CellPress
REVIEWS

Review

Whole-Genome Duplication and Yeast's Fruitful Way of Life

Ximena Escalera-Fanjul,[1] Héctor Quezada,[2] Lina Riego-Ruiz,[3] and Alicia González[1,*]

Studies on the fate of *Saccharomyces cerevisiae* paralogous gene pairs that arose through a whole-genome duplication event have shown diversification of retained duplicated genes. Paralogous functional specialization often results in improved function and/or novel function that could contribute to the adaptation of the organism to a new lifestyle. Here, we analyze and discuss particular case studies of paralogous functional diversification that could have played a role in the acquisition of yeast fermentative metabolism.

Highlights

S. cerevisiae genomic duplication resulting from an allopolyploidization event gave rise to a hybrid allotetraploid.

Retention and diversification of WGD paralogous genes allowed the establishment of fermentative metabolism.

Global analysis and case-specific studies are beginning to uncover the molecular mechanisms by which WGD paralogous diversification enabled the development and establishment of a 'new way of life'.

Origin and Significance of Yeast's Whole-Genome Duplication

Saccharomyces cerevisiae was the first eukaryote whose genome was fully sequenced. Genome analysis revealed the presence of duplicated gene blocks, indicating that this yeast lineage arose from a **whole-genome duplication** (**WGD**) (see Glossary) [1–3]. WGDs are a recurring phenomenon in evolution. Some fungi, most species of flowering plants, and vertebrates descend from ancestors that doubled their genomes. WGDs have been associated with the acquisition of novel traits as well as increased species richness [4].

It has been proposed that the selective conservation of duplicated genes in *S. cerevisiae* has been essential for the acquisition of a predominantly fermentative lifestyle, also termed the 'make-accumulate-consume' strategy (i.e., upon sugar availability, ethanol is produced, accumulated, and finally consumed). This strategy largely relies on the Crabtree effect (Box 1) that allows yeast to out-compete other microorganisms [5]. The retention of 551 WGD gene pairs involved in different cellular processes and the complexity of the Crabtree effect make it difficult to demonstrate whether the selective retention of WGD paralogous genes led to 'a new way of life'. *S. cerevisiae* is presently the best suited model to study this phenomenon due to the vast amount of technical resources and information available for this organism [6]. In addition, the advancement of whole-genome sequencing and phenotyping of 1011 *S. cerevisiae* isolates from diverse geographical locations has provided valuable information into the ecological and evolutive niche of this yeast [7]. This review examines evidence provided by different experimental approaches addressing the role of WGD in the acquisition of fermentative metabolism in *S. cerevisiae*.

Mapping the Beginning of the Fermentative Lifestyle

After the publication of the complete *S. cerevisiae* genome sequence in 1996, several studies focused on understanding its origins and evolution. While some groups supported the idea that the duplicated regions observed in this yeast's genome were the signatures of an ancient WGD [3,5], others proposed that these duplicated regions arose independently by local duplication events [8]. This controversy was not solved until 2004, when the genomic sequences of other yeast species were published and used for comparative genomic analyses. Independent studies revealed that the genome of *S. cerevisiae* mapped in 2:1 manner to non-WGD yeast genomes, supporting the WGD model [9–11]. Further scrutiny and the development of the

[1]Departamento de Bioquímica y Biología Estructural, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, Mexico City, Mexico
[2]Laboratorio de Investigación en Inmunología y Proteómica, Hospital Infantil de México Federico Gómez, Mexico City, Mexico
[3]División de Biología Molecular, Instituto Potosino de Investigación Científica y Tecnológica AC, San Luis Potosí, Mexico

*Correspondence: amanjarr@ifc.unam.mx (A. González).

Box 1. *Saccharomyces cerevisiae* Fermentative Metabolism

The majority of cells only use fermentation when respiration is impaired, for example, when oxygen concentration is low (Pasteur effect). However, in the presence of high glucose concentrations and regardless of oxygen availability, *Saccharomyces cerevisiae* prefers to ferment glucose to ethanol. This physiological phenomenon is named the Crabtree effect [64]. In this metabolic state, the cytoplasmic pyruvate produced by glycolysis is transformed to acetaldehyde and then to ethanol, which diffuses to the extracellular medium [21].

During fermentation, the energy for growth is mainly provided by glycolysis and not by the oxidative respiration pathway. However, the amount of energy produced from a glucose molecule by fermentation is lower than that produced by its complete oxidation. Consequently, during fermentation *S. cerevisiae* displays a high glycolytic flux to match its growth requirements. The increased glycolytic rate under the fermentative regime helps yeasts grow much faster than when they grow in non-fermentable carbon sources, such as ethanol, where the oxidative respiration pathway is used to produce energy [65].

During fermentation, glucose repression reduces the mass and energy flux through the tricarboxylic acid (TCA) cycle, the respiratory chain, and the oxidative phosphorylation. However, the TCA cycle displays enough activity to provide the necessary precursors to produce the macromolecular components to sustain rapid growth rate [65,66]. Under these conditions, the TCA cycle does not work as a cycle, but rather in a branched manner, mainly because the succinate dehydrogenase complex does not operate under these conditions [67]. When glucose becomes limiting, cells undergo a so-called diauxic shift, a transient phase in which they switch to a fully respiratory metabolism, catabolizing the released ethanol via the TCA cycle [65].

S. cerevisiae response to sugar availability is dependent on different time scales and biological levels, from the reprogramming of gene expression to enzyme allosteric regulation, including post-transcriptional and post-translational regulatory mechanisms. It also involves a variety of cellular processes as different signal transduction pathways, energy and fatty acid metabolism, carbohydrate storage, and amino acid biosynthesis [68].

Yeast Gene Order Browser, an online tool for visualizing gene order context (synteny) of any gene from several yeast genomes [9,10,12], showed that the whole *S. cerevisiae* genome is duplicated [1–3,12]. Furthermore, it has now been confirmed that as well as *S. cerevisiae*, several other yeasts such as *Candida glabrata* arose from WGDs [13]. When WGD was initially proposed [3], it was not possible to tell whether it was the result of an **autopolyploidization** or an **allopolyploidization** event [14,15]. However, autopolyploidization constituted the most parsimonious hypothesis. Later, evidence was presented in favor of allopolyploidization, indicating that the *S. cerevisiae* lineage arose from an interspecies hybridization between a strain related to the *Kluyveromyces*, *Lachancea*, and *Eremothecium* (KLE) clade and another related to *Zygosaccharomyces rouxii* and *Torulaspora delbrueckii* (ZT) clade [9,16]. This proposal was extensively analyzed and accepted [2]. Therefore, the pair of **paralogous genes** that are observed today could have both originated from one of the parental species (ohnologous pairs), or each one of the parental species involved in the allopoliplodization event could have independently generated one member of the pair (homeologous genes) [16].

Even before the WGD controversy was settled (1996–2004), the fact that the WGD could have allowed *S. cerevisiae* to develop an oxygen-independent metabolism, through the retention and further specialization of duplicated genes/products, was considered for the first time [3,5] (Box 2). This proposition was based on the observation that the majority of yeasts were unable to grow in absence of oxygen [17].

