SCUOLA DI DOTTORATO

UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Department of Biotechnology and Biosciences

PhD program: Biology and Biotechnology

Cycle: XXX

Curriculum: Biotechnology

Modulation of yeast hub elements of nitrogen and mRNA metabolism for eliciting industrially relevant phenotypes

The *Saccharomyces cerevisiae* glutamate synthase (Glt1) and the poly(A) binding protein (Pab1) as case studies

PhD student: Brambilla Marco

Registration number: 726122

Tutor: Professor Paola Branduardi

Coordinator: Professor Marco Ercole Vanoni

ACADEMIC YEAR

2016-2017

Index

Abstract

Abstract

Nowadays, for the production of energy and materials, our society mainly relies on fossil sources, but many concerns arise from their utilization, like greenhouse gases emission and non-renewability within the time of their consumption. Hence, biorefineries, which convert renewable biomasses into products and energy, could be a promising alternative. Despite some biorefineries are now at commercial scale, many challenges must be overcome to implement competitive bio-based industries, such as improving the performances of microorganisms (named cell factory) used during the production processes. Above all, the yeast *Saccharomyces cerevisiae* is the most prominent cell factory for producing bioethanol as biofuel.

The main objective of this thesis was to engineer *S. cerevisiae* strains with biotechnological interesting traits, among which improved growth performances and increased thermotolerance. For this purpose, we investigated the possibility to apply a cellular rewiring, by selecting two targets: the glutamate synthase (Glt1) that is an enzyme of the central nitrogen metabolism (CNM), and the main poly(A) binding protein (Pab1), a master regulator of mRNA metabolism.

Regarding *GLT1*, the physiological effects of its deletion and over-expression were assessed by growing yeasts in the presence of different nitrogen sources. Results showed that, independently on *GLT1* deletion or over-expression, the supplementation of ammonium sulfate, glutamate or glutamine considerably affects growth, protein content, viability and reactive oxygen species (ROS) accumulation. Conversely, *GLT1* modulation does not significantly influence these parameters. Overall, these data highlight the plasticity of the *S. cerevisiae* CNM in respect to the environment and confirm its robustness against internal perturbation. Moreover, even though the sole modulation of *GLT1* expression might not rewire the entire cell, the physiological characterization of this study might be helpful to guide the selection of other more promising candidate for the application of the rewiring approaches.

Then, the induction of cellular rewiring was assessed selecting Pab1. To this purpose, a strain carrying the unaltered *PAB1* chromosomal allele was transformed with a *PAB1* plasmid mutant library and then screened for isolating strains with high thermotolerance. The isolated clones showed growth improvement at both high temperatures and ethanol concentration by drop tests. Among all, the *PAB1* S40.7 variant was further characterized because, strikingly, it dominantly confers higher thermotolerance by expressing just the first 20 amino acidic residues of Pab1. This improved phenotype was also confirmed in bioreactor at 40°C. Remarkably, the S40.7 strain accumulates less ROS compared to the control strain, thus possibly explaining its increased thermotolerance. Overall, these results demonstrated that Pab1 is a powerful candidate to evoke complex phenotypes with improved traits, among which higher thermotolerance.

Finally, Pab1 was characterized to uncover the role of its six domains in the recruitment of the protein within stress granules (aggregates of untranslated mRNPs that form during stressful conditions). This characterization shows that Pab1 association into these aggregates relies mainly on its RNA recognition motifs (RRM), whose number is important for an efficient recruitment. Although its proline-rich (P) domain and its C-terminus do not directly participate in Pab1 association to stress granules, their presence strengthens or decreases, respectively, the distribution of synthetic Pab1 variants lacking at least one RRM into these aggregates.

Overall, the results of this thesis confirm and add novel demonstrations that the rewiring of yeast physiology has the potential to evoke industrially relevant phenotypes. At the same time, they also confirm that the selection of the target is not trivial and it should be carefully considered to induce a cellular rewiring in the desired direction. Pab1 undoubtedly emerged as a pivotal element that deserves more attention for future strain design and tailoring.

Riassunto

Le attuali produzioni di energia e materiali si basano soprattutto sullo sfruttamento di fonti fossili, ma la loro non rinnovabilità e la necessità di una maggiore sostenibilità ambientale han contribuito ad aumentare l'interesse verso risorse alternative. In questo contesto si collocano le bioraffinerie, che sono volte alla conversione di biomasse rinnovabili in prodotti commerciabili. Per implementare bioindustrie competitive, molte sfide devono essere superate, tra cui il miglioramento delle prestazioni dei microrganismi (*cell factory*) utilizzati nei processi produttivi. Tra tutte, il lievito *Saccharomyces cerevisiae* è la principale *cell factory* per la produzione di bioetanolo come biocarburante.

L'obiettivo principale di questa tesi è stato di ingegnerizzare ceppi di *S. cerevisiae* per ottenere una *cell factory* con caratteristiche rilevanti a livello industriale, tra cui aumentate *performances* di crescita e un incremento della termotolleranza. In particolare, è stata studiata la possibilità di innescare un profondo *rewiring* cellulare, selezionando due proteine: la glutammato sintasi (Glt1), un enzima del metabolismo centrale dell'azoto (CNM) e la principale *poly(A) binding protein* (Pab1), un fattore chiave nel metabolismo dell'mRNA.

Gli effetti fisiologici della delezione e overespressione di *GLT1* sono stati valutati coltivando *S. cerevisiae* con diverse fonti di azoto. I risultati mostrano che l'aggiunta di solfato d'ammonio, glutammato o glutammina, ma non la modulazione dell'espressione di *GLT1*, influenza notevolmente la crescita, il contenuto proteico, la vitalità e l'accumulo di specie reattive dell'ossigeno (ROS). Questi dati evidenziano la plasticità del CNM di *S. cerevisiae* rispetto a variazioni ambientali e confermano la sua robustezza contro perturbazioni interne. Inoltre, sebbene la modulazione dell'espressione di *GLT1* potrebbe non indurre un *rewiring* cellulare, la caratterizzazione fisiologica descritta può essere utile per selezionare altri *target* più promettenti per riprogrammazioni cellulari.

L'induzione del *rewiring* cellulare è stato poi valutato selezionando Pab1 come *target*. A tal scopo, un ceppo recante la copia endogena di *PAB1* è stato

trasformato con una *mutant library* plasmidica di *PAB1* e poi sottoposto a *screening* per isolare ceppi con maggior termotolleranza. I cloni isolati han mostrato, in *drop test*, una maggior crescita ad alte temperature e alte concentrazioni di etanolo. Tra tutte, la variante *PAB1* S40.7 conferisce una maggiore termotolleranza tramite l'espressione dei soli primi 20 amminoacidi di Pab1. Tale fenotipo è stato anche confermato in bioreattore a 40°C. Una possibile ragione della maggior termotolleranza del ceppo S40.7 potrebbe essere collegata ad un minor accumulo di ROS rispetto al ceppo di controllo. Nel complesso, questi risultati han dimostrato che Pab1 è un promettente *target* per indurre fenotipi complessi con tratti migliorati, tra cui una maggiore termotolleranza.

Infine, Pab1 è stata caratterizzata per determinare il ruolo dei suoi domini nel suo reclutamento all'interno di *stress granules* - SG (aggregati di mRNP non tradotti che si formano in condizioni stressanti). Questa caratterizzazione ha mostrato che l'associazione di Pab1 negli SG è principalmente dovuta agli *RNA Recognition Motifs* (RRM), il cui numero è importante per un efficiente reclutamento. Sebbene il dominio ricco di proline (P) e il dominio C-terminale non partecipino direttamente all'associazione di Pab1 negli SG, la loro presenza rafforza o diminuisce, rispettivamente, la localizzazione di varianti sintetiche di Pab1 prive di almeno un RRM in questi aggregati.

Nel complesso, i risultati di questa tesi evidenziano la potenzialità del *rewiring* cellulare nell'indurre fenotipi industrialmente rilevanti. Allo stesso tempo, la selezione di un determinante molecolare ideale per tale scopo non è scontata e deve essere attentamente valutata. In particolare, Pab1 è emerso come un *target* promettente per questo approccio.

Introduction

From fossil-based to bio-based economy: towards a more sustainable society

Nowadays, our economy mainly relies on fossil carbon sources for energy and material production. With the improvement in oil refinery technologies during the 20th century, this dependency strongly increased together with an unprecedented growth in population and in the size of the global economy (Schaffartzik *et al.*, 2014). However, many environmental, social and economic issues arose from oil-based economy. For example, increased greenhouse gases emissions such as carbon dioxide (CO_2), methane (CH_4) and nitrous oxide (N_2O), are strictly correlated with the utilization of fossil resources (Cherubini, 2010). Further, their uneven distribution results in a non-equitable access, which could led to geopolitics problems and poverty in many countries (Schaffartzik *et al.*, 2014). Moreover, being fossil carbon resources non-renewable within the time of their utilization, their depletion could cause price volatility (de Jong and Jungmeier, 2015), as well as risks in their supply due to the increasing demand of energy and goods that is directly correlated with world population increment (Lieder and Rashid, 2016). Altogether, these limitations highlight that this economic model, which is well-described by the "linear economy" of Fig.1A, is no longer sustainable. Therefore, a change of paradigm to a more sustainable economic system that takes into account the environmental, social and economic needs is required (Fig.1B).

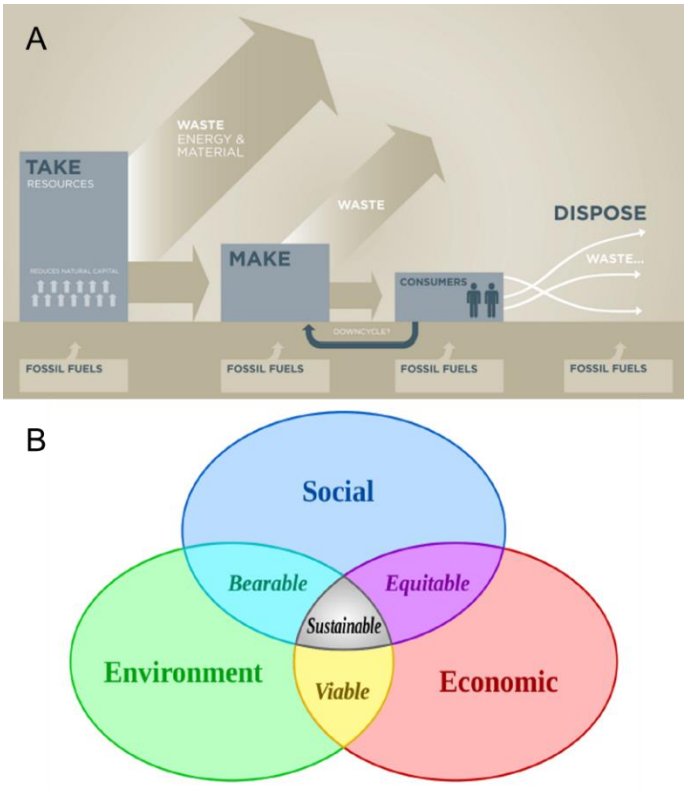


Fig.1 (A) Representation of linear economy (from https://www.ellenmacarthurfoundation.org). (B) Venn diagram of sustainable development redrawn from (Adams, 2006).

In this context, the alternative to the linear fossil-based economy is represented by the circular economy (Fig.2) which aims at *i)* realizing a closed loop material flow in the whole economic system through the 3R approach (Reduce, Reuse, Recycle), *ii)* minimizing environmental deterioration and wastes, *iii)* without restricting the economic growth or social and technical progress (Lieder and Rashid, 2016).

Fig.2 Representation of circular economy (from https://ec.europa.eu)

The concept of circularity is well represented by a bio-based economy, which mainly relies on green chemistry and biorefinery. The latter is defined by the International Energy Agency (IEA) as "the sustainable processing of biomass into a spectrum of marketable products and energy" (de Jong and Jungmeier, 2015). Indeed, differently from fossil sources, biomasses are both more widespread worldwide and renewable within the time of their consumption. Thanks to the latter characteristic, a higher energy security can be achieved, leading to a more sustainable socio-economic growth (Mengal *et al.*, 2017). Moreover, biorefineries have the ability to valorize both industrial and biogenic side streams (*e.g.* CO_2), thus reducing greenhouse gases emission on one side and turning wastes into resources on the other, which are key pillars of circular economy (Dupont-Inglis and Borg, 2017).

Considering these benefits, many actions have been undertaken to develop biorefineries worldwide. For example, a public-private partnership between the European Commission (EC) and the Bio-based Industries Consortium (BIC) originated the Bio-based Industries Joint Undertaking (BBI JU - Fig.3A) in 2014. The purposes are *i)* supporting research and innovation in Europe for establishing bio-based value chains and *ii)* delivering bio-based products superior (or at least comparable) to oil-based ones in terms of price,

performance, availability and sustainability (Mengal *et al.*, 2017). In other words, this partnership wishes to re-industrialize Europe by consolidating already existing biorefineries and by developing new ones. Similarly, the USDA BioPreferred® Program (Fig.3B) aims to facilitate the development and expansion of markets for bio-based products and to create jobs in rural America. Also in China, bio-based industries have seen a considerable expansion in the last decade thanks to several policy supports, such as the *11th "Five-Year Plan" for Bioindustry Development* in 2006 and the more recent *Bioindustry Development Plan* in 2013 (Wang *et al.*, 2017).

Fig.3 A) Bio-based Industries Joint Undertaking (BBI JU) (from https://bbi-europe.eu/) and B) BioPreferred® Program logo (from https://www.biopreferred.gov/BioPreferred/).

Therefore, as also witnessed by positive market trends for bio-based products (Fig.4), biorefineries are currently playing a fundamental role in the development of a more sustainable society, and they are predicted to further increase their importance in the future.

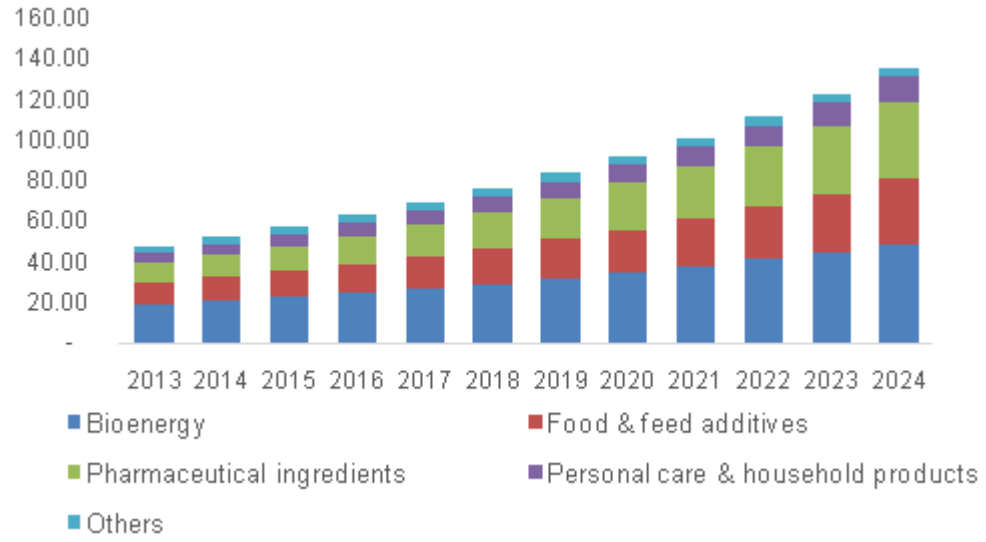


Fig.4 U.S. white biotechnology market volume, by application, 2013 - 2024 (USD Billion) (from White Biotechnology Market Analysis by products, by application and segment forecasts to 2024).

Biorefinery

As described before, a key pillar of bio-based economy is represented by the biorefinery, which aims to produce marketable materials and energy carriers starting from renewable biomasses. Currently, to categorize a biorefinery production process, the most recent and appropriate classification method is based on the "Bioenergy Task 42" developed in 2008 by IEA (de Jong and Jungmeier, 2015). According to this system, a biorefinery can be defined schematically on the basis of three main components: *i)* the starting biomass (also referred as feedstock), *ii)* the platforms and *iii)* the final products (Fig.5) (de Jong and Jungmeier, 2015). These elements are connected by the biomass conversion pathway, which refers to the upstream, midstream, and downstream processing of biomass into platforms and products (de Jong and Jungmeier, 2015).

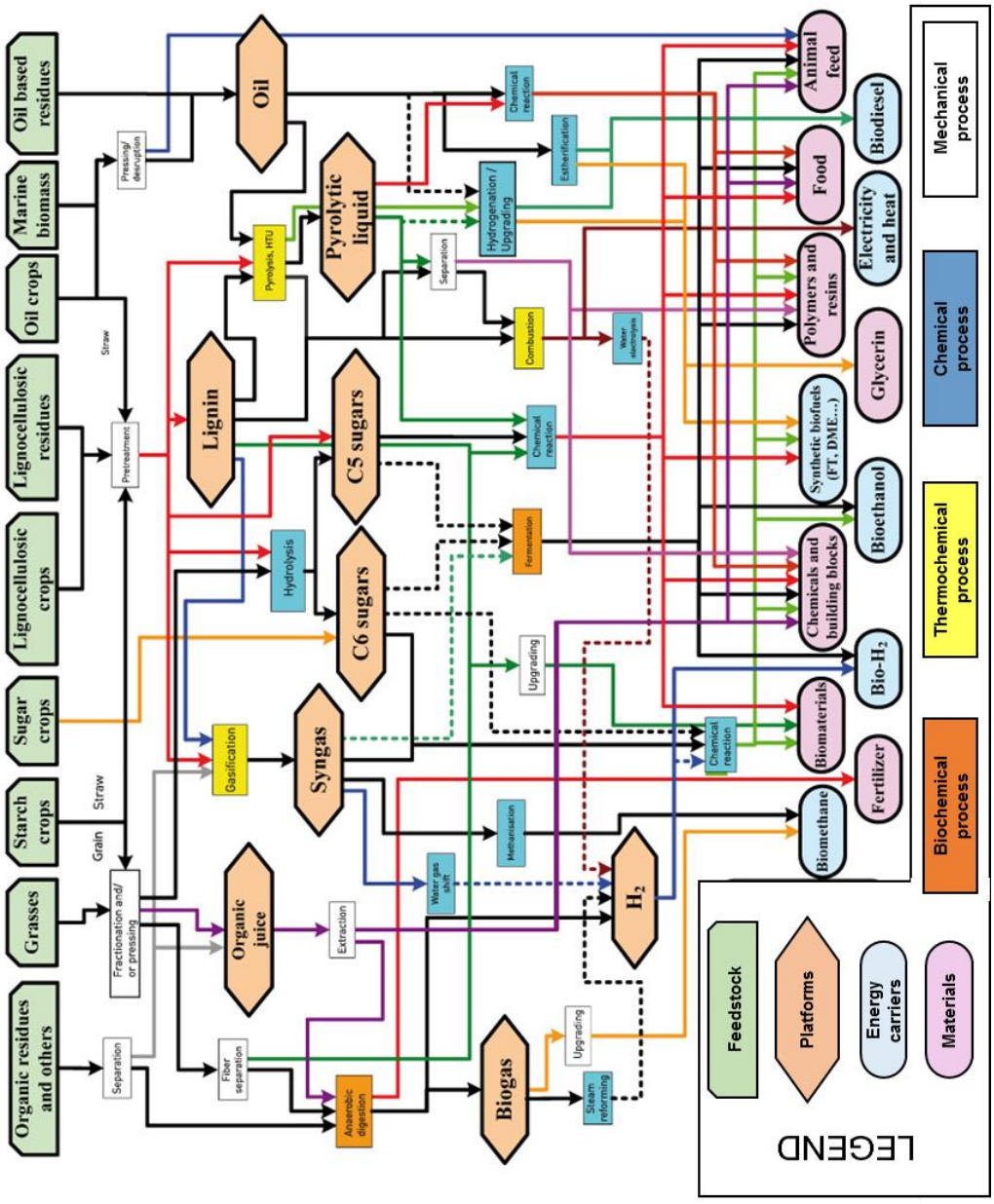


Fig.5 Representation of biorefinery network (redrawn from https://www.iea-bioenergy.task42).

The two main groups of feedstock are "dedicated energy crops" from agriculture (*e.g.* starch crops) and "biomass residues" from agriculture (*e.g.*

straw), forestry (*e.g.* wood chips), industry (*e.g.* process residues), households (*e.g.* municipal solid wastes) and aquaculture (*e.g.* algae) (Cherubini, 2010, de Jong and Jungmeier, 2015).

Platforms are intermediates compounds (sometimes can be already final products) that are obtained from the pretreatment of starting biomasses and are used for the production of final products. For example, C5 and C6 sugars derive from the pretreatment of carbohydrate-rich and lignocellulosic biomasses and are typically used for microorganisms' growth and biofuels production (de Jong and Jungmeier, 2015).

The final products of biorefinery can be mainly classified in materials and energy carriers. Some examples of materials are fine and bulk chemicals, building blocks, organic acids (*e.g.* succinic acid), polymers (*e.g.* plastics), fertilizers, food and animal feed products. Regarding energy carriers, gaseous (*e.g.* biomethane), solid (*e.g.* lignin) and liquid (*e.g.* bioethanol) biofuels, as well as power and heat, are the most representative (Cherubini, 2010).

Finally, biomasses can be converted into platforms and final products according to four types of processes that can be synergistically used: thermochemical (*e.g.* gasification, pyrolysis), chemical (*e.g.* acid hydrolysis), mechanical (*e.g.* fractionation, pressing) and biochemical (de Jong and Jungmeier, 2015). In particular, microorganisms and enzymes represent the core of biochemical processes and, more in general, of white biotechnology, which is defined as their utilization to produce chemicals, materials and bio-energy from renewable resources (Frazzetto, 2003).

A representative example of biorefinery, with the indication of biomasses, platforms, final products and biomass conversion pathway, is shown in Fig.6, which represents a typical bioethanol production process starting from lignocellulosic feedstocks.

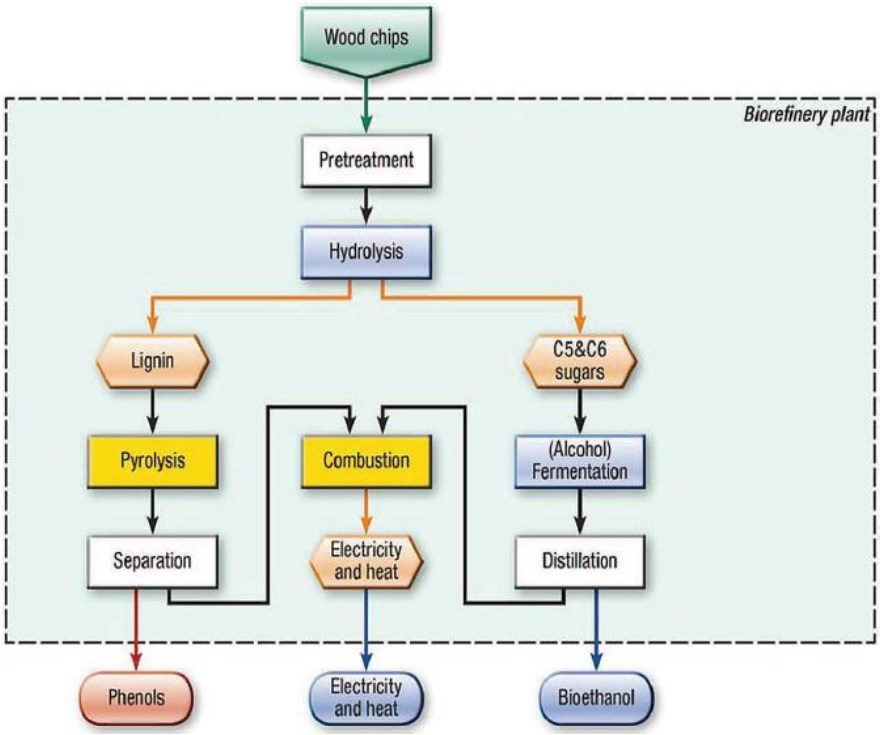


Fig.6 Example of biorefinery for the production of bioethanol from lignocellulose. The process includes one biomass (wood chips), three platforms (C5 and C6 sugars, electricity and heat, lignin) and three products (phenols, electricity and heat, bioethanol). Three types of biomass conversion pathways are used: thermochemical (pyrolysis, combustion), mechanical (pretreatment, separation, distillation) and biochemical (hydrolysis, fermentation) (de Jong and Jungmeier, 2015).

Cell factories

The microorganisms employed in biorefinery are named cell factories because, thanks to their metabolic capabilities, they are able to convert the platforms obtained from biomass pretreatment into products of interest such as native metabolites, heterologous compounds or proteins (Davy *et al.*, 2017). Currently, many novel microorganisms are receiving great attention as promising

for industrial applications, such as microalgae (*e.g. Schizochytrium limacinum*) (Fu *et al.*, 2016) and oleaginous yeasts (*e.g. Lipomyces starkeyi*) (Ageitos *et al.*, 2011) for fatty acids production, or non-*Saccharomyces* yeasts (*e.g. Zygosaccharomyces bailii*, *Kluyveromyces marxianus*) for their unusual tolerance to industrial stresses (Radecka *et al.*, 2015). Nevertheless, only few cell-factories are nowadays well-consolidated and used in industry. For example, the bacteria *Escherichia coli* (Baeshen et al., 2015), the yeasts *Pichia pastoris* (Ahmad *et al.*, 2014) and *Saccharomyces cerevisiae* (Hong and Nielsen, 2012), insect cells (baculovirus expression system) (Sari *et al.*, 2016) and mammalian cells (especially Chinese hamster ovary - CHO) (Wurm, 2004) are widely used for recombinant protein production. The filamentous fungi *Aspergillus niger* and *Aspergillus oryzae* are widely employed for large-scale production of enzymes and organic acids (*e.g.* citric acid) (Park *et al.*, 2017), while the bacteria *E. coli* and *Corynebacterium glutamicum* are used for the industrial amino acid production (D'Este *et al.*, 2017).

Among all, the yeast *S. cerevisiae* is a prominent cell factory for biorefinery applications. Traditionally, this yeast is used for brewing, bakery and winemaking (Kavšček *et al.*, 2015). Nowadays, chemicals for different applications have been produced in engineered strains, such as fuels, bulk chemicals, pharmaceutical and nutraceutical ingredients (Fig.7) (Hong and Nielsen, 2012, Li and Borodina, 2015). Further, *S. cerevisiae* is currently the major workhorse for bio-ethanol production, which is the most widely used biofuel in transportation.

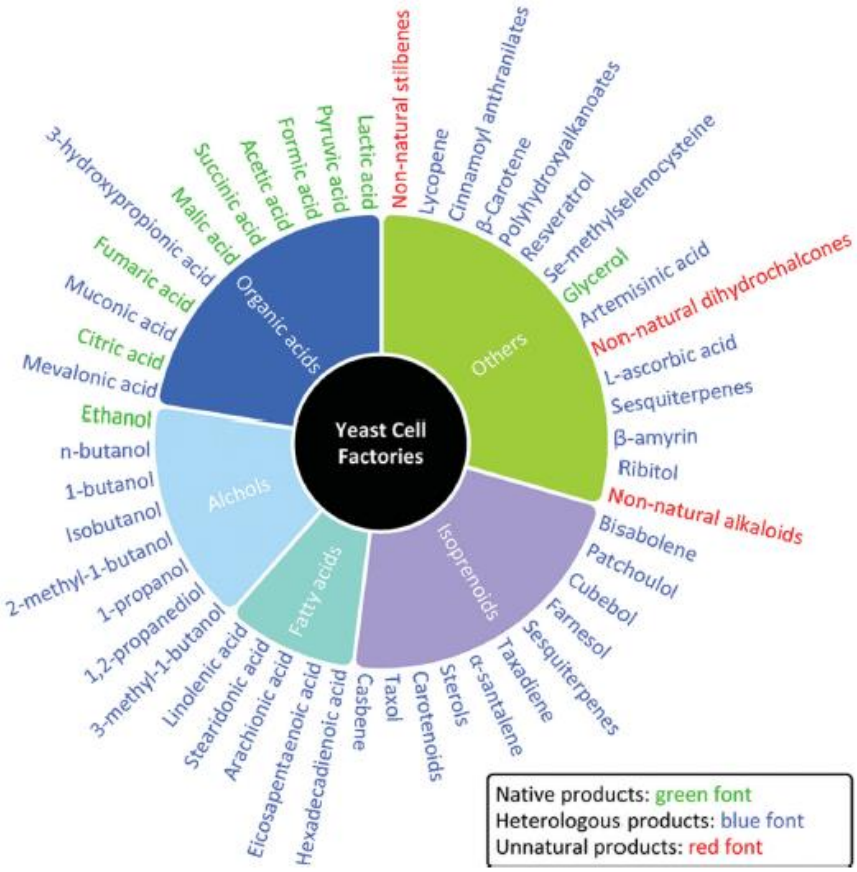


Fig.7 Examples of native, heterologous, and unnatural chemicals that have been produced in yeast (Li and Borodina, 2015)

First and second-generation ethanol production

Currently, bio-ethanol production mainly derives from fermentation of first generation biomasses (*e.g.* cane sugar and hydrolyzed cornstarch), which have the advantages of a high sugar content of the raw materials and their easy conversion into biofuel (Cherubini, 2010). In fact, high ethanol yields on fermentable sugars (> 90% of the theoretical maximum yield), ethanol titers of up to 21% (w/w), and volumetric productivities of 2–3 $kg \cdot m^{-3} \cdot h^{-1}$ are currently

achieved with first-generation production (Jansen *et al.*, 2017). Nevertheless, first-generation bioethanol is in competition with food and feed industries for the feedstock and/or for the use of agricultural land, giving rise to ethical implications. Moreover, their potential availability is restricted by soil fertility and the effective savings of greenhouse gases emissions and fossil energy consumption are limited by the high energy required for crop cultivation and conversion (Cherubini, 2010).

These problems have stimulated the development of second-generation ethanol production from lignocellulosic non-edible plant biomasses and municipal solid wastes. Interestingly, some processes are now at demonstration or full commercial scale (Table. 1) and are mostly based on fermentation by engineered *S. cerevisiae* strains (Jansen *et al*., 2017).

Company/plant	Country (state)	Feedstock	Capacity ML·year^{-1}
DuPont Cellulosic Ethanol LLC—Nevada	USA (IA)	Corn stover	113.6
Poet-DSM Advanced Biofuels LLC—Project Liberty[a]	USA (IA)	Corn cobs/corn stover	75.7
Quad County Cellulosic Ethanol Plant	USA (IA)	Corn fiber	7.6
Fiberight Demonstration Plant	USA (VA)	Waste stream	1.9
ICM Inc. Pilot integrated Cellulosic Biorefinery	USA (MO)	Biomass crops	1.2
American Process Inc.—Thomaston Biorefinery	USA (GA)	Other	1.1
ZeaChem Inc.—demonstration plant	USA (OR)	Biomass crops	1.0
Enerkem Alberta Biofuels LP	Canada (AB)	Sorted municipal solid waste	38.0
Enerkem Inc.—Westbury	Canada (QC)	Woody biomass	5.0
Iogen Corporation	Canada (ON)	Crop residue	2.0
Woodlands Biofuels Inc.—demonstration plant	Canada (ON)	Woody biomass	2.0
GranBio	Brazil	Bagasse	82.4
Raizen	Brazil	Sugarcane bagasse/straw	40.3
Longlive Bio-technology Co. Ltd—commercial demo	China	Corn cobs	63.4
Mussi Chemtex/Beta Renewables	Italy	*Arundo donax*, rice straw, wheat straw	75.0
Borregaard Industries AS—ChemCell Ethanol	Norway	Wood pulping residues	20.0

[a]With expansion of capacity to 94.6 ML per year.

Table 1. Operational commercial-scale plants for second-generation bioethanol production (Jansen *et al.*, 2017).

Many advantages arises with these feedstocks, among which the lack in competition with food and feed industries, the possibility to convert the whole biomass to the final product, a higher energy security in fossil-fuel importing

countries and development of rural economies (Cherubini, 2010, Jansen *et al.*, 2017).

However, some limitations are connected with second-generation feedstocks. Firstly, they are heterogeneous regarding bulk components (*e.g.* carbohydrates, lignin, proteins and oils) and most of the starting material is present in polymeric form (*e.g.* cellulose, starch, hemicellulose, proteins, lignin) (de Jong and Jungmeier, 2015). Therefore, these biomasses require harsher pretreatments than first generation processes to allow sugar release for *S. cerevisiae* fermentation (see below).

Another limitation is represented by stressors associated to the production process. In fact, during fermentation cell factories are exposed to severe conditions such as oxidative, osmotic and high temperatures stresses, as well as the presence of inhibitors in the medium. The latter's are strictly associated to lignocellulosic biomass pretreatments, which often release organic acid (*e.g.* acetic, formic and levulinic acid), furan derivatives (*e.g.* furfural and hydroxymethylfurfural) and phenols (*e.g.* vanillin) in different amounts depending on the starting feedstock (Fig.8) (Ling *et al.*, 2014). Collectively, stressors impair cell viability, growth and, thus, the final ethanol yield, production and productivity (Ling *et al.*, 2014).

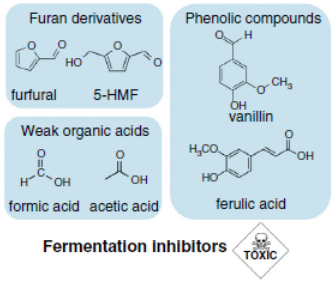


Fig.8 Three major classes of toxic compounds generated during lignocellulosic biomass pretreatment to release single sugars for yeast growth (Ling *et al.*, 2014).

An additional drawback is the difference between the available substrate molecules and the ones effectively exploitable by *S. cerevisiae*. For example, this yeast is not able to grow on xylose and arabinose (two pentose sugars derived from hemicellulose) as carbon sources and, thus, the effective yield of ethanol is restricted (Jansen *et al.*, 2017).

Considering these limitations, over the past two decades, many academia and industrial researches have been done to improve both *S. cerevisiae* as a cell factory and the entire operative process. Improving performances of *S. cerevisiae* during bioethanol production encompasses *i)* the increase of exploitable substrates spectrum, *ii)* the maximization of yeast tolerance to stressors associated to the process and *iii)* the enhancement of its ethanol production capacity (Steensels *et al.*, 2014, Jansen *et al.*, 2017). Rational and random strategies can be synergistically used to achieve these goals (Steensels *et al.*, 2014, Gao *et al.*, 2016). In the next paragraphs, an overview of the main strategies that can be applied to ameliorate yeast performances and therefore the whole process will be described.

Rational strategies for S. cerevisiae strain improvement

Rational approaches rely on genetic modifications in targeted loci to alter cell traits. Key in this strategy is the availability of detailed information about the relevant pathways, enzymes and their regulation. These approaches include genetic engineering, metabolic engineering and synthetic biology (Steensels *et al.*, 2014) that are very synergistic and not easily separable disciplines (Li and Borodina, 2015).

Genetic engineering relies on the principles of recombinant DNA technology and is used to directly manipulate (*e.g.* through deletion and/or over-

expression) the expression of single regulatory proteins, transporters or enzymes (Steensels *et al.*, 2014). This strategy was successfully employed to increase the robustness of *S. cerevisiae* towards different stressors associated to fermentation (Ling *et al.*, 2014). As a representative example, during fermentations on pretreated spruce, the increase of intracellular level of glutathione (*i.e.* an antioxidant tripeptide composed of cysteine, glutamate and glycine) by over-expressing genes involved in its metabolism (*GSH1*, *CYS3* and *GLR1*), improves both ethanol yield and conversion of furfural and HMF to less toxic compounds (Ask *et al.*, 2013).

Metabolic engineering is the field dedicated to design microbial metabolism to convert cheap raw materials into fuels and chemicals (Hong and Nielsen, 2012). As representative examples, it was successfully used to design *S. cerevisiae* strains able to exploit the xylose released after lignocellulosic biomass pretreatment, as well as to develop strains with improved ethanol production. In the first case, xylose-fermenting strains were obtained with the introduction of heterologous pathways for converting D-xylose into D-xylulose by *i)* simultaneous expression of a heterologous xylose reductase (*XR*) and xylitol dehydrogenase (*XDH*) or *ii)* expression of a heterologous xylose isomerase (*XI*) (Fig.9) (Jansen *et al.*, 2017). In the second case, many strategies were investigated to improve ethanol yield, production and productivity, such as the elimination of competitive pathway (*e.g.* glycerol biosynthesis) (Wang *et al.*, 2013) and the unbalance of intracellular redox state (*e.g.* through the engineering of central nitrogen metabolism) (Nissen *et al.*, 2000).

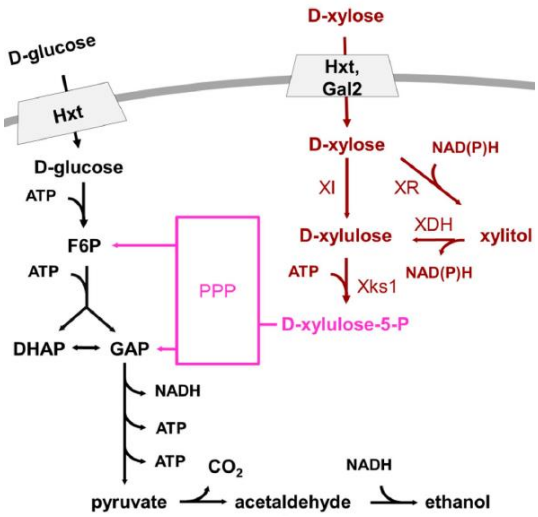


Fig.9 Representation of the heterologous pathways for D-xylose utilization in red, of the native glycolysis and alcoholic fermentation in black and of the native non-oxidative pentose-phosphate pathway (PPP) in magenta (Jansen *et al.*, 2017).

Finally, synthetic biology aims to design and construct new biological parts, devices, and systems (from single genes to completely new organisms), or to re-design existing, natural biological systems for useful purposes (Steensels *et al.*, 2014). In other words, it provides the tools for the optimization of an engineered pathway. To reach this purpose, synthetic biology uses individual genetic elements referred to as parts, preferably synthetized to be assembled and exchanged, therefore also named "biobricks": these are typically DNA fragments, enzymes, localization signals, transcription and translation control elements etc. (Li and Borodina, 2015). They allow, for example, an ameliorated transcription regulation of a pathway, the possibility to bring enzymes close together by scaffolding, compartmentalization or fusion and, finally, the possibility to alter functional properties of native enzymes, regulatory proteins and transporters via rational or random protein engineering (Fig.10) (Li and Borodina, 2015). As representative example, the possibility to engineer *S.*

cerevisiae membrane surface with heterologous cellulases onto chimeric scaffolds, allowed this yeast to directly hydrolyze polymeric cellulose into glucose monomers, thus paving the way for its possible use in consolidated bioprocessing (CBP) (Fan *et al.*, 2012).

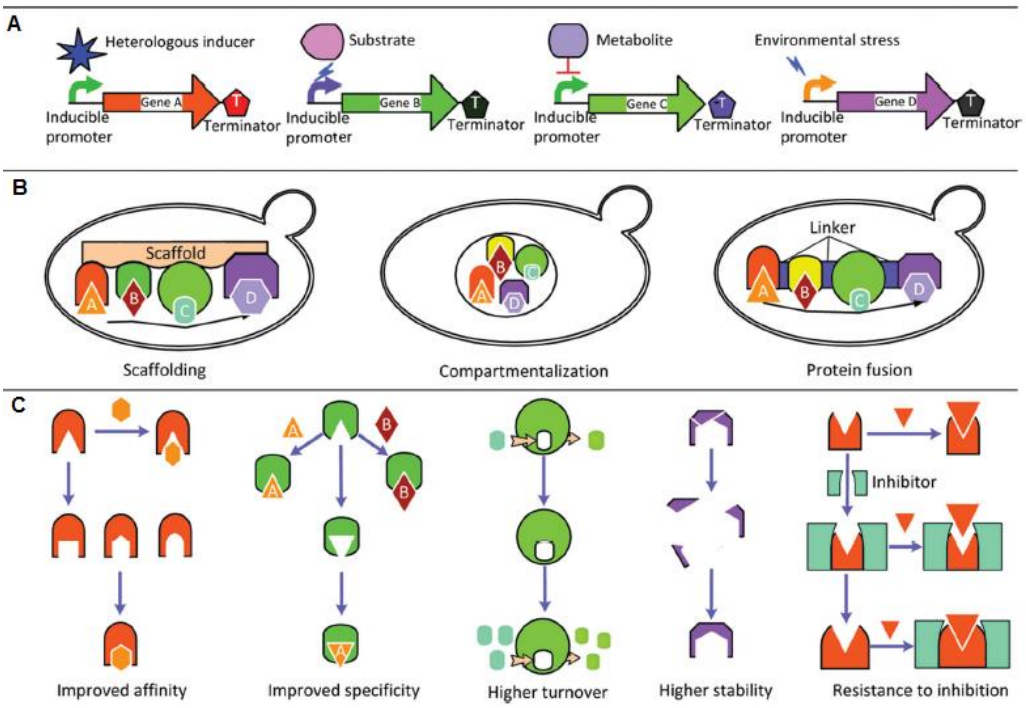


Fig.10 Approaches for optimizing a synthetic pathway. A) Advanced transcription regulation for balancing the levels of individual genes using constitutive regulated promoters. B) Change in enzymes localization by fusion, scaffolding, or expression in specific cellular compartments. C) Modification of enzyme properties by protein engineering (Li and Borodina, 2015).

Random strategies for S. cerevisiae strain improvement

As described before, rational approaches need a full knowledge of the system to perform targeted modifications. However, especially with novel and

still uncharacterized cell factories, these approaches can be of difficult applicability. Random strategies overcome this limitation because they do not necessarily require deepen knowledges of the selected cell factory. Broadly speaking, they consist in the generation of variability in the preferred microorganism followed by screening to select a desired trait and, finally, by inverse metabolic engineering to uncover novel potential target factors (Ling *et al.*, 2014). Random approaches include natural breeding, mutagenesis, adaptive laboratory evolution, genome shuffling and global transcription machinery engineering.

Natural breeding is the process of sexual hybridization (also referred as crossing or mating) used to generate artificial diversity in yeasts. This practice has the advantage of creating novel strains that can be classified as non-GMO for food and beverage applications, because the process of hybridization can occur spontaneously in nature (Steensels *et al.*, 2014). Thanks to technological advances, this process can be applied at large-scale even to promote hybridization across species. In this way, novel superior strains, with optimized particular features, can be obtained by combining selected parents characterized by specific phenotypes (Steensels *et al.*, 2014). As a representative example for bioethanol production, a multiple-tolerant *S. cerevisiae* strain was obtained by hybridizing a haploid strain characterized by high-temperature tolerance, with a haploid strain displaying high-ethanol productivity. After sporulation of the resulting hybrids, spores were screened and some of them showed an ameliorated ethanol productivity at high temperature compared to the parental strains (Benjaphokee *et al.*, 2012).

Mutagenesis, induced by X-rays, UV or chemicals, is a classical method to generate diversity in the entire genome of a strain. However, some drawbacks arise from this technique, among which significant cell-damage (which can often be underestimated during the initial process of selection) and time-consumption.

In fact, several round of mutagenesis are required for improving the performances of a microorganism for a specific trait (Steensels *et al.*, 2014). Therefore, although this technique was successfully used in some cases, such as for the isolation of more thermotolerant yeasts after ethidium bromide mutagenesis (Cha *et al.*, 2015), it is often coupled with other approaches (*e.g.* whole genome shuffling – see below) to improve performances of microorganisms.

Adaptive laboratory evolution (ALE), also referred to evolutionary engineering or directed evolution, is an approach to evolve a microorganism towards the acquisition of biotechnological interesting phenotypes (Dragosits and Mattanovich, 2013). This system can be used also to further optimize specific characteristics of strains created through other methods (*e.g.* genome shuffling – see below). To this purposes, microorganisms are cultured from weeks to years in defined media by repeated batch (Fig.11A) or chemostat (Fig.11B) cultivations under selective pressure. In this way, only those variants carrying mutations (spontaneous or induced with mutagens) that correlate with increased fitness will be positively selected (Dragosits and Mattanovich, 2013). This strategy was successfully employed in *S. cerevisiae* to enhance substrates utilization, products formation and stress tolerance. As a representative example, evolved *S. cerevisiae* strains with improved acetic acid resistance were obtained by serial batch cultivations, with alternating transfers to fresh medium with and without acetic acid (González-Ramos *et al.*, 2016).

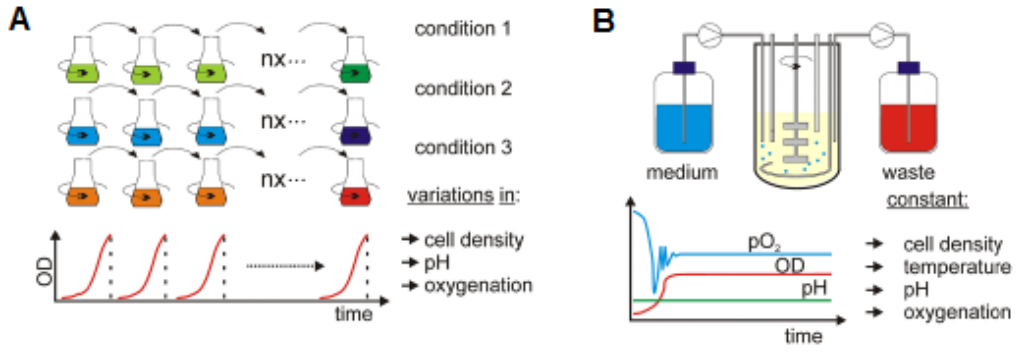


Fig.11 Different ALE strategies. A) serial batch cultures in shake flasks; B) chemostat cultures (Dragosits and Mattanovich, 2013).

Compared to mutagenesis and adaptive laboratory evolution, whole genome shuffling allows a faster and relatively easier phenotypic improvement in microorganisms (Biot-Pelletier and Martin, 2014). This technique consists in recursive recombination events among cells at genome level, starting from a genetically diverse population, which can be naturally available or artificially obtained for example by mutagenesis (Fig.12A). Recombination can be performed by protoplast fusion, sexual recombination, phage-mediated transduction or direct transformation (Fig.12B). After each steps of recombination, screening protocols are applied to the obtained population in order to isolate improved mutants, which can then be further recombined. Each time a mutant is isolated, it will be submitted to characterization (Biot-Pelletier and Martin, 2014). As example, this approach was successfully employed for selecting *S. cerevisiae* strains with improved ethanol tolerance starting from a natural collection of *Saccharomyces* yeasts as genetically diverse population (Snoek *et al.*, 2015).

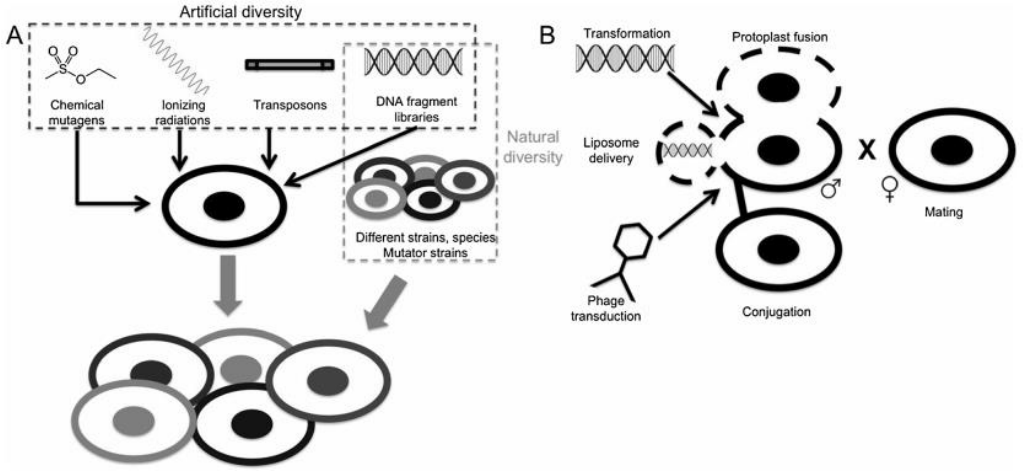


Fig.12 A) Sources of diversity and B) recombination methods for genome shuffling (Biot-Pelletier and Martin, 2014).

Global transcription machinery engineering (gTME) aims at reprogramming the whole transcriptome of cells to unlock complexes phenotypes, among which strains with biotechnological interesting traits (Alper and Stephanopoulos, 2007). In this approach, a mutant library of a "hub" cellular element (*e.g.* a transcription factor that controls transcription of many genes) is constructed and then transformed in a microorganism of interest, with the goal of evoking pleiotropic phenotypic effects (Fig.13). Next, a screening protocol is applied to isolate mutant with the improved desired phenotype (Alper and Stephanopoulos, 2007). This strategy is well-suited for example for improving multiple-stress resistance, which is a complex polygenic traits not easily modifiable with classical approaches (Steensels *et al.*, 2014). As example, it was successfully employed for improving ethanol tolerance and production in *S. cerevisiae*, by choosing the transcription factor Spt15 (Alper *et al.*, 2006) or the Rpb7 subunit of RNA polymerase II (Qiu and Jiang, 2017) as hub elements, and then selecting specific variants able to elicit the desired phenotype.

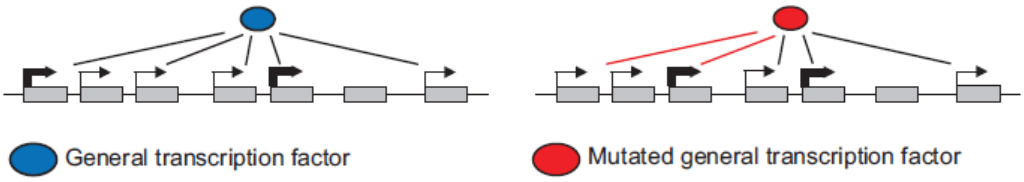


Fig.13 Global transcription machinery engineering (Steensels *et al.*, 2014).

Strategies for whole process improvement

The process of ethanol production from lignocellulosic biomasses can be divided into three main steps: the obtainment of a solution that contains fermentable sugars, the conversion of sugars to ethanol by fermentation and finally the ethanol separation and purification (Mohd Azhar *et al.*, 2017). Each of these steps must be optimize to increase the competitiveness of the whole process.

The first step is critical for the overall performance of the process. Since lignocellulosic biomasses (*e.g.* wheat straw, corn stover, rice straw and sugar cane bagasse) are heterogeneous and recalcitrant to degradation (de Jong and Jungmeier, 2015), severe pretreatments are required to release the sugars contained in the cellulose (*i.e.* polymer of glucose) and in the hemicellulose (*i.e.* polymer of different sugars, among which glucose, mannose, galactose, xylose and arabinose) fractions. In fact, the accessibility to these polysaccharides is limited by lignin, a recalcitrant aromatic polymer that needs to be (at least partially) degraded to ensure cellulose and hemicellulose release for the next saccharification step (Kumar and Sharma, 2017). To break down the lignin structure, the current pretreatment methods employ the use of harsh chemicals and/or physico-chemical conditions that lead to waste generation, environmental pollution and high costs. Furthermore, some of these pretreatments have a negative impact on the enzymes efficiency during the next saccharification and,

as already described before, they can produce inhibitors for microorganisms growth during fermentation (Capolupo and Faraco, 2016). Therefore, intense researches are focusing on the optimization of "green" pretreatment technologies to reduce energy demands, environmental impacts, the use of chemicals and catalysts, formation of inhibitors and wastes and, in parallel, to improve lignocellulose digestibility (Capolupo and Faraco, 2016). Four main types of pretreatments can be employed, among which physical (milling, grinding and extrusion), physico-chemical (steam-explosion, liquid-hot-water, ammonia fiber explosion and supercritical CO_2 explosion), chemical (*e.g.* organosolv, ozonolysis, acid and alkaline treatment) and biological (usage of cellulolytic, hemicellulolytic, and ligninolytic enzymes of fungi or bacteria) (Capolupo and Faraco, 2016).

Regarding the conversion of sugars into ethanol, many process configurations can be used, mainly separate hydrolysis and fermentation (SHF), simultaneous saccharification and fermentation (SSF), simultaneous saccharification and co-fermentation (SSCF) and consolidated bioprocessing (CBP) (Fig.14) (Paulova *et al.*, 2015, Mohd Azhar *et al.*, 2017). In SHF, cellulose and hemicellulose hydrolysis takes place in a separate tank in respect to fermentation. In this way, optimal operative conditions can be set up for both the processes (Paulova *et al.*, 2015, Mohd Azhar *et al.*, 2017). Currently, this configuration is largely used and ethanol yield usually exceeds the 80% of theoretical value (Paulova *et al.*, 2015). Nevertheless, one of the major drawback is the inhibition of hydrolytic enzymes due to the increasing concentrations of the end products (sugars released after the reaction) that lead to incomplete cellulose and hemicellulose hydrolysis; moreover, the need of different tanks increases the investment costs (Paulova *et al.*, 2015). Therefore, coupling hydrolysis of pretreated lignocellulosic material with fermentation in the same tank, with the other configuration processes, allows overall benefits to the process

because of lower cost, no enzymes inhibition (sugars are immediately utilized by yeast as they are produced) and shorter processing time (Paulova *et al.*, 2015, Mohd Azhar *et al.*, 2017). SSF and SSCF are very similar configurations, apart that in the former hexose fermentation is coupled with saccharification and separated from pentose fermentation, while in SSCF there is a co-fermentation of these sugars in the same tank. The main drawback of both SSF and SSCF is the difference between the optimal temperature of enzymes (usually ~50 °C) and that of yeast (~30 °C), which affects the efficiency of both processes and, thus, the final ethanol yields (Wallace-Salinas and Gorwa-Grauslund, 2013, Paulova *et al.*, 2015). Therefore, the availability of thermotolerant yeast to carry out fermentation at temperature higher than 30 °C is desired (Gao *et al.*, 2016). Finally, the CBP is the most intriguing configuration because it is a one-step process in which a feedstock is directly converted into a product by one microorganism or by a microbial consortium, without requiring pre-treatment of the biomass (Paulova et al., 2015). The most difficult task in CBP regards the cell factory, which must contemporary express appropriate hydrolytic enzymes and produce ethanol. Different microorganisms were tested for this purpose, among which cellulolytic thermophilic bacteria (*e.g. Clostridium thermocellum*), cellulolytic fungi (*e.g. Trichoderma reesei*) and engineered ethanol producers (*e.g. S. cerevisiae*), both alone and in consortium. However, the competitiveness of the CBP with traditional *S. cerevisiae* ethanol production is far too low, especially regarding process productivity (Paulova *et al.*, 2015).

