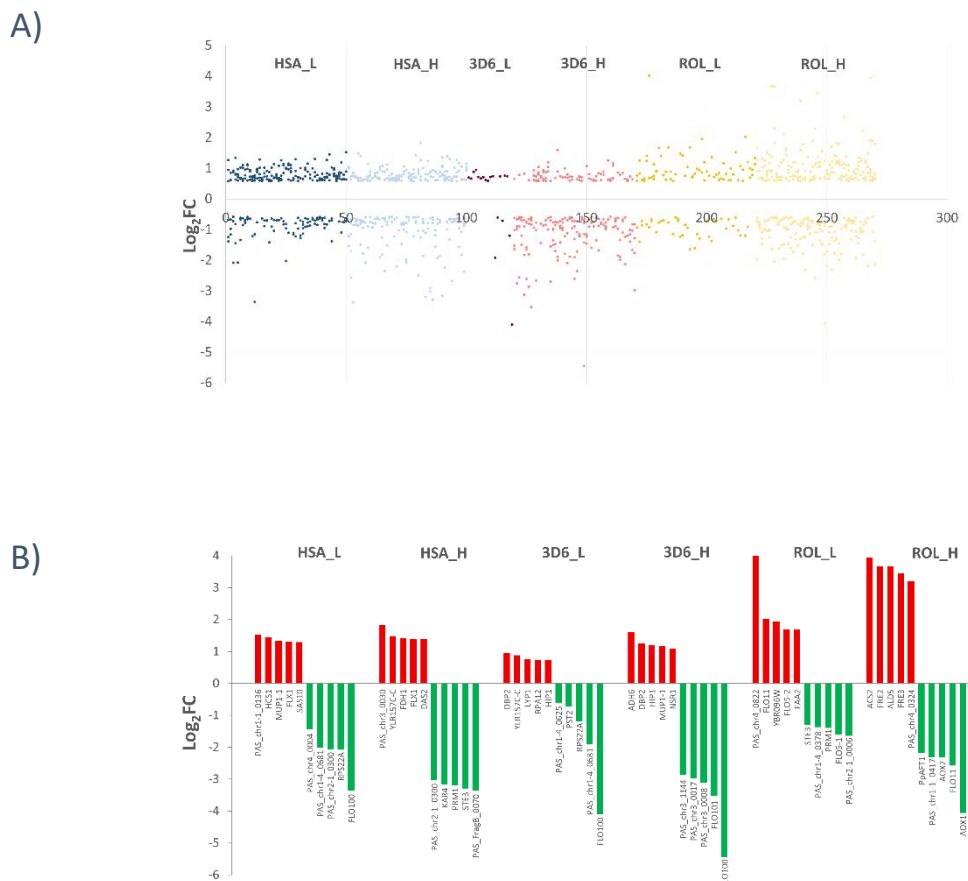


Figura 1. Anàlisi global dels gens regulats en comparació a la soca no productora. A) Distribució dels gens regulats B) Top5 dels gens regulats

5.2.3 La font de carboni té un impacte major en l'expressió gènica que els nivells de producció recombinant

Amb l'objectiu d'avaluar els patrons d'expressió, es va realitzar un anàlisi de components principals (PCA, de l'anglès *Principal Component Analysis*) dels nivells transcripcionals de les diferents soques (Figura 2). Per aquest anàlisi, cada soca productora es va tractar com una variable. El primer i el segon component principal (PC1 i PC2) explicaven conjuntament el 80% de la variabilitat total (54.7 i 25.9%, respectivament), i incloent el tercer component (PC3) representaven aproximadament el 90% (Figura 2A).

Es va observar una elevada correlació entre PC1 i les soques Hsa i 3d6, mentre que les soques productores de Rol estaven estretament associades a PC2, senyalant una elevada independència entre Hsa/3d6 i les soques Rol (Figura 2B, esquerra). Això podria suggerir la importància de la font de carboni per tal d'explicar la variabilitat de l'expressió gènica, sent així el factor discriminadori principal per separar clarament les soques depenent del creixement amb glucosa o en glicerol:metanol. De manera consistent, els gens més regulats descrits anteriorment (Figura 1B) van ser els majors contribuents a la tendència dels components principals. Remarcablement, no es va trobar cap separació depenent dels nivells d'expressió recombinant en la representació de PC1 vs PC2. Per tal d'analitzar l'impacte de l'expressió heteròloga es van representar els components 2 i 3 (Figura 2B, dreta). PC3 va permetre agrupar les soques Hsa_L, Hsa_H i Rol_L en una banda del component, mentre que Rol_H, 3d6_L i 3d6_H van quedar localitzades a la banda oposada. El tercer component podria representar l'impacte dels nivells de proteïna recombinant al perfil transcriptòmic de les soques. En efecte, les soques Hsa i Rol_L van ser les soques amb una major facilitat a l'hora de produir la proteïna corresponent (degut a la poca complexitat proteica o els nivells baixos de producció), mentre que la producció recombinant a les soques 3d6 i Rol_H seria més crítica.

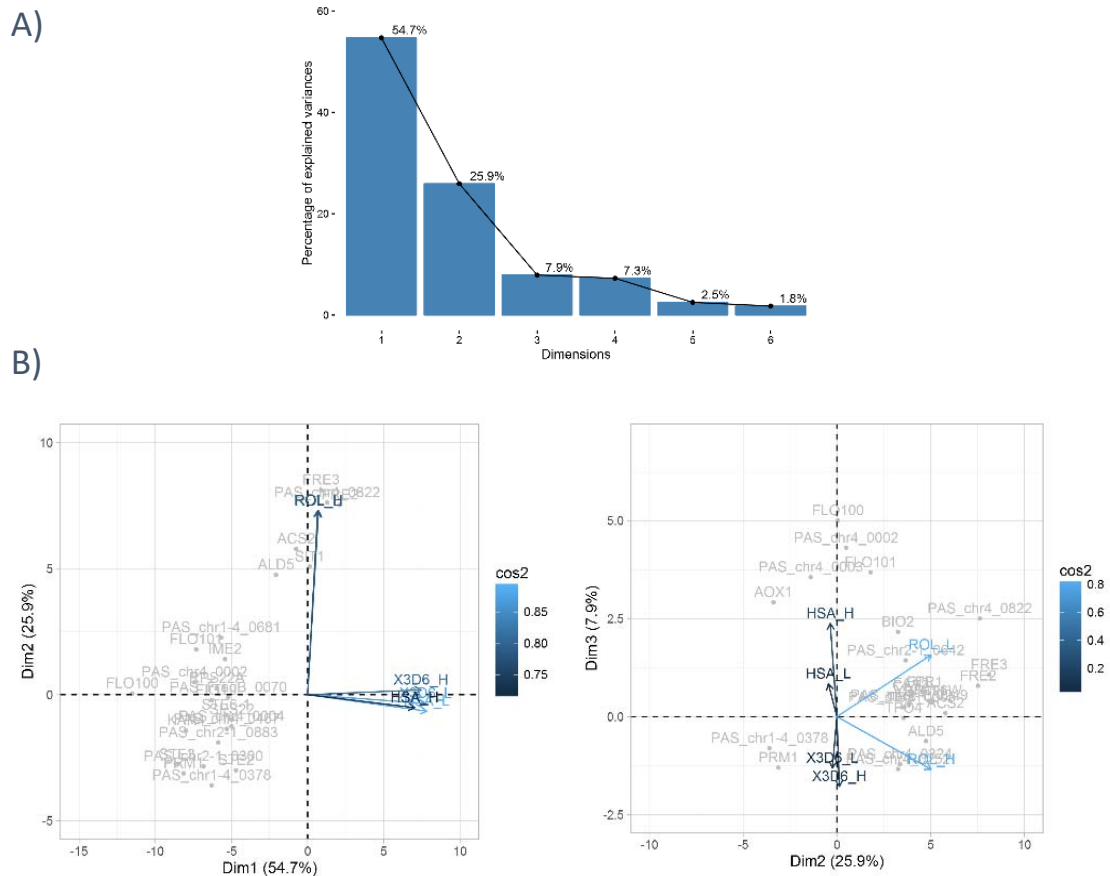


Figura 2. Anàlisi de l'expressió gènica mitjançant l'Anàlisi Principal de Components. A) Distribució de la contribució de cada component B) Anàlisi dels components PC1 vs PC2 (esquerra) i PC2 vs PC3 (dreta). El gradient de color es va establir en funció del valor \cos^2 (representatiu de la qualitat de les dades) i només es van representar al gràfic els 50 gens més contribuents. Valors elevats de \cos^2 asseguren una representació acurada de les variables i la significança de l'anàlisi.

D'una altra banda, la matriu de covariància va confirmar l'elevada correspondència observada al PCA entre les soques Hsa i 3d6, i la independència entre les soques Rol i les soques crescudes en glucosa (Figura 3A). La matriu també va revelar que l'increment de la dosi gènica tenia un impacte major en el transcriptoma de les soques Rol, mentre que les soques amb baixa i alta dosi gènica presentaven una semblança més gran en el cas de les soques Hsa i 3d6. Aquest fet podria indicar un impacte major de la dosi gènica en l'expressió recombinant dirigida per P_{AOX1} en comparació amb el sistema P_{GAP} .