The progress in comparative genomics in the 2000s allowed a better understanding of the evolution of the *S. cerevisiae* genome and a proper classification of the phylogenetic relationships among yeasts [13,18]. The new phylogenetic trees provided a framework to locate some major evolutionary events that could have contributed to the improved fermentation ability of

Glossary

Adaptive radiation: evolution of organisms into a wide variety of types adapted to specialized modes of life.
Allopolyploidization: a mechanism that generates an individual or strain with two complete sets of chromosomes, each one from a different species.
Autopolyploidization: in contrast to allopolyploidization, the two different sets of chromosomes are from the same species.
Biased fractionation: a phenomenon observed after WGDs where the ohnologues of a subgenome are preferentially lost.
Feedback inhibition: the phenomenon where the output of a process is used as an input to control the behavior of the process itself.
Genome dominance: a phenomenon observed after WGD where, in general, the ohnologues from one subgenome display higher transcriptional levels than their onhologous pair derived from the other subgenome.
Hetero-oligomers: a complex made of different protein subunits is called a hetero-oligomer; when only one type of protein subunit is used in the complex, it is called a homo-oligomer.
Moonlighting functions: protein moonlighting function is a phenomenon by which a protein can perform more than one biological role.
Orthologous: genes in different species that evolved from a common ancestral gene by speciation; orthologs retain the same function in the course of evolution.
Paralogous or ohnologous genes: duplicated genes in the same organism that were evolved from either a gene duplication, an autopolyploidization event, or an allopolyploidization event.
Robustness: the ability of a system to maintain its function despite a perturbation.
Small-scale duplication (SSD): doubling of a small-scale genomic regions involving one to a few genes.
Subgenome: set of genomic regions of a WGD species inherited from one of the parental species.

Whole-genome duplication (WGD): mechanism by which an organism doubles its chromosome number.

Box 2. Fate of Duplicated Genes

Gene duplicates, also called paralogs, are found in all three domains of life, suggesting that an important amount of genes were generated by gene duplication [69]. Although the immediate consequence of gene duplication is the doubling of the ancestral function (functional redundancy), after duplication genes can experience divergence that may result in the generation of new or specialized functions and thus shape evolution. This box summarizes the possible scenarios determining the fate of duplicated genes.

Loss of function: Loss of function occurs when functional redundancy is not advantageous and therefore one of the paralogous genes gradually becomes a pseudogene, through the accumulation of loss-of-function mutations [70]. Interestingly, soon after WGDs polyploids experience a rapid loss of genetic redundancy. The current hypothesis is that when WGDs result from an autopolyploidization, the loss of duplicated genes is randomly distributed between the two parental **subgenomes** (non-biased fractionation). By contrast, when the WGD derives from an allopoliploidization, loss of duplicated genes is biased towards a certain parental subgenome (**biased fractionation**), frequently towards the one displaying lower gene expression, a phenomenon called **genome dominance** [71,72].

Gene-dosage amplification: Gene-dosage amplification holds that after gene duplication, both gene copies are conserved since either incrementing the ancestral function represents a selective advantage or both copies are necessary to maintain the appropriate stoichiometry of higher complexes (e.g., protein complexes) [16,71,73,74]. However, it must be considered that increased gene copy number does not necessarily lead to an equivalent function increment, since function/product dosage can be negatively regulated [75–77]. Thus, the gene-dosage amplification requires that the increment in copy number produce an increase in the codified function.

Subfunctionalization: Subfunctionalization occurs when the ancestral function(s) is distributed between the two copies or is differentially specialized in each one, as a result of either the accumulation of degenerative mutations (duplication-degeneration-complementation model) or by positive selection (escape from adaptive conflict model). In both cases, both paralogs are necessary to perform the function that, in the ancestor, was carried out by a single gene [78–81].

Neofunctionalization: Neofunctionalization is the mechanism explaining the origin of novel functions through gene duplication. It considers that while one of the copies conserves the ancestral function, the other is free to acquire a new adaptative function through the accumulation of neutral mutations [82].

yeasts, such as the loss of the respiratory complex I, the horizontal transfer of the bacterial *URA1*, and the rewiring of rapid growth elements [19].

Starting in 2007, the role of the WGD in the acquisition of yeast fermentative metabolism began to be recognized, mainly through the characterization of the phenotypic responses of various representative species of the Saccharomycetaceae [19–21]. It was found that the Crabtree effect was present in several species of the above-mentioned family. However, yeasts closer to *Saccharomyces* in the phylogenetic tree displayed a more pronounced Crabtree effect, and with the exception of the *Tetrapisispora* group, were able to generate non-respiring mutants, while most of the non-WGD species could not [19,21]. As for the ability to grow under anaerobiosis, the post-WGD species were able to generate sufficient energy to grow under strictly anaerobic conditions, while non-WGD species showed a reduced Crabtree effect and ability to grow in the absence of oxygen [20]. These contributions support the proposition that yeasts' transition to a fermentative lifestyle began before the WGD in non-WGD yeasts in a rather multistep process. However, for Saccharomycetae, this metabolic trait was established after the WGD event [19]. This comparative physiological evidence, and the dating of both the WGD (from 150 to 100 Mya) and the **adaptive radiation** of flowering plants (from 101 to 66 Mya) in the Late Cretaceous [3,22,23], led to the hypothesis that the WGD allowed yeasts to compete for the increased amount of sugars present in fruits [19]. It could thus be expected that retained WGD genes could be enriched in cellular processes related to the make-accumulate-consume strategy.

WGD and Small-Scale Duplications Contribution to Fermentative Metabolism

The contribution of WGD to *S. cerevisiae* fermentative lifestyle could be assessed comparing WGD and **small-scale duplication** (SSD) gene sets in terms of genetic interactions and divergence in expression and functions. One of the first studies on this matter proposed that WGD, but not SSD, must have provided the raw material for conservation of dosage-sensitive genes and groups of genes expressed at a precise stoichiometric ratio, such as those encoding ribosomal proteins [24]. It has been proposed that such dosage and stoichiometric effects could have been involved in the conservation of glycolytic genes [25]. Consistent with this it was found that gene duplicates (combined WGD and SSD) were overrepresented in central carbon metabolism and hexose transport [26]. However, the separate analysis of the gene sets revealed that enrichment in genes involved in carbohydrate metabolism was higher for WGD than for SSD [27]. Divergence in the expression patterns between members of a pair has also been reported to be higher for the WGD than for the SSD set of paralogs [27]. It may be possible that divergent regulation between members of some duplicated pairs could have contributed to the specialization of one paralog for glycolysis-ethanol production and the other for gluconeogenesis-ethanol consumption phases. The impact of such differential regulation on the acquisition of fermentative metabolism could be estimated by comparing the rates of negative expression correlations between members of a pair during growth on glucose or ethanol as carbon sources. To explore this possibility, data from a quantitative proteomic study were used where the relative abundances of 3609 proteins were determined [28]. Comparison of the abundance of both paralogs during fermentation, diauxic shift, and glucose exhaustion phases indicated that 29% of WGD and 19% of SSD duplicates showed differential regulation during fermentative and respiratory metabolism. The same trend was observed using data from a transcriptomic study that determined the fold change in transcript levels for 5800 genes from cells cultured on rich media with glucose or ethanol as carbon sources [29]; the estimated percentage of divergent expression was 26% for the WGD and 14% for the SSD duplicates (for analysis details of proteomic and transcriptomic data, see the legend of Figure S1 in the supplemental information online). Interestingly, the enzyme-coding subset of genes with differential expression during fermentative and respiratory metabolism among WGD genes clustered around carbohydrate metabolism (Figure 1A; Figure S1A). By contrast, enzymes encoded by SSD paralogs seem to be randomly dispersed around various metabolic pathways (Figure 1B; Figure S1B) [28,29].