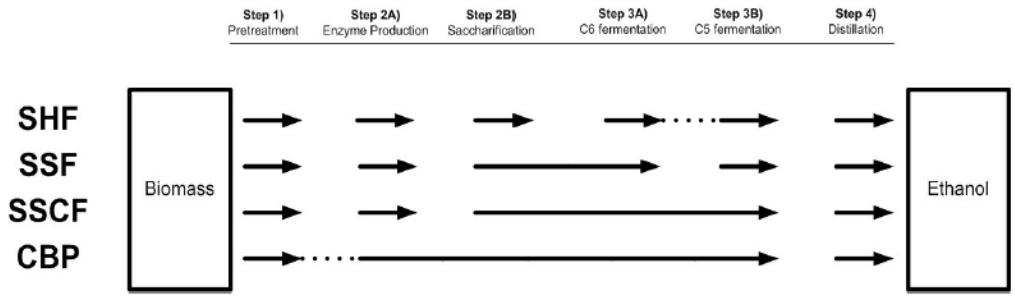


Fig.14 Different configurations of an ethanol production process (Scully and Orlygsson, 2015)

The final step to obtain ethanol is the distillation of the stream that comes out from the fermenters. Once ethanol is recovered, the semi-solid residue from the distillation process (named whole stillage), is separated into a liquid stream (thin stillage) and a soil cake using a filter press; the cake is a fraction enriched in lignin, which is often burned to provide additional energy to the whole plant (Valdivia *et al.*, 2016). Currently, to further optimize the entire process and reduce costs, attempts to develop efficient *in-situ* vacuum distillation (carried out at low temperatures and pressures) are performed, because of the possibility to reuse cellulases (one of the most expensive components) and yeast cells in the stillage for other rounds of fermentation (Zhang *et al.*, 2017).

Scope of the thesis

As extensively described in this introduction, biorefineries can provide a substantial contribution to the development of a more sustainable society. However, many hurdles must be overcome to allow both the implementation of bio-based industries and their competitiveness, among which the development of cell factories characterized by industrially relevant phenotypes. To this purpose, a possible strategy is to rewire the entire cellular metabolism to evoke complex

phenotypes, among which strains with traits that can match biotechnological needs. For this purpose, in this thesis two targets were investigated: the glutamate synthase (Glt1) and the major poly(A) binding protein (Pab1) of the yeast *S. cerevisiae*.

In the first case, the effects of *GLT1* deletion and over-expression on cell physiology were assessed aiming to verify whether this gene could be a possible candidate for the induction of a cellular rewiring. Indeed, since it encodes for an enzyme of the central nitrogen metabolism (a metabolic node through which almost all nitrogen sources are catabolized) (Magasanik and Kaiser, 2002), whose activity is NADH-dependent, the modulation of its expression was supposed to induce changes in the cellular redox state and physiology. Inorganic and organic nitrogen sources were supplied to the fermentation media and their effects on cells were assessed. This is a relevant topic in biorefinery, since the nature of the nitrogen source can exerts a considerable effects on final product production (Tesfaw and Assefa, 2014).

In the second case, Pab1 was selected as a target for obtaining thermotolerant yeast strains, a desirable trait to improve the productivity of a bioethanol production process from lignocellulose feedstocks. The rationale of choosing *PAB1* as a target came from previous studies conducted in our laboratory (Martani *et al.*, 2015), which demonstrated that the mild over-expression of this gene and the screening of a *PAB1* mutant library allowed the isolation of yeast strains with improved robustness towards acetic acid, an inhibitor in lignocellulosic hydrolysates. Hence, in this thesis we repeated the approach but oriented to the selection of thermotolerant strains, not only to retrieve novel promising hosts, but also to confirm that Pab1 can be a powerful candidate for obtaining industrially relevant phenotypes.

In parallel, the contribution of Pab1 domains (four RNA recognition motifs, a proline-rich and a C-terminal domains) in the protein recruitment within

stress granules (aggregates of mRNPs stalled in the process of translation initiation) was investigated. In fact, despite the availability of many information about protein domains' interactions and functions, this aspect has never been explored. This work can be of future utility in case of considering the different Pab1 domains as parts for synthetic biology purposes.

Sitography

https://www.ellenmacarthurfoundation.org

https://ec.europa.eu

https://bbi-europe.eu

https://www.biopreferred.gov/BioPreferred

References

Adams WM (2006) The future of sustainability: re-thinking environment and development in the twenty-first century. *Report of the IUCN Renowned Thinkers Meeting*, 29-31 January 2006

Ageitos JM, Vallejo JA, Veiga-Crespo P and Villa TG (2011) Oily yeasts as oleaginous cell factories. *Appl Microbiol Biotechnol* **90**: 1219-1227.

Ahmad M, Hirz M, Pichler H and Schwab H (2014) Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. *Appl Microbiol Biotechnol* **98**: 5301-5317.

Alper H and Stephanopoulos G (2007) Global transcription machinery engineering: a new approach for improving cellular phenotype. *Metab Eng* **9**: 258-267.

Alper H, Moxley J, Nevoigt E, Fink GR and Stephanopoulos G (2006) Engineering yeast transcription machinery for improved ethanol tolerance and production. *Science* **314**: 1565-1568.

Ask M, Mapelli V, Höck H, Olsson L and Bettiga M (2013) Engineering glutathione biosynthesis of *Saccharomyces cerevisiae* increases robustness to inhibitors in pretreated lignocellulosic materials. *Microb Cell Fact* **12**: 87.

Baeshen MN, Al-Hejin AM, Bora RS, Ahmed MM, Ramadan HA, Saini KS, Baeshen NA and Redwan EM (2015) Production of Biopharmaceuticals in *E. coli*: Current Scenario and Future Perspectives. *J Microbiol Biotechnol* **25**: 953-962.

Benjaphokee S, Hasegawa D, Yokota D, Asvarak T, Auesukaree C, Sugiyama M, Kaneko Y, Boonchird C and Harashima S (2012) Highly efficient bioethanol production by a *Saccharomyces cerevisiae* strain with multiple stress tolerance to high temperature, acid and ethanol. *N Biotechnol* **29**: 379-386.

Biot-Pelletier D and Martin VJ (2014) Evolutionary engineering by genome shuffling. *Appl Microbiol Biotechnol* **98**: 3877-3887.

Capolupo L and Faraco V (2016) Green methods of lignocellulose pretreatment for biorefinery development. *Appl Microbiol Biotechnol* **100**: 9451-9467.

Cha Y-L, An GH, Yang J, Moon Y-H, Yu G-D and Ahn J-W (2015) Bioethanol production from Miscanthus using thermotolerant *Saccharomyces cerevisiae* mbc 2 isolated from the respiration-deficient mutants. *Renewable Energy* **80**: 259-265.

Cherubini F (2010) The biorefinery concept: Using biomass instead of oil for producing energy and chemicals. *Energy Conversion and Management* **51**: 1412-1421.

D'Este M, Alvarado-Morales M and Angelidaki I (2017) Amino acids production focusing on fermentation technologies - A review. *Biotechnol Adv*. 2017.09.001

Davy AM, Kildegaard HF and Andersen MR (2017) Cell Factory Engineering. *Cell Syst* **4**: 262-275.

de Jong E and Jungmeier G (2015) Chapter 1 - Biorefinery Concepts in Comparison to Petrochemical Refineries A2 - Pandey, Ashok. *Industrial Biorefineries and White Biotechnology*, (Höfer R, Taherzadeh M, Nampoothiri KM and Larroche C, eds.), p.^pp. 3-33. Elsevier, Amsterdam.

Dragosits M and Mattanovich D (2013) Adaptive laboratory evolution -- principles and applications for biotechnology. *Microb Cell Fact* **12**: 64.

Dupont-Inglis J and Borg A (2017) Destination bioeconomy - The path towards a smarter, more sustainable future. *N Biotechnol*.

Fan LH, Zhang ZJ, Yu XY, Xue YX and Tan TW (2012) Self-surface assembly of cellulosomes with two miniscaffoldins on *Saccharomyces cerevisiae* for cellulosic ethanol production. *Proc Natl Acad Sci U S A* **109**: 13260-13265.

Frazzetto G (2003) White biotechnology. *EMBO Rep* **4**: 835-837.

Fu W, Chaiboonchoe A, Khraiwesh B, Nelson DR, Al-Khairy D, Mystikou A, Alzahmi A and Salehi-Ashtiani K (2016) Algal Cell Factories: Approaches, Applications, and Potentials. *Mar Drugs* **14**: 225

Gao L, Liu Y, Sun H, Li C, Zhao Z and Liu G (2016) Advances in mechanisms and modifications for rendering yeast thermotolerance. *J Biosci Bioeng* **121**: 599-606.

González-Ramos D, Gorter de Vries AR, Grijseels SS, van Berkum MC, Swinnen S, van den Broek M, Nevoigt E, Daran JM, Pronk JT and van Maris AJ (2016) A new laboratory evolution approach to select for constitutive acetic acid tolerance in *Saccharomyces cerevisiae* and identification of causal mutations. *Biotechnol Biofuels* **9**: 173.

Hong KK and Nielsen J (2012) Metabolic engineering of *Saccharomyces cerevisiae*: a key cell factory platform for future biorefineries. *Cell Mol Life Sci* **69**: 2671-2690.

Jansen MLA, Bracher JM, Papapetridis I, Verhoeven MD, de Bruijn H, de Waal PP, van Maris AJA, Klaassen P and Pronk JT (2017) *Saccharomyces cerevisiae* strains for second-generation ethanol production: from academic exploration to industrial implementation. *FEMS Yeast Res* **17**. fox044.

Kavšček M, Stražar M, Curk T, Natter K and Petrovič U (2015) Yeast as a cell factory: current state and perspectives. *Microb Cell Fact* **14**: 94.

Kumar AK and Sharma S (2017) Recent updates on different methods of pretreatment of lignocellulosic feedstocks: a review. *Bioresour Bioprocess* **4**: 7.

Li M and Borodina I (2015) Application of synthetic biology for production of chemicals in yeast *Saccharomyces cerevisiae*. *FEMS Yeast Res* **15**: 1-12.

Lieder M and Rashid A (2016) Towards circular economy implementation: a comprehensive review in context of manufacturing industry. *Journal of Cleaner Production* **115**: 36-51.

Ling H, Teo W, Chen B, Leong SS and Chang MW (2014) Microbial tolerance engineering toward biochemical production: from lignocellulose to products. *Curr Opin Biotechnol* **29**: 99-106.

Magasanik B and Kaiser CA (2002) Nitrogen regulation in *Saccharomyces cerevisiae*. *Gene* **290**: 1-18.

Martani F, Marano F, Bertacchi S, Porro D and Branduardi P (2015) The *Saccharomyces cerevisiae* poly(A) binding protein Pab1 as a target for eliciting stress tolerant phenotypes. *Sci Rep* **5**: 18318.

Mengal P, Wubbolts M, Zika E, Ruiz A, Brigitta D, Pieniadz A and Black S (2017) Bio-based Industries Joint Undertaking: The catalyst for sustainable bio-based economic growth in Europe. *N Biotechnol*.

Mohd Azhar SH, Abdulla R, Jambo SA, Marbawi H, Gansau JA, Mohd Faik AA and Rodrigues KF (2017) Yeasts in sustainable bioethanol production: A review. *Biochem Biophys Rep* **10**: 52-61.

Nissen TL, Kielland-Brandt MC, Nielsen J and Villadsen J (2000) Optimization of ethanol production in *Saccharomyces cerevisiae* by metabolic engineering of the ammonium assimilation. *Metab Eng* **2**: 69-77.

Park HS, Jun SC, Han KH, Hong SB and Yu JH (2017) Diversity, Application, and Synthetic Biology of Industrially Important *Aspergillus* Fungi. *Adv Appl Microbiol* **100**: 161-202.

Paulova L, Patakova P, Branska B, Rychtera M and Melzoch K (2015) Lignocellulosic ethanol: Technology design and its impact on process efficiency. *Biotechnol Adv* **33**: 1091-1107.

Qiu Z and Jiang R (2017) Improving *Saccharomyces cerevisiae* ethanol production and tolerance via RNA polymerase II subunit Rpb7. *Biotechnol Biofuels* **10**: 125.

Radecka D, Mukherjee V, Mateo RQ, Stojiljkovic M, Foulquié-Moreno MR and Thevelein JM (2015) Looking beyond *Saccharomyces*: the potential of non-conventional yeast species for desirable traits in bioethanol fermentation. *FEMS Yeast Res* **15**. fov053.

Sari D, Gupta K, Thimiri Govinda Raj DB, Aubert A, Drncová P, Garzoni F, Fitzgerald D and Berger I (2016) The MultiBac Baculovirus/Insect Cell Expression Vector System for Producing Complex Protein Biologics. *Adv Exp Med Biol* **896**: 199-215.

Schaffartzik A, Mayer A, Gingrich S, Eisenmenger N, Loy C and Krausmann F (2014) The global metabolic transition: Regional patterns and trends of global material flows, 1950-2010. *Glob Environ Change* **26**: 87-97.

Scully SM, Orlygsson J (2015) Recent Advances in Second Generation Ethanol Production by Thermophilic Bacteria. *Energies* **8**: 1-30.

Snoek T, Picca Nicolino M, Van den Bremt S, Mertens S, Saels V, Verplaetse A, Steensels J and Verstrepen KJ (2015) Large-scale robot-assisted genome shuffling yields industrial *Saccharomyces cerevisiae* yeasts with increased ethanol tolerance. *Biotechnol Biofuels* **8**: 32.

Steensels J, Snoek T, Meersman E, Picca Nicolino M, Voordeckers K and Verstrepen KJ (2014) Improving industrial yeast strains: exploiting natural and artificial diversity. *FEMS Microbiol Rev* **38**: 947-995.

Tesfaw A and Assefa F (2014) Current Trends in Bioethanol Production by *Saccharomyces cerevisiae*: Substrate, Inhibitor Reduction, Growth Variables, Coculture, and Immobilization. *Int Sch Res Notices* **2014**: 532852.

Valdivia M, Galan JL, Laffarga J and Ramos JL (2016) Biofuels 2020: Biorefineries based on lignocellulosic materials. *Microb Biotechnol* **9**: 585-594.

Wallace-Salinas V and Gorwa-Grauslund MF (2013) Adaptive evolution of an industrial strain of *Saccharomyces cerevisiae* for combined tolerance to inhibitors and temperature. *Biotechnol Biofuels* **6**: 151.

Wang J, Liu W, Ding W, Zhang G and Liu J (2013) Increasing ethanol titer and yield in a *gpd1Δ gpd2Δ* strain by simultaneous overexpression of *GLT1* and *STL1* in *Saccharomyces cerevisiae*. *Biotechnol Lett* **35**: 1859-1864.

Wang R, Cao Q, Zhao Q and Li Y (2017) Bioindustry in China: An overview and perspective. *N Biotechnol*.

White Biotechnology Market Analysis By Product (Biofuels, Biomaterials, Biochemicals, Industrial Enzymes), By Application (Bioenergy, Food and Feed Additives, Pharmaceutical ingredients, Personal Care and Household Products) And Segment Forecasts To 2024 - Published: October 2016 - Report ID: GVR-1-68038-135-1 4

Wurm FM (2004) Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat Biotechnol* **22**: 1393-1398.

Zhang J, Lei C, Liu G, Bao Y, Balan V and Bao J (2017) In-Situ Vacuum Distillation of Ethanol Helps To Recycle Cellulase and Yeast during SSF of Delignified Corncob Residues. *ACS Sustainable Chemistry and Engineering*.

Could the modulation of GLT1 expression rewire yeast physiology?

Previous studies in our laboratory showed that *GLT1*, encoding for the glutamate synthase enzyme (Glt1) of the central nitrogen metabolism (CNM), played an important role in *S. cerevisiae* cellular metabolism. Indeed, its deletion provoked an accumulation of succinate, ethanol and glycerol in a BY4741 strain cultured in glutamate-supplemented medium (unpublished data). However, a great variability in growth and metabolic profiles between the wild type and the *glt1Δ* strains was obtained, depending on the strategy used to complement BY4741 auxotrophies. When strains were complemented with integrative plasmids for uracil and methionine and multicopy plasmids for leucine and histidine, *GLT1* deletion improved growth in glutamate-supplemented medium, while the opposite situation occurred in the presence of ammonium sulfate as nitrogen source. Conversely, no differences in growth were detected when strains were complemented with sole integrative plasmids, independently on the nitrogen source supplied for fermentation. Moreover, strains are characterized by a huge clonal variability, because a targeted integration of plasmids is impossible to achieve in the BY4741 genetic background (its auxotrophies are obtained by full gene deletions). Therefore, to better elucidate the role of *GLT1* in cell physiology, the CEN.PK 102-5B genetic backgrounds was used in the following chapter, because its auxotrophies are due to point mutations and can be therefore used as target for plasmid(s) integration.

The physiological effects on prototrophic *GLT1* deleted and over-expressing strains were studied to evaluate whether cellular rewiring could be triggered by the modulation of its expression. In parallel, inorganic and organic nitrogen sources were supplied to the fermentation media to better understand their influence on cell growth and physiology.

Chapter 1

Physiological effects of *GLT1* modulation in *Saccharomyces cerevisiae* strains growing on different nitrogen sources

Published in

JOURNAL OF MICROBIOLOGY AND BIOTECHNOLOGY

ISSN:1017-7825 vol. 26 (2)

Brambilla Marco‡; Adamo Giusy Manuela‡; Frascotti Gianni; Porro Danilo; Branduardi Paola. (2016)

‡ These authors contributed equally to the work

DOI:10.4014/jmb.1508.08002. pp.326-336.

Introduction

Natural and engineered yeast cell factories are today extensively used for commercial productions (Hong and Nielsen, 2012). Based upon their innate metabolic abilities, yeasts have been employed since several decades for large-scale production of different natural compounds. In this respect, there are more than 600 yeast factories in operation in the world (Branduardi and Porro 2012). Furthermore, with the advent of recombinant DNA technology, it has become possible to introduce traits for the production of desired non-natural compounds: heterologous proteins and metabolites (Porro *et al.*, 2011). The scientific and technological platforms leading to the production of recombinant proteins seem under consolidation, with the exception of membrane proteins, with *Escherichia coli* and *Saccharomyces cerevisiae* being the two microbial workhorses mainly exploited for products under commercialization. Conversely, the production of heterologous and endogenous metabolites by engineered cell factories always and strongly suffers from extensive regulation of cellular metabolism, which easily evolves to ensure robustness. Indeed, the main research effort today is probably related to the design of robust and stable strains to match the limiting conditions often occurring during industrial fermentation. In this respect, central carbon metabolism (CCM), central nitrogen metabolism (CNM) and energy metabolism (EM, which includes redox metabolism), including their interconnections, play a crucial role for every microbial process and are emerging targets for eliciting profound cellular rewiring (see, as example (Edwards *et al.*, 2011)).

Non-engineered *S. cerevisiae* strains are mainly addicted to monomeric hexose sugars as carbon source, glucose being the favorite one, while the spectrum of nitrogen sources is wider. CNM in *S. cerevisiae* (Guillamón *et al.*,

2001, Magasanik and Kaiser, 2002, Ljungdahl and Daignan-Fornier, 2012) is based on five important enzymatic reactions (Fig. 1). The first is catalyzed by glutamate dehydrogenase 1 (Gdh1p), the *GDH1* gene product, which converts α-ketoglutarate and ammonia in glutamate, oxidizing NADPH:

$$\alpha\text{-Ketoglutarate} + NH4^{+} + NADPH + H^{+} \leftrightarrow \text{Glutamate} + NADP^{+}$$

Another isoform of this enzyme is Gdh3p, encoded by *GDH3*; its expression is negatively regulated in the presence of glucose, and is induced in the presence of ethanol (DeLuna *et al.*, 2001).

The second reaction is catalyzed by glutamine synthetase (Gln1p – GS), the *GLN1* gene product, which converts glutamate and ammonia in glutamine, consuming one ATP:

$$\text{Glutamate} + NH4^{+} + ATP \rightarrow \text{Glutamine} + ADP + Pi$$

This is the sole reaction synthesizing glutamine in the cells; as a consequence, the lack of this enzymatic activity results in glutamine auxotrophic strains (Magasanik and Kaiser, 2002).

Glutamate dehydrogenase 2 (Gdh2p), the *GDH2* gene product, catalyzes the reaction opposite to Gdh1p:

$$\text{Glutamate} + NAD^{+} \leftrightarrow \alpha\text{-Ketoglutarate} + NH4^{+} + NADH + H^{+}$$

When glutamate is utilized as the sole nitrogen source, cells can obtain the required ammonia to synthesize glutamine only through the reaction catalyzed by Gdh2p (Magasanik and Kaiser, 2002).

Glutamate synthase (Glt1p, also named as GOGAT), the *GLT1* gene product, converts α-ketoglutarate and glutamine into two molecules of glutamate, oxidizing NADH (Nissen *et al.*, 2000, Magasanik and Kaiser, 2002):

$$\alpha\text{-Ketoglutarate} + \text{Glutamine} + \text{NADH} + H^+ \rightarrow 2\ \text{Glutamate} + NAD^+$$

Finally, glutaminases A and B catalyze the deamination of glutamine to glutamate and ammonium (Soberón and González, 1987, Guillamón *et al.*, 2001).

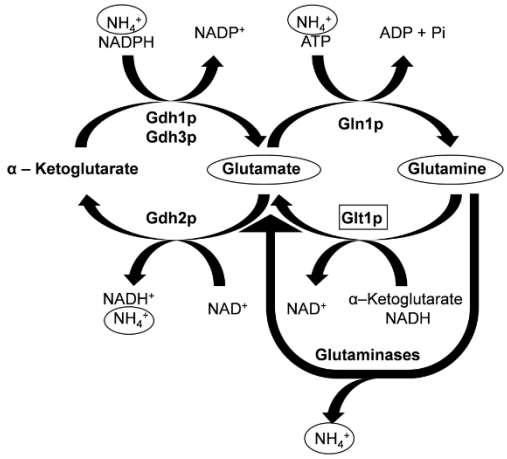


Fig.1 Representation of central nitrogen metabolism in *S. cerevisiae* (redrawn from (Magasanik and Kaiser, 2002)).

In addition to regulations operating at the transcriptional level, the redox state of the cell contributes to determine the equilibrium among these reactions.

The main purpose of this work was to analyze the effects of modulation of one of the key elements of CNM. More in detail, we analyzed *glt1Δ* and *GLT1* overexpressing CEN.PK strains during batch-flask growth on ammonium, glutamate, or glutamine, in aerobic conditions. To the best of our knowledge, this is the first description of the physiology of *S. cerevisiae* strains that exclusively modulate *GLT1* expression.

Results

Growth of wild-type, *glt1Δ*, and *GLT1* overexpressing strains on media containing $(NH_4)_2SO_4$, glutamate, or glutamine

To characterize the effects of internal and external modulations in one of the key enzyme (glutamate synthase, encoded by *GLT1*) of the CNM on the growth properties of *S. cerevisiae*, we grew the wild type CEN.PK *C*, the CEN.PK *glt1Δ*, and the CEN.PK *TPI-GLT1* strains in flask-batch culture on $(NH_4)_2SO_4$, glutamate, or glutamine as the nitrogen source. The nitrogen source was supplied at two different sets of concentrations, normalized to release 5.1 or 30.6 mM of $NH4^+$ (see Materials and Methods), to evaluate the effects of a low and a very high nitrogen amount on cellular growth; in particular, 30.6 mM is the highest possible concentration based on maximum glutamine solubility. Glucose was always supplied at 150 g/l, reported as the maximum concentration beyond which growth inhibition occurs (Murthy *et al.*, 2012). A high glucose concentration was selected since this is the condition very often applied when yeasts are employed as cell factories for the industrial production of biobased chemicals.

As initial control, we determined the activity of Glt1p for cells in exponential phase of growth in aerobic conditions. Since Glt1p localization is still dubious (Moreira dos Santos *et al.*, 2003), the enzymatic assay was performed on both soluble and organelle-associated protein fractions, but the activity was detectable only in the first one. Data are summarized in Fig. 2. In agreement with literature data (Valenzuela *et al.*, 1998), the activity appeared only slightly down-regulated for wild-type strains growing in glutamate and glutamine, when compared with cells growing on ammonium (black columns). *GLT1* overexpression was functional, as confirmed by fluorescence microscope

observation of CEN.PK *GLT1-GFP* and CEN.PK *TPI-GLT1-GFP* (data not shown) and by the higher measured activity (Fig. 2, white columns), but with no significant differences among the different nitrogen sources.

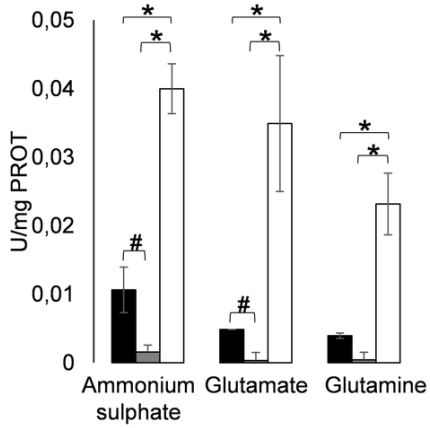


Fig.2 Glutamate synthase activity. Assays were performed on the soluble protein fraction of CEN.PK *C* (black columns), CEN.PK *glt1Δ* (grey columns), and CEN.PK *TPI-GLT1* (white columns). Values represent the average and the standard deviation of three independent experiments (*$p \leq 0.01$; #$p \leq 0.05$, Student's *t*-test).

Notably, both the deletion and the overexpression of the *GLT1* gene did not determine remarkable differences on the growth profiles (data not shown).

The observed differences are essentially ascribable to the different nitrogen source. As a representative example, Fig. 3 shows the growth kinetics of the wild-type CEN.PK *C* strain during aerobic growth on low (Panel A) or high (Panel B) nitrogen concentration. Yeast cells reached the highest optical density in the presence of high quantities of glutamine, whereas the lowest OD was obtained in the presence of ammonium sulfate. Moreover, when glutamate was supplemented, cells showed different trends of growth depending on the amino acid concentration, growing significantly faster when it was supplied at the lower concentration (Panel A versus Panel B).

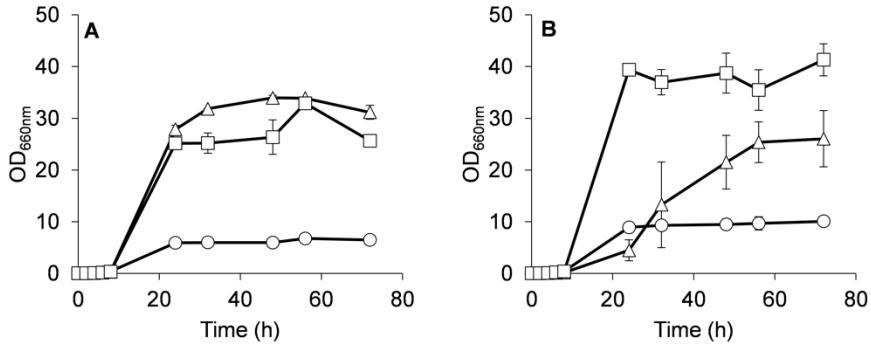


Fig.3 Growth curve of CEN.PK *C* strain (representative also for the *GLT1*-deleted and -overexpressing strains). Strains were cultivated in aerobic conditions and supplemented with ammonium sulfate (circle), glutamate (triangle), or glutamine (square), which were supplied to release (A) 5.1 mM or (B) 30.6 mM of ammonium. Results are the average and the standard deviation of three independent experiments.

Forward scatter and protein content of wild-type, *glt1Δ*, and *GLT1* overexpressing strains during growth on $(NH_4)_2SO_4$, glutamate or glutamine

For all the experiments described in the rest of the paper, the wild-type CEN.PK *C*, the CEN.PK *glt1Δ*, and the CEN.PK *TPI-GLT1* strains were grown in aerobic conditions in flask-batch culture on ammonium, glutamate, and glutamine supplied to release 30.6 mM of ammonium, as described in Fig. 3B. For cells harvested both in the exponential and stationary phases of growth, we determined the cell protein content and the cell volume at the single cell level by flow cytometry (Fig. 4). The first parameter was estimated using FITC, a typical marker of proteins (Miller and Quarles, 1990) that binds their N-terminal amine group, while the cell volume is related to forward scatter signal (FSC). Indeed this parameter measures the light scattered when each cell passes through the laser beam, providing a measurement that is related to cell shape, cell orientation, and cell volume.

As for the data described in the previous paragraph, remarkable differences are ascribable to the different media and not to the modulation of *GLT1*. The lower FITC and FSC signals were registered during growth on glutamine, and the higher FSC signals for cells grown in glutamate-supplemented medium (Fig. 4). This indicates that the cell dimension is minimal when cells are growing in the presence of glutamine and maximal in the presence of glutamate, under the tested conditions. Independently from the nitrogen source, both cell sizes and cell volume contents tend to increase in the stationary phase of growth (Panels B-D).

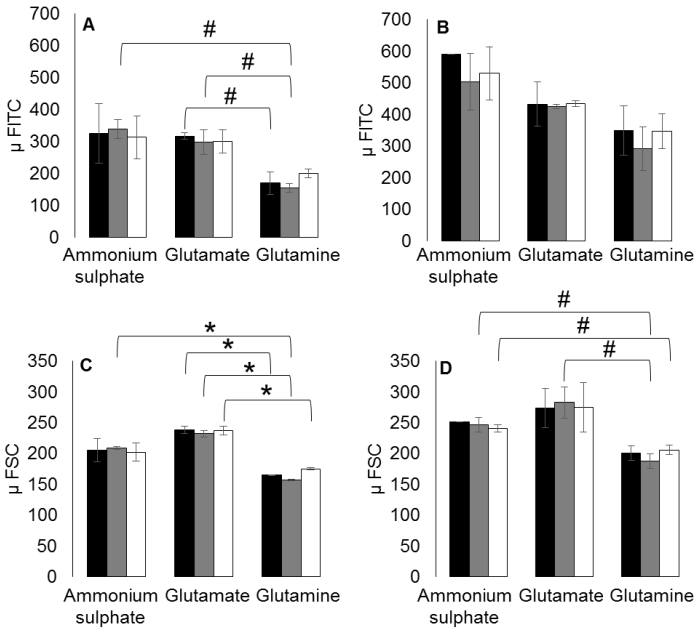


Fig.4 Protein content and FSC determination. The first parameter was measured considering the average of FITC associated fluorescence in the (A) exponential or (B) stationary phase of growth. The average FSC was evaluated in both the (C) exponential or (D) stationary phase of growth. Strains shown are CEN.PK *C* (black columns), CEN.PK *glt1Δ* (grey columns), and CEN.PK *TPI-GLT1* (white columns). Nitrogen sources were supplied to release 30.6 mM of ammonium. Results represent the average and the standard deviation of three independent experiments (*$p \leq 0.01$; #$p \leq 0.05$; Student's *t*-test).

PI and ROS accumulation in wild-type, *glt1Δ*, and *GLT1* overexpressing strains during growth on $(NH_4)_2SO_4$, glutamate or glutamine

The potential of the flow cytometry platform was utilized to determine the viability and reactive oxygen species (ROS) accumulation at the single cell level. Cells in the exponential and stationary phases of growth were first stained with propidium iodide, a marker for determining injured/dead cells (Hohenblum *et al.*, 2003). In the exponential phase of growth, no big differences were observed among the strains or media (Fig. 5A), and the fraction of PI-positive cells was very low. Remarkably, in the very late stationary phase, positivity to PI was particularly high for cells grown in the presence of ammonium sulfate, diversely to the data obtained for cells growing in the presence of glutamate and glutamine (Fig. 5B). Notably, in the presence of glutamine, the PI signals were higher than in the presence of glutamate, with CEN.PK *TPI-GLT1* being the most PI-positive strain (Fig. 5B).

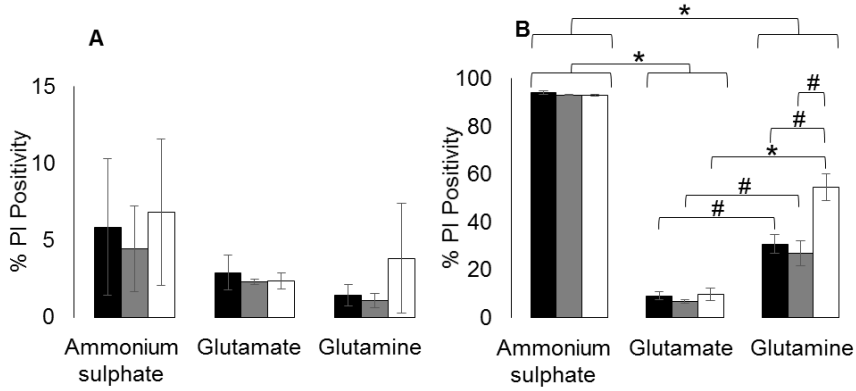


Fig.5 Determination of PI accumulation. This parameter was measured in the (A) exponential and (B) late stationary phases of growth as the percentage of PI-positive cells, for the strains CEN.PK *C* (black columns), CEN.PK *glt1Δ* (grey columns), and CEN.PK *TPI-GLT1* (white columns). Values represent the average and the standard deviation of three independent experiments (*$p \leq 0.01$; #$p \leq 0.05$; Student's *t*-test). Scale of Y-axis are different.

Considering that significant changes in viability have been observed in the different conditions, we sought to determine whether a correlation with ROS accumulation could exist. Therefore, the three strains of this study were challenged with H_2O_2 for triggering ROS production and accumulation inside the cells that were subsequently double stained, as previously reported (Branduardi *et al.*, 2007), with both PI and DHR123 (see Materials and Methods). The uncharged and non-fluorescent DHR123 passively diffuses across membranes into the cell where, in the presence of ROS, it is oxidized to cationic Rhodamine 123, its fluorescent counterpart, through the Fenton reaction (Kim *et al.*, 1996, Madeo *et al.*, 1999).

Fig. 6A shows the DHR123 versus PI cytogram for unchallenged cells. Each dot represents a single cell. The low DHR123 and low PI signals are indicative of a healthy growing yeast population (Panel A, circle with continuous line). Challenging the cells in the exponential phase of growth on ammonium sulfate (control - Panel B) or glutamine (Panel C) for 2 h with H_2O_2 resulted in a yeast population completely ROS positive (circle with double line), also independently from *GLT1* modulation. Furthermore, some of the cells showed a very strong PI signals (circle with dotted line) compared with the unstressed ones. As described above, similar results have been obtained independently from the genetic background of the yeast strain (CEN.PK *C*, CEN.PK *glt1Δ*, and CEN.PK *TPI-GLT1*), indicating that the observed patterns are only nitrogen source dependent. Differently, challenging cells in the exponential phase of growth in glutamate medium resulted in two clear distinct subpopulations, one ROS positive and the other negative (Panels D–F). Also in this case, in each cytogram it is possible to identify some yeast cells with higher PI signals than the unstressed ones, independently from the genetic background. A comparative analysis indicated that the highest fraction of DHR123-positive cells was observed in the

CEN.PK *glt1Δ* strain (~40%, versus ~24% and ~21% for the wild-type and CEN.PK *TPI-GLT1* strains, respectively; Panels D–F).

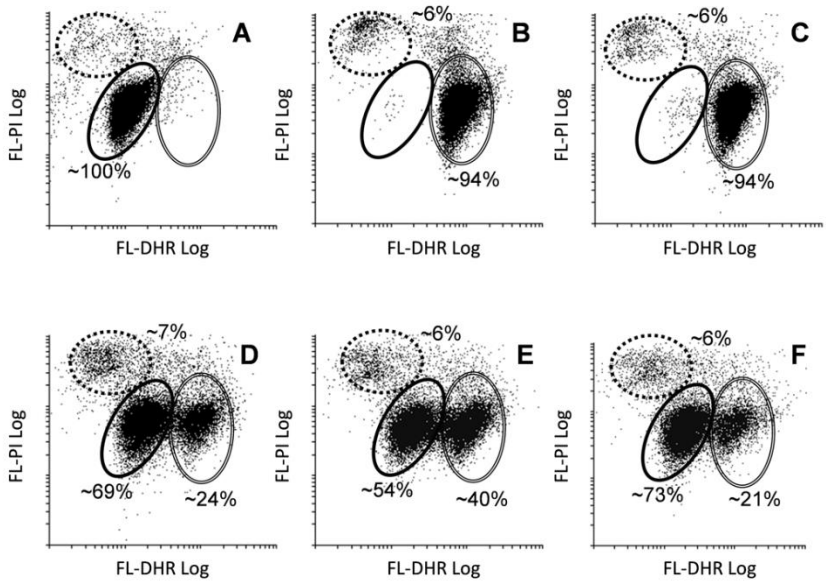


Fig.6 PI and cellular ROS accumulations. Samples were harvested in the exponential phase of growth and shocked for 2 h with 2 mM of hydrogen peroxide. (A) Typical dot plot of untreated cells. (B) Typical dot plot of shocked cells in the presence of ammonium sulfate. (C) Typical dot plot of shocked cells in the presence of glutamine. (D, E, F) Dot plots of shocked CEN.PK *C*, CEN.PK *glt1Δ*, and CEN.PK *TPI-GLT1* strains in the presence of glutamate. The circle with continuous line indicates cells in a healthy state, the circle with dotted line PI-positive cells, and the circle with double line DHR123-positive cells. Each panels show the average percentage values of two independent experiments.

Overall, flow cytometric data confirmed a pro-active role of glutamate in preventing the accumulation of ROS, probably with glutamate required for the synthesis of glutathione (Penninckx, 2002), one of the main radical scavengers in *S. cerevisiae*. Indeed, it has been shown that the supply of glutamate for glutathione biosynthesis was likely to be a factor affecting ROS accumulation and cell death in yeast (Perrone *et al.*, 2008). Furthermore, it has to be mentioned

that an involvement of the GABA pathway, starting from glutamate, has been described in the literature to be associated with scavenging properties against ROS (Coleman *et al.*, 2001, Cao *et al.*, 2013).

CNM enzymatic assays

The activities of the four CNM enzymes (Glt1p, Gln1p, Gdh2p, and Gdh1p) were then determined in wild-type CEN.PK *C* and CEN.PK *TPI-GLT1* strains. Gln1p, Gdh2p, and Gdh1p localization is cytosolic (Moreira dos Santos *et al.*, 2003) and, therefore, only the soluble protein fraction was considered for assays. Data are summarized in Fig. 7.

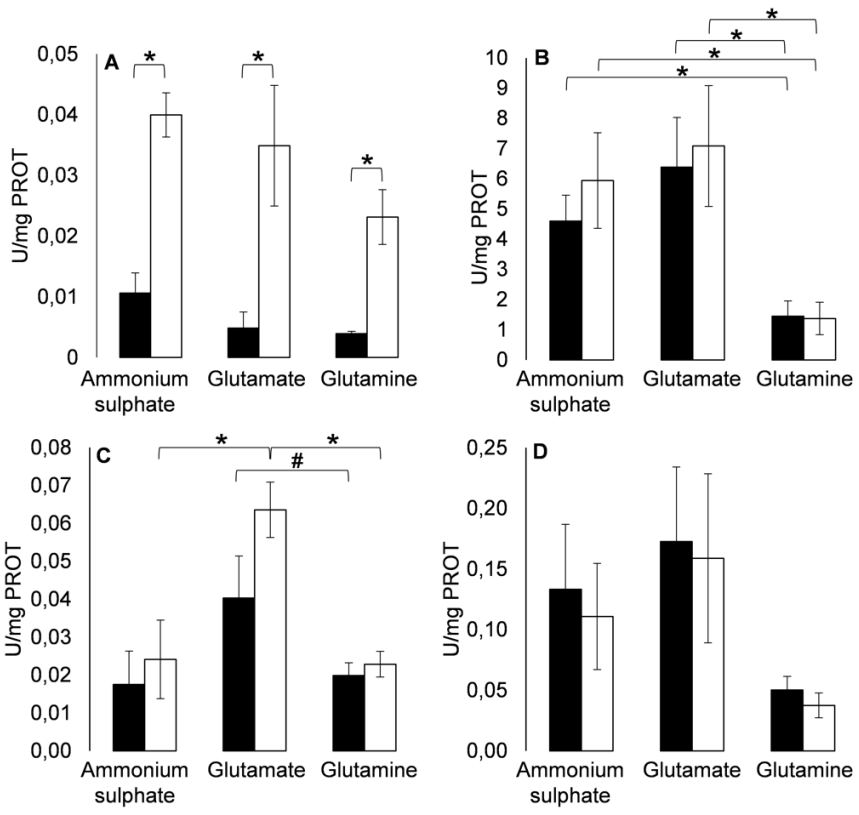


Fig.7 Activities of enzymes of central nitrogen metabolism, tested on the soluble protein fraction of CEN.PK *C* (black columns) and CEN.PK *TPI-GLT1* (white columns), cultured in media with the indicated nitrogen source. (A) Glutamate synthase activity. (B) Glutamine synthetase activity. (C) Glutamate dehydrogenase 2 activity. (D) Glutamate dehydrogenase activity. Nitrogen sources were supplied to release 30.6 mM of ammonium. Values represent the average and the standard deviation of three independent experiments (*$p \leq 0.01$; #$p \leq 0.05$; Student's *t*-test). Scale on Y-axis are different

Results show that *GLT1* overexpression led to an effectively higher Glt1p activity compared with the parental strain (CEN.PK *C*), independently from the media (Panel A). However, this perturbation does not affect the other three enzymatic activities (Panels B-D). Indeed, values tend to be similar between the two strains. Nevertheless, significant changes in activity values could be noticed comparing the three media, independently from the genetic background. Both Gln1p (Panel B) and Gdh2p (Panel C) activities were lower in glutamine medium than in glutamate one, whereas, despite present, this reduction was not statistically relevant for Gdh1p (Panel D) activity. Comparing ammonium sulfate versus glutamine, only Gln1p activity (Panel B) was significantly higher in the presence of the first nitrogen source. Finally, despite that all the activities tend to be slightly lower in the presence of ammonium sulfate than in glutamate medium, only the Gdh2p activity of the CEN.PK *TPI-GLT1* strain was significantly reduced (Panel C). Trying to find a possible explanation, we speculated that *GLT1* overexpression might increase glutamate levels that in turn, as a direct substrate of Gdh2p, could raise its activity. Therefore, in the presence of this amino acid, this effect might be even stronger. Furthermore, considering that the transcription factor Gln3p in the presence of glutamate can cross the nuclear membrane, stimulating CNM genes expression (Valenzuela *et al.*, 1998, Magasanik and Kaiser, 2002, Riego *et al.*, 2002, Zhao *et al.*, 2013), all the activities were generally increased in the presence of this nitrogen source. On the

contrary, since glutamine is a strong activator of the TOR pathway (target of rapamycin), which indirectly prevents the expression of all the four genes (Nissen *et al.*, 2000, Crespo *et al.*, 2002, Ljungdahl and Daignan-Fornier, 2012, Zhao *et al.*, 2013), CNM activities (but in our case with the exception of Glt1p) tend to be very low compared with glutamate medium, significantly in the case of Gln1p and Gdh2p (Panels B and C).

Discussions

S. cerevisiae is one of the successful workhorses for industrial applications (see also Introduction). This yeast is able to metabolize a wide variety of nitrogen sources via enzymatic reactions that are indirectly linked in a network of physiological responses. These responses and their regulations are crucial to the yeast for optimizing the exploitation of the environment and are therefore crucial to improve the biotransformation of protein-rich raw materials, such as whey or exhausted biomasses derived from fermentation processes (Huo *et al.*, 2011, Abdel-Rahman *et al.*, 2013). Indeed, the replacement of petrochemistry-based transport fuels and bulk chemicals by yeast industrial biotechnology requires cost-effective fermentation processes, where yields of substrate conversion into product must approach the maximal theoretical values.

Only very few manuscripts looking over the engineering of the central nitrogen metabolism of *S. cerevisiae* have been published (Nissen *et al.*, 2000, Moreira dos Santos *et al.*, 2003, Roca *et al.*, 2003, Wang *et al.*, 2013). In all these works, the redox intracellular status was unbalanced to decrease glycerol levels and to increase the ethanol production in aerobic or anaerobic batch, using ammonium sulfate as a sole nitrogen source. To reach this purpose, *GLT1* has always been overexpressed, with a contemporary deletion of *GPD1* and *GPD2* (encoding for glycerol-3-phosphate dehydrogenases) (Wang *et al.*, 2013), or with an additional *GDH1* deletion and *GLN1* overexpression (Nissen *et al.*, 2000, Moreira dos Santos *et al.*, 2003, Roca *et al.*, 2003). Improvements in terms of ethanol yield and production have been reached in some cases (Nissen *et al.*, 2000, Wang *et al.*, 2013), but the ethanol productivity resulted negatively affected by the reduced specific growth rate of the strains.

To focus the attention on the *GLT1* contribution in the CNM, we analyzed the effects of *GLT1* modulation by growing wild-type, *glt1Δ*, and *GLT1* overexpressing strains under different physiological conditions. To the best of our knowledge, this work is the first describing the physiology of *S. cerevisiae* strains exclusively modulating the *GLT1* expression.

At the cellular level, almost all nitrogen sources are catabolized to glutamate, which plays a key role in the direct biosynthesis of all the others amino acids, except for asparagine and tryptophan, which are synthesized starting from glutamine (Ljungdahl and Daignan-Fornier, 2012). In the experiments, high or low ammonium sulfate (as control), glutamate, or glutamine concentrations have been used. Furthermore, a high glucose concentration (150 g/l) has been chosen with a double purpose; on the one hand to balance and support the high nitrogen supply and on the other hand to simulate a condition often applied for industrial yeast fermentations. It is important to underline that such concentration is undoubtedly high for typical laboratory scale experiments and it could furthermore cause glucose repression on different genes, including those of the CNM. For example, an implication of Snf1p kinase, one of the major cytoplasmic glucose sensors, has been discovered in the regulation cascade of *GLT1* expression (Usaite *et al.*, 2008). High glucose concentrations can inactivate this kinase (Kayikci and Nielsen, 2015), which, in turn, prevents Gln3p nuclear accumulation (Bertram *et al.*, 2002). Consequently, it is important to consider that in the presented experimental setting the expression of genes controlled by this transcription factor, as *GDH1* (Riego *et al.*, 2002), *GDH2*, and *GLN1* (Magasanik and Kaiser, 2002), could be affected by such high glucose concentration.

Data obtained indicate that a different *GLT1* background does not interfere with the growth properties of the strains when cultivated on ammonia, glutamate, or glutamine as nitrogen sources, even if in the overexpressing strain

the Glt1p activity is greatly enhanced compared with the control strain. However, as an important exception, the *GLT1* overexpressing strain was less viable only in the presence of glutamine as nitrogen source. These data taken together confirm and highlight the robustness of the CNM of yeast against internal perturbation, and at the same time its plasticity with respect to the environment. Indeed, a strong modulation of yeast growth is obtained by using different nitrogen sources. The highest biomass productions are obtained by formulating media with glutamine, whereas glutamate confers higher cellular viability and robustness compared with the other media.

Considering the strong metabolic dependency of the two amino acids at the CNM node (see Fig. 1), a difference between glutamate and glutamine (as shown in Fig. 6) was not trivial to anticipate. This difference suggests a proactive role of glutamate to guarantee a generally higher cellular robustness. Nevertheless, data shown in Fig. 6 (Panels D–F) clearly indicate that with this nitrogen source, two distinct yeast subpopulations, which react to the external stressor in a very different way, appear in exponential phase of growth. Moreover, one subpopulation can easily prevent the formation of ROS species. In this respect, *GLT1* deletion seems to have a negative consequence, since it doubles (from ~24% to ~40%) the fraction of cells accumulating ROS. Diversely, there is no variation in the abundance of this subpopulation when *GLT1* is overexpressed. Currently, we do not have speculations for the interpretation of the data and future experiments will help in elucidating these aspects. For example, it will be interesting to monitor ROS formation also in early and late stationary phase of growth, to evaluate whether the effect of *GLT1* deletion is specific or not to the sole exponential phase. In general, understanding these aspects might foster the design of much robust yeast cell factories. Finally, when cells are supported with ammonium sulfate-based medium (the nitrogen source mainly used at laboratory level), the lowest optical density associated to the

lowest viability is observed. This needs to be taken into account when results obtained at laboratory scale are important for speculating about industrial scale-up, and opens further consideration on media formulation.

Materials and methods

S. cerevisiae strains construction

All the primers and *S. cerevisiae* strains employed and developed in this study are listed in Tables 1 and 2, respectively.

Table.1 List of primers used in this study

Name	Sequence
dGLT1 KAN FW	**ATGCCAGTGTTGAAATCAGACAATTTCGATCCATTGGAAG***CGG ATCCCCGGGTTAATTAA*
dGLT1 KAN REV2	**GGTCACTTTGAACTTTCAAAGACAAATGAATCCTCTGATAT***GAA TTCGAAGCTCGTTTAAACTG*
dGLT1 CNTR FW 2	TAGAAAAGAAAGCATGCCAGT
KAN RV	TTAGAAAAACTCATCGAGCATCAAATG
KAN FW	ATGGGTAAGGAAAAGACTCACGTT
dGLT1 CNTR RV 2	GCAATAGACGTGGGTCACTTTGA
GLT1_SMAI_FW	GCGCCG***CCCGGG*-ATGCCAGTGTGAAATCAGAC**
GLT1_SMAI_REV	GCATAT***CCCGGG*-GCTGGAACCATCCCAAGGTTC**C
TPICNTR_FW	GATCTACGTATGGTCAT
GLT1CNTR_REV	GTCATCACTAGTGATGT
GLT1-GFP FW	**ACGTGATTACAAACTATTGAAAGAATTAGCTAGTCAAGTC***CGGA TCCCCGGGTTAATTAA*
GLT1-GFP RV	**ATAATATACGATCATAAAATAAATAATAACTCAAGCTTT***GAATTC GAGCTCGTTTAAACTGG*
GFP RW	ATATTATTTGTATAGTTCATCCATGC
D GLT1 cntr fw2	TAGAAAAGAAAGCATGCCAGTGTTA

Bold: sequence annealing to the genome. Italics: sequence annealing to the plasmid. Bold and italics: *Xma*I restriction site.

Table.2 List of strains used in this study

Strains	Relevant genotype	Source
CEN.PK 102-5B	*MATa, ura3-52, his3-11, leu2-3/112, TRP1, MAL2-8C, SUC2* (reference strain)	A
CEN.PK *C*	CEN.PK 102-5B (*pYX012; pYX022; pYX042*)	B
CEN.PK *glt1Δ*	CEN.PK 102-5B *glt1::kanMX4* (*pYX012; pYX022; pYX042*)	B
CEN.PK *TPI-GLT1*	CEN.PK 102-5B *GLT1::pYX042(TPI-GLT1),* (*pYX012; pYX022*)	B
CEN.PK *GLT1-GFP*	CEN.PK 102-5B *GLT1::GLT1-GFP,* (*pYX012; pYX042*)	B
CEN.PK *TPI-GLT1-GFP*	CEN.PK 102-5B *GLT1::pYX042(TPI-GLT1), GLT1::GLT1-GFP,* (*pYX012*)	B

A) Kindly provided by Dr. P. Kotter (Institute of Microbiology, Johann Wolfang Goethe-University, Frankfurt, Germany). B) This sutdy.

The CEN.PK *C* control strain was obtained by transforming the reference strain CEN.PK 102-5B with the integrative plasmids pYX012, pYX022, and pYX042 (from R&D System, Wiesbaden, Germany).

The strain CEN.PK *glt1Δ* was constructed by deleting the first 1,650 bp of gene *GLT1*. First the *KanMX4* cassette was amplified from the pFA6-a-*KanMX4* (Wach *et al.*, 1994) plasmid with the primers dGLT1 KAN FW and dGLT1 KAN REV2. Then the reference strain CEN.PK 102-5B was transformed and selected on YPD plates with 200 mg/l of G418. Deletion was verified by PCR with the pairs of primers dGLT1 CNTR FW 2 / KAN RV and KAN FW / dGLT1 CNTR RV 2, respectively.

The *GLT1* overexpressing strain was created as described: the first 1,191 bp of the *GLT1* genomic ORF was amplified by PCR, including the ATG starting codon, with primers GLT1_SMAI_FW and GLT1_SMAI_REV. Then, both the

PCR product and the pYX042-ATG plasmid (RandD System) were *Xma*I digested and a ligation reaction was performed, resulting in plasmid pYX042-ATG (*TPI-GLT1*). This plasmid was *Bgl*II digested inside the insert and, once linearized, the reference strain CEN.PK 102-5B was transformed with it. The proper recombination in the *GLT1* locus was verified by PCR with primers TPICNTR_FW and GLT1CNTR_REV.

The strains CEN.PK *GLT1-GFP* and CEN.PK *TPI-GLT1-GFP* were created as follows: a fragment containing the genes *GFP* and *HIS3* (separated by a linker) was amplified by PCR from pFA6a-GFP(S65T)-His3MX6 with the primers GLT1-GFP FW and GLT1-GFP RV. With this construct, strains CEN.PK 102-5B and CEN.PK *TPI-GLT1* were transformed and selected on appropriate plates without histidine. Analytical PCR with primers GFP RW and DGLT1 cntr fw2 was performed on transformants to verify proper recombination.

When necessary, all the strains of this study were complemented to obtain the corresponding prototrophic strain with the integrative plasmids previously indicated.

All the transformations were performed according to the LiAc/ PEG/ss-DNA protocol (Gietz and Woods, 2002).

Media and growth kinetics

Yeast cultures were shake-flask grown in modified Verduyn medium (Verduyn *et al.*, 1992) composed of 150 g/l glucose, vitamins and traces (1,000×), water, and saline solution at pH 5. The latter were prepared with 0.5 g/l $MgSO_4 \cdot 7H_2O$ and 3 g/l KH_2PO_4, and subsequently supplied with different nitrogen sources as follows: 15 g/l $(NH_4)_2SO_4$ or 48.96 g/l glutamate or 18 g/l glutamine (corresponding to 30.6 mM of ammonium released); 2.5 g/l $(NH_4)_2SO_4$ or 8.16 g/l glutamate or 3 g/l glutamine (corresponding to 5.1 mM of ammonium

released). Pre-cultures were performed in minimal medium with 20 g/l glucose and 6.7 g/l yeast nitrogen base. Cells were inoculated at the starting optical density (OD_{660nm}) of 0.05 in 250 ml flasks containing 50 ml of medium. Fermentations were performed at 30°C under continuous shaking (160 rpm) and cellular growth was followed by measuring the OD_{660nm} using the spectrophotometer UV-1601 (Shimadzu).

