

RESEARCH ARTICLE

Aneuploidy influences the gene expression profiles in *Saccharomyces pastorianus* group I and II strains during fermentation

Roberto de la Cerda Garcia-Caro¹, Karsten Hokamp², Fiona Roche², Georgia Thompson¹, Soukaina Timouma³, Daniela Delneri³, Ursula Bond^{1*}

1 Moynes Institute, School of Genetics and Microbiology, Trinity College Dublin, College Green, Dublin, Ireland, **2** Smurfit Institute of Genetics, School of Genetics and Microbiology, Trinity College Dublin, College Green, Dublin, Ireland, **3** Manchester Institute of Biotechnology, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, United Kingdom

* ubond@tcd.ie



OPEN ACCESS

Citation: de la Cerda Garcia-Caro R, Hokamp K, Roche F, Thompson G, Timouma S, Delneri D, et al. (2022) Aneuploidy influences the gene expression profiles in *Saccharomyces pastorianus* group I and II strains during fermentation. PLoS Genet 18(4): e1010149. <https://doi.org/10.1371/journal.pgen.1010149>

Editor: Justin C. Fay, University of Rochester, UNITED STATES

Received: December 3, 2021

Accepted: March 15, 2022

Published: April 7, 2022

Copyright: © 2022 de la Cerda Garcia-Caro et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All data underlying the results are available as Supplemental Material and [Supporting Information](#) files. The WGS are available at <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA810967> RNAseq data is available at <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE197654>.

Funding: RC and ST were supported by European Union's Horizon 2020, Marie Skłodowska-Curie

Abstract

The lager yeasts, *Saccharomyces pastorianus*, are hybrids of *Saccharomyces cerevisiae* and *Saccharomyces eubayanus* and are divided into two broad groups, Group I and II. The two groups evolved from at least one common hybridisation event but have subsequently diverged with Group I strains losing many *S. cerevisiae* chromosomes while the Group II strains retain both sub-genomes. The complex genomes, containing orthologous alleles from the parental chromosomes, pose interesting questions regarding gene regulation and its impact on the fermentation properties of the strains. Superimposed on the presence of orthologous alleles are complexities of gene dosage due to the aneuploid nature of the genomes. We examined the contribution of the *S. cerevisiae* and *S. eubayanus* alleles to the gene expression patterns of representative Group I and II strains during fermentation. We show that the relative expression of *S. cerevisiae* and *S. eubayanus* orthologues is positively correlated with gene copy number. Despite the reduced *S. cerevisiae* content in the Group I strain, *S. cerevisiae* orthologues contribute to biochemical pathways upregulated during fermentation which may explain the retention of specific chromosomes in the strain. Conversely, *S. eubayanus* genes are significantly overrepresented in the upregulated gene pool in the Group II strain. Comparison of the transcription profiles of the strains during fermentation identified both common and unique gene expression patterns, with gene copy number being a dominant contributory factor. Thus, the aneuploid genomes create complex patterns of gene expression during fermentation with gene dosage playing a crucial role both within and between strains.

Author summary

Saccharomyces pastorianus are yeasts used for making lager type beers and are natural hybrids of two other yeasts, *Saccharomyces cerevisiae* and *Saccharomyces eubayanus*. The hybrids formed just 500–600 years ago, and the combined parental genomes are

Innovative Training Networks Programme, Project Aromagenesis (grant number 76436), <https://ec.europa.eu/research/mariecurieactions/actions/doctoral-networks>. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

responsible for the clean crisp flavours associated with lager beers. There are two types of lager yeasts: Group I strains have lost a significant portion of *S. cerevisiae* chromosomes, while the Group II strains contain the full *S. cerevisiae* complement. Both contain the full set of *S. eubayanus* chromosomes. An unusual consequence of the hybridisation is that the genomes of lager yeasts are aneuploid with the copy numbers of chromosomes ranging from 1–6. Aneuploidy is often associated with cancer in humans and therefore an understanding of how aneuploidy contributes to gene expression in lager yeasts may provide insights into its role in tumour cells. Here, we show that gene expression patterns are influenced by chromosomal aneuploidy with transcript levels directly correlated with gene dosage. We also examined the role played by the parental genomes in the gene expression profiles under fermentation conditions and show that while both genomes contribute to the transcript pools, *S. eubayanus* genes are over-represented during fermentation.

Introduction

The yeasts used in lager beer production have long been recognised as being unique and distinct from those used for making ales [1]. In 1870, in recognition of their unique physiological qualities, this group of yeast were given a taxonomical classification of *Saccharomyces pastorianus* by Max Reess [2]. After more than a century of genetic studies, we now know that *S. pastorianus* strains are natural hybrids of *Saccharomyces cerevisiae* and *Saccharomyces eubayanus* [3,4]. While *S. cerevisiae* has long been associated with beer production, especially for ales, *S. eubayanus* was only discovered in Patagonia, South America, in 2011 [5]. While Patagonia is considered to be the primary radiation source of *S. eubayanus*, natural isolates have also been found in North America, China, Tibet and New Zealand but to date, no isolates have been identified in Europe [6–9] with only limited metagenomic evidence for the presence of this species in Italian Alps [10]. The hybridisation events that generated the current strains of *S. pastorianus* are estimated to have occurred some 500–600 years ago following the introduction of *S. eubayanus* into Central Europe from China or Tibet, most likely along trade routes such as the Silk Road [9]. An inherited property of cryotolerance from the *S. eubayanus* genome allows *S. pastorianus* strains to ferment at temperatures as low as 7–13°C [11]. This property together with the robust fermentation kinetics inherited from the *S. cerevisiae* genome created yeast hybrid strains capable of producing a crisp clean tasting lager that is now the most favoured alcohol-containing beverage.

Genome analysis of *S. pastorianus* strains identified two distinct types, based on genome content and chromosome composition [4,12–17]. Group I, or Saaz strains, are typically triploid in DNA content, retaining all the parental *S. eubayanus* chromosomes but have lost many *S. cerevisiae* chromosomes. This group includes several strains isolated from breweries in Bavaria (Germany) and Bohemia (Czechia) and includes the strains CBS1538, CBS1513, and CBS1503 that were originally isolated in The Carlsberg Laboratory in the late 19th century [18]. The Group II, or Froberg strains, include isolates from Dutch, Danish and North American breweries and are mainly tetraploid in DNA content, containing approximately 2n *S. cerevisiae* and 2n *S. eubayanus* genome content. Both groups display chromosomal aneuploidy with chromosome numbers ranging from one to six [19,20]. In addition to the parental chromosomes, *S. pastorianus* strains contain several hybrid chromosomes containing both *S. cerevisiae* and *S. eubayanus* genes that resulted from recombination, at precise locations, between

the parental chromosomes. Some of the recombination breakpoints are located within coding regions, creating a set of hybrid genes unique to lager yeasts [4,12,17,21].

Based on genome analysis, the current understanding of the origin and evolution of *S. pastorianus* strains is that both Group I and II strains evolved from a common hybridisation event between *S. eubayanus* and *S. cerevisiae* strains to generate a progenitor hybrid. Subsequently, this progenitor strain underwent a second hybridisation event with a second *S. cerevisiae* strain to generate Group II strains [4,22]. This model is supported by shared recombination sites on hybrid chromosomes, evidence of Single Nucleotide Polymorphisms (SNPs) in the *S. cerevisiae* genome of Group II strains and differences in telomeric regions in Group I and II strains [22]. The genome data is also consistent with a scenario in which both groups emerged from a single hybridisation event between a diploid *S. eubayanus* strain and a heterozygous *S. cerevisiae* diploid strain, with Group I strains experiencing a selective loss of a significant proportion of the heterogeneous *S. cerevisiae* genome [16,20,23]. Both groups then evolved independently with each undergoing further recombination events between the sub-genomes [3,17,20,23]. Both groups have subsequently diverged to create distinct sub-groups, each with their own unique physiological and biological properties. Fermentation analysis of Group I and II strains reveals that each group produces distinctive aroma and flavour profiles [24].

The complex genome of *S. pastorianus*, containing orthologous alleles emanating from different parental chromosomes, poses interesting questions regarding gene regulation and its impact on the fermentation properties of the strains. Superimposed on the presence of orthologous alleles are complexities of gene dosage due to the aneuploid nature of the genomes. Aneuploidy has been shown to influence gene expression patterns in eukaryotic cells [25–28]. The presence of gene orthologues from two different parental *Saccharomyces* species, together with copy number differences between the orthologues has the potential to affect the cellular proteome and specifically the stoichiometry of *S. cerevisiae* and *S. eubayanus* proteins within protein complexes [29]. Furthermore, gene copy number differences between the Group I and Group II strains may lead to differences in cellular physiology and thus the fermentation properties of the two types of *S. pastorianus*.

Previous transcriptome analyses of *S. pastorianus* strains under fermentation conditions were limited by technology and mainly focussed on the analysis of *S. cerevisiae* genes [30–33]. Such studies predated the discovery of *S. eubayanus* as a contributing parent to *S. pastorianus* and advances in RNA sequencing technologies [34]. More recent transcriptome studies focussed on the analysis of sub-sets of genes with specific roles in fermentation or in specific physiological conditions such as cold storage or responses to temperature [35–37]. Important findings regarding the gene expression of genes involved in maltose utilisation, carbohydrate metabolism, glycerol mobilisation, anaerobiosis, and protein biosynthesis have emerged from these studies, providing a road map for a more detailed transcriptome analysis of these industrial strains.

Here, we analysed the transcriptomes of the Group I, CBS1538 and the Group II, W34/70 *S. pastorianus* strains and specifically examined the contribution of the two sub-genomes to the gene expression patterns under fermentations conditions. Using *de novo* genome sequencing of these strains, we related gene expression patterns to gene copy number. We show that the relative expression of *S. cerevisiae* and *S. eubayanus* orthologues is directly correlated to gene copy number in the W34/70 strain. Gene copy number plays a smaller role in gene expression patterns in the CBS1538 strain, most likely due to the more limited genetic redundancy. Despite the reduced *S. cerevisiae* content in the CBS1538 strain, *S. cerevisiae* orthologues contribute significantly to biochemical pathways upregulated during fermentation, specifically to amino acid metabolism. Finally, a comparison of the transcript patterns in the CBS1538 and

W34/70 strains identified both common and unique gene expression patterns during fermentation.

Results

Fermentation profiles of CBS1538 and W34/70 strains

Two *S. pastorianus* strains, CBS1538 and W34/70, representative of Group I and Group II strains respectively, were chosen for analysis. Both strains displayed similar fermentation profiles in 10% wort up to Day 4 but thereafter, strain W34/70 fermented faster than strain CBS1538 and reached a lower final attenuation (Fig 1).