Per tal d'aprofundir en la correlació entre les soques amb baixa i alta dosi gènica, es van representar tots els gens regulats en alguna de les soques (descartant aquells sense anotació funcional) en cada parell de soques productores d'una mateixa proteïna heteròloga. Amb l'objectiu de visualitzar clarament els principals processos implicats, els gens es van agrupar en 23 categories generals. Tal

i com es pot observar a la Figura 3B, mentre que es van mesurar nivells similars d'expressió gènica a les soques 3d6_L i 3d6_H, les soques productores d'Hsa i Rol van mostrar nivells significativament diferents entre les soques amb baixa i alta dosi gènica en diversos gens codificants per floculines. Addicionalment, es va detectar que les soques Rol estaven fortament desaparellades en alguns gens relacionats amb el metanol, en consonància amb l'alteració del metabolisme del metanol reportat prèviament a les soques multicòpia (Cámara et al., 2016a, 2016b, 2016c).

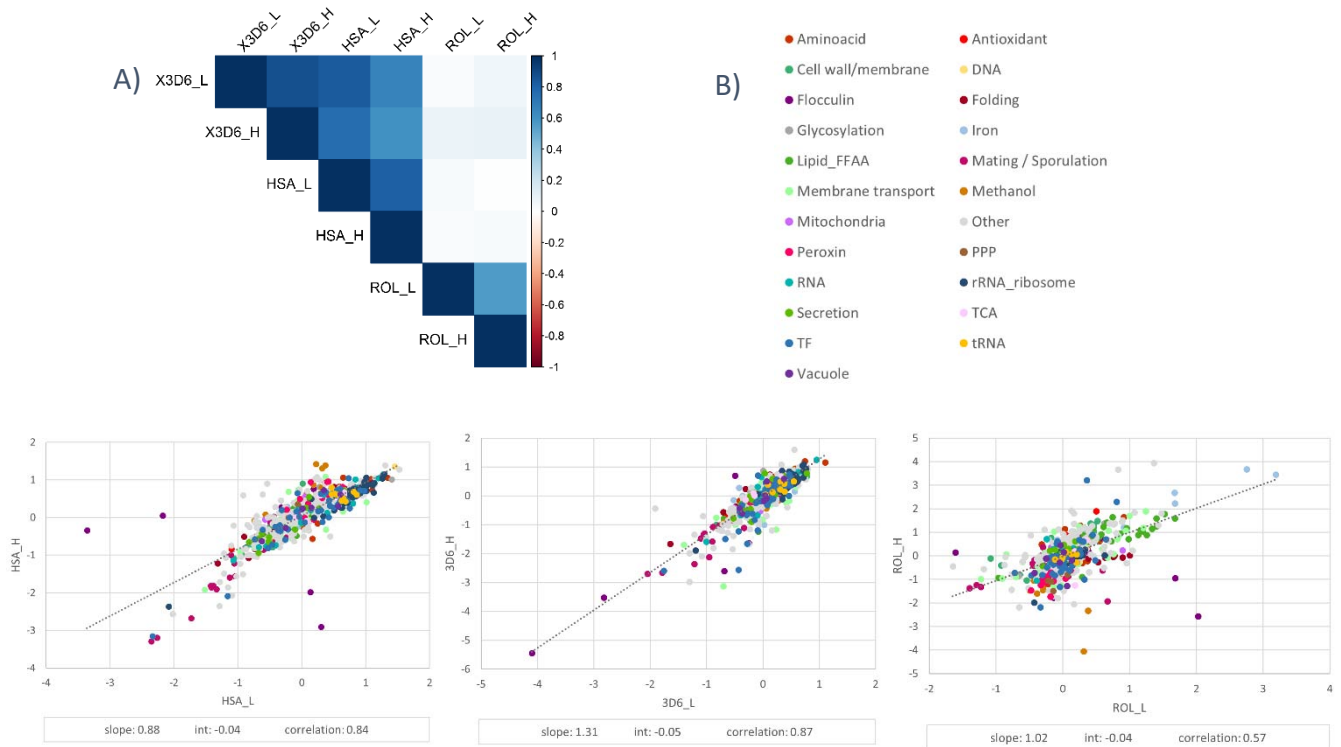


Figure 3. Correlació entre les diferents soques productores. A) Correlograma. Les correlacions positives es mostren en blau, mentre que les negatives ho fan en vermell. La intensitat del requadre és proporcional als coeficients de correlació. B) Correlació de l'expressió dels gens regulats (en comparació amb la soca no productora) entre les soques amb baixa i alta dosi gènica a les soques productores d'Hsa, 3d6 i Rol.

5.2.4 Anàlisi de *heatmaps*

Mitjançant l'anàlisi de *heatmaps* es van agrupar els diferents gens segons la seva regulació i la seva categoria funcional. Considerant només els gens amb alguna anotació funcional, la figura 4A mostra la regulació mitjana de les famílies funcionals a cada soca (la regulació individual de cada gen es mostra a la figura 4B). D'acord amb els resultats comentats anteriorment, diversos processos es

trobaven regulats de manera similar a les soques Hsa i 3d6. Concretament, les rutes metabòliques relacionades amb els aminoàcids, glicosilació, RNA, ribosomes i el cicle dels àcids tricarboxílics (TCA, de l'anglès *Tricarboxylic Acid Cycle*) van presentar una sobreexpressió en comparació amb la soca control no productora, mentre que els processos antioxidants i altres implicats en la paret cel·lular van resultar regulats negativament. D'altra banda, a les soques productores de Rol les rutes metabòliques relacionades amb el ferro i els àcids grassos van aparèixer sobreexpressats, mentre que els processos del metanol i les peroxines estaven regulats negativament (Cámara et al., 2016b). Destacablement, es va detectar una regulació global negativa a totes les soques productores en gens implicats en l'esperulació i la reproducció sexual, juntament amb les flocculines (amb l'excepció de la soca Rol_L).

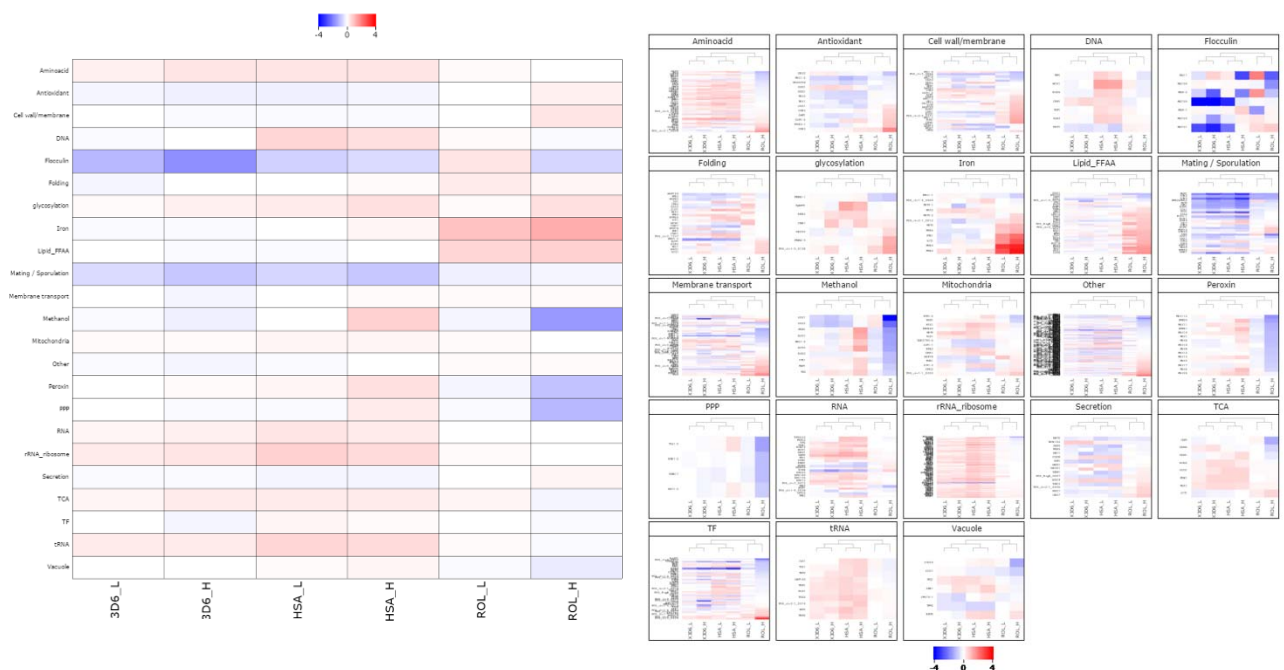


Figure 4. Anàlisi de *heatmap* dels principals processos regulats

5.2.5 La producció de proteïna recombinant disminueix les vies de reproducció sexual i afavoreix la biogènesi ribosomal i la síntesi d'aminoàcids a les soques P_{GAP}

Per tal d'estudiar detalladament els canvis d'expressió derivats de la producció de proteïna recombinant, es va triar la comparativa de les soques amb alta dosi gènica respecte les soques no productores. Primer, els gens amb tendències similars es van agrupar mitjançant clústers utilitzant l'algoritme k-means (k=12). Segons el perfil de regulació obtingut, els clústers es van classificar en 5 categories: I) Regulació específica de P_{GAP}, II) Regulació específica de P_{AOX1}, III) Regulació específica

de la producció d'Hsa, IV) Regulació depenent de la complexitat proteica i V) Regulació depenent dels nivells de producció.

