

# Genomic saturation mutagenesis and polygenic analysis identify novel yeast genes affecting ethyl acetate production, a non-selectable polygenic trait

Tom Den Abt<sup>1,2</sup>, Ben Souffriau<sup>1,2</sup>, Maria R. Foulquié-Moreno<sup>1,2</sup>, Jorge Duitama<sup>3</sup>, and Johan M. Thevelein<sup>1,2,\*</sup>

<sup>1</sup>Laboratory of Molecular Cell Biology, Institute of Botany and Microbiology, KU Leuven.

<sup>2</sup>Department of Molecular Microbiology, VIB, Kasteelpark Arenberg 31, B-3001 Leuven-Heverlee, Flanders, Belgium.

<sup>3</sup>Agrobiodiversity Research Area, International Center for Tropical Agriculture (CIAT), Cali, Colombia.

\* Corresponding Author:

Johan M. Thevelein, Laboratory of Molecular Cell Biology, Institute of Botany and Microbiology, KU Leuven; Department of Molecular Microbiology, VIB, Kasteelpark Arenberg 31, B-3001 Leuven-Heverlee, Flanders, Belgium; Tel: +32-16-321507; Fax.: +32-16-321979; E-mail: johan.thevelein@mmbio.vib-kuleuven.be

**ABSTRACT** Isolation of mutants in populations of microorganisms has been a valuable tool in experimental genetics for decades. The main disadvantage, however, is the inability of isolating mutants in non-selectable polygenic traits. Most traits of organisms, however, are non-selectable and polygenic, including industrially important properties of microorganisms. The advent of powerful technologies for polygenic analysis of complex traits has allowed simultaneous identification of multiple causative mutations among many thousands of irrelevant mutations. We now show that this also applies to haploid strains of which the genome has been loaded with induced mutations so as to affect as many non-selectable, polygenic traits as possible. We have introduced about 900 mutations into single haploid yeast strains using multiple rounds of EMS mutagenesis, while maintaining the mating capacity required for genetic mapping. We screened the strains for defects in flavor production, an important non-selectable, polygenic trait in yeast alcoholic beverage production. A haploid strain with multiple induced mutations showing reduced ethyl acetate production in semi-anaerobic fermentation, was selected and the underlying quantitative trait loci (QTLs) were mapped using pooled-segregant whole-genome sequence analysis after crossing with an unrelated haploid strain. Reciprocal hemizyosity analysis and allele exchange identified *PMA1* and *CEM1* as causative mutant alleles and *TPS1* as a causative genetic background allele. The case of *CEM1* revealed that relevant mutations without observable effect in the haploid strain with multiple induced mutations (in this case due to defective mitochondria) can be identified by polygenic analysis as long as the mutations have an effect in part of the segregants (in this case those that regained fully functional mitochondria). Our results show that genomic saturation mutagenesis combined with complex trait polygenic analysis could be used successfully to identify causative alleles underlying many non-selectable, polygenic traits in small collections of haploid strains with multiple induced mutations.

doi: 10.15698/mic2016.04.491

Received originally: 18.10.2015;

in revised form: 05.02.2016,

Accepted 08.02.2016,

Published To be determined.

**Keywords:** Mutant screen, genomic mutagenesis, non-selectable trait, polygenic trait, QTL mapping, genetic basis, flavor production, ethyl acetate production.

#### Abbreviations:

AUX – auxotrophic,  
EMS – ethyl methanesulphonate,  
QTL – quantitative trait locus,  
RD – respiratory deficiency,  
RHA – reciprocal hemizyosity analysis,  
SNP – single nucleotide polymorphism,  
ts+ – temperature sensitivity.

## INTRODUCTION

Random mutagenesis has been a powerful technique in microbial genetics for decades and has aided in identifying structural and regulatory genes involved in many biochem-

ical pathways and cellular processes [1-3]. A major drawback of this methodology is that mutants can only be isolated in phenotypes that in one way or another can be selected under a specific condition, in which only the mu-

tants of interest can multiply and not the wild type or any other mutant strains. In addition, mutant screens have focused for many years on mutants carrying single mutations causing a specific phenotype because there were no efficient methodologies to identify multiple mutant genes responsible for establishing a single polygenic trait. Also in programs for improvement of industrial yeast strains, random approaches, like population mutagenesis, have often been more successful than rational approaches targeting specific genes [4-6]. This likely reflects the lack of insight in the regulation of many metabolic pathways and cellular processes, especially in industrial yeast strains. In addition, the genetic basis of many commercially important traits is often poorly understood in laboratory yeast strains and even more in industrial yeast strains. Most traits of yeast that have industrial importance are non-selectable and polygenic. As a result, it has been difficult to obtain mutants in commercially-important traits of yeast and to identify the genes involved using classical complementation approaches.

The advent of high-throughput genome sequencing methodologies has allowed simultaneous mapping of QTLs underlying a polygenic trait-of-interest, while reciprocal hemizyosity analysis and allele exchange have allowed to identify the causative genes in the QTLs [7, 8]. This has now been performed with selectable or easily identifiable traits, allowing very high numbers of segregants to be used [9, 10], but also with small pools of selected segregants for difficult-to-score commercially important traits [11-26]. These new methodologies have allowed exploration of the huge *Saccharomyces cerevisiae* biodiversity to identify genes affecting commercially-important traits. This approach is limited, however, to genes sufficiently affected by natural variation to exert a significant effect on a trait-of-interest. There is no evidence that identification of all genes responsible for the natural variation in a trait-of-interest will reveal the complete genetic basis of this trait.

For non-selectable traits, on the other hand, all segregants need to be phenotyped individually, for instance in small scale fermentations, which can be quite laborious. There is however a tremendous interest in these non-selectable phenotypes [19, 26-28] and recently Hubmann *et al.* [17] and Pais *et al.* [12] showed that even with a limited number of segregants causative alleles could be identified for maximal ethanol accumulation and low glycerol production. For similar reasons, also random mutagenesis techniques have been limited to selectable or at least easily identifiable phenotypes. The possibility to simultaneously identify several causative alleles, in the midst of many irrelevant mutations, has provided the opportunity to also expand mutagenesis treatments to polygenic non-selectable phenotypes.

A non-selectable trait that attracted much attention is the production of volatile esters because these esters are responsible for the highly desired fruity aroma character of alcoholic beverages, like beer, wine and sake [29-33]. Although a number of structural genes involved in ester biosynthesis have been identified [34-41], little is known about the regulation of the biosynthetic pathways and

about other systems influencing the production of specific flavor esters positively or negatively. Ethyl acetate is such a flavor-active ester, which has a desirable, fruity flavor in low concentrations, but an undesirable solvent-like flavor in higher concentrations. Atf1 and Atf2 have been shown to catalyze the synthesis of ethyl acetate from its two substrates, ethanol and acetyl-CoA, but an *atf1Δ atf2Δ* strain still produces considerable amounts of ethyl acetate [38]. The genetic basis of the remaining ethyl acetate production remains unknown and provides an interesting non-selectable, likely polygenic phenotype to investigate with the genomic saturation mutagenesis technology. Identification of novel genes causing modified ethyl acetate production would allow for targeted strain improvement by genetic engineering but also provide new genetic targets for marker-assisted breeding [42], in order to refine the flavor profile of alcoholic beverages produced by yeast.

The polygenic nature of most traits requires many mutations to be introduced to create an appropriate base for their genetic analysis and an adaptation of current mutagenesis protocols is therefore necessary. Ultraviolet light (UV) and ethyl methanesulphonate (EMS) mutagenesis have been the most popular classical mutagenesis techniques in yeast but are limited in the number of mutations they introduce. Recently, Shiwa *et al.* [43] proposed an alternative mutagenesis technique using a mutant error-prone DNA polymerase  $\delta$ . Their study showed, however, that, although creating a more diverse mutation profile, classical mutagenesis techniques were still superior in introducing multiple genomic mutations as assessed by whole genome sequencing. Therefore adaptation of existing UV or EMS protocols was considered to be a more appropriate approach to maximize mutation load.

In this work we have evaluated the two most commonly used mutagen treatments, UV and EMS, for maximal introduction of random mutations in a single haploid yeast strain. Using an optimized protocol, we show that these haploid strains can accumulate ~900 mutations while retaining genetic proficiency, i.e. maintenance of mating capacity, sporulation capacity of the diploid strain and spore viability of the segregants. The mutant strains can thus be used for QTL mapping of polygenic, non-selectable traits. As a proof-of-principle we have applied pooled-segregant whole-genome sequence analysis to identify QTLs for low ethyl acetate production. Further analysis of the QTLs with the strongest linkage has identified causative mutations in *PMA1*, and *CEM1* as well as *TPS1* as a causative genetic background allele. These results show that genomic saturation mutagenesis combined with polygenic analysis can be used successfully to identify the genetic basis of non-selectable polygenic traits.

## RESULTS

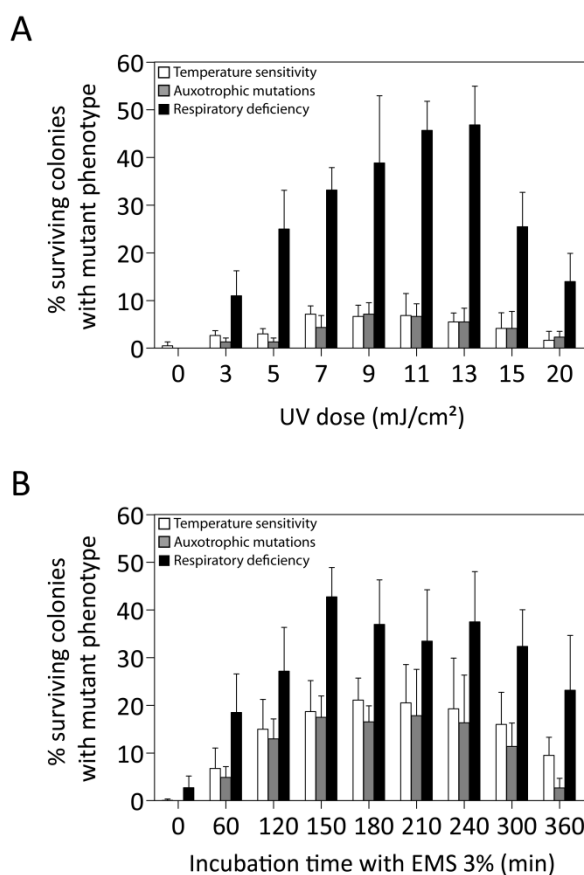
### Comparison of the efficiency of UV and EMS mutagenesis in enhancing mutation load in the absence of phenotypic selection.

We first evaluated the efficiency of UV and EMS mutagenesis to cause accumulation of multiple mutations in single

haploid yeast strains without selection for any phenotype. Haploid S288c *S. cerevisiae* cells were irradiated with UV light at different energy levels or treated with 3% EMS for different time intervals and subsequently allowed to grow to single colonies on rich YPD medium. Mutagenesis efficiency was first estimated by assessing the presence of auxotrophic mutations on eight drop-out media (AUX), temperature sensitivity at 37°C (ts+) and respiratory deficiency (RD). For both mutagen treatments an optimum curve was obtained with increasing dose for the three types of mutations (Fig. 1). For the AUX and ts+ mutations, the maximum efficiency of EMS (Fig. 1B) was nearly three times higher than that of UV (Fig. 1A), with maxima of about 20% for EMS and about 7% for UV. The frequency of RD mutations was much higher than that of AUX or ts+ mutations, with maxima of 46.8% (at 13 mJ/cm<sup>2</sup>) and 42.77% (at 150 min) for UV and EMS, respectively. Mating of 160 RD mutants with BY4742  $\rho 0$  showed that 73% of RD mutants were caused by mitochondrial mutations. EMS was chosen as the preferred mutagen to maximize mutation load without selection and AUX and ts+ mutations were considered the most reliable measure for accumulation of genomic mutations. Although RD does not appear to be a good measure for genomic mutation load, it still appears to be a valuable read-out for the mutagenesis process.