Chromosome composition of strains CBS1538 and W34/70

To compare the transcriptomes of CBS1538 and W34/70 lager yeasts, RNA was extracted on Day 2 and Day 4 of small-scale fermentations in 10% wort. These time points were chosen as previous studies have shown that this is the period of maximum metabolism during the fermentation [31]. RNA was also extracted from the same strains grown in minimal medium without amino acids to provide a baseline for comparison. To map the transcripts, the genomes of CBS1538 and W34/70 isolates were sequenced *de novo* and mapped to the annotated and fully assembled reference genome *S. pastorianus* 1483 (Group II strain) as well as to a combined genome assembled from the parental reference genomes *S. cerevisiae* and *S. eubayanus*. Both approaches yielded highly similar results, however since the *S. pastorianus* 1483 genome lacked some information for *S. cerevisiae* genes on chromosomes III and VII, the data from mapping to the combined parental genomes was used for this analysis. Information from genome sequencing and mapping of the two strains (Fig 2A and 2B) confirmed the absence of *S. cerevisiae* chromosomes II, III, IV, VI, VII, VIII, XII, XIII, XV and XVI in CBS1538 (Fig 2A). Furthermore, we observed alterations in the copy number of chromosomes and differences in hybrid chromosomes to the previously reported chromosome content of

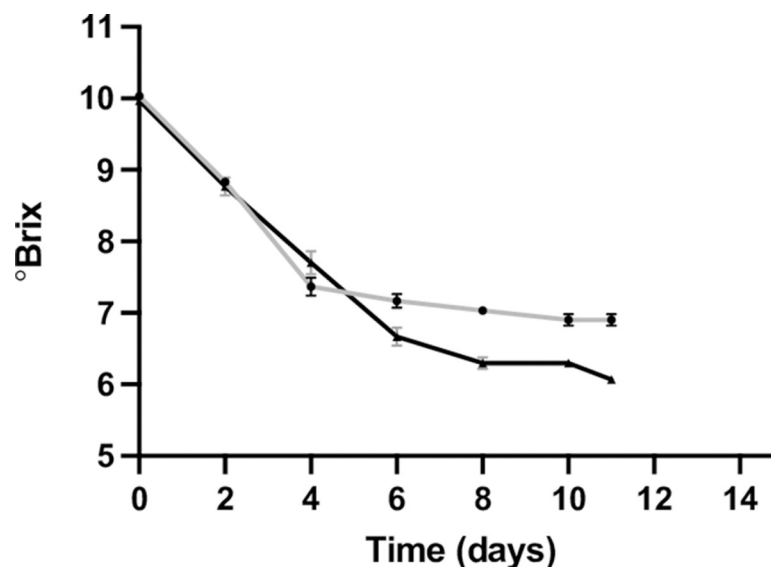


Fig 1. Fermentation profiles of the Group I CBS1538 and Group II W34/70 strains. Fermentations were carried out in 10% Wort at 13°C in 15 mL cylindrical tubes. Sugar consumption was measured with a Brixometer (°Brix). CBS1538: grey line, W34/70, black line. Error bars represent the standard deviations from the mean of triplicate fermentations. Supporting data is available in the supporting information file [S1 Data](#).

<https://doi.org/10.1371/journal.pgen.1010149.g001>

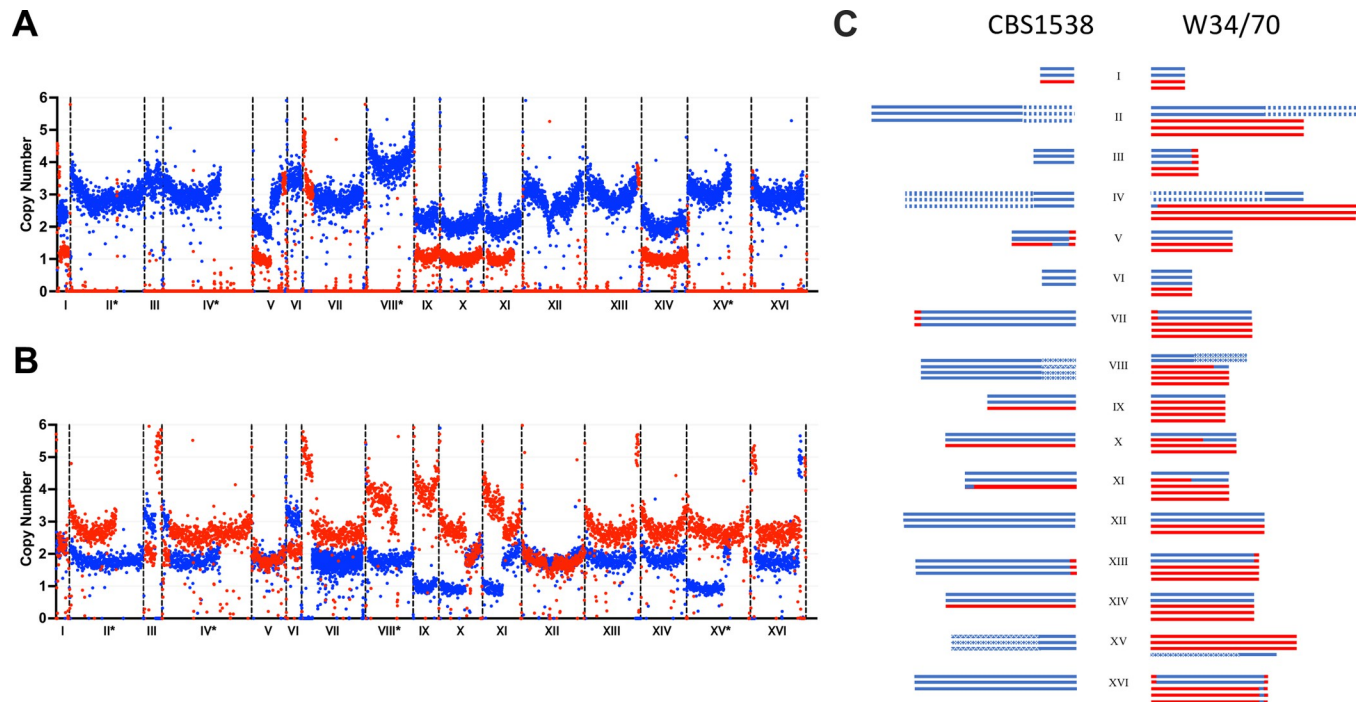


Fig 2. Chromosome copy number of CBS1538 (A) and W34/70 (B) genomes. Estimated copy number and chromosome types of *S. eubayanus* (blue) and *S. cerevisiae* (red) from *de novo* sequencing mapped to the combined parental genomes. The start of each chromosome is shown by a vertical dotted line. *, denotes translocations between *S. eubayanus* chromosomes II/IV, IV/II, VIII/XV, XV/VIII present in both strains. Hybrid chromosomes are evident from the change in copy number within a chromosome. C. Estimated types and copy number of *S. cerevisiae* (red) and *S. eubayanus* (blue) and hybrid (red/blue) chromosomes in CBS1538 and W34/70. Supporting data for A and B are available in the supporting information files [S2 Data](#) and [S3 Data](#) respectively.

<https://doi.org/10.1371/journal.pgen.1010149.g002>

both strains [16,17,23] (Fig 2C). The estimated chromosome copy numbers were 48 and 76 for CBS1538 and W34/70 respectively.

Gene expression profiles of CBS1538 and W34/70 strains during fermentation

Transcripts were either assigned as *S. cerevisiae* (Sc) or *S. eubayanus* (Se) based on sequence identity and chromosome assignment. A total of 11335 transcripts/Open Reading Frames (ORFs) were detected for the W34/70 strain of which 5083 were assigned as *S. eubayanus* and 6252 as *S. cerevisiae* (ratio Se:Sc = 0.81:1) (Table A in [S1 File](#)). For the CBS1538 strain, a total of 6798 transcripts were detected, 5163 were *S. eubayanus* and 1635 were *S. cerevisiae* (ratio Se:Sc = 3.16:1). These values agree well with the previously determined gene count for the two strains [38]. We examined the differential gene expression (DEG) of transcripts mapped to each strain in three different experimental conditions (minimal medium without amino acids, wort Day 2 and wort Day 4) (Tables B-D in [S1 File](#)). In the W34/70 strain, the greatest number of genes displaying DEG was detected in the comparison of cells grown in minimal medium and wort on Day 2 where 42% of transcripts were differentially expressed, while 28% of genes were differentially expressed between growth in minimal medium and in wort on Day 4. A much lower level of DEG was detected between cells grown in wort on Days 2 and 4 where just 14% of genes were differentially expressed. The pattern of DEG in CBS1538 was similar to that observed with the W34/70 strain with 32% of transcripts differentially expressed between

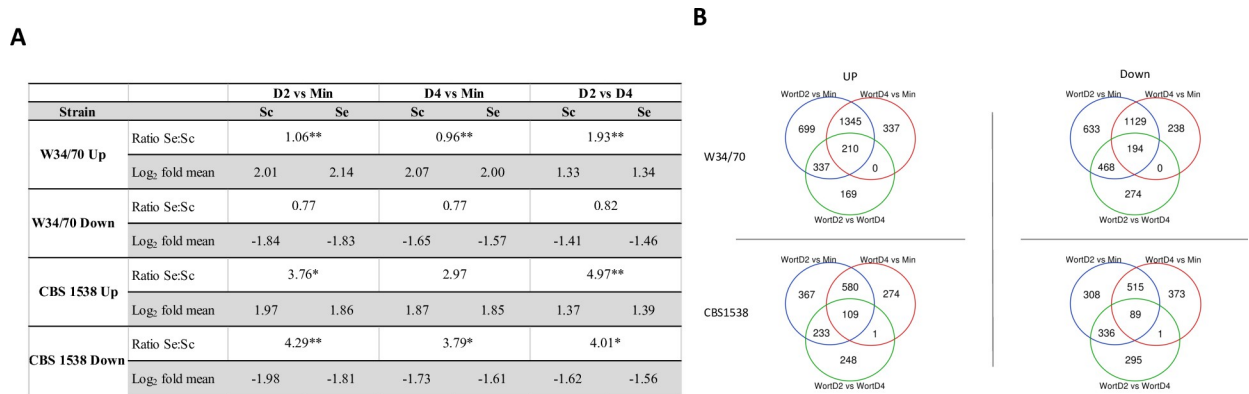


Fig 3. Ratio of *S. cerevisiae* (Sc) and *S. eubayanus* (Se) in differentially up-regulated and down-regulated transcript pools (log₂-fold change ≥ 1 or ≤ -1) and mean log₂-fold changes under the three conditions compared in CBS1538 and W34/70. A. Ratios of differentially expressed Se:Sc genes with significant differences from the expected are shown with an asterisk, $p \leq 0.001^{}$, $p \leq 0.05^{*}$. B. Venn diagrams showing the relationship of differentially up-regulated and down-regulated transcript pools under the comparisons of the three conditions shown.**

<https://doi.org/10.1371/journal.pgen.1010149.g003>

growth in minimal medium and wort on Day 2, 23% between growth in minimal medium and wort on Day 4 and 16% between growth in wort on Days 2 and 4.