Preparation of soluble and organelle-associated protein extract

Cells in exponential phase of growth were washed twice with cold deionized water and resuspended in extraction buffer composed of 0.1 M potassium phosphate buffer (pH 7.5), 1 mM ethylenediaminetetraacetic acid (EDTA), 1 mM dithiothreitol (DTT), 1 mM protease inhibitor cocktail, and 1 mM phenylmethanesulfonylfluoride (PMSF). The cell suspension was subjected to three cycles of mechanical disruption with the FastPrep-24 (MP Biomedical). Cellular lysate was first centrifuged at 700 ×g f or 10 min at 4°C to separate supernatants from glass beads and cellular debris and then at 20,817 ×g for 20 min at 4°C to obtain the soluble protein fraction; the organelle-associated protein fraction in the pellet was solubilized by the detergent DS1 composed of 7 M urea, 2 M thiourea, 40 g/l 3-((3-cholamidopropyl)dimethylammonium)-1-propanesulfonate (CHAPS), 60 mM DTT, and 20 mM 2-iodoacetamide (IAA). Protein concentration was estimated according to (Bradford, 1976), using bovine serum albumin as the reference.

Enzymatic assays

Glutamate synthase enzymatic assay was performed as previously described (Cogoni *et al.*, 1995), slightly modified as follows: the reaction mixture (1 ml final volume) contained buffer phosphate 50 mM (pH 7), NADH 10 mM, α-ketoglutarate 50 mM, and the sample. To start the reaction, glutamine 100 mM

was added. In parallel, this assay was repeated with azaserine 5 mM, a competitive inhibitor of Glt1p used as a control. NADPH glutamate dehydrogenase 1 and NADH glutamate dehydrogenase 2 activities were determined as previously described (Holmes *et al.*, 1989), slightly modified as follows: the reaction mixture (1 ml final volume) contained buffer phosphate 50 mM (pH 7), NADPH or NADH 10 mM, α-ketoglutarate 50 mM, and the sample. To start the reaction, ammonium chloride 100 mM was added. For all the assays, NADH (or NADPH) oxidation was monitored by following the decrease in absorbance at 340 nm for 10 min. The *ΔOD/min* was obtained using the maximum linear rate of the reaction. One unit (U) was defined as the amount of enzyme that oxidized 1 nmol of NADH (or NADPH) in one minute. Glutamine synthetase activity was determined according to (Kingdon *et al.*, 1968).

Flow cytometry analysis

Experiments were carried out with the Beckman Coulter CYTOMICS-FC 500. For propidium iodide (PI) staining, an amount of cells corresponding to 0.2 OD was washed with 1× PBS and then with Tris-HCl 50 mM / $MgCl_2$ 15 mM (pH 7.7) buffer. Subsequently, cells were resuspended in 1 ml of PI 0.23 mM (dissolved in Tris-HCl 50 mM / $MgCl_2$ 15 mM, pH 7.7). After 20min of incubation on ice and in the dark, samples were analyzed with excitation wavelength at 535 nm and emission wavelength at 617 nm.

For DHR123 (dihydrorhodamine 123) and PI double-staining experiments, samples were prepared as previously described (Branduardi *et al.*, 2007) but modified as follows: cells at 0.2 OD were washed and resuspended in 1 ml of 1× PBS. Samples were incubated at 30°C, in the dark, at 160 rpm for 2 h in the presence of 5 μg/ml of DHR123 and 2 mM of hydrogen peroxide. Cells were washed with 1× PBS, then with buffer Tris-HCl 50 mM / $MgCl_2$ 15 mM (pH 7.7), and finally resuspended in 1 ml of 5× PI. Samples were analyzed with

excitation wavelength at 535 nm for PI and 500 nm for DHR123 and emission wavelength at 617 nm and 536 nm, respectively.

For FITC (fluorescein isothiocyanate) experiments, cells at 2 OD were resuspended in 1 ml of cold ethanol 70% (v/v), and incubated for 15 min at −20°C first and then for 20 min at 4°C. Then 100 µl of the cellular suspension was washed once with 1× PBS, resuspended in a FITC solution 5 ng/ml (dissolved in $NaHCO_3$ 50 µg/ml), and incubated in the dark on ice for 30 min. Samples were analyzed with excitation wavelength at 495 nm and emission wavelength at 519 nm.

The data obtained from all the experiments were analyzed with the software Cyflogic ver. 1.2.1.

Fluorescence microscopy

Culture samples of strains CEN.PK *GLT1-GFP* and CEN.PK *TPI-GLT1-GFP*, corresponding to 0.5 OD, were washed once with 1× PBS, and then resuspended in 30 µl of 1× PBS and observed under the fluorescence microscope Nikon ECLIPSE 90i (Nikon), using the 100× objective. Images were acquired using CoolSnap CCD camera and then analyzed with the software Metamorph 6.3.

References

Abdel-Rahman MA, Tashiro Y and Sonomoto K (2013) Recent advances in lactic acid production by microbial fermentation processes. *Biotechnol Adv* **31**: 877-902.

Bertram PG, Choi JH, Carvalho J, Chan TF, Ai W and Zheng XF (2002) Convergence of TOR-nitrogen and Snf1-glucose signaling pathways onto Gln3. *Mol Cell Biol* **22**: 1246-1252.

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248-254.

Branduardi P, Fossati T, Sauer M, Pagani R, Mattanovich D and Porro D (2007) Biosynthesis of vitamin C by yeast leads to increased stress resistance. *PLoS One* **2**: e1092.

Branduardi P and Porro D (2012) Yeasts in Biotechnology. *Yeast: Molecular and Cell Biology*, (Feldmann H, ed.) p.^pp. 347-370. John Wiley and Sons Inc, Germany.

Cao J, Barbosa JM, Singh NK and Locy RD (2013) GABA shunt mediates thermotolerance in *Saccharomyces cerevisiae* by reducing reactive oxygen production. *Yeast* **30**: 129-144.

Cogoni C, Valenzuela L, González-Halphen D, Olivera H, Macino G, Ballario P and González A (1995) *Saccharomyces cerevisiae* has a single glutamate synthase gene coding for a plant-like high-molecular-weight polypeptide. *J Bacteriol* **177**: 792-798.

Coleman ST, Fang TK, Rovinsky SA, Turano FJ and Moye-Rowley WS (2001) Expression of a glutamate decarboxylase homologue is required for normal oxidative stress tolerance in *Saccharomyces cerevisiae*. *J Biol Chem* **276**: 244-250.

Crespo JL, Powers T, Fowler B and Hall MN (2002) The TOR-controlled transcription activators *GLN3*, *RTG1*, and *RTG3* are regulated in response to intracellular levels of glutamine. *Proc Natl Acad Sci U S A* **99**: 6784-6789.

DeLuna A, Avendano A, Riego L and Gonzalez A (2001) NADP-glutamate dehydrogenase isoenzymes of *Saccharomyces cerevisiae*. Purification, kinetic properties, and physiological roles. *J Biol Chem* **276**: 43775-43783.

Edwards AN, Patterson-Fortin LM, Vakulskas CA, *et al.* (2011) Circuitry linking the Csr and stringent response global regulatory systems. *Mol Microbiol* **80**: 1561-1580.

Gietz RD and Woods RA (2002) Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol* **350**: 87-96.

Guillamón JM, van Riel NA, Giuseppin ML and Verrips CT (2001) The glutamate synthase (GOGAT) of *Saccharomyces cerevisiae* plays an important role in central nitrogen metabolism. *FEMS Yeast Res* **1**: 169-175.

Hohenblum H, Borth N and Mattanovich D (2003) Assessing viability and cell-associated product of recombinant protein producing *Pichia pastoris* with flow cytometry. *J Biotechnol* **102**: 281-290.

Holmes AR, Collings A, Farnden KJ and Shepherd MG (1989) Ammonium assimilation by *Candida albicans* and other yeasts: evidence for activity of glutamate synthase. *J Gen Microbiol* **135**: 1423-1430.

Hong KK and Nielsen J (2012) Metabolic engineering of *Saccharomyces cerevisiae*: a key cell factory platform for future biorefineries. *Cell Mol Life Sci* **69**: 2671-2690.

Huo YX, Cho KM, Rivera JG, Monte E, Shen CR, Yan Y and Liao JC (2011) Conversion of proteins into biofuels by engineering nitrogen flux. *Nat Biotechnol* **29**: 346-351.

Kayikci Ö and Nielsen J (2015) Glucose repression in *Saccharomyces cerevisiae*. *FEMS Yeast Res* **15**. fov068.

Kim YM, Hong SJ, Billiar TR and Simmons RL (1996) Counterprotective effect of erythrocytes in experimental bacterial peritonitis is due to scavenging of nitric oxide and reactive oxygen intermediates. *Infect Immun* **64**: 3074-3080.

Kingdon HS, Hubbard JS and Stadtman ER (1968) Regulation of glutamine synthetase. XI. The nature and implications of a lag phase in the *Escherichia coli* glutamine synthetase reaction. *Biochemistry* **7**: 2136-2142.

Ljungdahl PO and Daignan-Fornier B (2012) Regulation of amino acid, nucleotide, and phosphate metabolism in *Saccharomyces cerevisiae*. *Genetics* **190**: 885-929.

Madeo F, Fröhlich E, Ligr M, Grey M, Sigrist SJ, Wolf DH and Fröhlich KU (1999) Oxygen stress: a regulator of apoptosis in yeast. *J Cell Biol* **145**: 757-767.

Magasanik B and Kaiser CA (2002) Nitrogen regulation in *Saccharomyces cerevisiae*. *Gene* **290**: 1-18.

Miller JS and Quarles JM (1990) Flow cytometric identification of microorganisms by dual staining with FITC and PI. *Cytometry* **11**: 667-675.

Moreira dos Santos M, Thygesen G, Kötter P, Olsson L and Nielsen J (2003) Aerobic physiology of redox-engineered *Saccharomyces cerevisiae* strains modified in the ammonium assimilation for increased NADPH availability. *FEMS Yeast Res* **4**: 59-68.

Murthy GS, Johnston DB, Rausch KD, Tumbleson ME and Singh V (2012) A simultaneous saccharification and fermentation model for dynamic growth environments. *Bioprocess Biosyst Eng* **35**: 519-534.

Nissen TL, Kielland-Brandt MC, Nielsen J and Villadsen J (2000) Optimization of ethanol production in *Saccharomyces cerevisiae* by metabolic engineering of the ammonium assimilation. *Metab Eng* **2**: 69-77.

Penninckx MJ (2002) An overview on glutathione in *Saccharomyces* versus non-conventional yeasts. *FEMS Yeast Res* **2**: 295-305.

Perrone GG, Tan SX and Dawes IW (2008) Reactive oxygen species and yeast apoptosis. *Biochim Biophys Acta* **1783**: 1354-1368.

Porro D, Gasser B, Fossati T, Maurer M, Branduardi P, Sauer M and Mattanovich D (2011) Production of recombinant proteins and metabolites in yeasts: when are these systems better than bacterial production systems? *Appl Microbiol Biotechnol* **89**: 939-948.

Riego L, Avendaño A, DeLuna A, Rodríguez E and González A (2002) *GDH1* expression is regulated by *GLN3*, *GCN4*, and *HAP4* under respiratory growth. *Biochem Biophys Res Commun* **293**: 79-85.

Roca C, Nielsen J and Olsson L (2003) Metabolic engineering of ammonium assimilation in xylose-fermenting *Saccharomyces cerevisiae* improves ethanol production. *Appl Environ Microbiol* **69**: 4732-4736.

Soberón M and González A (1987) Physiological role of glutaminase activity in *Saccharomyces cerevisiae*. *J Gen Microbiol* **133**: 1-8.

Usaite R, Wohlschlegel J, Venable JD, Park SK, Nielsen J, Olsson L and Yates Iii JR (2008) Characterization of global yeast quantitative proteome data generated from the wild-type and glucose repression *Saccharomyces cerevisiae* strains: the comparison of two quantitative methods. *J Proteome Res* **7**: 266-275.

Valenzuela L, Ballario P, Aranda C, Filetici P and González A (1998) Regulation of expression of *GLT1*, the gene encoding glutamate synthase in *Saccharomyces cerevisiae*. *J Bacteriol* **180**: 3533-3540.

Verduyn C, Postma E, Scheffers WA and Van Dijken JP (1992) Effect of benzoic acid on metabolic fluxes in yeasts: a continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast* **8**: 501-517.

Wach A, Brachat A, Pöhlmann R and Philippsen P (1994) New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae*. Yeast, **10**: 1793–1808.

Wang J, Liu W, Ding W, Zhang G and Liu J (2013) Increasing ethanol titer and yield in a *gpd1Δ gpd2Δ* strain by simultaneous overexpression of *GLT1* and *STL1* in *Saccharomyces cerevisiae*. *Biotechnol Lett* **35**: 1859-1864.

Zhao X, Zou H, Fu J, Chen J, Zhou J and Du G (2013) Nitrogen regulation involved in the accumulation of urea in *Saccharomyces cerevisiae*. *Yeast* **30**: 437-447.

Looking for a new candidate to induce cellular rewiring

As described in the previous chapter, neither the deletion, nor the over-expression of *GLT1* induced profound physiological changes in yeast. These results prompted that glutamate synthase might not be a promising candidate to induce a consistent cellular rewiring aimed at obtaining industrially relevant phenotypes. Indeed, the choice of an appropriate target for this purpose is crucial. In this context, one can look for candidates at different cellular levels, such as in the central carbon metabolism (CCM) or in the energy metabolism (EM). Another interesting possibility is to select a cellular element involved in the control of the whole genome expression like a transcription factor. Alternatively, a hub element that controls gene expression at post-transcriptional level could be chosen. Looking for candidates in this direction, the principal poly(A) binding protein (Pab1) was previously selected in our laboratory, because it plays a prominent role in all the processes related to the mRNA metabolism. In the following chapter, a detailed description of Pab1 intracellular functions and structure are presented.

Chapter 2

The *Saccharomyces cerevisiae* poly(A) binding protein (Pab1): master regulator of mRNA metabolism and cell physiology

Submitted to YEAST Journal

Brambilla Marco, Martani Francesca, Bertacchi Stefano, Branduardi Paola.

Introduction

Eukaryotic mRNAs are characterized at their 3' end by poly(A) tails that associate to the so-called poly(A) binding proteins (generically termed PABPs). PABPs role was initially ascribed to the sole mRNAs protection from degradation; afterwards, it became clear that PABPs, being scaffolds for the recruitment of different interactors on poly(A) tails, are involved in many events regulating gene expression at post-transcriptional level (Goss and Kleiman, 2013). Interestingly, higher eukaryotes are characterized by different specialized PABPs showing diverse localizations and functions (Smith *et al.*, 2014). For example, humans possess three cytosolic and one nuclear PABPs, while plants display eight PABPs that can be divided into four classes with different expression level and/or tissue-specificity (Goss and Kleiman, 2013). The majority of PABPs display a conserved structure with four RNA Recognition Motifs (RRM) at the N-terminus, a linker and a globular C-terminal domain, whereas few poly(A) binding proteins lack both the linker and the C-terminal domain or are characterized by the presence of just one or two RRMs (Goss and Kleiman, 2013, Smith *et al.*, 2014). A recent addition to the PABP family is the CCCH-type zinc finger class of proteins, which bind polyadenosine tracts thanks to seven tandem CCCH-type zinc fingers (ZnF) localized at their C-terminal region (Soucek *et al.*, 2012).

The yeast *Saccharomyces cerevisiae* possesses two poly(A) binding proteins, Pab1 and Nab2, that are essential and do not share sequence or structure homology (Schmid *et al.*, 2012), being Nab2 a CCCH-type zinc finger poly(A) binding protein (Soucek *et al.*, 2012). Moreover, they display different localization, with Pab1 predominantly distributed in the cytosol and Nab2 in the nucleus (Schmid *et al.*, 2012). Apart from a hypothesized role in modulating

deadenylation at the cytosolic level, Nab2 functions are mainly associated to the mRNA processing in the nuclear compartment and to the mRNA nuclear export (Soucek *et al.*, 2012). Conversely, Pab1 not only influences these nuclear processes, but it plays a broader role in the cytosol, where it regulates deadenylation and translation initiation and termination. In addition, recent findings revealed unexpected and novel functions associated to Pab1, as well as its potentiality as a target for studying human diseases and for increasing yeast stress tolerance. The goal of this review is to sum up the available information on the *S. cerevisiae* Pab1, in terms of structure and intracellular functions.

Pab1 structure

As mentioned above, the *S. cerevisiae* Pab1 is composed of four RRMs, a linker and a globular C-terminal domain like the majority of poly(A) binding proteins (Figure 1A) (Richardson *et al.*, 2012).

RRMs are highly conserved across species: they generally contain from 90 to 100 residues that fold into a 3D structure consisting of four antiparallel β-sheets flanked by two α-helices (Figure 1B) (Maris *et al.*, 2005, Eliseeva *et al.*, 2013). Each RRM is more similar to the corresponding domain in PABP proteins from other organisms compared to the other RRMs of the same protein (Burd *et al.*, 1991). Nevertheless, each RRM contains the two highly conserved sequences RNP-1 (K/R)-G-(F/Y)-(G/A)-(F/Y)-(V/I/L)-X-(F/Y) and RNP-2 (I/V/L)-(F/Y)-(I/V/L)-X-N-L (where X can be any amino acid) that are involved in poly(A) tail binding and lay in the β3 and β1 sheets, respectively (Figure 1D) (Maris *et al.*, 2005). Four residues majorly contribute to the RRM-RNA interaction: the hydrophobic residues in RNP-1 (positions 3 and 5) and in RNP-2 (position 2), as well as the first positively charged residue in RNP-1 (Maris *et al.*, 2005). In

particular, the second aromatic residue in RNP-1 (position 5) is involved in binding both poly(A) and poly(U) tracts, while the first charged residue in RNP-1 (position 1) is specific only for poly(U) binding (Deardorff and Sachs, 1997). Interestingly, sequence analysis of the *S. cerevisiae* RRMs reveals that the two RNPs in the RRM1 slightly diverges from the consensus sequences: the first residue of RNP-1 is a non-polar leucine instead of a positively charged amino acid, and the fifth residue of RNP-2 is a negatively charged aspartate instead of an asparagine (Figure 1D) (Deardorff and Sachs, 1997). Looking at the RNA binding specificity of the single RRMs by site specific mutagenesis of the second aromatic residue in RNP-1, the RRM2 resulted the one primarily involved in poly(A) binding, whereas the RRM4 displayed a higher affinity for non-specific polypyrimidine tracts, indicating different substrates specificities for these Pab1 domains (Deardorff and Sachs, 1997).

Differently from RRMs, the linker lacks a well-defined globular structure. It is significantly divergent in length across species, and pattern of conserved residues can be found only by aligning the linker sequences of closely related organisms. This linker domain is poor of charged amino acids and it is extremely enriched in prolines (for this reason it is called P domain), methionine, glycine, glutamine and asparagine compared to the average of the yeast proteome (Riback *et al.*, 2017). Therefore, it exhibits an intrinsically disordered behavior that results in a lack of a fixed 3D structure. Among the ensemble of its possible conformations, small-angle X-ray scattering (SAXS) analysis showed that the P domain tends to be mainly compact *in vitro*, with an estimated radius of gyration (R_g) of 20Å (Riback *et al.*, 2017).

Finally, among all the domains of the *S. cerevisiae* Pab1, the C-terminal domain is the sole whose 3D structure was solved by NMR (Kozlov *et al.*, 2002). Interestingly, this domain is globular and, differently from the human one, it folds into four α-helices instead of five and it contains two extra amino acids between

the 2nd and the 3rd helices, as well as a strongly bent C-terminal helix (Figure 1C) (Kozlov *et al.*, 2002). In literature, C-terminal domains of PABPs are also known as MLLE (*i.e.* mademoiselle) because of the presence of the amino acidic sequence KITGMLLE, which is responsible for the interaction with proteins containing the PAM2 motif (PABP-interacting motif 2) (Xie *et al.*, 2014, Roque *et al.*, 2015). However, in *S. cerevisiae* this sequence is slightly divergent from the canonical one, because one leucine and the glutamate are substituted by isoleucine and aspartate, respectively (KITGMILD) (Roque *et al.*, 2015).

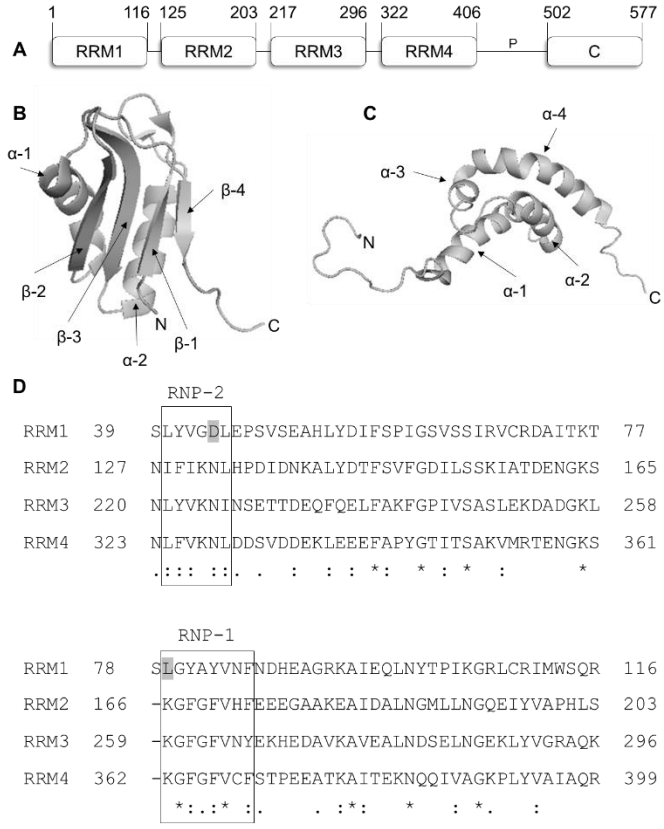


Fig.1 Pab1 structure. A) Schematic representation of the modular structure of the *S. cerevisiae* poly(A) binding protein Pab1. Numbers above the scheme indicate the beginning and the end of the corresponding domain according to (Richardson *et al.*, 2012). RRM: RNA recognition motifs. P: P domain. C: C domain. B) 3D structure of the RRM2 domain of human PABPC generated

using the program PyMOL (version 1.0.0.0) with Protein Data Bank (PDB) code 4F25. C) 3D structure of the C domain of the *S. cerevisiae* Pab1 generated using the program PyMOL (version 1.0.0.0) with PDB code 1IFW. D) Multiple alignment among portions of the RRM domains of the *S. cerevisiae* Pab1. Sequences were aligned using CLUSTAL W Multiple Sequence Alignment Program (version 1.83). Gray background indicates non-conserved residues in RNP-1 and RNP-2. Asterisk (*) indicates positions with fully conserved residue. Colon (:) indicates conservation among residues with strongly similar properties - scoring > 0.5 in the Gonnet PAM 250 matrix. Period (.) indicates conservation among residues with weakly similar properties - scoring =< 0.5 in the Gonnet PAM 250 matrix.

It was recently found that Pab1 is subjected to several post-translational modifications such as six methyl-glutamates, one methyl-arginine, five acetyl-lysines and one phosphoserine (Figure 2) (Low *et al.*, 2013, Low *et al.*, 2014). These modifications might influence the protein structure and/or specific protein-protein interactions and/or the ability to bind RNA. However, limited information on their physiological significance are available and future studies will be required to completely address their role.

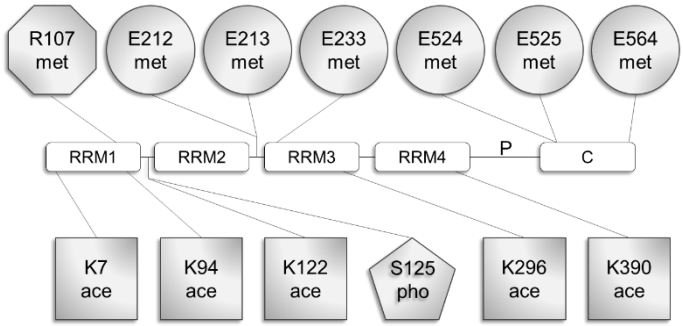


Fig.2 Post-translational modifications of Pab1. Octagon: methyl-arginine residue. Circle: methyl-glutamate residues. Square: acetyl-lysine residues. Pentagon: phospho-serine residue. RRM: RNA Recognition Motifs. P: P domain. C: C domain.

Pab1 localization

Pab1 is predominantly localized in the cytosol, where it regulates many processes such as translation, mRNA stability, deadenylation and stress granules formation. However, this protein is present to less extent also in the nucleus, where it regulates poly(A) tail length and mRNA export. Its ability to shuttle between these two compartments is linked to localization signals in its primary sequence, as well as to interactions with members of the importin-β/karyopherin family of transport receptors (Brune *et al.*, 2005).

Pab1 enters the nucleus in association with the Kap108/Sxm1 importin (Figure 3A) (Brune *et al.*, 2005). The minimal putative Nuclear Localization Sequence (NLS), which is necessary and sufficient for Pab1 to shuttle into the nucleus, encompasses the amino acids 281 – 337 (Brune *et al.*, 2005) between the RRM3 and the RRM4 domains. In addition, recent findings suggested that also the P domain, in conjunction with the presence of the RRM3, might be involved in the enhancement of Pab1 nuclear localization and/or in the prevention of its export (Brambilla *et al.*, 2017). Two mechanisms are responsible for Pab1 export from the nucleus. One of them requires the presence of three elements: the Xpo1/Crm1 exportin, which interacts with the RRM1 domain of Pab1 (Brune *et al.*, 2005), the RanGTP that stabilizes the interaction between Xpo1/Crm1 and Pab1, and the minimal putative Nuclear Export Sequence (NES) located between residues 12 and 17 ("LENLNI") of Pab1 (Figure 3B) (Dunn *et al.*, 2005). The second route of export, which is less characterized, requires Mex67 (*i.e.* an mRNA export receptor (Niño *et al.*, 2013)), and/or ongoing mRNA export (Figure 3C) (Brune *et al.*, 2005). Interestingly, when this pathway is compromised, no significant accumulation of Pab1 was observed in the nucleus, suggesting that Xpo1/Crm1 is sufficient for the efficient export of Pab1 to the cytoplasm (Brune *et al.*, 2005). On the other hand, Pab1 nuclear localization

increases when the Xpo1/Crm1 nuclear export pathway is compromised and it is exacerbated in the absence of also the Mex67/mRNA-dependent route of export (Brune *et al.*, 2005).

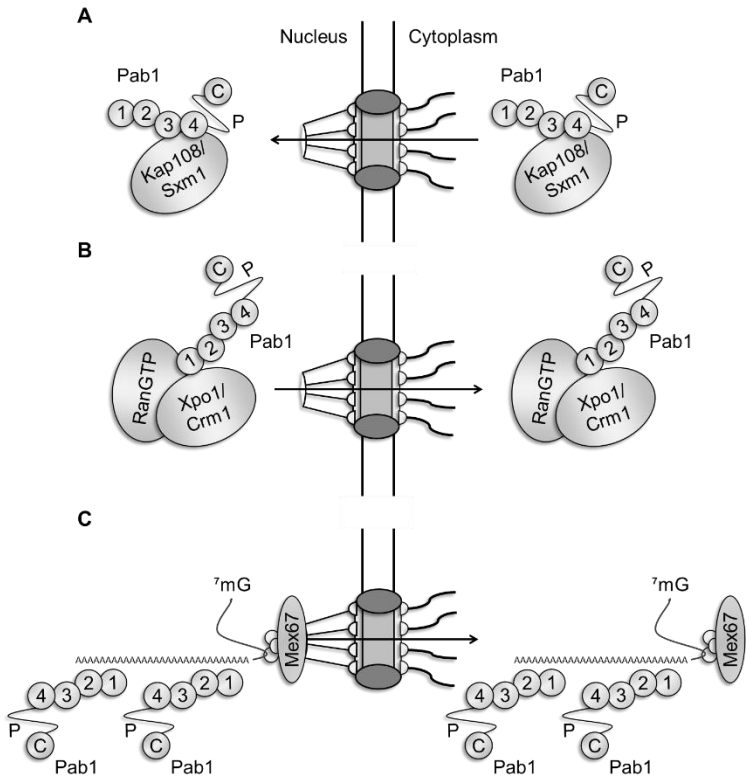


Fig.3 Representation of Pab1 shuttling across the nucleus and the cytoplasm. A) Pab1 enters the nucleus in association with the Kap108/Sxm1 importin. B) The major route of Pab1 export to the cytoplasm is mediated by the exportin Xpo1/Crm1. C) The minor pathway of Pab1 export requires Mex67 and/or ongoing mRNA.

Bypass suppressors of PAB1

The first loss of function studies on Pab1, performed by inactivating the promoter or using a temperature-sensitive mutation, revealed the important role

of this protein in the control of poly(A) tail length and translation initiation (Sachs and Davis, 1989). Successively, the discovery of many bypass suppressor mutations of *PAB1* deletion contributed to better characterize these functions. Suppressors identified until now can be mainly divided into two groups: mutations or deletion of genes affecting the formation of the 60S ribosomal subunit (*SPB1*, *RPL39*, *SPB4*, *RPL35A*), and mutations or deletions of genes encoding for enzymes responsible for mRNA turnover (*LSM1*, *DCP1*, *XRN1*, *RRP6*, *PAT1*, *MRT3*) (Table 1). Regarding the first group of suppressors, it has been proposed that the aberrant production of the 60S subunit may rescue the viability of *pab1Δ* cells by allowing Pab1-independent translation. In the second case, the increased stability of transcripts due to defects in mRNA decapping process and, in general, in mRNA turnover, may counteract the lower translation rate resulting from the absence of Pab1 (Boeck *et al.*, 1998, Wyers *et al.*, 2000). Interestingly, *pbp1Δ* allele suppresses the lethality associated with *PAB1* deletion through a still uncharacterized mechanism not linked to the previous ones. In fact, the deletion of *PBP1* did not alter translation rates, the accumulation of ribosomal subunits, mRNA poly(A) tail lengths or mRNA decay (Mangus *et al.*, 1998).

Loss of function studies using bypass suppressor mutations allowed the association of diverse phenotypes to the absence of Pab1, thus contributing to uncover the functions of this protein, described in details in the next paragraph.

Table. 1 Bypass suppressor mutations of *pab1Δ*

Gene	Description	Main function	Suppressor mutation	Mutant phenotype	Reference
SPB1	AdoMet-dependent methyltransferase	RNA processing and 60S ribosomal subunit maturation	*spb1-1* cold-sensitive allele *spb1-2* thermosensitive allele	-	Sachs and Davis, 1989 Pintard *et al.,* 2000
RPL39 (SPB2, RPL46)	Ribosomal 60S subunit protein L39	Ribosome biogenesis	*spb2-1* cold-sensitive allele *spb2Δ*	-	Sachs and Davis, 1989 Sachs and Davis, 1990
SPB4	ATP-dependent RNA helicase	synthesis of 60S ribosomal subunit	*spb4-1* cold-sensitive allele	-	Sachs and Davis, 1989
LSM1 (SPB8)	Lsm protein	Degradation of cytoplasmic mRNAs	*spb8-2*	accumulation of capped, poly(A)-deficient mRNAs	Boeck *et al.*, 1998
DCP1 (MRT2)	Subunit of the Dcp1p-Dcp2p decapping enzyme complex	mRNA decapping	*dcp1-1* *dcp1Δ*	reduce decapping activity; full length deadenylated mRNAs accumulation no decapping activity; full length deadenylated mRNAs accumulation	Beelman *et al.*, 1996; Hatfield *et al.,* 1996 Beelman et al., 1996
XRN1	5'-3' exonuclease	mRNA decay	*xrn1Δ*	increased levels of full-length mRNA	Caponigro and Parker, 1995
RRP6	Nuclear exosome exonuclease component	3'-5' exonuclease involved in RNA processing, maturation, surveillance, degradation, tethering, and export	*rrp6Δ*	nuclear accumulation of polyadenylated RNA	Dunn *et al.*, 2005
PAT1 (MRT1)	Deadenylation-dependent mRNA-decapping factor	mRNA decapping	*pat1-2* *pat1-3* *pat1Δ*	accumulation of full-length deadenylated transcripts reduced translation initiation	Hatfield *et al.*, 1996 Wyers *et al.*, 2000
MRT3	?	mRNA decay	*mrt3-1*	accumulation of full-length deadenylated transcripts	Hatfield *et al.*, 1996
PBP1	Pab1p-Binding Protein	mRNA polyadenylation	*pbp1Δ*	reduced polyadenylation	Mangus *et al.*, 1998
RPL35A (SOS1)	Ribosomal 60S subunit protein L35A	Ribosome biogenesis	*sos1Δ*	-	Zhong and Arndt, 1993

Pab1 influences mRNA export

Nuclear export of mRNAs is a fundamental step for the regulation of eukaryotic gene expression at post-transcriptional level (Niño *et al.*, 2013). Many proteins are involved in this process, among which the export receptor Mex67/Mtr2, docking mRNAs to the nuclear basket of the Nuclear Pore Complex (NPC), as well as the mRNAs-binding adaptors Yra1, Yra2, Nab2, Npl3, the THO/TREX and TREX-2 complexes that stimulate the association between Mex67 and mRNAs (Niño *et al.*, 2013).

Interestingly, Pab1 indirectly interacts, via RNA molecules, with different nucleoporins of the NPC and with Mex67 (Allen *et al.*, 2002); furthermore, as previously indicated, it can shuttle to the cytoplasm in association with Mex67/mRNA (Figure 3C) (Brune *et al.*, 2005). These observations suggest a possible involvement of Pab1 in mRNAs export. Indeed, yeast strains lacking Pab1, or carrying mutant variants unable to enter the nucleus, showed a delay in the export kinetic of *HAC1* mRNA (Brune *et al.*, 2005). Similarly, an accumulation of *SSA4* mRNA in intranuclear foci was observed in different *pab1Δ* suppressor strains (Dunn *et al.*, 2005). Therefore, the presence of Pab1 in the nucleus might promote an efficient mRNA export, possibly by linking this process to proper 3' end formation (see below) (Brune *et al.*, 2005, Dunn *et al.*, 2005, Niño *et al.*, 2013). However, *in situ* hybridization experiments revealed that, in the absence of Pab1, no significant accumulation of polyadenylated mRNAs in the nucleus occurs (Brune *et al.*, 2005, Dunn *et al.*, 2005). Therefore, its physiological relevance in this process is not completely clear and further studies will be required to fully elucidate it.

Pab1 regulates 3' end processing by controlling deadenylation

Pab1 plays a key role in the control of poly(A) tail length by regulating the association of the deadenylase complexes Pan2/Pan3 (PAN) and Ccr4/Pop2/NOT to mRNAs. In the nucleus, the biological significance of deadenylation is the trimming of poly(A) to achieve the proper length and to make mRNAs ready for export, whereas in the cytoplasm deadenylation is the first step of the general mRNA degradation pathway. It has been proposed a sequential model of deadenylation in which Pan2/Pan3 complex is responsible for poly(A) tail shortening during the first distributive phase of deadenylation and, successively, the more processive Ccr4/Pop2/NOT complex hydrolyzes the remaining 20-25 residues (Brown and Sachs, 1998, Tucker *et al.*, 2002, Wahle and Winkler, 2013, Wolf *et al.*, 2014).

The PAN complex interacts with poly(A) tails thanks to the interaction of the C-terminal domain of Pab1 with the PAM2 motif of the Pan3 subunit (Siddiqui *et al.*, 2007). Pab1 residues R506, G510, G528 and M533 are essential for this interaction (Mangus *et al.*, 2004). The recruitment of the PAN complex on mRNAs by Pab1 is essential for stimulating the efficient exonucleolytic activity of the Pan2 subunit. In fact, yeast strains in which the Pab1 C-terminal domain is absent, or mutated to prevent interaction with Pan3, contain a pool of mRNA with longer poly(A) tails compared to the wild type (Brown and Sachs, 1998, Simón and Séraphin, 2007, Martani *et al.*, 2015). Moreover, although Pan2/Pan3 complex can bind poly(A) *in vitro* independently from Pab1, the deadenylation rate in the absence of this protein is reduced (Wolf *et al.*, 2014). In addition to promote it, Pab1 is also implicated in the negative regulation of the PAN-mediated deadenylation through the recruitment of Pbp1 (Pab1-binding

protein 1). Although the residues involved in the binding with Pan3 and Pbp1 are different (localized at the C domain and P domain, respectively), Pab1 cannot bind both proteins at the same time because of steric hindrance. Consequently, Pbp1 binding to Pab1 displaces Pan3 and *vice versa*; therefore, Pan2/Pan3 mediated deadenylation depends on which protein is interacting with Pab1 (Figure 4) (Mangus *et al.*, 2004).

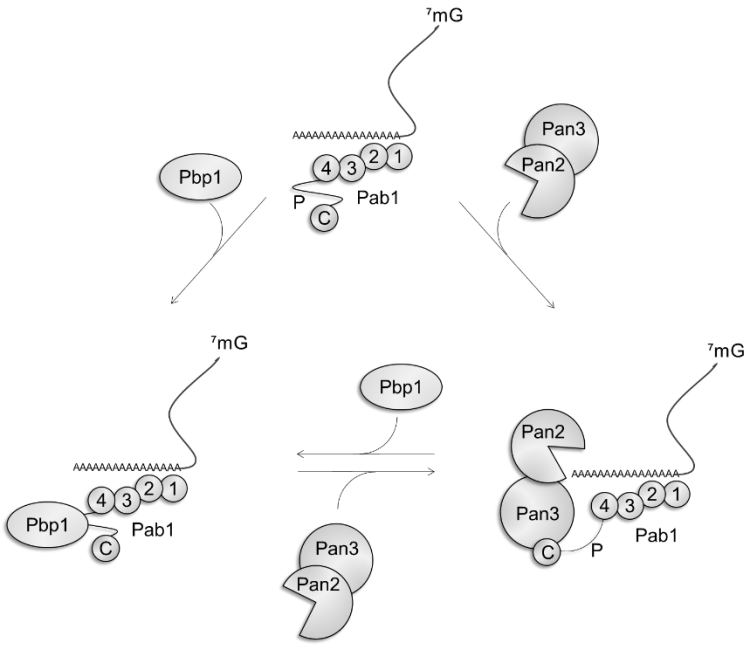


Fig.4 Representation of Pab1 positive and negative regulation of PAN-deadenylation. Pab1 promotes PAN-deadenylation by recruiting the Pan2/Pan3 complex to the poly(A) tail. This interaction is hindered by Pbp1 binding to Pab1. The rate of deadenylation is therefore dependent on the mutually exclusive interaction of Pab1 with Pan3 or Pbp1.

In addition to participate in the mechanisms of mRNA decay, the stimulation of PAN activity by Pab1 plays a role in the trimming of poly(A) tails and the release of mRNAs from transcription sites, thus coupling pre-mRNA processing to export (Dunn *et al.*, 2005, Niño *et al.*, 2013). These processes are

also influenced by other nuclear factors, such as the poly(A) binding protein Nab2, the nuclear exosome subunit Rrp6 and the subunits of the TRAMP complex Trf4 or Trf5 (Schmid *et al.*, 2012). Pre-mRNA processing consists of an endonucleolytic cleavage, which removes part of the 3' end of the pre-mRNA, followed by polyadenylation. The cleavage of mRNA precursors is tightly controlled by specific sequences in the pre-mRNA, as well as by the two multi-subunits cleavage factors I and II (CF I and CF II). Successively, poly(A) tail elongation is carried out by polyadenylation factor I (PF I) and poly(A) polymerase (Pap1), together with CF I (Amrani *et al.*, 1997, Minvielle-Sebastia *et al.*, 1997, Niño *et al.*, 2013). Next, the newly synthetized poly(A) tails are shortened from an initial length of 70 – 90 A's to a final length of 55 – 71 A's by the deadenylase activity of PAN, thus allowing the formation of export-competent mRNPs (Brown and Sachs, 1998, Mangus *et al.*, 2004, Dunn *et al.*, 2005). Pab1 seems to be involved in the control of poly(A) tail length also during the polyadenylation step, since it co-fractionates with CF I and physically interacts with Rna15, an essential subunit for the cleavage and polyadenylation activities of CF I (Amrani *et al.*, 1997, Minvielle-Sebastia *et al.*, 1997). Therefore, Pab1 might regulate poly(A) tail length by two possibly complementary mechanisms: the inhibition of Pap1 activity (Amrani *et al.*, 1997, Brown and Sachs, 1998), and/or the recruitment of the PAN deadenylation complex on mRNAs poly(A) tails (Brown and Sachs, 1998, Mangus *et al.*, 2004).

The other yeast deadenylase complex Ccr4/Pop2/NOT is composed of the scaffold subunits Not1, Not2 and Not3/5 and the two catalytic subunits Ccr4 and Pop2/Caf1. Pab1 can inhibit Ccr4 deadenylation *in vitro* (Tucker *et al.*, 2002), but there are no evidences of a direct contact between the two proteins. *In vivo* it has been observed that Ccr4 shifts from a slow distributive rate to a rapid processive rate of deadenylation when the poly(A) tail is reduced by about 20-25 A's, which corresponds to the number of residues bound by a Pab1 molecule (Yao

et al., 2007). This observation suggested that Pab1 impairs Ccr4 deadenylation by binding poly(A) tails, thus preventing Ccr4 association to mRNAs (Yao *et al.*, 2007). Therefore, Ccr4 deadenylation may be promoted by the removal of Pab1 from the poly(A) tail. Among the mechanisms responsible for Pab1 dissociation from the mRNA, protein self-circularization (Figure 5A) and oligomerization (Figure 5B) allow the formation of Pab1 structures unable to bind poly(A) tails (Yao *et al.*, 2007, Richardson *et al.*, 2012). The RRM1 and the P domains were identified as mainly responsible for protein self-assembly (Yao *et al.*, 2007). Another possible mechanism that allows Pab1 removal from poly(A) and, thus, the switch from a slow distributive to a rapid processive deadenylation, is mediated by Upf1 (Figure 5C), a factor involved in nonsense mediated decay (see below), that subtracts Pab1 molecules by interacting with their RRM1 domains (Richardson *et al.*, 2012). This domain is also involved in the acceleration of Ccr4 deadenylation by Puf3, a 3' UTR binding protein that controls degradation of specific mRNAs (Lee *et al.*, 2010). Altogether, these results highlight that the RRM1 might be the most critical domain of Pab1 for the control of deadenylation mediated by the Ccr4/Pop2/NOT complex. Accordingly, mutagenesis analyses of the RRM1 revealed that the replacement of some key residues (Y41, C70, C109, and Y83) with other amino acids led to a reduced deadenylation rate (Zhang *et al.*, 2013).

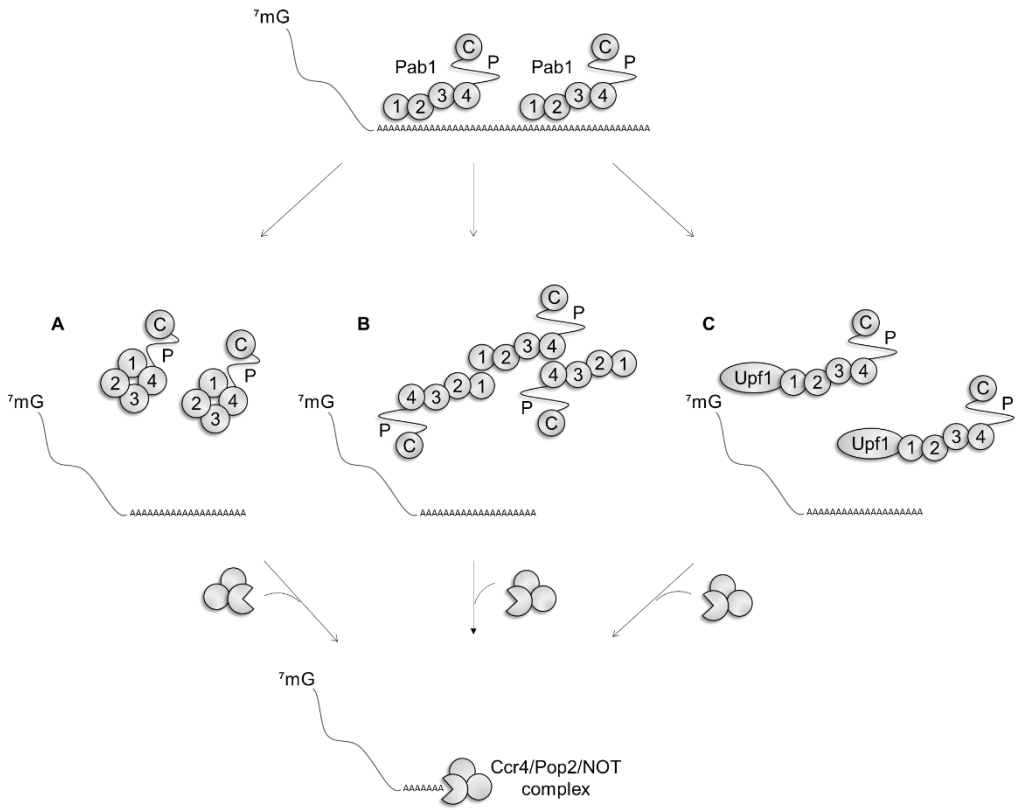


Fig.5 Mechanisms of Pab1 removal from poly(A) tails to promote Ccr4 mediated deadenylation. Pab1 self-association into A) circular species or B) multimers, or C) Pab1 interaction with Upf1, prevent its binding to poly(A) and stimulate Ccr4 dependent deadenylation. (Figures A and B have been redrawn from (Yao *et al.*, 2012)).

Pab1 has an ambiguous role in Nonsense Mediated Decay

A first proposal of Pab1 implication in the nonsense mediated decay (NMD), which is a surveillance pathway for the degradation of mRNAs with premature termination codon (PTC) from which deleterious truncated proteins could be synthetized (Kervestin and Jacobson, 2012, Parker, 2012), came from the so-called *faux*-UTR model. According to it, the 3' UTR downstream to a PTC lacks those termination factors that are normally present on canonical 3' UTRs,

leading to aberrant translation termination and, thus, to NMD activation (Amrani *et al.*, 2004). Transcripts that are targeted to NMD could be subjected to translation repression, and/or acceleration of deadenylation followed by increased rate of 3' to 5' degradation, and/or rapid decapping independent from deadenylation (Parker, 2012). This mechanism requires mainly the three proteins Upf1, Upf2 and Upf3, which are the core components of the surveillance complex, as well as the two releasing factors eRF1 and eRF3 that are required for stop codon recognition (Kervestin and Jacobson, 2012, Parker, 2012).

In this context, Pab1 was proposed to play a key role in discriminating normal and PTC-containing mRNAs because of its differential proximity to stop codons. Indeed, when Pab1 is tethered 37-73 nt downstream to the 3' of a PTC, reporter mRNAs were particularly stable *in vivo*, while they are subjected to NMD when Pab1 is tethered 164 nt downstream to the PTC (Amrani *et al.*, 2004). Consistent with these results, mRNAs with excessively long 3' UTR are usually substrates for NMD; moreover, the removal of the major part of the coding sequence downstream to a premature stop codon led to a stabilization of mRNAs (Kervestin and Jacobson, 2012). Nevertheless, it has been also demonstrated that PTC-containing mRNAs from suppressor *pab1Δ* strains, as well as PTC-containing mRNAs lacking the poly(A) tail, are still recognized and destabilized in a Upf1-dependent manner (Meaux *et al.*, 2008). These observations pointed out that Pab1 is not essential for discriminating between premature and normal stop codon, and suggest that other factors are likely involved in targeting mRNAs to NMD (Meaux *et al.*, 2008).

Pab1 loading on mRNAs influences translation initiation

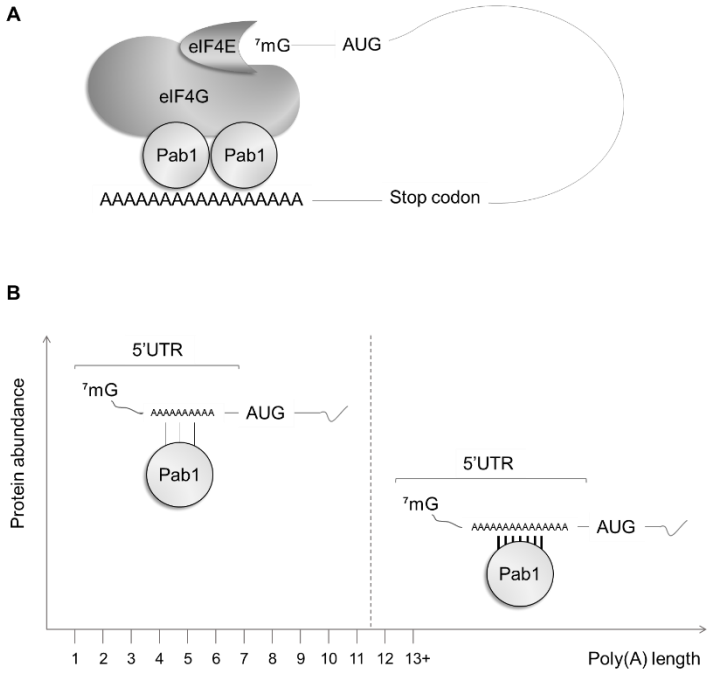
First evidences of the requirement of poly(A) tail and PABPs for translation were discovered in the 80's (Jacobson and Favreau, 1983, Grossi de Sa *et al.*, 1988, Sachs and Davis, 1989). Subsequent studies in *S. cerevisiae* revealed that Pab1 is necessary for translation initiation because it promotes the assembly of ribosomal subunits on mRNAs structure through diverse mechanisms. One of them is based on the closed-loop model in which a circular mRNA structure is formed thanks to the joining of the 7-methylguanosine (m7G) 5' cap with the 3' poly(A) tail (Figure 6A) (Archer *et al.*, 2015, Costello *et al.*, 2015, Melamed *et al.*, 2015, Dever *et al.*, 2016). The minimal factors required for closed-loop formation and integrity are the 5' cap, the 3' poly(A) tail, eIF4E, eIF4G1 or eIF4G2 (two redundant isoforms), eIF3 (a multiprotein complex) and at least two Pab1 molecules. eIF4G bridges the two ends of the mRNA by binding eIF4E and Pab1 that are associated to the 5' cap and the 3' poly(A) tail, respectively, enabling the formation of a stable closed loop. Interestingly, Pab1 association to eIF4G, which primarily occurs through the RRM2 domain of Pab1 (Melamed *et al.*, 2015, Dever *et al.*, 2016), significantly stimulates the binding between the mRNA and the entire eIF4F complex (composed of eIF4E, eIF4G and the RNA helicase eIF4A) (O'Leary *et al.*, 2013). This structure is suggested to restrict efficient translation to intact mRNAs, to help polysomes formation, to facilitate ribosome recycling and to protect mRNAs from degradation (Huch and Nissan, 2014, Archer *et al.*, 2015, Costello *et al.*, 2015). More in details, the assembly of the closed-loop facilitates the recruitment at the initiation codon of the 43S preinitiation complex (composed of 40S ribosomal subunit, eIF1, eIF2, eIF3, eIF5 and the methionyl-tRNA), leading to formation of the 48S complex (Komar *et al.*, 2012, Wang *et al.*, 2012, Dever *et al.*, 2016). Next, ribosome

scanning occurs and, once the start AUG codon is reached, the 60S subunit joins the complex to form the 80S ribosome, leading to the initiation of protein synthesis (Farley *et al.*, 2011, Komar *et al.*, 2012, Wang *et al.*, 2012, Costello *et al.*, 2015, Dever *et al.*, 2016).

Recently, it has been demonstrated that the closed-loop model is relevant for the translation of only a subset of mRNAs (Costello *et al.*, 2015). In fact, another subset of efficiently translated transcripts (encoding, for example, for proteins involved in amino acid and nucleotide metabolism, carbohydrate metabolism, energy generation, tRNA amino-acylation and translation) are enriched in Pab1 but under-represented in all the other factors required for the formation of the closed-loop. Therefore, the circular structure might be less relevant for these mRNAs, and alternative mechanisms of translation dependent on Pab1 might take place (Costello *et al.*, 2015). At the 5' UTR, some mRNAs contain Internal Ribosome Entry Sites (IRESs) that can recruit ribosomes for translation initiation through a cap-independent mechanism (reviewed in (Plank and Kieft, 2012)). For this reason, IRES-mediated translation initiation has been proposed to substitute the cap-dependent initiation translation when it is repressed, such as during cellular stresses, and thus represents an important strategy used by cells to regulate gene expression (Reineke and Merrick, 2009). Some IRESs in *S. cerevisiae* contain poly(A) tracts that can be bound by Pab1; it has been suggested that Pab1 binding at the 5' UTR may be a requisite for IRES activity and may recruit eIF4G in the absence of cap and eIF4E, allowing the recruitment of 40S subunit (Gilbert *et al.*, 2007). However, the role of Pab1 at the 5' UTR in favoring translation initiation is ambiguous and seems to depend on the strength with which it binds A-rich stretches. In fact, it was found that mRNAs with poly(A) stretches shorter than 12 adenosines are weakly bound by Pab1 and are translated more efficiently than mRNAs with poly(A) tracts equal to or longer than 12 adenosines (Figure 6B) (*i.e.* the minimal length for an

efficient Pab1 binding (Sachs *et al.*, 1987)) (Xia *et al.*, 2011). Consequently, protein abundance and synthesis are dependent on the poly(A) length at 5' UTR and, thus, on the Pab1 binding efficiency. Interestingly, the mRNA encoding for Pab1 contains an IRES site in its 5' UTR with a poly(A) tract of 11 consecutive A's that can escape an eventual negative auto-regulation of the protein itself (Xia *et al.*, 2011).

It has been suggested that Pab1 promotes translation initiation also by favoring the efficient formation of the 48S pre-initiation complex thanks to its interaction with the eIF3 complex (Farley *et al.*, 2011). Moreover, it has been proposed a model of action of poly(A) to promote 60S ribosomal subunit joining in which Pab1 binding to poly(A) tail inhibits the two RNA helicases Ski2 and Slh1 and, consequently, stimulates the activity of the translation initiation factors eIF5 and eIF5B (Figure 6C) (Searfoss *et al.*, 2001).