#### Accumulation of multiple mutations in single haploid strains by consecutive rounds of EMS mutagenesis.

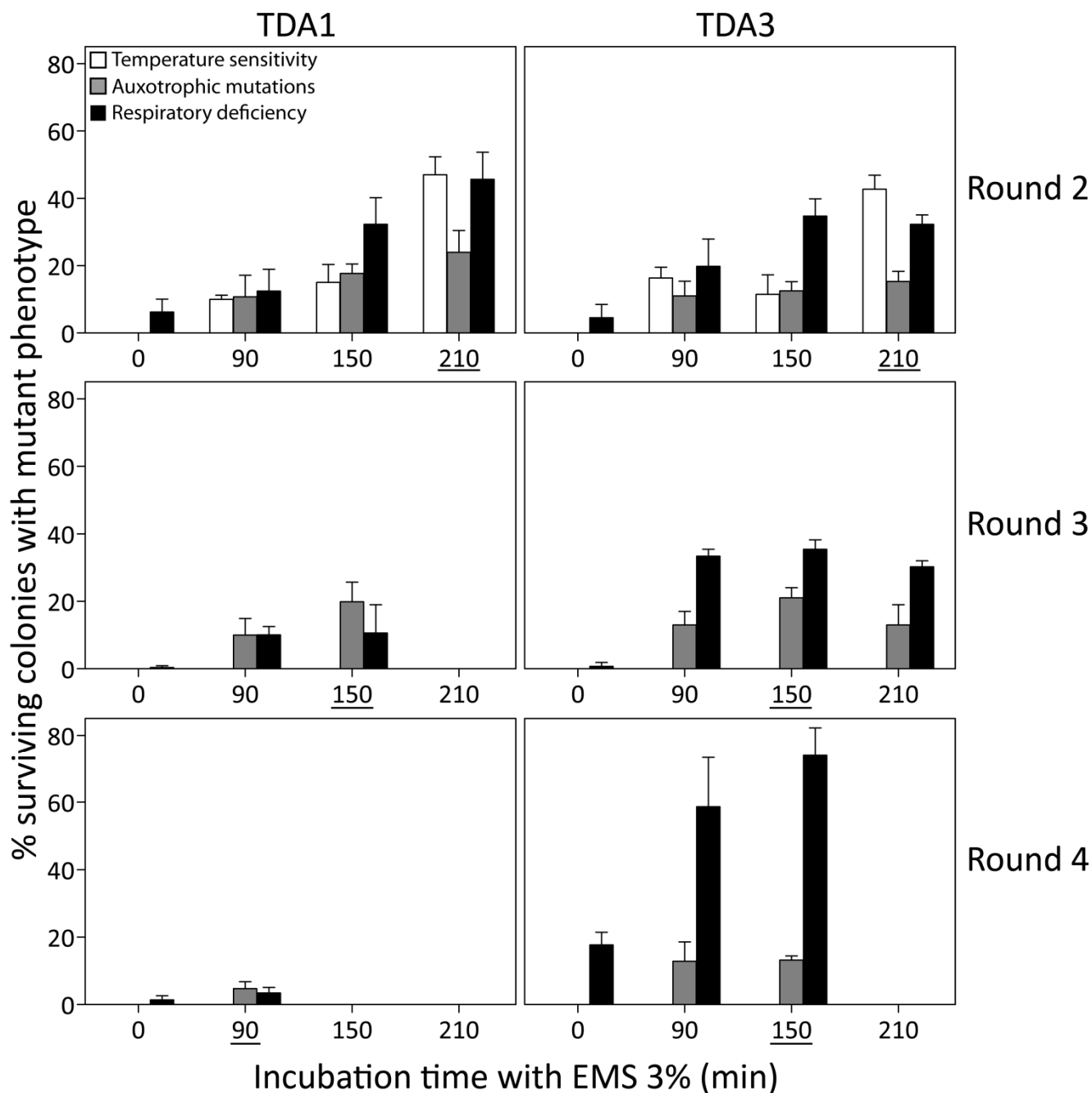
Since single treatments with EMS result in a limited number of induced mutations [43–46], we performed multiple rounds of EMS mutagenesis to increase the mutation frequency further. Two mutants, TDA1 and TDA3, were selected after the first round of 210 min EMS mutagenesis of the haploid S288c strain based on maintenance of respiration, mating capacity (with S288c *MAT $\alpha$  ura3::KanMX*), sporulation and a spore viability greater than 70%. For double selection in mating assays, the mutants were also selected to be prototrophic for uracil. The two mutants were each submitted to three additional rounds of EMS mutagenesis. In each round a single mutant was selected for the next mutagenesis round, based on the same criteria as applied after the first mutagenesis round (Fig. 2). TDA1-derived mutants were selected after 210, 150 and 90 min of EMS treatment in round 2, 3 and 4, respectively. In each of the four rounds, the selected TDA1-derived mutant lacked auxotrophic mutations and the same number of genes was thus tested in every round for the occurrence of mutations. TDA3-derived mutants were selected after 210, 150 and 150 min of EMS treatment in round 2, 3 and 4, respectively. In round 3, the TDA3-derived mutant selected had auxotrophic mutations for methionine and histidine. Hence, these two auxotrophies could no longer be tested in round 4 and the total number of genes evaluated for mutations in this round was thus reduced from 80 to 50. Maximum mutagenesis efficiencies were in all rounds within the range of that initially observed in S288c for both ts+ and AUX mutations except for round 4. A severe drop in efficiency was observed for TDA1 (AUX 4.7% at 90 min)



**FIGURE 1. Mutation frequencies induced by different doses in UV and EMS mutagenesis.** The S288c haploid yeast strain was submitted to mutagenesis by different doses of UV light (A) or different time periods of incubation with EMS (B). The mutation frequency in surviving colonies was assessed through the presence of different types of mutations: temperature sensitivity (white bars), auxotrophic mutations on eight drop-out media (grey bars) and respiratory deficiency (black bars). No viable colonies were recovered after 420 min of incubation with EMS.

possibly due to the accumulation of synthetically lethal mutations. Both mutant lines became ts+ after the second mutagenesis round.

Whole genome sequence analysis of a single colony from each line, i.e. TDA1(4) and TDA3(4), was performed after four rounds of mutagenesis. Genomic DNA of both mutants and the original S288c strain was extracted and subjected to whole-genome sequence analysis using Illumina HiSeq2000 technology (BGI, Hong Kong, China). The sequence reads were mapped to the S288c reference sequence and variants were identified and filtered using the NGSEP pipeline [47]. TDA1(4) and TDA3(4) had accumulated respectively 815 and 967 single nucleotide polymorphisms (SNPs) and small indels compared to the original S288c strain. The genomic loci of the 1294 variants located in ORFs were assessed for natural variation with ‘Align Strain Sequences’ at SGD (<http://www.yeastgenome.org/cgi-bin/FUNGI/alignment.pl>). The ORF sequences of an average of 40 yeast strains were compared to the S288c



**FIGURE 2. Mutation frequency after multiple EMS mutagenesis rounds.** Two mutants, TDA1 and TDA3, selected after 210 min of EMS mutagenesis of the S288c strain, were submitted to three additional EMS mutagenesis rounds. In each round, the mutation frequency was assessed after 0, 90, 150, 210 and 360 min of incubation with EMS. It was based on temperature sensitivity (white bars), auxotrophic mutations on eight drop-out media (grey bars) and respiratory deficiency (black bars). The time period of incubation with EMS for each selected mutant is underlined in each mutagenesis round. For both mutants, no viable colonies were obtained after 360 min of EMS treatment in the second mutagenesis round and at the other time points at which no bars are present. Both mutant lines became temperature sensitive in round 2 and TDA3 acquired auxotrophic mutations for methionine and histidine in round 3.

reference strain (Supplementary Tables S1 and S2). The 40 yeast strains were a selection out of 48 possible strains for which the specific ORF sequence was available, i.e. AWRI1631, AWRI796, BC187, BY4741, BY4742, CBS7960, CEN.PK, CLIB215, CLIB324, D273-10B, DBVPG6044, EC1118,

EC9-8, FL100, FY1679, FosterB, FosterO, JAY291, JK9-3d, K11, Kyokai7, L1528, LalvinQA23, M22, PW5, RedStar, RM11-1a, SEY6210, SK1, Sigma1278b, T7, T73, UC5, UWOPS05\_217\_3, VL3, Vin13, W303, X2180-1A, Y10, Y55, YJM269, YJM339, YJM789, YPH499, YPS128, YPS163, YS9,

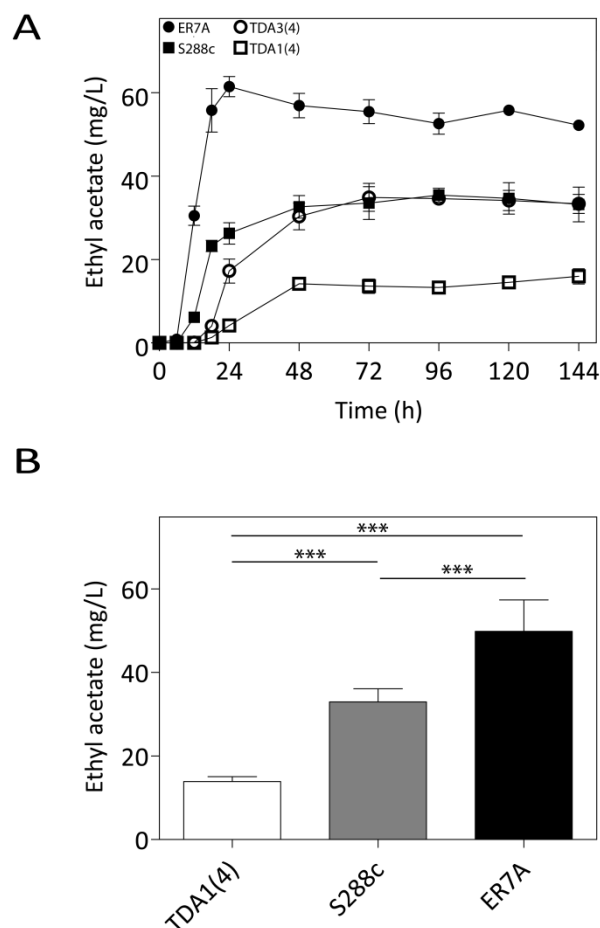
ZTW1. More than 93% of the assessed genomic loci did not display any natural variation. Hence, the induced mutations in TDA1(4) and TDA3(4) in these loci were all unique compared to the sequence present in the natural yeast strains evaluated and thus likely absent from all or most of the biodiversity of *S. cerevisiae*. The majority of the SNPs present in the 7% genomic loci with sequence variation were also different from SNPs present in the natural alleles. Hence, genomic mutagenesis induces mutations that are generally different from those present in the natural biodiversity.

#### Screening of TDA1(4) and TDA3(4) for non-selectable, polygenic phenotypes.

The two mutants, TDA1(4) and TDA3(4), were screened for flavor compound production as a typical example of a non-selectable and likely polygenic phenotype. This was performed in small scale (100 mL) semi-anaerobic fermentations with S288c and ER7A [17] as control strains. Fermentations were carried out in 10% YPD at 30°C with stirring at 80 rpm and a water lock placed on top. At the indicated times, samples were taken and flavor ester concentrations determined by headspace GC-FID. Fermentations finished in less than 96 h for all strains. Ethyl acetate production by TDA1(4) was considerably reduced at each time point compared to the control strains, while for TDA3(4) there was no significant difference in ethyl acetate production, although a slight delay in the onset was observed (Fig. 3A). We have subsequently chosen one fixed time point, i.e. 96 h, for determination of ethyl acetate levels in phenotyping the segregants for QTL mapping. After 96 h of fermentation, ethyl acetate concentrations were  $13.9 \pm 1.2$  mg/L and  $32.9 \pm 3.2$  mg/L for TDA1(4) and S288c, respectively (Fig. 3B). The low ethyl acetate production phenotype was chosen for further analysis as an example of a non-selectable, likely polygenic trait. For QTL mapping, ER7A ( $49.8 \pm 7.5$  mg/L) was chosen as inferior parent, instead of S288c, in order to have enough SNPs as genetic markers and because of the larger phenotypic difference in this case between the parent strains (Fig. 3B).