As *S. pastorianus* contains both *S. cerevisiae* and *S. eubayanus* sub-genomes, we examined the DEG patterns for *S. cerevisiae* or *S. eubayanus* alleles under the three different experimental conditions. In the W34/70 strain, *S. eubayanus* genes are significantly over-represented in the differentially upregulated gene pool in all three comparisons but were expressed at the expected ratio in the differentially downregulated gene pool (Fig 3A). Interestingly, on Day 2 relative to Day 4 in wort, there are twice as many Se alleles upregulated compared to Sc alleles. In the CBS1538 strain, Se alleles are overrepresented in both the upregulated and downregulated gene pools in cells in wort on Day 2 relative to Day 4 and in wort on Day 2 relative to minimal medium and in the downregulated gene pool in wort on Day 4 relative to minimal medium.

The mean log₂-fold change for *S. cerevisiae* and *S. eubayanus* alleles in both the W34/70 and CBS1538 strains are extremely consistent in all three comparisons indicating that both CBS1538 and W34/70 strains, and their Se and Sc sub-genomes, respond similarly to the physiological conditions. The mean log₂-fold change and the range of DEG on Days 2 and 4 wort was less than that observed for differential expression between minimum medium and wort on Days 2 and 4 for both strains with the exception for down regulated genes on Days 2 and 4 in CBS1538 (Figs 3A and S1).

We next examined the relationship of DEG under the three conditions in the two strains. As might be expected, a large number of genes are differentially expressed between growth in minimal medium and in wort on Days 2 and 4 in both the CBS1538 and W34/70 strains (Fig 3B). While fewer in number, there is an overlap in the genes differentially expressed in wort Day 2 relative to Day 4 and in wort Day 2 relative to minimal medium. On the other hand, there are no genes differentially expressed that are specific to wort Day 2 relative to Day 4 and to wort Day 4 relative to minimal medium in W34/70 and just 1 gene in this category in CBS1538.

Only a small number of genes are commonly upregulated or downregulated in all three conditions in the CBS1538 and W34/70 strains respectively (Fig 3B and Table E in S1 File). The commonly upregulated gene pool in W34/70 ($n = 210$), contains 23 pairs of Sc and Se orthologous alleles, including, *ADH1*, *ADH5*, *PGK1*, *GPH1*, *HXK1*, *ENO1*, that are central to carbohydrate metabolism and *ILV6*, required for branched chain amino acid biosynthesis.

Likewise, both Sc and Se alleles of *LEU4*, *LYS1* and *SPD1* are upregulated in all three conditions in CBS1538. Orthologous gene pairs are also observed in the commonly downregulated gene in both strains (Table E in [S1 File](#)).

Gene ontologies enriched in CBS1538 and W34/70 strains under fermentation conditions

Gene ontology analysis of the differentially expressed genes in the three conditions in CBS1538 and W34/70 revealed an enrichment in both carbohydrate and amino acid metabolism on Day 2 of fermentation in the upregulated gene pool ([Fig 4A](#)). Genes required for the utilisation of all major sugars including sucrose, fructose, galactose as well as pentose sugars, along with the genes required for glycolysis and pyruvate metabolism, are all upregulated in wort Day 2 relative to minimum medium and a subset of these genes are upregulated on Day 2 relative to Day 4 and on Day 4 relative to minimum medium in the W34/70 strain. Except for genes involved in starch and sucrose metabolism, the observed enriched gene set for carbohydrate metabolism in the W34/70 strain were not enriched in the CBS1538 strain ([Fig 4A](#)).

Genes for the biosynthesis and metabolism of amino acids are enriched during the fermentation in wort in both CBS1538 and W34/70 strains ([Fig 4A](#)). In addition to the enrichment of genes for these major metabolic activities in the two strains, we observed differences in specific gene ontologies between the two strains, with genes involved in glycerolipid metabolism, porphyrin metabolism, ABC transporters, peroxisome activity and longevity upregulated in the W34/70 strain but not the CBS1538 strain.

Genes involved in nucleic acid metabolism, DNA repair and DNA replication are downregulated under fermentation conditions and on Day 2 relative to Day 4 in wort in the W34/70 strain with a subset of these genes similarly regulated in the CBS1538 strain ([Fig 4B](#)). Likewise, genes involved in major anabolic pathways such as protein synthesis including genes required for ribosome biogenesis and rRNA, mRNA and tRNA processing are downregulated in wort and on Day 2 relative to Day 4 in both strains. Interestingly, a set of genes involved in fatty acid, glycerophospholipid and steroid metabolism, required for the production of membrane components, are uniquely downregulated in the CBS1538 strain ([Fig 4B](#)).

We examined the contribution of the Sc and Se sub-genomes to these enriched pathways. Se and Sc alleles were observed to contribute at the expected Se:Sc ratio or in a higher proportion than expected to the differentially regulated gene pool ([Fig 4A and 4B](#)). Se alleles contribute in a higher proportion than expected on Days 2 and 4 of fermentation to some gene ontologies in the W34/70 strain ([Fig 4A](#), purple-blue gradient shaded boxes). Conversely, we observed that Sc alleles contribute in a greater proportion than expected to amino acid metabolism in the CBS1538 strain under all three conditions ([Fig 4A](#), red gradient shaded boxes). For some specific pathways, we observed that Se alleles accounted for all the differentially regulated genes, for example Se alleles are used exclusively in histidine metabolism on Day 2 relative to Day 4 in the W34/70 strain and to phagosome associated genes in wort on Day 2 relative to Day 4 and to sphingolipid, butanoate and thiamine metabolism on Day 4 relative to minimum medium in CBS1538 ([Fig 4A](#), dark blue boxes).

For the downregulated genes, we observe that Sc alleles contribute to gene ontologies at the expected ratio, and in some pathways, in a higher proportion than expected to the ratio of Sc:Se for the whole genome in both strains while Se alleles exclusively contribute to three categories, namely one carbon pool by folate, fatty acid and steroid biosynthesis in the CBS1538 strain ([Fig 4B](#), dark blue boxes).

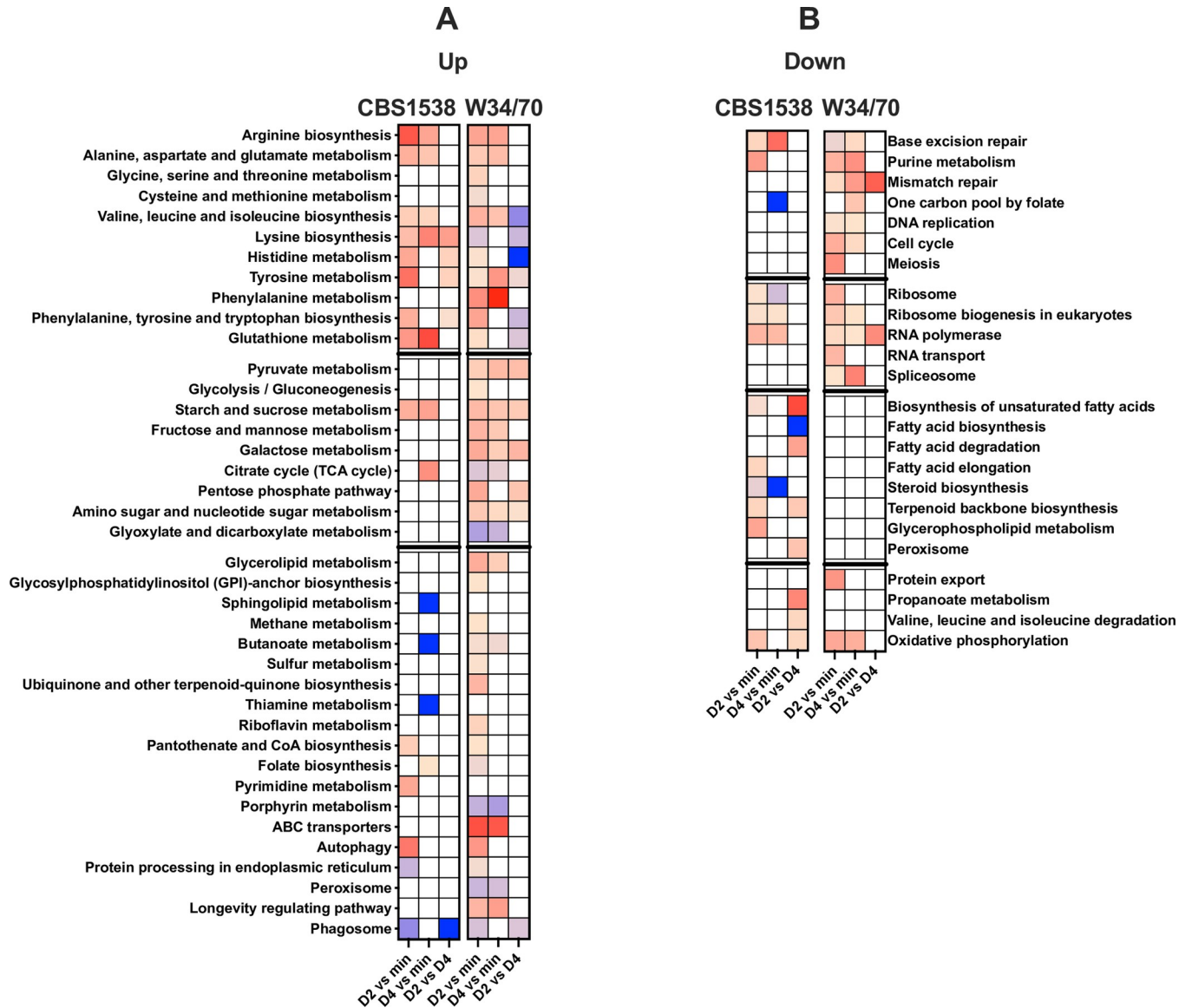


Fig 4. Gene ontology analysis of differentially expressed genes in W34/70 and CBS1538 strains. Genes displaying a log₂-fold change ≥1 (A) or ≤ -1 (B) were examined for enriched gene ontologies using Cluego. The heat map shows deviations from the expected ratio of Sc to Se alleles; expected ratio beige, >expected ratio; beige to red gradient, < expected ratio; beige to blue gradient. Blank cells, no enrichment. CBS1538 expected ratio Sc:Se = 0.24:0.76, W34/70, Sc:Se = 0.55:0.45.