Comparison of gene ontology (GO) categories' fold enrichment between divergent and complete sets of genes (Figure 2) shows that for the WGD group, glucose metabolism GO terms are overrepresented (Figure 2A). Furthermore, the WGD divergent set displays a higher fold enrichment in glucose metabolism GO terms than that exhibited by the WGD complete set (Figure 1A; Figure S1A). Moreover, glucose metabolism is not represented in the SSD group (Figure 2B), which for divergent and complete categories is enriched in biosynthetic processes, suggesting a different evolutionary trend. Overall, these results indicate that the WGD set of genes display a more relevant contribution than those of SSD set to the acquisition of a fermentative lifestyle in the *S. cerevisiae* lineage.

Duplicated Genes Case Studies Uncover Molecular Mechanisms behind Subfunctionalization

Detailed analysis of subfunctionalization of particular paralogous pairs is necessary to identify the molecular mechanisms involved in functional diversification and the roles of specific duplicate genes in the acquisition of fermentative metabolism. We first analyze four examples

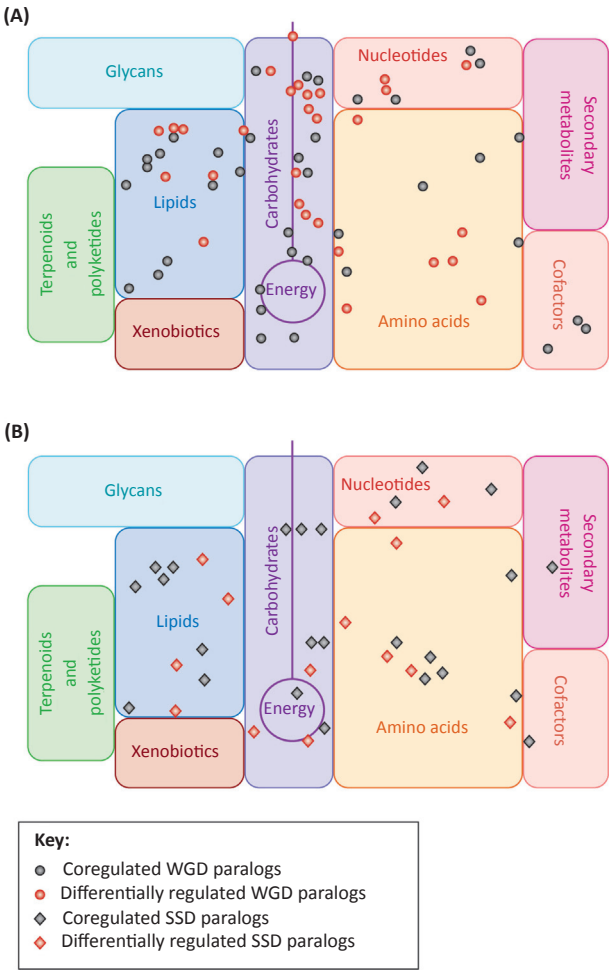

Figure 1. Duplicated Gene Pairs Originated from WGD Are Enriched for Genes Encoding Enzymes Involved in Carbohydrate Metabolism. Enzymes encoded by whole-genome duplication (WGD) (A) or small-scale duplication (SSD) (B) paralogs were mapped on the Kyoto Encyclopedia of Genes and Genomes refer pathway map (http://www.genome.jp/kegg/pathway/map/map01100.html) [84]. Coregulated or differentially regulated paralogs were selected according to their regulation pattern during fermentative or respiratory conditions from global studies [28,29]. For pathway details, gene identity, and analysis, please see Figure S1 in the supplemental information online.

of gene pairs that, according to the phylogeny-based prediction of PhylomeDB v4, could be proposed as ohnologous WGD pairs [30].

ENO1- and *ENO2*-encoded phosphopyruvate hydratases catalyze the reversible conversion of 2-phosphoglycerate to phosphoenolpyruvate. Under fermentative conditions, Eno2 constitutes the predominant form, while under respiratory conditions both polypeptides are present and contribute to gluconeogenesis [31]. Random assortment of Eno1/Eno2 monomers allows the formation of homo- and hetero-dimers: Eno1-Eno1, Eno2-Eno2, and Eno1-Eno2, whose peculiar organization could influence the opposed role these enzymes play in the glycolysis and gluconeogenesis [32] (Figure 3).

