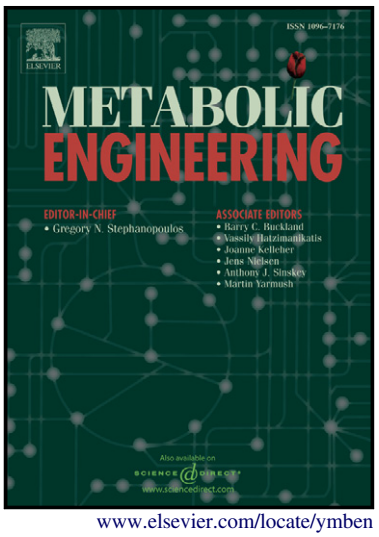



www.elsevier.com/locate/ymben

# Quantitative Trait Analysis of Yeast Biodiversity Yields Novel Gene Tools for Metabolic Engineering

Georg Hubmann, Maria R. Foulquié-Moreno, Elke Nevoigt, Jorge Duitama, Nicolas Meurens, Thiago M. Pais, Lotte Mathé, Sofie Saerens, Huyen Thi Thanh Nguyen, Steve Swinnen, Kevin J. Verstrepen, Luigi Concilio, Jean-Claude de Troostembergh, Johan M. Thevelein

PII: S1096-7176(13)00032-3
DOI: http://dx.doi.org/10.1016/j.ymben.2013.02.006
Reference: YMBEN785

To appear in: *Metabolic Engineering*

Received date: 4 September 2012
Revised date: 25 December 2012
Accepted date: 20 February 2013

Cite this article as: Georg Hubmann, Maria R. Foulquié-Moreno, Elke Nevoigt, Jorge Duitama, Nicolas Meurens, Thiago M. Pais, Lotte Mathé, Sofie Saerens, Huyen Thi Thanh Nguyen, Steve Swinnen, Kevin J. Verstrepen, Luigi Concilio, Jean-Claude de Troostembergh, Johan M. Thevelein, Quantitative Trait Analysis of Yeast Biodiversity Yields Novel Gene Tools for Metabolic Engineering, *Metabolic Engineering,* http://dx.doi.org/10.1016/j.ymben.2013.02.006

# Quantitative trait analysis of yeast biodiversity yields novel gene tools for metabolic engineering

**Georg Hubmann[1,2], Maria R. Foulquié-Moreno[1,2], Elke Nevoigt[1,2,3], Jorge Duitama[4,5], Nicolas Meurens[6], Thiago M. Pais[1,2], Lotte Mathé[1,2], Sofie Saerens[4,5,7], Huyen Thi Thanh Nguyen[1,2,3], Steve Swinnen[1,2,3], Kevin J. Verstrepen[4,5], Luigi Concilio[6], Jean-Claude de Troostembergh[6] and Johan M. Thevelein[1,2]***

[1]Laboratory of Molecular Cell Biology, Institute of Botany and Microbiology, KU Leuven, and [2]Department of Molecular Microbiology, VIB, Kasteelpark Arenberg 31, B-3001 Leuven-Heverlee, Flanders, Belgium. [3]School of Engineering and Science, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany. [4]Laboratory for Genetics and Genomics, Centre of Microbial and Plant Genetics (CMPG), KU Leuven, Kasteelpark Arenberg 22, B-3001 Leuven-Heverlee, Belgium. [5]Laboratory for Systems Biology, VIB, Bio-Incubator, Gaston Geenslaan 1, B-3001 Leuven-Heverlee, Belgium. [6]Cargill R&D Centre Europe, Havenstraat 84, B-1800 Vilvoorde, Belgium

[7]Present address: Wine Innovation Department, Cultures and Enzymes Division, Chr. Hansen A/S, 10-12 Bøge Allé, DK-2970 Hørsholm, Denmark

*Corresponding author (Phone: +32-16-321507, secr. +32-16-321500; Fax.: +32-16-321979; E-mail: johan.thevelein@mmbio.vib-kuleuven.be)

Running title: Genetic analysis of glycerol/ethanol yield

**Abstract**

Engineering of metabolic pathways by genetic modification has been restricted largely to enzyme-encoding structural genes. The product yield of such pathways is a quantitative genetic trait. Out of 52 *Saccharomyces cerevisiae* strains phenotyped in small-scale fermentations, we identified strain CBS6412 as having unusually low glycerol production and higher ethanol yield as compared to an industrial reference strain. We mapped the QTLs underlying this quantitative trait with pooled-segregant whole-genome sequencing using 20 superior segregants selected from a total of 257. Plots of SNP variant frequency against SNP chromosomal position revealed one major and one minor locus. Downscaling of the major locus and reciprocal hemizygosity analysis identified an allele of *SSK1*, $ssk1^{E330N...K356N}$, expressing a truncated and partially mistranslated protein, as causative gene. The diploid CBS6412 parent was homozygous for $ssk1^{E330N...K356N}$. This allele affected growth and volumetric productivity less than the gene deletion. Introduction of the $ssk1^{E330N...K356N}$ allele in the industrial reference strain resulted in stronger reduction of the glycerol/ethanol ratio compared to *SSK1* deletion and also compromised volumetric productivity and osmotolerance less. Our results show that polygenic analysis of yeast biodiversity can provide superior novel gene tools for metabolic engineering.



**Key words**

QTL analysis; yeast; *Saccharomyces cerevisiae*; glycerol yield; ethanol yield; reverse metabolic engineering

**Abbreviations**

GPDH, glycerol-3-phosphate dehydrogenase**;** PCR, polymerase chain reaction**;** RHA, reciprocal hemizygosity analysis

## 1. Introduction

Up to now, the targeted genetic engineering of microorganisms has concentrated largely on the modification of structural genes encoding enzymes in metabolic pathways. This has been done either by up- or downregulation of gene expression or by modification of kinetic characteristics, substrate specificity or regulatory properties of the constituent enzymes (Nevoigt, 2008). However, targeted engineering has shown only limited success when it comes to complex traits determined by multiple genes and largely unknown regulatory networks. Empirical approaches, on the other hand, such as mutagenesis and genome shuffling, have often been used more successfully as a strategy to alter such complex phenotypes (Nevoigt, 2008; Oud *et al.*, 2011). These approaches have mainly been applied to improve stress tolerance, which allows direct selection of improved mutants. Alternatively, improvement of stress tolerance in various species has been achieved using genomic, metagenomic or cDNA libraries, expressed in a sensitive host microorganism (Nicolaou *et al.*, 2010). Rational attempts have been carried out to engineer regulatory factors in order to simultaneously and randomly alter the regulation of many genes at a time. In bacterial systems, transcription of multiple genes can be modified easily through mutagenesis of σ factors, which has been referred to as global cellular transcription machinery engineering (Alper & Stephanopoulos, 2007). This strategy has also been used to improve ethanol tolerance and yield in *Saccharomyces cerevisiae* (Alper *et al.*, 2006). However, eukaryotic systems regulating gene expression are by far more complex than those of bacteria. In comparison to bacterial systems, the changes in the phenotype caused by mutations in transcription factors are still difficult to predict and may also cause unwanted side-effects on other important properties.

Genetic engineering of metabolic pathways in industrial eukaryotic microorganisms, such as yeast, is clearly limited by a lack of knowledge on regulatory factors and their mechanisms of action. This is particularly true under the conditions that occur in industrial applications. Moreover, the structural genes and regulatory factors involved in metabolic pathways contain many mutations in natural and industrial yeast strains, which create large phenotypic diversity, further complicating the understanding of the interplay between the functioning of the structural pathway and its regulatory systems. On the other hand, reverse engineering makes use of this phenotypic diversity by targeting genes identified by genetic analysis of natural and industrial strains with interesting traits (Bailey *et al.*, 1996; Nevoigt, 2008). Most of these traits, however, are complex and only recently methodologies have become available for efficient mapping and identification of the multiple mutant genes responsible for such complex traits (Swinnen *et al.*, 2012b).

The exceptional capacity of the yeast *S. cerevisiae* for anaerobic production of ethanol is the basis of nearly all industrial production of alcoholic beverages and fuel ethanol. Apart from carbon dioxide, glycerol is the most important byproduct in yeast ethanolic fermentation. Under anaerobic conditions, glycerol production is closely connected to the growth rate of the cells. The withdrawal of intermediates from glycolysis for biosynthetic purposes necessitates regeneration of $NAD^+$ to sustain the redox balance and in the absence of oxygen this is accomplished by formation of glycerol (Bakker *et al.*, 2001). A second function for glycerol production in yeast is its use as a compatible osmolyte under conditions of hyperosmotic stress (Blomberg & Adler, 1989; Hohmann, 2002).

Glycerol is synthesized in two steps from dihydroxyacetone phosphate by $NAD^+$ dependent glycerol-3-phosphate dehydrogenase (GPDH) and glycerol-3-phosphate phosphatase, encoded by *GPD1* and *GPD2*, and *GPP1* and *GPP2*, respectively (Albertyn *et al.*, 1994; Ansell *et al.*, 1997). Enhanced expression of *GPD1* is a major factor responsible for stimulation of glycerol production under osmostress (Albertyn *et al.*, 1994; Larsson *et al.*, 1993; Nevoigt & Stahl, 1997). The high osmolarity glycerol (HOG) pathway, responsible for osmostress-induced glycerol production and other cellular adaptations, has been characterized in great detail (Brewster *et al.*, 1993; Hohmann, 2002). Changes in extracellular osmolarity are sensed via two independent transmembrane proteins, Sho1 and Sln1, that both activate the HOG Map kinase pathway. The Sln1 branch plays the most prominent role and acts through a phosphotransfer system, composed of Sln1, Ypd1 and Ssk1. The two pathways converge on the phosphorylation of Pbs2, which activates the Map kinase Hog1. This causes translocation of Hog1 into the nucleus, where it associates with several transcriptional regulators, i.a. Sko1, Msn2, Smp1 and Hot1. These Hog1-regulator complexes induce *GPD1* expression to enhance the formation of glycerol under osmostress (Hohmann, 2002). Retention of glycerol within the cells and its efflux upon relief of osmostress are controlled by the Fps1 plasma membrane channel (Luyten *et al.*, 1995).