En relació als gens regulats únicament a les soques amb el sistema d'expressió P_{GAP} (Figura 5A), es van trobar 3 clústers amb una marcada regulació negativa (en comparació amb la soca no productora) mentre que només un clúster va mostrar una lleugera regulació positiva a les soques Hsa_H i 3d6_H.

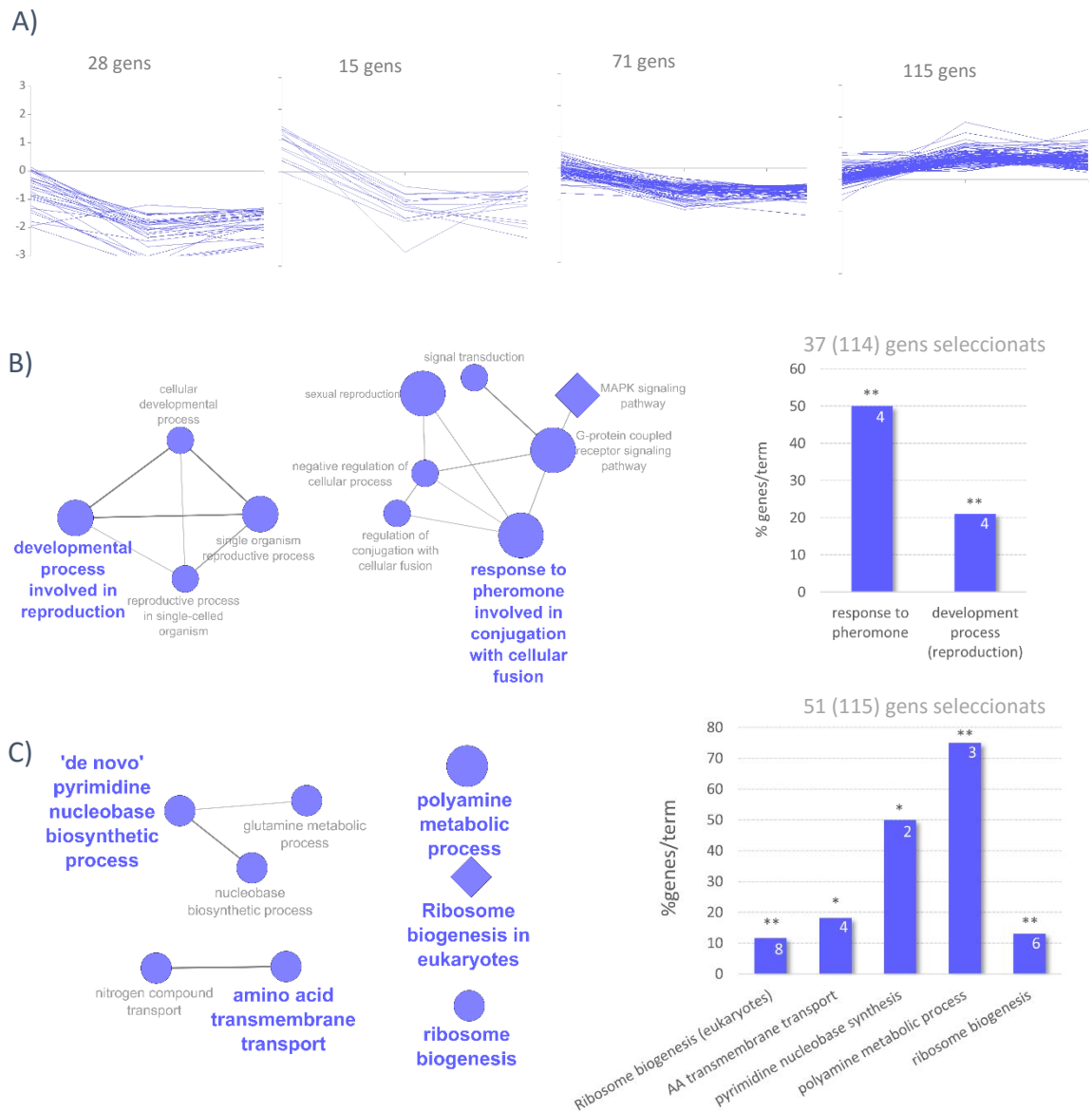


Figura 5. Anàlisi funcional dels gens regulats a les soques amb el sistema d'expressió P_{GAP}. A) Clústers amb una tendència d'expressió similar a les soques productores d'Hsa i 3d6. B) Anàlisi funcional dels gens regulats negativament a les soques Hsa i 3d6. C) Anàlisi funcional dels gens sobreexpressats a les soques Hsa i 3d6. Les diferents bases de dades utilitzades es representen amb un cercle (GO database) o un rombe (KEGG pathways). La mida dels nodes indica el nombre de gens implicats a cada funció. Només es representen aquells termes metabòlics amb un p-val<0.05. *p-val < 0.05; **p-val < 0.01

Els clústers amb una regulació negativa van mostrar un enriquiment significatiu en rutes implicades en la reproducció sexual, tal i com mostraven els termes “procés de desenvolupament implicat en reproducció”, “resposta a feromones” i “regulació de la conjugació”, entre d’altres (Figura 5B). Concretament, en aquests clústers es van localitzar gens codificants per proteïnes essencials per l’activació de l’aparellament, com *FIG1*, *FUS3*, *PRM1* i *PAS_chr4_1001*, receptors de feromones (*STE2* i *STE3*) i diversos gens implicats en l’esperulació (*RIM4* i *SPS22*). *P. pastoris* és un llevat ascomicets homotàl·lic amb reproducció per gemmació, que pot presentar estats haploides i diploides (Kurtzman, 2011). No obstant, a diferència del llevat *Saccharomyces cerevisiae*, *P. pastoris* és més estable en l’estat haploide, i es manté de manera predominant en aquest estat a menys que es vegi forçat a aparellar-se en certes circumstàncies (per exemple, en condicions de nitrogen limitant) i en conseqüència hagi de produir espores (Cregg et al., 1998). La regulació negativa del cicle sexual observat a les soques amb el sistema PGAP podria reflectir una reducció d’aquelles rutes no essencials en aquestes condicions de cultiu. D’acord amb aquestes observacions, Rebnegger et al. (2014) van reportar un comportament similar en cultius operats en continu de *P. pastoris* a taxes de creixement altes, equivalent a un subministrament ric de nutrients. En aquest cas, els autors també van descriure un enriquiment en els processos d’aparellament i resposta a feromones, mentre que a taxes intermèdies els nivells d’expressió d’aquestes vies era més elevat.

D’altra banda, el clúster amb una tendència positiva en les soques Hsa i 3d6 es trobava enriquit en termes relacionats amb la biogènesi dels ribosomes, el transport d’aminoàcids i el metabolisme de les poliamines (Figura 5C). És a dir, es va observar una sobreexpressió global en els processos implicats en la síntesi proteica. En particular, diversos gens necessaris per la maturació del rRNA, la biogènesi i l’assemblatge dels ribosomes (*NOP1*, *NOP4*, *NOP56*, *NOB1*, *PAB1*, *UTP4*, *NOG1*, *RLP24*, *NSR1*) juntament amb gens relacionats amb el tRNA (*LHP1*, *YEF3* i *DED81*) estaven regulats positivament a les soques Hsa_H i 3d6_H respecte la soca no productora. A més, es va observar que els gens inclosos en el terme “procés metabòlic de les poliamines” estaven implicats en la síntesi d’espermidina i hipusina. Tal i com va reportar Chattopadhyay et al. (2003), a *S. cerevisiae* l’espermidina és un factor essencial en la modificació de la hipusina del factor d’iniciació eIF5A, implicat en el metabolisme del mRNA, la biogènesi ribosomal i el procés de traducció. Remarcablement, el factor d’iniciació eIF4A de *P. pastoris* (codificat per *TEF1*) també es va trobar en aquest clúster, reforçant la hipòtesi d’una regulació global d’aquestes vies. Finalment, en relació amb el metabolisme dels aminoàcids, al terme “transport transmembrana d’aminoàcids” es van trobar diversos gens codificants per les permeases de l’arginina, la metionina, la lisina i la histidina (*CAN1*, *MUP1*, *LYP1*, *HIP1*, respectivament), a més d’altres gens relacionats amb la síntesi de

diferents aminoàcids (*ILV2*, *THR1*, *CYS3*). De manera interessant, la nul·la regulació d'aquests processos a la soca Rol_h (tot i presentar una producció de proteïna recombinant més elevada que la soca 3d6_H) podria indicar un mecanisme diferent per tal d'incrementar les taxes traduccionals en les soques amb el sistema d'expressió P_{AOX1} . Tal i com va descriure prèviament Prielhofer et al. (2015) la transcripció dels gens codificants pels constituents ribosomals i d'altres parts de la via traduccional no es veien afectades en les soques creixent en metanol en comparació a les soques amb glucosa com a font de carboni (tot i l'activació de la maquinària per l'assimilació del metanol i la proliferació dels peroxisomes). En efecte, aquests autors van observar un grau major de mRNAs associats a polisomes (és a dir, taxes més elevades de traducció) a les soques que utilitzaven metanol com a font de carboni. Aquests resultats podrien indicar que l'expressió heteròloga sota el control del promotor constitutiu P_{GAP} podria desencadenar una activació dels ribosomes i la síntesi d'aminoàcids, mentre que sota el control del promotor induïble P_{AOX1} es produiria un canvi en el perfil d'unió als polisomes.