<https://doi.org/10.1371/journal.pgen.1010149.g004>

The effect of gene dosage on gene expression profiles

As *S. pastorianus* strains are aneuploid in nature and have different copy numbers of *S. cerevisiae* and *S. eubayanus* chromosomes, we were interested to determine if gene expression profiles were influenced by the copy number of Sc and Se alleles. The copy number for each Sc and Se gene was calculated from the sequence coverage from the *de novo* genome sequencing. The DEG of Sc and Se orthologues within each strain was determined for the three experimental conditions and was correlated to the Sc:Se gene copy number ratio. The data for Sc and Se orthologues expressed on Day 2 is shown in Fig 5. The data for Day 4 and minimal medium is shown in S2 Fig. Sc and Se orthologues are present in varying ratios in the W34/70 strain (Fig 5A). The analysis of some

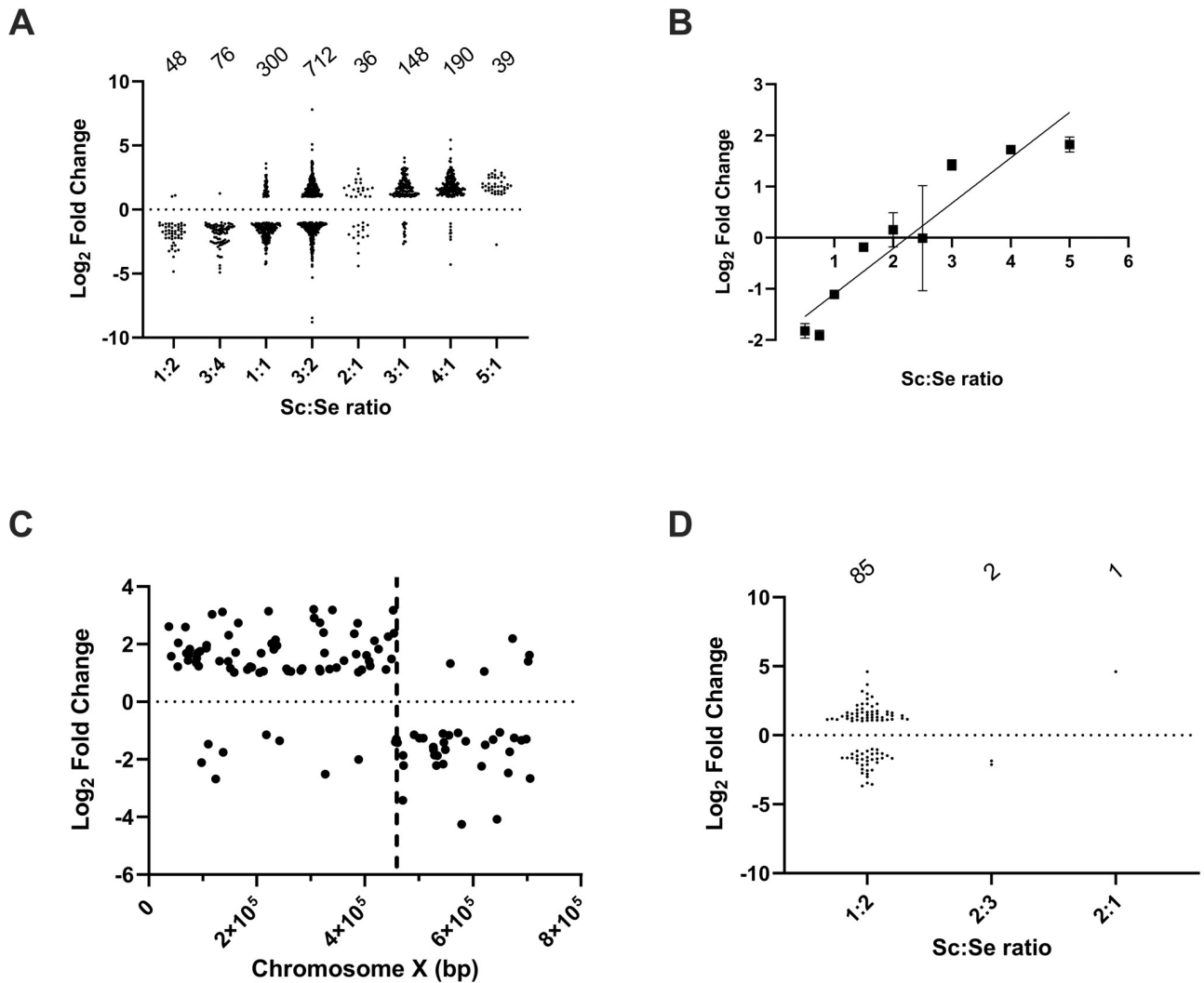


Fig 5. Effect of gene copy number on gene expression patterns. The DEG of Sc to Se orthologues were grouped according to the Sc to Se ratio of gene copy number in W34/70 (A) and CBS1538 (D). Data for Day 2 in wort is shown. Data for Day 4 and minimal medium are shown in the S2 Fig. The number of paired orthologues at each copy number ratio for genes showing $\log_2 \leq -1$ and $\log_2 \geq 1$ is shown above the graphs. B. Correlation of differential expression of Sc to Se gene orthologues in W34/70 with gene orthologue copy number ratio. The mean of \log_2 fold changes in expression between Sc and Se orthologues is plotted against Sc:Se gene orthologue copy number ratio. Error bars represent the standard error of the mean. C. \log_2 -fold gene expression difference between Sc and Se orthologues on chromosome X in W34/70. The dotted line marks the recombination site (*THD2*) on the hybrid chromosome. Sc:Se ratio to left of vertical dotted line is 3:1 and after 1:1.

<https://doi.org/10.1371/journal.pgen.1010149.g005>

1555 Sc and Se orthologues indicates that the DEG of Sc to Se alleles was positively correlated to the ratio of the copy number of Sc:Se alleles (Fig 5B, $p < 0.0001$, $R^2 = 0.90$). As the Sc gene copy number increases, so too does the levels of Sc transcripts and vice versa for Se alleles. At an Sc:Se ratio of 1:2 and 3:4, Se transcripts predominate while at ratios, 3:1, 4:1 and 5:1 Sc alleles predominate (Fig 5A). Interestingly, at a 1:1 ratio, Se alleles predominate in the W34/70 strain while at 3:2, 2:1 there is almost equal levels of both transcripts. The relationship between the DEG of Sc to Se orthologues and the ratio of gene copy number is observed for genes on both parental and hybrid chromosomes as exemplified in Fig 5C for chromosome X. The W34/70 strain has four copies of

chromosome X, two Sc, one Se and one hybrid chromosome (Sc/Se) with the recombination point occurring at the gene *THD2* (YJR009C) (Fig 2C). The ratio of Sc:Se to the left of *TDH2* is 3:1 and to the right 1:1 (Fig 2C). For genes to the left of *TDH2*, expression of Sc alleles predominates due to the higher Sc copy number while to the right, Se alleles predominate despite the equal copy number (Fig 5C). The correlation of DEG patterns to orthologue gene copy number ratio was observed in all three experimental conditions (S2 Fig).

The analysis of the gene expression of Sc and Se orthologues in the CBS1538 strain presented a different scenario. Firstly, due to the reduction of the Sc genome content in the CBS1538 strain, the number of total orthologues is much smaller ($n = 237$). Secondly, the variations in copy numbers are much smaller, with most orthologues with a \log_2 fold change ≥ 1 or ≤ -1 present in a ratio of Sc:Se, 1:2 (Fig 5D). There are just two alleles present in a 2:3 ratio and one at a 2:1 ratio (Fig 5D). Surprisingly, we see a different expression pattern to what was observed in W34/70; at the Sc:Se 1:2 ratio, a majority of Sc alleles display higher transcript levels than the orthologous Se alleles, although the reverse expression pattern is evident for some orthologues.

Comparison of the transcriptomes of CBS1538 and W34/70 strains

To directly compare the steady state mRNA landscapes in the CBS1538 and W34/70 strains, we generated a consolidated transcriptome for each strain by obtaining the sum of Sc and Se allele transcripts encoded for each gene (Table F in S1 File). We compared the expression of this gene set in the two strains under the three different experimental conditions (Fig 6 and Table G in S1 File). There is a high degree of overlap between the differentially expressed gene pool in both strains suggesting that both strains respond similarly to the environmental conditions (Fig 6A). Growth in minimal medium invoked the greatest condition-specific DEG between the two strains but a substantial number of genes were also specifically expressed in wort on Day 2 and Day 4 respectively. We identified 319 CBS1538-specific genes that are upregulated in all three conditions and likewise 299 W34/70-specific genes. The list of condition specific and strain-specific genes is shown in Table H in S1 File. An analysis of gene ontologies associated with the differentially expressed gene set reveals differences between the CBS1538 and W34/70 strains (Fig 6B). Firstly, we observed that genes associated with ribosome biosynthesis are upregulated on Day 2 and 4 in wort in the W34/70 strain while genes associated with metabolism of certain amino acid, pentose phosphate pathway and oxidative phosphorylation are also upregulated on Day 2 in this strain. Genes associated with DNA repair and protein processing and transport are upregulated in minimal medium in W34/70. Conversely, genes associated with membrane biosynthesis, sugar metabolism and amino acid metabolism are upregulated in minimal medium in CBS1538, and a subset of these genes are also upregulated on Day 4 in wort. Surprisingly, there are no gene ontologies enriched on Day 2 in the CBS1538 strain (Fig 6B).

As the gene copy number for the sum of gene orthologue transcripts still varied between the two strains and was shown to influence gene expression profiles within a strain, we looked to see if gene copy number influences the DEG observed between CBS1538 and W34/70. To do this, we compared the \log_2 fold changes in gene expression between the two strains to the ratio of the total gene copy number (sum of all Se and Sc alleles) for each gene. The data for gene expression on Day 2 in wort is shown in Fig 6C. There is a significant positive correlation between \log_2 fold change and ratio of total gene copy number between the two strains ($p < 0.001$, $R^2 = 0.8$). This correlation was observed across all three experimental conditions (S3 Fig). Thus, as observed for gene expression within a strain, gene dosage plays a significant role in the DEG between the CBS1538 and W34/70 strains.