Engineering of glycerol production in yeast has attracted considerable attention. Higher glycerol levels are desirable in wine and beer production as well as industrial glycerol production (Cambon *et al.*, 2006; Geertman *et al.*, 2006; Heux *et al.*, 2006; Nevoigt & Stahl, 1996; Remize *et al.*, 1999; Schuller & Casal, 2005). Multiple genetic modifications have been used to raise glycerol production and counteract the

side-effect of higher acetate production (Cambon *et al.*, 2006; Eglinton *et al.*, 2002; Ehsani *et al.*, 2009). Lower glycerol levels are highly desirable in ethanol fuel production because they are usually associated with increased ethanol yields (Basso *et al.*, 2008; Bro *et al.*, 2006; Nissen *et al.*, 2000a; Nissen *et al.*, 2000b). High ethanol yield is a key characteristic of bioethanol production strains, reaching approximately 90-93% of the theoretical maximum of 0.51 g ethanol per g glucose in current industrial processes (Bai *et al.*, 2008). Despite the high ethanol yield, part of the sugar is still used for yeast growth and glycerol production. Glycerol yield can reach up to 2.0 - 3.6 g per 100 g consumed glucose as already reported by Pasteur (Pasteur, 1858). Glycerol yields strongly depend on fermentation conditions (Alfenore *et al.*, 2004; Bideaux *et al.*, 2006; Gardner *et al.*, 1993) and medium composition, especially the type of nitrogen source used (Albers *et al.*, 1996).

A key challenge in industrial ethanol production is lowering glycerol yield without compromising osmostress tolerance and growth rate under anaerobic conditions. Osmotolerance is an important trait for industrial production, storage and utilization of yeast and growth rate is closely correlated with ethanol production rate under anaerobic conditions. Hence, diminution of *GPD1* and/or *GPD2* expression is not an option since it likely compromises osmostress tolerance and growth under anaerobic conditions (Ansell *et al.*, 1997; Bjorkqvist *et al.*, 1997; Nissen *et al.*, 2000a). Even strains with fine-tuned reduction in GPDH activity obtained with promotor engineering still showed a significant drop in osmotolerance and/or growth rate resulting in lower ethanol productivity (Hubmann *et al.*, 2011; Pagliardini *et al.*, 2010). Hence, despite considerable efforts, metabolic engineering of the structural genes for glycerol synthesis, *GPD1* and *GPD2*, has met with little final success

because of negative side-effects on growth, fermentation and osmotolerance (Bjorkqvist *et al.*, 1997; Bro *et al.*, 2006; Cao *et al.*, 2007; Guo *et al.*, 2011; Guo *et al.*, 2009; Guo *et al.*, 2010; Hubmann *et al.*, 2011; Kong *et al.*, 2007a; Kong *et al.*, 2006; Kong *et al.*, 2007b; Nissen *et al.*, 2000a; Nissen *et al.*, 2000b).

Reverse metabolic engineering is an attractive alternative (Bailey *et al.*, 1996), but the identification of the genetic basis of complex traits, such as glycerol yield in fermentation, has remained for many years an important bottleneck. The availability of genome-wide methods for scoring SNPs as genetic markers has facilitated simultaneous mapping of multiple linked loci referred to as quantitative trait loci (QTLs) (Brem *et al.*, 2002; Deutschbauer & Davis, 2005; Steinmetz *et al.*, 2002; Winzeler *et al.*, 1998). Next generation sequencing methods now allow very efficient QTL mapping using whole-genome sequence analysis of pooled segregants displaying the trait of interest (Ehrenreich *et al.*, 2010; Parts *et al.*, 2011; Swinnen *et al.*, 2012a).

We have now applied pooled-segregant whole-genome sequence analysis for identification of genetic elements determining glycerol yield in yeast fermentation. We were able to identify a yeast strain, CBS6412, with unusually low glycerol production and concomitantly higher ethanol production. Using 20-44 segregants, we succeeded in identifying major and minor QTLs and successfully identified *SSK1* as the causative allele in the locus with the strongest linkage. Introduction of the mutant *SSK1* allele in the industrial target strain lowered the glycerol/ethanol ratio more than deletion of *SSK1*, without compromising osmotolerance or ethanol productivity more.

Hence, our results show that QTL analysis of selected yeast strains can provide interesting new gene tools for reverse metabolic engineering.

## 2. Materials and methods

### 2.1 Microbial strains and cultivation conditions

All *S. cerevisiae* strains used are listed in Table 1. Strain CBS6412 was originally designated as sake yeast Kyokai No.7 in the CBS collection, but comparison of the genome and *SSK1* sequence with the published sequence of the strain Kyokai No.7 (*Saccharomyces cerevisiae* strain kyokai no. 7 genome sequencing project, accession no. PRJNA45827) revealed that this designation seems to be erroneous. *E. coli* strain DH5α$^{TM}$ (Invitrogen Corp., Carlsbad) was used for amplification of plasmids. The strain was grown in Luria-Bertani (LB) medium containing 0.5% w/v yeast extract, 1% w/v Bacto tryptone, 1% w/v NaCl, (pH 7.5) at 37˚C. *E. coli* transformation and isolation of plasmid DNA was carried out using standard techniques (Sambrook *et al.*, 1989). Transformants were selected on LB medium containing 100µg/ml ampicillin.

### 2.2 Mating, sporulation and tetrad analysis

Mating, sporulation and dissection of asci were carried out according to standard procedures (Sherman & Hicks, 1991). Mating type of segregants was determined by diagnostic PCR for the *MAT* locus (Huxley *et al.*, 1990).

### 2.3 Fermentation conditions

A selection of 52 *S. cerevisiae* wild type strains was screened in 250 ml oxygen-limited and stirred fermentations containing 1% w/v yeast extract, 2% w/v peptone and 12% w/v glucose. Screening of the selected parent strains and the segregants was performed in 15 ml falcon tubes containing 5 ml minimal medium containing 1.9 g $l^{-1}$ yeast nitrogen base (Difco), 5 g $l^{-1}$ ammonium sulphate, 250 mg $l^{-1}$ leucine, 50 mg $l^{-1}$ uracil, 100 mg $l^{-1}$ histidine, 30 mg $l^{-1}$ lysine, 20 mg $l^{-1}$ methionine and 50 g $l^{-1}$ glucose.

Fermentations were inoculated with an initial OD of 1 and their progress followed by weight loss. Selected segregants were also tested in 100 ml oxygen-limited stirred fermentations. All fermentations were carried out at 30°C.

High gravity fermentations were carried out in fermentation tubes containing 250 ml of YP and 33% w/v glucose. Precultures used as inoculum were first grown on YP 2% w/v glucose for 24 hours and then on YP 10% w/v glucose up to an $OD_{600}$ of 1. The fermentations were inoculated with $5 \cdot 10^7$ cells/ml and kept at 25°C. Stirring was applied for the first 4h (120 rpm). When the weight loss was stable for 2 consecutive days, the fermentation was considered to be finished.

SHF (Separate Hydrolysis and Fermentation) fermentations were carried out with wheat liquefact (24.5% dry mass content) acquired from a local ethanol plant. After adjustment of the pH to 4.5 with sulfuric acid, it was treated with Dextrozyme® (Novozyme, Denmark) for 24h at 60˚C to obtain hydrolysate. The latter was boiled at 100˚C for 20 min and then cooled. Oxygen-limited fermentations were carried out with 100 ml of this medium inoculated with 5 ml of yeast suspension. The fermentations were performed at 30˚C and were continuously stirred at 200 rpm.

Assessment of osmotolerance was performed in fermentations containing minimal medium with or without 0.7M or 1.4M NaCl, or 1M or 2M sorbitol. The fermentations were continuously stirred at 200 rpm.

**2.4 Determination of fermentation parameters**

In all fermentations weight loss was used to follow the progress of the fermentation. Glucose, glycerol and ethanol in the medium were determined by HPLC (Waters® isocratic BreezeTM HPLC, ion exchange column WAT010290). Column temperature was 75˚C, 5 mM $H_2SO_4$ was used as eluent with a flow rate of 1 ml $min^{-1}$ and refractive index detection was used (Waters, 2414 RI detector). Biomass was determined by $OD_{600}$ at the beginning and the end of fermentation and yeast dry mass also at the end. The product yield was calculated from the final product concentration (g. $l^{-1}$) and the difference in glucose concentration at the start and end of fermentation (consumed glucose in g. $l^{-1}$). The product yields in the SSF were based on the final product concentrations and the equivalent initial glucose concentration (the latter was measured in a completely hydrolyzed sample of wheat liquefact).

**2.5 Determination of maximum aerobic growth rate**

The aerobic cultivations were carried out in Erlenmeyer flasks containing 50 ml medium, which was composed of 1% [w/v] yeast extract and 2% [w/v] glucose. The strains were pre-cultured overnight at 30 °C and 200 rpm in small tubes containing 3 ml of the same medium. The pre-culture were then used to inoculate the main culture up to an initial $OD_{600}$ of 0.2. The Erlenmeyer flasks were incubated at 30°C in a rotary shaker at 200 rpm. Samples were taken every second hour in order to determine the $OD_{600}$. Non-linear regression was used to fit a polynomial curve to the experimental progression of the $OD_{600}$ values. The growth rate was calculated by differentiating the continuous curve. Specific growth rates were calculated by dividing the growth rate by the biomass estimated from the continuous curve at the same time. The maximum of the curve was used as the maxium growth rate.