5.2.6 Els nivells de complexitat proteica tenen un impacte mínim sobre el metabolisme global i l'estrès quan no es detecta proteïna acumulada

En termes de complexitat proteica, es van agrupar 51 gens amb una tendència lleugerament positiva a mesura que augmentava la complexitat estructural (és a dir, de nivells d'expressió baixos a la soca Rol_H fins nivells elevats per la soca 3d6_H) (Figura 6A). L'anàlisi funcional va revelar un mínim enriquiment en termes relacionats amb el metabolisme dels aminoàcids ("procés biosintètic cel·lular dels aminoàcids", "metabolisme de l'alanina, l'aspartat i el glutamat", entre d'altres), però no es va detectar una regulació global en aquest grup de gens (Figura 6B). Aquests resultats es podrien explicar pel fet que les soques utilitzades en aquest estudi presenten diferents nivells d'expressió. D'aquesta manera, la soca productora de 3d6 (la proteïna model més complexa) mostrava els nivells d'expressió més baixos, i conseqüentment, els principals processos teòricament afectats per la complexitat proteica (plegament, via secretora) no es veien regulats en comparació amb la soca no productora, impossibilitant la determinació d'una tendència reguladora característica enfront les altres soques productores. És més, la manca de regulació de la via ERAD (de l'anglès *ER-associated protein degradation*) o d'altres respostes d'estrès relacionat amb la degradació de proteïna recombinant (Hohenblum et al., 2004) indicaria una absència d'acumulació proteica significativa al citoplasma en aquestes condicions, tal i com s'havia reportat per les soques multicòpia de *ROL* anteriorment Cámara et

al. (2016b), on els nivells d'acumulació de lipasa no eren suficients per afectar de manera global els processos de degradació, plegament o secreció de proteïna. D'aquesta manera, tot i no disposar de dades dels nivells intracel·lulars d'Hsa i 3d6 a les soques amb alta dosi gènica, es podria descartar que els nivells baixos de producció d'anticòs fossin deguts a una limitació de la via secretora, sinó a un coll d'ampolla a nivells transcripcionals i/o traduccional.

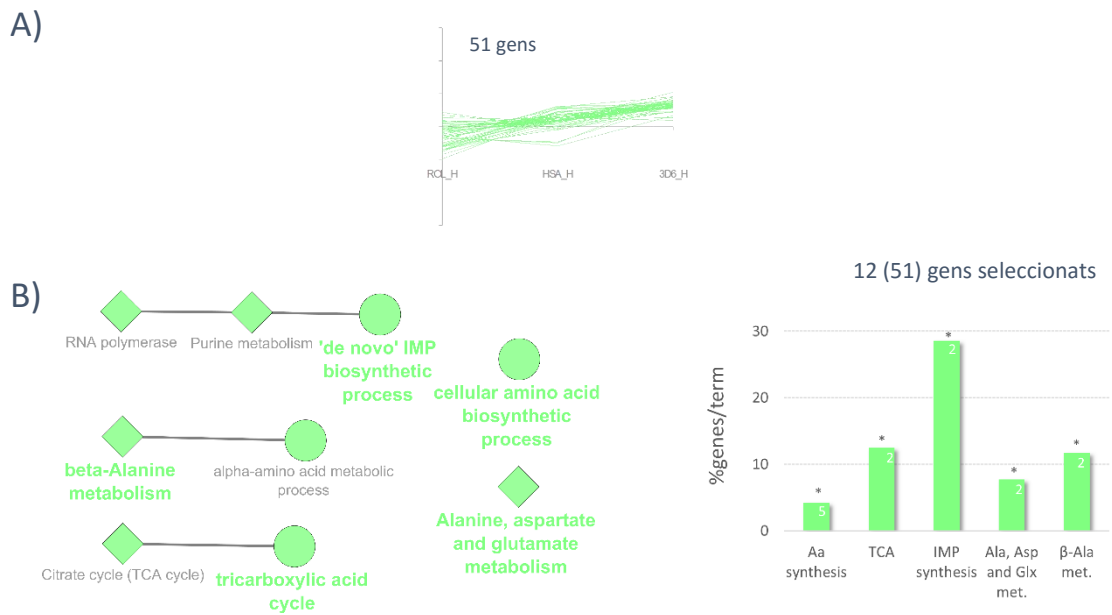


Figura 6. Anàlisi funcional dels gens amb una regulació dependent de la complexitat proteica. A) Clúster on s'observa un increment de l'expressió correlacionada amb la complexitat estructural de la proteïna recombinant B) Anàlisi funcional dels gens. Les diferents bases de dades utilitzades es representen amb un cercle (GO database) o un rombe (KEGG pathways). La mida dels nodes indica el nombre de gens implicats a cada funció. Només es representen aquells termes metabòlics amb un $p\text{-val} < 0.05$. * $p\text{-val} < 0.05$.

5.2.7 Els nivells elevats d'albumina de sèrum humà podrien ser deguts a una sobreexpressió de gens relacionats amb l'inici de la transcripció i la maquinària de reparació de DNA

Els gens regulats principalment a la soca Hsa_H (però no a les soques productores de 3d6 o Rol) podrien proporcionar informació sobre els mecanismes moleculars que estimulen una elevada producció de proteïna recombinant (Figura 7A). Considerant els gens únicament sobreexpressats a

la soca Hsa_H, l'anàlisi funcional només va assignar 20 dels 105 gens agrupats en aquest clúster a diversos termes. Tot i que sense ser estadísticament significatiu, es va observar un enriquiment general del metabolisme dels aminoàcids i el RNA (Figura 7B). Concretament, els gens localitzats en aquest clúster estaven implicats en els processos de replicació del DNA, la regulació de la transcripció, la biogènesi dels ribosomes i la biosíntesi i captació d'aminoàcids. En aquest context, es van trobar en aquest clúster diversos gens relacionats amb la qualitat del DNA i l'inici de la transcripció. *CRC39* i *NOT5*, ambdós codificants per diferents subunitats del complex CRC4-NOT (amb múltiples funcions en la transcripció, traducció i degradació de proteïnes en llevats; Miller i Reese, 2012) es van trobar sobreexpressats a la soca Hsa_H, juntament amb *BDF1*, codificant per a una proteïna que interacciona amb el complex TFIID (factor de transcripció II D). Notablement, tal i com van revisar Miller i Reese (2012), les interaccions entre les subunitats del complex CRC4-NOT i TFIID (entre d'altres) evidencia el paper d'aquestes molècules en la iniciació de la transcripció. De manera addicional, *TFB2* (codificant per una subunitat del complex TFIID (factor de transcripció H)), també implicat en l'inici de la transcripció i essencial en la reparació de DNA als llevats (Kainov et al., 2010) va presentar una regulació positiva, reforçant la idea que a la soca Hsa_H hi ha un increment dels factors addicionals necessaris per l'activació acurada de la transcripció per part de la polimerasa II (Dujon et al., 2012). Tanmateix, diferents gens codificants per proteïnes implicades en la reparació i replicació del DNA es van agrupar en aquest clúster (*RAD5*, *SLD2*, *HCS1*, *NSE4*). Aquestes observacions podrien indicar una eficiència major per part de la soca Hsa_H a l'hora d'activar la maquinària transcripcional, incloent els processos de reparació del DNA, resultant en nivells d'expressió més elevats d'albúmina.