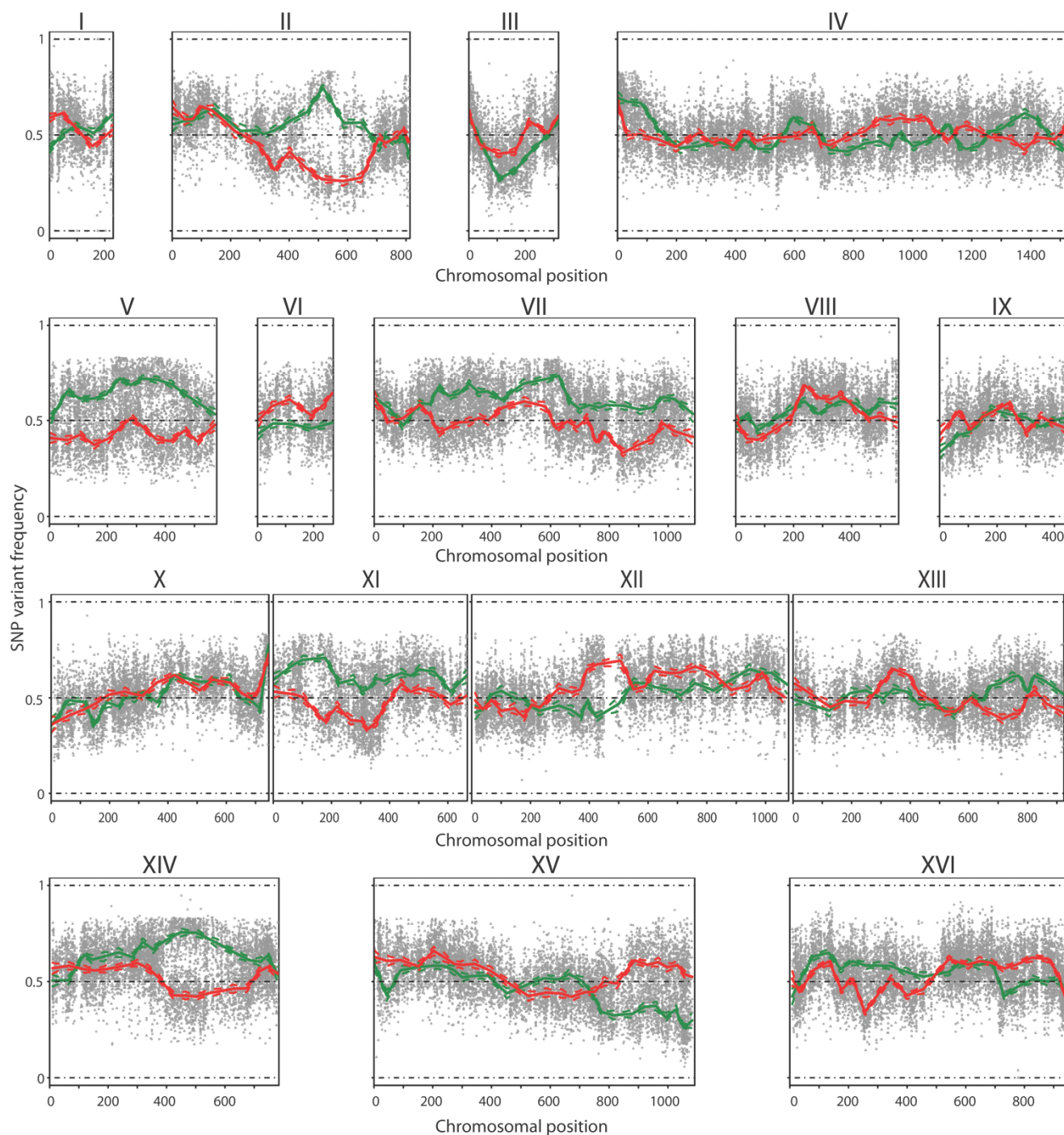
#### Selection of superior segregants with low ethyl acetate production and QTL mapping by pooled-segregant whole-genome sequence analysis.

The TDA1(4)/ER7A hybrid diploid strain was constructed, sporulated and 386 segregants were isolated and characterized for ethyl acetate production after 96 h in small scale semi-anaerobic fermentations (Supplementary Table S3). Among these, 41 segregants showed a low ethyl acetate production (<20 mg/L) and were selected for QTL mapping with pooled-segregant whole-genome sequence analysis (Supplementary Fig. S1). As a control, 41 randomly selected segregants were pooled and also subjected to pooled-segregant whole-genome sequence analysis. Genomic DNA of the selected and random pools was extracted and sequenced as described for the mutant strains. The sequence reads were aligned with the S288c reference genome and quality-filtered SNPs were used as genetic markers, essentially as described previously [18]. The ER7A



**FIGURE 3. Ethyl acetate production in small-scale semi-anaerobic fermentations. (A)** Ethyl acetate concentration during the course of a semi-anaerobic fermentation in 100 mL 10% YPD. Strains: (■) S288c, (●) ER7A, (□) TDA1(4) and (○) TDA3(4). **(B)** Ethyl acetate concentrations produced after 96 h, were analysed by a one-way ANOVA for TDA1(4) (white bar), S288c (grey bar) and ER7A (black bar) (p-value  $\leq 0.001$ ). Subsequent multiple comparison was carried out by a post hoc Tukey test with p-value  $\leq 0.001$  (\*\*\*).

parent had been sequenced previously [17]. The SNP variant frequency of the two pools was then plotted against the SNP position throughout the genome (Fig. 4). Smoothing of the SNP variant frequency in both pools, under the form of smoothing splines, was accomplished using a linear mixed model framework [18]. Seven regions showed a significant deviation from 50% (i.e. <30% or >70%). Two of these regions were located on chromosome III and XV and showed 27.2 and 25.9% linkage to ER7A, the inferior parent, respectively. These two QTLs were not further analysed since they were unlikely due to linkage with SNPs induced by mutagenesis in the non-selectable trait of low ethyl acetate production, which was the aim of this study. The other five regions were located on chromosome II, V, VII, XI and XIV and showed 74.4, 71.9, 73.2, 70.8 and 75.7% linkage, respectively, to the genome of TDA1(4), the superior parent.



**FIGURE 4. QTL mapping by pooled-segregant whole-genome sequence analysis of low ethyl acetate production.** The SNP variant frequency was plotted versus the SNP chromosomal position for the 16 yeast chromosomes (raw data: grey circles for selected pool and grey triangles for random pool; smoothed data: solid green line for selected pool and solid red line for random pool; statistical confidence interval: dashed green and red lines). Significant upward deviation from 50% (>70%), indicating linkage to the genome of TDA1(4), was observed on chromosome II, V, VII, XI and XIV for the selected pool. Significant downward deviation from 50% (<30%), indicating linkage to the genome of ER7A, was observed on chromosome III and XV.

#### Fine-mapping of QTLs by allele specific PCR in individual segregants.

Fine-mapping of the five QTLs was performed by scoring selected individual SNPs by allele specific PCR in individual

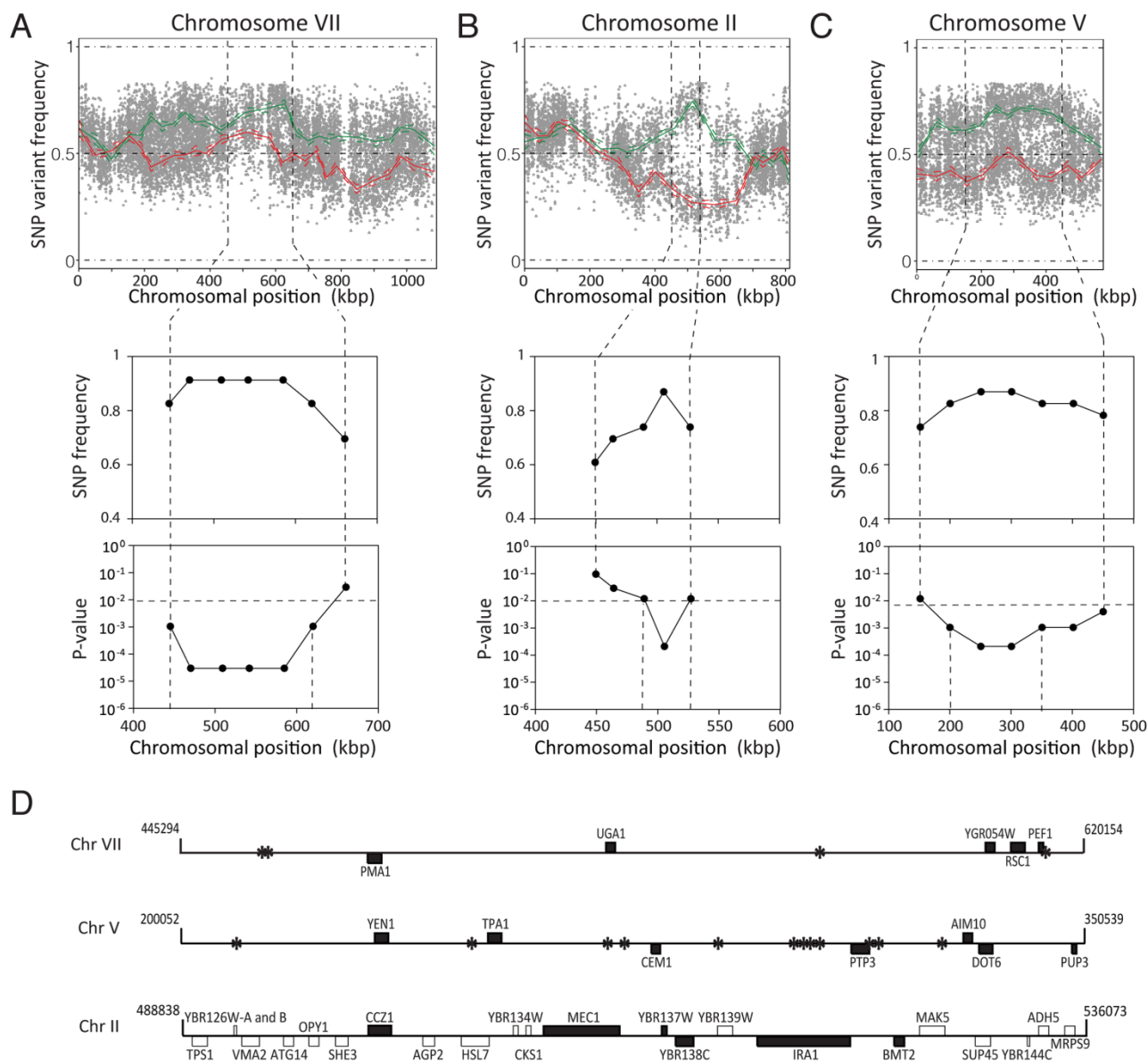
segregants. A more stringent cut-off was used for low ethyl acetate production in the individual segregants since in a separate S288c/ER7A cross, one out of 44 segregants accumulated 19.5 mg/L ethyl acetate, which is lower than the

20 mg/L cut-off used for QTL mapping. We chose 16.3 mg/L as the new cut-off since it was the highest measured ethyl acetate concentration in TDA1(4) after 96 h of fermentation. As a result, 18 TDA1/ER7A segregants were excluded from the fine-mapping. When the SNP variant frequency was determined in each QTL with the 23 remaining segregants of the selected pool, the strongest link to low ethyl acetate production was located in the QTL on chromosome VII, with a p-value of  $3.02 \times 10^{-5}$  (Fig. 5A). The two QTLs on chromosome II and V also showed strong linkage with both p-values being  $2.11 \times 10^{-4}$  (Fig. 5B, C). The

two putative QTLs on chromosome XI and XIV were not significant with this number of segregants and at this stringency level since their p-values were above the 0.01 threshold (both 0.012). Further analysis was thus focused on the QTLs on chromosome VII, II and V.