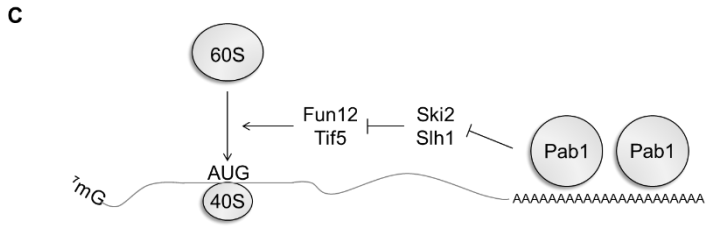


Fig.6 Pab1 promotes translation initiation. A) Representation of the closed loop structure: eIF4G bridges the two ends of the mRNA by binding eIF4E (associated to the 5' cap) and two molecules of Pab1 (associated to the 3' poly(A) tail), enabling the formation of a stable closed loop. B) The strength of Pab1 binding at 5' UTR regulates IRES-mediated translation initiation. When poly(A) tracts at the 5'UTR of a mRNA are longer than or equal to 12 adenosines, Pab1 tightly binds them and the corresponding protein abundance is low; conversely, poly(A) tracts shorter than 12 adenosines are weakly bound by Pab1 and protein abundance is high. C) Pab1 binding to poly(A) tail blocks the Ski2/Slh1 inhibition on the translation initiation factors eIF5/eIF5B, thus stimulating 60S joining (redrawn from (Searfoss *et al.*, 2001)).

Pab1 can also stimulate translation independently on its poly(A) tail bound-state. *In vitro* studies on mRNAs lacking the 3' poly(A) tail showed that Pab1 significantly increases the apparent affinity of eIF4E for the 5' cap; this effect is independent on the binding to both 3' poly(A) and eIF4G, and it likely correlates with the formation of a more compact and stable eIF4E-mRNA complex (O'Leary *et al.*, 2013). Therefore, Pab1 may *trans*-activate translation by directly interacting with initiation factors or other regulatory proteins (Otero *et al.*, 1999, Eliseeva *et al.*, 2013).

Pab1 might also negatively regulate translation initiation by binding specific 3' UTR sequences (Vasudevan *et al.*, 2005, Chritton and Wickens, 2011). In fact, *in vitro* cell-free translation assays demonstrated that Pab1 can repress translation of *MFA2* mRNA when glucose is used as carbon source for yeast growth; this regulation depends on Pab1 binding to AU-rich elements (AREs) at the 3' UTR (Vasudevan *et al.*, 2005). Further, Pab1 is implicated in the Puf5-mediated translation repression of several mRNAs, which depends on

the presence of both the PUF binding site and oligo(A) tracts at the 3' UTR (Chritton and Wickens, 2011).

Pab1 negatively regulates translation termination

The first indication of a possible role for the *S. cerevisiae* poly(A) binding protein in translation termination came in 2002, when it was demonstrated its physical interaction with the translation termination factor eRF3 (Sup35) (Cosson *et al.*, 2002). Upon GTP hydrolysis, eRF3 enhances stop codon recognition by eRF1 (Sup45) and then dissociates from the ribosome allowing the ATPase Rli1 to bind eRF1 and to promote both peptide and 60S subunit release (Dever *et al.*, 2016). The domains involved in the interaction of Pab1 with eRF3 are the P and the C domains (Roque *et al.*, 2015). Several hypotheses have been proposed for explaining the biological function of this association, among which a role in nonsense mediated decay (NMD) (Amrani *et al.*, 2004) and, at least in human, in deadenylation (Funakoshi *et al.*, 2007). However, both *in vitro* and *in vivo* experiments in *S. cerevisiae* demonstrated that the absence of this interaction does not affect neither the mRNA stability nor the NMD, but it influences translation termination (Roque *et al.*, 2015). More in detail, the lack of Pab1-eRF3 association significantly enhances the efficacy of stop codon recognition and, thus, of translation termination (Figure 7A) (Roque *et al.*, 2015). This result suggest that, at least partially and through a still uncharacterized mechanism, Pab1-eRF3 interaction might negatively regulate translation termination, thus allowing a basal level of translational readthrough (Figure 7B) (Roque *et al.*, 2015). This is a process that enables the ribosome to pass through the STOP codon by random insertion of a tRNA and to continue translation to the next STOP codon in the same reading frame (Dabrowski *et al.*, 2015).

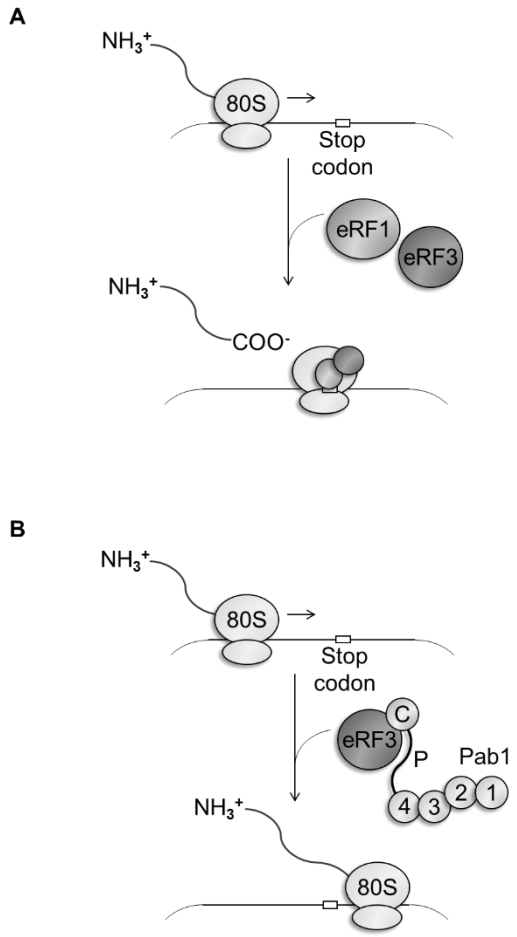


Fig.7 Hypothetical Pab1 effect on translation termination. A) eRF1-eRF3 association allows efficient stop codon recognition and translation arrest; B) Pab1 binding to eRF3 might impair translation termination, thus allowing a basal level of translational readthrough.

Pab1 affects cell physiology

Because of its broad functions in controlling global mRNA translation and turnover, Pab1 affects several mechanisms regulating the entire cell physiology.

As an example, a role of Pab1 in the cell cycle progression has been proposed, since its overexpression suppresses the conditional lethality of *bck2 swi6-ts* cells, which are characterized by cell cycle arrest (Flick and Wittenberg, 2005). Currently, the mechanisms responsible for this suppression are not known and a direct interaction of Pab1 with cell cycle-related factors has not been (already) identified. However, its general role in regulating proper deadenylation and mRNAs translation might indirectly contribute to the control of the cell cycle progression by temporally limiting the expression of cell cycle-related genes (Beilharz and Preiss, 2007). In fact, Ccr4 and Pan2 deadenylases, whose activity is regulated by Pab1 (see above), have been described to genetically interact with cell cycle regulators (Hammet *et al.*, 2002, Westmoreland *et al.*, 2004, Woolstencroft *et al.*, 2006),

By regulating mRNA metabolism, Pab1 has also an important role in cell physiology during stress. *S. cerevisiae* reacts to stressful conditions by reprogramming gene expression at transcriptional and post-transcriptional levels (Arribere *et al.*, 2011). Post-transcriptional regulation includes mRNA export, translation and decay (Arribere *et al.*, 2011), as well as the formation of intracellular protein aggregates with different size, composition and physical properties (Buchan and Parker, 2009, Wallace *et al.*, 2015, Protter and Parker, 2016, Riback *et al.*, 2017). Interestingly, Pab1 influences these phenomena by forming quinary assemblies, by affecting stress granules (SGs) formation and by stabilizing some mRNAs to allow rapid translation re-initiation after stress.

Quinary structures are transient protein assemblies that lack a fixed-stoichiometry and are kept together thanks to multivalent interactions and low-complexity sequences. (Wallace *et al.*, 2015). Their formation, which is dependent on phase-separation phenomena, is promoted by mild heat-stress conditions and, differently from deleterious misfolded-protein aggregates, they can maintain activity and re-solubilize, thus resulting in improved cells fitness

(Wallace *et al.*, 2015, Riback *et al.*, 2017). The Pab1 ability to undergo *in vivo* phase separation and to form these structures is mainly related to electrostatic forces associated to RRM domains. Despite the P domain is dispensable for the assembly of the described structure, it might have the role as a biophysical regulator, since mutations that reduce its hydrophobicity decrease Pab1 ability to phase separate and correlate with a reduced fitness in mild heat-stress conditions (Riback *et al.*, 2017). These results point out a key role of Pab1 in cellular adaptation to environmental stresses.

While quinary structures do assemble in mild heat-stress conditions, when challenged with more severe stress (e.g. heat-shock at 46°C or glucose starvation) cells accumulate mRNAs in untranslated mRNP granules referred to as stress granules (SGs) and processing bodies (PBs) (Buchan and Parker, 2009, Protter and Parker, 2016). SGs are composed of factors involved in translation initiation (e.g. 40S ribosomal subunit and eIF4G) and are believed to be sites for a rapid protein synthesis re-initiation after stress recovery, while PBs contain mRNAs that will be principally degraded, being composed of proteins involved in mRNA decay and translation repression (Buchan and Parker, 2009, Grousl *et al.*, 2009, Buchan *et al.*, 2011, Giménez-Barcons and Díez, 2011, Decker and Parker, 2012). Interestingly, according to the so-called "mRNA cycle", transcripts can be exchanged between these two aggregates and/or return to translation (Buchan and Parker, 2009, Protter and Parker, 2016).

As arguable from above, Pab1 is a typical component of SGs, commonly used as their marker (Swisher and Parker, 2010, Martani *et al.*, 2015, Jain *et al.*, 2016, Wheeler *et al.*, 2016, Brambilla *et al.*, 2017). Its recruitment into SGs is ascribable to the presence of all the four RRM domains, while both the P and the C domains are not strictly required for this process (Brambilla *et al.*, 2017). However, the P domain significantly enhances the association into SGs of Pab1 synthetic versions lacking one or more RRMs, while the C domain has the

opposite effect (Brambilla *et al.*, 2017). The role of Pab1 in SGs formation has not been fully elucidated yet, because the need of suppressors to overcome the lethality of its deletion limits the possibility to accurately define it. In fact, depending on the suppressor strains, different phenotypes were obtained in terms of SGs and PBs number after glucose starvation. For example, in the *pab1Δ spb2Δ* strain, a lower number of SGs and higher number of PBs compared to the *spb2Δ* single deleted strain were detected, suggesting that Pab1 might stimulate the transition of poly(A)$^+$ mRNAs from PBs to SGs. However, since SGs were still present in this suppressor strain, and since no changes in their number were detected in another bypass suppressor strain (*pab1Δ pat1-2*), it was concluded that Pab1 is dispensable for SGs formation (Swisher and Parker, 2010).

Finally, a link between Pab1 and translation re-initiation after glucose starvation was proposed. When cells are glucose-starved, the level of "bulk" translation is reduced, even though some mRNAs are still translated (possibly to adapt to the low-glucose environment) (Arribere *et al.*, 2011). Translationally repressed mRNAs encoding for mitochondrial proteins display a positive correlation among *i)* longer poly(A) tails, *ii)* Pab1 association and *iii)* the capacity to be translated again when glucose is added to cells. These results suggest a role of Pab1 in stabilizing non-translating mRNAs from decapping during stress, and highlight its importance for translation reestablishment (Arribere *et al.*, 2011).

Pab1 at the service of applied research

The *S. cerevisiae* Pab1 has been especially studied to unveil the cellular functions and the corresponding mechanisms in which it is implicated. Recently, it emerged also as a promising target for studying human diseases and for boosting yeast robustness. In one study, yeasts expressing Pab1 proteins with

polyalanine tracts longer than the wild type Pab1 have been used as models for studying aggregation and cellular dysfunctions associated to polyalanine-expansion, which occur in some human disorders such as the oculopharyngeal muscular dystrophy (OPMD) (Konopka *et al.*, 2011). In this work, the substitution of RNA-binding residues in Pab1 polyalanine-expanded species allowed the description of a model for the toxicity of polyalanine-expansions. In another work, Pab1 has been proposed as promising target for increasing yeast stress tolerance. Through the modulation of *PAB1* expression and the screening of *pab1* mutant libraries, some mutants with increased robustness towards acetic acid have been isolated (Martani *et al.*, 2015).

Conclusions

In the last years, the *S. cerevisiae* poly(A) binding protein has been extensively studied to disclose its cellular functions, pointing out the complex network in which Pab1 plays a pivotal role as a master regulator of the mRNA metabolism and of the entire cell physiology. Interestingly, novel findings about its functions are continuously emerging. For example, a possible role in the regulation of transcription and/or in the epigenetic control of chromatin function has been proposed, due to the discovery of Pab1 interactions with DNA at the *ENO2* and *GAL1* promoter regions (Byrum *et al.*, 2012, Guillen-Ahlers *et al.*, 2016).

Overall, the data reported in this review highlight that the physiological importance of Pab1 and its potential applications are undoubtedly larger than expected and far from being described.

References

Allen NP, Patel SS, Huang L, Chalkley RJ, Burlingame A, Lutzmann M, Hurt EC and Rexach M (2002) Deciphering networks of protein interactions at the nuclear pore complex. *Mol Cell Proteomics* **1**: 930-946.

Amrani N, Minet M, Le Gouar M, Lacroute F and Wyers F (1997) Yeast Pab1 interacts with Rna15 and participates in the control of the poly(A) tail length *in vitro*. *Mol Cell Biol* **17**: 3694-3701.

Amrani N, Ganesan R, Kervestin S, Mangus DA, Ghosh S and Jacobson A (2004) A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. *Nature* **432**: 112-118.

Archer SK, Shirokikh NE, Hallwirth CV, Beilharz TH and Preiss T (2015) Probing the closed-loop model of mRNA translation in living cells. *RNA Biol* **12**: 248-254.

Arribere JA, Doudna JA and Gilbert WV (2011) Reconsidering movement of eukaryotic mRNAs between polysomes and P bodies. *Mol Cell* **44**: 745-758.

Beelman CA, Stevens A, Caponigro G, LaGrandeur TE, Hatfield L, Fortner DM and Parker R (1996) An essential component of the decapping enzyme required for normal rates of mRNA turnover. *Nature* **382**: 642-646.

Beilharz TH and Preiss T (2007) Widespread use of poly(A) tail length control to accentuate expression of the yeast transcriptome. *RNA*. **13**(7):982-97.

Boeck R, Lapeyre B, Brown CE and Sachs AB (1998) Capped mRNA degradation intermediates accumulate in the yeast spb8-2 mutant. *Mol Cell Biol* **18**: 5062-5072.

Brambilla M, Martani F and Branduardi P (2017) The recruitment of the *Saccharomyces cerevisiae* poly(A)-binding protein into stress granules: new insights into the contribution of the different protein domains. *FEMS Yeast Res* **17**.

Brown CE and Sachs AB (1998) Poly(A) tail length control in *Saccharomyces cerevisiae* occurs by message-specific deadenylation. *Mol Cell Biol* **18**: 6548-6559.

Brune C, Munchel SE, Fischer N, Podtelejnikov AV and Weis K (2005) Yeast poly(A)-binding protein Pab1 shuttles between the nucleus and the cytoplasm and functions in mRNA export. *RNA* **11**: 517-531.

Buchan JR and Parker R (2009) Eukaryotic stress granules: the ins and outs of translation. *Mol Cell* **36**: 932-941.

Buchan JR, Yoon JH and Parker R (2011) Stress-specific composition, assembly and kinetics of stress granules in *Saccharomyces cerevisiae*. *J Cell Sci* **124**: 228-239.

Burd CG, Matunis EL and Dreyfuss G (1991) The multiple RNA-binding domains of the mRNA poly(A)-binding protein have different RNA-binding activities. *Mol Cell Biol* **11**: 3419-3424.

Byrum SD, Raman A, Taverna SD and Tackett AJ (2012) ChAP-MS: a method for identification of proteins and histone posttranslational modifications at a single genomic locus. *Cell Rep* **2**: 198-205.

Caponigro G and Parker R (1995) Multiple functions for the poly(A)-binding protein in mRNA decapping and deadenylation in yeast. *Genes Dev* **9**: 2421-2432.

Chritton JJ and Wickens M (2011) A role for the poly(A)-binding protein Pab1p in PUF protein-mediated repression. *J Biol Chem* **286**: 33268-33278.

Cosson B, Couturier A, Chabelskaya S, Kiktev D, Inge-Vechtomov S, Philippe M and Zhouravleva G (2002) Poly(A)-binding protein acts in translation termination via eukaryotic release factor 3 interaction and does not influence [PSI(+)] propagation. *Mol Cell Biol* **22**: 3301-3315.

Costello J, Castelli LM, Rowe W, *et al.* (2015) Global mRNA selection mechanisms for translation initiation. *Genome Biol* **16**: 10.

Dabrowski M, Bukowy-Bieryllo Z and Zietkiewicz E (2015) Translational readthrough potential of natural termination codons in eucaryotes--The impact of RNA sequence. *RNA Biol* **12**: 950-958.

Deardorff JA and Sachs AB (1997) Differential effects of aromatic and charged residue substitutions in the RNA binding domains of the yeast poly(A)-binding protein. *J Mol Biol* **269**: 67-81.

Decker CJ and Parker R (2012) P-bodies and stress granules: possible roles in the control of translation and mRNA degradation. *Cold Spring Harb Perspect Biol* **4**: a012286.

Dever TE, Kinzy TG and Pavitt GD (2016) Mechanism and Regulation of Protein Synthesis in *Saccharomyces cerevisiae*. *Genetics* **203**: 65-107.

Dunn EF, Hammell CM, Hodge CA and Cole CN (2005) Yeast poly(A)-binding protein, Pab1, and PAN, a poly(A) nuclease complex recruited by Pab1, connect mRNA biogenesis to export. *Genes Dev* **19**: 90-103.

Eliseeva IA, Lyabin DN and Ovchinnikov LP (2013) Poly(A)-binding proteins: structure, domain organization, and activity regulation. *Biochemistry (Mosc)* **78**: 1377-1391.

Farley AR, Powell DW, Weaver CM, Jennings JL and Link AJ (2011) Assessing the components of the eIF3 complex and their phosphorylation status. *J Proteome Res* **10**: 1481-1494.

Flick K and Wittenberg C (2005) Multiple pathways for suppression of mutants affecting G1-specific transcription in *Saccharomyces cerevisiae*. *Genetics* **169**: 37-49.

Funakoshi Y, Doi Y, Hosoda N, Uchida N, Osawa M, Shimada I, Tsujimoto M, Suzuki T, Katada T and Hoshino S (2007) Mechanism of mRNA deadenylation: evidence for a molecular interplay between translation termination factor eRF3 and mRNA deadenylases. *Genes Dev* **21**: 3135-3148.

Gilbert WV, Zhou K, Butler TK and Doudna JA (2007) Cap-independent translation is required for starvation-induced differentiation in yeast. *Science* **317**: 1224-1227.

Giménez-Barcons M and Díez J (2011) Yeast processing bodies and stress granules: self-assembly ribonucleoprotein particles. *Microb Cell Fact* **10**: 73.

Goss DJ and Kleiman FE (2013) Poly(A) binding proteins: are they all created equal? *Wiley Interdiscip Rev RNA* **4**: 167-179.

Grossi de Sa MF, Standart N, Martins de Sa C, Akhayat O, Huesca M and Scherrer K (1988) The poly(A)-binding protein facilitates *in vitro* translation of poly(A)-rich mRNA. *Eur J Biochem* **176**: 521-526.

Grousl T, Ivanov P, Frýdlová I, *et al.* (2009) Robust heat shock induces eIF2alpha-phosphorylation-independent assembly of stress granules containing eIF3 and 40S ribosomal subunits in budding yeast, *Saccharomyces cerevisiae*. *J Cell Sci* **122**: 2078-2088.

Guillen-Ahlers H, Rao PK, Levenstein ME, *et al.* (2016) HyCCAPP as a tool to characterize promoter DNA-protein interactions in *Saccharomyces cerevisiae*. *Genomics* **107**: 267-273.

Hammet A, Pike BL and Heierhorst J (2002) Posttranscriptional regulation of the *RAD5* DNA repair gene by the Dun1 kinase and the Pan2-Pan3 poly(A)-nuclease complex contributes to survival of replication blocks. *J Biol Chem* **277**: 22469-22474.

Hatfield L, Beelman CA, Stevens A and Parker R (1996) Mutations in trans-acting factors affecting mRNA decapping in *Saccharomyces cerevisiae*. *Mol Cell Biol* **16**: 5830-5838.

Huch S and Nissan T (2014) Interrelations between translation and general mRNA degradation in yeast. *Wiley Interdiscip Rev RNA* **5**: 747-763.

Jacobson A and Favreau M (1983) Possible involvement of poly(A) in protein synthesis. *Nucleic Acids Res* **11**: 6353-6368.

Jain S, Wheeler JR, Walters RW, Agrawal A, Barsic A and Parker R (2016) ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. *Cell* **164**: 487-498.

Kervestin S and Jacobson A (2012) NMD: a multifaceted response to premature translational termination. *Nat Rev Mol Cell Biol* **13**: 700-712.

Komar AA, Mazumder B and Merrick WC (2012) A new framework for understanding IRES-mediated translation. *Gene* **502**: 75-86.

Konopka CA, Locke MN, Gallagher PS, Pham N, Hart MP, Walker CJ, *et al*. (2011) A yeast model for polyalanine-expansion aggregation and toxicity. *Mol Biol Cell*. **22**(12):1971-84.

Kozlov G, Siddiqui N, Coillet-Matillon S, Trempe JF, Ekiel I, Sprules T and Gehring K (2002) Solution structure of the orphan PABC domain from *Saccharomyces cerevisiae* poly(A)-binding protein. *J Biol Chem* **277**: 22822-22828.

Lee D, Ohn T, Chiang YC, Quigley G, Yao G, Liu Y and Denis CL (2010) PUF3 acceleration of deadenylation *in vivo* can operate independently of CCR4 activity, possibly involving effects on the PAB1-mRNP structure. *J Mol Biol* **399**: 562-575.

Low JK, Hart-Smith G, Erce MA and Wilkins MR (2013) Analysis of the proteome of *Saccharomyces cerevisiae* for methylarginine. *J Proteome Res* **12**: 3884-3899.

Low JK, Hart-Smith G, Erce MA and Wilkins MR (2014) The *Saccharomyces cerevisiae* poly(A)-binding protein is subject to multiple post-translational modifications, including the methylation of glutamic acid. *Biochem Biophys Res Commun* **443**: 543-548.

Mangus DA, Amrani N and Jacobson A (1998) Pbp1p, a factor interacting with *Saccharomyces cerevisiae* poly(A)-binding protein, regulates polyadenylation. *Mol Cell Biol* **18**: 7383-7396.

Mangus DA, Evans MC, Agrin NS, Smith M, Gongidi P and Jacobson A (2004) Positive and negative regulation of poly(A) nuclease. *Mol Cell Biol* **24**: 5521-5533.

Maris C, Dominguez C and Allain FH (2005) The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J* **272**: 2118-2131.

Martani F, Marano F, Bertacchi S, Porro D and Branduardi P (2015) The *Saccharomyces cerevisiae* poly(A) binding protein Pab1 as a target for eliciting stress tolerant phenotypes. *Sci Rep* **5**: 18318.

Meaux S, van Hoof A and Baker KE (2008) Nonsense-mediated mRNA decay in yeast does not require PAB1 or a poly(A) tail. *Mol Cell* **29**: 134-140.

Melamed D, Young DL, Miller CR and Fields S (2015) Combining natural sequence variation with high throughput mutational data to reveal protein interaction sites. *PLoS Genet* **11**: e1004918.

Minvielle-Sebastia L, Preker PJ, Wiederkehr T, Strahm Y and Keller W (1997) The major yeast poly(A)-binding protein is associated with cleavage factor IA and functions in premessenger RNA 3'-end formation. *Proc Natl Acad Sci U S A* **94**: 7897-7902.

Niño CA, Hérissant L, Babour A and Dargemont C (2013) mRNA nuclear export in yeast. *Chem Rev* **113**: 8523-8545.

O'Leary SE, Petrov A, Chen J and Puglisi JD (2013) Dynamic recognition of the mRNA cap by *Saccharomyces cerevisiae* eIF4E. *Structure* **21**: 2197-2207.

Otero LJ, Ashe MP and Sachs AB (1999) The yeast poly(A)-binding protein Pab1p stimulates *in vitro* poly(A)-dependent and cap-dependent translation by distinct mechanisms. *EMBO J* **18**: 3153-3163.

Parker R (2012) RNA degradation in *Saccharomyces cerevisae*. *Genetics* **191**: 671-702.

Pintard L, Kressler D and Lapeyre B (2000) Spb1p is a yeast nucleolar protein associated with Nop1p and Nop58p that is able to bind S-adenosyl-L-methionine *in vitro*. *Mol Cell Biol* **20**: 1370-1381.

Plank TD and Kieft JS (2012) The structures of nonprotein-coding RNAs that drive internal ribosome entry site function. *Wiley Interdiscip Rev RNA* **3**: 195-212.

Protter DS and Parker R (2016) Principles and Properties of Stress Granules. *Trends Cell Biol* **26**: 668-679.

Reineke LC and Merrick WC (2009) Characterization of the functional role of nucleotides within the URE2 IRES element and the requirements for eIF2A-mediated repression. *RNA* **15**: 2264-2277.

Riback JA, Katanski CD, Kear-Scott JL, Pilipenko EV, Rojek AE, Sosnick TR and Drummond DA (2017) Stress-Triggered Phase Separation Is an Adaptive, Evolutionarily Tuned Response. *Cell* **168**: 1028-1040.e1019.

Richardson R, Denis CL, Zhang C, Nielsen ME, Chiang YC, Kierkegaard M, Wang X, Lee DJ, Andersen JS and Yao G (2012) Mass spectrometric

identification of proteins that interact through specific domains of the poly(A) binding protein. *Mol Genet Genomics* **287**: 711-730.

Roque S, Cerciat M, Gaugué I, Mora L, Floch AG, de Zamaroczy M, Heurgué-Hamard V and Kervestin S (2015) Interaction between the poly(A)-binding protein Pab1 and the eukaryotic release factor eRF3 regulates translation termination but not mRNA decay in *Saccharomyces cerevisiae*. *RNA* **21**: 124-134.

Sachs AB and Davis RW (1989) The poly(A) binding protein is required for poly(A) shortening and 60S ribosomal subunit-dependent translation initiation. *Cell* **58**: 857-867.

Sachs AB and Davis RW (1990) Translation initiation and ribosomal biogenesis: involvement of a putative rRNA helicase and RPL46. *Science* **247**: 1077-1079.

Sachs AB, Davis RW and Kornberg RD (1987) A single domain of yeast poly(A)-binding protein is necessary and sufficient for RNA binding and cell viability. *Mol Cell Biol* **7**: 3268-3276.

Schmid M, Poulsen MB, Olszewski P, Pelechano V, Saguez C, Gupta I, Steinmetz LM, Moore C and Jensen TH (2012) Rrp6p controls mRNA poly(A) tail length and its decoration with poly(A) binding proteins. *Mol Cell* **47**: 267-280.

Searfoss A, Dever TE and Wickner R (2001) Linking the 3' poly(A) tail to the subunit joining step of translation initiation: relations of Pab1p, eukaryotic translation initiation factor 5b (Fun12p), and Ski2p-Slh1p. *Mol Cell Biol* **21**: 4900-4908.

Siddiqui N, Mangus DA, Chang TC, Palermino JM, Shyu AB and Gehring K (2007) Poly(A) nuclease interacts with the C-terminal domain of polyadenylate-binding protein domain from poly(A)-binding protein. *J Biol Chem* **282**: 25067-25075.

Simón E and Séraphin B (2007) A specific role for the C-terminal region of the Poly(A)-binding protein in mRNA decay. *Nucleic Acids Res* **35**: 6017-6028.

Smith RW, Blee TK and Gray NK (2014) Poly(A)-binding proteins are required for diverse biological processes in metazoans. *Biochem Soc Trans* **42**: 1229-1237.

Soucek S, Corbett AH and Fasken MB (2012) The long and the short of it: the role of the zinc finger polyadenosine RNA binding protein, Nab2, in control of poly(A) tail length. *Biochim Biophys Acta* **1819**: 546-554.

Swisher KD and Parker R (2010) Localization to, and effects of Pbp1, Pbp4, Lsm12, Dhh1, and Pab1 on stress granules in *Saccharomyces cerevisiae. PLoS One* **5**: e10006.

Tucker M, Staples RR, Valencia-Sanchez MA, Muhlrad D and Parker R (2002) Ccr4p is the catalytic subunit of a Ccr4p/Pop2p/Notp mRNA deadenylase complex in *Saccharomyces cerevisiae*. *EMBO J* **21**: 1427-1436.

Vasudevan S, Garneau N, Tu Khounh D and Peltz SW (2005) p38 mitogen-activated protein kinase/Hog1p regulates translation of the AU-rich-element-bearing MFA2 transcript. *Mol Cell Biol* **25**: 9753-9763.

Wahle E and Winkler GS (2013) RNA decay machines: deadenylation by the Ccr4-not and Pan2-Pan3 complexes. *Biochim Biophys Acta* **1829**: 561-570.

Wallace EW, Kear-Scott JL, Pilipenko EV, *et al.* (2015) Reversible, Specific, Active Aggregates of Endogenous Proteins Assemble upon Heat Stress. *Cell* **162**: 1286-1298.

Wang X, Zhang C, Chiang YC, *et al.* (2012) Use of the novel technique of analytical ultracentrifugation with fluorescence detection system identifies a 77S monosomal translation complex. *Protein Sci* **21**: 1253-1268.

Westmoreland TJ, Marks JR, Olson JA, Thompson EM, Resnick MA and Bennett CB (2004) Cell cycle progression in G1 and S phases is CCR4 dependent following ionizing radiation or replication stress in *Saccharomyces cerevisiae. Eukaryot Cell* **3**: 430-446.

Wheeler JR, Matheny T, Jain S, Abrisch R and Parker R (2016) Distinct stages in stress granule assembly and disassembly. *Elife* **5**.

Wolf J, Valkov E, Allen MD, *et al.* (2014) Structural basis for Pan3 binding to Pan2 and its function in mRNA recruitment and deadenylation. *EMBO J* **33**: 1514-1526.

Woolstencroft RN, Beilharz TH, Cook MA, Preiss T, Durocher D and Tyers M (2006) Ccr4 contributes to tolerance of replication stress through control of *CRT1* mRNA poly(A) tail length. *J Cell Sci* **119**: 5178-5192.

Wyers F, Minet M, Dufour ME, Vo LT and Lacroute F (2000) Deletion of the *PAT1* gene affects translation initiation and suppresses a *PAB1* gene deletion in yeast. *Mol Cell Biol* **20**: 3538-3549.

Xia X, MacKay V, Yao X, Wu J, Miura F, Ito T and Morris DR (2011) Translation initiation: a regulatory role for poly(A) tracts in front of the AUG codon in *Saccharomyces cerevisiae. Genetics* **189**: 469-478.

Xie J, Kozlov G and Gehring K (2014) The "tale" of poly(A) binding protein: the MLLE domain and PAM2-containing proteins. *Biochim Biophys Acta* **1839**: 1062-1068.

Yao G, Chiang YC, Zhang C, Lee DJ, Laue TM and Denis CL (2007) PAB1 self-association precludes its binding to poly(A), thereby accelerating CCR4 deadenylation *in vivo*. *Mol Cell Biol* **27**: 6243-6253.

Zhang C, Lee DJ, Chiang YC, Richardson R, Park S, Wang X, Laue TM and Denis CL (2013) The RRM1 domain of the poly(A)-binding protein from *Saccharomyces cerevisiae* is critical to control of mRNA deadenylation. *Mol Genet Genomics* **288**: 401-412.

Zhong T and Arndt KT (1993) The yeast SIS1 protein, a DnaJ homolog, is required for the initiation of translation. *Cell* **73**: 1175-1186.

Is PAB1 an effective target for evoking industrially relevant phenotypes?

As previously anticipated, the selection of an appropriate candidate for the induction of a cellular rewiring, aimed at obtaining industrially relevant phenotypes, is not a trivial task to accomplish. In this context, our laboratory previously demonstrated that Pab1 modulation(s) allowed the isolation of strains with increased robustness to acetic acid, a potent inhibitor in lignocellulosic hydrolysates (Martani *et al.*, 2015). Hence, the possibility to rewire yeast cells through Pab1 was further evaluated in the following chapter. The main goal was to generate complex phenotypes and to isolate among them thermotolerant *Saccharomyces cerevisiae* strains, which are desired for the development of efficient SSF and SSCF processes (see Introduction).

Chapter 3

Selected mutants in the *Saccharomyces cerevisiae* poly(A) binding protein Pab1 can improve yeast thermotolerance

Manuscript in preparation

Brambilla Marco, Vitangeli Ilaria, Mento Alfredo, Martani Francesca, Branduardi Paola

Introduction

The yeast *Saccharomyces cerevisiae* is currently one of the most prominent and utilized cell factory. Indeed, apart from its widely known usage in bakery, brewery and winemaking (Kavščck *et al.*, 2015), nowadays it is employed for the sustainable industrial production of bioethanol, some pharma products, and nutraceuticals (Hong and Nielsen, 2012). Moreover, several recombinant *S. cerevisiae* strains have been constructed for the production of many heterologous compounds such as lactic and succinic acid, vanillin, polyketides, isoprenoids, penicillin and opiates (Mattanovich *et al.*, 2014, Kavščck *et al.*, 2015, Galanie *et al.*, 2015). Aiming to scale-up these (as well as other) productions at industrial level, high yield, production and productivity must be achieved (Hong and Nielsen, 2012). However, when cultured at industrial scale, *S. cerevisiae* usually face harsh conditions such as oxidative, osmotic and thermal stress, together with the stress related to final product accumulation and nutrient depletion (Kim *et al.*, 2013, Wallace-Salinas and Gorwa-Grauslund, 2013, Doğan *et al.*, 2014, Mattanovich *et al.*, 2014). Moreover, due to the pretreatment of lignocellulose, which is the most widespread and utilized biomass in 2nd generation biorefinery processes, toxic compounds such as organic acid, furan derivatives and phenols are generated. Altogether, these factors can amply affect *S. cerevisiae* viability or metabolic efficiency, thus limiting the possibility to scale-up processes that are successful at laboratory-scale (Ling *et al.*, 2014). Therefore, the development of robust yeasts is indispensable.

Of note, in order to improve the productivity of a bioethanol process from lignocellulose feedstocks, one of the most desirable trait that *S. cerevisiae* should display is a high thermotolerance. Indeed, carrying fermentation at temperature

higher than 30°C could reduce risks of contamination, the costs of cooling down the medium after sterilization and during fermentation (especially in tropical countries), as well as the cost of heating for the following distillation phase (Kim *et al.*, 2013, Gao *et al.*, 2016). Moreover, in a Simultaneous Saccharification and Fermentation (SSF) process, high temperature could improve the efficiency of saccharification, which is optimal around 50°C (Wallace-Salinas and Gorwa-Grauslund, 2013, Gao *et al.*, 2016). Considering all these benefits, many efforts have been done for improving *S. cerevisiae* thermotolerance, including both random and rational strategies, as well as heterologous expression of heat-inducible genes from thermophiles (reviewed in (Gao *et al.*, 2016)). However, to the best of our knowledge, the obtainment of more thermotolerant phenotypes has never been explored with a rewiring-based strategy.

The basic idea of this approach is to create a mutant library of a "hub" cellular element that controls expression of many genes as first, and then to transform a microorganism of interest, aiming to evoke pleiotropic effects to access novel cellular phenotypes. Next, a screening protocol is applied to isolate mutants with the desired improved features (Alper and Stephanopoulos, 2007). This strategy, named global Transcription Machinery Engineering (gTME), was already successfully employed for improving ethanol tolerance and production in *S. cerevisiae*, by selecting the transcription factor Spt15 (Alper *et al.*, 2006) or the Rpb7 subunit of RNA polymerase II (Qiu and Jiang, 2017) as hub elements.

Interestingly, we previously followed a similar approach to isolate *S. cerevisiae* strains with an improved tolerance to acetic acid, one of the most dangerous inhibitor in lignocellulosic hydrolysates. As hub element, the principal poly(A) binding protein (Pab1) was selected, as it is a master regulator of mRNA metabolism involved in the regulation of gene expression at the post-transcriptional level (Martani *et al.*, 2015). Indeed, Pab1 plays a role in mRNA 3' end processing and export from the nucleus (Brune *et al.*, 2005), in translation

initiation and termination (Costello *et al.*, 2015, Roque *et al.*, 2015), in mRNA deadenylation and decay (Parker, 2012), and in stress granules formation (Swisher and Parker, 2010). Interestingly, we proved that a rewiring methodology with a factor involved in the control of mRNA metabolism has the potential for the development of robust cell factories (Martani *et al.*, 2015). Hence, in the presented work we applied the same approach but to isolate *S. cerevisiae* strains with increased thermotolerance.

Results

Screening of thermotolerant *S. cerevisiae* strains

As already demonstrated by (Martani *et al.*, 2015), a *S. cerevisiae* strain, named Pab1(+), in which *PAB1* was over-expressed at low-dosage, was characterized by growth improvement in the presence of hydrogen peroxide or acetic acid. This strain was then considered as benchmark for the selection of further ameliorated strains obtained by transforming the parental strain with a mutant library for *PAB1* (mutagenized by error-prone PCR) (Martani *et al.*, 2015).

Here, a similar approach was repeated aiming to isolate strains with higher thermotolerance. After yeast transformation with either YCplac33*PAB1* or the plasmid mutant library, all the CFU were retrieved and cultured in liquid minimal medium. After 4 and 24 hours, cells were plated onto minimal medium and incubated at increasing temperatures from 39°C to 41°C. In parallel, 30°C was used as control. Interestingly, huge differences were observed in the CFU number depending both on the temperature and on the growth phase (Table 1). Indeed, cells plated after 4 hours of cultivation always resulted in lower CFU number compared to the corresponding cells cultivated for 24 hours: this might refer to the observation that cells approaching the stationary phase are intrinsically and generally more tolerant (Werner-Washburne *et al.*, 1993). Among all the temperature tested, 40°C was selected for further analysis, being the condition where the mutant library resulted in a higher number of CFU compared to the control Pab1(+) strain.

	Screening of cells grown for 4 hours		Screening of cells grown for 24 hours	
Conditions	**Mutant Library**	**Pab1(+)**	**Mutant Library**	**Pab1(+)**
30°C	≈ 300	≈ 300	≈ 500	≈ 500
39°C	≈ 150	≈ 150	≈ 250	≈ 250
40°C	3	0	12	5
41°C	0	0	0	0

Table 1 Numbers of colony forming units (CFU) obtained from the screening at high temperatures. CEN.PK113-5D cells, transformed with the control plasmid YCplac33*PAB1* (indicated as Pab1(+)) or the *PAB1* mutant library, were cultivated in minimal YNB medium for 4 and 24 hours, plated onto minimal YNB agar plates and then incubated at increasing temperatures (30, 39, 40 and 41°C). CFUs were counted after 3 days.

Next, the 15 isolated colonies from the mutant library were picked up and tested for their ability to grow at high temperatures by drop test. Interestingly, while no differences were observed among strains at permissive temperature, all the 15 isolated clones (indicated with "I" in Fig.1) displayed a consistent and similar growth advantage at 40, 41 and 42°C compared to the Pab1(+) strain (Fig.S1). To prove that the observed growth improvement was effectively associated to the presence of *PAB1* mutated variants, and not to other mutations eventually occurred in the plasmid or in the genome during the screening, strains were reconstructed as described in "Material and Methods". Drop tests confirmed a comparable growth advantage at high temperature also for the reconstructed strains (indicated with "R" in Fig.1), while no significant clonal differences were detected among the mutants.

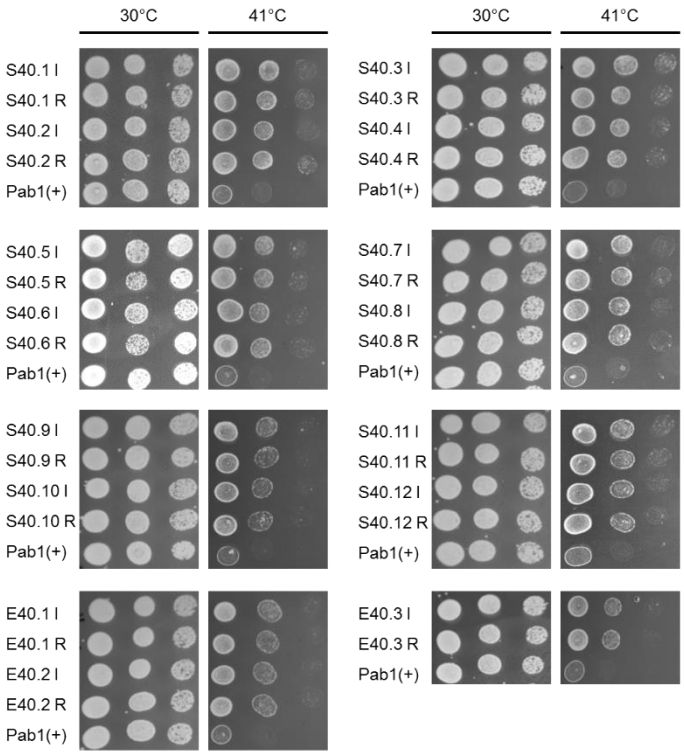


Fig.1 Evaluation of thermotolerance in selected Pab1 mutant strains by drop test. Mutants isolated from the screening (indicated with the letter "I"), reconstructed strains (indicated with the letter "R") and the control Pab1(+) strain were cultivated in minimal medium and then normalized to an OD_{660nm} of 0.4. Aliquots (5 µL) from the normalized yeast cultures or 10-fold serial dilutions were spotted onto minimal plates and incubated at 30 and 41°C for 3 days. Images are representative of three independent experiments.

Evaluation of multiple tolerance

High temperatures induce a multiplicity of cellular damages especially to proteins and membranes (Morano *et al.*, 2012), which could be caused also by other typical industrial stressors (Auesukaree, 2017). Therefore, looking for eventual multiple tolerance associated to the presence of the selected Pab1 variants, drop tests were repeated on plates supplemented with different toxic agents. No differences in growth were detected on plates supplemented with acetic acid, formic acid, hydroxymethylfurfural or hydrogen peroxide (Fig.S2),

thus confirming the specificity of the screening protocol for the isolation of more thermotolerant strains. Nevertheless, all the reconstructed strains displayed a minimal but reproducible growth improvement in ethanol-supplemented plates, which is visible after two and three days, while it is lost after six days of cultivation (Fig.2), where presumably ethanol evaporation restored permissive conditions.

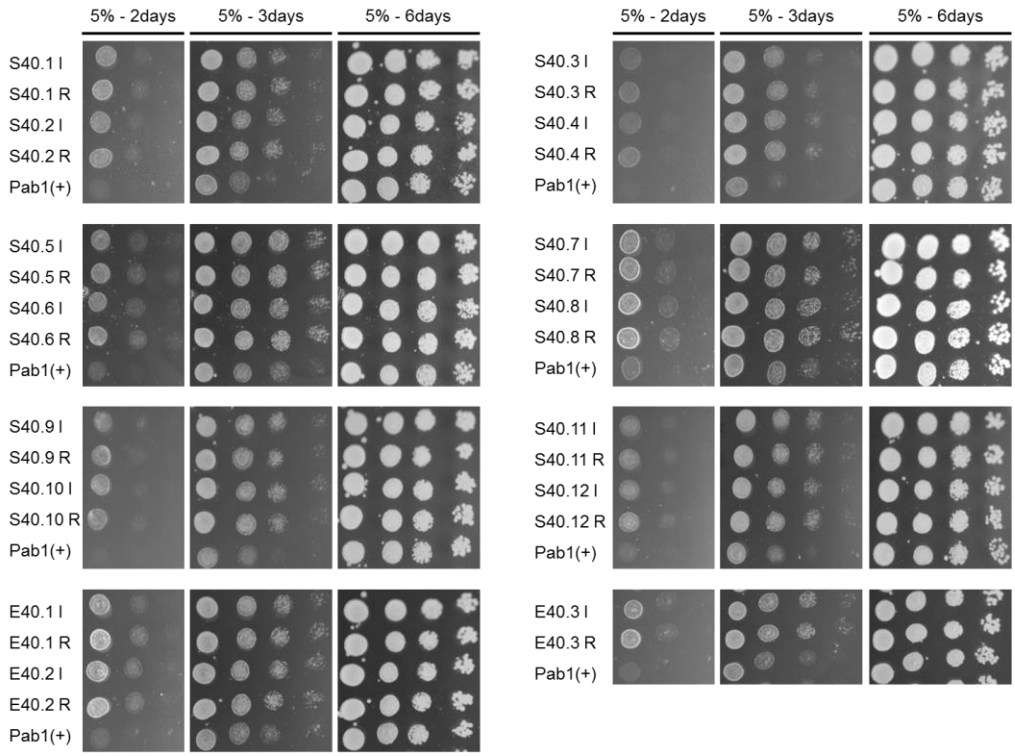


Fig.2 Evaluation of ethanol tolerance in selected Pab1 mutant strains by drop test. Mutants isolated from the screening (indicated with the letter "I"), reconstructed strains (indicated with the letter "R") and the control Pab1(+) strain were cultivated in minimal medium and then normalized to an OD_{660nm} of 0.4. Aliquots (5 µL) from the normalized yeast cultures or 10-fold serial dilutions were spotted onto minimal plates supplemented with 5% (v/v) ethanol. The strains were grown at 30°C. Images are representative of three independent experiments.

Sequencing of the *PAB1* mutated variants

Next, the *PAB1* ORF, 5'UTR and promoter of all the plasmids isolated from ameliorated strains were sequenced (Fig.S3). Strikingly, huge varieties of mutations were detected along the entire sequence and, by analyzing the resulting final proteins, no specific common patterns were observed (Fig.S4). This result suggested that the increased thermotolerance was linked to different mechanisms that were likely activated in cells depending on which variant of the protein they contain.

Interestingly, five Pab1 versions were characterized by nonsense mutations that leaded to truncations at different level of the protein. Of note, it was surprising and particularly interesting the *PAB1*-S40.7 variant, which encoded for a peptide including just the first 20 residues of the protein: MADITDKTAEQLENLNIQDD.

Analysis of the Pab1 S40.7 variant

The significant growth improvement provided by such a short peptide was quite unexpected. Therefore, the possibility that a translation readthrough event might occur (Dabrowski *et al.*, 2015), leading to the synthesis of a full-length protein ($Pab1_{FL}$), was evaluated. In this case, the resulting Pab1 variant would anyhow displayed the subsequent three missense mutations R403G, Q436H and Q572L (Fig.3).

Fig.3 Analysis of Pab1-S40.7 protein. $Pab1_{FL}$ version resulting from eventual translational readthrough at the premature stop codon (blue star), with the resulting missense mutations (yellow stars). RRM: RNA recognition motifs, P: P domain, C: C domain.

To test this possibility, the GFP sequence was cloned in frame downstream to the *PAB1* S40.7 sequence (to which the final stop codon was removed), so that a green signal would be expected in case of translational readthrough at the premature stop codon. To this purpose, the YCplac33*PAB1*-S40.7$_{FL}$-GFP plasmid was constructed and used to transform CEN.PK113-5D. Since no signal was detected by fluorescence microscopy in the resulting strain (data not shown), we confirmed that the translation of *PAB1* S40.7 effectively terminates after 20 residues.

Next, the intracellular localization of this peptide was assessed and compared to that of the endogenous Pab1 tagged with mCherry. For this purpose, the YCplac33*PAB1*-S40.7$_{20AA}$-GFP plasmid, which encodes for the S40.7 peptide tagged with GFP, was constructed and used to transform CEN.PK113-11C *PAB1-mCherry*. Interestingly, in the absence of stress the green signal resulted diffused in the entire cell, as opposed to the endogenous Pab1 that was primarily localized in the cytosol (Fig.4A). These results are expected because a Nuclear Export Sequence (NES), which is involved in the interaction with the exportin XpoI/CrmI, is present between residues 12 and 17 at the amino-terminus of the protein and it keeps Pab1 predominantly cytosolic (Brune *et al.*, 2005). However, at least the first 61 residues of Pab1 are required for this efficient interaction (Brune *et al.*, 2005); therefore, the S40.7 peptide is too short for interacting with XpoI/CrmI and, thus, it localized both in the nucleus and in the cytosol. Probably for the same reason, its possible re-localization into stress granules (*i.e.* aggregates of untranslated mRNPs that form during stressful conditions), which are induced by heat shock at 46°C (Grousl *et al.*, 2009) and that we monitored through the endogenous Pab1-mCherry marker (Swisher and Parker, 2010, Wheeler *et al.*, 2016), was not observed (Fig.4B). Therefore, this peptide is likely too short to be recruited into these aggregates.

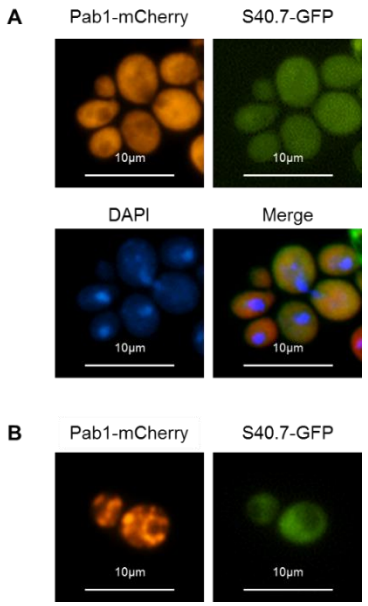


Fig.4 Intracellular localization of Pab1-mCherry and the GFP-tagged S40.7 peptide. Cells were grown until the exponential phase of growth (OD_{660nm} = 0.3–0.6) and then (A) stained with DAPI to a final concentration of 1 µg/mL, or (B) heat shocked for 50 min at 46°C.

Growth and metabolic characterization of the mutant Pab1-S40.7

To confirm the increased thermotolerance of the S40.7 reconstructed strain, batch growth kinetics in aerobic conditions were performed in bioreactors at 40°C and pH 5. Notably, while the control Pab1(+) strain did not grow at all, the S40.7 mutant reached an optical density between 3 and 3,5 at the end of the kinetic (Fig.5).

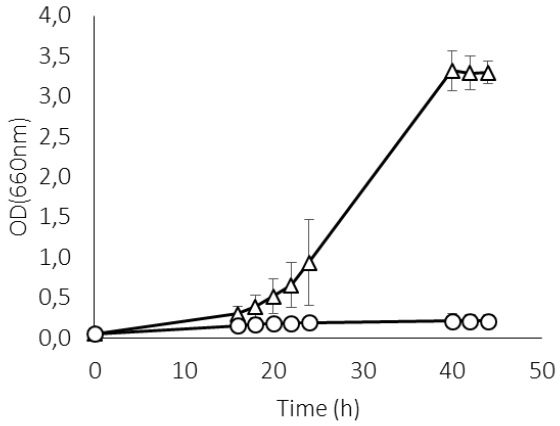


Fig.5 Growth kinetics monitored with optical density at 660_{nm}. Strains S40.7 (triangles) and Pab1(+) (circles) were cultivated in bioreactor in minimal YNB medium at 40°C, pH 5 and 25% of dissolved oxygen. Results are the mean and the standard deviation of two independent experiments.

Being *S. cerevisiae* a Crabtree positive yeast (De Deken, 1966), it can ferment glucose to ethanol even in the presence of oxygen, and started to respire it when glucose is exhausted. The S40.7 strain also show this diauxic behavior, starting to consume ethanol after its accumulation, at around 40 hours once glucose is depleted (Fig.6B-6C). In parallel, glycerol and small amounts of acetate were detected in the medium during the kinetic (Fig.6D-6E). These results not only confirmed the increased thermotolerance of the mutant, but also suggested that its metabolic capabilities were not compromised.

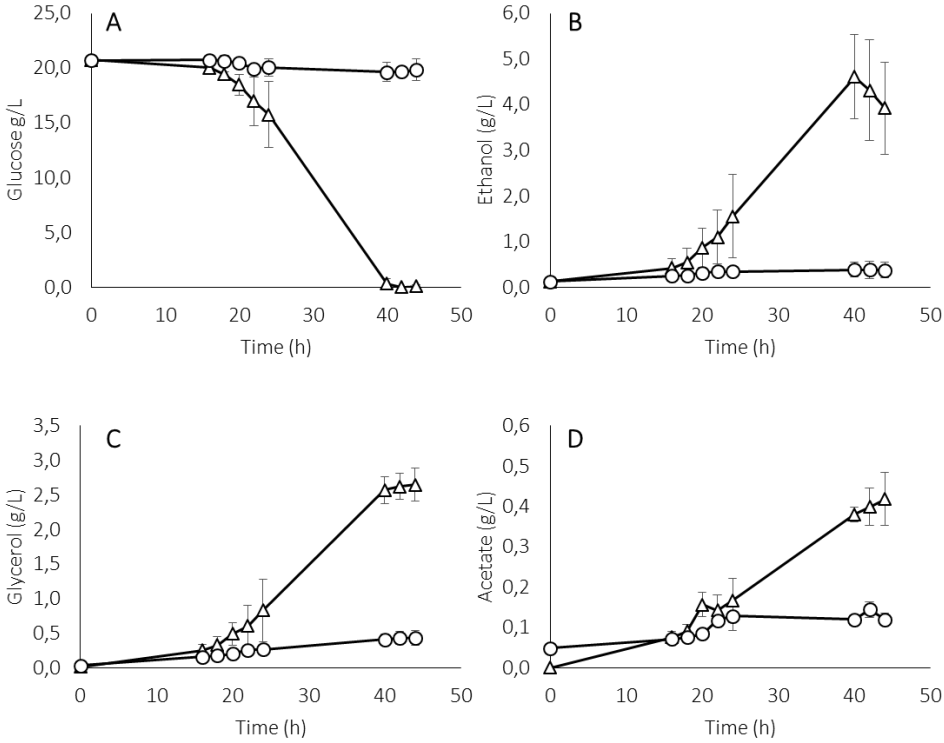


Fig.6 Metabolic profiles related to kinetics of Fig.5. Strains S40.7 (triangles) and Pab1(+) (circles) were cultivated in bioreactor in minimal YNB medium at 40°C, pH 5 and 25% of dissolved oxygen. (A) Glucose consumption, (B) ethanol production, (C) glycerol production and (D) acetate production. Results are the mean and the standard deviation of two independent experiments. Scale on Y-axis are different.

Reactive Oxygen Species (ROS) accumulation

Because high temperature are often correlated to increased oxidative stress (Morano *et al.*, 2012, Verghese *et al.*, 2012), Reactive Oxygen Species (ROS) were determined by flow cytometry using dihydrorhodamine 123 (DHR 123), which passively diffuses into the cell and, in the presence of ROS, is oxidized to fluorescent rhodamine 123. ROS accumulation was determined in exponential phase of growth during kinetics in baffled-shake flasks. Similarly to bioreactors, the S40.7 strain showed a growth advantage at 40°C compared to the

Pab1(+), which displayed a moderate growth (Fig.7B). The fact that the Pab1(+) showed different growth in bioreactor and baffled shake flasks at 40°C could be explained by the different instrumentation used and/or by the different temperature sensors of the two instruments. While no differences in growth (Fig.7A) and in ROS accumulation (Fig.7C) were detected at permissive temperature, ROS levels of the S40.7 were lower than those of the Pab1(+) strain at 40°C (Fig.7D). Moreover, in the latter it is possible to identify a subpopulation with higher ROS accumulation (Fig.7D). Altogether, these results suggested that some intracellular detoxifying mechanisms against ROS might be upregulated in this mutant.