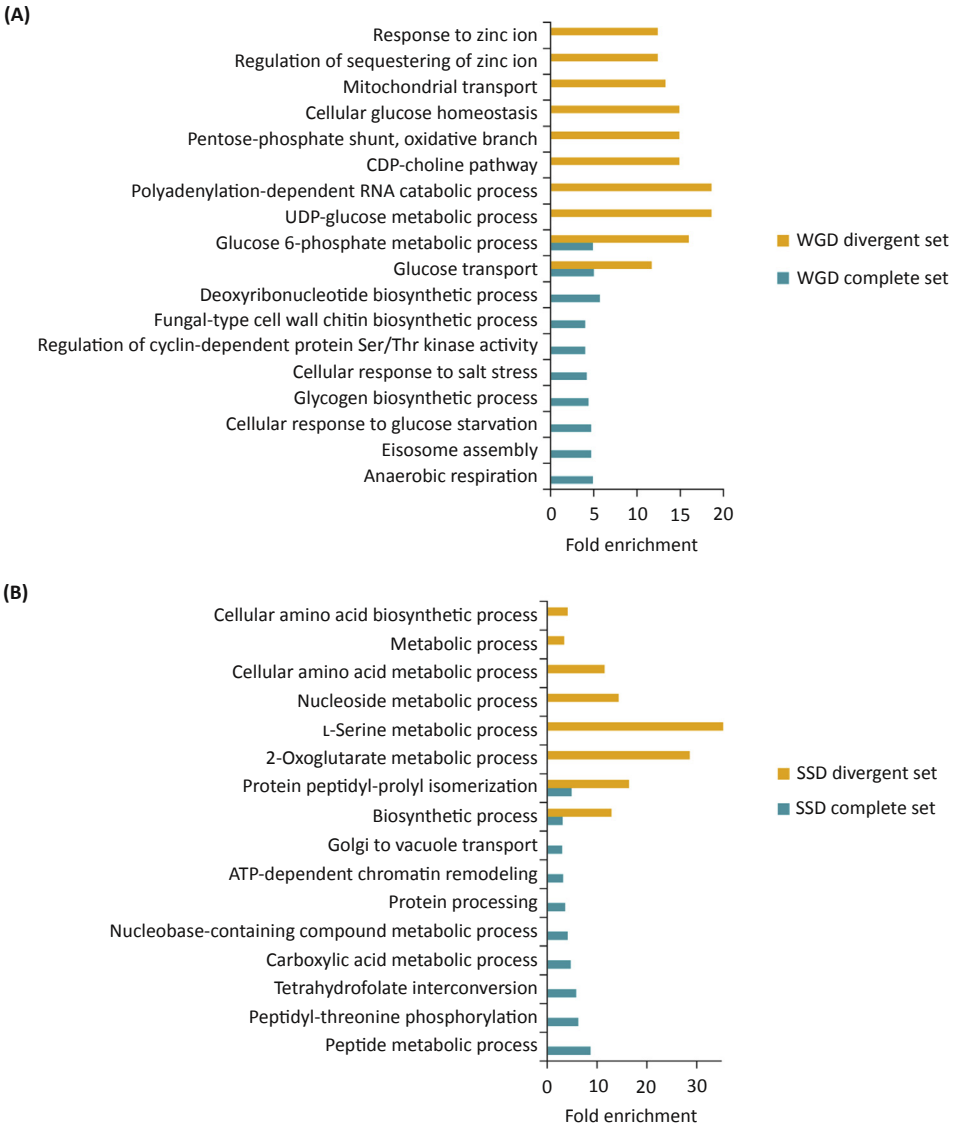

Figure 2. WGD and SSD Gene Sets Contribute to Different Cellular Processes during Fermentative and Respiratory Metabolism. Top ten GOTERM_BP_DIRECT categories with higher enrichment fold values among whole-genome duplication (WGD) (A) or small-scale duplication (SSD) (B) gene sets are shown ($P < 0.01$). The complete gene sets include 1094 and 674 genes for WGD and SSD, respectively. The sets with divergent expression patterns include 328 and 106 genes for WGD and SSD, respectively. Paralogs with divergent regulation between members of a pair were obtained from global studies [28,29] analyzed as described in Figure S1 in the supplemental information online. For the set of SSD paralogs with divergent regulation, only eight terms result with $P < 0.01$. Gene annotation enrichment analysis was made using the DAVID 6.7 website (https://david.ncifcrf.gov/) with default values. CDP, cytidine diphosphate; UDP, uridine diphosphate.

Formation of heteromeric isozymes plays a role in functional diversification of the two ohnologous WGD-pair [33] encoding cytosolic NADP-GDHs (*GDH1* and *GDH3*), catalyzing glutamate biosynthesis from ammonia and α-ketoglutarate (α-KG). While the homo-hexameric Gdh1 is the sole isoform present under fermentative conditions, the homo-hexameric Gdh3

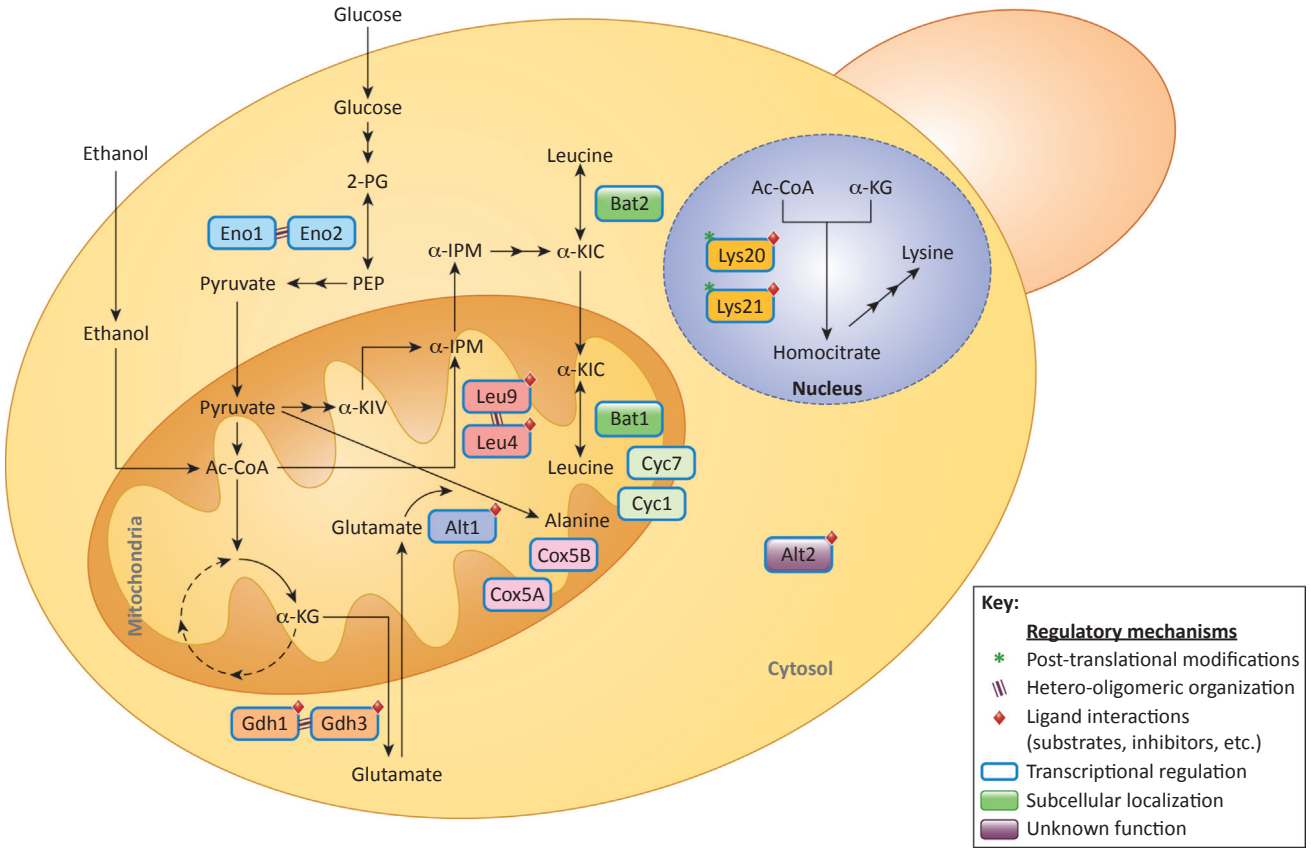

Figure 3. Molecular Mechanisms of Paralog Subfunctionalization. Enzyme-coding whole-genome duplication paralogs can diverge and become specialized in different metabolic conditions. Illustrative examples and the corresponding mechanisms of subfunctionalization are shown. Ac-CoA, acetyl coenzyme A; α-IPM, alpha isopropylmalate α-KG, alpha ketoglutarate; α-KIC, alpha ketoisocaproate; KIV, alpha ketoisovalerate; PEP, phosphoenolpyruvate; 2-PG, 2-phosphoglycerate.

is only present in a *gdh1*Δ mutant grown on ethanol as sole carbon source and the Gdh1-Gdh3 hetero-hexameric isoform(s) prevail under respiratory conditions [34]. Accordingly, *GDH1* expression is maintained through the different growth phases, while that of *GDH3* is null during fermentation and due to chromatin remodeling increases 70-fold when cells enter diauxic shift and respiratory growth [35].

The affinity of the three isoforms towards α-KG (Gdh1-Gdh1 > Gdh1-Gdh3 > Gdh3-Gdh3) agrees with the idea that under fermentative growth, energy is provided by glycolysis; thus, the high rate activity of Gdh1 does not compromise energy yield. Under respiratory conditions when the tricarboxylic acid (TCA) cycle contributes to energy production, *GDH1* and *GDH3* are simultaneously expressed and a variety of isozymes can be built up depending on the Gdh1:Gdh3 ratio. Therefore, the modulable heteromeric isoforms constitute a pacemaker mechanism determining glutamate biosynthesis and balanced α-KG utilization [34,36] (Figure 3).

The kinetic properties of the NADP-*Lk*Gdh1 present in the non-WGD yeast *Lachancea kluyveri* are closer to those of Gdh1, while kinetics of the *Kluyveromyces lactis* NADP-*Kl*Gdh1 resemble those of Gdh3, suggesting there is a correlation between the NADP-GDHs kinetic properties

and the level of adaptation to the fermentative or respiratory lifestyle [33]. The above-mentioned results indicate that both *GDH1* and *GDH3* originated from the KLE lineage, supporting predictions obtained through PhylomeDB v4 [30].