**2.6 Aerobic growth spot assay**

The aerobic growth behaviour was tested on solid medium, which contained 1% [w/v] yeast extract, 2% [w/v] glucose and 1.5% [w/v] agar. The strains were pre-cultured overnight at 30°C and 200 rpm in small tubes containing 3 ml of the same liquid medium. The $OD_{600}$ was determined for all pre-cultures and the amount of cells was harvested to obtain an $OD_{600}$ of 1 in 1 ml distilled water. After re-suspension of the cells in 1 ml water, these samples were subsequently diluted six times with a dilution factor of 1:5. All dilutions were transferred to a microtiter plate, and 5 µl of each dilution was transferred to the solid medium. The plates were incubated for two days at 30°C. Growth was monitored after 24 h and at the end.

**2.7 DNA methods**

Yeast genomic DNA was extracted with Phenol/Chloroform/Isoamyl-alcohol (25:24:1) (Hoffman & Winston, 1987) and further purified with diethyl-ether extraction or ethanol precipitation if required. PCR was performed with high-fidelity polymerases Phusion™ (Finnzymes) or ExTaq™ (TaKaRa) for cloning, amplification of deletion or insertion cassettes, and sequencing purposes. Sequencing was carried out using the dideoxy chain-termination method (Sanger & Coulson, 1975) at the VIB Genetic Service Facility (Antwerp). The sequences were analyzed with geneious (Geneious Basic 5.3.4), SeqMan (Lasergene Coresuite 8) or CLC DNA workbench (CLC bio) software. Lists of primers and plasmids are provided as Supplementary Tables 1 and 2, respectively.

**2.8 Pooled-segregant whole-genome sequence analysis**

After crossing the two parent strains CBS4C and ER7A, the 20 most superior segregants (lowest glycerol production) were assembled in the 'selected pool' while 20 random segregants were used to assemble the 'unselected pool'. The two pools were made by combining equal amounts of cells based on $OD_{600}$. High molecular weight DNA (3 µg, ~ 20kb fragments) was isolated from the pools and parent strains according to Johnston and Aust (1994). The purity of the DNA sample was estimated from UV measurement (260/280 = 1.7-2.0). The DNA samples were provided to GATC Biotech AG (Konstanz, Germany) and BGI (Hong Kong, China) for whole-genome sequence analysis by Illumina technology.

QTL analysis based on the distribution of SNP variant frequency over the length of the chromosomes was carried out as described by Swinnen et al. (2012a). The short read sequences obtained from the parental strains and the pools were mapped against the known S288c reference sequence using the mapping software Bfast (Homer *et al.*, 2009). After pairing, unique alignments for the CBS4C strain were selected and homozygous variants, i.e. SNPs and small indels, were called using SNVQ (Duitama *et al.*, 2012). In addition, regions with coverage below 0.5 or above 1.5 of the average coverage were identified and SNPs of those regions were filtered out. For each polymorphic position the variant calls in the aligned reads for the ER7A strain were then extracted and variants were filtered out for which the coverage of the reference variant was too small (<20x) or too large (>150x) or SNPs of both parents coincided but were different from the reference. Finally, the number of calls to the reference and the alternative variant of each selected polymorphic position was determined from the

set of aligned reads corresponding with the segregant pools. The SNP variant frequencies were calculated by dividing the number of the alternative variant by the total number of aligned reads. A very high or a very low frequency was a sign of a one-sided SNP segregation preferentially coming from one parent, indicating a genetic linkage to the trait of interest. Genetic linkage was statistically confirmed using the methods described earlier (Swinnen *et al.*, 2012a).

**2.9 Detection of SNP markers by allele specific PCR**

Individual SNPs were scored by **allele specific** PCR. The forward and reverse primer contained the nucleotide of ER7A or CBS4C as the 3' terminal nucleotide. The annealing temperature was optimized using DNA extracts of ER7A and CBS4C so as to allow only hybridization with primers containing a complete match. A list of primers is provided in Supplementary Table 1.

**2.10 Reciprocal hemizygosity analysis (RHA)**

For RHA analysis (Steinmetz *et al.*, 2002), two diploid strains were constructed by crossing CBS4C and ER7A wild type or *ssk1*Δ strains, so that the resulting diploids only contained a single *SSK1* allele, either CBS4C derived $ssk1^{E330N...K356N}$ or ER7A derived *SSK1*. Deletion cassettes were constructed essentially as described by Gueldener *et al.* (2002) with the phleomycin resistance marker $ble^{R}$ and *SSK1* gene deletion was confirmed by PCR. The selection marker was removed using the *Cre-loxP* system. The removal of the selection marker was verified by phleomycin sensitivity as well as by PCR. RHA was performed with three independent isolates of all tested diploids.

**2.11 Construction of *SSK1* insertion cassettes**

The repeat region H1 was PCR amplified with the primers A-6101 and A-6103 using genomic DNA of CBS4C and ER7A as template. The resulting PCR fragment was digested with *KpnI* and *SalI* and purified from an agarose gel. *SSK1* was PCR amplified from genomic DNA of CBS4C and ER7A and the primers A-6100 and A-6102. The obtained product of around 2800bp was digested with *SalI* and *XmaI*. The cloning vector pBluescriptII SK(+) (Fermentas) was digested with *KpnI* and *XmaI* and ligated with the repeat region H1 and the *SSK1* allele of the respective strain. The construct was verified using Sanger sequencing. The two selectable and counter-selectable systems, *AMD1* and *NAT1-GIN11*, were used to introduce the insertion cassette. During counter-selection, the marker genes spontaneously looped out via the H1 repeat region, leaving no scars of non-*S. cerevisiae* DNA in the genome. The *AMD1* marker of *Zygosaccharomyces rouxii* was cut out of the plasmid pF6a-AMD1-MX6 using SacI and BglII (Shepherd & Piper, 2010). The fragment was gel purified and ligated with pUG66, which was also digested with the same two enzymes. The resulting plasmid pUG-AMD was used for PCR amplification of the *AMD1* marker using the primers A-5166 and A-6770. The PCR product as well as the H1-SSK1 plasmids were digested with *SalI* and ligated, resulting in plasmids pBluescriptII_*AMD1_ ssk1*$^{E330N...K356N}$ and pBluescriptII_*AMD1_SSK1*. The selection marker *NAT1* was amplified from pAG25 using primers A-7116 and A-7117. The *GIN11* counter-selection marker (Akada *et al.*, 2002) was amplified from pG119 using primers A-7118 and A-7119. Both fragments were sequentially digested with *DraIII* and *SalI* and ligated with the H1-SSK1 plasmid, which was previously digested with *SalI* resulting in the plasmids pBluescriptII_*NAT1_GIN11_ssk1*$^{E330N...K356N}$ and pBluescriptII_*NAT1_GIN11_SSK1*.

The insertion cassette *H1-loxP-AMD1-loxP-SSK1* and *H1-loxP-AMD1-loxP-ssk1*$^{E330N...K356N}$ were amplified from pBluescriptII_*AMD1_SSK1* and pBluescriptII_*AMD1_ssk1*$^{E330N...K356N}$ using the outside flanking primers matching the M13 primer binding sites of the plasmid pBluescriptII SK(+). The PCR product was purified and used for transformation. Cassettes with *H1-NAT1-GIN11-SSK* or *H1-NAT1-GIN11- ssk1*$^{E330N...K356N}$ were digested with *BspHI* and digestion products were used for transformation. Yeast was transformed with the LiAc/PEG method (Gietz *et al.*, 1992).

**2.12 Reciprocal *SSK1* allele replacement in CBS4C and ER7A**

Site-directed modification of the CBS4C and ER7A *SSK1* locus was carried out using a two step-method. In the first step, the *SSK1* insertion cassettes (see above) were transferred to the *SSK1* deletion strains, CBS4C *ssk1*Δ and ER7A *ssk1*Δ. After transformation, positive clones were selected on YD agar plates containing 200µg/ml ClonNat. The presence of the insertion cassette was verified by PCR using the primers A-5168 and A-7301. In the second step, the marker genes were removed by selection of spontaneous loop-outs on galactose-containing medium after induction of the counter-selectable marker *GIN11* (Akada *et al.*, 2002; Akada *et al.*, 1999; Olesen *et al.*, 2000). Positive looped-out clones were identified by ClonNat sensitivity and verified by PCR using the forward primer A-5168 and the *SSK1* allele specific reverse primers A-5126 and A-5127. The inserted *SSK1* alleles were verified by Sanger sequencing.

**2.13 *SSK1* allele replacement in the industrial strain Ethanol Red**

The *ssk1*$^{E330N...K356N}$ allele of CBS4C was inserted twice in the Ethanol Red derivative, HG5 (see Table 1), which had both *SSK1* alleles deleted. The latter strain was constructed by introducing a disruption cassette flanked with loxP sites using homologous recombination (Gueldener *et al.*, 2002; Kotaka *et al.*, 2009). The disruption cassette was constructed with homologous sequences (H1/H2) corresponding to the 5' and 3' end of the *SSK1* ORF surrounding the phleomycin resistance-gene *ble*$^{R}$ used as selectable marker. The selectable marker, *ble*$^{R}$, was removed by Cre recombinase. A second disruption cassette was constructed with the recombination sites H1$^{*}$ and H2$^{*}$, which were located inside the first homologous integration sites, H1 and H2, enabling specific recombination into the 2$^{nd}$ *SSK1* allele of the diploid strain. Gene disruption was verified by PCR. The *ble*$^{R}$ marker gene was again removed using the Cre/loxP system. The double deletion was confirmed by PCR using primers located outside the integration site.

The two *ssk1*$^{E330N...K356N}$ insertion cassettes, *H1-loxP-AMD1-loxP-ssk1*$^{E330N...K356N}$ and *H1-NAT1-GIN11-ssk1*$^{E330N...K356N}$, were successively transformed. After the 1$^{st}$ transformation of the *H1-loxP-AMD1-loxP-ssk1*$^{E330N...K356N}$ cassette, transformants were selected based on hydrolysis by Amd1 of acetamide used as sole nitrogen source. The correct integration of the insertion cassette was verified by PCR using the primers A-5168 and A-5894 inside the *AMD* gene. In the 2$^{nd}$ transformation, the *H1-NAT1-GIN11-ssk1*$^{E330N...K356N}$ was transferred. Positive transformants were selected on acetamide medium, containing 200µg/ml ClonNat. Correct integration in the 2$^{nd}$ chromosome was verified by PCR using the primers A-5168 and A-7301 to test the

presence of the insertion cassette as well as the primers A-5168 and A-5169 to verify the disappearance of the two ORF deletions of the Ethanol Red *ssk1*Δ/Δ. Counter-selection was simultaneously applied for both marker systems using medium with 100mM fluoroacetamide and 0.04% galactose (induction of *GIN11*).