Adicionalment, diversos gens codificants per sintetases de tRNA també es van trobar regulades de manera positiva (*ILS1*, *TYS1*, *PAS_chr2-1_0314*), juntament amb gens implicats en la biogènesi dels ribosomes i el seu assemblatge (*NOP6*, *NOP14*, *NOP58*, *SAS10*, *RRP14*, *JJJ1*, *TSR1*, *ECM1*, *RPL3*, *MAK21*, *ENP2*). A més, en aquest grup també es van trobar diferents permeases d'aminoàcids i d'altres transportadors (*PUT4-1*, *VBA5*, *TPO2*, *AGC1*, *PAS_chr3_0182*) així com diferents gens relacionats amb algunes etapes de la síntesi d'aminoàcids (*CPA2*, *GLT1*, *LYS4*, *MET1*, *MET13*). Aquests resultats podrien reflectir una activació global de la maquinària proteica, degut a la producció elevada d'albúmina. De manera interessant, aquesta regulació lleugera (però global) no s'observa a les altres soques recombinants quan es comparen amb la soca no productora. Això es podria explicar pels nivells baixos de producció obtinguts en condicions de cultiu en continu, insuficients per desencadenar una resposta detectable a les soques, excepte per Hsa_H.

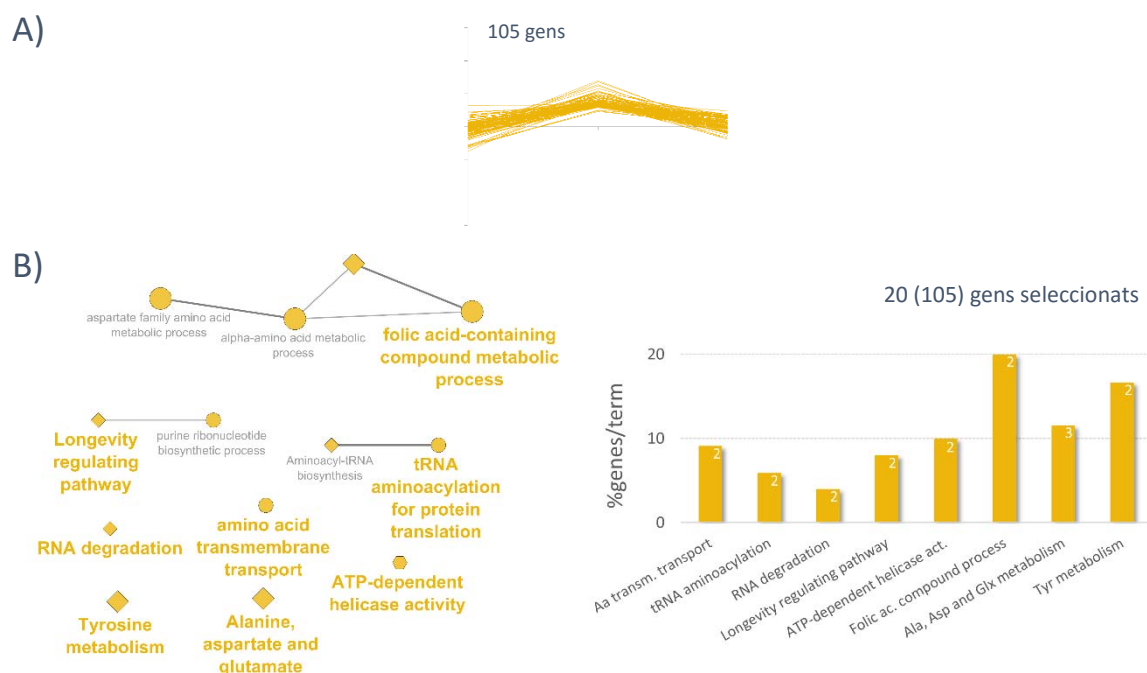


Figura 7. Anàlisi funcional dels gens regulats específicament a la soca Hsa_H. A) Clúster on s’observa un increment de l’expressió només a les soques Hsa_h en comparació amb la soca no productora B) Anàlisi funcional dels gens. Les diferents bases de dades utilitzades es representen amb un cercle (GO database) o un rombe (KEGG pathways). La mida dels nodes indica el nombre de gens implicats a cada funció.

5.2.8 La limitada anotació funcional del genoma de *P. pastoris* dificulta l’anàlisi de diversos grups d’interès

En relació al clúster amb una regulació negativa únicament a la soca Hsa_H, 33 dels 53 gens no estaven anotats funcionalment (Figura 8A). En aquest clúster es van trobar *LSM3* i *LSM5*, gens codificants per proteïnes relacionades amb la degradació del mRNA, juntament amb dos components del complex SAGA (codificats per *SPT20* i *SGF29*). Malauradament, no es van poder extreure patrons generals d’expressió dels gens regulats negativament únicament a la soca Hsa_H.

Entre els gens regulats negativament únicament a la soca 3d6 (Figura 8B) només 32 dels 70 gens disposaven d’alguna anotació funcional. En aquests clústers es van trobar agrupats diferents transportadors de la membrana plasmàtica, implicats en la captació de glucosa, zinc i ferro (codificats pels gens *GTH1*, *PAS_chr3_0516* i *FET4-1*, respectivament). Addicionalment, diversos

factores de transcripció del tipus zinc van aparèixer en aquest grup (*STB4*, *CAT8-2*, *PAS_chr1-4_0038*, *PAS_chr1-3_0015*, *PAS_chr3_0012*, *YPR022C-1*) tot i que seria necessària informació addicional per tal d'esbrinar en quins processos reguladors estan implicats.

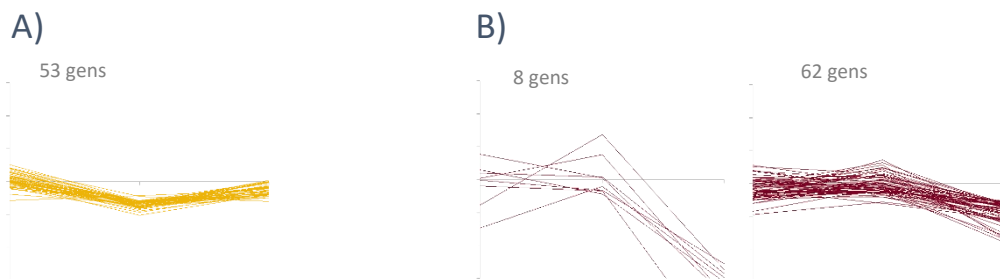


Figure 8. Clústers amb anotació funcional limitada. A) Cúster dels gens amb regulació negativa únicament a la soca Hsa_H. B) Clústers dels gens amb regulació negativa a la soca 3d6_H (en comparació a la soca no productora)

5.2.9 La comparativa dels perfils transcriptòmics de les soques d'alta dosi gènica envers les de baixa dosi pot proporcionar diverses dianes per futures estratègies d'enginyeria genètica

Amb l'objectiu d'analitzar més detalladament l'impacte d'elevades dosis gèniques (i conseqüentment, produccions de proteïna recombinant més elevades) a les soques, es va dur a terme una comparativa entre les soques amb alta i baixa dosi gènica, per tal d'estudiar únicament l'impacte de l'expressió heteròloga al transcriptoma.

D'aquesta manera, un total de 336 gens es van trobar regulats en alguna de les soques. Tal i com es pot observar a la Figura 9A i 9B, el nombre de gens regulats en aquesta comparativa va ser significativament més elevat a la soca Rol que a les soques productores d'Hsa i 3d6, indicant un major impacte de l'increment de la dosi gènica en les soques amb el sistema P_{AOX1} , en consonància amb el que reflectien els paràmetres fisiològics de creixement i consum (Taula I).

Mitjançant l'algoritme *k-means* ($k=9$) es van obtenir les principals tendències reguladores dels gens expressats de manera diferencial (Figura 9C). Els clústers 1 i 2, on s'agrupaven els gens principalment regulats a la soca Rol_H (en comparació amb la soca Rol_L), presentaven un enriquiment amb termes relacionats amb el metabolisme del metanol i els àcids grassos, respectivament, tal i com s'havia descrit prèviament (Cámara et al., 2016b).

En el cas dels gens únicament regulats a la soca Hsa_H (Figura 9C, clúster 3), 22 dels 44 gens estaven anotats funcionalment, i només 9 estaven significativament regulats en comparació a la soca amb dosi gènica baixa (Hsa_L). De manera remarcable, es va observar una sobreexpressió 7 vegades superiors al gen *SGA1*, implicat en la degradació del glicogen. A més, es va mesurar un augment de 2 i 3 vegades pels gens codificants de la S-formilglutamina hidrolasa i la dihidroxiacetona quinasa (*FGH1* i *DAK2*, respectivament), relacionats amb processos de detoxificació (Molin i Blomberg, 2003; Yurimoto et al., 2011). Finalment, es va observar una regulació 3 vegades menor al gen *PAS_chr3_0284* (amb un domini característic de factor de transcripció) a les soques Rol_H i 3d6_H en comparació amb les soques de dosi gènica baixa corresponent, mentre que aquest mateix gen presentava una sobreexpressió del doble a la soca Hsa_H envers Hsa_L, sent així una diana prometedora per futures estratègies d'enginyeria de soques.