#### Identification of a causative allele in the QTL on chromosome VII.

Sequence analysis of the 175 kb region with strongest linkage on chromosome VII revealed seven intragenic and four intergenic SNPs in the TDA1(4) sequence compared to the

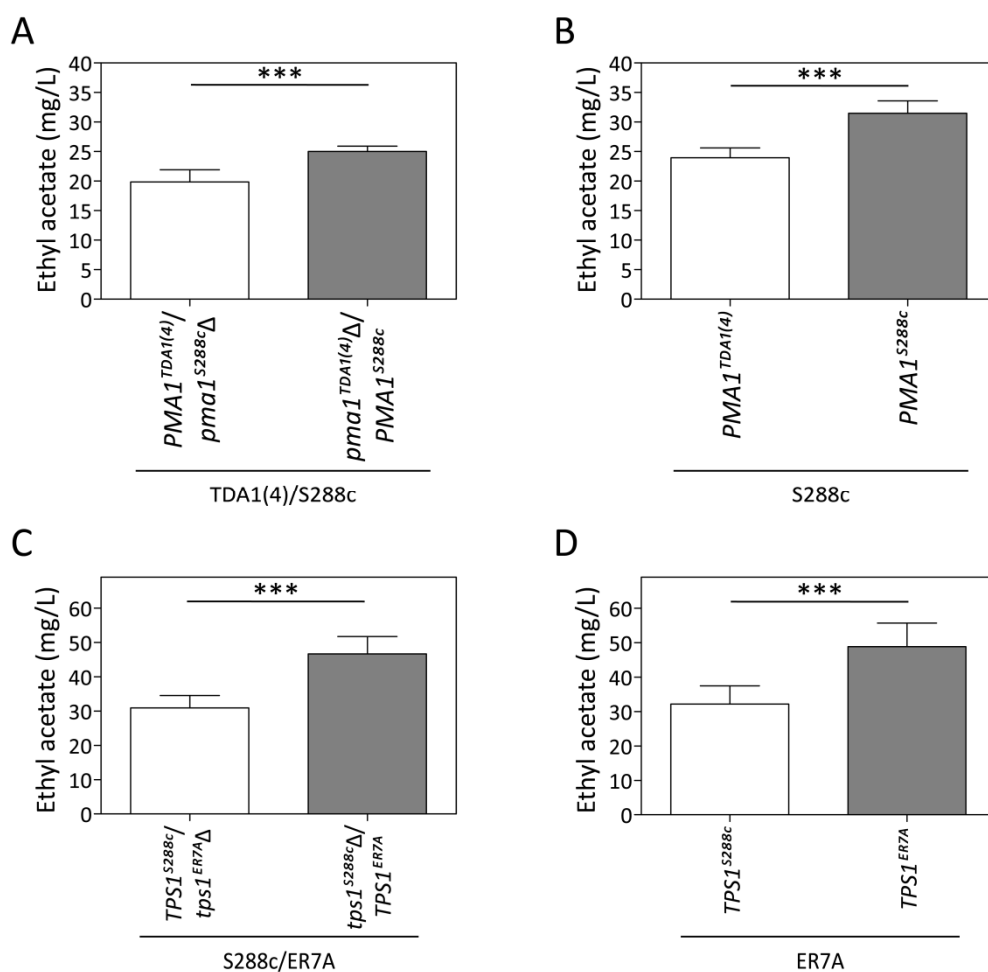


**FIGURE 5. Fine mapping of QTLs on chromosome VII, II and V.** Top: SNP variant frequency plotted versus SNP chromosomal position for chromosomes VII (**A**), II (**B**) and V (**C**) (raw data: grey circles for selected pool and grey triangles for random pool; smoothed data: solid green line for selected pool and solid red line for random pool; statistical confidence interval: dashed green and red lines). Middle and bottom: SNP frequency and associated p-value as determined by allele specific PCR in the 23 selected individual segregants (<16.3 mg/L ethyl acetate), in the indicated regions on chromosome VII (**A**), II (**B**) and V (**C**). The statistical confidence line (p-value  $\leq 0.01$ ) is also indicated. (**D**) Overview of intergenic regions (\*) and genes present on chromosomes VII, II and V with a mutation in TDA1(4) (black bars) compared to S288c. For chromosome II the genes with a mutation in ER7A (white bars) compared to S288c are also indicated.

sequence of the S288c parent strain (Fig. 5D). Except for one intergenic SNP located in an autonomously replicating sequence, these were evaluated as possible causative genetic elements by reciprocal hemizyosity analysis (RHA) [21]. For each of the candidate causative variants, we constructed two hybrid hemizygous TDA1(4)/S288c strains differing only in the candidate locus, either by deleting the precise mutant ORF or the two neighboring ORFs for intra- and intergenic SNPs, respectively. Comparison of ethyl acetate production by the two hybrid strains for each of the eleven variants revealed a significant difference only for *PMA1* (Fig 6A and supplementary Fig S2). The hybrid with the *PMA1*<sup>TDA1(4)</sup> allele showed a significantly reduced ethyl acetate production ( $19.8 \pm 2.1$  mg/L) compared to the hybrid with the *PMA1*<sup>S288c</sup> allele ( $25.0 \pm 0.9$  mg/L). We further confirmed *PMA1* as causative allele for low ethyl acetate production by allele replacement in the S288c strain (Fig 6B). A significant reduction in ethyl acetate production was observed when *PMA1*<sup>S288c</sup> was replaced by *PMA1*<sup>TDA1(4)</sup> in the S288c strain ( $24.0 \pm 1.7$  mg/L compared to  $31.5 \pm 2.1$  mg/L). Sequence comparison of the *PMA1*<sup>TDA1(4)</sup> and *PMA1*<sup>S288c</sup> alleles revealed one SNP in the ORF (position 482216 bp of chromosome VII), resulting in an M152I change in the amino acid sequence.

### Identification of the causative alleles in the QTLs on chromosome II and V.

Similar to the QTL on chromosome VII, RHA was performed for six and sixteen genes with a mutation in the QTLs on chromosome II and V, respectively (two intergenic SNPs were excluded on chromosome V because of their location between a TY element and a tRNA or an autonomously replicating sequence) (Fig. 5D). For none of the candidate mutant genes on chromosome II did we observe an allele-specific difference in ethyl acetate production (Supplementary Fig S3). This could be due to the fact that the causative gene in this QTL is not one of the genes with an induced mutation in TDA1(4) but rather a background allele of the S288c parent strain, causing lower ethyl acetate production compared to the ER7A parent. Hence, we performed RHA with all 23 ORFs containing SNPs between S288c and ER7A in the ER7A/S288c hybrid strain (Supplementary Fig S4). Only the hybrid expressing the *TPS1*<sup>S288c</sup> allele showed a significantly reduced ethyl acetate production ( $30.9 \pm 3.6$  mg/L) compared to the hybrid expressing the *TPS1*<sup>ER7A</sup> allele ( $46.7 \pm 5.1$  mg/L) (Fig. 6C). We further confirmed the *TPS1*<sup>S288c</sup> involvement in ethyl acetate reduction by allele replacement in the ER7A background (Fig. 6D). A significant reduction in ethyl acetate production was observed when *TPS1*<sup>ER7A</sup> was replaced by *TPS1*<sup>S288c</sup> in the ER7A strain ( $48.9$



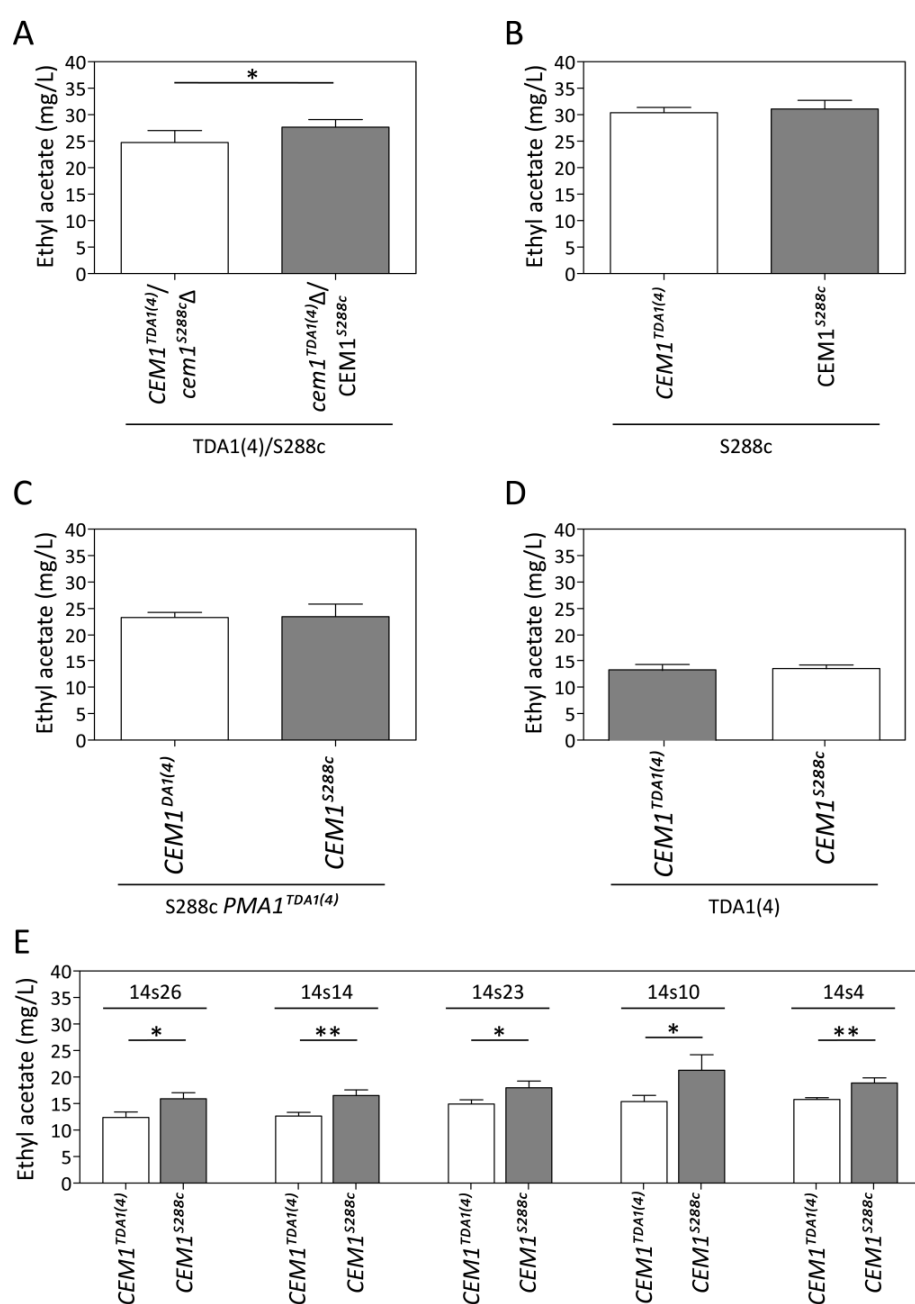
**FIGURE 6. Reciprocal hemizyosity analysis (RHA) and allele exchange of the *PMA1* and *TPS1* alleles.** Ethyl acetate production in semi-anaerobic fermentations after 96 h. **(A)** RHA of *PMA1* in the TDA1(4)/S288c hybrid strain. **(B)** Allele replacement of *PMA1*<sup>S288c</sup> by *PMA1*<sup>TDA1(4)</sup> in the S288c strain. **(C)** RHA of *TPS1* in the ER7A/S288c hybrid strain. **(D)** Allele replacement of *TPS1*<sup>ER7A</sup> by *TPS1*<sup>S288c</sup> in the ER7A strain. The results were analysed with an unpaired, two-sided Student t-test with p-value  $\leq 0.001$  (\*\*\*).



$\pm 6.8$  mg/L compared to  $32.2 \pm 5.3$  mg/L ethyl acetate). The *TPS1* allele of ER7A contains only one single and synonymous mutation compared to the S288c allele, located 1152 bp (V384V) from the start codon. *TPS1*<sup>ER7A</sup>, however, also contained SNPs upstream (-226, relative to the start codon) and downstream (+61 and +130, relative to the stop codon) of the ORF. This suggests that differences in the expression level of *TPS1*<sup>ER7A</sup> might be responsible for the higher ethyl acetate production in ER7A.