**2.14 Data deposition**

Sequencing data have been deposited at the SRA database (NCBI), http://www.ncbi.nlm.nih.gov/sra, with the account number SRA054394.

## 3. Results

### 3.1 Selection of parent strains for genetic mapping of low glycerol yield

We have evaluated 52 diploid *S. cerevisiae* strains from diverse origins for the ratio between the amount of glycerol and ethanol produced in small-scale (250 ml) fermentations with complex medium containing 12% glucose. A continuous and normal distribution of the trait was observed (Figure 1). The CBS6412 strain showed the lowest glycerol yield (0.043 g. $g^{-1}$) of all strains tested, which was about 63% of that of the reference industrial strain Ethanol Red (0.068 g. $g^{-1}$) (Figure 2A), an industrial strain commonly used for bioethanol production with corn and wheat starch hydrolysate. As it had both a low glycerol/ethanol ratio and a low glycerol yield, CBS6412 was chosen as the superior strain and Ethanol Red was used as the inferior strain.

In order to obtain haploid strains for genetic mapping analysis, the two diploid strains were sporulated and segregants were tested in small-scale fermentations. Glycerol yields of the segregants were normally distributed around those of the diploid parents (Figure 2B), indicating a highly heritable phenotype. The CBS6412 segregant, CBS4C, had an even lower glycerol yield than its parental diploid (Figure 2A), indicating acquirement of one or more beneficial, recessive alleles present in heterozygous form in the diploid strain. CBS4C was selected as the superior parent strain for the genetic mapping. The Ethanol Red segregant ER7A had a glycerol yield closest to its parental diploid and served as inferior parent strain.

### 3.2 Construction of the CBS4C/ER7A hybrid and selection of superior segregants with low glycerol yield

The CBS4C and ER7A haploid strains were crossed with each other and 257 segregants were isolated and first characterized for glycerol and ethanol yield in 5 ml fermentations with 5% glucose in minimal medium. Figure 2C shows a histogram of the glycerol yield in the segregant population in comparison with that of the CBS4C and ER7A haploid parents. The glycerol yield showed a normal distribution and most segregants had a glycerol yield close to the average (0.063 g. $g^{-1}$, ±142% of the CBS4C glycerol yield). We re-tested the 48 segregants with a glycerol yield below 120% of the CBS4C parent in 100 ml small-scale fermentations; 44 segregants showed the same low glycerol yield also under these conditions. Among these, the 20 segregants showing the lowest glycerol yield (≤ 0.054 g $g^{-1}$) were selected for QTL mapping with pooled-segregant whole-genome sequence analysis. The 24 remaining segregants were used for subsequent validation of the results as described below. A second pool with 20 randomly selected segregants was also subjected to pooled-segregant whole-genome sequence analysis and used as control.

### 3.3 QTL mapping using pooled-segregant whole-genome sequence analysis

The genomic DNA of the selected and random pools, as well as the two parent strains, was extracted and submitted to custom sequence analysis using Illumina HiSeq 2000 technology (GATC Biotech AG, Konstanz, Germany; BGI, Hong Kong, China). The sequence reads of the CBS4C and ER7A parent strains were aligned with the S288c standard sequence, which allowed to identify 21,818 SNPs between CBS4C and ER7A. The SNPs were filtered as described previously (Duitama *et al.*, 2012). The

variant frequency of the quality-selected SNPs in the DNA of the two pools was then plotted against the SNP position on the chromosome. The scattered raw data were smoothened by fitting smoothing splines in the generalized linear mixed model framework as previously described (Swinnen *et al.*, 2012a). The results are shown in Figure 3. A prominent QTL with strong linkage was present on chromosome XII (between 135,000 and 200,000 bp) and is shown in more detail in Figure 4B. Individual SNPs from that region, as well as from the QTL with lower linkage on chromosome II (Figure 4A), were scored by allele specific PCR detection in the 20 individual segregants of the selected pool (Figure 4A, B). The precise SNP variant frequency determined in this way was used to verify the linkage of the two regions on chromosome II and XII, respectively. This revealed a very strong linkage with low glycerol yield for the QTL on chromosome XII with the minimal P-value being $1.45 \cdot 10^{-4}$, while the P-values for the QTL on chromosome II only dropped just below the 0.05 threshold for significance (0.009). The same SNPs were also scored by allele specific PCR in the 24 remaining segregants with a glycerol yield below 120% of the CBS4C parent. Calculation of the P-values for the whole group of 44 segregants no longer revealed significant linkage for the QTL on chromosome II. On the other hand, the P-values for the QTL on chromosome XII dropped to $9 \cdot 10^{-11}$, strongly increasing significance of the linkage. Hence, we concentrated the further analysis on the QTL of chromosome XII.

### 3.4 Identification of *SSK1* as a causative gene in the QTL on chromosome XII

The 20,000 bp region with the strongest linkage in the QTL on chromosome XII contained 13 genes, of which four genes contained non-synonymous mutations in the ORF (Figure 4B). One of those four genes, *SSK1,* was located in the centre of the

QTL, which had a slightly stronger linkage. Ssk1 has a known function in the HOG pathway. Sequence comparison of the *SSK1* alleles of the parental strains CBS4C and ER7A with the allele of the reference strain S288c revealed ten polymorphisms between the sequence of the *SSK1* ORF in CBS4C and ER7A. A single base pair deletion at position 162,907 bp of Chr. XII was the most prominent mutation in the CBS4C *SSK1* ORF, since it caused a reading frame shift and a new stop codon at position 357 in the protein. This resulted in a new primary amino acid sequence from position 330 until 356, while the wild type Ssk1 protein had a total length of 712 amino acids. Hence, we named the new allele $ssk1^{E330N...K356N}$. The dramatic change in amino acid sequence and the truncation would normally be expected to result in a completely inactive protein and therefore in a phenotype similar to that of the *ssk1*Δ strain. However, this was not the case. The $ssk1^{E330N...K356N}$ allele caused a different phenotype compared to deletion of *SSK1* (see below).

Next, we evaluated *SSK1* as possible causative gene using reciprocal hemizygosity analysis (RHA) (Steinmetz *et al.*, 2002). For that purpose, two CBS4C/ER7A hybrid and hemizygous diploid strains were constructed differing only in a single *SSK1* allele, the other allele being deleted. The diploid strain with the single $ssk1^{E330N...K356N}$ allele derived from CBS4C showed a significantly reduced glycerol yield and a significantly higher ethanol yield than the diploid strain with the *SSK1* allele from the ER7A strain (Figure 5). This showed that $ssk1^{E330N...K356N}$ was a causative gene in the QTL on chromosome XII.

To evaluate whether the $ssk1^{E330N...K356N}$ allele of CBS4C behaved as a recessive allele and whether it caused the same phenotype as deletion of *SSK1*, we also constructed a

CBS4C/ER7A hybrid diploid strain with both *SSK1* alleles deleted and compared its phenotype with that of CBS4C/ER7A with its original *SSK1* alleles. The glycerol and ethanol yields of these strains were similar to that of the hemizygous diploid strain with the *SSK1* allele from ER7A or the $ssk1^{E330N...K356N}$ allele from CBS4C, respectively (Figure 6A). This indicates that the $ssk1^{E330N...K356N}$ allele from CBS4C is a recessive allele and that $ssk1^{E330N...K356N}$ behaves as a loss of function allele, at least in the hybrid background and the fermentation conditions used (100 ml anaerobic fermentations in minimal medium containing 5% glucose). When the glycerol yield (0.043 g. $g^{-1}$) of the CBS4C parent strain was normalized to 100%, the glycerol yield of ER7A (147%) and that of the diploids ER7A/ CBS4C (145%) and ER7A/ CBS4C *ssk1*Δ (148%) was very similar (Figure 6A). In contrast, the strains ER7A *ssk1*Δ/ CBS4C $ssk1^{E330N...K356N}$ and ER7A *ssk1*Δ/ CBS4C *ssk1*Δ had a glycerol yield of 119% and 122%, respectively, (Figure 6A) suggesting that $ssk1^{E330N...K356N}$ was responsible for the majority of the reduction in glycerol yield in CBS4C compared to ER7A. This agrees with the result of the pooled-segregant whole-genome mapping, which revealed the *SSK1* locus as the only QTL with significant linkage.

To confirm the importance of *SSK1* in an alternative way, we reciprocally exchanged the *SSK1* alleles of CBS4C and ER7A by homologous recombination. The $ssk1^{E330N...K356N}$ allele was introduced as new allele in ER7A *ssk1*Δ. In the same manner, *SSK1* of ER7A was introduced in CBS4C *ssk1*Δ. The ER7A $ssk1^{E330N...K356N}$ strain showed reduced glycerol yield and enhanced ethanol yield comparable to the ER7A *ssk1*Δ strain (Figure 6B). Deletion of $ssk1^{E330N...K356N}$ in CBS4C resulted in a further reduction in glycerol yield when compared to the original CBS4C. Concomitantly, ethanol yield increased in the CBS4C *ssk1*Δ strain when compared to

the original CBS4C strain, which may be due to a growth deficiency caused by *ssk1*Δ in that strain background. This growth defect was suggested also by the maximal volumetric ethanol production rate, which was strongly reduced in the CBS4C *ssk1*Δ strain compared to the original CBS4C strain with the $ssk1^{E330N...K356N}$ allele (Figure 6B). In the ER7A background there was no difference. This again shows that at least in the CBS4C background the effect of *ssk1*Δ is different from that of $ssk1^{E330N...K356N}$.