Paralogous mitochondrial homo- and heteromeric isozymes were formed by the *LEU4*/*LEU9* ohnologous WGD-pair [30], encoding α-isopropylmalate synthases that synthesize α-isopropylmalate (α-IPM) from acetyl-coenzyme A (CoA) and α-ketoisovalerate, an intermediate of leucine biosynthesis and co-activator of the Leu3 transcriptional regulator [37]. Leu4-Leu4 and Leu9-Leu9 homodimers are leucine sensitive or resistant isoforms, respectively, while the heterodimeric isoform Leu4-Leu9 shows intermediate leucine sensitivity. On ethanol as carbon source, *LEU4* expression is 3-fold increased compared to that found in glucose, while that of *LEU9* is low on both carbon sources. Leu4-Leu9 and Leu4-Leu4 formation prevents leucine and α-IPM over-synthesis, mediating a more balanced distribution of pyruvate and acetyl-CoA to intermediary biosynthesis and energy yielding metabolism particularly under respiratory conditions. Exclusive expression of Leu9, in a *leu4Δ* genetic background, results in slow growth on either glucose or ethanol as carbon sources and leucine accumulation, indicating that Leu9-Leu9 resistance to leucine **feedback inhibition** impairs flux control through the biosynthetic pathway [37] (Figure 3).

Lys20 and Lys21 nuclear homocitrate synthases participate in lysine biosynthesis by catalyzing homocitrate production, condensing acetyl-CoA with α-KG [36]. Lysine is a strong inhibitor for Lys21, inducing positive cooperativity for α-KG binding. Under fermentative conditions, Lys20 and Lys21 play redundant roles, and even when *LYS20* and *LYS21* ohnologous WGD-pair are overexpressed, and lysine pools increase 4- to 10-fold, growth rate is not reduced, confirming the conclusion that α-KG drainage is not deleterious under fermentative conditions [36]. Lys21 activity is crucial for growth under respiratory conditions and can sustain wild-type growth by itself. Subfunctionalization includes differential *LYS20* and *LYS21* expression and a post-transcriptional mechanism that determines the abundance of each enzyme. Differential lysine inhibition, α-KG modulation, and the control of the intracellular amount of the two enzymes result in a regulatory mechanism determining the rate at which α-KG is diverted to either lysine biosynthesis or to other metabolic pathways (Figure 3) [38].

The above-mentioned presented cases show that the combined action of the formation of paralogous homo- and hetero-oligomeric isozymes with peculiar kinetic properties and contrasting gene expression results in functional diversification, providing a means to differentially control metabolic fluxes under fermentative or respiratory conditions.

The four next examples correspond to WGD-pairs that could be tentatively proposed as homeologous, according to PhylomeDB v4 prediction [30]. Bat1 and Bat2 branched chain aminotransferases are located in mitochondria and cytosol, respectively; thus, formation of **hetero-oligomers** is impeded. Diversification of the *BAT1* and *BAT2* relies on the acquisition of opposed expression profiles [39]. Bat1 has a biosynthetic character, since *BAT1* expression is privileged in the absence of branched chain amino acids, while Bat2 displays a catabolic role. *BAT2* expression is maximal in the presence of branched chain amino acids [39,40]. Acquisition of environmentally regulated expression profiles could allow operation of differential or redundant roles depending on the peculiar environments found in *S. cerevisiae* natural habitats [7,41,42].

In regard to the contribution of Bat1 and Bat2 adaptation to fermentative metabolism, *BAT1* decreased expression under respiratory conditions and resulted in diminished metabolic flow to

amino acid biosynthesis, favoring energy-yielding pathways (Figure 3) [40]. The *KlBAT1* **orthologous gene**, present in *K. lactis* non-WGD yeast, encodes an enzyme displaying the dual biosynthetic-catabolic character that was later specialized among the *S. cerevisiae BAT1/BAT2* paralogous pair [39].

Functional characterization of the *ALT1/ALT2* WGD-pair gene/products established that only Alt1 possessed the presumed role of alanine transaminase activity [43]. The analysis of the only orthologous gene/product present in *L. kluyveri* and *K. lactis* non-WGD type yeasts suggested that Alt2 functional diversification resulted in the loss of its alanine transaminase activity [44]. Probably, as the result of a structural change modifying the interaction with its cofactor pyridoxal 5-phosphate (PLP), allowing the formation of only one of the two tautomeric PLP isomers necessary for the transamination reaction. The fact that Alt2 forms a catalytically active Schiff base with PLP and is found in a phylogenetic branch with several *Saccharomyces* genera, which do not harbor Alt1, suggests this enzyme has a yet undiscovered function (Figure 3) [45].

Cytochrome *c* oxidase (COXC) is a complex composed of several proteins that transfer electrons from cytochrome *c* to molecular oxygen [46]. Cox5 is essential for complex assembly and has two paralogous isoforms: Cox5A and Cox5B. *CYC1* and *CYC7* encode paralogous isoforms of the electron carrier protein cytochrome *c*, promoting the final transfer of electrons to oxygen by COXC. Under aerobic conditions, *COX5A* and *CYC1* are expressed, whereas *COX5B* and *CYC7* are Rox repressed and are only expressed under hypoxia. Differential expression of these genes had been previously considered to be important in yeast evolutionary adaptation to anaerobic growth [3]. As for *BAT1* and *BAT2*, functional diversification is achieved through the acquisition of opposed transcriptional patterns [40].

For some cases, orthologous *K. lactis* and/or *L. kluyveri* counterparts have been analyzed; however, it would be important to study both a KLE and a ZT representative for each case. This is mandatory to clarify whether sub- or neofunctionalization has occurred.

WGD Paralogous Genes Impart Robustness and Phenotypic Plasticity

As mentioned above, duplicated genes allow *S. cerevisiae* to adjust its phenotype to properly respond to a changing environment. The role of paralogous homo- and hetero-oligomers as key players regulating metabolic fluxes leading to new phenotypes has also been highlighted. In this section, we discuss this evidence in terms of **robustness** and phenotypic plasticity (Box 3). Gene duplicates have been associated to genetic robustness [47–49]; however, a recent study found opposing evidence suggesting that in some cases gene duplication can impart fragility [50]. By analyzing the protein–protein interaction network of 56 pairs of duplicated genes (WGD or SSD), it was found that in one-third of the cases, upon deletion of a paralogous copy, the remaining copy conserved its capacity to interact with its native binding partners and increased its association with those of the deleted copy, thus compensating paralogous pair loss [50]. In a similar amount of cases, after deletion of its cognate pair, the remaining copy decreased interactions even with its own binding partners, demonstrating dependency on its paralogous copy. In these cases, the lower expressed copy of the homologous pair was the only one displaying the dependency phenotype, indicating asymmetry and showing a stronger decrease in fitness (fragility) than that displayed by the independent copies. Furthermore, in the absence of its pair, the protein concentration of the dependent copy decreased, although its mRNA showed wild-type levels, suggesting protein destabilization [50,51].

Considering that, in general, dependent pairs form paralogous hetero-oligomers, it was proposed that the different paralogous isoforms could be degraded at different rates

Box 3. Duplicated Genes Contribute to Genetic Robustness

Robustness is defined as the ability of a system to maintain its function upon a perturbation. In living systems, perturbations can either be environmental or genetic. Multiple mechanisms that confer robustness to biological systems have been described. (i) Distributed robustness. Upon a perturbation, specific biological functions can be accomplished by alternative pathways, as in signaling and metabolic routes [49,83]. (ii) Redundancy. Multiple functional copies of a given component of the biological system can compensate the failure of another highly similar component, for example, paralogous genes [49,83]. (iii) Feedback control. By sensing the output of a function, it is possible to modulate the input to either narrow the output function (negative feedback) or to amplify and maintain signals against noise (positive feedback). This type of control contributes to preservation of cellular function (homeostasis) [47–49]. (iv) Modularity. The confinement of functions minimize the effects of disturbances by containing perturbations and damage locally [47–49].