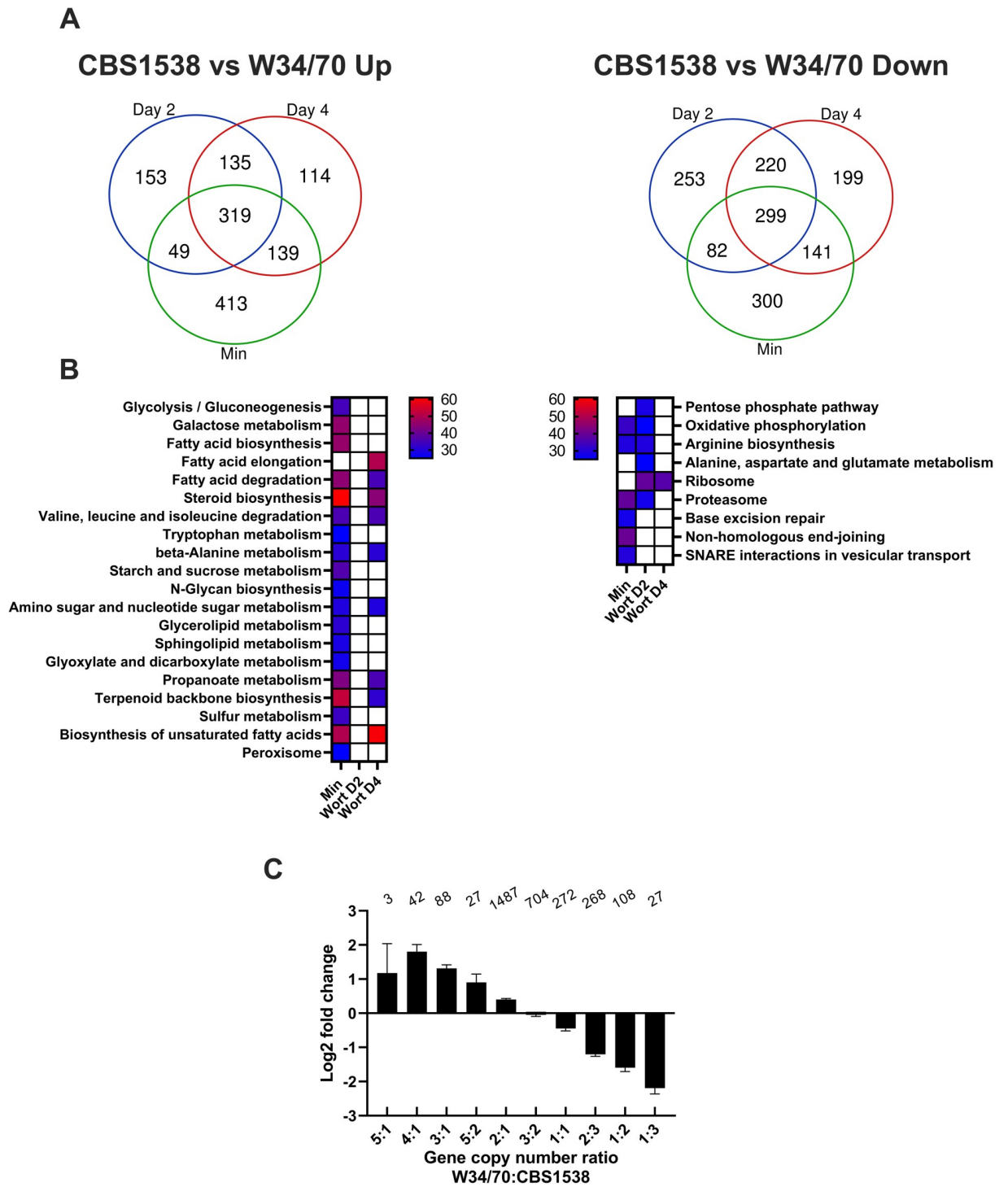


Fig 6. Comparison of gene expression in CBS1538 and W34/70 strains. Transcript reads for Sc and Se alleles were summed in CBS1538 and W34/70 respectively to create a combined transcriptome. Gene expression profiles in the three conditions, growth in minimal medium, in wort on Day 2 and in wort on Day 4 were compared between the two strains. **A.** Venn diagram of genes upregulated in the CBS1538 and W34/70 strains. **B.** Gene ontologies (GOs) enriched in the CBS1538 and W34/70 strains respectively. Left panel, GOs upregulated in CBS1538, right panel, GOs upregulated in W34/70. The parameters for selection of GO terms were as follows: p-value less than 0.05, minimum number of four genes per pathway, pathways containing more than 25% of the genes differentially regulated. Heat map shows the percentage of genes associated with each GO term: gradient blue to red, where blue represents a minimum of 25% of associated genes and red represents a maximum of 60% of associated genes. **C.** Relationship of gene copy number and gene expression in W34/70 and CBS1538. The ratio of the gene copy number for the sum of all alleles (Sc+Se) was compared to Log₂ fold differences in gene expression in W34/70 and CBS1538. The number of genes in each ratio category is shown above the bars. Data for Day 2 in wort is shown. Data for Day 4 and minimal medium are shown in the S3 Fig.

<https://doi.org/10.1371/journal.pgen.1010149.g006>

Discussion

The origin and evolution of lager yeasts are widely debated, and several hypotheses have been proposed to account for the divergence in chromosome composition and copy number between strains [20,22,23]. The major distinction between Group I and II strains is the loss a significant portion of the *S. cerevisiae* sub-genome through the loss of whole chromosomes in the Group I strains. In addition, differences in chromosome composition, copy number and the number and type of hybrid chromosomes exist between strains within each group. Despite these differences, strains from both groups ferment sugars and produce aromatic beers, with individual strains displaying variations in the final aromatic volatile profile [24]. A major difference between Group I and II strains lies in the ability to ferment maltotriose. While all Group II strains can uptake and ferment maltotriose, only a subset of Group I strains have the necessary transporters to import this trisaccharide [39–41]. Here we have examined the gene expression profiles during fermentation in wort and in growth in minimal medium of representative Group I and II strains. The Group I strain used in this study, CBS1538, does not ferment maltotriose. We confirm here the superior fermentation rate of the Group II strain W34/70: the two strains consume sugars at the same rate up to Day 4 and thereafter the W34/70 strain fermented at a faster rate.

We were interested in understanding how the two sub-genomes, present in the two lager yeast strains, contribute to the overall transcriptome during fermentation and furthermore what are the consequences for the reduction in the *S. cerevisiae* sub-genome to the gene expression patterns between CBS1538 and W34/70. The complex genomes of lager yeasts pose several challenges to the analysis of the transcriptomes of these strains. At present, just one *S. pastorianus* Group II genome, that of strain CBS1483, has been fully annotated and assembled into chromosomes and this serves as the reference genome for *S. pastorianus* at NCBI [20], although up to 16 annotated genomes are also available <https://www.ncbi.nlm.nih.gov/genome/browse/#!/eukaryotes/342/>. Due to differences in chromosome copy numbers, and chromosomal rearrangements, the reference genome may not be the ideal genome for transcript mapping, for example, it differs from the two strains used in this study, as it lacks hybrid chromosomes III and VII. There are also chromosome copy number differences between the strains used here and the reference strain.

To avoid such issues, we re-sequenced the W34/70 and CBS1538 strains which were mapped to a *de novo* generated genome consisting of the combined reference genomes of the parental strains *S. cerevisiae* and *S. eubayanus*. The copy numbers for each chromosome and genes therein were determined from the sequence coverage depth from *de novo* sequencing. The RNA transcripts were also mapped to this combined genome.

We noted some differences in the chromosome composition of both W34/70 and CBS1538 to the published data [16,23,42]. Specifically, we noted one extra copy of chromosomes Sc III, VII, XIV, and XVI in relation to the most recently published W34/70 genome sequence [23]. Additionally, we reconfirm the presence of a hybrid chromosome, ScVIII/SeVIII, recombining at *YHR165C*, that was previously identified in the W34/70 strain and in another Group II strain, 7012, by our group [12,21] but which is not documented in other published sequences [16,23]. For the CBS1538, we noted just two copy number differences, an extra copy of both Se VI and SeVIII. Chromosome copy number differences are also noted between published genomes for W34/70 [16,23,42]. The differences in copy numbers between studies may arise from differences in methods used to determine sequence coverage depth or may reflect genuine differences between strain isolates. The lager yeast strains emerged just some 500–600 years ago and thus may still be experiencing genomic flux in this early stage of evolution. We previously showed that the

genomes of lager yeasts are dynamic and can undergo chromosome rearrangements to produce new hybrid chromosomes as well as chromosome copy number changes following exposure to stress such as high temperatures [43]. Furthermore, fermentations carried out in high specific gravity wort (22°P) at an ambient temperature of 20°C, which is higher than that used for routine fermentations (13°C), led to chromosome copy number changes in a single round of fermentation [43]. Thus, differences in propagation and culturing of *S. pastorianus* strains may contribute to the differences observed between strain isolates.

The analysis of the gene expression patterns reveals that the CBS1538 and W34/70 strains respond similarly to the physiological conditions imposed with both strains showing similar mean log₂ fold changes under all compared conditions. Interestingly, we observed that the Sc and Se sub-genomes are differentially utilised under the different physiological conditions with Se alleles contributing significantly to gene expression on Day 2 of fermentation in both strains and on Day 4 in the W34/70 strain.

Gene ontology analysis revealed the upregulation of pathways associated with carbohydrate and amino acid metabolism on Day 2 of fermentation although there was less enrichment of genes associated with sugar metabolism in the CBS1538 strain. The upregulation of carbohydrate and amino acid metabolism in the early stages of fermentation is consistent with what is known about the overall metabolic activity of yeast during fermentation. Cells undergo 1–2 doublings during the first three days of fermentation and thereafter, cell numbers remain unchanged or slightly decrease. The lack of representation of genes associated with carbohydrate metabolism in the CBS1538 strain, except for those associated with starch and sucrose metabolism, distinguishes the two strains and appears to reflect differences in metabolic activity. As overall sugar consumption is similar in the two strains on Day 2, it is possible genes associated with carbohydrate metabolism in the CBS1538 strains are induced but did not reach the cut off threshold for analysis. It does not appear that sugar metabolism is slowed down in this strain as otherwise we may have expected to see such genes upregulated on Day 4. The observed differences in carbohydrate metabolism may reflect the previously noted differences in the types and copy numbers of genes encoding for maltose and maltotriose transporters between the two strains [39–41].

The upregulation of genes associated with amino acid metabolism is significant as the secondary metabolites associated with flavour such as higher alcohols and esters are produced from the catabolism of amino acids. The gene ontology also revealed differences in metabolism between the two strains. For example, genes associated with methane and sulphur metabolism, which can also contribute to flavour profiles, are upregulated in W34/70 but not in CBS1538.

Gene ontology analysis confirm the over representation of Se alleles in the upregulated genes in the W34/70 strain and additionally identified specific pathways where either Sc or Se alleles are exclusively enriched. Considering the reduced Sc content in the CBS1538 strain, it is surprising that Sc alleles appear to be over-represented in genes enriched in several pathways. Specifically, we see Sc alleles contributing to Arginine and Lysine biosynthesis and Tyrosine metabolism at a higher level than expected from the Sc:Se ratio for the whole genome. We also observed that Sc alleles are more likely to be downregulated than Se alleles in both CBS1538 and W34/70 strains however specific pathways here also contain exclusively Se alleles.

Previous studies in the haploid *S. cerevisiae* showed that the steady state levels of mRNA transcripts are tightly controlled to maintain homeostasis. Disruption of mRNA turnover, for example by deletion of the major 5' to 3' exonuclease *XRN1*, is compensated by changes to RNA polymerase II transcription to restore homeostasis [44,45]. Conversely, increased transcription rates resulting from increased gene copy number lead to compensatory changes in mRNA turnover to preserve the expected steady state levels of mRNAs [46]. However, a recent analysis of laboratory and wild *S. cerevisiae* strains revealed no evidence of dosage

compensation and instead described a direct correlation between gene copy number and gene expression levels [47].