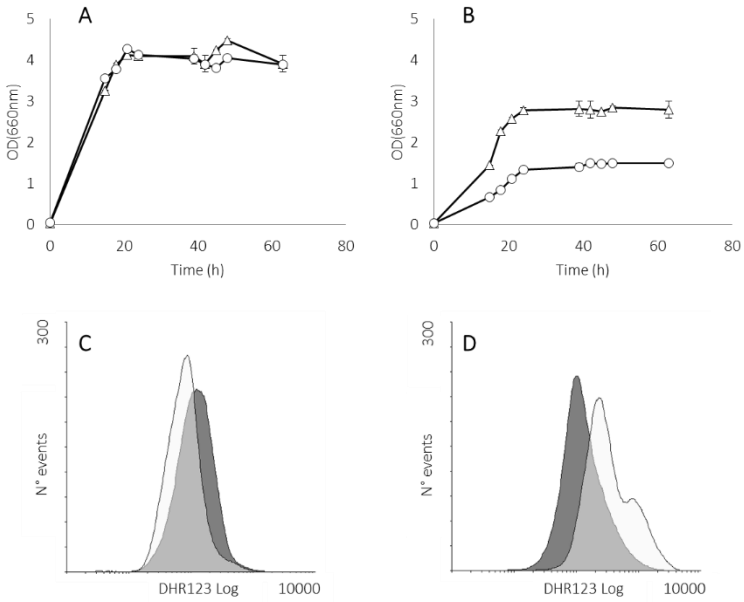


Fig.7 Growth kinetics and ROS accumulation. Growth was monitored with optical density at 660_{nm}. Strains S40.7 (triangles) and Pab1(+) (circles) were cultivated in baffled shake flasks in minimal YNB medium at 30°C (A) and 40°C (B). In parallel, ROS accumulation was measured after 16 hours from the inoculum at 30°C (C) and 40°C (D) for the S40.7 (dark grey) and Pab1(+) (light grey) strains. Dihydrorhodamine 123 (DHR123) was used as a probe. Growth kinetics are the mean and the standard deviation of two independent experiments and histograms are representative of the same two independent experiments.

Discussions

In 2nd generation bioethanol production, thermotolerant *S. cerevisiae* strains are attractive to achieve the coupling of enzyme saccharification and fermentation, as well to reduce risks of contamination, costs of cooling after sterilization and during fermentation and costs of heating for distillation (Kim *et al.*, 2013, Wallace-Salinas and Gorwa-Grauslund, 2013, Gao *et al.*, 2016). Therefore, the development of thermotolerant yeasts received great attention and was successfully explored with both rational and random approaches (Gao *et al.*, 2016). In the last decade, rewiring-based strategies have been investigated to improve phenotypes in terms of both tolerance to inhibitors and production of desired metabolites by selecting hub elements involved in the regulation of gene expression at transcriptional (Alper *et al.*, 2006, Alper and Stephanopoulos, 2007, Qiu and Jiang, 2017) or post-transcriptional levels (Martani *et al.*, 2015). Interestingly, it was recently proven that the development of thermotolerant yeast strains can be achieved through a rewiring strategy, in which the whole transcription was modulated by an artificial zinc finger protein (AZFP) library (Khatun *et al.*, 2017).

In this work, we demonstrated that the development of more thermotolerant yeast strains is effective also by reprogramming cells at the post-transcriptional level. In particular, Pab1, which is a master regulator of mRNA metabolism, is a powerful candidate for unlocking more thermotolerant *S. cerevisiae* strains. Indeed, the contemporary expression of the genomic *PAB1* and the mutated *PAB1* sequences isolated from the presented screening, improved *S. cerevisiae* resistance at high temperatures. Interestingly, a slight growth improvement was also detected when cells were challenged with ethanol, likely because similar cellular damages are induced by high temperatures and high

ethanol concentration, among which increased membrane permeability, reduced proton motive force and intracellular pH, as well as glycolysis inhibition (Gibson *et al.*, 2007).

Remarkably, sequence analysis of all the 15 *PAB1* mutated versions revealed the absence of a common pattern of mutations, which suggests that each variants engage diverse mechanisms to ensure thermotolerance, presumably by involving different interactors. For example, the Pab1-S40.3 version is characterized by three missense mutations very close to each other that fall in the P domain, which is implicated in regulating the ability of Pab1 to phase separate into quinary assemblies. The formation of these structures improves yeast fitness in mild heat-stress conditions (Riback *et al.*, 2017) and, therefore, a correlation between increased thermotolerance and this phenomenon might not be excluded. Furthermore, the releasing factor of translation termination eRF3 and the negative regulator of PAN-dependent deadenylation Pbp1 physically associate with the P domain (Roque *et al.*, 2015), thus suggesting that the modulation of these interactions, and consequently of these processes, might correlate with the increased thermotolerance. Similar considerations can be done for Pab1 versions S40.5, S40.6 and S40.16, which display higher amount of mutations in the RRM1 and the RRM2 domains. These ones are involved in the interaction with the poly(A) tails, as well as with the translation initiation factor eIF4G, whose association enhances the rate of protein synthesis (Richardson *et al.*, 2012). Moreover, the RRM1 domain is involved in the processes of Pab1 self-association (Yao *et al.*, 2007) and PUF3-mediated deadenylation (Lee *et al.*, 2010), which are both engaged in stimulating the rate of Ccr4/Pop2/NOT deadenylation. Therefore, mutations in these domains might influence these processes, thus determining improved thermotolerance with a still unknown mechanism. These variants were not characterized in this study and will surely deserve future investigations.

Strikingly, the *PAB1* version S40.7 displays a nonsense mutation that leads to a peptide comprising the first 20 residues of the protein. As demonstrated by growth kinetics in bioreactor and baffled shake-flasks, a strain expressing this peptide is characterized by a growth advantage at high temperature compared to the control Pab1(+). Conversely, no differences were detected at permissive temperature in baffled shake-flasks, while growth profiles still need to be determined at 30°C in bioreactors. This is not the first case in which a small peptide exerts a significant effect on cell physiology. Indeed, 299 ORFs of the *S. cerevisiae* genome encode for small peptides of less than hundred residues that play roles in stress response, energy metabolism, as mating pheromones, transcriptional regulators, transporters, chaperonins and so on (Kastenmayer *et al.*, 2006). As a representative example related to thermotolerance, the deletion of the *TSC3* gene, which encodes for a short peptide of 80 residues involved in the first step of sphingolipids biosynthesis (Gable *et al.*, 2000), resulted in a *S. cerevisiae* strain with increased sensitivity both to high temperatures and to the genotoxic agents hydroxyurea and bleomycin (Kastenmayer *et al.*, 2006). Similarly, the deletion of *TFB5*, encoding for a 72-residues component of the transcription factor TFIIH, resulted in a growth inhibition at 37°C and in an increased sensitivity to UV irradiation (Ranish *et al.*, 2004). Here it is not easily identifiable the mechanism through which the S40.7 peptide increases yeast thermotolerance. One can speculate that it might act at an unprecise stage of the heat-shock response, for example at a transcriptional level by stimulating the expression of heat-shock and/or antioxidant proteins or downstream to the final effectors of the response. Currently, we determined that a strain expressing the S40.7 peptide is characterized by a lower ROS content at high temperature in exponential phase of growth compared to the control. Future experiments will be required to investigate which antioxidant defenses might be upregulated to counteract ROS accumulation in the S40.7 strain.

Finally, it is important to underline that this genetic modification was realized in a prototrophic CEN.PK genetic background, which is a common laboratory strain. It is not possible to exclude that different results might be obtained in different genetic backgrounds. Currently, the BY4741, which is characterized by leucine, methionine, histidine and uracil auxotrophies, is under investigation in our laboratory to confirm or to contradict the S40.7 capability in increasing thermotolerance. As a future perspective, more robust industrial strains will be also tested.

In conclusion, at the best of our knowledge, the presented work shows for the first time the development of thermotolerant *S. cerevisiae* strains by rewiring cells at post-transcriptional level, and further confirms Pab1 as a powerful target to evoke complex phenotypes with improved features.

Materials and methods

Yeast strains

The *S. cerevisiae* genetic backgrounds already available in this study were CEN.PK113-5D (*MATa; MAL2-8c; SUC2; ura3-52*) (Entian and Kötter, 1998), Pab1(+) (Martani *et al.*, 2015) and CEN.PK113-11C *PAB1-mCherry* (*MATa; MAL2-8c; SUC2; ura3-52; his3Δ1; PAB1::PAB1-mCherry-HIS3*) (Brambilla *et al.*, 2017). All the strains constructed in this study are listed in Table 2. Yeast transformation were performed with the LiAc/PEG/ss-DNA protocol (Gietz and Woods, 2002).

Table 2. List of strains constructed in this work

Strains	Genotype	Source
CEN.PK113-5D	*MATa; MAL2-8c; SUC2; ura3-52*	A
Pab1(+)	CEN.PK113-5D [YCplac33 *PAB1*]	B
CEN.PK113-11C PAB1-mCherry	*MATa; MAL2-8c; SUC2; ura3-52; his3Δ1; PAB1::PAB1-mCherry-HIS3*	C
S40.1	CEN.PK113-5D [YCplac33 *PAB1*-S40.1]	D
S40.2	CEN.PK113-5D [YCplac33 *PAB1*-S40.2]	D
S40.3	CEN.PK113-5D [YCplac33 *PAB1*-S40.3]	D
S40.4	CEN.PK113-5D [YCplac33 *PAB1*-S40.4]	D
S40.5	CEN.PK113-5D [YCplac33 *PAB1*-S40.5]	D
S40.6	CEN.PK113-5D [YCplac33 *PAB1*-S40.6]	D

Table. 2 Continue

Strains	Genotype	Source
S40.7	CEN.PK113-5D [YCplac33 *PAB1*-S40.7]	D
S40.8	CEN.PK113-5D [YCplac33 *PAB1*-S40.8]	D
S40.9	CEN.PK113-5D [YCplac33 *PAB1*-S40.9]	D
S40.10	CEN.PK113-5D [YCplac33 *PAB1*-S40.10]	D
S40.11	CEN.PK113-5D [YCplac33 *PAB1*-S40.11]	D
S40.12	CEN.PK113-5D [YCplac33 *PAB1*-S40.12]	D
E40.1	CEN.PK113-5D [YCplac33 *PAB1*-E40.1]	D
E40.2	CEN.PK113-5D [YCplac33 PAB1-E40.2]	D
E40.3	CEN.PK113-5D [YCplac33 PAB1-E40.3]	D
$S40.7_{FL}$-GFP	CEN.PK113-5D [YCplac33PAB1-$S40.7_{FL}$-GFP]	D
$S40.7_{20AA}$-GFP	CEN.PK113-11C *PAB1-mCherry* [YCplac33*PAB1*-$S40.7_{20AA}$-GFP YCplac33*PAB1*-$S40.7_{20AA}$-GFP]	D

A) Entian and Kötter, 1998. B) Martani *et al.*, 2015. C) Brambilla *et al.*, 2017. D) This study

Plasmids construction

The plasmid mutant library was constructed as described in (Martani *et al.*, 2015), with the GeneMorph II EZClone Domain Mutagenesis Kit (Agilent Technologies, catalog #200552). To produce a low mutational frequency (0 - 4.5 mutations/kb), the first mutagenic PCR was performed with ~1µg of the plasmid YCplac33*PAB1* used as template, as suggested by manufacturer's instructions.

To reconstruct the strains isolated from the screening, plasmids were first retrieved from the isolated colonies. The isolated plasmids were *Pst*I/*Sal*I digested to recover the *PAB1* mutated sequences, which were cloned in the original YCplac33, thus yielding the reconstructed plasmids. Next, the CEN.PK113-5D was transformed with them to obtain the reconstructed strains.

To verify premature termination codon recognition, the GFP was fused in frame downstream to the *PAB1*-S40.7 sequence. This sequence was amplified from YCplac33*PAB1*-S40.7 to exclude the final stop codon with primers Pab1_PstI_FW - GAT TTA CTG CAG GTA TAT ATA TTT GCG TGT AAG TGT GTG T and Pab1_BamHI_RV - AAT ATT GGA TCC GAG CTT GCT CAG TTT GTT GTT CTT GCT C. The PCR product was *Pst*I/*Bam*HI digested and cloned in the YCplac33*PAB1*-GFP plasmid, which was previously digested with the same enzymes to remove the *PAB1* sequence and to maintain the GFP. The resulting YCplac33*PAB1*-S40.7$_{FL}$-GFP plasmid was used to transform the CEN.PK113-5D background.

To evaluate the intracellular distribution of the Pab1-S40.7 peptide, the first 20 codons of the wild type *PAB1* sequence were amplified using the YCplac33*PAB1* plasmid as template and Pab1_PstI_FW (see above) and S40.7_BamHI_REV – ACT GTC GGA TCC GGT CAT CTT GAA TAT TCA AGT TTT CCA ATT GTT C as primers. Next, the PCR product was *Pst*I/*Bam*HI digested and cloned in the YCplac33*PAB1*-GFP plasmid, which was previously digested with the same enzymes to remove the *PAB1* sequence and to maintain the GFP. The corresponding YCplac33*PAB1*-S40.7$_{20AA}$-GFP plasmid was used to transform the CEN.PK113-11C *PAB1-mCherry*.

Media, growth kinetics and metabolites determination

Yeast cultures were grown in in minimal synthetic medium with 6.7g/L Yeast Nitrogen Base without Amino Acids (YNB - BD™ Difco™) and 20g/L glucose.

For growth kinetics in bioreactors, pre-cultures were grown in 250 mL Erlenmeyer flasks containing 50 mL of medium for 24 hours at 30°C and under continuous shaking (160rpm). Next cells were inoculated at OD_{660nm} 0.05 in 2.0 L bioreactors (Sartorius Stedim BIOSTAT® Bplus, Germany) containing 1L of medium. The aeration rate, agitation and temperature were set to 1vvm, 300rpm (in cascade to 25% of dissolved oxygen) and 40°C, respectively. The pH was maintained at 5 by automatic pumping of 4 M NaOH. When needed, antifoam emulsion (Sigma-Aldrich, MO, USA) was added to prevent excessive formation of foam.

Cellular growth was followed by measuring the OD_{660nm} using the spectrophotometer UV-1601 (Shimadzu). Culture supernatants were collected after centrifugation at 14.000 rpm for 10 minutes at 4°C and metabolites determination was performed by HPLC using H_2SO_4 0.005 N as mobile phase and Rezex ROA-Organic Acid H+ (8%) column 300 × 7.8 mm (Phenomenex®). Analysis were performed at 40°C with a flux of 0.5 mL min-1.

Phenotype selection

The CEN.PK113-5D was transformed with either YCplac33*PAB1* or the plasmid mutant library and plated onto minimal YNB agar plates. After two days of growth at 30°C, all the CFU were retrieved and cultured in liquid minimal YNB medium at 30°C and 160rpm. After 4 and 24 hours, cell cultures were diluted to a final OD_{660nm} of 0.001 and 250 µL were plated onto minimal YNB agar plates in triplicate. Plates were incubated in at 39, 40 or 41°C. In parallel, plates were incubated at 30°C to verify that an equal amount of cells transformed

with the control plasmid and the mutant library were plated, and also to check whether the whole library was plated. After three days, the colony forming units (CFU) were counted.

Drop test

Cells were cultured in minimal YNB medium in 100 mL Erlenmeyer shake flasks until an OD_{660nm} of 0.5 - 0.6 and then normalized to an OD_{660nm} of 0.4. Aliquots (5 µL) from the normalized yeast cultures or 10-fold serial dilutions were spotted onto minimal YNB agar plates and incubated at 30 or 41°C. Acetic and formic acid supplemented plates were prepared from a 2X agar stock and a 10X YNB stock at pH3. Plates supplemented with hydroxymethylfurfural, H_2O_2 or ethanol were prepared from a 2X agar stock and a 10X YNB stock. Ethanol plates were incubated in a lidded box containing a beaker with 5% ethanol to minimize evaporation of this alcohol from the plates (van Voorst *et al.*, 2006).

Flow Cytometry Analysis

Experiments were carried out with the Beckman Coulter CYTOMICS-FC 500. For DHR123 (dihydrorhodamine 123) staining experiments, samples from the growth kinetics in bioreactor were prepared as follows: 0.2 OD_{660nm} aliquots were harvested and brought to 1mL with PBS. Samples were incubated at 40°C, in the dark, for 2 h in the presence of 5 µg/ml of DHR123 prior to the analysis.

Preparation of Soluble Protein Extract and Enzymatic Assay

After 20 hours from the inoculum in bioreactor, 50 OD_{660nm} of cells were harvested and washed twice with cold deionized water prior to resuspension in 50mM sodium phosphate buffer (pH7.0) for catalase assay, or Tris–HCl (pH7.5) for SOD assay, both containing 0.5mM PMSF and protease inhibitors (complete

Protease Inhibitor Cocktail, Roche) (Martani *et al.*, 2013). Cells were subjected to three cycles of mechanical disruption with the FastPrep-24 (MP Biomedical). Lysates were centrifuged for 20 min at 4°C and the obtained supernatants were used for catalase and SOD assays that were performed as described in (Martani *et al.*, 2013).

Briefly, catalase activity was measured discontinuously by correlating the H_2O_2 consumption to catalase activity within an appropriate time interval. To this purpose, 30 µL 2.2mM H_2O_2 was added to 300 µL cell-free extract and rapidly mixed. Every 2 minutes between 0 and 10 minutes, 50 µL aliquots of the mixture were rapidly mixed with 950 µL FOX reagent [ammonium ferrous sulphate (Sigma-Aldrich) 250 µM, xylenol orange (Sigma-Aldrich) 100 µM, sorbitol 100mM, H_2SO_4 25mM]. After 30 min of incubation at room temperature, the mixtures were centrifuged for 3 min at 12.000 rpm and the supernatants were read at 560 nm. Using a standard calibration curve, the H_2O_2 concentration in the sample was calculated and, from the obtained values, catalase activities were calculated based on another reference curve, previously obtained with serial dilutions of a catalase standard (catalase from bovine liver, Sigma-Aldrich, cat. no. C30).

SOD activities were assayed using the Sigma SOD assay kit (#19160) according to manufacturer's instructions. The assay was performed in 96 multiwell plates and measurements were obtained using Victor™ X3 (PerkinElmer®) multilabel plate reader.

Enzymatic activities were expressed as U/mg protein. The total protein concentration was determined as described in (Bradford, 1976), using bovine serum albumin as a standard.

References

Alper H and Stephanopoulos G (2007) Global transcription machinery engineering: a new approach for improving cellular phenotype. *Metab Eng* **9**: 258-267.

Alper H, Moxley J, Nevoigt E, Fink GR and Stephanopoulos G (2006) Engineering yeast transcription machinery for improved ethanol tolerance and production. *Science* **314**: 1565-1568.

Auesukaree C (2017) Molecular mechanisms of the yeast adaptive response and tolerance to stresses encountered during ethanol fermentation. *J Biosci Bioeng* **124**: 133-142.

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248-254.

Brambilla M, Martani F and Branduardi P (2017) The recruitment of the *Saccharomyces cerevisiae* poly(A)-binding protein into stress granules: new insights into the contribution of the different protein domains. *FEMS Yeast Res* **17**: fox059

Brune C, Munchel SE, Fischer N, Podtelejnikov AV and Weis K (2005) Yeast poly(A)-binding protein Pab1 shuttles between the nucleus and the cytoplasm and functions in mRNA export. *RNA* **11**: 517-531.

Costello J, Castelli LM, Rowe W, *et al.* (2015) Global mRNA selection mechanisms for translation initiation. *Genome Biol* **16**: 10.

Dabrowski M, Bukowy-Bieryllo Z and Zietkiewicz E (2015) Translational readthrough potential of natural termination codons in eucaryotes--The impact of RNA sequence. *RNA Biol* **12**: 950-958.

De Deken RH (1966) The Crabtree effect: a regulatory system in yeast. *J Gen Microbiol* **44**: 149-156.

Doğan A, Demirci S, Aytekin A and Şahin F (2014) Improvements of tolerance to stress conditions by genetic engineering in *Saccharomyces cerevisiae* during ethanol production. *Appl Biochem Biotechnol* **174**: 28-42.

Entian K-D and Kötter P (1998) 23 Yeast Mutant and Plasmid Collections. *Methods in Microbiology,* Vol. 26 (Brown AJP and Tuite M, eds.), p.^pp. 431-449. Academic Press.

Gable K, Slife H, Bacikova D, Monaghan E and Dunn TM (2000) Tsc3p is an 80-amino acid protein associated with serine palmitoyltransferase and required for optimal enzyme activity. *J Biol Chem* **275**: 7597-7603.

Galanie S, Thodey K, Trenchard IJ, Filsinger Interrante M and Smolke CD (2015) Complete biosynthesis of opioids in yeast. *Science* **349**: 1095-1100.

Gao L, Liu Y, Sun H, Li C, Zhao Z and Liu G (2016) Advances in mechanisms and modifications for rendering yeast thermotolerance. *J Biosci Bioeng* **121**: 599-606.

Gibson BR, Lawrence SJ, Leclaire JP, Powell CD and Smart KA (2007) Yeast responses to stresses associated with industrial brewery handling. *FEMS Microbiol Rev* **31**: 535-569.

Gietz RD and Woods RA (2002) Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol* **350**: 87-96.

Grousl T, Ivanov P, Frýdlová I, *et al.* (2009) Robust heat shock induces eIF2alpha-phosphorylation-independent assembly of stress granules containing eIF3 and 40S ribosomal subunits in budding yeast, *Saccharomyces cerevisiae*. *J Cell Sci* **122**: 2078-2088.

Hong KK and Nielsen J (2012) Metabolic engineering of *Saccharomyces cerevisiae*: a key cell factory platform for future biorefineries. *Cell Mol Life Sci* **69**: 2671-2690.

Kastenmayer JP, Ni L, Chu A, *et al.* (2006) Functional genomics of genes with small open reading frames (sORFs) in *S. cerevisiae*. *Genome Res* **16**: 365-373.

Kavšček M, Stražar M, Curk T, Natter K and Petrovič U (2015) Yeast as a cell factory: current state and perspectives. *Microb Cell Fact* **14**: 94.

Khatun MM, Yu X, Kondo A, Bai F and Zhao X (2017) Improved ethanol production at high temperature by consolidated bioprocessing using *Saccharomyces cerevisiae* strain engineered with artificial zinc finger protein. Bioresour Technol **245**: 1447-1454.

Kim IS, Kim YS, Kim H, Jin I and Yoon HS (2013) *Saccharomyces cerevisiae* KNU5377 stress response during high-temperature ethanol fermentation. *Mol Cells* **35**: 210-218.

Lee D, Ohn T, Chiang YC, Quigley G, Yao G, Liu Y and Denis CL (2010) PUF3 acceleration of deadenylation in vivo can operate independently of CCR4 activity, possibly involving effects on the PAB1-mRNP structure. *J Mol Biol* **399**: 562-575.

Ling H, Teo W, Chen B, Leong SS and Chang MW (2014) Microbial tolerance engineering toward biochemical production: from lignocellulose to products. *Curr Opin Biotechnol* **29**: 99-106.

Martani F, Marano F, Bertacchi S, Porro D and Branduardi P (2015) The *Saccharomyces cerevisiae* poly(A) binding protein Pab1 as a target for eliciting stress tolerant phenotypes. *Sci Rep* **5**: 18318.

Martani F, Fossati T, Posteri R, Signori L, Porro D and Branduardi P (2013) Different response to acetic acid stress in *Saccharomyces cerevisiae* wild-type and L-ascorbic acid-producing strains. *Yeast* **30**: 365-378.

Mattanovich D, Sauer M and Gasser B (2014) Yeast biotechnology: teaching the old dog new tricks. *Microb Cell Fact* **13**: 34.

Morano KA, Grant CM and Moye-Rowley WS (2012) The response to heat shock and oxidative stress in *Saccharomyces cerevisiae*. *Genetics* **190**: 1157-1195.

Parker R (2012) RNA degradation in *Saccharomyces cerevisiae*. *Genetics* **191**: 671-702.

Qiu Z and Jiang R (2017) Improving *Saccharomyces cerevisiae* ethanol production and tolerance via RNA polymerase II subunit Rpb7. *Biotechnol Biofuels* **10**: 125.

Ranish JA, Hahn S, Lu Y, Yi EC, Li XJ, Eng J and Aebersold R (2004) Identification of *TFB5*, a new component of general transcription and DNA repair factor IIH. *Nat Genet* **36**: 707-713.

Riback JA, Katanski CD, Kear-Scott JL, Pilipenko EV, Rojek AE, Sosnick TR and Drummond DA (2017) Stress-Triggered Phase Separation Is an Adaptive, Evolutionarily Tuned Response. *Cell* **168**: 1028-1040.e1019.

Richardson R, Denis CL, Zhang C, Nielsen ME, Chiang YC, Kierkegaard M, Wang X, Lee DJ, Andersen JS and Yao G (2012) Mass spectrometric identification of proteins that interact through specific domains of the poly(A) binding protein. *Mol Genet Genomics* **287**: 711-730.

Roque S, Cerciat M, Gaugué I, Mora L, Floch AG, de Zamaroczy M, Heurgué-Hamard V and Kervestin S (2015) Interaction between the poly(A)-binding protein Pab1 and the eukaryotic release factor eRF3 regulates translation termination but not mRNA decay in *Saccharomyces cerevisiae*. *RNA* **21**: 124-134.

Swisher KD and Parker R (2010) Localization to, and effects of Pbp1, Pbp4, Lsm12, Dhh1, and Pab1 on stress granules in *Saccharomyces cerevisiae*. *PLoS One* **5**: e10006.

van Voorst F, Houghton-Larsen J, Jønson L, Kielland-Brandt MC and Brandt A (2006) Genome-wide identification of genes required for growth of *Saccharomyces cerevisiae* under ethanol stress. *Yeast* **23**: 351-359.

Verghese J, Abrams J, Wang Y and Morano KA (2012) Biology of the heat shock response and protein chaperones: budding yeast (*Saccharomyces cerevisiae*) as a model system. *Microbiol Mol Biol Rev* **76**: 115-158.

Wallace-Salinas V and Gorwa-Grauslund MF (2013) Adaptive evolution of an industrial strain of *Saccharomyces cerevisiae* for combined tolerance to inhibitors and temperature. *Biotechnol Biofuels* **6**: 151.

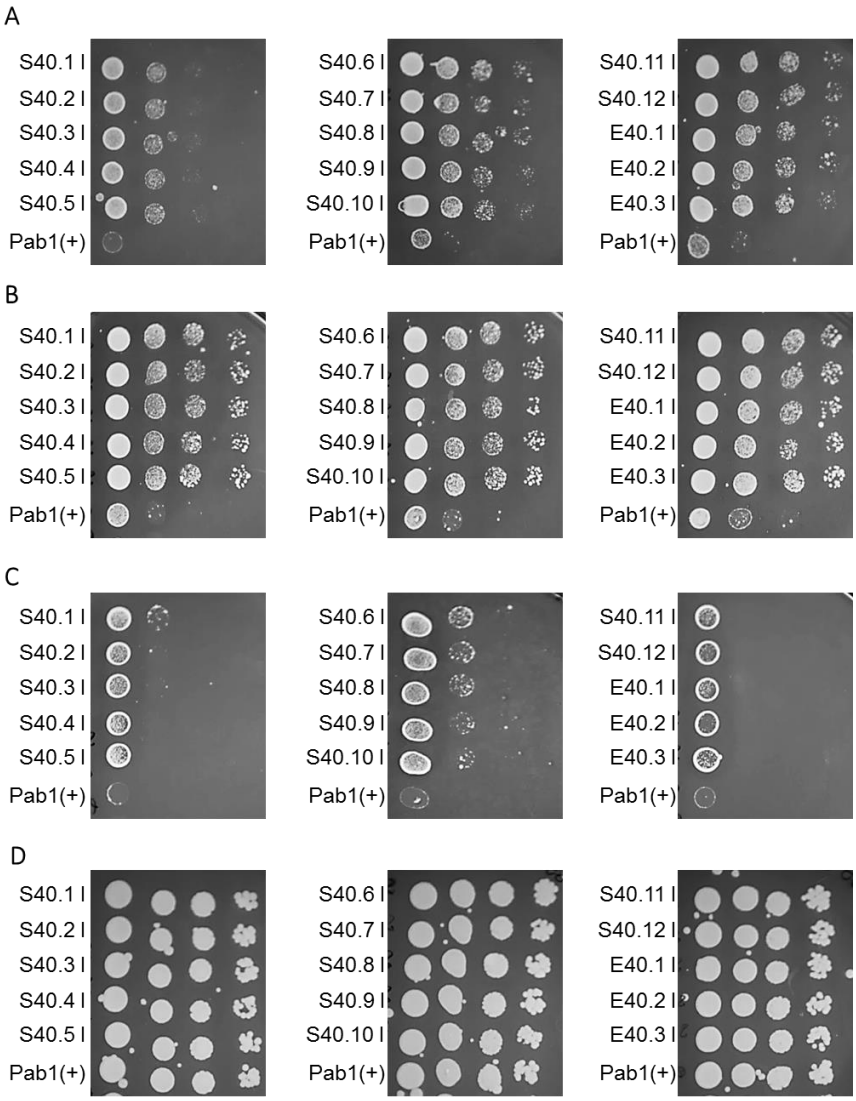
Werner-Washburne M, Braun E, Johnston GC and Singer RA (1993) Stationary phase in the yeast *Saccharomyces cerevisiae*. *Microbiol Rev* **57**: 383-401.

Wheeler JR, Matheny T, Jain S, Abrisch R and Parker R (2016) Distinct stages in stress granule assembly and disassembly. *Elife* **5**: 18413

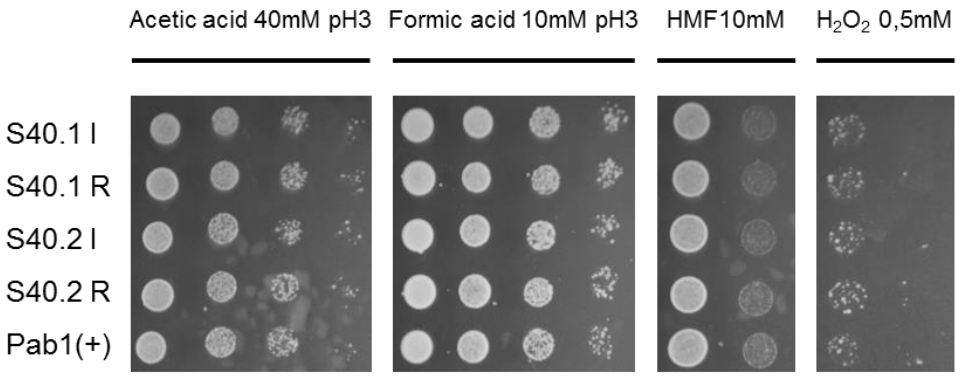
Yao G, Chiang YC, Zhang C, Lee DJ, Laue TM and Denis CL (2007) PAB1 self-association precludes its binding to poly(A), thereby accelerating CCR4 deadenylation in vivo. *Mol Cell Biol* **27**: 6243-6253.

Supplementary data

Supplementary Fig.S1 Evaluation of multiple tolerances in selected Pab1 mutant strains by drop test. Mutants isolated from the screening (indicated with the letter "I") and the control Pab1(+) strain were cultivated in minimal medium and then normalized to an OD_{660nm} of 0.4. Aliquots (5 µL) from the normalized yeast cultures or 10-fold serial dilutions were spotted onto minimal plates and cultured at 40°C for 4 days (A), 41°C for 6 days (B), 42°C for six days (C) and 30°C for four days (D). Images are representative of three independent experiments.



Supplementary Fig.S2 Evaluation of multiple tolerances in selected Pab1 mutant strains by drop test. Mutants isolated from the screening (indicated with the letter "I"), reconstructed strains (indicated with the letter "R") and the control Pab1(+) strain were cultivated in minimal medium and then normalized to an OD_{660nm} of 0.4. Aliquots (5 µL) from the normalized yeast cultures or 10-fold serial dilutions were spotted onto minimal plates supplemented with the indicated toxic agents. The strains were grown for 3 days at 30°C. Images are representative for all the strains evaluated in this study.