It is considered that duplicated genes contribute to genetic robustness since the paralogous copies maintain a certain degree of functional redundancy. Thus, when one of the gene duplicates is deleted, the remaining copy can provide enough function to compensate the deletion of its paralogous pair. In several cases, the wild-type function of the retained copy is enough to compensate the deleted pair (passive paralogous compensation). In many other cases, the remaining paralog needs to modify its performance, to avoid the phenotypic or fitness defects produced by the deletion of its pair (active paralogous compensation). Different molecular mechanisms of active paralogous compensation have been identified: (i) Transcriptional reprograming. Some paralogous genes are upregulated upon the genetic perturbation, functionally compensating the deleted copy. (ii) Protein relocalization. Even though there are limited examples, the remaining copy could experience relocalization to the subcellular compartment were the deleted copy used to reside. (iii) Protein interaction rewiring. Some paralogous genes can increase the association with the binding partners of the missing copy [51].

Another important attribute of biological systems is phenotypic plasticity. This plasticity is recognized as the ability of a single genotype to produce different phenotypes, allowing a best performance within changing environments. Often phenotypic plasticity is considered to be the opposite of robustness [54]. However, paralogous genes have also been associated with phenotypic plasticity [62].

(independent homo-oligomers < hetero-oligomers < dependent homo-oligomers < monomers) [52,53]. However, the above-mentioned analysis did not account for self-interactions [50]. Interestingly, for the paralogs proteins Leu4 and Leu9 the formation of the Leu4-Leu9 heterodimer is preferred over that of the homodimers, possibly as the result of a higher affinity between different monomers than between self-monomers. However, both paralogous copies can form homodimers; nevertheless, it cannot be ruled out that when the concentration of the independent form is too low, oligomer:monomer equilibrium may be displaced towards the monomeric forms, resulting in higher degradation [52]. Besides Leu4-Leu9 heterodimers, other paralogous genes are capable of forming hetero-oligomeric complexes such as Gdh1-Gdh3 and Eno1-Eno2. For Leu4/Leu9 and Gdh1/Gdh3, in all cases the three different isoforms show specific kinetic parameters [32,34,37]. The coexistence of more than one isoform allows coping with a wider range of metabolite concentrations (substrates, inhibitors, etc.) than when there is only one isoform, providing a buffering effect and thus contributing to robustness (feedback control).

Another important attribute of biological systems is the phenotypic plasticity. This plasticity is recognized as the ability of a single genotype to produce different phenotypes, allowing a best performance within changing environments. Often phenotypic plasticity is considered to be the opposite of robustness [54]. However, since different isozymes are used in distinct metabolic context [34,37], regulating carbon flux to biomass production. It is possible that the ability to form hetero-oligomers contribute to the acquisition of the Crabtree effect and thus to phenotypic plasticity. If this is the case, heteromeric forming paralogous proteins could be considered as major regulators of both robustness and phenotypic plasticity, providing optimal fitness within changing environments [55].

Concluding Remarks and Future Directions

Recent advances in entire genome sequencing have generated an increasing interest in understanding the role of polyploidy on ecological shifts. In particular, different approaches have been used to elucidate the relationship between WGD and the acquisition of fermentative metabolism in *S. cerevisiae*. By comparing some physiological traits among the Saccharomycetaceae, it was possible to determine the central role of the WGD in the establishment of a new way of life for *S. cerevisiae*. The localization of the origin of fermentative metabolism before the WGD [19] suggests the prior existence of a selection pressure towards an oxygen-independent metabolism. The evolutionary trends followed by paralogous genes originated from WGD, which were identified through global analyses, have unveiled the general strategies that favored the establishment of a new lifestyle [26,29]. Moreover, the detailed analysis of particular duplicated pairs has shown some of the mechanisms behind the functional diversification of paralogous genes. Although detailed studies are scarce, some general rules on how particular changes in paralogous pairs have contributed to the acquisition of fermentative metabolism are starting to emerge (Figure 3).

Particularly interesting has been the finding that formation of hetero-oligomeric complexes contributes to robustness and phenotypic plasticity by enhancing the role of paralogous enzymes influencing the intracellular concentration of key metabolites such as α-IPM controlling the Leu3 transcriptional network [40]; leucine influencing cell growth control [56,57]; lysine determining response to oxidative stress [58]; and α-KG playing a functional role in aging [59].

Although the neolocalization of a paralogous copy is frequently observed [60], the evolutionary and physiological role of diversification of subcellular localization remains to be fully addressed (see Outstanding Questions) and certainly constitutes an open pathway to understand functional diversification of paralogs. In this respect, the analysis of **moonlighting functions** could govern the diversification of subcellular localization.

Future studies of paralogous pairs diversification must imperatively include specific analyses to understand whether a particular pattern is the outcome of sub- or neofunctionalization. To this end, pertinent paralogous functional characterization will also have to be addressed in non-WGD yeasts.

The rapidly increasing amount of *S. cerevisiae* isolates and the availability of their entire genome sequences will promote the systematical study of WGD genes from yeast strains collected from natural habitats [7,41,42,61–63] that could lead to the understanding of the involvement of duplicated genes in other processes besides the fermentative metabolism observed in controlled environments.

Acknowledgments

We acknowledge Beatriz Aguirre-López for advice and productive discussions during the development of this review. Our research is supported by grant from Dirección de Asuntos del Personal Académico (IN204018) and Consejo Nacional de Ciencia y Tecnología (CB-2014-239492-B). This work was supported by the Deutsche Forschungsgemeinschaft (RTG1772 'Computational Systems Biology').

Supplemental Information

Supplemental information associated with this article can be found online at https://doi.org/10.1016/j.tig.2018.09.008.

Outstanding Questions

What is the role of subcellular localization in paralogues functional diversification? Although it has been '*a priori*' considered that subcellular localization could have an important influence in the physiological role of a protein, no experiments have been performed that result in re-localization or co-localization of paralogous enzymes.

Are moonligthing properties inherited from the ancestor or did they arise after WGD? Do moonlighting properties diversify between paralogs? There are several examples of paralogous genes generated through WGD in which one of the members of the paralogous pair has an identified additional function (moonlight). However detailed analysis of its counterpart is compelling to definitively establish whether the moonlighting property has been conserved in both paralogs. Differential localization of enzymes such as Bat1 and Bat2 placed in mitochondria and cytosol, respectively, could be considered as a mechanism preventing formation hetero-oligomers. However, since Bat1 has a moonlighting function forming a complex with aconitase, its mitochondrial localization could have an important role in TCA regulation and not in its aminotransferase function.

Mathematical models have proven useful in understanding complex system behavior. Models may help answer the following questions at a systemic level: How does gene duplication alter metabolic fluxes and transcriptional regulation? What are the underlying regulatory mechanisms that allowed yeasts to acquire a fermentative lifestyle? How do WGD homologous genes orchestrate energy metabolism and biomass production to sustain the rapid growth rate observed during the Crabtree effect?

What are the most appropriate culture conditions to address the role of duplicated genes?

What is the relative contribution of post-translational modifications to divergence in regulation of proteins encoded by duplicated genes?