Here we show that the aneuploid nature of the *S. pastorianus* genomes directly contributes to the gene expression patterns during fermentation. We observed a positive correlation between orthologue copy number and the steady state levels of orthologue transcripts in the W34/70 strain: as the copy number of Sc or Se genes increases so too does the associated levels of transcripts. Interestingly, when the orthologue gene copy number ratio is 1:1, we observed that the Se alleles showed higher levels of transcripts than the Sc alleles in all conditions tested. Surprisingly, the pattern of usage of orthologues is different in the CBS1538 strain. While there is less variation in the ratios of Sc:Se genes with most orthologues present in a 1:2 ratio, here we observed that Sc orthologues produce relatively higher levels of transcripts than the Se counterpart although there is a greater variation in the distribution of gene expression patterns amongst orthologues at this ratio in the CBS1538 strain. At the same ratio in the W34/70 strain, Se alleles predominate. The over-representation of *S. cerevisiae* alleles in some biochemical processes such as amino acid metabolism may explain the selective retention of *S. cerevisiae* chromosomes in the CBS1538 strain as there may have been a selective evolutionary pressure to retain specific genes required during fermentation.

The *S. eubayanus* parent of the lager yeasts is a cryotolerant strain and thus it might be hypothesised that Se alleles will be favoured during the fermentations which are carried out at 13°C. We did observe a greater contribution of Se alleles to the transcriptome in the W34/70 strain during fermentation conditions, however, it is not universal, and this contribution is tempered by gene copy number effects. A transcriptome analysis for the Group I strain, CBS1513, had also observed an overrepresentation of *S. eubayanus* when cultures were grown at low temperatures [37].

To compare the overall steady state levels of protein encoding transcripts in the CBS1538 and W34/70 strains, we created combined transcriptomes for each strain by summing the Sc and Se transcripts that encode for the same protein. Using this approach, we observed that that overall transcription patterns were similar between the two strains indicating that despite the significant differences at a genome level and pathways of evolution, the two lager yeast groups display similar transcription patterns in fermentation conditions. Nevertheless, gene ontology analysis identified some unique transcription patterns. Consistent with what was observed for the differential expression of Sc and Se alleles within a strain, the gene expression patterns in both strains were significantly influenced by gene copy number. Thus, it might be anticipated that gene expression profiles may differ in other *S. pastorianus* strains due to the variations in chromosome copy number between strains.

Our analysis of the expression of gene orthologues within a strain and between strains indicates that the aneuploid genomes of the lager yeasts create complex patterns of gene expression during fermentation and that gene copy number plays a crucial role in the gene expression patterns both within a strain and between strains.

Materials and methods

Yeast strains and growth conditions

The Group I strain, CBS 1538, was obtained from the Collection de Levures d'Interet Biotechnologique, Paris, France and the Group II strain Weihenstephan 34/70 (W34/70) was kindly supplied by Dr. Jurgen Wendland, Geisenheim Hoch Universitat, Germany. For propagation, strains were grown in YPDM medium (1% (w/v) yeast extract, 2% (w/v) peptone and 1–2% (w/v) of both dextrose and maltose at 20–25°C; 2% YPDM contains 1% dextrose and 1% maltose.

Small-scale fermentations (10mL) were carried out in 10% wort containing 1mM ZnSO₄ (Spraymalt, Brewferm, Amber 18EBC, Brouland, Belgium) in 15 mL glass test tubes.

Cells were first propagated in 200 mL of 4% YPDM at 25°C for two days and were pitched into wort at a cell density of 1.5x10⁷/mL. For the small scale fermentations, 10mL aliquots of the starting fermentation culture were added to 15mL glass tubes, each fitted with a water trap airlock attached to a bung. Individual tubes were used for sampling at time intervals. All fermentations were carried out at 13°C without shaking. After sampling, the specific gravity of the wort was measured using a refractometer (HANNA, Romania). The residual cultures from the small scale fermentations were centrifuged and the cells saved for RNA extraction.

Cells were also grown in 2% YPDM overnight, washed with sterile water and inoculated at a cell density of 1x10⁶ cells/mL into 60 mL of minimal medium (0.17% Yeast Nitrogen Base w/o amino acids and ammonium salts supplemented with 1% of dextrose and 1% of maltose, and 0.5% (NH₄)₂SO₄ as a nitrogen source. The cultures were grown at 20°C, in triplicate, and cells were harvested in the exponential phase.

DNA extraction, *de novo* genome sequencing and copy number variation analysis

DNA extraction was carried out following the protocol described by [48]. Briefly, yeast cells were grown in an overnight culture of 5 ml of 2% YPDM. Cells were pelleted and washed with distilled water and centrifuged. The pellet was resuspended in 500 µL of 0.9 M sorbitol and 0.1 M EDTA, pH 7.5 and 30 µL of a solution of lyticase (1.5 mg/mL) was added and incubated for 20–40 min. After this, spheroplasts were recovered centrifuged and resuspended in 500 µL of 50 mM Tris HCl and 20 mM EDTA, pH 7.4 and 13 µL of 10% SDS. Samples were incubated for 65°C for 5–10 min. 200 µL of 5M potassium acetate were added and tubes were incubated on ice for 10 min. After centrifugation, 700 µL of 100% isopropanol was added to the supernatant, the samples incubated at room temperature for 10 min and then recentrifuged. The pelleted DNA was washed twice with 70% ethanol, air-dried at room temperature, and resuspended in 50 µL of water.

Total DNA was treated with RNase A (10 µg/mL) for 30 min at 37°C. DNA was precipitated and resuspended in 50 µL of water. The quality of the DNA was determined by agarose gel electrophoresis and by determining the OD₂₆₀:OD₂₈₀ ratio using a Nanodrop spectrophotometer (Epoch 2, Agilent BioTek). *De novo* genome sequencing was carried out by Novogene (www.novogene.com) with Illumina technology on paired-end reads (150bp). The genome sequences were deposited into the SRA database at the National Centre for Biotechnology Information <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA810967>.

After trimming paired-end reads with trim_galore (version 0.6.6, calling cutadapt version 2.10), the reads were mapped with bowtie2 (version 2.4.2) [49,50] to the parental genomes of *S. pastorianus* CBS1483, (assembly ASM1102231v1) <https://www.ncbi.nlm.nih.gov/genome/?term=Saccharomyces+pastorianus>, or to a combined genome of *S. cerevisiae* and *S. eubayanus* derived from *S. cerevisiae* S288C, assembly R64 <https://www.ncbi.nlm.nih.gov/genome/?term=Saccharomyces+cerevisiae> and *S. eubayanus* (SEU3.0) <https://www.ncbi.nlm.nih.gov/genome/?term=Saccharomyces+eubayanus>.

In addition to the ‘—fast-local’ presets, parameters were set to limit the insert size to lengths between 200 and 1000 basepairs. Mapped reads were filtered through samtools (version 1.11) to exclude reads that were not mapped in pairs and were not the primary alignment (-F 268) and to include only reads with high mapping quality (-q 44).

For hybrid chromosomes, genes were assigned as *S. cerevisiae* or *S. eubayanus* based on their location relevant to known recombination sites and their % identity to the parental

genomes. Reads from *de novo* DNA sequencing, mapped against the combined genomes of the reference strains of *S. cerevisiae* and *S. eubayanus* were transformed into sorted BAM files using samtools and the data was extracted as reads/500 bp and then normalized by the size of the library (total number of reads) to give an estimated chromosome copy number [51].

RNA extraction and sequencing

RNA extraction was carried out following the protocol previously described [52] but with some modifications. Cells were resuspended in 400 μ L of buffer containing 50 mM of NaOAc, pH 5.2, 10 mM EDTA pH 8.0 and 40 μ L of 10% SDS, 400 μ L of Phenol-Chloroform-Isoamyl alcohol (PCA; 25:24:10), pH 4.2, and 100 μ L of acid washed glass beads were added. The tubes were placed in a heat block at 65°C for 10 min and vortexed intermittently. The samples were placed on ice for 5 min. After centrifugation, the supernatant was re-extracted with PCA, pH 4.2 and then once with chloroform. Nucleic acids in the supernatant were precipitated, washed twice with 70% ethanol and then resuspended in 100 μ L RNase-free H₂O. The RNA concentration was measured using a Nanodrop spectrophotometer and RNA integrity was checked by Agilent 2100 Bioanalyzer as per manufacturer's instructions.

RNA sequencing was conducted on cDNA libraries using the HiSeq4000 Illumina Platform at the Genomic Technologies Core Facility at the University of Manchester. The libraries were generated using the TruSeq Stranded mRNA assay (Illumina, Inc.) according to the manufacturer's protocol. The RNAseq dataset was deposited to the GEO database at NCBI <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE197654>.

As a reference genome for the mapping, the combined *S. cerevisiae* and *S. eubayanus* genomes were used but with variations as present in the re-sequenced CBS1538 and W34/70 strains. For that, consensus Fasta sequences were generated from the WGS data mapped to *S. cerevisiae* and *S. eubayanus* through an in-house Perl script. To transfer gene annotation from those genomes, insertions and deletions were ignored and only nucleotide replacements were taken into account. The paired RNA-Seq data were mapped against these strain-specific reference genomes using STAR [53]. Multimapping reads are dealt with through the outSAMmultNmax parameter set to 1 and outMultimapperOrder set to 'random'. Additionally, only the highest quality alignments were kept by filtering with samtools (-q 255). To aggregate reads counts, the tool "featureCounts" was used on the mapped data together with the gene annotation from the *S. cerevisiae* and *S. eubayanus* genome [54]. With the -B and the -P options set, only read counts that have both ends aligned within a distance of 50–600 basepairs were considered.

Read counts from the RNA mapping were uploaded into iDEP9.1 [55]. Data were transformed for clustering and PCA by EdgeR using a minimum CPM of 0.5 in 1 library and adding a pseudocount of 4. PCA analysis was used to check the reproducibility of replicates. Outlier replicates were removed and data reprocessed. Differentially Expressed Genes (DEG) were calculated using DESeq2 with a False Discovery Rate (FDR) cutoff of 0.05 and a minimum log₂ fold change of ≥ 1 or ≤ -1 . Venn diagrams were made using the online tool (<http://bioinformatics.psb.ugent.be/>). The copy number for each gene was extracted from the reads coverage of the *de novo* genome sequencing of the *S. pastorianus* strains.

Enrichment of DEG was carried out using ClueGO [56]. Parameters used on ClueGO were set as follows: *Saccharomyces cerevisiae* S228C was used as a Load Maker List. KEGG Ontology was used for the analysis with the following filters, p-value less than 0.05, minimum number of four genes per pathway, pathways containing more than 25% of the genes regulated, Kappa score was set at 0.4. A two-sided hypergeometric test and a Benjamini-Hochberg pV correction was applied. *S. eubayanus*-like allele names were translated into *S. cerevisiae* gene names for ClueGO analysis.

A combined transcriptome of *S. pastorianus* genes was generated by summing the reads for *S. cerevisiae* and *S. eubayanus* orthologues using the consolidate function in Excel. Unmatched genes were retained in the combined transcriptome with reads mapped respectively to the *S. cerevisiae* or *S. eubayanus* reference genomes. The combined genome contained 7568 unique gene ids. The copy number for each gene was extracted from the reads coverage of the *de novo* genome sequencing of the *S. pastorianus* strains and were summed for the matched orthologues to give a total gene copy number for each gene id.