Supplementary Fig.S3 Multiple sequence alignment of *PAB1* promoter, 5'UTR and coding sequence of variants isolated from the screening. Yellow-highlighted letters indicate nucleotide substitution. Green-highlighted letters indicate start codons. Red letters indicate premature stop codons leading to truncation in the encoded proteins.

```
PAB1-S40.5        GCGTCTAAGTGTGTGTACTATAGGGCACCGTAAAGTAATAATGCTTAATTAGTTACTACT 60
PAB1-S40.10       GCGTGTAAGTGTGTGTACTATAGGGCACCGTAAAGTAATAATGCTTAATTAGTTACTACT 60
PAB1-S40.9        GCGTGTAAGTGTGTGTACTAAAGGGCACCGTAAAGTAATAATGCTTAATTAGTTACTACT 60
PAB1-S40.4        GCGTGTAAGTGTGTGTACTATAGGGCACCGTAAAGTAATAATGCTTAATTAGTTACTACT 60
PAB1-S40.3        GCGTGTAAGTGTGTGTACTATAGGGCACCGTAAAGTAATAATGCTTAATTAGTTACTACT 60
PAB1-S40.2        GCGTGTAAGTGTGTGTACTATAGGGCACCGTAAAGTAATAATGCTTAATTAGTTACTACT 60
PAB1-S40.1        GCGTGTAAGTGTGTGTACTATAGGGCACCGTAAAGTAATAATGCTTAATTAGTTACTACT 60
PAB1-wt           GCGTGTAAGTGTGTGTACTATAGGGCACCGTAAAGTAATAATGCTTAATTAGTTACTACT 60
PAB1-S40.7        GCGTGTAAGTGTGTGTACTATAGGGCACCGTAAAGTAATAATGCTTAATTAGTTACTACT 60
PAB1-S40.8        GCGTGTAAGTGTGTGTACTATAGGGCACCGTAAAGTAATAATGCTTAATTAGTTACTACT 60
PAB1-S40.11       GCGTGTAAGTGTGTGTACTATAGGGCACCGTAAAGTAATAATGCTTAATTAGTTACTACT 60
PAB1-E40.2        GCGTGTAAGTGTGTGTACTATAGGGCACCGTAAAGTAATAATGCTTAATTAGTTACTACT 60
PAB1-S40.6        GCGTGTAAGTGTGTGTACTATAGGGCACCGTAAAGTAATAATGCTTAATTAGTTACTACT 60
PAB1-S40.12       GCGTGTAAGTGTGTGTACTATAGGGCACCGTAAAGTAATAATGCTTAATTAGTTACTACT 60
PAB1-E40.3        GCGTGTAAGTGTGTGTACTATAGGGCACCGTAAAGTAATAATGCTTAATTAGTTACTACT 60
PAB1-E40.1        GCGTGTAAGTGTGTGTACTATAGGGCACCGTAAAGTAATAATGCTTAATTAGTTACTACT 60
                  **** *************** ****************************************

PAB1-S40.5        ATGACCATATAAGAGGTCATACTGTATGAAGCCACAAAGCAGATAGATCAATCATGTTTA 120
PAB1-S40.10       ATGTCCATATAAGAGGTCATACTGTATGAAGCCACTAAGCAGATAGATCAATCATGTTTA 120
PAB1-S40.9        ATGACCATATAAGAGGTCATACTGTATGAAGCCACAAAGCAGATGGATCAATCATGTTTA 120
PAB1-S40.4        ATGACCATATAAGAGGTCATACAGTATGAAGCCACAAAGCAGATAGATCAATCATGTTTA 120
PAB1-S40.3        ATGACCATATAAGAGGTCATACTGTATGAAGCCACAAAGCAGATAGATCAATCATGTTTA 120
PAB1-S40.2        ATGACCATATAAGAGGTCATACTGTATGAAGCCACAAAGCAGATAGATCAATCATGTTTA 120
PAB1-S40.1        ATGACCATATAAGAGGTCATACTGTATGAAGCCACAAAGCAGATAGATCAATCATGTTTA 120
PAB1-wt           ATGACCATATAAGAGGTCATACTGTATGAAGCCACAAAGCAGATAGATCAATCATGTTTA 120
PAB1-S40.7        ATGACCATATAAGAGGTCATACTGTATGAAGCCACAAAGCAGATAGATCAATCATGTTTA 120
PAB1-S40.8        ATGACCATATAAGTGGTCATACTGTATGAAGCCACAAAGCAGATAGATCAATCATGTTTA 120
PAB1-S40.11       ATGACCATATAAGAGGTCATACTGTATGAAGCCACAAAGCAGATAGATCAATCATGTTTA 120
PAB1-E40.2        ATGACCATATAAGAGGTCATACTGTATGAAGCCACAAAGCAGATAGATCAATCATGTTTA 120
PAB1-S40.6        ATGACCATATAAGAGGTCATACTGTATGAAGCCACAAAGCAGATAGATCAATCATGTTTA 120
PAB1-S40.12       ATGACCATATAAGAGGTCATACTGTATGAAGCCACAAAGCAGATAGATCAATCATGTTTA 120
PAB1-E40.3        ATGACCATATAAGAGGTCATACTGTATGAAGCCACAAAGCAGATAGATCAATCATGTTTA 120
PAB1-E40.1        ATGACCATATAAGAGGTCATACTGTATGAAGCCACAAAGCAGATAGATCAATCATATTTA 120
                  *** ********* ******** ************ ******** ********** ****

PAB1-S40.5        ACGAAAACTGTTAATCGAAGATTATTTCTTTTTTTTTTTTCTCTTTCCTTTTTACAAAGAA 180
PAB1-S40.10       ACGAAAACTGTTAATCGAAGATTATTTCTTTTTTTTTTTTCTCTTTCCTTTTTACAAAGAA 180
PAB1-S40.9        ACGAAAACTGTTAATCGAAGATTATTTCTTTTTTTTTTTTCTCTTTCCTTTTTACAAAGAA 180
PAB1-S40.4        ACGAAAACTGTTAATCGAAGATTATTTCTTTTTTTTTTTTCTCTTTCCTTTTTACAAAGAA 180
PAB1-S40.3        ACGAAAACTGTTAATCGAAGATTATTTCTTTTTTTTTTTTCTCTTTCCTTTTTACAAAGAA 180
PAB1-S40.2        ACGAAAACTGTTAATCGAAGATTATTTCTTTTTTTTTTTTCTCTTTCCTTTTTACAAAGAA 180
PAB1-S40.1        ACGAAAACTGTTAATCGAAGATTATTTCTTTTTTTTTTTTCTCTTTCCTTTTTACAAAGAA 180
PAB1-wt           ACGAAAACTGTTAATCGAAGATTATTTCTTTTTTTTTTTTCTCTTTCCTTTTTACAAAGAA 180
PAB1-S40.7        ACGAAAACTGTTAATCGAAGATTATTTCTTTTTTTTTTTTCTCTTTCCTTTTTACAAAGAA 180
PAB1-S40.8        ACGAAAACTGTTAATCGAAGATTATTTCTTTTTTTTTTTTCTCTTTCCTTTTTACAAAGAA 180
PAB1-S40.11       ACGAAAACTGTTAATCGAAGATTATTTCTTTTTTTTTTTTCTCTTTCCTTTTTACAAAGAA 180
PAB1-E40.2        ACGAAAACTGTTAATCGAAGATTATTTCTTTTTTTTTTTTCTCTTTCCTTTTTACAAAGAA 180
PAB1-S40.6        ACGAAAACTGTTAATCGAAGATTATTTCTTTTTTTTTTTTCTCTTTCCTTTTTACAAAGAA 180
PAB1-S40.12       ACGAAAACTGTTAATCGAAGATTATTTCTTTTTTTTTTTTCTCTTTCCTTTTTACAAAGAA 180
PAB1-E40.3        ACGAAAACTGTTAATCGAAGATTATTTCTTTTTTTTTTTTCTCTTTCCTTTTTACAAAGAA 180
PAB1-E40.1        ACGAAAACAGTTAATCGAAGATTATTTCTTTTTTTTTTTTCTCTTTCCTTTTTACAAAGAA 180
                  ******** ************************************************

PAB1-S40.5        AATTTTTTTTGCGCTTTTTGCCATCACCATCGCAAGTTCTGGGACAATTGTTCTCTTTCG 240
PAB1-S40.10       AATTTTTTTTGCGCTTTTTGCCATCTCCATCGCAAGTTCTGGGACAATTGTTCTCTTTCG 240
PAB1-S40.9        AATTTTTTTTGCGCTTTTTGCCATCACCATCGCAAGTTCTGGGACAATTGTTCTCTTTCG 240
PAB1-S40.4        AATTTTTTTTGCGCTTTTTGCCATCACCATCGCAAGTTCTGGGACAATTGTTCTCTTTCG 240
PAB1-S40.3        AATTTTTTTTGCGCTTTTTGCCATCACCATCGCAAGTTCTGGGACAATTGTTCTCTTTCG 240
PAB1-S40.2        AATTTTTTTTGCGCTTTTTGCCATCACCATCGCAAGTTCTGGGACAATTGTTCTCTTTCG 240
PAB1-S40.1        AATTTTTTTTGCGCTTTTTGCCATCACCATCGCAAGTTCTGGGACAATTGTTCTCTTTCG 240
PAB1-wt           AATTTTTTTTGCGCTTTTTGCCATCACCATCGCAAGTTCTGGGACAATTGTTCTCTTTCG 240
```

```
PAB1-S40.7      AATTTTTTTTGCGCTTTTTGCCATCACCATCGCAAGTTCTGGGACAATTGTTCTCTTTCG 240
PAB1-S40.8      AATTTTTTTTGCGCTTTTTGCCATCACCATCGCAAGTTCTGGGACAATTGTTCTCTTTCG 240
PAB1-S40.11     AATTTTTTTTGCGCTTTTTGCCATCACCATCGCAAGTTCTGGGACAATTGTTCTCTTTCG 240
PAB1-E40.2      AATTTTTTTTGCGCTTTTTGCCATCACCATCGCAAGTTCTGGGACAATTGTTCTCTTTCG 240
PAB1-S40.6      AATTTTTTTTGCGCTTTTTGCCATCACCATCGCAAGTTCTGGGACAATTGTTCTCTTTCG 240
PAB1-S40.12     AATTTTTTTTGCGCTTTTTGCCATCACCATCGCAAGTTCTGGGACAATTGTTCTCTTTCG 240
PAB1-E40.3      AATTTTTTTTGCGCTTTTTGCCATCACCATCGCAAGTTCTGGGACAATTGTTCTCTTTCG 240
PAB1-E40.1      AATTTTTTTTGCGCTTTTTGCCATCACCATCGCAAGTTCTGGGACAATTGTTCTCTTTCG 240
                ************************* ********************************

PAB1-S40.5      CTCCAGTTCCAAGGAAAGAGGTTTCTGTTTTACTTAATAGAAAGTGTCATCTTGTATTTT 300
PAB1-S40.10     CTCCAGTTCCAAGGAAAGAGGTTTCTGTTTTACTTAATAGAAAGTGTCATCTTGTATTTT 300
PAB1-S40.9      CTCCAGTTCCAAGGAAAGAGGTTTCTGTTTTACTTAATAGAAAGTGTCATCTTGTATTTT 300
PAB1-S40.4      CTCCAGTTCCAAGGAAAGAGGTTTATGTTTTACTTAATAGAAAGTGTCATCTTGTATTTT 300
PAB1-S40.3      CTCCAGTTCCAAGGAAAGAGGTTTCTGTTTTACTTAATAGAAAGTGTCATCTTGTATTTT 300
PAB1-S40.2      CTCCAGTTCCAAGGAAAGAGGTTTCTGTTTTACTTAATAGAAAGTGTCATCTTGTATTTT 300
PAB1-S40.1      CTCCAGTTCCAAGGAAAGAGGTTTCTGTTTTACTTAATAGAAAGTGTCATCTTGTATTTT 300
PAB1-wt         CTCCAGTTCCAAGGAAAGAGGTTTCTGTTTTACTTAATAGAAAGTGTCATCTTGTATTTT 300
PAB1-S40.7      CTCCAGTTCCAAGGAAAGAGGTTTCTGTTTTACTTAATAGAAAGTGTCATCTTGTATTTT 300
PAB1-S40.8      CTCCAGTTCCAGGGAAAGAGGTTTCTGTTTTACTTAATAGAAAGTGTCATCTTGTATTTT 300
PAB1-S40.11     CTCCAGTTCCAAGGAAAGAGGTTTCTGTTTTACTTAATAGAAAGTGTCATCTTGTATTTT 300
PAB1-E40.2      CTCCAGTTCCAAGGAAAGAGGTTTCTGTTTTACTTAATAGAAAGTGTCATCTTGTATTTT 300
PAB1-S40.6      CTCCAGTTCCAAGGAAAGAGGTTTCTGTTTTACTTAATAGAAAGTGTCATCTTGTATTTT 300
PAB1-S40.12     CTCCAGTTCCAAGGAAAGAGGTTTCTGTTTTACTTAATAGAAAGTGTCATCTTGTATTTT 300
PAB1-E40.3      CTCCAGTTCCAAGGAAAGAGGTTTCTGTTTTACTTAATAGAAAGTGTCATCTTGTATTTT 300
PAB1-E40.1      CTCCAGTTCCAAGGAAAGAGGTTTCTGTTTTACTTAATAGAAAGTGTCATCTTGTATTTT 300
                *********** ************ ********************************

PAB1-S40.5      ATATCTCTTCTTTCTAGTGTAAAATTCTTTAGTTTTGATTTTGTATTTTTAGGACAGTGA 360
PAB1-S40.10     ATATCTCTTCTTTCTTGTGTAAAATTCTTTAGTTTTGATTTTGTATTTTTAGGACAGTGA 360
PAB1-S40.9      ATATCTCTTCTTTCTTGTGTAAAATTCTTTAGTTTTGATTTTGTATTTTTAGGACAGTGA 360
PAB1-S40.4      ATATCTCTTCTTTCTTGTGTAAAATTCTTTAGTTTTGATTTTGTATTTTTAGGACAGTGA 360
PAB1-S40.3      ATATCTCTTCTTTCTTGTGTAAAATTCTTTAGTTTTGATTTTGTATTTTTAGGACAGTGA 360
PAB1-S40.2      ATATCTCTTCTTTCTTGTGTAAGATTCTTTAGTTTTGATTTTGTATTTTTAGGACAGTGA 360
PAB1-S40.1      ATATCTCTTCTTTCTTGTGTAAAATTCTTTAGTTTTGATTTTGTATTTTTAGGACAGTGA 360
PAB1-wt         ATATCTCTTCTTTCTTGTGTAAAATTCTTTAGTTTTGATTTTGTATTTTTAGGACAGTGA 360
PAB1-S40.7      ATATCTCTTCTTTCTTGTGTAAAATTCTTTAGTTTTGATTTTGTATTTTTAGGACAGTGA 360
PAB1-S40.8      ATATCTCTTCTTTCTTGTGTAAAATTCTATAGTTTTGATTTTGTATTTTTAGGACAGTGA 360
PAB1-S40.11     ATATCTCTTCTTTCTTGTGTAAAATTCTTTAGTTTTGATTTTGTATTTTTAGGACAGTGA 360
PAB1-E40.2      ATATCTCTTCTTTCTTGTGTAAAATTCTTTAGTTTTGATTTTGTATTTTTAGGACAGTGA 360
PAB1-S40.6      ATATCTCTTCTTTCTTGTGTAAAATTCTTTAGTTTTGATTTTGTATTTTTAGGACAGTGA 360
PAB1-S40.12     ATATCTCTTCTTTCTTGTGTAAAATTCTTTAGTTTTGATTTTGTATTTTTAGGACAGTGA 360
PAB1-E40.3      ATATCTCTTCTTTCTTGTGTAAAATTCTTTAGTTTTGATTTTGTATTTTTAGGACAGTGA 360
PAB1-E40.1      ATATCTCTTCTTTCTTGTGTAAAATTCTGTAGTTTTGATTTTGTATTTTTAGGACAGTGA 360
                *************** ****** ***** ********************************

PAB1-S40.5      GCTACGAAGTAACATTTTTACTTAATAACCGTTTGAAGCATAGAGCAGGCCCTGGTACCA 420
PAB1-S40.10     GCTACGAAGTAACATTTTTACTTAATAACCGTTTGAAGCATAGAGCAGGCCCTGGTACCA 420
PAB1-S40.9      GCTACGAAGTAACATTTTTACTTAATAACCGTTTGAAGCATAGAGCAGGCCCTGGTACCA 420
PAB1-S40.4      GCTACGAAGTAACATTTTTACTTAATAACCGTTTGAAGCATAGAGCAGGCCCTGGTACCA 420
PAB1-S40.3      GCTACGAAGTAACATTTTTACTTAATAACCGTTTGAAGCATAGAGCAGGCCCTGGTACCA 420
PAB1-S40.2      GCTACGAAGTAACATTTTTACTTAATAACCGTTTGAAGCATAGAGCAGGCCCTGGTACCA 420
PAB1-S40.1      GCTACGAAGTAACATTTTTACTTAATAACCGTTTGAAGCATAGAGCAGGCCCTGGTACCA 420
PAB1-wt         GCTACGAAGTAACATTTTTACTTAATAACCGTTTGAAGCATAGAGCAGGCCCTGGTACCA 420
PAB1-S40.7      GCTACGAAGTAACATTTTTACTTAATAACCGTTTGAAGCATAGAGCAGGCCCTGGTAGCA 420
PAB1-S40.8      GCTACGAAGTAACACTTTTACTTAATAACCGTTTGAAGCATAGAGCAGGCCCTGGTACCA 420
PAB1-S40.11     GCTACGAAGTAACATTTTTACTTAATAACCGTTTGAAGCATAGAGCAGGTCCTGGTACCA 420
PAB1-E40.2      GCTACGAAGTAACATTTTTACTTAATAACCGTTTGAAGCATAGAGCAGGCCCTGGTACCA 420
PAB1-S40.6      GCTACGAAGTAACATTTTTACTTAATAACCGTTTGAAGCATAGAGCAGGCCCTGGTACCA 420
PAB1-S40.12     GCTACGAAGTAACATTTTTACTTAATAACCGTTTGAAGCATAGAGCAGGCCCTGGTACCA 420
PAB1-E40.3      GCTACGAAGTAACATTTTTACTTAATAACCGTTTGAAGCATAGAGCAGGCCCTGGTACCA 420
PAB1-E40.1      GCTACGAAGTAACATTTTTACTTAATAACCGTTTGAAGCATAGAGCAGGCCCTGGTACCA 420
                ************** ********************************** ******* **

PAB1-S40.5      CCACCTAATATCTGGCTTTTTATTCAATAAAAACTCAAAAAAAAAAAATCCAAAAAAAACT 480
PAB1-S40.10     CCACCTAATATCTGGCTTTTTATTCAATACAAACTCAAAAAAAAAAAATCCAAAAAAAACT 480
PAB1-S40.9      CCACCTAATATCTGGCTTTTTATTCAATAAAAACTCAAAAAAAAAAAATCCAAAAAAAACT 480
PAB1-S40.4      CCACCTAATATCTGGCTTTTTATTCAATAAAAACTCAAAAAAAAAAAATCCAAAAAAAACT 480
PAB1-S40.3      CCACCTAATATCTGGCTTTTTATTCAATAAAAACTCAAAAAAAAAAAATCCAAAAAAAACT 480
PAB1-S40.2      CCACCTAATATCTGGCTTTTTATTCAATAAAAACTCAAAAAAAAAAAATCCAAAAAAAACT 480
PAB1-S40.1      CCACCTAATATCTGGCTTTTTATTCAATAAAAACTCAAAAAAAAAAAATCCAAAAAAAACT 480
PAB1-wt         CCACCTAATATCTGGCTTTTTATTCAATAAAAACTCAAAAAAAAAAAATCCAAAAAAAACT 480
```

```
PAB1-S40.7      CCACCTAATATCTGGCTTTTTATTCAATAAAAACTCAAAAAAAAAAATCCAAAAAAAACT 480
PAB1-S40.8      CCACCTAATATCTGGCTTTTTATTCAATAAAAACTCAAAAAAAAAAATCCAAAAAAAACT 480
PAB1-S40.11     CCACCTAATATCTGGCTTTTTATTCAATAAAAACTCAAAAAAAAAAATCCAAAAAAAACT 480
PAB1-E40.2      CCACCTAATATCTGGCTTTTTATTCAATAAAAACTCAAAAAAAAAAATCCAAAAAAAACT 480
PAB1-S40.6      CCACCTAATATCTGGCTTTTTATTCAATAAAAACTCAAAAAAAAAAATCCAAAAAAAACT 480
PAB1-S40.12     CCACCTAATATCTGGCTTTTTATTCAATAAAAACTCAAAAAAAAAAATCCAAAAAAAACT 480
PAB1-E40.3      CCACCTAATATCTGGCTTTTTATTCAATAAAAACTCAAAAAAAAAAATCCAAAAAAA-CT 480
PAB1-E40.1      CCACCTAATATCTGGCTTTTTATTCAATAAAAACTCAAAAAAAAAAATCCAAAAAAAACT 480
                **************************** ************************** **

PAB1-S40.5      AAAAAACCAATAAAAATAAAATGGCTGATATTACTGATAAGACAGCTGAACAATTGGAAA 540
PAB1-S40.10     AAAAAACCAATAAAAATAAAATGGCTGATATTACTGATAAGACAGCTGAACAATTGGAAA 540
PAB1-S40.9      AAAAAACCAATAAAAATAAAATGGCTGATATTACTGATAAGACAGCTGAACAATTGGAAA 540
PAB1-S40.4      AAAAAACCAATAAAAATAAAATGGCTGATATTACTGATAAGACAGCTGAACAATTGGAAA 540
PAB1-S40.3      AAAAAACCAATAAAAATAAAATGGCTGATATTACTGATAAGACAGCTGAACAATTGGAAA 540
PAB1-S40.2      AAAAAACCAATAAAAATAAAATGGCTGATATTACTGATAAGACAGCTGAACAATTGGAAA 540
PAB1-S40.1      AAAAAACCAATAAAAATAAAATGGCTGATATTACTGATAAGACAGCTGAACAATTGGAAA 540
PAB1-wt         AAAAAACCAATAAAAATAAAATGGCTGATATTACTGATAAGACAGCTGAACAATTGGAAA 540
PAB1-S40.7      AAAAAACCAATAAAAATAAAATGGCTGATATTACTGATAAGACAGCTGAACAATTGGAAA 540
PAB1-S40.8      AAAAAACCAATAAAAATAAAATGGCTGATATTACTGATAAGACAGCTGAACAATTGGAAA 540
PAB1-S40.11     AAAAAACCAATAAAAATAAAATGGCTGATATTACTGATAAGACAGCTGAACTATTGGAAA 540
PAB1-E40.2      AAAAAACCAATAAAAATAAAATGGCTGATATTACCGATAAGACAGCTGAACAATTGGAAA 540
PAB1-S40.6      AAAAAACCAATAAAAATAAAATGGCTGATATTACTGATAAGACAGCTGAACAATTGGAAA 540
PAB1-S40.12     AAAAAACCAATAAAAATAAAATGGCTGATATTACTGATAAGACAGCTGAACAATTGGAAA 540
PAB1-E40.3      AAAAAACCAATAAAAATAAAATGGCTGATATTACTGATAAGACAGCTGAACAATTGGAAA 540
PAB1-E40.1      AAAAAACCAATAAAAATAAAATGGCTGATATTACTGATAAGACAGCTGAACAATTGGAAA 540
                ********************************** ***************** ********

PAB1-S40.5      ACTTGAATATTCAAGATGCCCAAAAGCAAGCCGCCACTGGTTCAGAAAGCCAATCTGTTG 600
PAB1-S40.10     ACTTGAATATTCAAGATGACCAAAAGCAAGCCGCCACTGGTTCAGAAAGCCAATCTGTTG 600
PAB1-S40.9      ACTTGAATATTCAAGATGACCAAAAGCAAGCCGCCACTGGTTCAGAAAGCCAATCTGTTG 600
PAB1-S40.4      ACTTGAATATTCAAGATGACCAAAAGCAAGCCGCCACTGGTTCAGAAAGCCAATCTGTTG 600
PAB1-S40.3      ACTTGAATATTCAAGATGACCAAAAGCAAGCCGCCACTGGTTCAGAAAGCCAATCTGTTG 600
PAB1-S40.2      ACTTGAATATTCAAGATGACCAAAAGCAAGCCGCCACTGGTTCAGAAAGCCAATCTGTTG 600
PAB1-S40.1      ACTTGAATATTCAAGATGACCAAAAGCAAGCTGCCACTGGTTCAGAAAGCCAATCTGTTG 600
PAB1-wt         ACTTGAATATTCAAGATGACCAAAAGCAAGCCGCCACTGGTTCAGAAAGCCAATCTGTTG 600
PAB1-S40.7      ACTTGAATATTCAAGATGACTAAAAGCAAGCCGCCACTGGTTCAGAAAGCCAATCTGTTG 600
PAB1-S40.8      ACTTGAATATTCAAGATGACCAAAAGCAAGCCGCCACTGGTTCAGAAAGCCAATCTGTTG 600
PAB1-S40.11     ACTTGAATATTCAAGATGACCAAAAGCAAGCCGCCACTGGTTCAGAAAGCCAATCTGTTG 600
PAB1-E40.2      ACTTGAATATTCAAGATGACCAAAAGCAAGCCGCCACTGGTTCAGCAAGCCAATCTGTTG 600
PAB1-S40.6      ACTTGAATATTCAAGATGACCAAAAGCAAGCCGCCACTGGTTCAGAAAGCCAATCTGTTG 600
PAB1-S40.12     ACTTGAATATTCAAGATGACCAAAAGCAAGCCGCCACTGGTTCAGAAAGCCAATCTGTTG 600
PAB1-E40.3      ACTTGAATATTCAAGATGACCAAAAGCAAGCCGCCACTGGTTCAGAAAGCCAATCTGTTG 600
PAB1-E40.1      ACTTGAATATTCAAGATGACCAAAAGCAAGCCGCCACTGGTTCAGAAAGCCAATCTGTTG 600
                ****************** * ********** ************* ***************

PAB1-S40.5      AAAACTCTTCTGCATCATTATATGTTGGTGACTTAGAACCTTCTGTTTCCGAAGCCCACT 660
PAB1-S40.10     AAAACTCTTCTGCATCATTATATGTTGGTGACTTAGAACCTTCTGTTTCCGAAGCCCACT 660
PAB1-S40.9      AAAACTCTACTGCATCATTATATGTTGGTGACTTAGAACCTTCTGTTTCCGAAGCCCACT 660
PAB1-S40.4      AAAACTCTTCTGCATCATTATATGTTGGTGACTTAGAACCTTCTGTTTCCGAAGCCCACT 660
PAB1-S40.3      AAAACTCTTCTGCATCATTATATGTTGGTGACTTAGAACCTTCTGTTTCCGAAGCCCACT 660
PAB1-S40.2      AAAACTCTTCTGCATCATTATATGTTGGTGACTTAGAACCTTCTGTTTCCGAAGCCCACT 660
PAB1-S40.1      AAAACTCTTCTGCATCATTATATGTTGGTGACTTAGAACCTTCTGTTTCCGAAGCCCACT 660
PAB1-wt         AAAACTCTTCTGCATCATTATATGTTGGTGACTTAGAACCTTCTGTTTCCGAAGCCCACT 660
PAB1-S40.7      AAAACTCTTCTGCATCATTATATGTTGGTGACTTAGAACCTTCTGTTTCCGAAGCCCACT 660
PAB1-S40.8      AAAACTCTTCTGCATCATTATATGTTGGTGACTTAGAACCTTCTGTTTCCGGAGCCCACT 660
PAB1-S40.11     AAAACTCTTCTGCATCATTATATGTTGGTGACTTAGAACCTTCTGTTTCCGAAGCCCACT 660
PAB1-E40.2      AAAACTCTTCTGCATCATTATATGTTGGTGACTTAGAACCTTCTGTTTCCGAAGCCCACA 660
PAB1-S40.6      AAAACTCTTCTGCATCATTATATGTTGGTGACTTAGAACCTTCTGTTTCCGAAGCCCACT 660
PAB1-S40.12     AAAACTCTTCTGCATCATTTTATGTTGGTAACTTAGAACCTTCTGTTTCCGAAGCCCACT 660
PAB1-E40.3      AAAACTCTTCTGCATCATTATATGTTGGTGACTTAGAACCTTCTGTTTCCGAAGCCCACT 660
PAB1-E40.1      AAAACTCTTCTGCATCATTATATGTTGGTGACTTAGAACCTTCTGTTTCCGAAGCCCACT 660
                ******** ********** ********* ********************** *******

PAB1-S40.5      TATATGATATCTTCTCTCCAATCGGTTCAGTCTCCTCCATTCGTGTCTGTCGTGATGCTA 720
PAB1-S40.10     TATATGATATCTTCTCTCCAATCGGTTCAGTCTCCTCCATTCGTGTCTGTCGTGATGCCA 720
PAB1-S40.9      TATATGATATCTTCTCTCCAATCGGTTCAGTCTCCTCCATTCGTGTCTGTCGTGATGCCA 720
PAB1-S40.4      TATATGATATCTTCTCTCCAATCGGTTCAGTCTCCTCCATTCGTGTCTGTCGTGATGCCA 720
PAB1-S40.3      TATATGATATCTTCTCTCCAATCGGTTCAGTCTCCTCCATTCGTGTCTGTCGTGATGCCA 720
PAB1-S40.2      TATATGATATCTTCTCTCCAATCGGTTCAGTCTCCTCCATTCGTGTCTGTCGTGATGCCA 720
PAB1-S40.1      TATATGATATCTTCTCTCCAATCGGTTCAGTCTCCTCCATTCGTGTCTGTCGTGATGCCA 720
PAB1-wt         TATATGATATCTTCTCTCCAATCGGTTCAGTCTCCTCCATTCGTGTCTGTCGTGATGCCA 720
```

```
PAB1-S40.7      TATATGATATCTTCTCTCCAATCGGTTCAGTCTCCTCCATTCGTGTCTGTCGTGATGCCA 720
PAB1-S40.8      TATATGATATCTTCTCTCCAATCGGTTCAGTCTCCTCCATTCGTGTCTGTCGTGATGCCA 720
PAB1-S40.11     TATATGATATCTTCTCTCCAATCGGTTCAGTCTCCTCCATTCGTGTCTGTCGTGATGCCA 720
PAB1-E40.2      TATATGATATCTTCTCTCCAATCGGTTCAGTCTCCTCCATTCGTGTCTGTCGTGATGCCA 720
PAB1-S40.6      TATATGATATCTTCTCTCCAATCGGTTCAGTCTCCTCCATTCGTGTCTGTCGTGATGCCA 720
PAB1-S40.12     TATATGATATCTTCTCTCCAATCGGTTCAGTCTCCTCCATTCGTGTCTGTCGTGATGCCA 720
PAB1-E40.3      TATATGATATCTTCTCTCCAATCGGTTCAGTCTCCTCCATTCGTGTCTGTCGTGATGCCA 720
PAB1-E40.1      TATATGATATCTTCTCTCCAATCAGTTCAGTCTCCTCCATTCGTGTCTGTCGTGATGCCA 720
                ********************** ********************************** *

PAB1-S40.5      TCACTAAGACCTCTTTGGGCTATGCTTATGTTAACTTTAACGACCATGAAGCCGGCAGAA 780
PAB1-S40.10     TCACTAAGACCTCTTTGGGCTATACTTATGTTAACTTTAACGACCATGAAGCCGGCAGAA 780
PAB1-S40.9      TCACTAAGACCTCTTTGGGCTATGCTTATGTTAACTTTAACGACCATGAAGCCGGCAGAA 780
PAB1-S40.4      TCACTAAGACCTCTTTGGGCTATGCTTATGTTAACTTTAACGACCATGAAGCCGGCAGAA 780
PAB1-S40.3      TCACTAAGACCTCTTTGGGCTATGCTTATGTTAACTTTAACGACCATGAAGCCGGCAGAA 780
PAB1-S40.2      TCACTAAGACCTCTTTGGGCTATGCTTATGTTAACTTTAACGACCATGGAGCCGGCAGAA 780
PAB1-S40.1      TCACTAAGACCTCTTTGGGCTATGCTTATGTTAACTTTAACGACCATGAAGCCGGCAGAA 780
PAB1-wt         TCACTAAGACCTCTTTGGGCTATGCTTATGTTAACTTTAACGACCATGAAGCCGGCAGAA 780
PAB1-S40.7      TCACTAAGACCTCTTTGGGCTATGCTTATGTTAACTTTAACGACCATGAAGCCGGCAGAA 780
PAB1-S40.8      TCACTAAGACCTCTTTGGGCTATGCTTATGTTAACTTTAACGACCATGAAGCCGGCAGAA 780
PAB1-S40.11     TCACTAAGACCTCTTTGGGCTATGCTTATGTTAACTTTAACGACCATGAAGCCGGCAGAA 780
PAB1-E40.2      TCACTAAGACCTCTTTGGGCTATGCTTATGTTAACTTTAACGACCATGAAGCCGGCAGAA 780
PAB1-S40.6      TCACTAAGACCTCTTTGGGCTATGCTTATGTTAACTTTAACGACCATGAAGCCGGCAGAA 780
PAB1-S40.12     TCACTAAGACCTCTATGGGCTATGCTTATGTTAACTTTAACGACCATGAAGCCGGCAGAA 780
PAB1-E40.3      TCACTAAGACCTCTTTGGGCTATGCTTATGTTAACTTTAACGACCATGAAGCCGGCAGAA 780
PAB1-E40.1      TCACTAAGACCTCTTTGGGCTATGCTTATGTTAACTTTAACGATCATGAAGCCGGCAGAA 780
                ************** ******** ******************* **** ***********

PAB1-S40.5      AAGCAATTGAGCAATTGAACTACACTCCAATCAAGGGTAGATTATGCCGTATTATGTGGT 840
PAB1-S40.10     AAGCAATTGAGCAATTGAACTACACTCCAATCAAGGGTAGATTATGCCGTATTATGTGGT 840
PAB1-S40.9      AAGCAATTGAGCAATTGAACTACACTCCAATCAAGGGTAGATTATGCCGTATTATGTGGT 840
PAB1-S40.4      AATCAATTGAGCAATTGAACTACACTCCAATCAAGGGTAGATTATGCCGTATTATGTGGT 840
PAB1-S40.3      AAGCAATTGAGCAATTGAACTACACTCCAATCAAGGGTAGATTATGCCGTATTATGTGGT 840
PAB1-S40.2      AAGCAATTGAGCAATTGAACTACACTCCAATCAAGGGTAGATTATGCCGTATTATGTGGT 840
PAB1-S40.1      AAGCAATTGAGCAATTGAACTACACTCCAATCAAGGGTAGATTATGCCGTATTATGTGGT 840
PAB1-wt         AAGCAATTGAGCAATTGAACTACACTCCAATCAAGGGTAGATTATGCCGTATTATGTGGT 840
PAB1-S40.7      AAGCAATTGAGCAATTGAACTACACTCCAATCAAGGGTAGATTATGCCGTATTATGTGGT 840
PAB1-S40.8      AAGCAATTGAGCAATTGAACTACACTCCAATCAAGGGTAGATTATGCCGTATTATGTGGT 840
PAB1-S40.11     AAGCAATTGAGCAATTGAACTACACTCCAATCAAGGGTAGATTATGCCGTATTATGTGGT 840
PAB1-E40.2      AAGCAATTGAGCAATTGAACTACACTCCAATCAAGGGTAGATTATGCCGTATTATGTGGT 840
PAB1-S40.6      AAGCAATTGAGCAATTGAACTACACTCCAATCAAGGGTAGATTATGCCGTATTATGTGGT 840
PAB1-S40.12     AAGCAATTGAGCAATTGAACTACACTCCAATCAAGGGTAGATTATGCCGTATTATGTGGT 840
PAB1-E40.3      AAGCAATTGAGCAATTGAACTACACTCCAATCAAGGGTAGATTATGCCGTATTATGTGGT 840
PAB1-E40.1      AAGCAATTGAGCAATTGAACTACACTCCAATCAAGGGTAGATTATGCCGTATTATGTGGT 840
                ** ************************************************************

PAB1-S40.5      CTCAACGTGACCCATCATTGAGAAAGAAGGGTTCTGGTAACATCTTTATCAAGAACTTGC 900
PAB1-S40.10     CTCAACGTGACCCATCATTGAGAAAGAAGGGTTCTGGTAACATCTTTATCAAGAACTTGC 900
PAB1-S40.9      CTCAACGTGACCCATCATTGAGAAAGAAGGGTTCTGGTAACATCTTTATCAAGAACTTGC 900
PAB1-S40.4      CTCAACGTGACCCATCATTGAGAAAGAAGGGTTCTGGTAACATCTTTATCAAGAACTTGC 900
PAB1-S40.3      CTCAACGTGACCCATCATTGAGAAAGAAGGGTTCTGGTAACATCTTTATCAAGAACTTGC 900
PAB1-S40.2      CTCAACGTGACCCATCATTGAGAAAGAAGGGTTCTGGTAACATCTTTATCAAGAACTTGC 900
PAB1-S40.1      CTCAACGTGACCCATCATTGAGAAAGAAGGGTTCTGGTAACATCTTTATCAAGAACTTGC 900
PAB1-wt         CTCAACGTGACCCATCATTGAGAAAGAAGGGTTCTGGTAACATCTTTATCAAGAACTTGC 900
PAB1-S40.7      CTCAACGTGACCCATCATTGAGAAAGAAGGGTTCTGGTAACATCTTTATCAAGAACTTGC 900
PAB1-S40.8      CTCAACGTGACCCATCATTGAGAAAGAAGGGTTCTGGTAACATCTTTATCAAGAACTTGC 900
PAB1-S40.11     CTCAACGTGACCCATCATTGAGAAAGAAGGGTTCTGGTAACATCTTTATCAAGAACTTGC 900
PAB1-E40.2      CTCAACGTGACCCATCATTGAGAAAGAAGGGTTCTGGTAACATCTTTATCAAGAACTTGC 900
PAB1-S40.6      CTCAACGTGACCCATCATTGAGAAAGAAGGGTTCTGGTAACATCTTTATCAAGAACTTGC 900
PAB1-S40.12     CTCAACGTGACCCATCATTGAGAAAGAAGGGTTCTGGTAACATCTTTATCAAGAACTTGC 900
PAB1-E40.3      CTCAACGTGACCCATCATTGAGAAAGAAGGGTTCTGGTAACATCTTTATCAAGAACTTGC 900
PAB1-E40.1      CTCACCGTGACCCATCATTGAGAAAGAAGGGTTCTGGTAACATCTTTATCAAGAACTTGC 900
                **** ************************************************************

PAB1-S40.5      ACCCTGATATTGACAACAAGGCTTTATATGACACTTTCTCTGTGTTTGGTGACATCTTGT 960
PAB1-S40.10     ACCCTGATATTGACAACAAGGCTTTGTATGACACTTTCTCTGTGTTTGGTGACATCTTGT 960
PAB1-S40.9      ACCCTGATATTGACAACAAGGCTTTGTAAGACACTTTCTCTGTGTTTGGTGACATCTTGT 960
PAB1-S40.4      ACCCTGATATTGACAACAAGGCTTTGTATGACACTTTCTCTGTGTTTGGTGACATCTTGT 960
PAB1-S40.3      ACCCTGATATTGACAACAAGGCTTTGTATGACACTTTCTCTGTGTTTGGTGACATCTTGT 960
PAB1-S40.2      ACCCTGATATTGACAACAAGGCTTTGTATGACACTTTCTCTGTGTTTGGTGACATCTTGT 960
PAB1-S40.1      ACCCTGATATTGACAACAAGGCTTTGTATGACACTTTCTCTGTGTTTGGTGACATCTTGT 960
PAB1-wt         ACCCTGATATTGACAACAAGGCTTTGTATGACACTTTCTCTGTGTTTGGTGACATCTTGT 960
```

```
PAB1-S40.7      ACCCTGATATTGACAACAAGGCTTTGTATGACACTTTCTCTGTGTTTGGTGACATCTTGT 960
PAB1-S40.8      ACCCTGATATTGACAACAAGGCTTTGTATGACACTTTCTCTGTGTTTGGTGACATCTTGT 960
PAB1-S40.11     ACCCTGATATTGACAACAAGGCTTTGTATGACACTTTCTCTGTGTTTGGTGACATCTTGT 960
PAB1-E40.2      ACCCTGATATTGACAACAAGGCTTTGTATGACACTTTCTCTGTGTTTGGTGACATCTTGT 960
PAB1-S40.6      ACCCTGATATTGACAACAAGGCTTTGTATGACACTTTCTCTGTGTTTGGTGACATCTTGT 960
PAB1-S40.12     ACCCTGATATTGACAACAAGGCTTTGTATGACACTTTCTCTGTGTTTGGTGACATCTTGT 960
PAB1-E40.3      ACCCTGATATTGACAACAAGGCTTTGTATGACACTTTCTCTGTGTTTGGTGACATCTTGT 960
PAB1-E40.1      ACCCTGATATTGACAACAAGGCTTTGTATGACACTTTCTCTGTGTTTGGTGACATCTTGT 960
                ************************* ** ******************************

PAB1-S40.5      CCAGCAAGATTGCCACCGACGAAAACGGAAAATCCAAGGGTTTTGGGTTTGTTCACTTCG 1020
PAB1-S40.10     CCAGCAAGATTGCCACCGACGAAAACGGAAAATCCAAGGGTTTTGGGTTTGTTCACTTCG 1020
PAB1-S40.9      CCAGCAAGATTGCCACCGACGAAAACGGAAAATCCAAGGGTTTTGGGTTTGTTCACTTCG 1020
PAB1-S40.4      CCAGCAAGATTGCCACCGACGAAAACGGAAAATCCAAGGGTTTTGGGTTTGTTCACTTCG 1020
PAB1-S40.3      CCAGCAAGATTGCCACCGACGAAAACGGAAAATCCAAGGGTTTTGGGTTTGTTCACTTCG 1020
PAB1-S40.2      CCAGCAAGATTGCCACCGACGAAAACGGAAAATCCAAGGGTTTTGGGTTTGTTCACTTCG 1020
PAB1-S40.1      CCAGCAAGATTGCCACCGACGAAAACGGAAAATCCAAGGGTTTTGGGTTTGTTCACTTCG 1020
PAB1-wt         CCAGCAAGATTGCCACCGACGAAAACGGAAAATCCAAGGGTTTTGGGTTTGTTCACTTCG 1020
PAB1-S40.7      CCAGCAAGATTGCCACAGACGAAAACGGAAAATCCAAGGGTTTTGGGTTTGTTCACTTCG 1020
PAB1-S40.8      CCAGCAAGATTGCCACCGACGAAAACGGAAAATCCAAGGGTTTTGGGTTTGTTCACTTCG 1020
PAB1-S40.11     CCAGCAAGATTGCCATCGACGAAAACGGAAAATCCAAGGGTTTTGGGTTTGTTCACTTCG 1020
PAB1-E40.2      CCAGCAAGATTGCCACCGACGAAAACGGAAAATCCAAGGGTTTTGGGTTTGTTCACTTCG 1020
PAB1-S40.6      CCAGCAAGATTGCCACCGACGAAAACGGAAAATCCAAGGGTATTGGGTTTGTTCACTTCG 1020
PAB1-S40.12     CCAGCAAGATTGCGACCGACGAAAACGGAAAATCCAAGGGTTTTGGGTTTGTTCACATCG 1020
PAB1-E40.3      CCAGCAAGATTGCCACCGACGAAAACGGAAAATCCAAGGGTTTTGGGTTTGTTCACTTCG 1020
PAB1-E40.1      CCAGCAAGATTGCCACCGACGAAAACGGAAAATCCAAGGGTTTTGGGTTTGTTCACTTCG 1020
                ************* *  ************************ ************** ***

PAB1-S40.5      AAGAAGAAGGTGCTGCCAAGGAAGCTATTGATGCTTTGAATGGTATGATGTTTGAACGGTC 1080
PAB1-S40.10     AAGAAGAAGGTGCTGCCAAGGAAGCTATTGATGCTTTGAATGGTATGCTGTTGAACGGTC 1080
PAB1-S40.9      AAGAAGAAGGTGCTGCCAAGGAAGCTATTGATGCTTTGAATGGTATGCAGTTGAACGGTC 1080
PAB1-S40.4      AAGAAGAAGGTGCTGCCAAGGAAGCTTTTGATGCTTTGAATGGTATGCTGTTGAACGGTC 1080
PAB1-S40.3      AAGAAGAAGGTGCTGCCAAGGAAGCTATTGATGCTTTGAATGGTATGCTGTTGAACGGTC 1080
PAB1-S40.2      AAGAAGAAGGTGCTGCCAAGGAAGCTATTGATGCTTTGAATGGTATGCTGTTGAACGGTC 1080
PAB1-S40.1      AAGAAGAAGGTGCTGCCAAGGAAGCTATTGATGCTTTGAATGGTATGCTGTTGAACGGTC 1080
PAB1-wt         AAGAAGAAGGTGCTGCCAAGGAAGCTATTGATGCTTTGAATGGTATGCTGTTGAACGGTC 1080
PAB1-S40.7      AAGAAGAAGGTGCTGCCAAGGAAGCTATTGATGCTTTGAATGGTATGCTGTTGAACGGTC 1080
PAB1-S40.8      AAGAAGAAGGTGCTGCCAAGGAAGCTATTGATGCTTTGAATGGTATGCTGTTGAACGGTC 1080
PAB1-S40.11     AAGAAGAAGGTGCTGCCAAGGAAGCTATTGATGCTTTGAATGGTATGCTGTTGAACGGTC 1080
PAB1-E40.2      AAGAAGAAGGTGCTGCCAAGGAAGCTATTGATGCTTTGAATGGTATGCTGTTGAACGGTC 1080
PAB1-S40.6      AAGAAGAAGGTGCTGCCAAGGAAGCTATTGATGCTTTGAATGGTATGCTGTTGAACGGTC 1080
PAB1-S40.12     AAGAAGAAGGTGCTGCCAAGAAAGCTATTGATGCTTTGAATGGTATGCTGTTGAACGGTC 1080
PAB1-E40.3      AAGAAGAAGGTGCTGCCAAGGAAGCTATTGATGCTTTGAATGGTATGCTGTTGAACGGTC 1080
PAB1-E40.1      AAGAAGAAGGTGCTGCCAAGGAAGCTATTGATGCTTTGAATGGTATGCTGTTGAACGGTC 1080
                ******************** ***** ********************  ***********

PAB1-S40.5      AAGAAACTTATGTTGCTCCTCACTTGTCCAGAAAGGTACGTGACTCTCAATTGGAAGAGA 1140
PAB1-S40.10     AAGAAATTTATGTTGCTCCTCACTTGTCCAGAAAGGAACGTGACTCTCAATTGGAAGAGA 1140
PAB1-S40.9      AAGAAATTTATGTTGCTCCTCACTTGTCCAGAAAGGAACGTAACTCTCAATTGGAAGAGA 1140
PAB1-S40.4      AAGAAATTTATGTTGCTCCTCACTTGTCCAGAAAGGAACGTGACTCTCAATTGGAAGAGA 1140
PAB1-S40.3      AAGAAATTTATGTTGCTCCTCACTTGTCCAGAAAGGAACGTGACTCTCAATTGGAAGAGA 1140
PAB1-S40.2      AAGAAATTTATGTTGCTCCTCACTTGTCCAGAAAGGAACGTGACTCTCAATTGGAAGAGA 1140
PAB1-S40.1      AAGAAATTTATGTTGCTCCTCACTTGTCCAGAAAGGAACGTGACTCTCAATTGGAAGAGA 1140
PAB1-wt         AAGAAATTTATGTTGCTCCTCACTTGTCCAGAAAGGAACGTGACTCTCAATTGGAAGAGA 1140
PAB1-S40.7      AAGAAATTTATGTTGCTCCTCACTTGTCCAGAAAGGAACGTGACTCTCAATTGGAAGAGA 1140
PAB1-S40.8      AAGAAATTTATGTTGCTCCTCACTTGTCCAGAAAGGAACGTGACTCTCAATTGGAAGAGA 1140
PAB1-S40.11     AAGAAATTTATGTTGCTCCTCACTTGTCCAGAAAGGAACGTGACTCTCAATTGGAAGAGA 1140
PAB1-E40.2      AAGAAATTTATGTTGCTCCTCACTTGTCCAGAAAGGAACGTGACTCTCAATTGGAAGAGA 1140
PAB1-S40.6      AAGAAATTTATGTTGCTCCTCACTCGTCCAGAAAGGAACGTGACTCTCAATTGGAAGAGA 1140
PAB1-S40.12     AAGAAATTTATGTTGCTCCTCACTTGTCCAGAAAGGAACGTGACTCTCAATTGGAAGAGA 1140
PAB1-E40.3      AAGAAATTTATGTTGCTCCTCACTTGTCCAGAAAGGAACGTGACTCTCAATTGGAAGAGA 1140
PAB1-E40.1      AAGAAATTTATGTTGCTCCTCACTTGTCCAGAAAGGAACGTGACTCTCAATTGGAAGAGA 1140
                ****** ***************** ************ **** ******************

PAB1-S40.5      CTAAGGCACATTACACTAACCTTTATGTGAAAAACATCAACTCCGAAACTACTGACGAAC 1200
PAB1-S40.10     CTAAGGCACATTACACTAACCTTTATGTGAAAAACATCAACTCCGAAACTACTGACGAAC 1200
PAB1-S40.9      CTAAGGCACATTACACTAACCTTTATGTGAAAAACATCAACTCCGAAACTACTGACGAAC 1200
PAB1-S40.4      CTAAGGCACATTACACTAACCTTTATGTGAAAAACATCGACTCCGAAACTACTGACGAAC 1200
PAB1-S40.3      CTAAGGCACATTACACTAACCTTTATGTGAAAAACATCAACTCCGAAACTACTGACGAAC 1200
PAB1-S40.2      CTAAGGCACATTACACTAACCTTTATGTGAAAAACATCAACTCCGAAACTACTGACGAAC 1200
PAB1-S40.1      CTAAGGCACATTACACTAACCTTTATGTGAAAAACATCAACTCCGAAACTACTGACGAAC 1200
PAB1-wt         CTAAGGCACATTACACTAACCTTTATGTGAAAAACATCAACTCCGAAACTACTGACGAAC 1200
```

```
PAB1-S40.7      CTAAGGCACATTACACTAACCTTTATGTGAAAAACATCAACTCCGAAACTACTGACGAAC 1200
PAB1-S40.8      CTAAGGCACATTACACTAACCTTTATGTGAAAAACATCAACTCCGAAACTACTGACGAAC 1200
PAB1-S40.11     CTAAGGCACATTACACTAACCTTTATGTGAAAAACATCAACTCCGAAACTACTGACGAAC 1200
PAB1-E40.2      CTAAGGCACATTACACTAACCTTTATGTGAAAAACATCAACTCCGAAACTACTGACGAAC 1200
PAB1-S40.6      CTAAGGCACATTACACTAACCTTTATGTGAAAAACATCAACTCCGAAACTACTGACGAAC 1200
PAB1-S40.12     CTAAGGCACATTACACTAACCTTTATGTGAAAAACATCAACTCCGAAACTACTGACGAAC 1200
PAB1-E40.3      CTAAGGCACATTACACTAACCTTTATGTGAAAAACATCAACTCCGAAACTACTGACGAAC 1200
PAB1-E40.1      CTAAGGCACATTACACTAACCTTTATGTGAAAAACATCAACTCCGAAACTACTGACGAAC 1200
                ************************************ ********************

PAB1-S40.5      AATTCCAAGAATTGTTTGCCAAATTTGGTCCAATTGTTTCTGCCTCTTTGGAAAAGGATG 1260
PAB1-S40.10     AATTCCAAGAATTGTTTGCCAAATTTGGTCCAATTGTTTCTGCCTCTTTGGAAAAGGATG 1260
PAB1-S40.9      AATTCCAAGAATTGTTTGCCAAATTTGGTCCAATTGTTTCTGCCTCTTTGGAAAAGGATG 1260
PAB1-S40.4      AATTCCAAGAATTGTTTGCCAAATTTGGTCCAATTGTTTCTGCCTCTTTGGAAAAGGATG 1260
PAB1-S40.3      AATTCCAAGAATTGTTTGCCAAATTTGGTCCAATTGTTTCTGCCTCTTTGGAAAAGGATG 1260
PAB1-S40.2      AATTCCAAGAATTGTTTGCCAAATTTGGTCCAATTGTTTCTGCCTCTTTGGAAAAGGATG 1260
PAB1-S40.1      AATTCCAAGAATTGTTTGCCAAATTTGGTCCAATTGTTTCTGCCTCTTTGGAAAAGGATG 1260
PAB1-wt         AATTCCAAGAATTGTTTGCCAAATTTGGTCCAATTGTTTCTGCCTCTTTGGAAAAGGATG 1260
PAB1-S40.7      AATTCCAAGAATTGTTTGCCAAATTTGGTCCAATTGTTTCTGCCTCTTTGGAAAAGGATG 1260
PAB1-S40.8      AATTCCAAGAATTGTTTGCCAAATTTGGTCCAATTGTTTCTGCCTCTTTGGAAAAGGATG 1260
PAB1-S40.11     AATTCCAAGAATTGTTTGCCAAATTTGGTCCAATTGTTTCTGCCTCTTTGGAAAAGGATG 1260
PAB1-E40.2      AATTCCAAGAATTGTTTGCCAAATTTGGTCCAATTGTTTCTGCCTCTTTGGAAAAGGATG 1260
PAB1-S40.6      AATTCCAAGAATTGTTTGCCAAATTTGGTCCAATTGTTTCTGCCTCTTTGGAAAAGGATG 1260
PAB1-S40.12     AATTCCAAGAATTGTTTGCCAAATTTGGTCCAATTGTTTCTGCCTCTTTGGAAAAGGATG 1260
PAB1-E40.3      AATTCCAAGAATTGTTTGCCAAATTTGGTCCAATTGTTTCTGCCTCTTTGGAAAAGGATG 1260
PAB1-E40.1      AATTCCAAGAATTGTTTGCCAAATTTGGTCCAATTGTTTCTGCCTCTTTGGAAAAGGATG 1260
                ************************************************************

PAB1-S40.5      CTGATGGAAAATTGAAGGGTTTCGGGTTTGTTAACTACGAAAAGCATGAAGACGCTGTGA 1320
PAB1-S40.10     CTGATGGAAAATTGAAGGGTTTCGGGTTTGTTAACTACGAAAAGCATGAAGACGCTGTGA 1320
PAB1-S40.9      CTGATGGAAAATTGAAGGGTTTCGGGTTTGTTAACTACGAAAAGCATGAAGACGCTGTGA 1320
PAB1-S40.4      CTGATGGAAAATTGAAGGGTTTCGGGTTTGTTAACTACGAAAAGCATGAAGACGCTGTGA 1320
PAB1-S40.3      CTGATGGAAAATTGAAGGGTTTCGGGTTTGTTAACTACGAAAAGCATGAAGACGCTGTGA 1320
PAB1-S40.2      CTGATGGAAAATTGAAGGGTTTCGGGTTTGTTAACTACGAAAAGCATGAAGACGCTGTGA 1320
PAB1-S40.1      CTGATGGAAAATTGAAGGGTTTCGGGTTTGTTAACTACGAAAAGCATGAAGACGCTGTGA 1320
PAB1-wt         CTGATGGAAAATTGAAGGGTTTCGGGTTTGTTAACTACGAAAAGCATGAAGACGCTGTGA 1320
PAB1-S40.7      CTGATGGAAAATTGAAGGGTTTCGGGTTTGTTAACTACGAAAAGCATGAAGACGCTGTGA 1320
PAB1-S40.8      CTGATGGAAAATTGAAGGGTTTCGGGTTTGTTAACTACGAAAAGCATGAAGACGCTGTGA 1320
PAB1-S40.11     CTGATGGAAAATTGAAGGGTTTCGGGTTTGTTAACTACGAATAGCATGAAGACGCTGTGA 1320
PAB1-E40.2      CTGATGGAAAATTGAAGGGTTTCGGGTTTGTTAACCACGAAAAGCATGAAGACGCTGTGA 1320
PAB1-S40.6      CTGATGGAAAATTGAAGGGTTTCGGGTTTGTTAACTACGAAAAGCATGAAGACGCTGTGA 1320
PAB1-S40.12     CTGATGGAAAATTGAAGGGTTTCGGGTTTGTTAACTACGAAAAGCATGAAGACGCTGTGA 1320
PAB1-E40.3      CTGATGGAAAATTGAAGGGTTTCGGGTTTGTTAACTACGAAAAGCATGAAGACGCTGTGA 1320
PAB1-E40.1      CTGATGGAAAATTGAAGGGTTTCGGGTTTGTTAACTACGAAAAGCATGAAGACGCTGTGA 1320
                ************************************ ***** ******************

PAB1-S40.5      AAGCTGTTGAAGCTTTGAATGACTCTGAACTAAATGGAGAAAAGTTATACGTTGGTCGTG 1380
PAB1-S40.10     AAGCTGTTGAAGCTTTGAATGACTCTGAACTAAATGGAGAAAAGTTATACGTTGGTCGTG 1380
PAB1-S40.9      AAGCTGTTGAAGCTTTGAATGACTCTGAACTAAATGGAGAAAAGTTATACGTTGGTCGTG 1380
PAB1-S40.4      AAGCTGTTGAAGCTTTGAATGACTCTGAACTAAATGGAGAAAAGTTATACGTTGGTCGTG 1380
PAB1-S40.3      AAGCTGTTGAAGCTTTGAATGACTCTGAACTAAATGGAGAAAAGTTATACGTTGGTCGTG 1380
PAB1-S40.2      AAGCTGTTGAAGCTTTGAATGACTCTGAACTAAATGGAGAAAAGTTATACGTTGGTCGTG 1380
PAB1-S40.1      AAGCTGTTGAAGCTTTGAATGACTCTGAACTAAATGGAGAAAAGTTATACGTTGGTCGTG 1380
PAB1-wt         AAGCTGTTGAAGCTTTGAATGACTCTGAACTAAATGGAGAAAAGTTATACGTTGGTCGTG 1380
PAB1-S40.7      AAGCTGTTGAAGCTTTGAATGACTCTGAACTAAATGGAGAAAAGTTATACGTTGGTCGTG 1380
PAB1-S40.8      AAGCTGTTGAAGCTCTGAATGACTCTGAACTAAATGGAGAAAAGTTATACGTTGGTCGTG 1380
PAB1-S40.11     AAGCTGTTGAAGCTTTGAATGACTCTGACCTAAATGGAGAAAAGTTGTACGTTGGTCATG 1380
PAB1-E40.2      TAGCTGTT-AAGCTTTGAATGACTCTGACCTAAATGGAGAAAAGTTATACGTTGGTCGTG 1380
PAB1-S40.6      AAGCTGTTGAAGCTTTGAATGACTCTGACCTAAATGGAGAAAAGTTATACGTTGGTCGTG 1380
PAB1-S40.12     AAGCTGTTGAAGCTTTGAATGACTCTGACCTAAATGGAGAAAAGTTATACGTTGGTCGTG 1380
PAB1-E40.3      AAGCTGTTGAAGCTTTGAATGACTCTGACCTAAATGGAGAAAAGTTATACGTTGGTCGTG 1380
PAB1-E40.1      AAGCTGTTGAAGCTTTGAATGACTCTGACCTAAATGGAGAAAAGTTATACGTTGATCGTG 1380
                 ******* ***** ************* ***************** ******* ** **

PAB1-S40.5      CCCAAAAGGAGAATGAACGTATGCATGTCTTGAAGAAGCAATACGAAGCTTACAGATTGG 1440
PAB1-S40.10     CCCAAAAGAAGAATGAACGTATGCATGTCTTGAAGAAGCAATACGAAGCTTACAGATTGG 1440
PAB1-S40.9      CCCAAAAGAAGAATGAACGTATGCATGTCTTGAAGAAGCAATACGAAGCTTACAGATTGG 1440
PAB1-S40.4      CCCAAAAGAAGAATGAACGTATGCATGTCTTGAAGAAGCAATACGAAGCTTACAGATTGG 1440
PAB1-S40.3      CCCAAAAGAAGAATGAACGTATGCATGTCTTGAAGAAGCAATACGAAGCTTACAGATTGG 1440
PAB1-S40.2      CCCAAAAGAAGAATGAACGTATGCATGTCTTGAAGAAGCAATACGAAGCTTACAGATTGG 1440
PAB1-S40.1      CCCAAAAGAAGAATGAACGTATGCATGTCTTGAAGAAGCAATACGAAGCTTACAGATTGG 1440
PAB1-wt         CCCAAAAGAAGAATGAACGTATGCATGTCTTGAAGAAGCAATACGAAGCTTACAGATTGG 1440
```

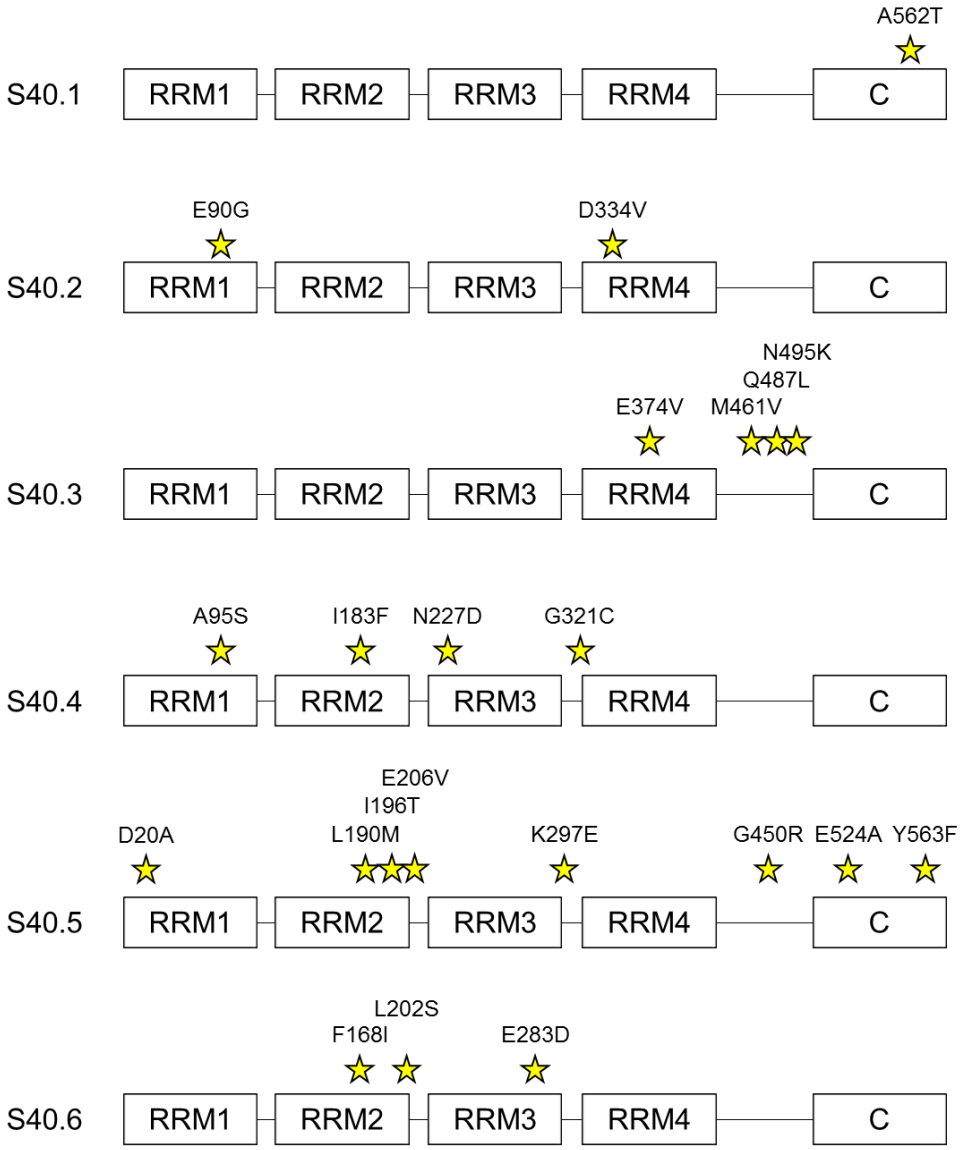
```
PAB1-S40.7      CCCAAAAGAAGAATGAACGTATGCATGTCTTGAAGAAGCAATACGAAGCTTACAGATTGG 1440
PAB1-S40.8      CCCAAAAGAAGAATGAACGTATGCATGTCTTGAAGAAGCAATACGAAGCTTACAGATTGG 1440
PAB1-S40.11     CCCAAAAGAAGAATGAACGTATGCATGTCTTGAAGAAGCAATACGAAGCTTACAGATTGG 1440
PAB1-E40.2      CCCAAAAGAAGAATGAACGTATGCATGTCTTGAAGAAGCAATACGAAGCTTACAGATTGG 1440
PAB1-S40.6      CCCAAAAGAAGAATGAACGTATGCATGTCTTGAAGAAGCAATACGAAGCTTACAGATTGG 1440
PAB1-S40.12     CCCAAAAGAAGAATGAACGTATGCATGTCTTGAAGAAGCAATACGAAGCTTACAGATTGG 1440
PAB1-E40.3      CCCAAAAGAAGAATGAACGTATGCATGTCTTGAAGAAGCAATACGAAGCTTACAGATTGG 1440
PAB1-E40.1      CCCAAAAGAAGAATGAACGTATGCATGTCTTGAAGAAGCAATACGAAGCTTACAGATTGG 1440
                ******** ************************************************** 

PAB1-S40.5      AAAAAATGGCCAAGTACCAAGGTGTTAATTTGTTTGTTAAGAACTTAGATGACAGCGTTG 1500
PAB1-S40.10     AAAAAATGGCCAAGTACCAAGGTGTTAATTTGTTTGTTAAGAACTTAGATGACATCGTTG 1500
PAB1-S40.9      AAAAAATGGCCAAGTACCAAGGTGTTAATTTGTTTGTTAAGAACTTAGATGACAGCGTTG 1500
PAB1-S40.4      AAAAAATGGCCAAGTACCAATGTGTTAATTTGTTTGTTAAGAACTTAGATGACAGCGTTG 1500
PAB1-S40.3      AAAAAATGGCCAAGTACCAAGGTGTTAATTTGTTTGTTAAGAACTTAGATGACAGCGTTG 1500
PAB1-S40.2      AAAAAATGGCCAAGTACCAAGGTGTTAATTTGTTTGTTAAGAACTTAGATGACAGCGTTG 1500
PAB1-S40.1      AAAAAATGGCCAAGTACCAAGGTGTTAATTTGTTTGTTAAGAACTTAGATGACAGCGTTG 1500
PAB1-wt         AAAAAATGGCCAAGTACCAAGGTGTTAATTTGTTTGTTAAGAACTTAGATGACAGCGTTG 1500
PAB1-S40.7      AAAAAATGGCCAAGTACCAAGGTGTTAATTTGTTTGTTAAGAACTTGGATGACAGCGTTG 1500
PAB1-S40.8      AAAAAATGGCCAAGTACCAAGGTGTTAATTTGTTTGTTAAGAACTTAGATGACAGCGTTG 1500
PAB1-S40.11     AAAAAATGGCCAAGTACCAAGGTGTTAATTTGTTTGTTAAGAACTTAGATGACAGCGTTG 1500
PAB1-E40.2      AAAAAATGGCCAAGTACCAAGGTGTTAATTTGTTTGTCAAGAACTTAGATGACAGCGTTG 1500
PAB1-S40.6      AAAAAATGGCCAAGTACCAAGGTGTTAATTTGTTTGTTAAGAACTTAGATGACAGCGTTG 1500
PAB1-S40.12     AAAAAATGGCCAAGTACCAAGGTGTTAATTTGTTTGTTAAGAACTTAGATGACAGCGTTG 1500
PAB1-E40.3      AAAAAATGGCCAAGTACCAAGGTGTTAATTTGTTTGTTAAGAACTTAGATGACAGCGTTG 1500
PAB1-E40.1      AAAAAATGGCCAAGTACCAAGGTGTTAATTTGTTTGTTAAGAACTTAGATGACAGCGTTG 1500
                ******************** ***************** ******** ******* ***** 

PAB1-S40.5      ATGACGAAAAGTTGGAAGAAGAATTTGCTCCATATGGTACTATCACTTCTGCAAAGGTTA 1560
PAB1-S40.10     ATGACGAAAAGTTGGAAGAAGAATTTGCTCCATATGGTACTATCACTTCTGCAAAGGTTA 1560
PAB1-S40.9      ATGACGAAAAGTTGGAAGAAGAATTTGCTCCATATGGTACTATCACTTCTGCAAAGGTTA 1560
PAB1-S40.4      ATGACGAAAAGTTGGAAGAAGAATTTGCTCCATATGGTACTATCACTTCTGCAAAGGTTA 1560
PAB1-S40.3      ATGACGAAAAGTTGGAAGAAGAATTTGCTCCATATGGTACTATCACTTCTGCAAAGGTTA 1560
PAB1-S40.2      TTGACGAAAAGTTGGAAGAAGAATTTGCTCCATATGGTACTATCACTTCTGCAAAGGTTA 1560
PAB1-S40.1      ATGACGAAAAGTTGGAAGAAGAATTTGCTCCATATGGTACTATCACTTCTGCAAAGGTTA 1560
PAB1-wt         ATGACGAAAAGTTGGAAGAAGAATTTGCTCCATATGGTACTATCACTTCTGCAAAGGTTA 1560
PAB1-S40.7      ATGACGAAAAGTTGGAAGAAGAATTTGCTCCATATGGTACTATCACTTCTGCAAAGGTTA 1560
PAB1-S40.8      ATGACGAAAAGTTGGAAGAAGAATTTGCTCCATATGGTACTATCACTTCTGCAAAGGTTA 1560
PAB1-S40.11     ATGACGAAAAGTTGGAAGAAGAATTTGCTCCATATGGTACTATCACTTCTGCAAAGGTTA 1560
PAB1-E40.2      ATGACGAAAAGTTGGAAGAAGAATTTGCTCCATATGGTACTATCACTTCTGCAAAGGTTA 1560
PAB1-S40.6      ATGACGAAAAGTTGGAAGAAGAATTTGCTCCATATGGTACTATCACTTCTGCAAAGGTTA 1560
PAB1-S40.12     ATGACGAAAAGTTGGAAGAAGAATTTGCTCCATATGGTACTATCACTTCTGCAAAGGTTA 1560
PAB1-E40.3      ATGACGAAAAGTTGGAAGAAGAATTTGCTCCATATGGTACTATCACTTCTGCAAAGGTTA 1560
PAB1-E40.1      ATGACGAAAAGTTGGAAGAAGAATTTGCTCCATATGGTACTATCACTTCTGCAAAGGTTA 1560
                 ************************************************************ 

PAB1-S40.5      TGAGAACCGAAAACGGTAAGTCTAAGGGTTTTGGTTTTGTTTGTTTCTCAACTCCAGAGG 1620
PAB1-S40.10     TGAGAACCGAAAACGGTAAGTCTAAGGGTTTTGGTTTTGTTTGTTTCTCAACTCCAGAGG 1620
PAB1-S40.9      TGAGAACCGAAAACGGTAAGTCTAAGGGTTTTGGTTTTGTTTGTTTCTCAACTCCAGAGG 1620
PAB1-S40.4      TGAGAACCGAAAACGGTAAGTCTAAGGGTTTTGGTTTTGTTTGTTTCTCAACTCCAGAGG 1620
PAB1-S40.3      TGAGAACCGAAAACGGTAAGTCTAAGGGTTTTGGTTTTGTTTGTTTCTCAACTCCAGAGG 1620
PAB1-S40.2      TGAGAACCGAAAACGGTAAGTCTAAGGGTTTTGGTTTTGTTTGTTTCTCAACTCCAGAGG 1620
PAB1-S40.1      TGAGAACCGAAAACGGTAAGTCTAAGGGTTTTGGTTTTGTTTGTTTCTCAACTCCAGAGG 1620
PAB1-wt         TGAGAACCGAAAACGGTAAGTCTAAGGGTTTTGGTTTTGTTTGTTTCTCAACTCCAGAGG 1620
PAB1-S40.7      TGAGAACCGAAAACGGTAAGTCTAAGGGTTTTGGTTTTGTTTGTTTCTCAACTCCAGAGG 1620
PAB1-S40.8      TGAGAACCAAAAACGGTAAGTCTAAGGGTTTTGGTTTTGTTTGTTTCTCAACTCCAGAGG 1620
PAB1-S40.11     TGAGAACCGAAAACGGTAAGTCTAAGGGTTTTGGTTTTGTTTGTTTCTCAACTCCAGAGG 1620
PAB1-E40.2      TGAGAACCGAAAACGGTAAGTCTAAGGGTTTTGGTTTTGTTTGTTTCTCAACTCCAGAGG 1620
PAB1-S40.6      TGAGAACCGAAAACGGTAAGTCTAAGGGTTTTGGTTTTGTTTGTTTCTCAACTCCAGAGG 1620
PAB1-S40.12     TGAGAACCGAAAACGGTAAGTCTAAGGGTTTTGGTTTTGTTTGTTTCTCAACTCCAGAGG 1620
PAB1-E40.3      TGAGAACCGAAAACGGTAAGTCTAAGGGTTTTGGTTTTGTTTGTTTCTCAACTCCAGAGG 1620
PAB1-E40.1      TGAGAACCGAAAACGGTAAGTCTAAGGGTTTTGGTTTTGTTTGTTTCTCAACTCCAGAGG 1620
                ******** ************************************************** 

PAB1-S40.5      AAGCTACTAAGGCCATTACAGAAAAGAACCAACAAATTGTTGCTGGTAAGCCATTATACG 1680
PAB1-S40.10     AAGCTACTAAGGCCATTACAGAAAAGAACCAACAAATTGTTGCTGGTAAGCCATTATACG 1680
PAB1-S40.9      AAGCCACTAAGGCCATTACAGAAAAGAACCAACAAATTGTTGCTGGTAAGCCATTATACG 1680
PAB1-S40.4      AAGCTACTAAGGCGATTACAGAAAAGAACCAACAAATTGTTGCTGGTAAGCCATTATACG 1680
PAB1-S40.3      TAGCTACTAAGGCCATTACAGAAAAGAACCAACAAATTGTTGCTGGTAAGCCATTATACG 1680
PAB1-S40.2      AAGCTACTAAGGCCATTACAGAAAAGAACCAACAAATTGTTGCTGGTAAGCCATTATACG 1680
PAB1-S40.1      AAGCTACTAAGGCCATTACAGAAAAGAACCAACAAATTGTTGCTGGTAAGCCATTATACG 1680
PAB1-wt         AAGCTACTAAGGCCATTACAGAAAAGAACCAACAAATTGTTGCTGGTAAGCCATTATACG 1680
```

```
PAB1-S40.7      AAGCTACTAAGGCCATTACAGAAAAGAACCAACAAATTGTTGCTGGTAAGCCATTATACG 1680
PAB1-S40.8      AAGCTACTAAGGCCATTACAGAAAAGAACCAACAAATTGTTGCTGGTAAGCCATTATACG 1680
PAB1-S40.11     AAACTACTAAGGCCATTACAGAAAAGAACCAACAAATTGTTGCTGGTAAGCCATTATACG 1680
PAB1-E40.2      AAGCTACTAAGGCCATTACAGAAAAGAACCAACAAATTGTTGCTGGTAAGCCATTATACG 1680
PAB1-S40.6      AAGCTACTAAGGCCATTACAGAAAAGAACCAACAAATTGTTGCTGGTAAGCCATTATACG 1680
PAB1-S40.12     AAGCTACTAAGGCCATTACAGAAAAGAACCAACAAATTGTTGCTGGTAAGCCATTATACG 1680
PAB1-E40.3      AAGCTACTAAGGCCATTACAGAAAAGAACCAACAAATTGTTGCTGGTAAGCCATTATACG 1680
PAB1-E40.1      AAGCTACTAAGGCCAATACAGAAAAGAACCAACAAATTGTTGCTGGTAAGCCATTATACG 1680
                 * * ******** * ********************************************

PAB1-S40.5      TTGCCATTGCTCAAAGAAAAGACGTAAGACGTTCTCAATTGGCTCAACAAATCCAAGCCA 1740
PAB1-S40.10     TTGCCATTGCTCAAAGAAAAGACGTAAGACGTTCTCAATTGGCTCAACAAATCCAAGCCA 1740
PAB1-S40.9      TTGCCATTGCTCAAAGAAAAGATGTAAGACGTTCTCAGTTGGCTCAACAAATCCAAGCCA 1740
PAB1-S40.4      TTGCCATTGCTCAAAGAAAAGACGTAAGACGTTCTCAATTGGCTCAACAAATCCAAGCCA 1740
PAB1-S40.3      TTGCCATTGCTCAAAGAAAAGACGTAAGACGTTCTCAATTGGCTCAACAAATCCAAGCCA 1740
PAB1-S40.2      TTGCCATTGCTCAAAGAAAAGACGTAAGACGTTCTCAATTGGCTCAACAAATCCAAGCCA 1740
PAB1-S40.1      TTGCCATTGCTCAAAGAAAAGACGTAAGACGTTCTCAATTGGCTCAACAAATCCAAGCCA 1740
PAB1-wt         TTGCCATTGCTCAAAGAAAAGACGTAAGACGTTCTCAATTGGCTCAACAAATCCAAGCCA 1740
PAB1-S40.7      TTGCCATTGCTCAAAGAAAAGACGTAGGACGTTCTCAATTGGCTCAACAAATCCAAGCCA 1740
PAB1-S40.8      TTGCCATTGCTCAAAGAAAAGACGTAAGACGTTCTCAATTGGCTCAACAAATCCAAGCCA 1740
PAB1-S40.11     TTGCCATTGCTCAAAGAAAAGACGTAAGACGTTCTCAATTGGCTCAACAAATCCAAGCCA 1740
PAB1-E40.2      TTGCCATTGCTCAAAGAAAAGACGTAAGACGTTCTCAATTGGCTCAACAAATCCAAGCCA 1740
PAB1-S40.6      TTGCCATTGCTCAAAGAAAAGACGTAAGACGTTCTCAATTGGCTCAACAAATCCAAGCCA 1740
PAB1-S40.12     TTGCCATTGCTCAAAGAAAAGACGTAAGACGTTCTCAATTGGCTCAACAAATCCAAGCCA 1740
PAB1-E40.3      TTGCCATTGCTCAAAGAAAAGACGTAAGACGTTCTCAATTGGCTCAACAAATCCAAGCCA 1740
PAB1-E40.1      TTGCCATTGCTCAAAGAAAAGACGTAAGACGTTCTCAATTGGCTCAACAAATCCAAGCCA 1740
                ********************** *** ********** **********************

PAB1-S40.5      GAAATCAAATGAGATACCAGCAAGCTACTGCTGCCGCTGCCGCCGCCGCTGCCGGTATGC 1800
PAB1-S40.10     GAAATCAAATGAGATACCAGCAAGCTACTGCTGCCGCTGCCGCCGCCGCTGCCGGTATGC 1800
PAB1-S40.9      GAAATCAAATGAGATACCAGCAAGCTACTGCTGCCGCTGCCGCCGCCGCTGCCGGTATGC 1800
PAB1-S40.4      GAAATCAAATGAGATACCAGCAAGCTACTGCTGCCGCTGCCGCCGCCGCTGCCGGTATGC 1800
PAB1-S40.3      GAAATCAAATGAGATACCAGCAAGCTACTGCTGCCGCTGCCGCCGCCGCTGCCGGTATGC 1800
PAB1-S40.2      GAAATCAAATGAGATACCAGCAAGCTACTGCTGCCGCTGCCGCCGCCGCTGCCGGTATGC 1800
PAB1-S40.1      GAAATCAAATGAGATACCAGCAAGCTACTGCTGCCGCTGCCGCCGCCGCTGCCGGTATGC 1800
PAB1-wt         GAAATCAAATGAGATACCAGCAAGCTACTGCTGCCGCTGCCGCCGCCGCTGCCGGTATGC 1800
PAB1-S40.7      GAAATCAAATGAGATACCAGCAAGCTACTGCTGCCGCTGCCGCCGCCGCTGCCGGTATGC 1800
PAB1-S40.8      GAAATCAAATGAGATACCAGCAAGCTACTGCTGCCGCTGCCGCCGCCGCTGCCGGTATGC 1800
PAB1-S40.11     GAACTCAAATGAGATACCAGCAAGCTACTGCTGCCGCTGCCGCCGCCGCTGCCGGTATGC 1800
PAB1-E40.2      GAAATCAAATGAGATACCAGCAAGCTACTGCTGCCGCTGCCGCCGCCGCTGCCGGTATGC 1800
PAB1-S40.6      GAAATCAAATGAGATACCAGCAAGCTACTGCTGCCGCTGCCGCCGCCGCTGCCGGTATGC 1800
PAB1-S40.12     GAAATCAAATGAGATACCAGCAAGCTACTGCTGCCGCTGCCGCCGCCGCTGCCGGTATGC 1800
PAB1-E40.3      GAAATCAAATGAGATACCAGCAAGCTACTGCTGCCGCTGCCGCCGCCGCTGCCGGTATGC 1800
PAB1-E40.1      GAAATCAAATGAGATACCAGCAAGCTACTGCTGCCGCTGCCGCCGCCGCTGCCGGTATGC 1800
                *** ********************************************************

PAB1-S40.5      CAGGTCAATTCATGCCTCCAATGTTCTATGGTGTTATGCCACCAAGACGTGTTCCATTCA 1860
PAB1-S40.10     CAGGTCAATTCATGCCTCCAATGTTCTATGGTGTTATGCCACCAAGAGGTGTTCCATTCA 1860
PAB1-S40.9      CAGGTCAATTCATGCCTCCAATGTTCTATGGTGATATGCCACCAAGAGGTGTTCCATTCA 1860
PAB1-S40.4      CAGGTCAATTCATGCCTCCAATGTTCTATGGTGTTATGCCACCAAGAGGTGTTCCATTCA 1860
PAB1-S40.3      CAGGTCAATTCATGCCTCCAATGTTCTATGGTGTTATGCCACCAAGAGGTGTTCCATTCA 1860
PAB1-S40.2      CAGGTCAATTCATGCCTCCAATGTTCTATGGTGTTATGCCACCAAGAGGTGTTCCATTCA 1860
PAB1-S40.1      CAGGTCAATTCATGCCTCCAATGTTCTATGGTGTTATGCCACCAAGAGGTGTTCCATTCA 1860
PAB1-wt         CAGGTCAATTCATGCCTCCAATGTTCTATGGTGTTATGCCACCAAGAGGTGTTCCATTCA 1860
PAB1-S40.7      CAGGTCATTTCATGCCTCCAATGTTCTATGGTGTTATGCCACCAAGAGGTGTTCCATTCA 1860
PAB1-S40.8      CAGGTCAATTCATGCCTCCAATGTTCTATGGTGTTATGCCACCAAGAGGTGTTCCATTCA 1860
PAB1-S40.11     CAGGTCAATTCATGCCTCCAATGTTCTATGGTGTTATGCCACCAAGAGGTGTTCCATTCA 1860
PAB1-E40.2      CAGGTCAATTCATGCCTCCAATGTTCTATGGTGTTATGCCACCAAGAGGTGTTCCATTCA 1860
PAB1-S40.6      CAGGTCAATTCATGCCTCCAATGTTCTATGGTGTTATGCCACCAAGAGGTGTTCCATTCA 1860
PAB1-S40.12     CAGGTCAATTCATGCCTCCAATGTTCTATGGTGTTGTGCCACCAAGAGGTGTTCCATTCA 1860
PAB1-E40.3      CAGGTCAATTCATGCCTCCAATGTTCTATGGTGTTATGCCACCAAGAGGTGTTCCATTCA 1860
PAB1-E40.1      CAGGTCAATTCATGCCTCCAATGTTCTATGGTGTTATGCCACCAAGAGGTGTTCCATTCA 1860
                ******* ************************* * *********** ************

PAB1-S40.5      ACGGTCCAAACCCACAACAAATGAACCCAATGGGCGGTATGCCAAAGAACGGCATGCCAC 1920
PAB1-S40.10     ACGGTCCAAACCCACAACAAATGAACCGAATGGGCGGTATGCCAAAGAACGGCATGCCAC 1920
PAB1-S40.9      ACGATCCAAACCCACAACAAATGAACCCAATGGGCGGTATGCCAAAGAACGGCATGCCAC 1920
PAB1-S40.4      ACGGTCCAAACCCACAACAAATGAACCCAATGGGCGGTATGCCAAAGAACGGCATGCCAC 1920
PAB1-S40.3      ACGGTCCAAACCCACAACAAGTGAACCCAATGGGCGGTATGCCAAAGAACGGCATGCCAC 1920
PAB1-S40.2      ACGGTCCAAACCCACAACAAATGAACCCAATGGGCGGTATGCCAAAGAACGGCATGCCAC 1920
PAB1-S40.1      ACGGTCCAAACCCACAACAAATGAACCCAATGGGCGGTATGCCAAAGAACGGCATGCCAC 1920
PAB1-wt         ACGGTCCAAACCCACAACAAATGAACCCAATGGGCGGTATGCCAAAGAACGGCATGCCAC 1920
```