As shown by the aerobic growth assay in Figure 6C, we observed that the CBS4C *ssk1*Δ strain grew more poorly than the original CBS4C strain, which contains the $ssk1^{E330N...K356N}$ allele. A clear difference in growth of these two strains was also observed in aerobic liquid cultures, where the two strains, CBS4C and CBS4C *ssk1*Δ, grew at a maximum rate of 0.47 $h^{-1}$ and 0.35 $h^{-1}$, respectively (Figure 6C). This difference in growth was again not observed in the ER7A strain background. The introduction of wild type *SSK1* in CBS4C *ssk1*Δ enhanced its glycerol yield and reduced its ethanol yield, and it increased the maximal volumetric ethanol production rate (Figure 6B). These results confirmed *SSK1* as a causative allele for reduced glycerol and enhanced ethanol production in CBS4C. They also show that the $ssk1^{E330N...K356N}$ allele causes a different effect compared to the deletion of *SSK1* and that it also differs in causing no apparent growth defect in the CBS4C background as opposed to the deletion of *SSK1*.

Given the recessive character of the $ssk1^{E330N...K356N}$ allele, we tested its presence in the original diploid strain CBS6412 and found it to be present in two copies (data not shown). This suggests that the unusual allele may provide a selective advantage in specific environmental niches.

### 3.5 Reduction of the glycerol/ethanol ratio in an industrial bioethanol strain using *ssk1*$^{E330N...K356N}$ as a novel gene tool

To test the functionality of *ssk1*$^{E330N...K356N}$ as a novel gene tool for reduction of glycerol yield under industrially relevant conditions, both *SSK1* alleles of the industrial bioethanol production strain, Ethanol Red, were replaced by the *ssk1*$^{E330N...K356N}$ variant using homologous recombination. In addition, an Ethanol Red *ssk1*Δ/*ssk1*Δ strain and an Ethanol Red *ssk1*$^{E330N...K356N}$/*ssk1*Δ strain were constructed. These strains were tested in fermentations with minimal medium (5% w/v glucose), high gravity medium (YP with 33% w/v glucose) and wheat hydrolysate (SHF: Separate Hydrolysis and Fermentation). The results are shown in Figure 7A. The double deletion of *SSK1* reduced the glycerol yield. Interestingly, further reduction of glycerol yield was observed after introduction of one copy of *ssk1*$^{E330N...K356N}$, while introduction of the second copy of *ssk1*$^{E330N...K356}$ lowered glycerol yield even more. Ethanol yields clearly increased in all Ethanol Red mutants compared to the wild type strain in the minimal medium. The reduction of glycerol yield under high gravity or SHF conditions was generally less pronounced compared to minimal medium. Thus, the concomitant increase in ethanol yield in the Ethanol Red mutants was less obvious. Nevertheless, particularly the result obtained in minimal medium indicated that in the Ethanol Red diploid background the *ssk1*$^{E330N...K356N}$ allele did not simply behave as a loss-of-function allele but had a stronger reducing effect on the glycerol/ethanol ratio than deletion of the *SSK1* gene. These results confirm the usefulness of the *ssk1*$^{E330N...K356N}$ allele as a novel gene tool for lowering glycerol production in industrial yeast strains. Identification of other mutant alleles in the CBS4C strain and introduction of these alleles in the Ethanol Red strain with two

*ssk1*$^{E330N...K356N}$ alleles, may allow to reduce glycerol yield even more, especially under the industrially-relevant conditions.

### 3.6 The novel gene tool *ssk1*$^{E330N...K356N}$ retains its positive effect under high osmolarity conditions

Several previous studies successfully reduced glycerol yield in *S. cerevisiae* with a concomitant increase in ethanol yield. However, many of the resulting strains showed a significantly reduced maximal volumetric ethanol production rate and increased sensitivity against osmotic stress (Bjorkqvist *et al.*, 1997; Guadalupe Medina *et al.*, 2010; Hubmann *et al.*, 2011; Nissen *et al.*, 2000a). In order to address this issue, we determined both the glycerol/ethanol ratio and the maximal volumetric ethanol production rate in the Ethanol Red strains containing one or two *ssk1*$^{E330N...K356N}$ alleles under conditions of high osmolarity. In general, the cells produced higher levels of glycerol under hyperosmotic stress, i.e. in the presence of 1.4 M and 2 M sorbitol or 0.7 M and 1 M NaCl (Figure 7B). In spite of this, a similar improvement in the glycerol/ethanol ratio was observed in the Ethanol Red strains containing one or two *ssk1*$^{E330N...K356N}$ alleles. The maximal volumetric ethanol production rate dropped with increasing osmolarity but this drop was not correlated with the presence or the number of *ssk1*$^{E330N...K356N}$ alleles. Hence, the *ssk1*$^{E330N...K356N}$ allele does not appear to cause an increase in osmosensitivity and retains its positive effect under conditions of high osmolarity. Close examination of the effect of *ssk1*$^{E330N...K356N}$ on glycerol production in the Ethanol Red background also allows to make a quantitative assessment of the contribution of this allele to the phenotype. The initial glycerol yield was 167% of the CBS4C yield while the double insertion of *ssk1*$^{E330N...K356N}$

caused a drop to 128% of the CBS4C yield. Hence, the *ssk1* mutation appears to determine 50-60% of the trait.

**4. Discussion**

Up to now, QTL analysis in *S. cerevisiae* has been performed mainly with selectable traits easily scorable in high throughput assays, so that many hundreds, thousands or, with appropriate selection, even millions (Parts *et al.*, 2011), of segregants can be used. Stress tolerance for instance is a trait that can be screened easily with large numbers of strains simultaneously. For many industrially-important traits, such as fermentation kinetics and product yields, on the other hand, segregants have to be phenotyped individually, for instance in small-scale fermentations. This requires much more work and is usually limited to a few hundred strains. Glycerol yield is a non-selectable metabolic trait and up to now no such quantitative traits have been submitted to polygenic analysis in yeast. In this work we have shown that genetic mapping of QTLs underlying this metabolic trait can be performed successfully using pooled-segregant whole-genome sequence analysis with only 20 segregants out of a total of 257 phenotyped. This shows that genetic analysis of polygenic metabolic traits can now be used as a novel tool for identifying interesting alleles for reverse metabolic engineering. This approach should in principle be useful for reverse metabolic engineering of any phenotype and any target industrial yeast strain of which a mating-competent haploid segregant can be obtained.

The conspicuous deficiency in our understanding of the interplay between metabolic pathways and cellular regulation has made successful metabolic engineering and synthetic biology a more daunting task than originally anticipated (Yadav *et al.*, 2012). An important issue in this respect is also the inability to mimic industrial-scale conditions precisely in the laboratory, and therefore the inability to test for all possible side-effects which may occur at industrial scale. This creates high risk for

taking multi-engineered strains into industrial scale. As shown in the present paper, appropriate screening of yeast biodiversity can identify strains with superior performance for a specific metabolic trait, like reduced glycerol production. We have shown that the responsible mutant alleles can be used successfully as new gene tools to reverse engineer an industrial yeast strain, minimizing the risk of side-effects on other industrially-important properties.

Application of pooled-segregant whole-genome sequence analysis resulted in the identification of one major QTL located on chromosome XII, and a minor QTL located on chromosome II. Besides this major QTL, two to three other QTLs were expected in CBS4C, which may have been revealed by testing more segregants or defining more stringent conditions for selection. For instance, the weakly linked region on chromosome II was initially also confirmed by scoring the SNPs in the 20 individual segregants, the inclusion of the 24 segregants with a somewhat weaker superior phenotype abolished the significance. This seems to indicate that the causative gene in this QTL is additive to the major mutation in *SSK1* in further lowering glycerol production.

We identified a mutant *SSK1* allele as a causative gene in the QTL on chromosome XII. Ssk1 is part of a “two component” signal transduction system, which signals via a multistep phosphorelay mechanism from the plasma membrane osmo-sensor Sln1 to the redundant MAP kinase kinase kinases (MAPKKK) Ssk2 and Ssk22 (Hohmann, 2002). Upon osmostress, Ssk1 is rapidly dephosphorylated and the resulting protein binds with and activates Ssk2 and Ssk22. Upon completion of osmoadaptation, dephosphorylated Ssk1 is degraded by the Ubc7-dependent ubiquitin-proteasome

system, causing downregulation of the HOG pathway (Sato *et al.*, 2003). The *ssk1*$^{E330N...K356N}$ protein has lost part of its carboxyterminal end including the response regulatory domain (aa. 505 – 647). However, at high osmolarity Ssk1 is unphosphorylated, which enables the protein to interact with the non-catalytic inhibitory domain of Ssk2 and Ssk22 MAPKKK (Maeda *et al.*, 1995). Hence, one would normally expect that absence of the functional regulatory domain causes overactivation of the HOG pathway. In fact, hyperactivation of Hog, by deletion of negative regulators like *PTC1* or *PTP2* (Warmka *et al.*, 2001; Wurgler-Murphy *et al.*, 1997), is lethal in *S. cerevisiae*. Therefore, the *ssk1*$^{E330N...K356N}$ is probably not just a simple loss-of-function allele. This is underscored by the observation that it does not cause the same effect as deletion of *SSK1* for glycerol/ethanol yield, ethanol productivity and growth in the CBS4C background.

A major goal of our work was to identify specific alleles, which would allow to reduce glycerol yield without or with less-pronounced side-effects on osmotolerance or volumetric ethanol productivity. The *ssk1*$^{E330N...K356N}$ allele apparently fulfills these criteria. Introduction of *ssk1*$^{E330N...K356N}$ in the target strain Ethanol Red caused an even stronger reduction of glycerol production than the deletion of *SSK1*, without causing a stronger effect on volumetric productivity. Under anaerobic conditions glycerol production is required for reoxidation of the NADH that is produced in glycolysis for the synthesis of intermediates that are withdrawn for biosynthetic purposes. Under osmostress conditions, glycerol is produced as a compatible osmolyte to prevent cellular dehydration. A possible explanation is that under both conditions, the Ethanol Red strain produces an excess of glycerol and that therefore a

partial reduction can be achieved without seriously compromising ethanol productivity or osmotolerance.