El clúster 5 comprenia aquells gens amb regulació negativament únicament a la soca 3d6_H en comparació amb 3d6_L. De manera interessant, 4 dels 14 gens anotats estaven relacionats amb transport. Concretament, els transportadors relacionats amb el ferro codificats per *FET4-1* i *PAS_chr3_0315*, el transportador amb alta afinitat pel zinc *PAS_chr3_0516*, i el transportador general *PAS_chr3_0008*. D'acord amb aquestes observacions, entre els diferents possibles factors de transcripció inclosos en aquest grup es va trobar *STB4*, codificant per un factor de transcripció amb una possible regulació de diferents transportadors. Finalment, es va mesurar una disminució de 2.5 vegades en el gen codificant per la biotina sintasa (*BIO2*), un dels dos gens regulats a totes les soques d'alta dosi gènica envers les de baixa dosi (diagrama de Venn, Figura 9B). Contràriament, a les soques Hsa_H i Rol_H aquest gen es trobava sobreexpressat 2 i 3 vegades en comparació amb les soques de baixa dosi gènica, respectivament. Donat el paper essencial de la biotina en el metabolisme dels llevats, actuant com a co-factor de diverses carboxilases (Tong, 2013) implicades en el metabolisme del piruvat, els àcids grassos i els aminoàcids, aquesta regulació negativa de *BIO2* a la soca 3d6_H podria reflectir una reestructuració del metabolisme central del carboni provocat pels elevats nivells d'expressió d'una proteïna tan complexa com 3d6 (tot i que no suficientment elevats per tal de veure's reflectit en el creixement cel·lular o les taxes de consum de substrat).

Analitzant el clúster 7, amb una tendència negativa en relació amb la complexitat proteica, dos grups funcionals van destacar sobre els altres: els gens relacionats amb els processos d'aparellament i la formació d'espores (*PAS_chr2-1_0550*, *CDA2*, *CHS1*, *MPT5*, *KCH1*, *CTS1*, *RIM4*, *YJL171C*, *STE11*), i els gens implicats en el transport a través de la membrana (*PDR5-2*, *PIC2*, *OPT1*, *PAS_chr2-1_006*, *TAT2*, *RSB1*). Sorprenentment, al clúster 9 (amb un increment de la sobreexpressió correlacionada amb l'augment de la complexitat estructural de la proteïna

recombinant) també es van trobar gens implicats en la resposta a feromones i esporulació (*STE6-1*, *SSP2*, *IME2*, *MATalpha1*), indicant una possible funció alternativa d'aquestes proteïnes que podria estar relacionada amb la producció de proteïna recombinant.

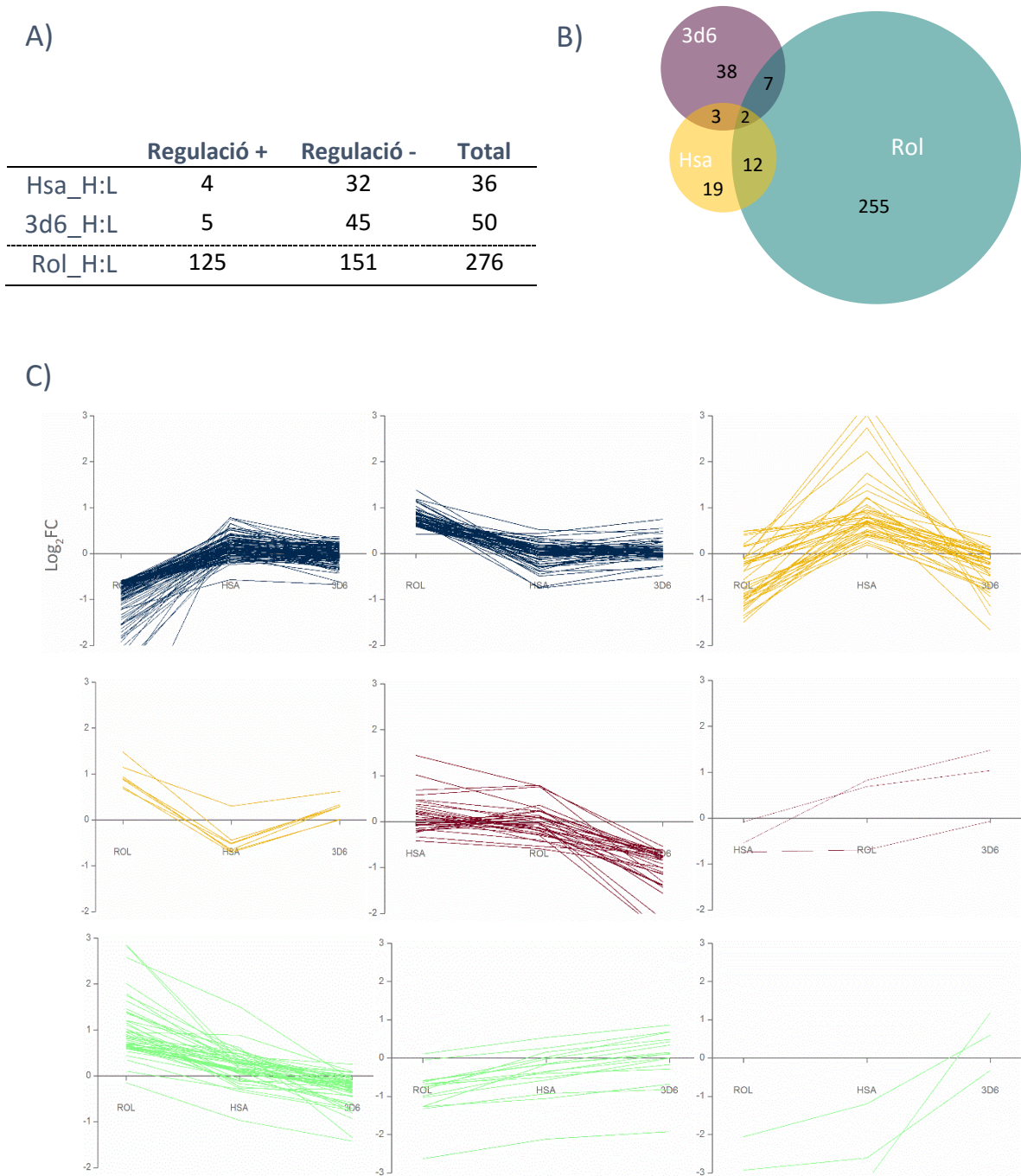


Figura 9. Anàlisi de l'expressió gènica de les soques amb alta dosi envers les soques amb baixa dosi gènica.

A) Taula amb els gens regulats de manera diferencial a la soca d'alta dosi gènica (Hsa_H, 3d6_H, Rol_H) envers les soques amb baix nombre de còpies (Hsa_L, 3d6_L, Rol_L) B) Diagrama de Venn proporcional (s'inclouen el nombre de gens comuns regulats en cada comparativa) C) Anàlisi de clústers dels 336 gens regulats en alguna de les diferents comparatives mitjançant l'algorisme k-means (k=9). Blau, regulació específica de les soques amb sistema P_{AOX1}; taronja, regulació específica de les soques productores d'Hsa; vermell, regulació dependent dels nivells de productivitat de les soques; verd; regulació dependent de la complexitat estructural de la proteïna recombinant expressada

5.3. REFERÈNCIES

Baumann K, Carnicer M, Dragosits M, Graf AB, Stadlmann J, Jouhten P, Maaheimo H, Gasser B, Albiol J, Mattanovich D, Ferrer P. 2010. A multi-level study of recombinant *Pichia pastoris* in different oxygen conditions. *BMC Syst. Biol.* 4:141.

Baumann K, Dato L, Graf AB, Frascotti G, Dragosits M, Porro D, Mattanovich D, Ferrer P, Branduardi P. 2011. The impact of oxygen on the transcriptome of recombinant *S. cerevisiae* and *P. pastoris* - a comparative analysis. *BMC Genomics* 12:218.

Cámara E, Albiol J, Ferrer P. 2016a. Droplet digital PCR-aided screening and characterization of *Pichia pastoris* multiple gene copy strains. *Biotechnol. Bioeng.* 113:1542–1551

Cámara E, Landes N, Albiol J, Gasser B, Mattanovich D, Ferrer P. 2016b. Increased dosage of *AOX1* promoter-regulated expression cassettes leads to transcription attenuation of the methanol metabolism in *Pichia pastoris*. Under submission of Scientific Reports.

Cámara E, Monforte S, Albiol J, Ferrer P. 2016c. Deregulation of methanol metabolism reverts transcriptional limitations of recombinant *Pichia pastoris* with multiple expression cassettes under control of *AOX1* promoter. Under submission.

Chattopadhyay MK, Tabor CW, Tabor H. 2003. Spermidine but not spermine is essential for hypusine biosynthesis and growth in *Saccharomyces cerevisiae*: spermine is converted to spermidine in vivo by the FMS1-amine oxidase. *Proc. Natl. Acad. Sci. U. S. A.* 100:13869–13874.

Cregg, J., Shen, S., Johnson, M., Waterham, H., Classical genetic manipulation. In: Higgins, D.R., Cregg, J.M. (Eds.), *Pichia Protocols*, Humana Press, New York, 1998, pp. 17–26.