Statistics

False discoveries rates for DEG and gene ontologies were determined using the Benjamini-Hochberg correction. Chi-squared tests were used to determine the enrichment of genes from sub-genomes under the different physiological conditions used.

Supporting information

S1 Fig. Box and whiskers plot of the differential expression of Se and Sc genes in comparisons between Day 2 vs min, Day 4 vs min and Day 2 vs Day 4. The box extends from the 25th to 75th percentiles, the mean is shown by a bar within the box. The whiskers show the minimum and maximum values in the data set.

(TIF)

S2 Fig. Effect of gene copy number on gene expression patterns on Day 4 and in minimal medium. The DEG of Sc to Se orthologues were grouped according to the Sc to Se ratio of gene copy number in W34/70 (A) and CBS1538 (D). Data for minimal medium and for Day 4 (D4) are shown. The number of paired orthologues at each copy number ratio is shown above the graphs. **B.** Correlation of differential expression of Sc:Se gene orthologues in W34/70 with gene orthologue copy number ratio. The mean of \log_2 fold changes in expression between Sc and Se orthologues is plotted against Sc:Se gene orthologue copy number ratio. Error bars represent the standard error of the mean. **C.** \log_2 -fold gene expression difference between Sc and Se orthologues on chromosome X in W34/70. The dotted line marks the recombination site (*THD2*) on the hybrid chromosome. Sc:Se ratio to left of vertical dotted line is 3:1 and after 1:1.

(TIF)

S3 Fig. Relationship of gene copy number and gene expression in W34/70 and CBS1538.

The ratio of the gene copy number for the sum of all alleles (Sc+Se) was compared to \log_2 -fold differences in gene expression in W34/70 and CBS1538. The number of genes in each ratio category is shown above the bars. Data for Day 4 in wort and for growth in minimal medium is shown. The error bars represent the standard error of the mean \log_2 -fold change in gene expression between WS34/70 and CBS1538 strains.

(TIF)

S1 File. Supplemental Tables. Table A in S1 File. Table of raw reads for triplicate samples of growth in minimal medium (min) in wort on Day 2 (Day 2) and in wort on Day 4 (Day 4). The Geneid correspond to the *S. eubayanus* or *S. cerevisiae* geneids from the reference genomes. Where gene names are known these have been added to column 2. Columns 3 and 4 show the copy number for each gene in CBS1538 and W34/70 respectively. **Table B in S1 File.** Differential gene expression for comparison of transcript reads in wort on Day 2 vs growth in minimal medium. The \log_2 -fold change and p-adjusted values for each gene is shown.

Table C in S1 File. Differential gene expression for comparison of transcript reads in wort on Day 4 vs growth in minimal medium. The \log_2 -fold change and p-adjusted values for each

gene is shown. **Table D in S1 File.** Differential gene expression for comparison of transcript reads in wort on Day 2 vs Day 4. The \log_2 -fold change and p-adjusted values for each gene is shown. **Table E in S1 File.** List of genes up-regulated and down-regulated in the comparisons of all three conditions, Day 2 vs min, Day 4 vs min and Day 2 vs Day 4. **Table F in S1 File.** Consolidated transcriptome of CBS1538 and W34/70. A combined transcriptome of *S. pastorianus* genes was generated by summing the reads for *S. cerevisiae* and *S. eubayanus* orthologues using the consolidate function in Excel. Unmatched genes were retained in the combined transcriptome with reads mapped respectively to the *S. cerevisiae* or *S. eubayanus* reference genomes. The raw reads for triplicate samples of growth in minimal medium (min) in wort on Day 2 (Day 2) and in wort on Day 4 (Day 4) are shown. The copy number for each gene was extracted from the reads coverage of the *de novo* genome sequencing of the *S. pastorianus* strains CBS1538 and W34/70 respectively and mapped to a combined genome of *S. cerevisiae* and *S. eubayanus* derived from *S. cerevisiae* S288C, assembly R64 and *S. eubayanus* (SEU3.0) and were summed for the matched orthologues to give a total gene copy number for each gene id (Columns D and E). Genes with zero reads were removed from the dataset. **Table G in S1 File.** Differential gene expression data from the consolidated transcriptome analysis. Differential gene expression for comparison of transcript reads in CBS1538 vs W34/70 in minimal medium and in wort on Day 2 and Day 4. The \log_2 -fold change and p-adjusted values for each gene is shown. **Table H in S1 File.** List of genes up-regulated and down-regulated in CBS1538 and W34/70 from the consolidated transcriptome dataset under the three physiological conditions examined in this study and their intersects.

(XLSX)

S1 Data. Brix readings for Fig 1.

(TXT)

S2 Data. Reads from *de novo* DNA sequencing of CBS1538, mapped against the combined genomes of the reference strains of *S. cerevisiae* and *S. eubayanus* were transformed into sorted BAM files using samtools and the data was extracted as reads/500 bp and then normalized by the size of the library (total number of reads) to give an estimated chromosome copy number.

(TXT)

S3 Data. Reads from *de novo* DNA sequencing of W34/70, mapped against the combined genomes of the reference strains of *S. cerevisiae* and *S. eubayanus* were transformed into sorted BAM files using samtools and the data was extracted as reads/500 bp and then normalized by the size of the library (total number of reads) to give an estimated chromosome copy number.

(TXT)

Acknowledgments

We thank the Genomic Technologies Core Facility and the Bioinformatic Core Facility at the University of Manchester for assistance with RNAseq analysis and Ian Donaldson for initial bioinformatic support.

Author Contributions

Conceptualization: Ursula Bond.

Data curation: Karsten Hokamp, Fiona Roche, Georgia Thompson.

Formal analysis: Roberto de la Cerda Garcia-Caro.

Funding acquisition: Daniela Delneri, Ursula Bond.

Methodology: Roberto de la Cerda Garcia-Caro, Karsten Hokamp, Fiona Roche, Soukaina Timouma.

Project administration: Ursula Bond.

Resources: Daniela Delneri.

Software: Fiona Roche.

Supervision: Ursula Bond.

Writing – original draft: Roberto de la Cerda Garcia-Caro, Ursula Bond.

Writing – review & editing: Roberto de la Cerda Garcia-Caro, Ursula Bond.

References

1. Hornsey IS. A history of beer and brewing. Cambridge, UK: The Royal Society of Chemistry; 2003.
2. Barnett JA. A history of research on yeasts 2: Louis Pasteur and his contemporaries, 1850–1880. *Yeast*. 2000; 16(8):755–771. [https://doi.org/10.1002/1097-0061\(20000615\)16:8<755::AID-YEA587>3.0.CO;2-4](https://doi.org/10.1002/1097-0061(20000615)16:8<755::AID-YEA587>3.0.CO;2-4) PMID: 10861901
3. Monerawela C, Bond U. Brewing up a storm: The genomes of lager yeasts and how they evolved. *Bio-technol Adv*. 2017; 35(4):512–519. <https://doi.org/10.1016/j.biotechadv.2017.03.003> PMID: 28284994
4. Dunn B, Sherlock G. Reconstruction of the genome origins and evolution of the hybrid lager yeast *Saccharomyces pastorianus*. *Genome Res*. 2008; 18(10):1610–1623. <https://doi.org/10.1101/gr.076075.108> PMID: 18787083
5. Libkind D, Hittinger C, Valério E, Gonçalves C, Dover J, Johnston M, et al. Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. *Proceedings of the National Academy of Sciences of the United States of America*. 2011; 108(35):14539–14544. <https://doi.org/10.1073/pnas.1105430108> PMID: 21873232
6. Peris D, Langdon QK, Moriarty RV, Sylvester K, Bontrager M, Charron G, et al. Complex Ancestries of Lager-Brewing Hybrids Were Shaped by Standing Variation in the Wild Yeast *Saccharomyces eubayanus*. *PLoS Genet*. 2016; 12(7):e1006155. <https://doi.org/10.1371/journal.pgen.1006155> PMID: 27385107
7. Langdon QK, Peris D, Eizaguirre JI, Oplente DA, Buh KV, Sylvester K, et al. Postglacial migration shaped the genomic diversity and global distribution of the wild ancestor of lager-brewing hybrids. *PLoS Genet*. 2020; 16(4):e1008680. <https://doi.org/10.1371/journal.pgen.1008680> PMID: 32251477
8. Eizaguirre JI, Peris D, Rodriguez ME, Lopes CA, De Los Rios P, Hittinger CT, et al. Phylogeography of the wild Lager-brewing ancestor (*Saccharomyces eubayanus*) in Patagonia. *Environ Microbiol*. 2018; 20(10):3732–3743. <https://doi.org/10.1111/1462-2920.14375> PMID: 30105823
9. Bing J, Han PJ, Liu WQ, Wang QM, Bai FY. Evidence for a Far East Asian origin of lager beer yeast. *Curr Biol*. 2014; 24(10):R380–381. <https://doi.org/10.1016/j.cub.2014.04.031> PMID: 24845661
10. Alsammar HF, Naseeb S, Brancia LB, Gilman RT, Wang P, Delneri D. Targeted metagenomics approach to capture the biodiversity of *Saccharomyces* genus in wild environments. *Environ Microbiol Rep*. 2019; 11(2):206–214. <https://doi.org/10.1111/1758-2229.12724> PMID: 30507071
11. Krogerus K, Seppanen-Laakso T, Castillo S, Gibson B. Inheritance of brewing-relevant phenotypes in constructed *Saccharomyces cerevisiae* x *Saccharomyces eubayanus* hybrids. *Microb Cell Fact*. 2017; 16(1):66. <https://doi.org/10.1186/s12934-017-0679-8> PMID: 28431563
12. Bond U, Neal C, Donnelly D, James T. Aneuploidy and copy number breakpoints in the genome of lager yeasts mapped by microarray hybridisation. *Current Genetics*. 2004; 45(6):360–370. <https://doi.org/10.1007/s00294-004-0504-x> PMID: 15103502
13. Kodama Y, Kielland-Brandt MC, Hansen J. Lager brewing yeast. In: Sunnerhagen P, Piskur J, editors. *Comparative Genomics. Topics in Current Genetics*. 15. Berlin, Heidelberg: Springer; 2006. p. 145–164.
14. Nakao Y, Kanamori T, Itoh T, Kodama Y, Rainieri S, Nakamura N, et al. Genome sequence of the lager brewing yeast, an interspecies hybrid. *DNA Research* 2009; 16(2):115–129. <https://doi.org/10.1093/dnares/dsp003> PMID: 19261625
15. Walther A, Hesselbart A, Wendland J. Genome Sequence of *Saccharomyces carlsbergensis*, the World's First Pure Culture Lager Yeast. G3 (Bethesda, Md). 2014; 4(5):783–793. <https://doi.org/10.1534/g3.113.010090> PMID: 24578374