The original diploid parent strain CBS6412 had two copies of the *ssk1*$^{E330N...K356N}$ allele, supporting that this mutation was also causative for the unusually low glycerol production in that strain. This suggests that the mutant protein may provide a specific advantage in certain environmental niches, for instance by enhancing ethanol production without compromising osmostress tolerance. In this genetic background, the *ssk1*$^{E330N...K356N}$ allele also did not compromise growth as opposed to the deletion of *SSK1*. The precise origin of strain CBS6412, however, is not known. The industrial strain Ethanol Red, with two copies of the *ssk1*$^{E330N...K356N}$ allele, showed a reduction in the glycerol/ethanol ratio, but not to the same extent as in CBS6412. Identification of causative mutations in the other QTLs may allow further reduction of the glycerol/ethanol ratio. The successful engineering of the industrial target strain makes us conclude that the huge diversity of natural and industrial *S. cerevisiae* strains available in culture collections may constitute a rich source of novel gene tools for reverse metabolic engineering. In contrast to classical direct metabolic engineering, these gene tools are not limited by our knowledge of the underlying molecular mechanisms.

**Acknowledgments**

We wish to thank Yudi Yang, Mekonnen Demeke, Jean Paul Meijnen, Ben Souffriau and all other members of the Genetic Analysis Group for stimulating discussions, Catherina Coun, Martine De Jonge, Eef Allegaert, Evy Vanderheyden and Paul Vandecruys for technical support and Nico Vangoethem for preparation of the figures and informatics support. We thank Prof. P. Piper (Sheffield) for providing the pFA6a-AMD1-MX6 plasmid. This work has been supported by a predoctoral fellowship to SSw, a post-doctoral fellowship to SSa and a research project (O&O 070626) to J-CdT from the Agency for Innovation by Science and Technology (IWT-Flanders) and SBO grant (IWT 50148) from IWT-Flanders, the EC 7th Framework program (NEMO project) and IOF-Knowledge platform (IKP/10/002 ZKC 1836) to JMT and SBO (IWT 90043) and BOF-Program financing (project NATAR) to KJV and JMT.

**References**

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**,** 393-402.

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**,** 43-8.

Albers, E., Larsson, C., Liden, G., Niklasson, C., Gustafsson, L., 1996. Influence of the nitrogen source on *Saccharomyces cerevisiae* anaerobic growth and product formation. Appl Environ Microbiol. 62**,** 3187-95.

Albertyn, J., Hohmann, S., Thevelein, J. M., Prior, B. A., 1994. GPD1, which encodes glycerol-3-phosphate dehydrogenase, is essential for growth under osmotic stress in *Saccharomyces cerevisiae*, and its expression is regulated by the high-osmolarity glycerol response pathway. Mol Cell Biol. 14**,** 4135-44.

Alfenore, S., Cameleyre, X., Benbadis, L., Bideaux, C., Uribelarrea, J. L., Goma, G., Molina-Jouve, C., Guillouet, S. E., 2004. Aeration strategy: a need for very high ethanol performance in *Saccharomyces cerevisiae* fed-batch process. Appl Microbiol Biotechnol. 63**,** 537-42.

Alper, H., Moxley, J., Nevoigt, E., Fink, G. R., Stephanopoulos, G., 2006. Engineering yeast transcription machinery for improved ethanol tolerance and production. Science. 314**,** 1565-8.

Alper, H., Stephanopoulos, G., 2007. Global transcription machinery engineering: a new approach for improving cellular phenotype. Metab Eng. 9**,** 258-67.

Ansell, R., Granath, K., Hohmann, S., Thevelein, J. M., Adler, L., 1997. The two isoenzymes for yeast NAD+-dependent glycerol 3-phosphate dehydrogenase encoded by *GPD1* and *GPD2* have distinct roles in osmoadaptation and redox regulation. Embo J. 16**,** 2179-87.

Bai, F. W., Anderson, W. A., Moo-Young, M., 2008. Ethanol fermentation technologies from sugar and starch feedstocks. Biotechnol Adv. 26**,** 89-105.

Bailey, J. E., Sburlati, A., Hatzimanikatis, V., Lee, K., Renner, W. A., Tsai, P. S., 1996. Inverse metabolic engineering: A strategy for directed genetic engineering of useful phenotypes. Biotechnol Bioeng. 52**,** 109-21.

Bakker, B. M., Overkamp, K. M., van Maris, A. J., Kotter, P., Luttik, M. A., van Dijken, J. P., Pronk, J. T., 2001. Stoichiometry and compartmentation of NADH metabolism in *Saccharomyces cerevisiae*. FEMS Microbiol Rev. 25**,** 15-37.

Basso, L. C., de Amorim, H. V., de Oliveira, A. J., Lopes, M. L., 2008. Yeast selection for fuel ethanol production in Brazil. FEMS Yeast Res. 8**,** 1155-63.

Bideaux, C., Alfenore, S., Cameleyre, X., Molina-Jouve, C., Uribelarrea, J. L., Guillouet, S. E., 2006. Minimization of glycerol production during the high-performance fed-batch ethanolic fermentation process in *Saccharomyces cerevisiae*, using a metabolic model as a prediction tool. Appl Environ Microbiol. 72**,** 2134-40.

Bjorkqvist, S., Ansell, R., Adler, L., Liden, G., 1997. Physiological response to anaerobicity of glycerol-3-phosphate dehydrogenase mutants of *Saccharomyces cerevisiae*. Appl Environ Microbiol. 63**,** 128-32.

Blomberg, A., Adler, L., 1989. Roles of glycerol and glycerol-3-phosphate dehydrogenase (NAD+) in acquired osmotolerance of *Saccharomyces cerevisiae*. J Bacteriol. 171**,** 1087-92.

Brem, R. B., Yvert, G., Clinton, R., Kruglyak, L., 2002. Genetic dissection of transcriptional regulation in budding yeast. Science. 296**,** 752-5.

Brewster, J., de Valoir, T., Dwyer, N., Winter, E., Gustin, M., 1993. An osmosensing signal transduction pathway in yeast. Science. 259**,** 1760-1763.

Bro, C., Regenberg, B., Forster, J., Nielsen, J., 2006. In silico aided metabolic engineering of *Saccharomyces cerevisiae* for improved bioethanol production. Metab Eng. 8**,** 102-11.

Cambon, B., Monteil, V., Remize, F., Camarasa, C., Dequin, S., 2006. Effects of *GPD1* overexpression in *Saccharomyces cerevisiae* commercial wine yeast strains lacking *ALD6* genes. Appl Environ Microbiol. 72**,** 4688-94.

Cao, L., Zhang, A., Kong, Q., Xu, X., Josine, T. L., Chen, X., 2007. Overexpression of *GLT1* in *fps1DeltagpdDelta* mutant for optimum ethanol formation by *Saccharomyces cerevisiae*. Biomol Eng. 24**,** 638-42.

Deutschbauer, A. M., Davis, R. W., 2005. Quantitative trait loci mapped to single-nucleotide resolution in yeast. Nat Genet. 37**,** 1333-40.

Duitama, J., Srivastava, P. K., Mandoiu, I. I., 2012. Towards accurate detection and genotyping of expressed variants from whole transcriptome sequencing data. BMC Genomics. 13(Suppl 2)**,** S6.

Eglinton, J. M., Heinrich, A. J., Pollnitz, A. P., Langridge, P., Henschke, P. A., de Barros Lopes, M., 2002. Decreasing acetic acid accumulation by a glycerol overproducing strain of *Saccharomyces cerevisiae* by deleting the ALD6 aldehyde dehydrogenase gene. Yeast. 19**,** 295-301.

Ehrenreich, I. M., Torabi, N., Jia, Y., Kent, J., Martis, S., Shapiro, J. A., Gresham, D., Caudy, A. A., Kruglyak, L., 2010. Dissection of genetically complex traits with extremely large pools of yeast segregants. Nature. 464**,** 1039-42.

Ehsani, M., Fernandez, M. R., Biosca, J. A., Julien, A., Dequin, S., 2009. Engineering of 2,3-butanediol dehydrogenase to reduce acetoin formation by glycerol-overproducing, low-alcohol *Saccharomyces cerevisiae*. Appl Environ Microbiol. 75**,** 3196-205.

Gardner, N., Rodrigue, N., Champagne, C. P., 1993. Combined effects of sulfites, temperature, and agitation time on production of glycerol in grape juice by *Saccharomyces cerevisiae*. Appl Environ Microbiol. 59**,** 2022-8.

Geertman, J. M., van Maris, A. J., van Dijken, J. P., Pronk, J. T., 2006. Physiological and genetic engineering of cytosolic redox metabolism in *Saccharomyces cerevisiae* for improved glycerol production. Metab Eng. 8**,** 532-42.

Gietz, D., St Jean, A., Woods, R. A., Schiestl, R. H., 1992. Improved method for high efficiency transformation of intact yeast cells. Nucleic Acids Res. 20**,** 1425.

Guadalupe Medina, V., Almering, M. J., van Maris, A. J., Pronk, J. T., 2010. Elimination of glycerol production in anaerobic cultures of a *Saccharomyces cerevisiae* strain engineered to use acetic acid as an electron acceptor. Appl Environ Microbiol. 76**,** 190-5.

Gueldener, U., Heinisch, J., Koehler, G. J., Voss, D., Hegemann, J. H., 2002. A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. Nucleic Acids Res. 30**,** e23.

Guo, Z. P., Zhang, L., Ding, Z. Y., Shi, G. Y., 2011. Minimization of glycerol synthesis in industrial ethanol yeast without influencing its fermentation performance. Metab Eng. 13**,** 49-59.