De Schutter K, Lin Y-C, Tiels P, Van Hecke A, Glinka S, Weber-Lehmann J, Rouzé P, Van de Peer Y, Callewaert N. 2009. Genome sequence of the recombinant protein production host *Pichia pastoris*. *Nat. Biotechnol.* 27:561–566.

Dong ZC, Chen Y. 2013. Transcriptomics: Advances and approaches. *Sci. China Life Sci.* 56:960–967.

Dragosits M, Stadlmann J, Graf A, Gasser B, Maurer M, Sauer M, Kreil DP, Altmann F, Mattanovich D. 2010. The response to unfolded protein is involved in osmotolerance of *Pichia pastoris*. *BMC Genomics* 11:207.

Dujon B, Branduardi P, Porro D, Gaillardin C, Feldmann H. 2012. Yeast. Ed. Horst Feldmann. Wiley-Blackwell 455 p.

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.

Smyth, G. K. (2005). Limma: linear models for microarray data. En: *Bioinformatics and Computational Biology Solutions using R and Bioconductor*, R. Gentleman, V. Carey, S. Dudoit, R. Irizarry, W. Huber (eds.), Springer, New York, pages 397–420.

Hesketh AR, Castrillo JI, Sawyer T, Archer DB, Oliver SG. 2013. Investigating the physiological response of *Pichia (Komagataella) pastoris* GS115 to the heterologous expression of misfolded proteins using chemostat cultures. *Appl. Microbiol. Biotechnol.* 97:9747–9762.

Hohenblum H, Gasser B, Maurer M, Borth N, Mattanovich D. 2004. Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant *Pichia pastoris*. *Biotechnol. Bioeng.* 85:367–75.

Hulsen T, de Vlieg J, Alkema W. 2008. BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* 9:488.

Kainov DE, Selth LA, Svejstrup JQ, Egly JM, Poterzsmann A. 2010. Interacting partners of the Tfb2 subunit from yeast TFIIF. *DNA Repair (Amst).* 9:33–39.

Kell DB, Oliver SG. 2004. Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. *BioEssays* 26:99–105.

Kobayashi K, Kuwae S, Ohya T, Ohda T, Ohyama M, Ohi H, Tomomitsu K, Ohmura T. 2000. High-level expression of recombinant human serum albumin from the methylotrophic yeast *Pichia pastoris* with minimal protease production and activation. *J. Biosci. Bioeng.* 89:55–61.

Kurtzman, C.P. *Komagataella Y.* Yamada, Matsuda, Maeda & Mikata 1995, In: *The Yeast. A Taxonomic Study*, Kurtzman, C.P., Fell, J., and Boekhout, T., Eds., 5th ed. Amsterdam: Elsevier, 2011, vol. 2, pp. 491–495.

Liang S, Wang B, Pan L, Ye Y, He M, Han S, Zheng S, Wang X, Lin Y. 2012. Comprehensive structural annotation of *Pichia pastoris* transcriptome and the response to various carbon sources using deep paired-end RNA sequencing. *BMC Genomics* 13:738.

Liu L, Feizi A, Österlund T, Hjort C, Nielsen J. 2014. Genome-scale analysis of the high-efficient protein secretion system of *Aspergillus oryzae*. *BMC Syst. Biol.* 8:73.

Love KR, Shah KA, Whittaker CA, Wu J, Bartlett MC, Ma D, Leeson RL, Priest M, Borowsky J, Young SK, Love JC. 2016. Comparative genomics and transcriptomics of *Pichia pastoris*. *BMC Genomics* 17:550.

Mattanovich D, Graf A, Stadlmann J, Dragosits M, Redl A, Maurer M, Kleinheinz M, Sauer M, Altmann F, Gasser B. 2009. Genome, secretome and glucose transport highlight unique features of the protein production host *Pichia pastoris*. *Microb. Cell Fact.* 8:29.

Miller JE, Reese JC. 2012. Ccr4-Not complex: the control freak of eukaryotic cells. *Crit. Rev. Biochem. Mol. Biol.* 47:315–333.

Mojzita D, Oja M, Rintala E, Wiebe M, Penttilä M, Ruohonen L. 2014. Transcriptome of *Saccharomyces cerevisiae* during production of D-xylonate. *BMC Genomics* 15:763.

Molin M, Norbeck J, Blomberg A. 2003. Dihydroxyacetone kinases in *Saccharomyces cerevisiae* are involved in detoxification of dihydroxyacetone. *J. Biol. Chem.* 278:1415–1423.

Ohya T, Ohyama M, Kobayashi K. 2005. Optimization of human serum albumin production in methylotrophic yeast *Pichia pastoris* by repeated fed-batch fermentation. *Biotechnol. Bioeng.* 90:876–887.

Prielhofer R, Cartwright SP, Graf AB, Valli M, Bill RM, Mattanovich D, Gasser B. 2015. *Pichia pastoris* regulates its gene-specific response to different carbon sources at the transcriptional, rather than the translational, level. *BMC Genomics* 16:1–17.

Rebnegger C, Graf AB, Valli M, Steiger MG, Gasser B, Maurer M, Mattanovich D. 2014. In *Pichia pastoris*, growth rate regulates protein synthesis and secretion, mating and stress response. *Biotechnol. J.* 9:511–525.

Rußmayer H, Buchetics M, Gruber C, Valli M, Grillitsch K, Modarres G, Guerrasio R, Klavins K, Neubauer S, Drexler H, Steiger M, Troyer C, Al Chalabi A, Krebiehl G, Sonntag D, Zellnig G, Daum G,

- Graf AB, Altmann F, Koellensperger G, Hann S, Sauer M, Mattanovich D, Gasser B. 2015. Systems-level organization of yeast methylotrophic lifestyle. *BMC Biol.* 13:80.
- Sauer M, Branduardi P, Gasser B, Valli M, Maurer M, Porro D, Mattanovich D. 2004. Differential gene expression in recombinant *Pichia pastoris* analysed by heterologous DNA microarray hybridisation. *Microb. Cell Fact.* 3:17.
- Sims AH, Gent ME, Lanthaler K, Nigel S, Oliver SG, Robson GD, Dunn-coleman NS. 2005. Transcriptome Analysis of Recombinant Protein Secretion by *Aspergillus nidulans* and the Unfolded-Protein Response In Vivo. *Appl. Environ. Microbiol.* 71:2737–2747.
- Spies D, Ciaudo C. 2015. Dynamics in Transcriptomics: Advancements in RNA-seq Time Course and Downstream Analysis. *Comput. Struct. Biotechnol. J.* 13:469–477.
- Tong L. 2013. Structure and function of biotin-dependent carboxylases. *Cell. Mol. Life Sci.* 70:863–891.
- Ventimiglia G, Petralia S. 2013. Recent Advances in DNA Microarray Technology: An Overview on Production Strategies and Detection Methods. *Bionanoscience* 3:428–450.
- Wang Z, Gerstein M, Snyder M. 2010. RNA-Seq: a revolutionary tool for transcriptomics. *Zhong* 10:57–63.
- Westerhoff H V, Palsson BO. 2004. The evolution of molecular biology into systems biology. *Nat. Biotechnol.* 22:1249–52.
- Yurimoto H, Oku M, Sakai Y. 2011. Yeast methylotrophy: Metabolism, gene regulation and peroxisome homeostasis. *Int. J. Microbiol.* 2011.

6. CONCLUSIONS GENERALS I PERSPECTIVES FUTURES

La creixent popularitat del llevat *Pichia pastoris* com a hoste per a la producció de proteïnes recombinants ha potenciat l'aparició de diferents estratègies per tal de millorar la seva expressió heteròloga. L'increment del nombre de còpies del gen a expressar és una de les tècniques més àmpliament utilitzades, tot i que altes dosis gèniques sovint deriven en sobrecàrregues metabòliques a la soca productora.

Fins el dia d'avui, diferents estudis s'havien centrat principalment en l'impacte d'una elevada producció de proteïna heteròloga en les vies de plegament i secreció de la cèl·lula. En aquest treball descrivim el primer estudi sistemàtic de l'impacte de l'increment de la dosi gènica al llevat *P. pastoris* mitjançant un anàlisi transcriptòmic global, utilitzant la lipasa de *Rhizopus oryzae* (Rol) com a proteïna model, expressada sota el control de P_{AOX1} .

Concretament, es va construir una col·lecció de soques amb un nombre creixent de còpies de *ROL*, i mitjançant cultius a diferent escala, es va determinar que la productivitat més elevada es trobava a la soca amb dues còpies de *ROL*, mentre que subseqüents increments provocaven una disminució en la producció de lipasa. D'altra banda, es va observar que l'assimilació de metanol (font de carboni als diferents cultius) es veia disminuïda a les soques multicòpia de *ROL*, en correlació amb el nombre de còpies de cada soca.