For chromosome V, seven intragenic and nine intergenic SNPs, located in the most strongly linked 150 kb region, were analyzed for possible causative character (Fig. 5C). None of the TDA1(4)/S288c RHA strains constructed for these SNPs showed a significant difference in ethyl acetate production (Supplementary Fig. S5), except for *CEM1* (24.7

$\pm 2.3$  mg/L compared to  $27.7 \pm 1.4$  mg/L) (Fig. 7A). Sequence comparison of the *CEM1* allele from TDA1(4) with that of the S288c reference strain revealed one SNP in the *CEM1* ORF (position 278369 bp on chromosome V) resulting in an A420T amino acid change. Allele replacement of *CEM1* in the S288c strain, however, did not cause any significant change in ethyl acetate production (Fig. 7B). We also constructed an S288c strain in which both *CEM1* and *PMA1* alleles were replaced with the TDA1(4) alleles, in order to investigate a possible dependence of the *CEM1*<sup>TDA1(4)</sup> allele on the *PMA1*<sup>TDA1(4)</sup> allele. Its ethyl acetate production was compared with that of the strain with only the *PMA1*<sup>TDA1(4)</sup> allele replaced (Fig. 7C). However, again no difference in ethyl acetate production was observed. We further explored a possible epistatic interaction with an-



other, yet unidentified, TDA1(4) allele. For this purpose, we replaced the *CEM1* allele in TDA1(4) by the S288c allele. Again no difference in ethyl acetate production could be observed between strains containing either *CEM1*<sup>TDA1(4)</sup> or *CEM1*<sup>S288c</sup> (Fig. 7D). The *CEM1* gene product is located in the mitochondria [48] and expression of the low ethyl acetate production phenotype might be dependent on proper mitochondrial function, which was lost after the 4<sup>th</sup> round of mutagenesis in TDA1(4). When we assessed the presence of the *CEM1*<sup>TDA1(4)</sup>/*CEM1*<sup>S288c</sup> and *PMA1*<sup>TDA1(4)</sup>/*PMA1*<sup>S288c</sup> alleles in 44 segregants of the hybrid TDA1(4)/S288c, the five lowest ethyl acetate producing segregants all possessed both the *PMA1*<sup>TDA1(4)</sup> and *CEM1*<sup>TDA1(4)</sup> alleles (Table 1). All five segregants also had proper mitochondrial function as assessed by growth on YPSEG. This suggests that *PMA1*<sup>TDA1(4)</sup> and *CEM1*<sup>TDA1(4)</sup> are both indeed important for low ethyl acetate production and that the contribution of *CEM1*<sup>TDA1(4)</sup> depends on mitochondrial function. Since these five segregants apparently contained most of the causative alleles for low ethyl acetate production present in TDA1(4), we replaced *CEM1*<sup>TDA1(4)</sup> in these segregants by *CEM1*<sup>S288c</sup>. This resulted in a significant increase in ethyl acetate production in all five segregants (Fig. 7E). This result confirmed that the induced mutation in *CEM1* is causative for low ethyl acetate production and that it is dependent on proper mitochondrial function. Replacement of the *CEM1* allele in the S288c strain, which has proper mitochondrial function, did not cause an observable effect. Hence, it is likely that the induced mutation in *CEM1* is not only dependent on proper mitochondrial function but also on another, yet unidentified, causative induced mutation. This induced mutation is not present in S288c but is assumed to be present in all five TDA1(4)/S288c segregants with the lowest ethyl acetate production. In addition, these results showed that causative induced mutations can still be identified in the genetic analysis in spite of the fact that the mutation is unable to cause an observable phenotypic effect in the original strain with multiple induced mutations obtained by genomic saturation mutagenesis.

## DISCUSSION

The aim of this work was to explore whether the main disadvantage of classical mutant screens, i.e. the inability of isolating mutants in non-selectable, polygenic phenotypes, could be addressed by combining genomic saturation mutagenesis and QTL mapping based on pooled-segregant whole-genome sequence analysis. The first challenge was to accumulate enough mutations in the genome of a single haploid strain so as to create a significant chance that the multiply mutated strain will be affected in many non-selectable phenotypes while maintaining genetic proficiency. For that purpose, we optimized a methodology to accumulate a high number of mutations in a single haploid strain by applying multiple rounds of EMS mutagenesis. Intuitively, one would expect the strains with higher numbers of mutations to become affected in viability or vitality and therefore at some point strains with lower numbers of

mutations to dominate in the surviving populations. Hence, genomic saturation mutagenesis would be limited by counterselection of the preferred multiply-mutated strains. In spite of this, we were able to obtain two haploid strains with about 900 mutations and given the steady increase in the number of mutations in the subsequent mutagenesis rounds we expect that considerably higher numbers of artificial mutations can be induced in a single genome, while still retaining the mating capacity required for genetic mapping. Moreover, most of the induced mutations were located at genomic loci not affected by natural variation, at least in the 48 strains for which genome sequence data are now available at SGD. At the genomic loci that were affected by natural variation, the induced mutations had low similarity with the nucleotide variation in these 48 strains as compared to the S288c sequence. These induced mutations can thus be used to unravel underlying genetic elements that would not be discovered by analyzing natural phenotypic diversity.

To estimate the number of mutations accumulated, we initially used the classical procedures based on phenotypic screening of auxotrophic, temperature sensitive and respiratory deficiency mutations. This is similar to the assays of mutation frequency in the literature, which are based on mutations causing a selectable phenotype in single genes [49, 50]. In our case, we selected single colonies from the whole population of surviving cells and tested each colony separately for the presence of auxotrophic, temperature sensitive and respiratory deficiency mutations. This had the advantage that many more genes could be assessed for the presence of mutations. Although we could use this approach to optimize the mutagenesis method and to demonstrate a steady increase in the number of mutations likely accumulated, we cannot exclude phenotypic bias in the number of mutations inferred. On the other hand, the strong drop in the cost of whole genome sequencing now allows to estimate the number of mutations accumulated directly and more reliably in multiple isolates after each mutagenesis round. Multiple rounds of EMS mutagenesis turned out to be much more efficient than multiple rounds of UV mutagenesis in inducing high numbers of mutations in single haploid strains when no direct selection for the mutants was applied.

The biggest challenge of the genomic saturation mutagenesis was the maintenance of genetic proficiency in the strains with multiple induced mutations. Since the goal of the new methodology is to use the multiply mutated strains for elucidating the genetic basis of non-selectable, polygenic traits by QTL mapping, the mutant haploid strains must maintain mating capacity with a haploid reference strain, sporulation capacity of the diploid strain and spore viability of the segregants. Strains that lost mating capacity could be eliminated quite easily by counterselection for a combination of a prototrophic marker and geneticin resistance. Sporulation capacity of the diploid was usually maintained. This is likely due to complementation by the partner genome of any recessive mutation that would compromise sporulation capacity. Only dominant mutations affecting sporulation capacity would present a

**TABLE 1.** Presence of the *PMA1* and *CEM1* alleles in 44 segregants of the TDA1(4)/S288c hybrid. The presence of the *CEM1*<sup>TDA1(4)</sup> or *CEM1*<sup>S288c</sup> and *PMA1*<sup>TDA1(4)</sup> or *PMA1*<sup>S288c</sup> alleles was assessed by Sanger sequencing in 44 segregants of the hybrid TDA1(4)/S288c strain. The level of ethyl acetate production in semi-anaerobic fermentations after 96 h is indicated for each segregant.

Segregant	<i>PMA1</i> allele	<i>CEM1</i> allele	Ethyl acetate (mg/L)
14s 26	TDA1(4)	TDA1(4)	11.6 ± 1.2
14s 14	TDA1(4)	TDA1(4)	12.9 ± 1.4
14s 23	TDA1(4)	TDA1(4)	13.8 ± 0.5
14s 10	TDA1(4)	TDA1(4)	13.9 ± 0.6
14s 4	TDA1(4)	TDA1(4)	14.9 ± 0.3
14s 3	TDA1(4)	S288c	15.7 ± 2.3
14s 32	TDA1(4)	S288c	15.9 ± 0.4
14s 37	TDA1(4)	TDA1(4)	16.0 ± 0.4
14s 39	S288c	S288c	16.2 ± 1.4
14s 5	S288c	TDA1(4)	17.5 ± 2.2
14s 8	TDA1(4)	TDA1(4)	17.6 ± 0.3
14s 21	TDA1(4)	S288c	18.0 ± 2.8
14s 18	TDA1(4)	S288c	18.4 ± 0.2
14s 33	TDA1(4)	S288c	19.0 ± 0.1
14s 19	S288c	TDA1(4)	19.2 ± 2.4
14s 31	TDA1(4)	TDA1(4)	19.4 ± 0.5
14s 35	S288c	TDA1(4)	19.5 ± 2.8
14s 15	TDA1(4)	S288c	19.5 ± 2.4
14s 44	TDA1(4)	S288c	19.7 ± 0.2
14s 27	S288c	S288c	19.9 ± 1.8
14s 6	S288c	S288c	20.0 ± 0.7
14s 7	TDA1(4)	S288c	21.0 ± 0.6
14s 42	TDA1(4)	S288c	21.0 ± 0.9
14s 9	S288c	S288c	21.2 ± 1.0
14s 24	S288c	TDA1(4)	21.4 ± 1.0
14s 1	S288c	TDA1(4)	22.4 ± 1.9
14s 25	TDA1(4)	S288c	22.8 ± 3.3
14s 43	S288c	TDA1(4)	23.0 ± 2.9
14s 30	S288c	TDA1(4)	23.3 ± 1.7
14s 34	TDA1(4)	TDA1(4)	23.6 ± 3.3
14s 17	TDA1(4)	S288c	23.8 ± 0.3
14s 11	S288c	TDA1(4)	24.1 ± 0.4
14s 12	TDA1(4)	S288c	25.4 ± 1.5
14s 40	TDA1(4)	TDA1(4)	25.7 ± 1.4
14s 28	S288c	TDA1(4)	27.9 ± 2.2
14s 2	S288c	S288c	28.2 ± 1.2
14s 41	S288c	TDA1(4)	28.2 ± 1.7
14s 38	S288c	S288c	30.0 ± 1.3
14s 13	S288c	S288c	30.5 ± 1.5
14s 16	S288c	TDA1(4)	30.5 ± 0.7
14s 29	S288c	S288c	30.8 ± 0.1
14s 22	S288c	S288c	32.3 ± 0.5
14s 20	S288c	TDA1(4)	34.1 ± 1.5
14s 36	S288c	S288c	37.0 ± 7.8

problem but these are likely much more rare than recessive mutations. Surprisingly, the most challenging problem was the maintenance of spore viability in the segregants. A possible explanation is that the persistent mutagenesis causes accumulation of multiple lethal mutations in essential genes, which are suppressed by a lower number or even by only one suppressor mutation. For instance, in the cAMP-PKA pathway, a single mutation in *BCY1*, the gene encoding the regulatory subunit of protein kinase A, is able to suppress lethality caused by inactivation of *CYR1*, encoding adenylate cyclase, *RAS1/RAS2* or *CDC25/SDC25*, but also of lethal mutations in all these genes together [51]. When this happens for multiple combinations of lethal and suppressor mutations, as soon as a spore inherits a single lethal mutation without its suppressor mutation, it will be dead. Hence, with an increasing number of mutations accumulated in further mutagenesis rounds, a single suppressor mutation present could suppress the lethal effect of an increasing number of lethal mutations introduced. When this happens with multiple suppressor mutations, it would rapidly compromise the possibility of isolating viable spores after crossing with another strain since two of the spores in a tetrad will always lack the suppressor mutation. In spite of this problem, we were able to maintain sufficient spore viability to allow genetic analysis with strains accumulating about 900 mutations. It is important to emphasize that mutations in essential genes are not necessarily completely inactivating mutations. They can also affect the expression level, the regulation or the cellular location of the gene product, and in this way cause specific phenotypes. Hence, we feel confident that the current number of about 900 mutations does not yet represent real saturation and that accumulation in a single strain of much higher numbers of useful mutations, i.e. mutations with a phenotypic effect, while maintaining genetic proficiency, is thus likely possible. This would create a strain collection useful for screening many non-selectable phenotypes, not only at the cellular but also at the subcellular or biochemical level, such as misregulation of cellular location or alteration of post-translational modification of specific proteins.