Guo, Z. P., Zhang, L., Ding, Z. Y., Wang, Z. X., Shi, G. Y., 2009. Interruption of glycerol pathway in industrial alcoholic yeasts to improve the ethanol production. Appl Microbiol Biotechnol. 82**,** 287-92.

Guo, Z. P., Zhang, L., Ding, Z. Y., Wang, Z. X., Shi, G. Y., 2010. Improving ethanol productivity by modification of glycolytic redox factor generation in glycerol-3-phosphate dehydrogenase mutants of an industrial ethanol yeast. J Ind Microbiol Biotechnol.

Heux, S., Sablayrolles, J. M., Cachon, R., Dequin, S., 2006. Engineering a *Saccharomyces cerevisiae* wine yeast that exhibits reduced ethanol production during fermentation under controlled microoxygenation conditions. Appl Environ Microbiol. 72**,** 5822-8.

Hoffman, C. S., Winston, F., 1987. A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. Gene. 57**,** 267-72.

Hohmann, S., 2002. Osmotic stress signaling and osmoadaptation in yeasts. Microbiol Mol Biol Rev. 66**,** 300-72.

Homer, N., Merriman, B., Nelson, S. F., 2009. BFAST: an alignment tool for large scale genome resequencing. PLoS One. 4**,** e7767.

Hubmann, G., Guillouet, S., Nevoigt, E., 2011. Gpd1 and Gpd2 fine-tuning for sustainable reduction of glycerol formation in *Saccharomyces cerevisiae*. Appl Environ Microbiol. 77**,** 5857-67.

Huxley, C., Green, E. D., Dunham, I., 1990. Rapid assessment of *S. cerevisiae* mating type by PCR. Trends Genet. 6**,** 236.

Johnston, C. G., Aust, S. D., 1994. Detection of *Phanerochaete chrysosporium* in soil by PCR and restriction enzyme analysis. Appl Environ Microbiol. 60**,** 2350-4.

Kong, Q. X., Cao, L. M., Zhang, A. L., Chen, X., 2007a. Overexpressing *GLT1* in *gpd1Delta* mutant to improve the production of ethanol of *Saccharomyces cerevisiae*. Appl Microbiol Biotechnol. 73**,** 1382-6.

Kong, Q. X., Gu, J. G., Cao, L. M., Zhang, A. L., Chen, X., Zhao, X. M., 2006. Improved production of ethanol by deleting *FPS1* and over-expressing *GLT1* in *Saccharomyces cerevisiae*. Biotechnol Lett. 28**,** 2033-8.

Kong, Q. X., Zhang, A. L., Cao, L. M., Chen, X., 2007b. Over-expressing *GLT1* in a *gpd2Delt*a mutant of *Saccharomyces cerevisiae* to improve ethanol production. Appl Microbiol Biotechnol. 75**,** 1361-6.

Kotaka, A., Sahara, H., Kondo, A., Ueda, M., Hata, Y., 2009. Efficient generation of recessive traits in diploid sake yeast by targeted gene disruption and loss of heterozygosity. Appl Microbiol Biotechnol. 82**,** 387-95.

Larsson, K., Ansell, R., Eriksson, P., Adler, L., 1993. A gene encoding sn-glycerol 3-phosphate dehydrogenase (NAD+) complements an osmosensitive mutant of *Saccharomyces cerevisiae*. Mol Microbiol. 10**,** 1101-11.

Luyten, K., Albertyn, J., Skibbe, W. F., Prior, B. A., Ramos, J., Thevelein, J. M., Hohmann, S., 1995. Fps1, a yeast member of the MIP family of channel proteins, is a facilitator for glycerol uptake and efflux and is inactive under osmotic stress. Embo J. 14**,** 1360-71.

Maeda, T., Takekawa, M., Saito, H., 1995. Activation of yeast PBS2 MAPKK by MAPKKKs or by binding of an SH3-containing osmosensor. Science. 269**,** 554-8.

Nevoigt, E., 2008. Progress in metabolic engineering of *Saccharomyces cerevisiae*. Microbiol Mol Biol Rev. 72**,** 379-412.

Nevoigt, E., Stahl, U., 1996. Reduced pyruvate decarboxylase and increased glycerol-3-phosphate dehydrogenase [NAD+] levels enhance glycerol production in *Saccharomyces cerevisiae*. Yeast. 12**,** 1331-7.

Nevoigt, E., Stahl, U., 1997. Osmoregulation and glycerol metabolism in the yeast *Saccharomyces cerevisiae*. FEMS Microbiol Rev. 21**,** 231-41.

Nicolaou, S. A., Gaida, S. M., Papoutsakis, E. T., 2010. A comparative view of metabolite and substrate stress and tolerance in microbial bioprocessing: From biofuels and chemicals, to biocatalysis and bioremediation. Metab Eng. 12**,** 307-31.

Nissen, T. L., Hamann, C. W., Kielland-Brandt, M. C., Nielsen, J., Villadsen, J., 2000a. Anaerobic and aerobic batch cultivations of *Saccharomyces cerevisiae* mutants impaired in glycerol synthesis. Yeast. 16**,** 463-74.

Nissen, T. L., Kielland-Brandt, M. C., Nielsen, J., Villadsen, J., 2000b. Optimization of ethanol production in *Saccharomyces cerevisiae* by metabolic engineering of the ammonium assimilation. Metab Eng. 2**,** 69-77.

Olesen, K., Franke Johannesen, P., Hoffmann, L., Bech Sorensen, S., Gjermansen, C., Hansen, J., 2000. The pYC plasmids, a series of cassette-based yeast plasmid vectors providing means of counter-selection. Yeast. 16**,** 1035-43.

Oud, B., van Maris, A.J., Daran, J.M., Pronk, J.T., 2012. Genome-wide analytical approaches for reverse metabolic engineering of industrially relevant phenotypes in yeast. FEMS Yeast Res 12, 183-96.

Pagliardini, J., Hubmann, G., Bideaux, C., Alfenore, S., Nevoigt, E., Guillouet, S. E., 2010. Quantitative evaluation of yeast's requirement for glycerol formation in very high ethanol performance fed-batch process. Microb Cell Fact. 9**,** 36.

Parts, L., Cubillos, F. A., Warringer, J., Jain, K., Salinas, F., Bumpstead, S. J., Molin, M., Zia, A., Simpson, J. T., Quail, M. A., Moses, A., Louis, E. J., Durbin, R., Liti, G., 2011. Revealing the genetic structure of a trait by sequencing a population under selection. Genome Res. 21**,** 1131-8.

Pasteur, M. L., 1858. Production constante de glycérine dans la fermentation alcoolique. C R Acad. Sci. 46**,** 857.

Remize, F., Roustan, J. L., Sablayrolles, J. M., Barre, P., Dequin, S., 1999. Glycerol overproduction by engineered *Saccharomyces cerevisiae* wine yeast strains leads to substantial changes in by-product formation and to a stimulation of fermentation rate in stationary phase. Appl Environ Microbiol. 65**,** 143-9.

Sambrook, J., Maniatis, T., Fritsch, E. F., 1989. Molecular cloning : a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Sanger, F., Coulson, A. R., 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J Mol Biol. 94**,** 441-8.

Sato, N., Kawahara, H., Toh-e, A., Maeda, T., 2003. Phosphorelay-regulated degradation of the yeast Ssk1p response regulator by the ubiquitin-proteasome system. Mol Cell Biol. 23**,** 6662-71.

Schuller, D., Casal, M., 2005. The use of genetically modified *Saccharomyces cerevisiae* strains in the wine industry. Appl Microbiol Biotechnol. 68**,** 292-304.

Shepherd, A., Piper, P. W., 2010. The Fps1p aquaglyceroporin facilitates the use of small aliphatic amides as a nitrogen source by amidase-expressing yeasts. FEMS Yeast Res. 10**,** 527-34.

Sherman, F., Hicks, J., 1991. Micromanipulation and dissection of asci. Methods Enzymol. 194**,** 21-37.

Steinmetz, L. M., Sinha, H., Richards, D. R., Spiegelman, J. I., Oefner, P. J., McCusker, J. H., Davis, R. W., 2002. Dissecting the architecture of a quantitative trait locus in yeast. Nature. 416**,** 326-30.

Swinnen, S., Schaerlaekens, K., Pais, T., Claesen, J., Hubmann, G., Yang, Y., Demeke, M., Foulquie-Moreno, M. R., Goovaerts, A., Souvereyns, K., Clement, L., Dumortier, F., Thevelein, J. M., 2012a. Identification of novel causative genes determining the complex trait of high ethanol tolerance in yeast using pooled-segregant whole-genome sequence analysis. Genome Research. 22**,** 975-984.

Swinnen, S., Thevelein, J. M., Nevoigt, E., 2012b. Genetic mapping of quantitative phenotypic traits in *Saccharomyces cerevisiae*. FEMS Yeast Res. 12**,** 215-27.

Warmka, J., Hanneman, J., Lee, J., Amin, D., Ota, I., 2001. Ptc1, a type 2C Ser/Thr phosphatase, inactivates the HOG pathway by dephosphorylating the mitogen-activated protein kinase Hog1. Mol Cell Biol. 21**,** 51-60.

Winzeler, E. A., Richards, D. R., Conway, A. R., Goldstein, A. L., Kalman, S., McCullough, M. J., McCusker, J. H., Stevens, D. A., Wodicka, L., Lockhart, D. J., Davis, R. W., 1998. Direct allelic variation scanning of the yeast genome. Science. 281**,** 1194-7.

Wurgler-Murphy, S. M., Maeda, T., Witten, E. A., Saito, H., 1997. Regulation of the *Saccharomyces cerevisiae* HOG1 mitogen-activated protein kinase by the PTP2 and PTP3 protein tyrosine phosphatases. Mol Cell Biol. 17**,** 1289-97.

Yadav, V. G., De Mey, M., Lim, C. G., Ajikumar, P. K., Stephanopoulos, G., 2012. The future of metabolic engineering and synthetic biology: towards a systematic practice. Metab Eng. 14**,** 233-41.