References

1. Goffeau, A. *et al.* (1996) Life with 6000 genes. *Science* 274, 546, 563–567
2. Wolfe, K.H. (2015) Origin of the yeast whole-genome duplication. *PLoS Biol.* 13, 1–7
3. Wolfe, K.H. and Shields, D.C. (1997) Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* 387, 708–713
4. Van De Peer, Y. *et al.* (2017) The evolutionary significance of polyploidy. *Nat. Rev. Genet.* 18, 411–424
5. Piškur, J. (2001) Origin of the duplicated regions in the yeast genomes. *Trends Genet.* 17, 302–303
6. Botstein, D. and Fink, G.R. (2011) Yeast: an experimental organism for 21st century biology. *Genetics* 189, 695–704
7. Peter, J. *et al.* (2018) Genome evolution across 1,011 *Saccharomyces cerevisiae* isolates. *Nature* 556, 339–344
8. Koszul, R. *et al.* (2004) Eucaryotic genome evolution through the spontaneous duplication of large chromosomal segments. *EMBO J.* 23, 234–243
9. Kellis, M. *et al.* (2004) Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* 428, 617–624
10. Dietrich, F.S. *et al.* (2004) The *Ashbya gossypii* genome as a tool for mapping the ancient *Saccharomyces cerevisiae* genome. *Science* 304, 304–307
11. Dujon, B. *et al.* (2004) Genome evolution in yeasts. *Nature* 430, 35–44
12. Byrne, K.P. and Wolfe, K.H. (2005) The yeast gene order browser: combining curated homology and syntenic context reveals gene fate in polyploid species. *Genome Res.* 15, 1456–1461
13. Kurtzman, C. (2003) Phylogenetic circumscription of *Saccharomyces*, *Kluyveromyces* and other members of the *Saccharomycetaceae*, and the proposal of the new genera *Lachancea*, *Nakaseomyces*, *Naumovia*, *Vanderwaltozyma* and *Zygotorulaspora*. *FEMS Yeast Res.* 4, 233–245
14. Gordon, J.L. *et al.* (2009) Additions, losses, and rearrangements on the evolutionary route from a reconstructed ancestor to the modern *Saccharomyces cerevisiae* genome. *PLoS Genet.* 5, 1–14
15. Gordon, J.L. *et al.* (2011) Mechanisms of chromosome number evolution in yeast. *PLoS Genet.* 7, 1–13
16. Marcet-Houben, M. and Gabaldón, T. (2015) Beyond the whole-genome duplication: phylogenetic evidence for an ancient interspecies hybridization in the baker's yeast lineage. *PLoS Biol.* 13, 1–26
17. Olsson, L. and Mo, K. (2001) Ability for anaerobic growth is not sufficient for development of the petite phenotype in *Saccharomyces kluyveri*. *J. Bacteriol.* 183, 2485–2489
18. Kurtzman, C. and Robnett, C. (2003) Phylogenetic relationships among yeasts of the "*Saccharomyces* complex" determined from multigene sequence analyses. *FEMS Yeast Res.* 3, 417–432
19. Hagman, A. *et al.* (2013) Yeast "make-accumulate-consume" life strategy evolved as a multi-step process that predates the whole genome duplication. *PLoS One* 8, 1–12
20. Merico, A. *et al.* (2007) Fermentative lifestyle in yeasts belonging to the *Saccharomyces* complex. *FEBS J.* 274, 976–989
21. Dashko, S. *et al.* (2014) Why, when, and how did yeast evolve alcoholic fermentation? *FEMS Yeast Res.* 14, 826–832
22. Augusto, L. *et al.* (2014) The enigma of the rise of angiosperms: can we untie the knot? *Ecol. Lett.* 17, 1326–1338
23. Lunau, K. (2004) Adaptive radiation and coevolution — pollination biology case studies. *Org. Divers. Evol.* 4, 207–224
24. Davis, J.C. and Petrov, D.A. (2005) Do disparate mechanisms of duplication add similar genes to the genome? *Trends Genet.* 21, 548–551
25. Conant, G.C. and Wolfe, K.H. (2007) Increased glycolytic flux as an outcome of whole-genome duplication in yeast. *Mol. Syst. Biol.* 3, 1–12
26. Kuepfer, L. *et al.* (2005) Metabolic functions of duplicate genes in *Saccharomyces cerevisiae*. *Genome* 15, 1421–1430
27. Guan, Y. *et al.* (2007) Functional analysis of gene duplications in *Saccharomyces cerevisiae*. *Genetics* 175, 933–943
28. Murphy, J.P. *et al.* (2015) Comprehensive temporal protein dynamics during the diauxic shift in *Saccharomyces cerevisiae*. *Mol. Cell. Proteomics* 14, 2454–2465
29. Mattenberger, F. *et al.* (2017) The phenotypic plasticity of duplicated genes in *Saccharomyces cerevisiae* and the origin of adaptations. *Genes Genomes Genet.* 7, 63–75
30. Huerta-Cepas, J. *et al.* (2014) PhylomeDB v4: zooming into the plurality of evolutionary histories of a genome. *Nucleic Acids Res.* 42, 897–902
31. Cohen, R. *et al.* (1987) Transcription of the constitutively expressed yeast enolase gene *ENO1* is mediated by positive and negative *cis*-acting regulatory sequences. *Mol. Cell. Biol.* 7, 2753–2761
32. Mcalister, L. and Holland, M.J. (1982) Targeted deletion of a yeast enolase structural gene. *J. Biol. Chem.* 257, 7181–7188
33. Campero-Basaldua, C. *et al.* (2017) Diversification of the kinetic properties of yeast NADP-glutamate-dehydrogenase isozymes proceeds independently of their evolutionary origin. *Microbiologyopen* 6, 1–18
34. DeLuna, A. *et al.* (2001) NADP-glutamate dehydrogenase isoenzymes of *Saccharomyces cerevisiae*. *J. Biol. Chem.* 276, 43775–43783
35. Avendaño, A. *et al.* (2005) Swi/SNF-GCN5-dependent chromatin remodelling determines induced expression of *GDH3*, one of the paralogous genes responsible for ammonium assimilation and glutamate biosynthesis in *Saccharomyces cerevisiae*. *Mol. Microbiol.* 57, 291–305
36. Quezada, H. *et al.* (2008) Specialization of the paralogue *LYS21* determines lysine biosynthesis under respiratory metabolism in *Saccharomyces cerevisiae*. *Microbiology* 154, 1656–1667
37. López, G. *et al.* (2015) Diversification of paralogous α-isopropylmalate synthases by modulation of feedback control and heterooligomerization in *Saccharomyces cerevisiae*. *Eukaryot. Cell* 14, 564–577
38. Quezada, H. *et al.* (2011) The Lys20 homocitrate synthase isoform exerts most of the flux control over the lysine synthesis pathway in *Saccharomyces cerevisiae*. *Mol. Microbiol.* 82, 578–590
39. Colón, M. *et al.* (2011) *Saccharomyces cerevisiae* Bat1 and Bat2 aminotransferases have functionally diverged from the ancestral-like *Kluyveromyces lactis* orthologous enzyme. *PLoS One* 6, 1–13
40. González, J. *et al.* (2017) Diversification of transcriptional regulation determines subfunctionalization of paralogous branched chain aminotransferases in the yeast *Saccharomyces cerevisiae*. *Genetics* 207, 975–991
41. Liti, G. *et al.* (2009) Population genomics of domestic and wild yeasts. *Nature* 458, 337–341
42. Naumov, G.I. *et al.* (2013) Molecular genetic diversity of the *Saccharomyces* yeasts in Taiwan: *Saccharomyces arboricola*, *Saccharomyces cerevisiae* and *Saccharomyces kudriavzevii*. *Antonie Van Leeuwenhoek* 103, 217–228
43. Peñalosa-Ruiz, G. *et al.* (2012) Paralogous *ALT1* and *ALT2* retention and diversification have generated catalytically active and inactive aminotransferases in *Saccharomyces cerevisiae*. *PLoS One* 7, 1–13
44. Escalera-Fanjul, X. *et al.* (2017) Evolutionary diversification of alanine transaminases in yeast: catabolic specialization and biosynthetic redundancy. *Front. Microbiol.* 8, 1–16
45. Rojas-Ortega, E. *et al.* (2018) *Saccharomyces cerevisiae* differential functionalization of presumed ScALT1 and ScALT2 alanine transaminases has been driven by diversification of pyridoxal phosphate interactions. *Front. Microbiol.* 9, 1–16
46. Dodia, R. *et al.* (2014) Comparisons of subunit 5A and 5B isoenzymes of yeast cytochrome c oxidase. *Biochem. J.* 464, 335–342