16. Wendland J. Lager yeast comes of age. *Eukaryotic Cell*. 2014; 13(10):1256–1265. <https://doi.org/10.1128/EC.00134-14> PMID: 25084862
17. Hewitt SK, Donaldson IJ, Lovell SC, Delneri D. Sequencing and characterisation of rearrangements in three *S. pastorianus* strains reveals the presence of chimeric genes and gives evidence of breakpoint reuse. *PLoS One*. 2014; 9(3):e92203. <https://doi.org/10.1371/journal.pone.0092203> PMID: 24643015
18. Barnett JA, Lichtenthaler FW. A history of research on yeasts 3: Emil Fischer, Eduard Buchner and their contemporaries, 1880–1900. *Yeast*. 2001; 18(4):363–388. [https://doi.org/10.1002/1097-0061\(20010315\)18:4<363::AID-YEA677>3.0.CO;2-R](https://doi.org/10.1002/1097-0061(20010315)18:4<363::AID-YEA677>3.0.CO;2-R) PMID: 11223946
19. Monerawela C, Bond U. The hybrid genomes of *Saccharomyces pastorianus*: A current perspective. *Yeast*. 2018; 35(1):39–50. <https://doi.org/10.1002/yea.3250> PMID: 28787090
20. Salazar AN, Gorter de Vries AR, van den Broek M, Brouwers N, de la Torre Cortes P, Kuijpers NGA, et al. Chromosome level assembly and comparative genome analysis confirm lager-brewing yeasts originated from a single hybridization. *BMC Genomics*. 2019; 20(1): <https://doi.org/10.1186/s12864-019-6263-3> PMID: 31791228
21. Monerawela C, Bond U. Recombination Sites on Hybrid Chromosomes in *S. pastorianus* share Common Sequence Motifs and define a Complex Evolutionary Relationship between Group I and II Lager Yeasts. *FEMS Yeast Research*. 2017; 17(5): <https://doi.org/10.1093/femsyr/fox1047>
22. Monerawela C, James TC, Wolfe KH, Bond U. Loss of lager specific genes and subtelomeric regions define two different *Saccharomyces cerevisiae* lineages for *Saccharomyces pastorianus* Group I and II strains. *FEMS Yeast Res*. 2015; 15(2): <https://doi.org/10.1093/femsyr/fou008> PMID: 25673756
23. Okuno M, Kajitani R, Ryusui R, Morimoto H, Kodama Y, Itoh T. Next-generation sequencing analysis of lager brewing yeast strains reveals the evolutionary history of interspecies hybridization. *DNA Research*. 2016; 23(1):67–80. <https://doi.org/10.1093/dnares/dsv037> PMID: 26732986
24. Gibson BR, Storgards E, Krogerus K, Vidgren V. Comparative physiology and fermentation performance of Saaz and Froberg lager yeast strains and the parental species *Saccharomyces eubayanus*. *Yeast*. 2013; 30(7):255–266. <https://doi.org/10.1002/yea.2960> PMID: 23695993
25. Ried T, Meijer G, Harrison DJ, Grech G, Franch-Expósito S, Briffard R, et al. The landscape of genomic copy number alterations in colorectal cancer and their consequences on gene expression levels and disease outcome. *Molecular Aspects of Medicine*. 2019; 69, (October 2019):48–61. <https://doi.org/10.1016/j.mam.2019.07.007> PMID: 31365882
26. Pavelka N, Rancati G, Zhu J, Bradford WD, Saraf A, Florens L, et al. Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. *Nature*. 2010; 468(7321):321–325. <https://doi.org/10.1038/nature09529> PMID: 20962780
27. Sheltzer JM, Torres EM, Dunham MJ, Amon A. Transcriptional consequences of aneuploidy. *Proc Natl Acad Sci U S A*. 2012; 109(31):12644–12649. <https://doi.org/10.1073/pnas.1209227109> PMID: 22802626
28. Torres EM, Sokolsky T, Tucker CM, Chan LY, Boselli M, Dunham MJ, et al. Effects of aneuploidy on cellular physiology and cell division in haploid yeast. *Science*. 2007; 317(5840):916–924. <https://doi.org/10.1126/science.1142210> PMID: 17702937
29. Piatkowska EM, Naseeb S, Knight D, Delneri D. Chimeric protein complexes in hybrid species generate novel phenotypes. *PLoS Genet*. 2013; 9(10): <https://doi.org/10.1371/journal.pgen.1003836> PMID: 24137105
30. Olesen K, Felding T, Gjermansen C, Hansen J. The dynamics of the *Saccharomyces carlsbergensis* brewing yeast transcriptome during a production-scale lager beer fermentation. *FEMS Yeast Res*. 2002; 2(4):563–573. <https://doi.org/10.1111/j.1567-1364.2002.tb00123.x> PMID: 12702272
31. James TC, Campbell S, Donnelly D, Bond U. Transcription profile of brewery yeast under fermentation conditions. *Journal of Applied Microbiology*. 2003; 94:432–448. <https://doi.org/10.1046/j.1365-2672.2003.01849.x> PMID: 12588552
32. James TC, Campbell SG, Bond UM. Comparative analysis of global gene expression in lager and laboratory yeast strains grown in wort. *Proceedings of the IEEE*. 2002; 90:1887–1899.
33. Smart KA. Brewing yeast genomes and genome-wide expression and proteome profiling during fermentation. *Yeast*. 2007; 24(11):993–1013. <https://doi.org/10.1002/yea.1553> PMID: 17879324
34. Horinouchi T, Yoshikawa K, Kawaide R, Furusawa C, Nakao Y, Hirasawa T, et al. Genome-wide expression analysis of *Saccharomyces pastorianus* orthologous genes using oligonucleotide microarrays. *J Biosci Bioeng*. 2010; 110(5):602–607. <https://doi.org/10.1016/j.jbiosc.2010.05.011> PMID: 20547377
35. Somani A, Box WG, Smart KA, Powell CD. Physiological and transcriptomic response of *Saccharomyces pastorianus* to cold storage. *FEMS Yeast Res*. 2019; 19(4): <https://doi.org/10.1093/femsyr/foz025> PMID: 31073596

36. Xu W, Wang J, Li Q. Comparative proteome and transcriptome analysis of lager brewer's yeast in the autolysis process. *FEMS Yeast Res.* 2014; 14(8):1273–1285. <https://doi.org/10.1111/1567-1364.12223> PMID: 25345722
37. Timouma S, Balarezo-Cisneros LN, Pinto J, De La Cerda R, Bond U, Schwartz JM, et al. Transcriptional profile of the industrial hybrid *Saccharomyces pastorianus* reveals temperature-dependent allele expression bias and preferential orthologous protein assemblies. *Mol Biol Evol.* 2021; 38(12):5437–5452. <https://doi.org/10.1093/molbev/msab282> PMID: 34550394
38. Timouma S, Schwartz JM, Delneri D. HybridMine: A Pipeline for Allele Inheritance and Gene Copy Number Prediction in Hybrid Genomes and Its Application to Industrial Yeasts. *Microorganisms.* 2020; 8(10):1554–1569. <https://doi.org/10.3390/microorganisms8101554> PMID: 33050146
39. Magalhaes F, Vidgren V, Ruohonen L, Gibson B. Maltose and maltotriose utilisation by group I strains of the hybrid lager yeast *Saccharomyces pastorianus*. *FEMS Yeast Res.* 2016; 16(5): <https://doi.org/10.1093/femsyr/fow1053>
40. Vidgren V, Londesborough J. Characterization of the *Saccharomyces bayanus*-type AGT1 transporter of lager yeast. *Journal of the Institute of Brewing.* 2012; 118(2):148–151.
41. Vidgren V, Ruohonen L, Londesborough J. Characterization and functional analysis of the MAL and MPH Loci for maltose utilization in some ale and lager yeast strains. *Appl Environ Microbiol.* 2005; 71(12):7846–7857. <https://doi.org/10.1128/AEM.71.12.7846-7857.2005> PMID: 16332759
42. van den Broek M, Bolat I, Nijkamp JF, Ramos E, Luttik MA, Koopman F, et al. Chromosomal Copy Number Variation in *Saccharomyces pastorianus* Is Evidence for Extensive Genome Dynamics in Industrial Lager Brewing Strains. *Appl Environ Microbiol.* 2015; 81(18):6253–6267. <https://doi.org/10.1128/AEM.01263-15> PMID: 26150454
43. James T, Usher J, Campbell S, Bond U. Lager yeasts possess dynamic genomes that undergo rearrangements and gene amplification in response to stress. *Current Genetics.* 2008; 53(3):139–152. <https://doi.org/10.1007/s00294-007-0172-8> PMID: 18183398
44. Haimovich G, Medina DA, Causse SZ, Garber M, Millan-Zambrano G, Barkai O, et al. Gene expression is circular: factors for mRNA degradation also foster mRNA synthesis. *Cell.* 2013; 153(5):1000–1011. <https://doi.org/10.1016/j.cell.2013.05.012> PMID: 23706738
45. Sun M, Schwalb B, Pirkl N, Maier KC, Schenk A, Failmezger H, et al. Global analysis of eukaryotic mRNA degradation reveals Xrn1-dependent buffering of transcript levels. *Mol Cell.* 2013; 52(1):52–62. <https://doi.org/10.1016/j.molcel.2013.09.010> PMID: 24119399
46. Osley MA, Hereford LM. Yeast histone genes show dosage compensation. *Cell.* 1981; 24(2):377–384. [https://doi.org/10.1016/0092-8674\(81\)90327-5](https://doi.org/10.1016/0092-8674(81)90327-5) PMID: 7016340
47. Torres EM, Springer M, Amon A. No current evidence for widespread dosage compensation in *S. cerevisiae*. *Elife.* 2016; 5:e10996. <https://doi.org/10.7554/eLife.10996> PMID: 26949255
48. Querol A, Barrio E, Ramón D. A Comparative Study of Different Methods of Yeast Strain Characterization. *Systematic and Applied Microbiology.* 1992; 15(3):439–446.
49. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012; 9(4):357–359. <https://doi.org/10.1038/nmeth.1923> PMID: 22388286
50. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal.* 2011; 17(1):10–12.
51. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics.* 2009; 25(16):2078–2079. <https://doi.org/10.1093/bioinformatics/btp352> PMID: 19505943
52. Ares M. Isolation of total RNA from yeast cell cultures. *Cold Spring Harb Protoc.* 2012; 2012(10):1082–1086. <https://doi.org/10.1101/pdb.prot071456> PMID: 23028070
53. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013; 29(1):15–21. <https://doi.org/10.1093/bioinformatics/bts635> PMID: 23104886
54. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics.* 2014; 30(7):923–930. <https://doi.org/10.1093/bioinformatics/btt656> PMID: 24227677
55. Ge SX, Son EW, Yao R. iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinformatics.* 2018; 19(1): <https://doi.org/10.1186/s12859-018-2486-6> PMID: 30567491
56. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics.* 2009; 25(8):1091–1093. <https://doi.org/10.1093/bioinformatics/btp101> PMID: 19237447