Els estudis transcriptòmics posteriors van confirmar l'impacte negatiu de la dosi gènica a la ruta metabòlica del metanol, a més de rutes relacionades com la biogènesi dels peroxisomes i la producció de cofactors necessaris pels enzims implicats en aquestes rutes. L'anàlisi a diferents nivells de la producció lipàsica va descartar les vies de plegament i secreció com a principals factors

limitants a les soques amb alta dosi gènica, apuntant a una limitació transcripcional com a principal factor de la caiguda de l'activitat lipàsica. D'aquesta manera, es va proposar el factor de transcripció Mxr1 (principal regulador de P_{AOX1} i altres promotors relacionats amb el metabolisme del metanol) com principal causa de la caiguda de l'expressió dels gens implicats en aquestes rutes (inclòs *ROL*), donada la seva expressió constitutiva a baixos nivells.

Aquesta hipòtesi es va verificar amb la modificació d'una soca amb 4 còpies de *ROL* co-expressant una variant desregulada de Mxr1. La millora en l'assimilació del metanol com a font de carboni i l'increment en la producció de lipasa van demostrar que l'increment del nombre de còpies dels cassets d'expressió sota el control de P_{AOX1} poden provocar una insuficiència en el nombre de molècules de Mxr1 per tal d'activar tots els promotors necessaris.

Finalment, l'estudi comparatiu a nivell transcriptòmic de l'expressió de diferents proteïnes a *P. pastoris* va revelar el fort impacte de la font de carboni en l'expressió gènica global, tal i com havien apuntat d'altres autors. A més, va donar noves evidències dels factors implicats en la replicació del DNA i la transcripció del RNA que s'expressen de manera diferencial només en aquelles soques que presenten una elevada productivitat, proporcionant així noves dianes per a futures estratègies d'enginyeria genètica.

Per tal d'explotar el potencial de l'increment de la dosi gènica com a estratègia de millora dels sistemes basats en P_{AOX1} , les futures estratègies de modificació s'haurien de centrar en la sobreexpressió de Mxr1 a les soques multicòpia de *ROL*, a banda de sobreexpressar d'altres factors reguladors de la via del metanol (per exemple, Trm1). A més, la coexpressió de factors de secreció i/o plegament podrien millorar encara més l'eficiència de les soques, per tal d'evitar l'acumulació intracel·lular de lipasa. Finalment, l'enriquiment del medi de cultiu amb aminoàcids i altres suplementes energètics permetrien esbrinar si la capacitat traduccional de les soques multicòpia es podria augmentar en aquestes condicions.

7. CONTRIBUCIONS CIENTÍFIQUES

Cámara E, Albiol J, Ferrer P. 2016. Droplet digital PCR-aided screening and characterization of *Pichia pastoris* multiple gene copy strains. *Biotechnol. Bioeng.* 113:1542–1551.

Contribució dels autors

Elena Cámara va construir la col·lecció de soques multicòpia, va realitzar tots els cultius, la posta a punt de la ddpCR i la comparativa amb la qPCR, els processos analítics, l'anàlisi de les dades i la interpretació dels resultats, a més de contribuir en el disseny experimental de l'estudi i escriure el manuscrit.

Joan Albiol va contribuir en el disseny de l'estudi, l'anàlisi de les dades i la revisió del manuscrit.

Pau Ferrer va participar en el disseny conceptual de l'experiment, la interpretació dels resultats i la revisió del manuscrit.

Cámara E, Landes N, Albiol J, Gasser B, Mattanovich D, Ferrer P. Increased dosage of *AOX1* promoter-regulated expression cassettes leads to transcription attenuation of the methanol metabolism in *Pichia pastoris*. Manuscrit enviat a la revista científica Scientific Reports.

Contribució dels autors

Elena Cámara va realitzar tots els cultius, la presa de mostres i els experiment d'hibridació, la ddpCR, els WB i tots els processos analítics, l'anàlisi de les dades i la interpretació dels resultats, a més de contribuir en el disseny experimental de l'estudi i escriure el manuscrit.

Nils Landes va donar assistència en la realització dels cultius i els microarrays.

Brigitte Gasser i Diethard Mattanovich van participar en el disseny conceptual d'aquest estudi, la interpretació de les dades òmiques i la revisió del manuscrit.

Joan Albiol va contribuir en el disseny de l'estudi, l'anàlisi de les dades i la revisió del manuscrit.

Pau Ferrer va participar en el disseny conceptual de l'experiment, la interpretació dels resultats i la revisió del manuscrit.

Cámara E, Monforte S, Albiol J, Ferrer P. Deregulation of methanol metabolism reverts transcriptional limitations of recombinant *Pichia pastoris* with multiple expression cassettes under control of *AOX1* promoter. Manuscrit enviat a revista científica.

Contribució dels autors

Elena Cámara va realitzar tots els cultius, l'anàlisi per ddPCR, i tots els processos analítics, l'anàlisi de les dades i la interpretació dels resultats, a més d'idear el disseny experimental de l'estudi i escriure el manuscrit.

Sergi Monforte va construir les soques, i va col·laborar en la realització dels cultius, l'anàlisi per ddPCR i els processos analítics.

Joan Albiol va contribuir en el disseny de l'estudi, l'anàlisi de les dades i la revisió del manuscrit.

Pau Ferrer va participar en el disseny conceptual de l'experiment, la interpretació dels resultats i la revisió del manuscrit.

Cámara E, Landes N, Albiol J, Gasser B, Mattanovich D, Ferrer P. Comparative transcriptomics analysis of the physiological response of *Pichia pastoris* to the expression of three different recombinant proteins. Manuscrit en preparació.

Contribució dels autors

Elena Cámara va construir la col·lecció de soques multicòpia de *ROL*, va realitzar tots els cultius de les soques *ROL* i la corresponent analítica i microarrays, i es va encarregar de l'anàlisi funcional, a més d'idear l'estudi comparatiu i escriure el manuscrit.

Nils Landes va realitzar els cultius de les soques productores d'HSA i 3D6, i va dur a terme la corresponent analítica i microarrays.

Brigitte Gasser i Diethard Mattanovich van participar en la interpretació de les dades òmiques.

Joan Albiol va contribuir en l'anàlisi de les dades i la revisió del manuscrit.

Pau Ferrer va participar en el disseny conceptual de l'estudi comparatiu, la interpretació dels resultats i la revisió del manuscrit.

8. ANNEX

Taula I. Seqüències dels oligonucleòtids utilitzats durant aquest treball. *T_m*, melting temperature; %GC, contingut de guanosina + citosina (en %) dels oligonucleòtids.

OLIGONUCLEÒTID	SEQÜÈNCIA (5'-3')	T _m	%GC	MIDA AMPLICON (pb)
sonda ROL	CCGGTCACTCACTCGGTGGTGCA	74.8	65.2	-
sonda ACT	TCCGTATGGATCGGTGGTTCTATCCTCGCT	75.4	53.3	-
ROL fwd1	CCTGTCGTCCAAGAACAAC	62	52.4	164
ROL rev1	GAGGACCACCAACAGTGAAG	62	53	
ACT1 fw	TGTCCGGTGGTACTACTATGTTCC	65	50	199
ACT1 rev	GATTCGTCGTA CTCTTGCTTTGA	62	43	
AOX1 fw	GACATTCACGGTTTCGAAGG	61	50	78
AOX1 rev	CCTCAAGAAGTCCTGGCAAAC	63	52	
DAS1 fw	TTGAACTGGGACGGAGTG	60	55	135
DAS1 rev	CAAACCGTTGGCAATAGCAC	62	50	
FAB1-2 fw	CCCTTGTTTGACGGAATG	60	52	97
FAB1-2 rev	TTCCTCCGACAGGTCTAAC	60	52	
FDH1 fw	GGTGCTGGAAGAATTGG	57	52	135
FDH1 rev	GACAGTGTGACTCTTC	55	52	
KAR2 fw	GATGAAGTCGGGTCGTGTAG	61	55	110
KAR2 rev	TCTTAGCAGCATCACCAACC	62	50	
MXR1 fw	TTCGCCATTCTACTCG	57	52	101
MXR1 rev	GGGCTTGACTCATTTAG	57	52	
pGAP_fw	GATTATTGGAAACCACCAGAATCG	61	41	450
MXR1*_int_rev	AGTCTTGTTATGGCCGCATC	62	50	
MXR1*_PmlI fw	GGCGACACGTGATGAGCAATCTACCCCAAC	75	58	3490
MXR1*_KpnI rev	CGGCAGGTACCCTAGACACCACCATCTAGTCGG	76	60	
MXR1*seq_1	CCGGCTGTATCTGGATTTAG	59	50	-
MXR1*seq_2	CTCTCGATGTGAACAGGAAC	59	50	-
MXR1*seq_3	GTGGCTACGATAGGTGCAG	61	57	-
MXR1*seq_4	CACAGTTGGGATAGGATG	56	50	-
MXR1*seq_5	CTTTCAGCTGCGGCATTTG	62	52	-
MXR1*seq_6	GCAAATGGCATTCTGACATCC	62	47	-

PEX6 fw	GGTTTGGATGTGGTCAAG	57	50	148
PEX6 rev	TGGCTTTAGCAAGCAGAG	59	50	
ROL fw	AAGTGGGACTGTGTCCAATG	61	50	158
ROL rev	GCACTTCTGAAGGAGTTG	57	50	