47. Stelling, J. *et al.* (2004) Robustness of cellular functions. *Cell* 118, 675–685
48. Kitano, H. (2004) Biological robustness. *Nat. Rev. Genet.* 5, 826–837
49. Frederik Nijhout, H. *et al.* (2017) Systems biology of phenotypic robustness and plasticity. *Integr. Comp. Biol.* 57, 171–184
50. Diss, G. *et al.* (2017) Gene duplication can impart fragility, not robustness, in the yeast protein interaction network. *Science* 355, 630–634
51. Diss, G. *et al.* (2014) Molecular mechanisms of paralogous compensation and the robustness of cellular networks. *J. Exp. Zool. Part B Mol. Dev. Evol.* 322, 488–499
52. Veitia, R.A. (2017) Gene duplicates: agents of robustness or fragility? *Trends Genet.* 33, 377–379
53. Veitia, R.A. (2017) Gene duplicates: agents of fragility? – a reply to Landry and Diss. *Trends Genet.* 33, 658–660
54. Padilla, D.K. *et al.* (2014) Addressing grand challenges in organismal biology: the need for synthesis. *Bioscience* 64, 1178–1187
55. Lachowiec, J. *et al.* (2016) Molecular mechanisms governing differential robustness of development and environmental responses in plants. *Ann. Bot.* 117, 795–809
56. Kingsbury, J.M. *et al.* (2015) Branched-chain aminotransferases control TORC1 signaling in *Saccharomyces cerevisiae*. *PLoS Genet.* 11, 1–24
57. Bonfils, G. *et al.* (2012) Leucyl-tRNA synthetase controls TORC1 via the EGO complex. *Mol. Cell* 46, 105–110
58. O'Doherty, P.J. *et al.* (2014) Transcriptomic and biochemical evidence for the role of lysine biosynthesis against linoleic acid hydroperoxide-induced stress in *Saccharomyces cerevisiae*. *Free Radic. Res.* 48, 1454–1461
59. Chin, R.M. *et al.* (2014) The metabolite α-ketoglutarate extends lifespan by inhibiting ATP synthase and TOR. *Nature* 510, 397–401
60. Marques, A.C. *et al.* (2008) Functional diversification of duplicate genes through subcellular adaptation of encoded proteins. *Genome Biol.* 9, 1–12
61. Stefanini, I. *et al.* (2016) Dynamic changes in microbiota and mycobiota during spontaneous 'Vino Santo Trentino' fermentation. *Microb. Biotechnol.* 9, 195–208
62. Legras, J.-L. *et al.* (2018) Adaptation of *S. cerevisiae* to fermented food environments reveals remarkable genome plasticity and the footprints of domestication. *Mol. Biol. Evol.* 35, 1712–1727
63. Liti, G. (2015) The fascinating and secret wild life of the budding yeast *S. cerevisiae*. *Elife* 4, 1–9
64. Kayikci, Ö. and Nielsen, J. (2015) Glucose repression in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* 15, 1–8
65. Compagno, C. *et al.* (2014) Introduction to carbon metabolism in yeast. In *Molecular Mechanisms in Yeast Carbon Metabolism*, pp. 1–19, Springer
66. Sokolov, S.S. *et al.* (2015) Negative feedback of glycolysis and oxidative phosphorylation: mechanisms of and reasons for it. *Biochemistry* 80, 559–564
67. Gombert, A.K. and Moreira, M. (2001) Network identification and flux quantification in the central metabolism of *Saccharomyces cerevisiae* under different conditions of glucose repression. *J. Bacteriol.* 183, 1441–1451
68. Tripodi, F. *et al.* (2015) Post-translational modifications on yeast carbon metabolism: regulatory mechanisms beyond transcriptional control. *Biochim. Biophys. Acta Gen. Subj.* 1850, 620–627
69. Zhang, J. (2003) Evolution by gene duplication: an update. *Trends Ecol. Evol.* 18, 292–298
70. Conrad, B. and Antonarakis, S.E. (2007) Gene duplication: a drive for phenotypic diversity and cause of human disease. *Annu. Rev. Genomics Hum. Genet.* 8, 17–35
71. Emery, M. *et al.* (2018) Preferential retention of genes from one parental genome after polyploidy illustrates the nature and scope of the genomic conflicts induced by hybridization. *PLoS Genet.* 14, 1–24
72. Liang, Z. and Schnable, J.C. (2018) Functional divergence between subgenomes and gene pairs after whole genome duplications. *Mol. Plant* 11, 388–397
73. Birchler, J.A. and Veitia, R.A. (2007) The gene balance hypothesis: from classical genetics to modern genomics. *Plant Cell Online* 19, 395–402
74. Wapinski, I. *et al.* (2007) Natural history and evolutionary principles of gene duplication in fungi. *Nature* 449, 54–61
75. Kondrashov, F.A. (2012) Gene duplication as a mechanism of genomic adaptation to a changing environment. *Proc. R. Soc. B Biol. Sci.* 279, 5048–5057
76. Kondrashov, F.A. and Kondrashov, A.S. (2006) Role of selection in fixation of gene duplications. *J. Theor. Biol.* 239, 141–151
77. Tang, Y.C. and Amon, A. (2013) Gene copy-number alterations: a cost-benefit analysis. *Cell* 152, 394–405
78. Hughes, A.L. (1994) The evolution of functionally novel proteins after gene duplication. *Proc. Biol. Sci.* 256, 119–124
79. Innan, H. and Kondrashov, F. (2010) The evolution of gene duplications: classifying and distinguishing between models. *Nat. Rev. Genet.* 11, 97–108
80. Sikosek, T. *et al.* (2012) Escape from adaptive conflict follows from weak functional trade-offs and mutational robustness. *Proc. Natl. Acad. Sci. U. S. A.* 109, 14888–14893
81. Espinosa-Cantú, A. *et al.* (2015) Gene duplication and the evolution of moonlighting proteins. *Front. Genet.* 6, 1–7
82. Ohno, S. (1970) *Evolution by Gene Duplication,* Springer
83. Wagner, A. (2008) Robustness and evolvability: a paradox resolved. *Proc. R. Soc. B Biol. Sci.* 275, 91–100
84. Kanehisa, M. *et al.* (2016) KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* 44, D457–D462
85. Fares, M.A. *et al.* (2013) The roles of whole-genome and small-scale duplications in the functional specialization of *Saccharomyces cerevisiae* Genes. *PLoS Genet.* 9, 1–13