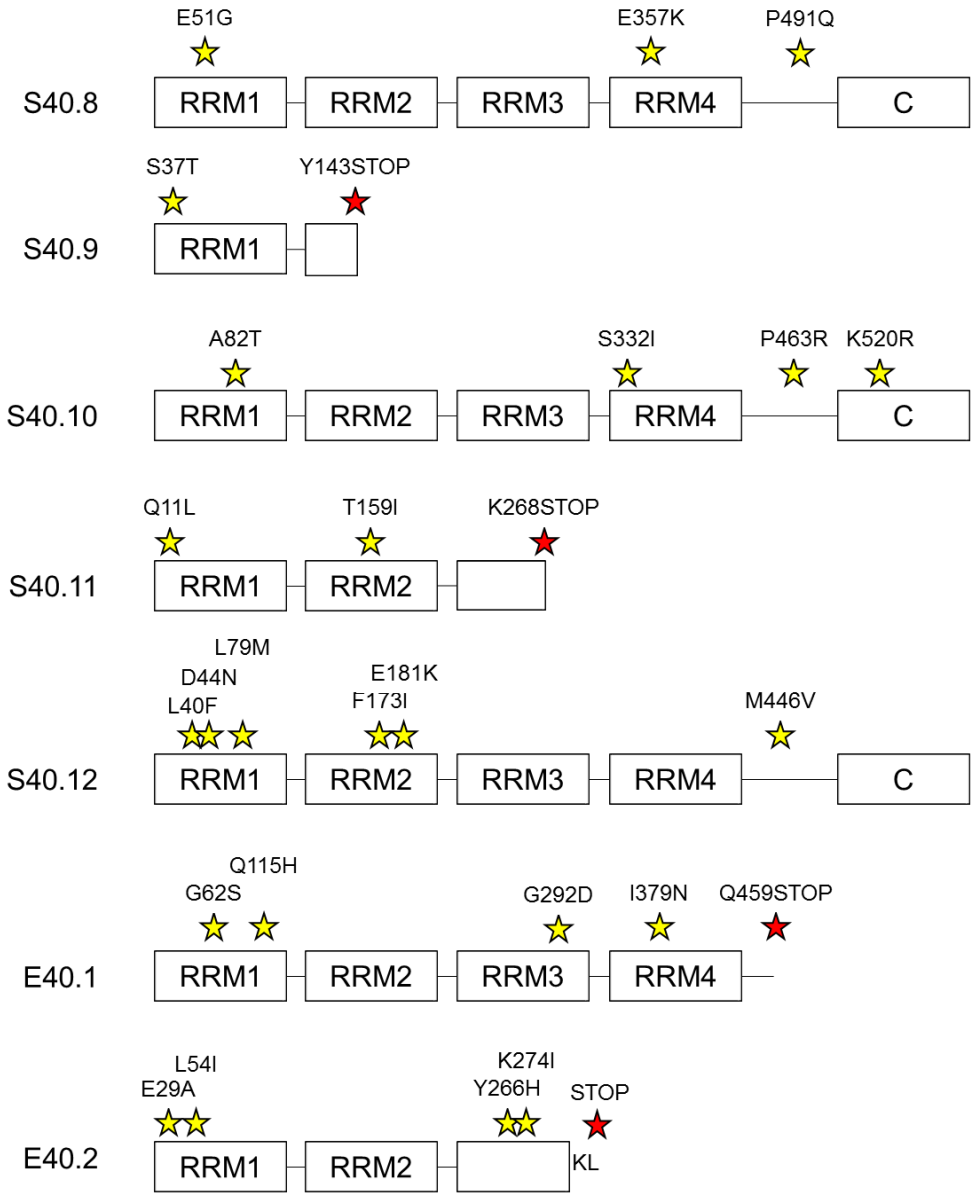
```
PAB1-S40.7      ACGGTCCAAACCCACAACAAATGAACCCAATGGGCGGTATGCCAAAGAACGGCATGCCAC 1920
PAB1-S40.8      ACGGTCCAAACCCACAACAAATGAACCCAATGGGCGGTATGCCAAAGAACGGCATGCCAC 1920
PAB1-S40.11     ACGGTCCAAACCCACAACAAATGAACCCAATGGGCGGTATGCCAAAGAACGGCATGCCAC 1920
PAB1-E40.2      ACGGTCCAAACCCACAACAAATGAACCCAATGGGCGGTATGCCAAAGAACGGCATGCCAC 1920
PAB1-S40.6      ACGGTCCAAACCCACAACAAATGAACCCAATGGGCGGTATGCCAAAGAACGGCATGCCAC 1920
PAB1-S40.12     ACGGTCCAAACCCACAACAAATGAACCCAATGGGCGGTATGCCAAAGAACGGCATGCCAC 1920
PAB1-E40.3      ACGGTCCAAACCCACAACAAATGAACCCAATGGGCGGTATGCCAAAGAACGGCATGCCAC 1920
PAB1-E40.1      ACGGTCCAAACCCATAACAAATGAACCCAATGGGCGGTATGCCAAAGAACGGCATGCCAC 1920
                *** ********** ***** ****** ********************************

PAB1-S40.5      CTCAATTTAGAAATGGTCCGGTTTACGGCGTCCCCCCACAAGGTGGTTTCCCAAGAAATG 1980
PAB1-S40.10     CTCAATTTAGAAATGGTCCGGTTTACGGCGTCCCCCCACAAGGTGGTTTCCCAAGAAATG 1980
PAB1-S40.9      CTCAATTTAGAAATGGTCCGGTTTACGGCGTCCCCCCACAAGGTGGTTTCCCAAGAAATG 1980
PAB1-S40.4      CTCAATTTAGAAATGGTCCGGTTTACGGCGTCCCCCCACAAGGTGGTTTCCCAAGAAATG 1980
PAB1-S40.3      CTCAATTTAGAAATGGTCCGGTTTACGGTGTCCCCCCCACTAGGTGGTTTCCCAAGAAATG 1980
PAB1-S40.2      CTCAATTTAGAAATGGTCCGGTTTACGGCGTCCCCCCACAAGGTGGTTTCCCAAGAAATG 1980
PAB1-S40.1      CTCAATTTAGAAATGGTCCGGTTTACGGCGTCCCCCCACAAGGTGGTTTCCCAAGAAATG 1980
PAB1-wt         CTCAATTTAGAAATGGTCCGGTTTACGGCGTCCCCCCACAAGGTGGTTTCCCAAGAAATG 1980
PAB1-S40.7      CTCAATTTAGAAATGGTCCGGTTTACGGCGTCCCCCCACAAGGTGGTTTCCCAAGAAATG 1980
PAB1-S40.8      CTCAATTTAGAAATGGTCCGGTTTACGGCGTCCCCCCACAAGGTGGTTTCCAAAGAAATG 1980
PAB1-S40.11     CTCAATTTAGAAATGGTCCGGTTTACGGCGTCCCCCCACAAGGTGGTTTCCCAAGAAATG 1980
PAB1-E40.2      CTCAATTTAGAAATGGTCCGGTTTACGGCGTCCCCCCACAAGGTGGTTTCCCAAGAAATG 1980
PAB1-S40.6      CTCAATTTAGAAATGGTCCGGTTTACGGCGTCCCCCCACAAGGTGGTTTCCCAAGAAATG 1980
PAB1-S40.12     CTCAATTTAGAAATGGTCCGGTTTACGGCGTCCCCCCACAAGGTGGTTTCCCAAGAAATG 1980
PAB1-E40.3      CTCAATTTAGAAATGGTCCGGTTTACGGCGTCCCCCCACAAGGTGGTTTCCCAAGAAATG 1980
PAB1-E40.1      CTCAATTTAGAAATGGTCCGGTTTACGGCGTCCCCCCACAAGGTGGTTTCCCAAGAAATG 1980
                **************************** ********** *********** ********

PAB1-S40.5      CCAACGATAACAACCAATTTTATCAACAAAAGCAAAGACAAGCTTTGGGTGAACAATTAT 2040
PAB1-S40.10     CCAACGATAACAACCAATTTTATCAACAAAAGCAAAGACAAGCTTTGGGTGAACAATTAT 2040
PAB1-S40.9      CCAACGATAACAACCAATTTTATCAACAAAAGCAAAGACAAGCTTTAGGTGAACAATTAT 2040
PAB1-S40.4      CCAACGATAACAACCAATTTTATCAACAAAAGCAAAGACAAGCTTTGGGTGAACAATTAT 2040
PAB1-S40.3      CCAAGGATAACAACCAATTTTATCAACAAAAGCAAAGACAAGCTTTGGGTGAACAATTAT 2040
PAB1-S40.2      CCAACGATAACAACCAATTTTATCAACAAAAGCAAAGACAAGCTTTGGGTGAACAATTAT 2040
PAB1-S40.1      CCAACGATAACAACCAATTTTATCAACAAAAGCAAAGACAAGCTTTGGGTGAACAATTAT 2040
PAB1-wt         CCAACGATAACAACCAATTTTATCAACAAAAGCAAAGACAAGCTTTGGGTGAACAATTAT 2040
PAB1-S40.7      CCAACGATAACAACCAATTTTATCAACAAAAGCAAAGACAAGCTTTGGGTGAACAATTAT 2040
PAB1-S40.8      CCAACGATAACAACCAATTTTATCAACAAAAGCAAAGACAAGCTTTGGGTGAACAATTAT 2040
PAB1-S40.11     CCAACGATAACAACCAATTTTATCAACAAAAGCAAAGACAAGCTTTGGGTGAACAATTAT 2040
PAB1-E40.2      CCAACGATAACAACCAATTTTATCAACAAAAGCAAAGACAAGCTTTGGGTGAACAATTAT 2040
PAB1-S40.6      CCAACGATAACAACCAATTTTATCAACAAAAGCAAAGACAAGCTTTGGGTGAACAATTAT 2040
PAB1-S40.12     CCAACGATAACAACCAATTTTATCAACAAAAGCAAAGACAAGCTTTGGGTGAACAATTAT 2040
PAB1-E40.3      CCAACGATAACAACCAATTTTATCAACAAAAGCAAAGACAAGCTTTGGGTGAACAATTAT 2040
PAB1-E40.1      CCAACGATAACAACCAATTTTATCAACAAAAGCAAAGACAAGCTTTGGGTGAACAATTAT 2040
                **** ****************************************** *************

PAB1-S40.5      ACAAGAAGGTTTCTGCTAAGACTTCAAATGCAGAAGCAGCTGGTAAAATTACTGGTATGA 2100
PAB1-S40.10     ACAAGAAGGTTTCTGCTAGGACTTCAAATGAAGAAGCAGCTGGTAAAATTACTGGTATGA 2100
PAB1-S40.9      ACAAGAAGGTTTCTGCTAAGACTTCAAATGAAGAAGCAGCTGGTAAAATTACTGGTATGA 2100
PAB1-S40.4      ACAAGAAGGTTTCTGCTAAGACTTCAAATGAAGAAGCAGCTGGTAAAATTACTGGTATGA 2100
PAB1-S40.3      ACAAGAAGGTTTCTGCTAAGACTTCAAATGAAGAAGCAGCTGGTAAAATTACTGGTATGA 2100
PAB1-S40.2      ACAAGAAGGTTTCTGCTAAGACTTCAAATGAAGAAGCAGCTGGTAAAATTACTGGTATGA 2100
PAB1-S40.1      ACAAGAAGGTTTCTGCTAAGACTTCAAATGAAGAAGCAGCTGGTAAAATTACTGGTATGA 2100
PAB1-wt         ACAAGAAGGTTTCTGCTAAGACTTCAAATGAAGAAGCAGCTGGTAAAATTACTGGTATGA 2100
PAB1-S40.7      ACAAGAAGGTTTCTGCTAAGACTTCAAATGAAGAAGCAGCTGGTAAAATTACTGGTATGA 2100
PAB1-S40.8      ACAAGAAGGTTTCTGCTAAGACTTCAAATGAAGAAGCAGCTGGTAAAATTACTGGTATGA 2100
PAB1-S40.11     ACAAGAAGGTTTCTGCTAAGACTTCAAATGAAGAAGCAGCTGGTAAAATTACTGGTATGA 2100
PAB1-E40.2      ACAAGAAGGTTTCTGCTAAGACTTCAAATGAAGAAGCAGCTGGTAAAATTACTGGTATGA 2100
PAB1-S40.6      ACAAGAAGGTTTCTGCAAAGACTTCAAATGAAGAAGCAGCTGGTAAAATTACTGGTATGA 2100
PAB1-S40.12     ACAAGAAGGTTTCTGCTAAGACTTCAAATGAAGAAGCAGCTGGTAAAATTACTGGTATGA 2100
PAB1-E40.3      ACAAGAAGGTTTCTGCTAAGACTTCAAATGAAGAAGCAGCTGGTAAAATTACTGGTATGA 2100
PAB1-E40.1      ACAAGAAGGTTTCTGCTAAGACTTCAAATGAAGAAGCAGCTGGTAAAATTACTGGTATGA 2100
                **************** * *********** ****************************

PAB1-S40.5      TTTTGGATTTGCCACCTCAAGAGGTCTTCCCATTGTTGGAAAGTGATGAATTGTTCGAAC 2160
PAB1-S40.10     TTTTGGATTTGCCACCTCAAGAGGTCTTCCCATTGTTGGAAAGTGATGAATTGTTCGAAC 2160
PAB1-S40.9      TTTTGGATTTGCCACCTCAAGAGGTCTTCCCATTGTTGGAAAGTGATGAATTGTTCGAAC 2160
PAB1-S40.4      TTTTGGATTTGCCACCTCAAGAGGTCTTCCCATTGTTGGAAAGTGATGAATTGTTCGAAC 2160
PAB1-S40.3      TTTTGGATTTGCCACCTCAAGAGGTCTTCCCATTGTTGGAAAGTGATGAATTGTTCGAAC 2160
PAB1-S40.2      TTTTGGATTTGCCACCTCAAGAGGTCTTCCCATTGTTGGAAAGTGATGAATTGTTCGAAC 2160
PAB1-S40.1      TTTTGGATTTGCCACCTCAAGAGGTCTTCCCATTGTTGGAAAGTGATGAATTGTTCGAAC 2160
PAB1-wt         TTTTGGATTTGCCACCTCAAGAGGTCTTCCCATTGTTGGAAAGTGATGAATTGTTCGAAC 2160
```

```
PAB1-S40.7      TTTTGGATTTGCCACCTCAAGAGGTCTTCCCATTGTTGGAAAGTGATGAATTGTTCGAAC 2160
PAB1-S40.8      TTTTGGATTTGCCACCTCAAGAGGTCTTCCCATTGTTGGAAAGTGATGAATTGTTCGAAC 2160
PAB1-S40.11     TTTTGGATTTGCCACCTCAAGAGGTCTTCCCATTGTTGGAAAGTGATGAATTGTTCGAAC 2160
PAB1-E40.2      TTCTGGATTTGCCACCTCAAGAGGTCTTCCCATTGTTGGAAAGTGATGAATTGTTCGAAC 2160
PAB1-S40.6      TTTTGGATTTGCCACCTCAAGAGGTCTTCCCATTGTTGGAAAGTGATGAATTGTTCGAAC 2160
PAB1-S40.12     TTTTGGATTTGCCACCTCAAGAGGTCTTCCCATTGTTGGAAAGTGATGAATTGTTCGAAC 2160
PAB1-E40.3      TTTTGGATTTGCCACCTCAAGAGGTCTTCCCATTGTTGGAAAGTGATGAATTGTTCGAAC 2160
PAB1-E40.1      TTTTGGATTTGCCACCTCAAGAGGTCTTCCCATTGTTGGAAAGTGATGAATTGTTCGAAC 2160
                ** ************************************************************

PAB1-S40.5      AACACTACAAAGAAGCTTCTGCTGCCTTTGAGTCTTTCAAAAAGGAGCAAGAACAACAAA 2220
PAB1-S40.10     AACACTACAAAGAAGCTTCTGCTGCCTATGAGTCTTTCAAAAAGGAGCAAGAACAACAAA 2220
PAB1-S40.9      AACACTACAAAGAAGCTTCTGCTGCCTATGAGTCTTTCAAAAAGGAGCAAGAACAGCAAA 2220
PAB1-S40.4      AACACTACAAAGAAGCTTCTGCTGCCTATGAGTCTTTCAAAAAGGAGCAAGAACAACAAA 2220
PAB1-S40.3      AACACTACAAAGAAGCTTCTGCTGCCTATGAGTCTTTCAAAAAGGAGCAAGAACAACAAA 2220
PAB1-S40.2      AACACTACAAAGAAGCTTCTGCTGCCTATGAGTCTTTCAAAAAGGAGCAAGAACAACAAA 2220
PAB1-S40.1      AACACTACAAAGAAGCTTCTGCTACCTATGAGTCTTTCAAAAAGGAGCAAGAACAACAAA 2220
PAB1-wt         AACACTACAAAGAAGCTTCTGCTGCCTATGAGTCTTTCAAAAAGGAGCAAGAACAACAAA 2220
PAB1-S40.7      AACACTACAAAGAAGCTTCTGCTGCCTATGAGTCTTTCAAAAAGGAGCAAGAACTACAAA 2220
PAB1-S40.8      AACACTACAAAGAAGCTTCTGCTGCCTATGAGTCTTTCAAAAAGGAGCAAGAACAACAAA 2220
PAB1-S40.11     AACACTACAAAGAAGCTTCTGCTGCCTATGAGTCTTTCAAAAAGGAGCAAGAACAACAAA 2220
PAB1-E40.2      AACACTACAAAGAAGCTTCTGCTGCCTATGAGTCTTTCAAAAAGGAGCAAGAACAACAAA 2220
PAB1-S40.6      AACACTACAAAGAAGCTTCTGCTGCCTATGAGTCTTTCAAAAAGGAGCAAGAACAACAAA 2220
PAB1-S40.12     AACACTACAAAGAAGCTTCTGCTGCCTATGAGTCTTTCAAAAAGGAGCAAGAACAACAAA 2220
PAB1-E40.3      AACACTACAAAGAAGCTTCTGCTGCCTATGAGTCTTTCAAAAAGGAGCAAGAACAACAAA 2220
PAB1-E40.1      AACACTACAAAGAAGCTTCTGCTGCCTATGAGTCTTTCAAAAAGGAGCAAGAACAACAAA 2220
                ********************** *** ************************  ****

PAB1-S40.5      CTGAGCAAGCTTAA 2333
PAB1-S40.10     CTGAGCAAGCTTAA 2333
PAB1-S40.9      CTGAGCAAGCTTAA 2333
PAB1-S40.4      CTGAGCAAGCTTAA 2333
PAB1-S40.3      CTGAGCAAGCTTAA 2333
PAB1-S40.2      CTGAGCAAGCTTAA 2333
PAB1-S40.1      CTGAGCAAGCTTAA 2333
PAB1-wt         CTGAGCAAGCTTAA 2333
PAB1-S40.7      CTGAGCAAGCTTAA 2333
PAB1-S40.8      CTGAGCAAGCTTAA 2333
PAB1-S40.11     CTGAGCAAGCTTAA 2333
PAB1-E40.2      CTGAGCAAGCTTAA 2333
PAB1-S40.6      CTGAGCAAGCTTAA 2333
PAB1-S40.12     CTGAGCAAGCTTAA 2333
PAB1-E40.3      CTGAGCAAGCTTAA 2333
PAB1-E40.1      CTGAGCAAGCTTAA 2333
                **************
```

Supplementary Fig.S4 List of Pab1 mutated proteins. Schematic representation of the proteins encoded by the *PAB1* variants isolated from the screening. Yellow stars: missense mutations. Red stars: premature stop codons. RRM: RNA recognition motifs. P: P-domain. C: C-domain.



Supplementary Fig.S4 Continue



Discovering new functions associated to Pab1 domains

Results described in the previous chapter confirmed that Pab1 is a powerful candidate to evoke complex phenotypes with improved traits. Interestingly, some of the isolated Pab1 variants that confer increased thermotolerance are truncated at different levels (because of premature stop codons). These data suggested that specific domains of the protein might be responsible for this phenotypic improvement and, thus, rationally exploitable as parts or eventually developed as "biobricks" (see Introduction) for synthetic biology. However, for this purpose a complete knowledge of their intracellular role is required. Despite the availability of many information about domains' interactions and functions, their contribution in the Pab1 recruitment within stress granules (aggregates of mRNPs stalled in the process of translation initiation) was never investigated. The results of this characterization are described in the following chapter.

Chapter 4

The recruitment of the *Saccharomyces cerevisiae* poly(A) binding protein into stress granules: new insights into the contribution of the different protein domains

Published in

FEMS YEAST RESEARCH

ISSN:1567-1356 vol. 17 (6)

Brambilla Marco ‡, Martani Francesca ‡ and Branduardi Paola. (2017)

‡These authors contributed equally to the work

DOI:10.1093/femsyr/fox059.

Introduction

Poly(A) binding proteins (generically termed PABPs) are ubiquitous and conserved among Eukarya, where they are involved in almost all mRNAs' post-transcriptional events (Eliseeva *et al.* 2013; Smith *et al.* 2014). Multiple PABPs can be found in higher eukaryotes: in humans and in *Arabidopsis thaliana* there are up to five and eight PABPs, respectively, while single celled eukaryotes have just one (Goss and Kleiman 2013). Apart from some specialized PABPs (Eliseeva *et al.* 2013; Smith *et al.* 2014), their structure is typically highly conserved. These modular proteins are composed of four RNA Recognition Motifs (RRM) for poly(A) tail binding that are linked to a carboxy-terminal domain (C domain) through a flexible linker rich in proline (for this reason named P domain), methionine, glycine, glutamine and asparagine (Yao *et al.* 2007; Richardson *et al.* 2012). The structure of RRMs (~90-100 residues) is typically highly conserved from yeast to higher eukaryotes: these globular domains consist of four antiparallel β-sheets flanked by two α-helices, with two highly conserved sequences (RNP-1 and RNP-2) in the two central β-sheets that are involved in poly(A) tail binding (Eliseeva *et al.* 2013; Smith *et al.* 2014). Differently from RRMs, the P domain is intrinsically disordered and it is significantly divergent in length across different species (Riback *et al.* 2017). In particular, the alignment of the P domain among closely related species revealed that many sites are perfectly conserved, while medium- and long-chain aliphatic residues often substitute from each other (Riback *et al.* 2017). Finally, the C-terminal domain is globular with approximately 75 residues that fold into four or five α-helices depending on the species (Smith *et al.* 2014).

The PABP of *Saccharomyces cerevisiae* (named Pab1) has been extensively studied over the last decades to describe its role in controlling mRNA

metabolism. It has been demonstrated that Pab1 binding to the poly(A) tail protects transcripts from degradation by inhibiting the deadenylase activity of the Ccr4/Pop2/Not complex (Yao *et al*. 2007; Zhang *et al*. 2013). On the contrary, Pab1 stimulates the activity of the minor deadenylation complex Pan2/Pan3 by recruiting it, through its C-terminal domain, to the poly(A) tail of mRNAs (Simón and Séraphin 2007; Wahle and Winkler 2013). The dissociation of Pab1 from the poly(A) tail, which is responsible for Ccr4 deadenylation, is induced by protein self-association, a process occurring through the formation of circular species and/or oligomers (Yao *et al.* 2007). The circularization of Pab1 is likely accomplished by the interaction of the RRM1 with the RRM4 and the P domain, whereas the structure of Pab1 oligomers is less clear (Yao *et al.* 2007). Furthermore, Pab1 is implicated in the control of translation initiation and termination. In fact, by binding eIF4G through the RRM1 and RRM2 domains, it allows the formation of a closed-loop structure of the mRNA that stimulates translation initiation (Richardson *et al.* 2012; Melamed *et al.* 2015). On the contrary, Pab1 negatively regulates translation termination by interacting, through its P and C domains, with the releasing factor eRF3 (Sup35); such interaction allows a basal level of translational readthrough (Roque *et al.* 2015). Finally, Pab1 is involved in mRNA export from nucleus to cytoplasm through the Mex67/mRNA-dependent pathway. To exert this function, Pab1 enters into the nucleus through the interaction between its Nuclear Localization Signal (NLS), localized in the RRM4 domain, and the importin Kap108/Sxm1 (Brune *et al.* 2005; Richardson *et al.* 2012). A second route through which Pab1 can be exported to the cytoplasm occurs through the Xpo1/Crm1-dependent pathway (Brune *et al.* 2005; Dunn *et al.* 2005).

Pab1 has also an important role in promoting the assembly of stress granules, of which it is a component but is not indispensable (Swisher and Parker 2010). Stress granules are cytoplasmic aggregates of untranslated mRNPs that

form when translation initiation is impaired, such as when cells encounter stressful conditions (Buchan and Parker 2009; Protter and Parker 2016). They contain, among others, factors involved in translation initiation, such as the 40S ribosomal subunit and eIF4G; therefore, they are thought to be sites for mRNA translation re-initiation (Buchan and Parker 2009). However, depending on the stress, their composition can be slightly variable (Buchan and Parker 2009; Grousl *et al.* 2009; Buchan *et al.* 2011; Giménez-Barcons and Díez 2011; Decker and Parker 2012). When not translated, mRNAs can be cleared by vacuolar autophagy together with the entire stress granule in which they are stored, or can be exchanged with processing bodies (P-bodies) where they are principally degraded (Buchan and Parker 2009; Buchan *et al.* 2013; Walters and Parker 2015; Protter and Parker 2016). However, mRNAs in P-bodies can also return to stress granules and/or to translation (Buchan and Parker 2009; Buchan *et al.* 2013; Walters and Parker 2015; Protter and Parker 2016). It has been proposed that Pab1 role in stress granules formation might be to promote the transition of polyadenylated mRNAs from P-bodies to stress granules. In fact, *PAB1* deletion in the bypass suppressor *spb2Δ* strain reduced the number of stress granules and increased the number and P-bodies. However, being Pab1 an essential protein, the need for suppressors limits the possibility to accurately depict its role in stress granules formation (Swisher and Parker 2010).

As described above, the domains of Pab1 are responsible for its multiple functions because of their specific interactions with different proteins involved in the regulation of mRNA metabolism. However, to the best of our knowledge little is known about their function in Pab1 recruitment within stress granules. In this work, we aimed to understand the contribution of each Pab1 domain in protein recruitment within stress granules. For this purpose, we created *S. cerevisiae* strains containing both the endogenous Pab1 tagged with mCherry, as a marker of stress granules, and different GFP-tagged versions of the protein,

lacking one or more domains. The intracellular distribution and the propensity of Pab1 variants to localize within stress granules, evoked by glucose deprivation or heat shock, were then determined by fluorescent microscopy observations.

Results

Contribution of RRM domains in the intracellular localization of Pab1

To examine the contribution of protein domains in the intracellular distribution of Pab1 and its association within stress granules, we analyzed the *in vivo* localization of Pab1 mutants in which the domains, starting from the RRM1, where sequentially removed (Fig.1). For this purpose, these protein versions were tagged with GFP and their localization was juxtaposed and compared to that of the endogenous Pab1 tagged with mCherry, used as a marker of stress granules (Buchan *et al.* 2010; Swisher and Parker 2010; Wheeler *et al.* 2016). To ensure high expression levels, the constitutive *TPI* promoter was chosen for the over-expression of each variant.

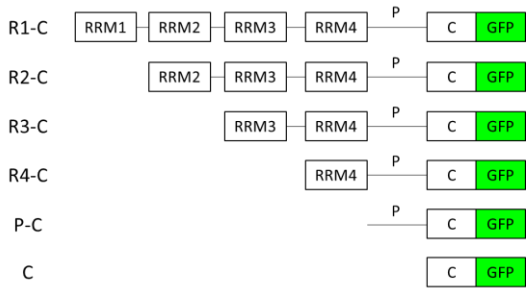


Fig.1 List of GFP-tagged Pab1 variants. Schematic representation of Pab1 variants R1-C, R2-C, R3-C, R4-C, P-C and C. RRM: RNA recognition motifs, P: P-domain, C: C-domain, GFP: green fluorescent protein. The DNA constructs encoding for each Pab1 variant were constitutively overexpressed with the *TPI* promoter.

First, exponentially growing cells were analyzed to detect whether the absence of one or more domains could influence Pab1 localization under unstressed conditions. As show in Fig.2, the full-length R1-C version, corresponding to the wild type protein, overlapped completely with the

endogenous Pab1 and localized predominantly in the cytoplasm as diffused protein, while almost no signal could be detected in the nucleus (see Table S1). On the contrary, cells carrying the N-truncated versions displayed green fluorescence also in the nucleus; remarkably, this phenomenon was more pronounced for the R3-C variant (Fig.2; Table S1).

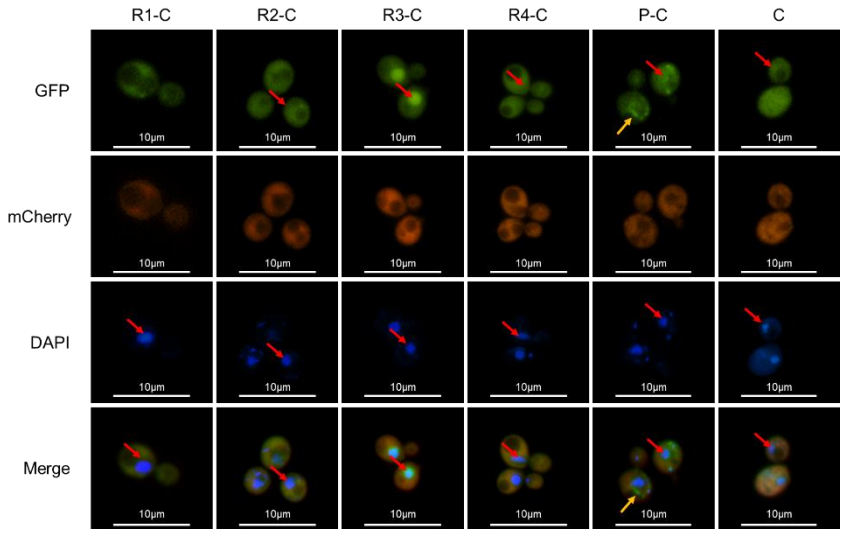


Fig.2 Effect of the removal of RRM on the localization of Pab1 in the absence of stress. Intracellular localization of Pab1-mCherry and GFP-tagged Pab1 variants shown in Fig.1. Cells were grown until the exponential phase of growth (OD_{660nm} = 0.3–0.6) and then stained with DAPI to a final concentration of 1 μg/mL. Yellow arrows indicate aggregates of the corresponding Pab1-GFP variant; red arrows indicate nuclei.

As previously reported, stress granules resulting from different type of stress display a different assembly and composition (Buchan *et al.* 2011). Therefore, we evoked stress granules by glucose starvation or heat shock to reinforce general observations and underline possible peculiarities. By scoring protein aggregates related to the wild type Pab1 (red fluorescent dots) and the mutant versions (green fluorescent dots), we evaluated: *i)* the percentage of cells with green fluorescent dots; *ii)* the mutant/wild type dots ratio, as an index of the

propensity of synthetic versions to aggregate similarly as the endogenous Pab1; *iii)* the overlap between green and red fluorescent dots to evaluate the localization of Pab1 variants within stress granules.

Under stressful conditions, cells overexpressing the full-length protein displayed both red and green fluorescent foci that completely overlaid, indicating that the R1-C protein resides in stress granules (Fig.3A-B; Fig.4A; Table S2). Similarly, the cytoplasmic fraction of all the protein variants carrying at least one RRM almost fully co-localized within stress granules (Fig.3A-B; Fig.4A; Table S2). However, in glucose starved cells, the propensity to aggregate into stress granules generally decreased by lowering the number of RRM domains: in fact, R2-C, R3-C and R4-C variants displayed fewer protein aggregates compared to the endogenous stress granules, as observed by the mutant/wild type dots ratio of 0.90, 0.67 and 0.68, respectively (Fig.3A; Fig.4B; Table S3). A similar but less pronounced effect was observed in heat shocked cells (Fig.3B; Fig.4B; Table S3). Although we cannot exclude a direct correlation between mutant/wild type dots ratio and mutant protein levels, in stressful conditions the R4-C showed the highest expression level (Fig.S1 and Table S4), but a low mutant/wild type dots ratio. Therefore, it is likely that, at least for these variants, the different expression levels do not influence the propensity of the synthetic versions to aggregate into stress granules.

Both in the absence and in the presence of stress, in more than 90% of the cells the P-C variant displayed protein aggregates that did not co-localize with stress granules (Fig.2; Fig.3; Fig.4A Table S2). This behavior was likely related to the P domain because no aggregates were detected when the C domain was present alone, independently from a possible imposed stressful condition (Fig.3).

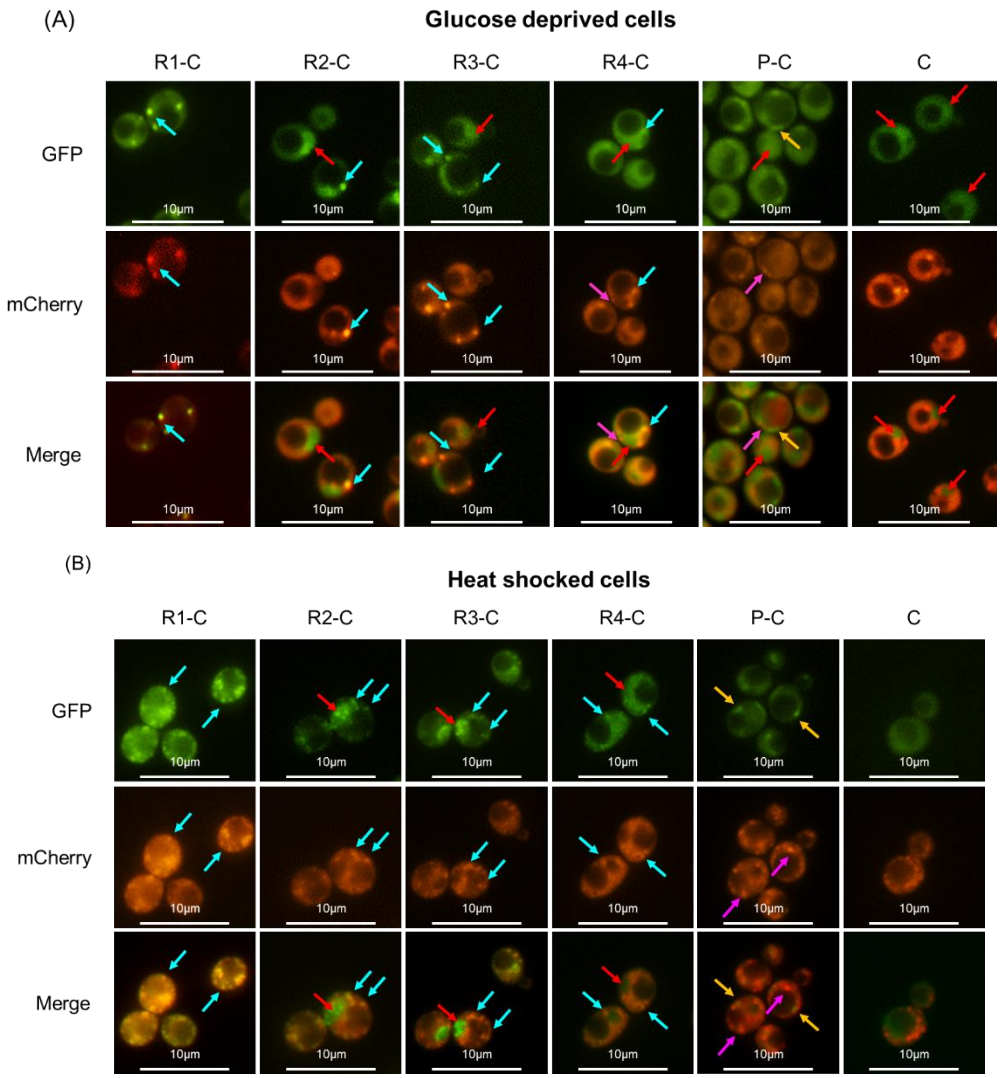


Fig.3 Effect of the removal of RRM on the localization of Pab1 in presence of stress. Intracellular localization of Pab1-mCherry and GFP-tagged Pab1 variants shown in Fig.1. Cells were grown until the exponential phase of growth (OD_{660nm} = 0.3–0.6) and then (A) glucose starved for 15min (see above) or (B) heat shocked for 50 min. Light blue arrows indicate co-localized green and red foci; fuchsia arrows indicate Pab1-mCherry stress granules without a co-localizing counterpart in the Pab1-GFP variant; yellow arrows indicate aggregates of the corresponding Pab1-GFP variant; red arrows indicate nuclei.

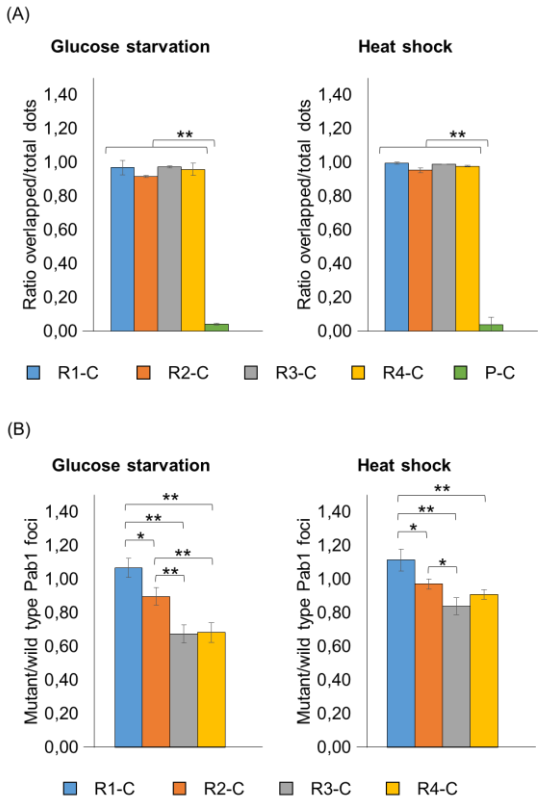


Fig.4 Analysis of the recruitment into stress granules of Pab1 variants carrying the C-terminal domain. Strains carrying the GFP-tagged Pab1 variants shown in Fig.1 were grown until the exponential phase of growth (OD_{660nm} = 0.3–0.6) and then subjected to glucose starvation for 15 min or heat shock for 50 min. Data in panel (A) show the ratio between the overlapped green-red fluorescent dots and the total number of dots. Data in panel (B) show the ratio between the number of mutant (green) and the number of wild type (red) foci. The variant C was not included in the histograms because no aggregates were detected in the corresponding strain, while the variant P-C was not included in panel (B) because the corresponding aggregates did not co-localize with endogenous stress granules. Values indicate the mean and the standard deviation of three independent experiments. Significance was calculated with two-tailed student t-test ($*P \leq 0.05$; $**P \leq 0.01$; see Table S6).

These data suggest that the presence of at least one RRM is indispensable to allow a specific recruitment of Pab1 within stress granules, and that the presence of all four RRM domains seems to be required to direct its efficient inclusion into these aggregates. Furthermore, in the absence of the four RRMs, the C terminal domain seems not directly involved in protein localization into stress granules; however, its presence might influence the association of Pab1 within stress granules.

Cellular localization of Pab1 variants lacking the C-terminal domain

To evaluate the contribution of the C domain in Pab1 localization and recruitment within stress granules, we analyzed the intracellular localization of the previously described versions lacking the C-terminus (Fig.5). The P domain alone was not detected by western blot analysis indicating that it may not be translated or it may undergo degradation (data not shown). In the absence of stress, the localization of all these variants was as the one observed in the presence of the C domain (see Fig. S2A).

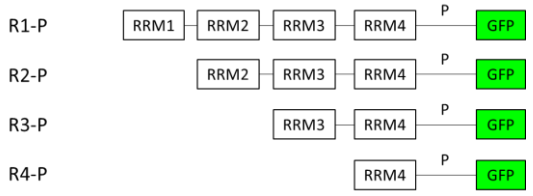


Fig.5 List of GFP-tagged Pab1 variants. Schematic representation of Pab1 variants R1-P, R2-P, R3-P and R4-P. RRM: RNA recognition motifs, P: P-domain, C: C-domain, GFP: green fluorescent protein. The DNA constructs encoding for each Pab1 variant were constitutively overexpressed with the *TPI* promoter.

In glucose starved or heat shocked cells, the green fluorescence associated to the R1-P, R2-P, R3-P and R4-P variants almost completely overlapped with the red fluorescent aggregates of the wild type Pab1 (Fig.6A; Fig.S2B-S2C; Table S2). Thus, the protein could localize into stress granules even in the absence of the C domain. However, the mutant/wild type dots ratio of these variants was slightly higher than the one measured in cells containing the R2-C, R3-C and R4-C versions (Fig.6B; Table S3). These results suggest that the C domain may have a negative effect on the Pab1 recruitment within stress granules in the absence of one or more RRM domains, even though we cannot exclude that the observation could be affected by different expression levels of the variants under stressful conditions (see Fig.S1 and Table S4).

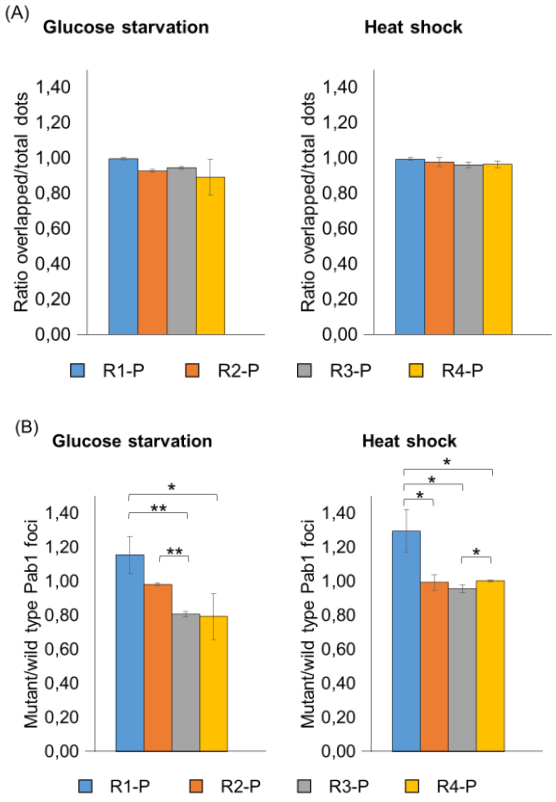


Fig.6 Analysis of the recruitment into stress granules of Pab1 variants lacking the C-terminal domain. Strains carrying the GFP-tagged Pab1 variants shown in Fig.5 were grown until the exponential phase of growth (OD_{660nm} = 0.3–0.6) and then subjected to glucose starvation for 15 min or heat shock for 50 min. Data in panel (A) (see above) show the ratio between the overlapped green-red fluorescent dots and the total number of dots. Data in panel (B) show the ratio between the number of mutant (green) and the number of wild type (red) foci. Values indicate the mean and the standard deviation of three independent experiments. Significance was calculated with two-tailed student t-test ($*P \leq 0.05$; $**P \leq 0.01$; see Table S6).

Cellular localization of Pab1 variants lacking the P and C domains

The P domain has a prominent role in Pab1 self-aggregation, a process responsible for favoring mRNA deadenylation by Ccr4 (Yao *et al.* 2007; Richardson *et al.* 2012) and that may affect the recruitment of the protein into stress granules. Therefore, the P domain was removed together with the C domain leading to protein versions composed of only one or more RRM domains (R1-R4, R2-R4, R3-R4 and R4, Fig.7) that were studied for their ability to localize into stress granules.

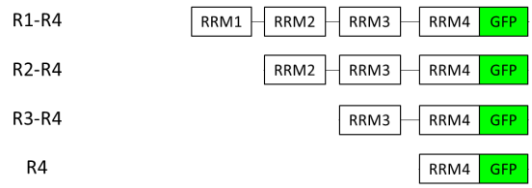


Fig.7 List of GFP-tagged Pab1 variants. Schematic representation of Pab1 variants R1-R4, R2-R4, R3- R4 and R4. RRM: RNA recognition motifs, P: P-domain, C: C-domain, GFP: green fluorescent protein. The DNA constructs encoding for each Pab1 variant were constitutively overexpressed with the *TPI* promoter.

In the absence of stress, the R1-R4 variant localized predominantly in the cytoplasm, while the removal of the RRM1 domain led to a Pab1 distribution in both the nucleus and the cytoplasm (Fig.S3A; Table S1), as previously observed.

Interestingly, the nuclear localization of the R3-R4 variant was not as severe as that observed in R3-C and R3-P (Fig.S3A; Fig.2; Fig.S2A).

When cells were glucose starved or heat shocked, proteins co-localization within stress granules was not significantly affected by the removal of the P and the C domains for all the four variants (Fig.8A; Fig. S3B-S3C; Table S2). Remarkably, the R4 strain showed a very low ratio of mutant/wild type dots (0.04 in glucose deprived and 0.12 in heat shocked cells) (Fig.8B; Table S3). Moreover, a poor number of cells displaying R4 associated-aggregates were observed compared to the others (13.2% in glucose deprived and 44.7% in heat shocked cells, Table S5), indicating that the RRM4 domain alone is not particularly prone to be delivered into stress granules. The propensity to be recruited into stress granules improved with the further addition of one and two RRMs: indeed, when cells were glucose starved, the ratio of mutant/wild type dots for the R3-R4 and R2-R4 variants increased up to 0.48 and to 0.59 respectively (Fig.8B; Table S3). Interestingly, when cells were subjected to heat shock, this parameter resulted 0.30 for R3-R4 and 0.95 for R2-R4 variants (Fig.8B; Table S3), thus indicating that in this condition the presence of a third RRM significantly enhanced recruitment within stress granules. Finally, the R1-R4 version fully associated to stress granules independently from the triggering stressful conditions, since cells displayed almost the same number of green and red fluorescent foci (Fig.8B; Table S3).