In the current study, we have screened the two strains with multiple induced mutations for deficiencies in flavor compound production, a non-selectable trait of crucial importance in the alcoholic beverage industry and of which the metabolic basis and regulation are only poorly understood [29, 52, 53]. We found that the TDA1(4) multiply mutated strain had noteworthy low ethyl acetate production. The ease with which a deficiency in a target trait was identified shows the usefulness of this approach. It can be easily predicted that with higher numbers of mutations accumulated in single haploid strains, a small collection of multiply mutated haploid strains may be affected in a very high number of phenotypes that are all compatible with the requirements for genetic analysis. Abolishment or even a drastic change in the phenotype is not required. A statistically significant difference that is reliable enough to perform the subsequent genetic analysis is sufficient. Remarkably, it is not even required that the phenotype is observa-

ble in the multiply mutated strain. This is shown in the current study by the mutation in *CEM1* that requires proper mitochondrial function to have effect, while the parent multiply mutated strain no longer had proper mitochondrial function. Since proper mitochondrial activity was recovered in all of the viable segregants, the *CEM1* mutation could exert its effect in these segregants. This shows that the accumulation of useful mutations in the multiply mutated strain is not limited by the fact that these mutations can exert a recognizable phenotypic effect. The only requirement is that the mutation can exert its effect in a sufficient number of segregants. If this is the case, the locus of the mutation will be identified in the QTL mapping. In practice, efficient QTL mapping by pooled-segregant whole genome sequence analysis requires only about 30 selected segregants [12, 18].

In our work we have crossed the multiply mutated haploid strain TDA1(4) with an unrelated haploid strain, ER7A. This provides the advantage that thousands of SNPs are available for efficient QTL mapping. On the other hand, it also makes it possible that genetic background mutations rather than induced mutations are responsible for part of the difference in the phenotype. This was the case in our study for the *TPS1* gene. In principle, however, this does not have to be a disadvantage since it simply provides an additional way to discover gene products affecting the phenotype of interest. Since the genome sequence of the multiply mutated strain and its parent strain are known, these mutations can be identified easily.

In the present study, we have identified three new genes affecting ethyl acetate production. On chromosome VII we identified *PMA1*<sup>TDA1(4)</sup> as causative allele for lowering ethyl acetate production. The mutation causes an M152I amino acid change in transmembrane domain II of Pma1, an essential H<sup>+</sup>-ATPase that maintains the electrochemical proton gradient over the plasma membrane. Transmembrane domain I and II are linked by a short intracellular loop and form a conformationally-sensitive structure. Many recessive mutations in transmembrane domain I or II are lethal or cause a reduction in H<sup>+</sup>-ATPase activity [54]. An amino acid change in this region could thus have an important effect on H<sup>+</sup>-ATPase activity and in this way affect many traits, including ethyl acetate production. The precise molecular mechanism, however, remains unclear.

As previously mentioned, the *CEM1*<sup>TDA1(4)</sup> causative allele presents an interesting case since it was identified as causative allele in the genetic analysis with the segregants, although its effect could not be observed in the original TDA1(4) strain. The *CEM1* gene product is a beta-keto-acyl synthase located in the mitochondria and required for mitochondrial respiration [48]. This may explain why the effect of the *CEM1*<sup>TDA1(4)</sup> allele on ethyl acetate production is dependent on proper mitochondrial function. The *CEM1* gene product has never been connected to ethyl acetate biosynthesis. It is a homolog of the *FAS2* gene product, the alpha subunit of fatty acid synthetase, which synthetizes aceto-acetyl(-ACP) in the fatty acid biosynthesis pathway. Since both fatty acid biosynthesis and ethyl acetate biosynthesis use acetyl-CoA as substrate, one possible explana-

tion for the reduction of ethyl acetate production by the *CEM1*<sup>TDA1(4)</sup> allele, is lowering of the acetyl-CoA level due to enhanced activity of the fatty acid biosynthesis pathway. This result indicates the potential of the novel methodology proposed in this paper to obtain new insight in the functioning and interaction between metabolic pathways.

A number of mutant alleles identified with the current methodology will undoubtedly only very indirectly affect the trait-of-interest, but this also applies to classical mutant screens using selectable phenotypes. This situation may be true for the *PMA1* gene product although it may also reflect our lack of insight in the regulation of metabolism and the presence of additional unanticipated functions of well-established proteins. This may also be true for the *TPS1* gene product, of which the S288c allele contributed to the low ethyl acetate production in the TDA1(4) strain. *TPS1* encodes trehalose-6-phosphate synthase, the first enzyme of trehalose biosynthesis. Its product, trehalose-6-phosphate, is also involved in allosteric regulation of glycolysis at the level of hexokinase [55]. *Tps1* may thus affect the levels of ethanol and acetyl-CoA, the substrates for ethyl acetate biosynthesis. However, multiple results suggest additional regulatory functions for *Tps1* [56-60].

In conclusion, our work has shown that genomic saturation mutagenesis combined with complex trait polygenic analysis can be used successfully to identify causative alleles underlying non-selectable, polygenic traits. Small collections of independently mutagenized haploid strains, containing very high numbers of induced mutations, could be screened for any non-selectable phenotype, including structural phenotypes at the subcellular level and biochemical phenotypes at the protein or metabolite level. The genetic basis can then be identified by the now well-established methodology for polygenic analysis of complex traits.

## MATERIALS AND METHODS

### Yeast strains, growth conditions, mating and sporulation

All *S. cerevisiae* strains used are listed in Table 2. S288c *ura3::KanMX* was constructed as described [61] and homologous integration verified by polymerase chain reaction (PCR). BY4742 *p0* was obtained according to Fox *et al.* [62]. Yeast strains were grown at 30°C and 200 rpm in YPD medium containing 1% yeast extract, 2% Bacto peptone and 2% glucose. Where appropriate the medium was supplemented with 200 µg/mL geneticin or 100 µg/mL nourseothricin. Other media used are SD [0.171% Yeast Nitrogen Base w/o amino acids (Difco), 0.5% ammonium sulphate (or 0.1% glutamate if supplemented with geneticin) and 2% glucose at pH 6.5], YPSEG (1% yeast extract, 2% Bacto peptone, 1% succinic acid, 2% glycerol and 2% ethanol at pH 5.5) and sporulation medium (0.5% potassium acetate at pH 6.0). 1.5% Bacto agar was used to make solid nutrient plates. All percentages indicate weight per volume, unless stated otherwise. Mating, sporulation and dissection of tetrads were carried out according to standard protocols [63].

### Mutagenesis

Yeast cells were grown to stationary phase, harvested, washed twice and re-suspended for EMS mutagenesis in 1 mL 0.1 M sodium phosphate buffer (pH 7.0) at a concentration of  $2 \times 10^8$  cells/mL. The samples were incubated with a final concentration of 3% (v/v) EMS for different time periods at 30°C and 200 rpm. EMS was neutralized by washing twice with 5% sodium thiosulphate [2]. The pelleted cells were re-suspended in 1 mL sterile water and spread on YPD plates at appropriate dilutions. For UV mutagenesis the samples were re-suspended in 1 mL sterile water at a concentration of  $2 \times 10^8$  cells/mL, spread on 120 mm x 120 mm YPD plates at appropriate dilutions and irradiated with UV light at different energy levels (GS Gene Linker, BIO-RAD) [66]. Single colonies were manually replated on YPD plates and allowed to grow for 3 days. Auxotrophic mutations were determined by re-plating the colonies

TABLE 2. *S. cerevisiae* strains used in this study.

Strain	Genotype/Description	Source
S288c	<i>MATa</i> prototroph	[74]
S288c <i>MATa</i>	Mating type switched S288c	This study
S288c 2n	Isogenic S288c diploid	This study
S288c <i>ura3::KANMX</i>	<i>MATa ura3::KANMX</i>	This study
BY4742 <i>p0</i>	<i>MATa his3Δ1 leu2Δ0 ura3Δ0 lys2Δ0 p0</i>	This study
ER7A	Segregant 7A of Ethanol Red, <i>Mata</i>	[17]
<i>ade2Δ</i>	(BY4741) <i>MATa his3Δ1 leu2Δ0 ura3Δ0 met15Δ0 ade2::KANMX</i>	[65]
<i>arg2Δ</i>	(BY4741) <i>MATa his3Δ1 leu2Δ0 ura3Δ0 met15Δ0 arg2::KANMX</i>	[65]
<i>his3Δ1</i>	(BY4741) <i>MATa his3Δ1 leu2Δ0 ura3Δ0 met15Δ0</i>	[64]
<i>Leu2Δ0</i>	(BY4741) <i>MATa his3Δ1 leu2Δ0 ura3Δ0 met15Δ0</i>	[64]
<i>lys2Δ0</i>	(BY4742) <i>MATa his3Δ1 leu2Δ0 ura3Δ0 lys2Δ0</i>	[64]
<i>met15Δ0</i>	(BY4741) <i>MATa his3Δ1 leu2Δ0 ura3Δ0 met15Δ0</i>	[64]
<i>ura3Δ0</i>	(BY4741) <i>MATa his3Δ1 leu2Δ0 ura3Δ0 met15Δ0</i>	[64]
<i>trp1Δ</i>	(BY4741) <i>MATa his3Δ1 leu2Δ0 ura3Δ0 met15Δ0 trp1::KANMX</i>	[65]

on eight drop-out media (SD-ADE, SD-ARG, SD-HIS, SD-LEU, SD-LYS, SD-MET, SD-TRP, SD-URA) at 30°C. To determine temperature sensitivity and respiratory deficiency the colonies were re-plated on YPD at 37°C and on YPSEG at 30°C, respectively. YPD and SD plates incubated at 30°C served as control. For EMS and UV mutagenesis of the S288c strain a total of respectively 1536 and 576 colonies were analyzed at each indicated time interval. For the strains with multiple induced mutations 288 colonies were analyzed. The same single colonies were allowed to mate with S288c *MAT $\alpha$  ura3::KanMX* in liquid YPD, diploids selected on SD-URA plates supplemented with 200  $\mu$ g/mL geneticin, sporulated for 7 days and 20 random tetrads dissected to assess spore viability.

### General molecular biology methods

Yeast cells were transformed with the LiOAc/SS-DNA/PEG method [67, 68]. Yeast genomic DNA was extracted with Phenol/Chloroform/Isoamyl-alcohol (25:24:1) [69] and further purified with ethanol precipitation if required. PCR was performed with high-fidelity polymerases Phusion and Q5 (New England Biolabs) or ExTaq (Takara) and Taq (New England Biolabs) for diagnostic purposes. Sanger sequencing [70] was performed by the VIB Genetic Service Facility (Antwerp). The sequences were analyzed with Vector NTI (Invitrogen) or CLC DNA workbench (CLC Bio) software. Lists of primers and plasmids are provided as supplementary tables S4 and S5, respectively.

### Fermentation conditions

The cells were pre-grown in 3 mL YPD medium for 24 h and subsequently grown to stationary phase for 2 days in 40 mL YPD medium. Cells were washed with sterile, distilled water and used to inoculate 100 mL 10% YPD medium in 100 mL flasks. Fermentations were carried out at 30°C with stirring at 80 rpm with a magnetic rod (30 x 6 mm) and a water lock on top to create semi-anaerobic conditions. Samples for gas chromatographic analysis were prepared by centrifugation and immediate cooling (on ice) of the supernatant in an air-tight container at indicated times.

### Headspace gas chromatography

Headspace gas chromatography coupled with flame ionization detection (GC-FID) was used for the measurement of ethyl acetate in the fermentation samples. Samples of 5 mL were collected in 15 mL pre-cooled glass tubes and analyzed with a calibrated Trace Ultra gas chromatograph (Thermo Scientific) equipped with a Stabilwax column (length 60 m, internal diameter 0.25 mm, layer thickness 0.25  $\mu$ m; Restek) and coupled to a Triplus RSH headspace autosampler (Thermo Scientific). Samples were heated for 10 min at 60 °C, with agitation, in the headspace autosampler. The injection block and flame ionization detector were kept at constant temperatures of 220°C and 250°C, respectively, and helium was used as a carrier gas. The oven temperature was held at 50°C for 5 min, then increased to 240°C at a rate of 15°C per min. Results were analyzed with Chromquest software (Thermo Scientific).

### Whole genome sequence analysis of mutant strains

High molecular weight DNA (5  $\mu$ g; ~20 kb fragments) was isolated according to Johnston and Aust [71]. The purity of the DNA sample was estimated from UV measurement (260/280=1.7-2.0). The DNA samples were provided to BGI

(Hong Kong, China) for whole genome sequence analysis by Illumina technology at ~60 fold nominal coverage. The NGSEP pipeline [47] version 2.1.3 was used to align reads to the reference genome S288c, and to identify and filter variants relative to S288c. In brief, reads were aligned to the reference genome using bowtie2 [72] with minimum insert size 350, maximum insert size 550 and allowing up to 3 valid alignments per read. Then alignments were sorted and variants were identified by the FindVariants command of NGSEP using its recommended parameters for WGS data: 1) Maximum base quality score 30; 2) Minimum genotype quality 40; 3) At most two alignments starting at each genomic site were considered; 4) Sample ploidy equal to 1; 5) Prior heterozygosity rate 10E-6. Variants falling into repetitive regions or regions with predicted copy number variation were filtered out using the FilterVCF command of NGSEP.

### Pooled-segregant whole-genome sequence analysis

After crossing the two parent strains, TDA1(4) and ER7A, 41 superior segregants (ethyl acetate production <20 mg/L after 96 h of fermentation in two biological replicates) were used to assemble the 'selected pool'. A 'non-selected', i.e. random pool of equal size was also assembled. All segregants were grown separately and the pools were made by combining equal amounts of cells based on OD<sub>600</sub>. Whole-genome sequence analysis was carried out in a similar way as for the mutant strains at ~40 fold nominal coverage. Sequence alignment was performed using SeqMan NGen 4 (DNASTAR) and assembly and mapping with Lasergene 10 (DNASTAR). A variant was called if it was covered by at least 25 sequence reads with a quality score of at least 10 and minimum 90% of the reads contained the variant. QTL analysis was carried out as described previously [18]. Essentially, SNP variant frequencies were calculated by dividing the number of the variant by the total number of reads at each SNP location. Deviation from 50%, either above or below, was a sign of a one-sided SNP variant segregation, indicating a genetic linkage to the trait of interest.

### Detection of SNP markers by allele specific PCR in individual segregants

SNPs were scored by allele specific PCR in individual segregants of TDA1(4)/ER7A, that accumulated <16.8 mg/L ethyl acetate after 96 h fermentation. The forward and reverse primer each contained either the S288c or ER7A nucleotide as the 3' terminal nucleotide and were spaced between 400 and 1700 bp apart. The annealing temperature was optimized using genomic DNA of both parents to allow only hybridization with primers containing the exact complement. The SNP data in the individual segregants were analyzed using the binomial exact test.

### Reciprocal hemizygosity analysis (RHA)

For RHA analysis [21] the exact ORF or the two neighboring ORFs for intragenic or intergenic variants, respectively, were deleted with *NatMX4* [61] in the TDA1(4)/S288c or S288c/ER7A hybrid diploid strains. Homologous integration was verified by PCR and the remaining, i.e. non-deleted, allele identified by Sanger sequencing. RHA analysis was performed with two independent isolates of all diploids tested and fermentations were repeated twice.

### Allele replacement

Reciprocal allele replacements at the *CEM1* locus in S288c and TDA1(4) and the *TPS1* locus in ER7A were achieved by a two step-method. In the first step, *CEM1* or *TPS1* was deleted with a deletion cassette amplified by PCR and containing both the selectable *NatMX4* and counter-selectable *GIN11* marker [17]. After transformation, positive clones were selected on nourseothricin containing YPD plates. The presence of the deletion cassette was verified by PCR. In the second step, either the S288c or TDA1(4) allele for *CEM1* or the S288c allele for *TPS1* were amplified from genomic DNA and inserted at the deletion site. Positive clones were selected on galactose containing medium by inducing the counter-selectable *GIN11* marker [73, 74]. Integration was assessed by nourseothricin sensitivity and the inserted allele verified by PCR and Sanger sequencing. The allele replacement of the essential gene *PMA1* was done similarly to *CEM1* in an isogenic S288c diploid. The latter strain was generated by mating type switching [75] and a subsequent backcross to the original S288c *MATa* strain. In the S288c *MATa* strain the HO-gene was expressed from the pFL39 GAL1 HO *KanMX* plasmid to induce mating type switch. Mating type was determined by PCR for the *MAT* locus [76]. The near-isogenic S288c *PMA1/PMA1*<sup>TDA1(4)</sup> diploid was sporulated to generate S288c *MATa PMA1*<sup>TDA1(4)</sup>. Double allele replacement of both *CEM1* and *PMA1* in S288c *MATa* was accomplished by mating S288c *MATa CEM1*<sup>TDA1(4)</sup> with S288c *MATa PMA1*<sup>TDA1(4)</sup> and subsequent sporulation of the diploid. Correct presence of the alleles was verified by Sanger sequencing. For all allele replacements, except the double replacement, 2 independent clones were tested and fermentations were repeated twice.

### Data access

Sequencing data have been deposited at the SRA database (NCBI), <http://www.ncbi.nlm.nih.gov/sra>, with the account number SRP044392.

### REFERENCES

- Forsburg SL (2001). The art and design of genetic screens: yeast. *Nat Rev Genet* 2(9): 659-668.
- Amberg DC, Burke DJ, Strathern JN (2005). 2005 edition. *Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Mortimer RK, Hawthorne DC (1975). Genetic mapping in yeast. *Methods Cell Biol* 11:221-233.
- Attfield PV (1997). Stress tolerance: the key to effective strains of industrial baker's yeast. *Nat Biotechnol* 15(13): 1351-1357.
- Pretorius IS, Curtin CD, Chambers PJ (2012). The winemaker's bug: From ancient wisdom to opening new vistas with frontier yeast science. *Bioeng Bugs* 3(3): 147-156.
- Steensels J, Snoek T, Meersman E, Nicolino MP, Voordeckers K, Verstrepen KJ (2014). Improving industrial yeast strains: exploiting natural and artificial diversity. *FEMS Microbiol Rev* 38(5): 947-995.
- Liti G, Louis EJ (2012). Advances in quantitative trait analysis in yeast. *PLoS Genet* 8(8): e1002912.
- Swinnen S, Thevelein JM, Nevoigt E (2012). Genetic mapping of quantitative phenotypic traits in *Saccharomyces cerevisiae*. *FEMS Yeast Res* 12(2): 215-227.
- Parts L, Cubillos FA, Warringer J, Jain K, Salinas F, Bumpstead SJ, Molin M, Zia A, Simpson JT, Quail MA, Moses A, Louis EJ, Durbin R, Liti G (2011). Revealing the genetic structure of a trait by sequencing a population under selection. *Genome Res* 21(7): 1131-1138.
- Ehrenreich IM, Torabi N, Jia Y, Kent J, Martis S, Shapiro JA, Gresham D, Caudy AA, Kruglyak L (2010). Dissection of genetically complex traits with extremely large pools of yeast segregants. *Nature* 464(7291): 1039-1042.
- Hubmann G, Mathe L, Foulquie-Moreno MR, Duitama J, Nevoigt E, Thevelein JM (2013). Identification of multiple interacting alleles conferring low glycerol and high ethanol yield in *Saccharomyces cerevisiae* ethanolic fermentation. *Biotechnol Biofuels* 6(1): 87.
- Pais TM, Foulquie-Moreno MR, Hubmann G, Duitama J, Swinnen S, Goovaerts A, Yang Y, Dumortier F, Thevelein JM (2013). Comparative polygenic analysis of maximal ethanol accumulation capacity and tolerance to high ethanol levels of cell proliferation in yeast. *PLoS Genet* 9(6): e1003548.

### ACKNOWLEDGMENTS

We wish to thank Georg Hubmann, Jean-Paul Meijnen and all other members of the Genetic Analysis Group for stimulating discussions, Ilse Palmans, Paul Vandecruys and Jan Wouters for technical support, Alessandro Fiori for providing the BY4742 p0 strain and Nico Vangoethem for informatics support. This work has been supported by SBO grants (IWT 50148 and IWT 90043) from IWT-Flanders, Concerted Research Actions (GOA/13/006), IOF-Knowledge platform (IKP/10/002 ZKC 1836) and BOF-Program financing (project NATAR) to JMT.

### SUPPLEMENTAL MATERIAL

All supplemental data for this article are available online at [www.microbialcell.com](http://www.microbialcell.com).

### CONFLICT OF INTEREST

The authors declare no conflict of interest.

### COPYRIGHT

© 2016 Abt *et al.* This is an open-access article released under the terms of the Creative Commons Attribution (CC BY) license, which allows the unrestricted use, distribution, and reproduction in any medium, provided the original author and source are acknowledged.

Please cite this article as: Tom Den Abt, Ben Souffriau, Maria R. Foulquié-Moreno, Jorge Duitama, and Johan M. Thevelein (2016). Genomic saturation mutagenesis and polygenic analysis identify novel yeast genes affecting ethyl acetate production, a non-selectable polygenic trait. *Microbial Cell* 3(4): 159-175. doi: 10.15698/mic2016.04.491