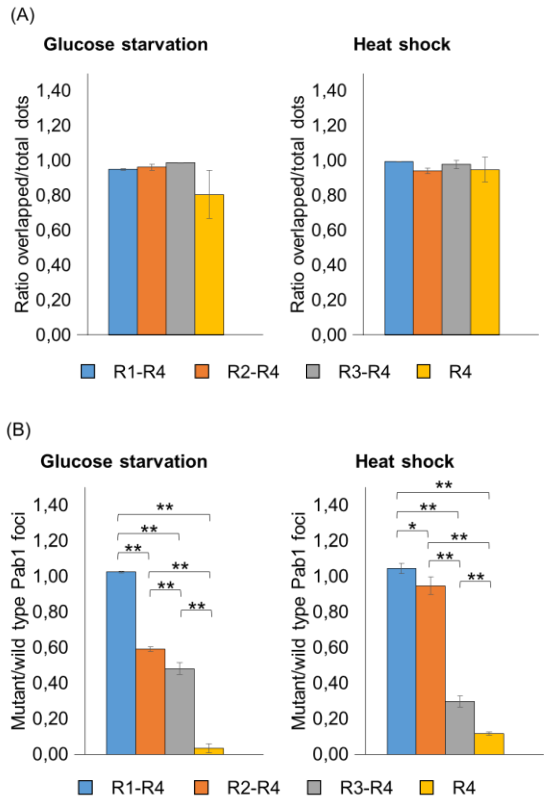


Fig.8 Analysis of the recruitment into stress granules of Pab1 variants lacking the P and the C domains. Strains carrying the GFP-tagged Pab1 variants shown in Fig.7 were grown until the exponential phase of growth (OD_{660nm} = 0.3–0.6) and then subjected to glucose starvation for 15 min or heat shock for 50 min. Data in panel (A) show the ratio between the overlapped green-red fluorescent dots and the total number of dots. Data in panel (B) show the ratio between the number of mutant (green) and the number of wild type (red) foci. Values indicate the mean and the standard deviation of three independent experiments. Significance was calculated with two-tailed student t-test ($*P \leq 0.05$; $**P \leq 0.01$; see Table S6).

Therefore, a synthetic version lacking the P and the C domain necessarily requires all the four RRMs to be efficiently recruited within stress granules. Furthermore, this result points out that the P domain is not strictly required for the inclusion of Pab1 into stress granules. However, the ratio of mutant/wild type dots for R2-P, R3-P and R4-P was consistently higher compared to the

corresponding variants lacking the P domain (Fig.6B-8B; Table S3), indicating that its presence strengthens the recruitment of the N-truncated variants into stress granules.

Discussions

The *S. cerevisiae* poly(A) binding protein Pab1 is one of the major actors in mRNA metabolism. By binding poly(A) tails, this protein can protect mRNAs from degradation (Simón and Séraphin 2007) and creates a scaffold for recruiting partners involved in mRNA polyadenylation (Zhao *et al.*, 1999), nuclear export (Brune *et al.* 2005; Dunn *et al.* 2005), deadenylation (Simón and Séraphin 2007; Yao *et al.* 2007; Parker 2012; Richardson *et al.* 2012; Wahle and Winkler 2013), translation initiation (Wyers *et al.* 2000) and termination (Richardson *et al.* 2012; Roque *et al.* 2015).

Despite several studies having been devoted to associate specific functions of this protein to its single domains, to the best of our knowledge their role in determining Pab1 association within stress granules is still unclear. Thus, in this work we sought to elucidate this point by observing the localization of several Pab1 GFP-tagged versions, lacking one or more domains, in stressful conditions known to evoke stress granules: glucose deprivation and heat shock. The endogenous Pab1 protein tagged with mCherry was used as a marker of these mRNPs aggregates (Buchan *et al.* 2010; Swisher and Parker 2010; Wheeler *et al.* 2016).

We observed that the number of RRM domains influence Pab1 distribution between the nucleus and the cytoplasm. The RRM1 domain is implicated in the cytosolic localization of Pab1, probably because it contains the amino-terminal Nuclear Export Sequence (NES) that has a predominant role in exporting Pab1 from the nucleus by Xpo1/Crm1 (Brune *et al.* 2005; Dunn *et al.* 2005).

The RRM1 domain contains the amino-terminal Nuclear Export Sequence (NES), which has a role in the export of Pab1 from the nucleus

mediated by Xpo1/Crm1 (Brune *et al.* 2005; Dunn *et al.* 2005). Therefore, variants carrying the RRM1 resulted diffused in the cytoplasm, while the dual cytosolic/nuclear distribution of the variants lacking this domain might be related to a compensation between the two minor pathways responsible for Pab1 import by Kap108/Sxm1 and export by Mex67-mRNA (Brune *et al.* 2005), whose effects could be accentuated in the absence of NES. Among all, the R3-C and R3-P variants showed the most intense nuclear localization, which was quite surprising since the Nuclear Localization Signal (NLS) resides in the RRM4; we should have therefore expected to find a marked nuclear localization of all the versions carrying this domain but lacking the RRM1. In addition, we observed a less intense localization of R3-R4 and R4-P variants in the nucleus. Hence, these results suggest that the presence of both the RRM3 and the P domain favors the nuclear localization or impedes the export to the cytoplasm.

We found out that the removal of one, two or three RRM domains did not affect Pab1 co-localization within stress granules, but it reduced the propensity of the protein to be recruited within them; this effect was particularly evident in the absence of two and three RRM domains. Therefore, the efficient Pab1 association to stress granules correlates with the number of RNA recognition motifs, possibly because the presence of more RRMs enables higher protein-protein interactions that boost mRNPs formation and, subsequently, the assembly of higher ultra-structures (Wheeler *et al.* 2016). However, we cannot exclude that the presence of a single RRM1, RRM2 or RRM3, or diverse combinations of them, might behave differently because of the specific interactions in which they are involved. In fact, several post-translational modifications such as seven methyl-glutamate, five acetyl-lysine and one phosphoserine (Low *et al.* 2014) are present along Pab1 sequence and may impact on many molecular aspects, among which the delivery of the protein into stress granules (Protter and Parker 2016).

Moreover, we discovered that the sole presence of RNA recognition motifs seems to be sufficient to allow Pab1 association within stress granules upon stress, since the variant carrying all the four RRMs fully co-localized with native stress granules even in the absence of P and C domains. Accordingly, it has been recently demonstrated that in mild stress conditions Pab1 molecules undergo phase separation even in the absence of the low-complexity P domain (Riback *et al.* 2017). Furthermore, in the absence of all RRM domains, the P-C and C domains alone were not recruited into stress granules, thus supporting the role of RRMs in determining the specific localization of Pab1 within these aggregates.

Although the recruitment of Pab1 within stress granules can occur independently from the P domain, in the absence of at least one RRM its presence favors Pab1 association to stress granules. Therefore, we suggest that this domain may strengthen the aggregation and, consequently, the delivery of Pab1 into stress granules, thanks to its intrinsically disordered nature (Riback *et al.* 2017) that may promote non-specific hydrophobic interactions. Indeed, variants without RRM domains formed irregularly shaped aggregates even in the absence of stress.

Interestingly, Pab1 variants lacking at least one RRM displayed a slight better recruitment within stress granules in the absence, rather than in the presence, of the C-terminal domain, even though we cannot exclude that small difference in expression levels might have influenced this result. Thus, the C domain might have a negative effect on Pab1 localization within stress granules. In particular, we hypothesize that the absence of the C domain might increase the P domain exposure, thus enhancing its intrinsically disordered nature that, in turn, will increase redundant interactions. A second possible explanation might reside in a reduced binding of Pab1 with eRF3 (involved in translation termination), Pbp1 and the PAN complex (both involved in deadenylation), whose activities

are dependent on their interaction with the C domain (Richardson *et al.* 2012; Roque *et al.* 2015). If confirmed, this last hypothesis would open new questions about the role of these proteins in stress granule formation.

Finally, Pab1 localization within stress granules evoked by glucose starvation or heat shock was similarly affected by the removal of modules, but in some cases, the extent of the effect was different depending on the stress. This might be related to a diverse composition of the resulting granules (Buchan *et al.* 2008; Grousl *et al.* 2009; Buchan *et al.* 2011; Grousl *et al.* 2013) and/or to different mechanisms of assembly. In fact, while glucose starvation is suggested to form stress granules in conjunction with pre-existing P-bodies, heat shock is described to promote assembly independently from them (Buchan *et al.* 2011). Remarkably, it has been recently found that the lysine acetyltransferase NuA4 complex, which is known to interact with Pab1, is required for regulating stress granules formation in glucose deprived cells but not in heat shocked ones (Rollins *et al.* 2017).

It is important to underline that Pab1 can undergo a process of self-association, which is dependent on the presence of the RRM1 and the P domain (Yao *et al.* 2007). Although this phenomenon has been demonstrated to influence Ccr4/Pop2/Not deadenylation (Yao *et al.* 2007), little is known about its role in Pab1 recruitment into stress granules. Therefore, we cannot exclude that self-association might affect the delivery into stress granules of Pab1 variants carrying or lacking the RRM1 or the P domain. In addition, in our experimental setting, the contemporary expression of the endogenous and the synthetic *PAB1* versions may result in cross-oligomerization among the corresponding proteins, thus possibly affecting their recruitment into stress granules. This aspect could be better elucidated by using different *pab1Δ* suppressor strains, such as *pab1Δ spb2Δ* and *pab1Δ pat1-2* (Swisher and Parker 2010), coupled with other markers of stress granules, such as Pbp1.

In conclusions, in this work we demonstrated that Pab1 domains differently contribute in the recruitment of Pab1 into stress granules. These novel findings suggest that the modularity of Pab1 may be an interesting candidate for module-based engineering, in which domains replacement or insertion can be rationally designed for the development of synthetic proteins. Surely, further studies including the modeling of Pab1 modules and their functions can be of help for understanding the dynamic equilibrium of the wild type protein as well as for tailoring novel constructs.

Materials and Methods

Strains and plasmids construction

The *S. cerevisiae* genetic background used in this study was CEN.PK113-11C (*MATa; MAL2-8c; SUC2; ura3-52; his3Δ1*). The *S. cerevisiae* strains created and used in this study are listed in Table 1.

Table.1 List of strains used in this study

Strains	Relevant genotype	Source
CEN.PK113-11C	*MATa; MAL2-8c; SUC2; ura3-52; his3Δ1* (reference strain)	A
PAB1-mCherry	*CEN.PK113-11C; PAB1::PAB1-mCherry-HIS3*	B
R1-C	*PAB1-mCherry [pYX012(-ATG)RRM1-C-GFP]*	B
R2-C	*PAB1-mCherry [pYX012(-ATG)RRM2-C-GFP]*	B
R3-C	*PAB1-mCherry [pYX012(-ATG)RRM3-C-GFP]*	B
R4-C	*PAB1-mCherry [pYX012(-ATG)RRM4-C-GFP]*	B
P-C	*PAB1-mCherry [pYX012(-ATG)P-C-GFP]*	B
C	*PAB1-mCherry [pYX012(-ATG)C-GFP]*	B
R1-P	*PAB1-mCherry [pYX012(-ATG)RRM1-P-GFP]*	B
R2-P	*PAB1-mCherry [pYX012(-ATG)RRM2-P-GFP]*	B
R3-P	*PAB1-mCherry [pYX012(-ATG)RRM3-P-GFP]*	B
R4-P	*PAB1-mCherry [pYX012(-ATG)RRM4-P-GFP]*	B
P	*PAB1-mCherry [pYX012(-ATG)P-GFP]*	B
R1-R4	*PAB1-mCherry [pYX012(-ATG)RRM1-RRM4-GFP]*	B
R2-R4	*PAB1-mCherry [pYX012(-ATG)RRM2-RRM4-GFP]*	B
R3-R4	*PAB1-mCherry [pYX012(-ATG)RRM3-RRM4-GFP]*	B
R4	*PAB1-mCherry [pYX012(-ATG)RRM4-GFP]*	B

A) Kindly provided by Dr. P. Kotter (Institute of Microbiology, Johann Wolfang Goethe-University, Frankfurt, Germany). B) This sutdy.

To fuse the genomic *PAB1* gene with mCherry, the reference strain was transformed with a fragment containing homologous flanking for *PAB1*, a 21 bp linker, the mCherry sequence, the 3'UTR of *PGK1* and the *HIS3* auxotrophic. This fragment was PCR amplified with primers Pab1-mCh-FW and Pab1-mCh-RV from *pYX022-mCherry* vector that was previously constructed as follows: a 934 bp fragment that comprised the linker, the mCherry sequence and the 3'UTR of *PGK1* was PCR amplified with primers mCh-FW and mCh-RV from the vector *YCplac33 EDC3-mCherry* (pRP1657), kindly provided by Prof. Parker (Buchan *et al.* 2008). The fragment was subsequently *Xma*I and *Xho*I cloned into *pYX022* plasmid (RandS system, Wiesbaden, Germany) to obtain the desired *pYX022-mCherry* vector.

The DNA sequences encoding for Pab1-GFP truncated versions were PCR amplified from *YCplac33 PAB1-GFP* (Martani *et al.* 2015) and cloned under the control of the *TPI* promoter into the *pYX012(-ATG)* plasmid, that was created by digesting with *Eco*RI and *Bam*HI, blunt-ending, and re-ligating the *pYX012* (RandS system, Wiesbaden, Germany). More in detail, the sequences encoding for all the synthetic variants carrying the C domain fused with GFP were PCR amplified using GFP-TER-RV as reverse primer and RRM1-TER-FW, or RRM2-TER-FW, or RRM3-TER-FW, or RRM4-TER-FW, or PC-TER-FW, or C-TER-FW-NEW as forward primers. The corresponding fragments were *Apa*I/*Sal*I cloned into *pYX012(-ATG)* to obtain vectors *pYX012(-ATG) RRM1-C-GFP*, *pYX012(-ATG) RRM2-C-GFP*, *pYX012(-ATG) RRM3-C-GFP*, *pYX012(-ATG) RRM4-C-GFP*, *pYX012(-ATG) P-C-GFP* and *pYX012(-ATG) C-GFP*.

Fragments lacking the sequence encoding for the C domain were PCR amplified using P-BamHI-RV as reverse primer and RRM1-TER-FW, or RRM2-TER-FW, or RRM3-TER-FW, or RRM4-TER-FW, or PC-TER-FW as forward primers. Fragments lacking the sequence encoding for the P and the C domains were PCR amplified using RRM4-BamHI-RV as reverse primer and RRM1-

TER-FW, or RRM2-TER-FW, or RRM3-TER-FW, or RRM4-TER-FW as forward primers. The corresponding fragments were cloned into *pYX012(-ATG) RRM1-C-GFP* plasmid (previously *Apa*I/*Bam*HI digested to remove the *RRM1-C* sequence and to maintain the *GFP*) to obtain vectors *pYX012(-ATG) RRM1-P-GFP*, *pYX012(-ATG) RRM2-P-GFP*, *pYX012(-ATG) RRM3-P-GFP*, *pYX012(-ATG) RRM4-P-GFP*, *pYX012(-ATG) P-GFP*, *pYX012(-ATG) RRM1-RRM4-GFP*, *pYX012(-ATG) RRM2-RRM4-GFP*, *pYX012(-ATG) RRM3-RRM4-GFP* and *pYX012(-ATG) RRM4-GFP*.

The recombinant vectors were *Stu*I linearized into the *URA3* marker and used to transform the *PAB1-mCherry* strain. Transformation was performed with the LiAc/PEG/ss-DNA protocol (Gietz and Woods 2002). The primers used in this study are listed in Table 2.

Table.2 List of primers used in this study

Primers	Sequence
mCh-FW	CTCGAG**CCCGGG**TTAATTAACATG
mCh-RV	TATGAA**CTCGAG**CATAAAGGCATTAAAAGAGGAGCG
Pab1-mCh-FW	GTCTTTCAAAAAGGAGCAAGAACAACAAACTGAGCAAGCTCGTCGACCCGGGTTAATTAAC
Pab1-mCh-RV	TAAGTTTGTTGAGTAGGGAAGTAGGTGATTACATAGAGCACGTTTTAAGAGCTTGGTGAGC
GFP-TER-RV	GGAAGC**GTCGAC**TTATTTGTATAGTTCATCCATGCC
RRM1-TER-FW	GTAAAA**GGGCCC**ATGGCTGATATTACTGATAAG
RRM2-TER-FW	GTAAAA**GGGCCC**ATGGGTAACATCTTTATCAAGAACT
RRM3-TER-FW	GTAAAA**GGGCCC**ATGTACACTAACCTTTATGTGAAAA
RRM4-TER-FW	GTAAAA**GGGCCC**ATGAATTTGTTTGTTAAGAACTTAGATG
PC-TER-FW	GTAAAA**GGGCCC**ATGTTGGCTCAACAAATCCAAGC

Table.2 Continue

Primers	Sequence
C-TER-FW	GTAAAA**GGGCCC**ATGCAAAAGCAAAGACAAGCTTTG
RRM4-BamHI-RV	GACTTT**GGATCC**GTTGAGAACGTCTTACGTC
P-BamHI-RV	GACTGG**GGATCC**GTTGATAAAATTGGTTGTTATCG

CCCGGG: *Xma*I restriction site; **CTCGAG**: *Xho*I restriction site; **GTCGAC**: *Sal*I restriction site; **GGGCCC**: *Apa*I restriction site; **GGATCC**: *Bam*HI restriction site.

Media and induction of stress granules

Strains were grown in shake flasks in Yeast Nitrogen Base (YNB) minimal medium, with 2% glucose, until the exponential phase of growth (OD_{660nm} = 0.3–0.6) prior to stress treatment. Heat shocked cells were prepared as previously described (Grousl *et al.* 2009): briefly, cell suspension was pelleted at 3000 rpm for 5 minutes, resuspended in pre-heated YNB minimal medium and then incubated at 46°C for 50 minutes at 160 rpm. Glucose deprived cells were prepared as previously described (Buchan *et al.* 2010): briefly, 2 mL of cells were centrifuged for 30 seconds at 13000 rpm, washed twice with YNB medium without glucose and incubated with the same medium at 30°C for 15 minutes at 160 rpm.

Protein extraction and western blot analysis

Cells were grown until the exponential phase (OD_{660nm} 0.3-0.6) prior to protein extraction: cells were centrifuged at 3000 rpm for five minutes and resuspended in extraction buffer composed of 0.1 M potassium phosphate buffer (pH 7.5), 1 mM ethylenediaminetetraacetic acid (EDTA), 1 mM dithiothreitol (DTT), 1 mM protease inhibitor cocktail, and 1 mM phenylmethane sulfonyl

fluoride (PMSF). Cells were subjected to three cycles of mechanical disruption with the FastPrep-24 (MP Biomedical). Cellular lysates were centrifuged at 14000 rpm for 20 minutes at 4°C. Supernatants were collected and protein concentration was estimated using bovine serum albumin (BSA) as reference (Bradford 1976). Proteins were resuspended in Laemmli buffer, boiled five minutes and resolved in a SDS-PAGE. Western blot was performed using the primary antibodies anti-GFP antibody JL-8 (Clontech) and anti-actin antibody C4 (Abcam). Pierce™ Rabbit anti-mouse IgG (Fc) conjugated with alkaline phosphatase (AP) (ThermoScientific) was used as secondary antibody. The membranes were developed adding CDP-Star Chemioluminescent Substrate (Sigma) and then exposed to Pierce CL-Xposure film to reveal signals. Bands were quantified using ImageJ 1.47v software.

Fluorescence microscopy analyses

Cells were observed with Nikon Eclipse 90i microscope and images were acquired with Digital Sight DS-U3 Nikon camera using NIS-Elements software (version 4.3). To visualize nuclei, cells in exponential phase of growth were stained with DAPI as described in (Forsburg and Rhind 2006) modified as follows: 200 µL of ethanol 70% was added to 1 mL of cells and incubated at room temperature for 5 minutes. Then cells were centrifuged at 6000 rpm for 2 minutes and resuspended in 100 µL of PBS with DAPI at a final concentration of 1 µg/mL. Quantitation of cells displaying both nuclear and cytosolic localization of the corresponding GFP tagged version was blindly determined for each cell by merging images of the green and the blue channels. Then, the percentage ratio between the number of cells with dual localization and the total number of cells was calculated. Quantification of stress granules and protein aggregates was performed in a blind manner for three independent experiments. For each experiment, green and red fluorescent dots were counted in a variable number of

cells between 60 and 100. For each cell, the ratio between the number of green and red dots was calculated and the overlap between green and red fluorescent dots was blindly determined by merging images of the green and red channels and looking for the resulting yellow dots; then, the ratio between the overlapped and the total number of dots was calculated. Finally, the mean and the standard deviation were calculated for all the parameters. Student *t*-test was used for evaluating significant differences between strains.

References

Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 1976;**72**:248-54

Brune C, Munchel SE, Fischer N *et al*. Yeast poly(A)-binding protein Pab1 shuttles between the nucleus and the cytoplasm and functions in mRNA export. *RNA* 2005;**11**:517-531.

Buchan JR, Parker R. Eukaryotic stress granules: the ins and outs of translation. *Mol Cell* 2009;**36**:932-941.

Buchan JR, Muhlrad D, Parker R. P bodies promote stress granule assembly in *Saccharomyces cerevisiae*. *J Cell Biol* 2008;**183**:441-455.

Buchan JR, Nissan T, Parker R. Analyzing P-bodies and stress granules in *Saccharomyces cerevisiae*. *Methods Enzymol* 2010;**470**:619-640.

Buchan JR, Yoon JH, Parker R. Stress-specific composition, assembly and kinetics of stress granules in *Saccharomyces cerevisiae*. *J Cell Sci* 2011;**124**:228-239.

Buchan JR, Kolaitis RM, Taylor JP *et al*. Eukaryotic stress granules are cleared by autophagy and Cdc48/VCP function. *Cell* 2013;**153**:1461-1474.

Decker CJ, Parker R. P-bodies and stress granules: possible roles in the control of translation and mRNA degradation. *Cold Spring Harb Perspect Biol* 2012;**4**:a012286.

Dunn EF, Hammell CM, Hodge CA *et al*. Yeast poly(A)-binding protein, Pab1, and PAN, a poly(A) nuclease complex recruited by Pab1, connect mRNA biogenesis to export. *Genes Dev* 2005;**19**:90-103.

Eliseeva IA, Lyabin DN, Ovchinnikov LP. Poly(A)-binding proteins: structure, domain organization, and activity regulation. *Biochemistry (Mosc)* 2013;**78**:1377-1391.

Forsburg SL, Rhind N. Basic methods for fission yeast. *Yeast* 2006;**23**:173-183.

Gietz RD, Woods RA. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol* 2002;**350**:87-96.

Giménez-Barcons M, Díez J. Yeast processing bodies and stress granules: self-assembly ribonucleoprotein particles. *Microb Cell Fact* 2011;**10**:73.

Goss DJ, Kleiman FE. Poly(A) binding proteins: are they all created equal? *Wiley Interdiscip Rev RNA* 2013;**4**:167-179.

Grousl T, Ivanov P, Malcova I *et al*. Heat shock-induced accumulation of translation elongation and termination factors precedes assembly of stress granules in *S. cerevisiae*. *PLoS One* 2013;**8**:e57083.

Grousl T, Ivanov P, Frýdlová I *et al.* Robust heat shock induces eIF2alpha-phosphorylation-independent assembly of stress granules containing eIF3 and 40S ribosomal subunits in budding yeast, *Saccharomyces cerevisiae*. *J Cell Sci* 2009;**122**:2078-2088.

Low JK, Hart-Smith G, Erce MA *et al*. The *Saccharomyces cerevisiae* poly(A)-binding protein is subject to multiple post-translational modifications, including the methylation of glutamic acid. *Biochem Biophys Res Commun* 2014;**443**:543-548.

Martani F, Marano F, Bertacchi S *et al*. The *Saccharomyces cerevisiae* poly(A) binding protein Pab1 as a target for eliciting stress tolerant phenotypes. *Sci Rep* 2015;**5**:18318.

Melamed D, Young DL, Miller CR *et al*. Combining natural sequence variation with high throughput mutational data to reveal protein interaction sites. *PLoS Genet* 2015;**11**:e1004918.

Parker R. RNA degradation in *Saccharomyces cerevisae*. *Genetics* 2012;**191**:671-702.

Protter DS, Parker R. Principles and Properties of Stress Granules. *Trends Cell Biol* 2016;**26**:668-679.

Riback JA, Katanski CD, Kear-Scott JL *et al*. Stress-Triggered Phase Separation Is an Adaptive, Evolutionarily Tuned Response. *Cell* 2017;**168**:1028-1040.e1019.

Richardson R, Denis CL, Zhang C *et al*. Mass spectrometric identification of proteins that interact through specific domains of the poly(A) binding protein. *Mol Genet Genomics* 2012;**287**:711-730.

Rollins M, Huard S, Morettin A *et al*. Lysine acetyltransferase NuA4 and acetyl-CoA regulate glucose-deprived stress granule formation in *Saccharomyces cerevisiae*. *PLoS Genet* 2017;**13**:e1006626.

Roque S, Cerciat M, Gaugué I *et al*. Interaction between the poly(A)-binding protein Pab1 and the eukaryotic release factor eRF3 regulates translation termination but not mRNA decay in *Saccharomyces cerevisiae*. *RNA* 2015;**21**:124-134.

Simón E, Séraphin B. A specific role for the C-terminal region of the Poly(A)-binding protein in mRNA decay. *Nucleic Acids Res* 2007;**35**:6017-6028.

Smith RW, Blee TK, Gray NK. Poly(A)-binding proteins are required for diverse biological processes in metazoans. *Biochem Soc Trans* 2014;**42**:1229-1237.

Swisher KD, Parker R. Localization to, and effects of Pbp1, Pbp4, Lsm12, Dhh1, and Pab1 on stress granules in *Saccharomyces cerevisiae*. *PLoS One* 2010;**5**:e10006.

Wahle E, Winkler GS. RNA decay machines: deadenylation by the Ccr4-not and Pan2-Pan3 complexes. *Biochim Biophys Acta* 2013;**1829**:561-570.

Walters RW, Parker R. Coupling of Ribostasis and Proteostasis: Hsp70 Proteins in mRNA Metabolism. *Trends Biochem Sci* 2015;**40**:552-559.

Wheeler JR, Matheny T, Jain S *et al*. Distinct stages in stress granule assembly and disassembly. *Elife* 2016;**5**.

Wyers F, Minet M, Dufour ME *et al*. Deletion of the PAT1 gene affects translation initiation and suppresses a *PAB1* gene deletion in yeast. *Mol Cell Biol* 2000;**20**:3538-3549.

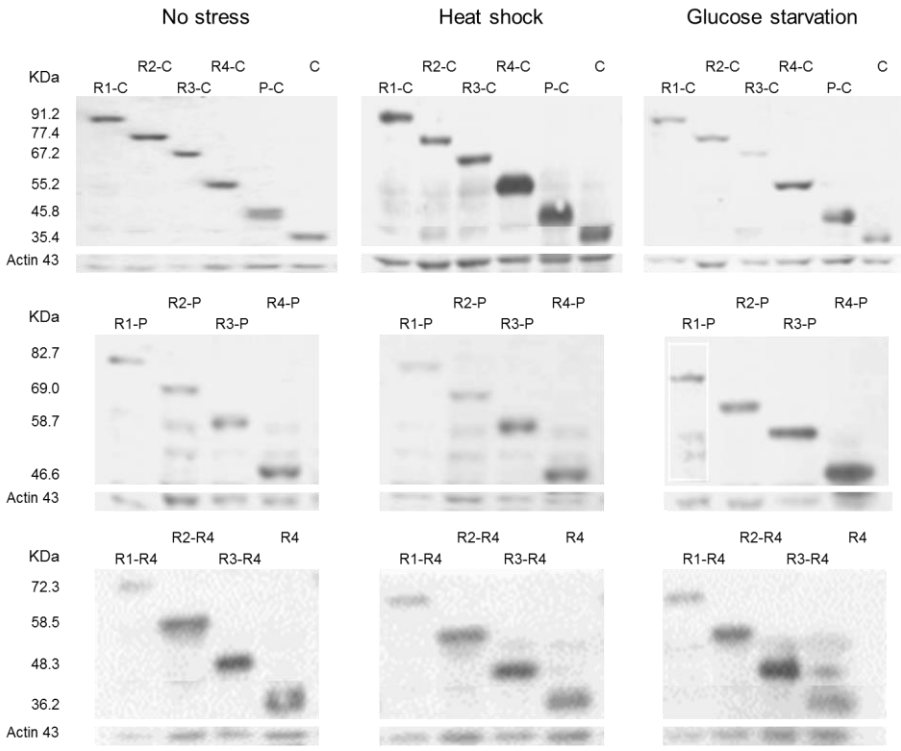
Yao G, Chiang YC, Zhang C *et al*. PAB1 self-association precludes its binding to poly(A), thereby accelerating CCR4 deadenylation in vivo. *Mol Cell Biol* 2007;**27**:6243-6253.

Zhang C, Lee DJ, Chiang YC *et al*. The RRM1 domain of the poly(A)-binding protein from *Saccharomyces cerevisiae* is critical to control of mRNA deadenylation. *Mol Genet Genomics* 2013;**288**:401-412.

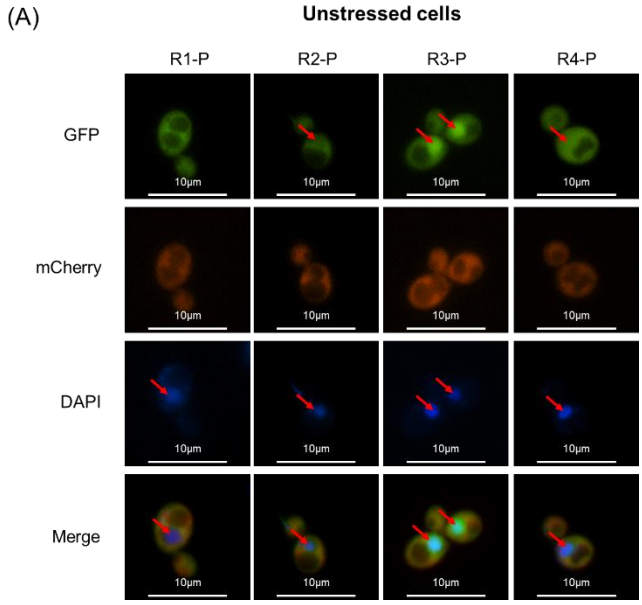
Zhao J, Hyman L, Moore C. Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol Mol Biol Rev* 1999;**63**:405-445.

Supplementary data

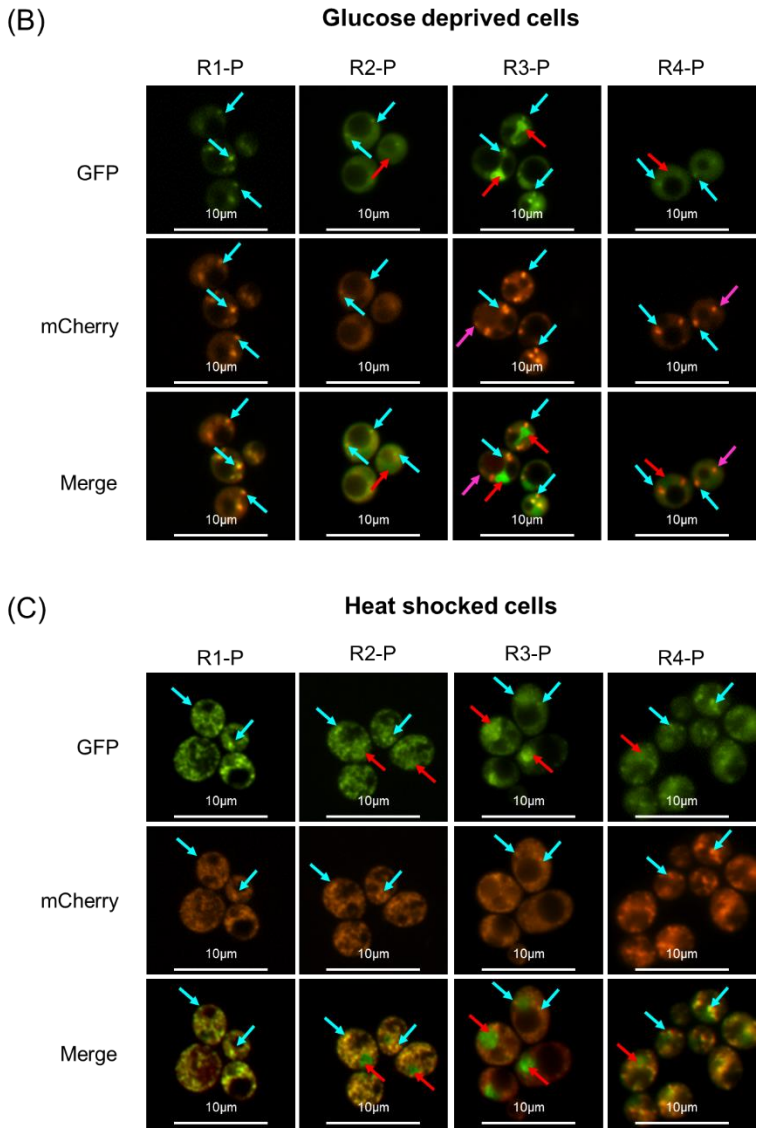
Supplementary Fig.S1 Western blot analysis of GFP-tagged Pab1 variants and actin levels in the absence and presence of stress. Anti-GFP and anti-actin were used as primary antibodies and AP conjugated rabbit anti-Mouse IgG (FC) as secondary antibody.



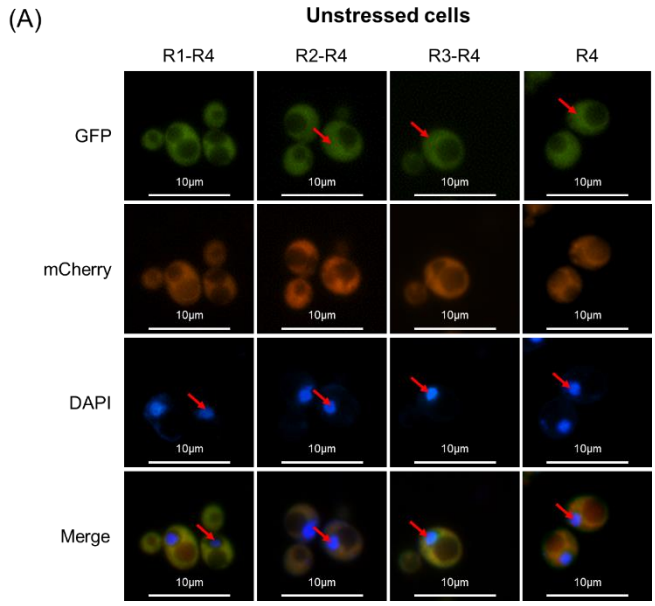
Supplementary Fig.S2 Effect of the removal of RRM and C domain on the localization of Pab1 in the absence and presence of stress. Intracellular localization of Pab1-mCherry and GFP-tagged Pab1 variants shown in Fig.5 Cells were grown until the exponential phase of growth (OD_{660nm} = 0.3–0.6) and then (A) stained with DAPI to a final concentration of 1 µg/mL or (B) glucose starved for 15 minutes or (C) heat shocked for 50 minutes. Light blue arrows indicate co-localized green and red foci; fuchsia arrows indicate Pab1-mCherry stress granules without a co-localizing counterpart in the Pab1-GFP variant; red arrows indicate nuclei.



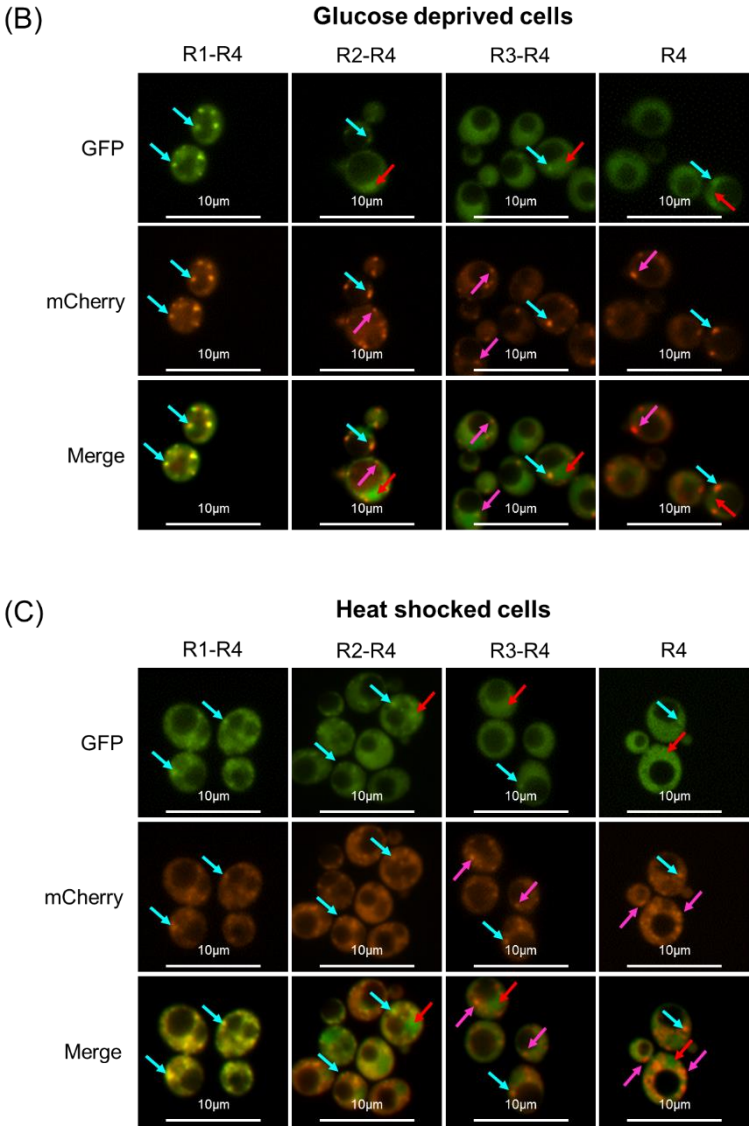
Supplementary Fig.S2 Continue



Supplementary Fig.S3 Effect of the removal of RRM, P and C domains on the localization of Pab1 in the absence and presence of stress. Intracellular localization of Pab1-mCherry and GFP-tagged Pab1 variants shown in Fig.7. Cells were grown until the exponential phase of growth (OD_{660nm} = 0.3–0.6) and then (A) stained with DAPI to a final concentration of 1 µg/mL or (B) glucose starved for 15 minutes or (C) heat shocked for 50 minutes. Light blue arrows indicate co-localized green and red foci; fuchsia arrows indicate Pab1-mCherry stress granules without a co-localizing counterpart in the Pab1-GFP variant; red arrows indicate nuclei.



Supplementary Fig.S3 Continue



Supplementary Table S1:

Percentage of cells displaying both nuclear and cytosolic localization. Values indicate the mean and the standard deviation of three independent experiments.

Strains	Ratio (%)
R1-C	9.25 (± 4.57)
R2-C	83.04 (± 7.00)
R3-C	95.27 (± 4.88)
R4-C	91.99 (± 1.53)
P-C	93.98 (± 2.07)
C	95.20 (± 2.32)
R1-P	10.70 (± 6.99)
R2-P	85.41 (± 7.45)
R3-P	96.71 (± 0.50)
R4-P	94.94 (± 1.83)
R1-R4	8.33 (± 1.19)
R2-R4	87.38 (± 6.42)
R3-R4	96.15 (± 1.51)
R4	90.43 (± 2.02)

Supplementary Table S2:

Values of overlapped/total dots ratio related to Figures 4A, 6A and 8A. Values indicate the mean and the standard deviation of three independent experiments.

Strains	Glucose starvation	Heat shock
R1-C	0.97 (± 0.04)	0.99 (± 0.01)
R2-C	0.92 (± 0.01)	0.95 (± 0.01)
R3-C	0.97 (± 0.01)	0.99 (± 0.00)
R4-C	0.96 (± 0.04)	0.98 (± 0.01)
P-C	0.04 (± 0.01)	0.04 (± 0.04)
C	-	-
R1-P	1.00 (± 0.01)	1.00 (± 0.01)
R2-P	0.93 (± 0.01)	0.98 (± 0.03)
R3-P	0.94 (± 0.01)	0.96 (± 0.02)
R4-P	0.89 (± 0.10)	0.96 (± 0.02)
R1-R4	0.95 (± 0.00)	0.99 (± 0.00)
R2-R4	0.96 (± 0.02)	0.94 (± 0.02)
R3-R4	0.99 (± 0.00)	0.98 (± 0.02)
R4	0.81 (± 0.14)	0.95 (± 0.07)

Supplementary Table S3:

Values of green/red fluorescent dots ratio related to Figures 4B, 6B and 8B. Values indicate the mean and the standard deviation of three independent experiments.

Strains	Glucose starvation	Heat shock
R1-C	1.07 (± 0.06)	1.11 (± 0.06)
R2-C	0.90 (± 0.05)	0.97 (± 0.03)
R3-C	0.67 (± 0.05)	0.84 (± 0.05)
R4-C	0.68 (± 0.06)	0.91 (± 0.03)
R1-P	1.15 (± 0.11)	1.30 (± 0.13)
R2-P	0.98 (± 0.01)	0.99 (± 0.05)
R3-P	0.81 (± 0.02)	0.96 (± 0.02)
R4-P	0.79 (± 0.13)	1.00 (± 0.01)
R1-R4	1.03 (± 0.01)	1.05 (± 0.03)
R2-R4	0.59 (± 0.01)	0.95 (± 0.05)
R3-R4	0.48 (± 0.03)	0.30 (± 0.03)
R4	0.04 (± 0.03)	0.12 (± 0.01)

Supplementary Table S4:

Western blot quantification. Values correspond to the ratio between GFP and actin signals quantified with ImageJ 1.47v software.

Strains	No stress	Heat shock	Glucose starvation
R1-C	3.27	1.78	2.00
R2-C	3.01	1.03	1.41
R3-C	2.60	1.23	1.46
R4-C	2.24	3.51	2.69
P-C	1.82	2.19	2.16
C	1.31	1.81	1.91
R1-P	1.81	2.91	1.05
R2-P	1.11	1.50	1.83
R3-P	2.05	3.55	2.80
R4-P	2.75	4.02	3.25
R1-R4	2.01	1.81	2.34
R2-R4	3.51	2.06	1.97
R3-R4	3.81	3.25	2.40
R4	2.80	1.62	1.69

Supplementary Table S5:

Percentage of glucose deprived and heat shocked cells containing aggregates of Pab1 variants. Values indicate the mean and the standard deviation of three independent experiments.

Strains	Glucose deprivation (%)	Heat shock (%)
R1-C	97.2 (± 3.9)	98.6 (± 2.0)
R2-C	95.6 (± 3.7)	99.3 (± 1.0)
R3-C	95.8 (± 4.5)	98.9 (± 1.6)
R4-C	81.2 (± 11.6)	98.9 (± 1.6)
P-C	81.0 (± 7.6)	95.5 (± 6.4)
C	-	-
R1-P	95.8 (± 5.9)	99.2 (± 1.1)
R2-P	98.9 (±1.5)	99.4 (± 0.8)
R3-P	95.2 (± 4.5)	99.4 (± 0.9)
R4-P	94.6 (± 0.6)	99.4 (± 0.8)
P	89.4 (± 9.9)	92.8 (±4.5)
R1-R4	99.2 (± 1.2)	99.4 (±0.8)
R2-R4	80.9 (± 1.3)	98.7 (± 0.0)
R3-R4	75.4 (± 11.1)	54.1 (± 10.7)
R4	13.2 (± 8.5)	44.7 (± 9.2)

Supplementary Table S6:

P-values related to the comparison between values reported in the Supplementary Table S3. *P*-values were calculated with 2-tailed student *t*-test.

Compared strains		Glucose deprivation (*P*-value)	Heat shock (*P*-value)
R1-C	R2-C	0.019	0.026
R1-C	R3-C	0.001	0.004
R1-C	R4-C	0.001	0.008
R2-C	R3-C	0.006	0.017
R2-C	R4-C	0.009	0.061
R3-C	R4-C	0.861	0.103
R1-P	R2-P	0.052	0.017
R1-P	R3-P	0.005	0.010
R1-P	R4-P	0.022	0.015
R2-P	R3-P	0.000	0.287
R2-P	R4-P	0.072	0.739
R3-P	R4-P	0.872	0.024
R1-R4	R2-R4	0.000	0.042
R1-R4	R3-R4	0.000	0.000
R1-R4	R4	0.000	0.000
R2-R4	R3-R4	0.005	0.000
R2-R4	R4	0.000	0.000
R3-R4	R4	0.000	0.001
R1-C	R1-P	0.280	0.089
R1-C	R1-R4	0.281	0.166
R1-P	R1-R4	0.108	0.028
R2-C	R2-P	0.049	0.544
R2-C	R2-R4	0.001	0.520
R2-P	R2-R4	0.000	0.319

Supplementary Table S6 Continue

Compared strains		Glucose deprivation (*P*-value)	Heat shock (*P*-value)
R3-C	R3-P	0.014	0.020
R3-C	R3-R4	0.006	0.000
R3-P	R3-R4	0.000	0.000
R4-C	R4-P	0.261	0.006
R4-C	R4	0.000	0.000
R4-P	R4	0.001	0.000

Conclusions and future perspectives

The main purpose of this thesis was to evaluate whether cellular rewiring approaches could be applied to isolate *S. cerevisiae* strains characterized by biotechnological interesting phenotypes. This possibility was evaluated by choosing the central nitrogen metabolism (CNM) and the mRNA metabolism as candidates to induce a profound cellular reorganization.

In the first case (see **Chapter 1**), the CNM was selected because almost all nitrogen sources are catabolized to glutamate (Magasanik and Kaiser, 2002) and, therefore, its modulation might have caused a cellular rearrangement with metabolic and physiological consequences. In particular, the glutamate synthase (Glt1), which is involved in the conversion of glutamine and α-ketoglutarate to glutamate, was selected as molecular target. In fact, since this enzyme uses NADH as a cofactor for catalysis, the modulation of *GLT1* expression might have provoked a redox intracellular unbalance and, thus, a profound cellular effect. Indeed, previous works in our laboratory partially supported this idea. However, as described in this thesis, no significant changes in growth, protein content, viability and ROS accumulation were associated to *GLT1* deletion or over-expression. As important exceptions, *GLT1* deletion negatively affected the scavenging activity of glutamate against ROS accumulation, when cells were treated with H_2O_2, whereas Glt1 overexpression led to lower viability in glutamine medium. While *GLT1* modulation does not significantly influence cell physiology, the nitrogen source supplied for fermentation (ammonium sulfate, glutamate or glutamine) considerably affected the parameters previously indicated. Overall, on one side these results highlight the importance of preliminarily evaluating the best nitrogen source for media formulation, and on the other, the awareness that the *S. cerevisiae* glutamate synthase is not a promising candidate for rewiring-based strategies.

These results underline that, to induce a marked cellular reprogramming, an appropriate target must be accurately selected, but due to lack of knowledge,

this is not a trivial task to accomplish. Interestingly, many works suggested the possibility to rewire cells by selecting a transcription factors (gTME approach) as "hub" element to induce pleiotropic effects and to isolate industrially interesting phenotypes, including more stress tolerant yeast (see, as example, Alper *et al.*, 2006). In this thesis, a similar approach was followed by selecting the main poly(A) binding protein (Pab1), which was previously demonstrated in our laboratory to be an effective target for isolating more acetic acid resistant strains (Martani *et al.*, 2015). This protein is implicated in many post-transcriptional event of the mRNA metabolism, such as poly(A) tail length regulation (Brown and Sachs, 1998), mRNA export (Brune *et al.*, 2005), translation initiation (Costello *et al.*, 2015), translation termination (Roque *et al.*, 2015), deadenylation (Wahle and Winkler, 2013) and stress granules formation (Swisher and Parker, 2010) (see **Chapter 2**). The final goal of the rewiring applied in this thesis was to isolate more thermotolerant *S. cerevisiae* strains that could find potential application for the development of efficient SSF and SSCF processes. Interestingly, some more thermotolerant strains, characterized by the presence of mutated *PAB1* variants in addition to the endogenous gene, were isolated with this approach (see **Chapter 3**). Therefore, results confirm not only that Pab1 is a powerful target to evoke complex phenotypes with improved traits, but also that cellular rewiring can be induced by acting at post-transcriptional level. Hence, in order to reprogram cells and to obtain industrially relevant phenotypes, the mRNA metabolism can be a promising candidate, whose potentialities will deserve future investigations through the selection of other elements involved in its regulation.

Remarkably, the identification of very diverse Pab1 truncated variants linked to improved yeast thermotolerance suggests that discrete portions of the protein may be associated to unexpected and unpredictable functions. This consideration implies that the single Pab1 domains (RNA recognition motifs, P

and C domains) might disclose novel surprising roles. Therefore, the modular structure of Pab1 may be an attractive platform for synthetic biology approaches aimed, for example, at rewiring the mRNA metabolism. Nevertheless, to exploit rationally single Pab1 domains as parts for synthetic biology purposes, a full understanding of their intracellular role is required.

To start, the contribution of the different domains in terms of Pab1 recruitment within stress granules (of which it is a specific component) was assessed (see **Chapter 4**). Stress granules play fundamental roles in cell physiology during stress, both by regulating gene expression (through the selective recruitment of specific mRNAs in an untranslated state (Zid and O'Shea, 2014)), and by selectively including regulatory proteins to temporary turn off their functions (*e.g.* TORC1) (Takahara and Maeda, 2012). Pab1 domains might be rationally exploited to design chimeric protein to allow, modulate or abolish their recruitment to stress granules, in order to induce complex physiological changes.

Overall, this thesis highlights that the development of cell factories with improved biotechnological features is not trivial to achieve, as shown in the work of Glt1. Conversely, a more promising expansion of the biorefinery concept towards a more sustainable society came from the studies on Pab1, which revealed huge potentialities of this protein both for developing yeast strains with desired industrial traits and for possible synthetic biology applications.

References

Alper H, Moxley J, Nevoigt E, Fink GR and Stephanopoulos G (2006) Engineering yeast transcription machinery for improved ethanol tolerance and production. *Science* **314**: 1565-1568.

Brown CE and Sachs AB (1998) Poly(A) tail length control in *Saccharomyces cerevisiae* occurs by message-specific deadenylation. *Mol Cell Biol* **18**: 6548-6559.

Brune C, Munchel SE, Fischer N, Podtelejnikov AV and Weis K (2005) Yeast poly(A)-binding protein Pab1 shuttles between the nucleus and the cytoplasm and functions in mRNA export. *RNA* **11**: 517-531.

Costello J, Castelli LM, Rowe W, *et al.* (2015) Global mRNA selection mechanisms for translation initiation. *Genome Biol* **16**: 10.

Magasanik B and Kaiser CA (2002) Nitrogen regulation in *Saccharomyces cerevisiae*. *Gene* **290**: 1-18.

Roque S, Cerciat M, Gaugué I, Mora L, Floch AG, de Zamaroczy M, Heurgué-Hamard V and Kervestin S (2015) Interaction between the poly(A)-binding protein Pab1 and the eukaryotic release factor eRF3 regulates translation termination but not mRNA decay in *Saccharomyces cerevisiae*. *RNA* **21**: 124-134.

Swisher KD and Parker R (2010) Localization to, and effects of Pbp1, Pbp4, Lsm12, Dhh1, and Pab1 on stress granules in *Saccharomyces cerevisiae*. *PLoS One* **5**: e10006.

Takahara T and Maeda T (2012) Transient sequestration of TORC1 into stress granules during heat stress. *Mol Cell* **47**: 242-252.

Wahle E and Winkler GS (2013) RNA decay machines: deadenylation by the Ccr4-not and Pan2-Pan3 complexes. *Biochim Biophys Acta* **1829**: 561-570.

Zid BM and O'Shea EK (2014) Promoter sequences direct cytoplasmic localization and translation of mRNAs during starvation in yeast. *Nature* **514**: 117-121.

Ringraziamenti

Giunto al termine di questi tre anni magnifici di Dottorato, mi sento in dovere di ringraziare molte persone speciali grazie alle quali ho raggiunto questo meraviglioso traguardo.

Il primo grazie va alla mia famiglia, mamma e papà, Daniele, Cristina e la new entry Anita, nonna Andrea, zia Virginia e zia Luigia, Corrado e Lella, perché ho sempre potuto contare su di voi e so che lo potrò fare sempre. Grazie mille!

Il secondo grazie va inevitabilmente a tutti quelle persone che sono passate dal Brandulab, perché se ogni mattino di questi tre anni mi sono svegliato felice di andare al lavoro, è stato soprattutto grazie alla sintonia che si è creata tra noi, ai confronti costruttivi su tutto, ai momenti di sfaso, alle uscite insieme e chi più ne ha più ne metta. Spero di rivivere tutto questo ovunque mi capiterà di lavorare e, in caso contrario (spero di no), allora sarebbe proprio vero che "C'è solo un Brandulab!".

Dei ringraziamenti un po' più speciali vanno a Nadia, Franci, Franci e Raffa, per il vostro supporto ogni volta che ho avuto dubbi, incertezze o momenti di sconforto.

Ovviamente, non posso non ringraziare i miei due ex-tesisti Alfredo e Ilaria, perché senza l'aiuto concreto e la passione che ci avete messo non sarebbe stata la stessa cosa.

Vorrei ringraziare i miei compagni di corso di Dottorato e Giancarlo Pornaccini per aver reso epici insieme i tre incontri di Verbania. In particolare vorrei ringraziare Giulia per la tua amicizia e per il sostegno che ci siamo dati a vicenda in questi anni.

Ringrazio poi tutti i miei amici di sempre: Fabio, Noe, Andre, Cecco, Ga, Albi, Cero, Mandella, Miki, Motti, Marghe, Walter, Massi, Ele, Ila, Anna, Pes, Sara, Elli, Tirel, Giada, Fede, Cri, Cone, Tommy, Chiara, Pool e, in particolare, Ale e Charlie per aver provveduto al mio "maledeto" stipendio.

Ovviamente ringrazio i miei amici di Bevera: Sma, Franca e Dan, Giulia e Jack, Laura, Clara, Silvia, Lolla, Alessia, Luca e Ele.

L'ultimo ringraziamento, lo voglio dedicare a Paola. Hai creduto in capacità che non credevo di avere, mi hai dato grande fiducia e spronato a dare il meglio anche quando pensavo di non essere all'altezza, cazziato quando serviva e mai a sproposito. Queste doti ti rendono unica e degna "Leader" di questo meraviglioso Brandulab. Grazie!