13. Yang Y, Foulquie-Moreno MR, Clement L, Erdei E, Tanghe A, Schaerlaekens K, Dumortier F, Thevelein JM (2013). QTL analysis of high thermotolerance with superior and downgraded parental yeast strains reveals new minor QTLs and converges on novel causative alleles involved in RNA processing. *PLoS Genet* 9(8): e1003693.
14. Brice C, Sanchez I, Bigey F, Legras JL, Blondin B (2014). A genetic approach of wine yeast fermentation capacity in nitrogen-starvation reveals the key role of nitrogen signaling. *BMC Genomics* 15: 495.
15. Lewis JA, Broman AT, Will J, Gasch AP (2014). Genetic architecture of ethanol-responsive transcriptome variation in *Saccharomyces cerevisiae* strains. *Genetics* 198(1): 369-382.
16. Steyer D, Ambroset C, Brion C, Claudel P, Delobel P, Sanchez I, Erny C, Blondin B, Karst F, Legras JL (2012). QTL mapping of the production of wine aroma compounds by yeast. *BMC Genomics* 13: 573.
17. Hubmann G, Foulquie-Moreno MR, Nevoigt E, Duitama J, Meurens N, Pais TM, Mathe L, Saerens S, Nguyen HT, Swinnen S, Verstrepen KJ, Concilio L, de Troostembergh JC, Thevelein JM (2013). Quantitative trait analysis of yeast biodiversity yields novel gene tools for metabolic engineering. *Metab Eng* 17: 68-81.
18. Swinnen S, Schaerlaekens K, Pais T, Claesen J, Hubmann G, Yang Y, Demeke M, Foulquie-Moreno MR, Goovaerts A, Souvereyns K, Clement L, Dumortier F, Thevelein JM (2012). Identification of novel causative genes determining the complex trait of high ethanol tolerance in yeast using pooled-segregant whole-genome sequence analysis. *Genome Res* 22(5): 975-984.
19. Marullo P, Aigle M, Bely M, Masneuf-Pomarede I, Durrens P, Dubourdieu D, Yvert G (2007). Single QTL mapping and nucleotide-level resolution of a physiologic trait in wine *Saccharomyces cerevisiae* strains. *FEMS Yeast Res* 7(6): 941-952.
20. Ibstedt S, Stenberg S, Bages S, Gjuvsland AB, Salinas F, Kourtchenko O, Samy JK, Blomberg A, Omholt SW, Liti G, Beltran G, Warringer J (2015). Concerted evolution of life stage performances signals recent selection on yeast nitrogen use. *Mol Biol Evol* 32(1): 153-161.
21. Steinmetz LM, Sinha H, Richards DR, Spiegelman JI, Oefner PJ, McCusker JH, Davis RW (2002). Dissecting the architecture of a quantitative trait locus in yeast. *Nature* 416(6878): 326-330.
22. Diezmann S, Dietrich FS (2011). Oxidative stress survival in a clinical *Saccharomyces cerevisiae* isolate is influenced by a major quantitative trait nucleotide. *Genetics* 188(3): 709-722.
23. Hu XH, Wang MH, Tan T, Li JR, Yang H, Leach L, Zhang RM, Luo ZW (2007). Genetic dissection of ethanol tolerance in the budding yeast *Saccharomyces cerevisiae*. *Genetics* 175(3): 1479-1487.
24. Sinha H, David L, Pascon RC, Clauder-Munster S, Krishnakumar S, Nguyen M, Shi G, Dean J, Davis RW, Oefner PJ, McCusker JH, Steinmetz LM (2008). Sequential elimination of major-effect contributors identifies additional quantitative trait loci conditioning high-temperature growth in yeast. *Genetics* 180(3): 1661-1670.
25. Lopez-Martinez G, Margalef-Catala M, Salinas F, Liti G, Cordero-Otero R (2015). *ATG18* and *FAB1* are involved in dehydration stress tolerance in *Saccharomyces cerevisiae*. *PLoS One* 10(3): e0119606.
26. Katou T, Namise M, Kitagaki H, Akao T, Shimoi H (2009). QTL mapping of sake brewing characteristics of yeast. *J Biosci Bioeng* 107(4): 383-393.
27. Salinas F, Cubillos FA, Soto D, Garcia V, Bergstrom A, Warringer J, Ganga MA, Louis EJ, Liti G, Martinez C (2012). The genetic basis of natural variation in oenological traits in *Saccharomyces cerevisiae*. *PLoS One* 7(11): e49640.
28. Ambroset C, Petit M, Brion C, Sanchez I, Delobel P, Guerin C, Chiapello H, Nicolas P, Bigey F, Dequin S, Blondin B (2011). Deciphering the molecular basis of wine yeast fermentation traits using a combined genetic and genomic approach. *G3 (Bethesda)* 1(4): 263-281.
29. Saerens SM, Delvaux FR, Verstrepen KJ, Thevelein JM (2010). Production and biological function of volatile esters in *Saccharomyces cerevisiae*. *Microb Biotechnol* 3(2): 165-177.
30. Meilgaard MC (1991). The flavor of beer. *MBAA Tech Q* 28(3): 132-141.
31. Suomalainen H (1981). Yeast esterases and aroma esters in alcoholic beverages. *J Inst Brew* 87(5): 296-300.
32. Peddie HAB (1990). Ester formation in brewery fermentations. *J Inst Brew* 96(5): 327-331.
33. Nykänen L (1986). Formation and occurrence of flavor compounds in wine and distilled beverages. *Am J Enol Vitic* 37(1): 84-96.
34. Saerens SM, Verstrepen KJ, Van Laere SD, Voet AR, Van Dijck P, Delvaux FR, Thevelein JM (2006). The *Saccharomyces cerevisiae* *EHT1* and *EEB1* genes encode novel enzymes with medium-chain fatty acid ethyl ester synthesis and hydrolysis capacity. *J Biol Chem* 281(7): 4446-4456.
35. Malcorps P, Dufour JP (1992). Short-chain and medium-chain aliphatic-ester synthesis in *Saccharomyces cerevisiae*. *Eur J Biochem* 210(3): 1015-1022.
36. Fujii T, Nagasawa N, Iwamatsu A, Bogaki T, Tamai Y, Hamachi M (1994). Molecular cloning, sequence analysis, and expression of the yeast alcohol acetyltransferase gene. *Appl Environ Microbiol* 60(8): 2786-2792.
37. Nagasawa N, Bogaki T, Iwamatsu A, Hamachi M, Kumagai C (1998). Cloning and nucleotide sequence of the alcohol acetyltransferase gene (*ATF2*) from *Saccharomyces cerevisiae* Kyokai No. 7. *Biosci Biotechnol Biochem* 62(10): 1852-1857.
38. Verstrepen KJ, Van Laere SD, Vanderhaegen BM, Derdelinckx G, Dufour JP, Pretorius IS, Winderickx J, Thevelein JM, Delvaux FR (2003). Expression levels of the yeast alcohol acetyltransferase genes *ATF1*, *Lg-ATF1*, and *ATF2* control the formation of a broad range of volatile esters. *Appl Environ Microbiol* 69(9): 5228-5237.
39. Fukuda K, Kiyokawa Y, Yanagiuchi T, Wakai Y, Kitamoto K, Inoue Y, Kimura A (2000). Purification and characterization of isoamyl acetate-hydrolyzing esterase encoded by the *IAH1* gene of *Saccharomyces cerevisiae* from a recombinant *Escherichia coli*. *Appl Microbiol Biotechnol* 53(5): 596-600.
40. Fukuda K, Yamamoto N, Kiyokawa Y, Yanagiuchi T, Wakai Y, Kitamoto K, Inoue Y, Kimura A (1998). Balance of activities of alcohol acetyltransferase and esterase in *Saccharomyces cerevisiae* is important for production of isoamyl acetate. *Appl Environ Microbiol* 64(10): 4076-4078.
41. Lilly M, Bauer FF, Lambrechts MG, Swiegers JH, Cozzolino D, Pretorius IS (2006). The effect of increased yeast alcohol acetyltransferase and esterase activity on the flavour profiles of wine and distillates. *Yeast* 23(9): 641-659.
42. Kitagaki H, Kitamoto K (2013). Breeding research on sake yeasts in Japan: history, recent technological advances, and future perspectives. *Annual review of food science and technology* 4: 215-235.
43. Shiwa Y, Fukushima-Tanaka S, Kasahara K, Horiuchi T, Yoshikawa H (2012). Whole-Genome Profiling of a Novel Mutagenesis Technique Using Proofreading-Deficient DNA Polymerase delta. *Int J Evol Biol* 2012(2012): 1-8.
44. Flibotte S, Edgley ML, Chaudhry I, Taylor J, Neil SE, Rogula A, Zapf R, Hirst M, Butterfield Y, Jones SJ, Marra MA, Barstead RJ, Moerman



- DG (2010). Whole-genome profiling of mutagenesis in *Caenorhabditis elegans*. **Genetics** 185(2): 431-441.
45. Timmermann B, Jarolim S, Russmayer H, Kerick M, Michel S, Kruger A, Bluemlein K, Laun P, Grillari J, Lehrach H, Breitenbach M, Ralsler M (2010). A new dominant peroxiredoxin allele identified by whole-genome re-sequencing of random mutagenized yeast causes oxidant-resistance and premature aging. **Aging (Albany NY)** 2(8): 475-486.
46. Blumenstiel JP, Noll AC, Griffiths JA, Perera AG, Walton KN, Gilliland WD, Hawley RS, Staehling-Hampton K (2009). Identification of EMS-induced mutations in *Drosophila melanogaster* by whole-genome sequencing. **Genetics** 182(1): 25-32.
47. Duitama J, Quintero JC, Cruz DF, Quintero C, Hubmann G, Foulque-Moreno MR, et al. (2014). An integrated framework for discovery and genotyping of genomic variants from high-throughput sequencing experiments. **Nucleic Acids Res** 42(6): e44.
48. Harington A, Herbert CJ, Tung B, Getz GS, Slonimski PP (1993). Identification of a new nuclear gene (*CEM1*) encoding a protein homologous to a beta-keto-acyl synthase which is essential for mitochondrial respiration in *Saccharomyces cerevisiae*. **Mol Microbiol** 9(3): 545-555.
49. Whelan WL, Gocke E, Manney TR (1979). The *CAN1* locus of *Saccharomyces cerevisiae*: fine-structure analysis and forward mutation rates. **Genetics** 91(1): 35-51.
50. Boeke JD, LaCrute F, Fink GR (1984). A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. **Mol Gen Genet** 197(2): 345-346.
51. Conrad M, Schothorst J, Kankipati HN, Van Zeebroeck G, Rubio-Teixeira M and Thevelein JM (2014). Nutrient sensing and signaling in the yeast *Saccharomyces cerevisiae*. **FEMS Microbiol Rev** 38(2): 254-299.
52. Boulton C, Quain D (2006). Brewing yeast and fermentation. 1<sup>st</sup> edition. **Blackwell Science, Oxford, U.K.**
53. Liu S-Q, Holland R, Crow VL (2004). Esters and their biosynthesis in fermented dairy products: a review. **Int Dairy J** 14(11): 923-945.
54. Morsomme P, Slayman CW, Goffeau A (2000). Mutagenic study of the structure, function and biogenesis of the yeast plasma membrane H(+)-ATPase. **Biochim Biophys Acta** 1469(3): 133-157.
55. Thevelein JM, Hohmann S (1995). Trehalose synthase: guard to the gate of glycolysis in yeast? **Trends Biochem Sci** 20(1): 3-10.
56. Hohmann S, Bell W, Neves MJ, Valckx D, Thevelein JM (1996). Evidence for trehalose-6-phosphate-dependent and -independent mechanisms in the control of sugar influx into yeast glycolysis. **Mol Microbiol** 20(5): 981-991.
57. Bonini BM, Van Vaec C, Larsson C, Gustafsson L, Ma P, Winderickx J, Van Dijck P, Thevelein JM (2000). Expression of *Escherichia coli* *otsA* in a *Saccharomyces cerevisiae* *tps1* mutant restores trehalose 6-phosphate levels and partly restores growth and fermentation with glucose and control of glucose influx into glycolysis. **Biochem J** 350 (Pt 1): 261-268.
58. Bonini BM, Van Dijck P, Thevelein JM (2003). Uncoupling of the glucose growth defect and the deregulation of glycolysis in *Saccharomyces cerevisiae* *Tps1* mutants expressing trehalose-6-phosphate-insensitive hexokinase from *Schizosaccharomyces pombe*. **Biochim Biophys Acta** 1606(1-3): 83-93.
59. van Vaec C, Wera S, van Dijck P, Thevelein JM (2001). Analysis and modification of trehalose 6-phosphate levels in the yeast *Saccharomyces cerevisiae* with the use of *Bacillus subtilis* phosphotrehalase. **Biochem J** 353(Pt 1): 157-162.
60. Walther T, Mtimet N, Alkim C, Vax A, Loret MO, Ullah A, Gancedo C, Smits GJ, Francois JM (2013). Metabolic phenotypes of *Saccharomyces cerevisiae* mutants with altered trehalose 6-phosphate dynamics. **Biochem J** 454(2): 227-237.
61. Guldener U, Heck S, Fielder T, Beinbauer J, Hegemann JH (1996). A new efficient gene disruption cassette for repeated use in budding yeast. **Nucleic Acids Res** 24(13): 2519-2524.
62. Fox TD, Folley LS, Mulero JJ, McMullin TW, Thorsness PE, Hedin LO, Costanzo MC (1991). Analysis and manipulation of yeast mitochondrial genes. **Methods Enzymol** 194: 149-165.
63. Sherman F, Hicks J (1991). Micromanipulation and dissection of asci. **Methods Enzymol** 194: 21-37.
64. Brachmann CB, Davies A, Cost GJ, Caputo E, Li J, Hieter P, Boeke JD (1998). Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. **Yeast** 14(2): 115-132.
65. Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, Chu AM, Connelly C, Davis K, Dietrich F, Dow SW, El Bakkoury M, Foury F, Friend SH, Gentalen E, Giaever G, Hegemann JH, Jones T, Laub M, Liao H, Liebundguth N, Lockhart DJ, Lucau-Danila A, Lussier M, M'Rabet N, Menard P, et al. (1999). Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. **Science** 285(5429): 901-906.
66. Winston F (2008). EMS and UV mutagenesis in yeast. **Curr Protoc Mol Biol** Chapter 13: Unit 13.13B.
67. Gietz RD, Schiestl RH (2007). High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. **Nat Protoc** 2(1): 31-34.
68. Gietz RD, Schiestl RH, Willems AR, Woods RA (1995). Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. **Yeast** 11(4): 355-360.
69. Hoffman CS, Winston F (1987). A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. **Gene** 57(2-3): 267-272.
70. Sanger F, Coulson AR (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. **J Mol Biol** 94(3): 441-448.
71. Johnston CG, Aust SD (1994). Detection of *Phanerochaete chrysosporium* in soil by PCR and restriction enzyme analysis. **Appl Environ Microbiol** 60(7): 2350-2354.
72. Langmead, B, Salzberg, SL (2012) Fast gapped-read alignment with Bowtie 2. **Nature methods** 9(4): 357-359.
73. Akada R, Hirosawa I, Kawahata M, Hoshida H, Nishizawa Y (2002). Sets of integrating plasmids and gene disruption cassettes containing improved counter-selection markers designed for repeated use in budding yeast. **Yeast** 19(5): 393-402.
74. Akada R, Matsuo K, Aritomi K, Nishizawa Y (1999). Construction of recombinant sake yeast containing a dominant *FAS2* mutation without extraneous sequences by a two-step gene replacement protocol. **J Biosci Bioeng** 87(1): 43-48.
75. Jensen R, Sprague GF, Jr., Herskowitz I (1983). Regulation of yeast mating-type interconversion: feedback control of *HO* gene expression by the mating-type locus. **Proc Natl Acad Sci U S A** 80(10): 3035-3039.
76. Huxley C, Green ED, Dunham I (1990). Rapid assessment of *S. cerevisiae* mating type by PCR. **Trends Genet** 6(8): 236.