**Figure 1. Variation in glycerol and ethanol yield.**

(**A**) Glycerol/ethanol ratio in 52 natural and industrial *S. cerevisiae* strains. The selected diploids used for quantitative trait analyis, Ethanol Red (inferior parent, target industrial bioethanol production strain) and CBS6412 (superior parent) are marked in black. (**B**) Normal distribution of the glycerol/ethanol ratio around a mean value of 5.7% for the 52 strains. Fermentations were carried out in 250 ml oxygen-limited and stirred medium containing 1% w/v yeast extract, 2% w/v peptone and 12% w/v glucose. The glucose was essentially fermented to completion by all strains.

**Figure 2. Glycerol and ethanol yield in the segregants of the diploid parent strains and in segregants from the cross between the selected haploid parent strains.**

(**A**) Glycerol and ethanol yield of the diploids, Ethanol Red and CBS6412, and the CBS6412 segregant, CBS4C, showing the lowest glycerol yield of all tested segregants of CBS6412. Fermentations were carried out in 100 ml minimal medium with 10% glucose. (**B**) Distribution of the glycerol yield in the haploid segregants of CBS6412 (black bars) and Ethanol Red (white bars). The distribution was normal around the value of the diploid parents, CBS6412 (left small black square on top) and Ethanol Red (right small black square on top). (**C**) Distribution of the glycerol yield in the segregants from the cross CBS4C x ER7A (white bars). All segregants were screened in 5 ml fermentations. After evaluation of the 48 segregants with the lowest glycerol yield in 100 ml minimal medium with 5% glucose, the 44 segregants with the lowest glycerol production were selected for pooled-segregant whole-genome sequence analysis (black bars). The glycerol yield of the haploid parents CBS4C and ER7A is indicated with small black squares on top.

**Figure 3. Plots of SNP variant frequency versus chromosomal position and corresponding P-values.**

The variation in SNP variant frequency is shown for all 16 yeast chromosomes (raw data: small grey circles; smoothened data: black line; statistical confidence interval: grey lines). Significant upward deviations from the average of 0.5 indicate linkage to the superior parent CBS4C, while significant downward deviations indicate linkage to the inferior parent ER7A. The smoothened line was determined as described previously (Swinnen *et al.*, 2012a). Strong candidate QTLs were found on chromosome II (at position 500,000 – 700,000 bp) and chromosome XII (at position 135,000 – 200,000 bp), but only for the latter the P-value dropped below the significance limit of 0.05.

**Figure 4. SNP variant frequency and P-values determined in individual segregants for downscaling of the QTLs.**

Top: SNP variant frequency map of chromosome II (**A**) and chromosome XII (**B**) determined with the pool of 20 selected segregants (raw data: small circles; smoothened data: black line; statistical confidence interval: black stippled lines) and the pool of 20 unselected segregants (raw data: small triangles; smoothened data: grey line; statistical confidence interval: grey stippled lines). Middle: SNP variant frequency of seven selected SNPs in the candidate regions on chromosome II (at position 500,000 – 700,000 bp) (**A**) and chromosome XII (at position 135,000 – 200,000 bp) (**B**), determined in the individual 20 most superior segregants ( ) and the individual 44 most superior segregants (●). Smoothened lines: SNP variant frequency determined with the pool of 20 selected segregants (black line) and the pool of 20 unselected segregants (grey line). Bottom: P-values for the same seven SNPs in the regions on chromosome II (**A**) and XII (**B**). The statistical confidence line (P-value ≤ 0.05) is also indicated. The region on chromosome XII was significantly linked. Lowest panel: overview of all genes present and all SNPs identified in the region with the highest linkage of the QTL on chromosome XII (154,000 bp – 175,000 bp). Genes marked with a star contained a non-synonymous mutation in the ORF.

**Figure 5. Identification of *SSK1* as the causative gene in the QTL on chromosome XII.**

Diploid strains constructed for reciprocal hemizygosity analysis (RHA) with either the deletion of the $ssk1^{E330N...K356N}$ allele of CBS4C or the deletion of the *SSK1* allele of ER7A. Glycerol and ethanol yield of the two hemizygous diploid strains. The difference in glycerol and ethanol yield for the two diploids was significant by the Student t-test. Fermentations were carried out in 100 ml minimal medium with 5% glucose.

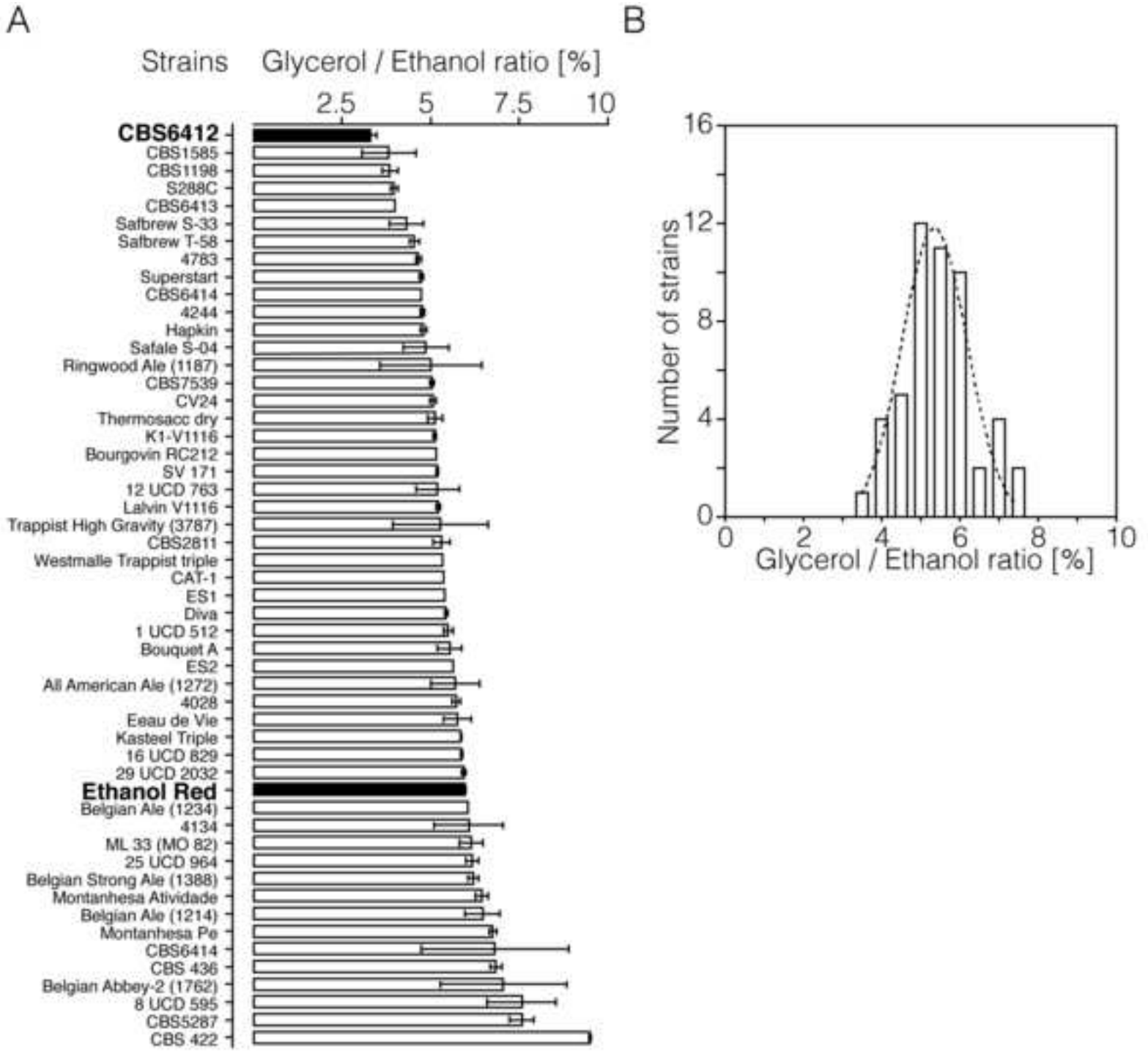
**Figure 6. Glycerol and ethanol yield after deletion or reciprocal exchange of the *SSK1* alleles.**

(**A**) Comparison of the glycerol and ethanol yield of the two hemizygous strains with that of the ER7A and CBS4C parental strains and the *SSK1*/*ssk1*$^{E330N...K356N}$ and *ssk1*Δ/*ssk1*Δ diploids. (**B**) Glycerol and ethanol yield and maximal volumetric ethanol production rate in the ER7A and CBS4C parental strains, *SSK1* deletion versions and the ER7A and CBS4C strains with reciprocal exchange of the *SSK1* and *ssk1*$^{E330N...K356N}$ alleles. Fermentations were carried out in 100 ml minimal medium with 5% glucose. (**C**) Growth comparison of the strains, as detailed in (**B**), under aerobic conditions. Fivefold dilutions of overnight pre-grown cells were spotted on YPD medium and incubated at 30°C for two days. The maximum growth rate in liquid cultures was determined from $OD_{600}$ measurements in 50 ml YPD shake flasks incubated at 30˚C.

**Figure 7. Glycerol and ethanol yields and osmostress tolerance in fermentations with the industrial bioethanol production strain Ethanol Red in which one or two copies of the *ssk1*$^{E330N...K356N}$ allele had been introduced.**

(**A**) Glycerol and ethanol yields in fermentations with minimal medium (5% w/v glucose), high gravity medium (YP with 33% w/v glucose) and wheat hydrolyzate (SHF: Separate Hydrolysis and Fermentation). (**B**) Glycerol/ethanol ratio and maximal volumetric ethanol production rate ($r_{max}$ in g $l^{-1}$ $h^{-1}$) in fermentations with minimal medium (5% glucose) in the presence of NaCl (0, 0.7 and 1M) or sorbitol (0, 1.4 and 2M).

Figure 1



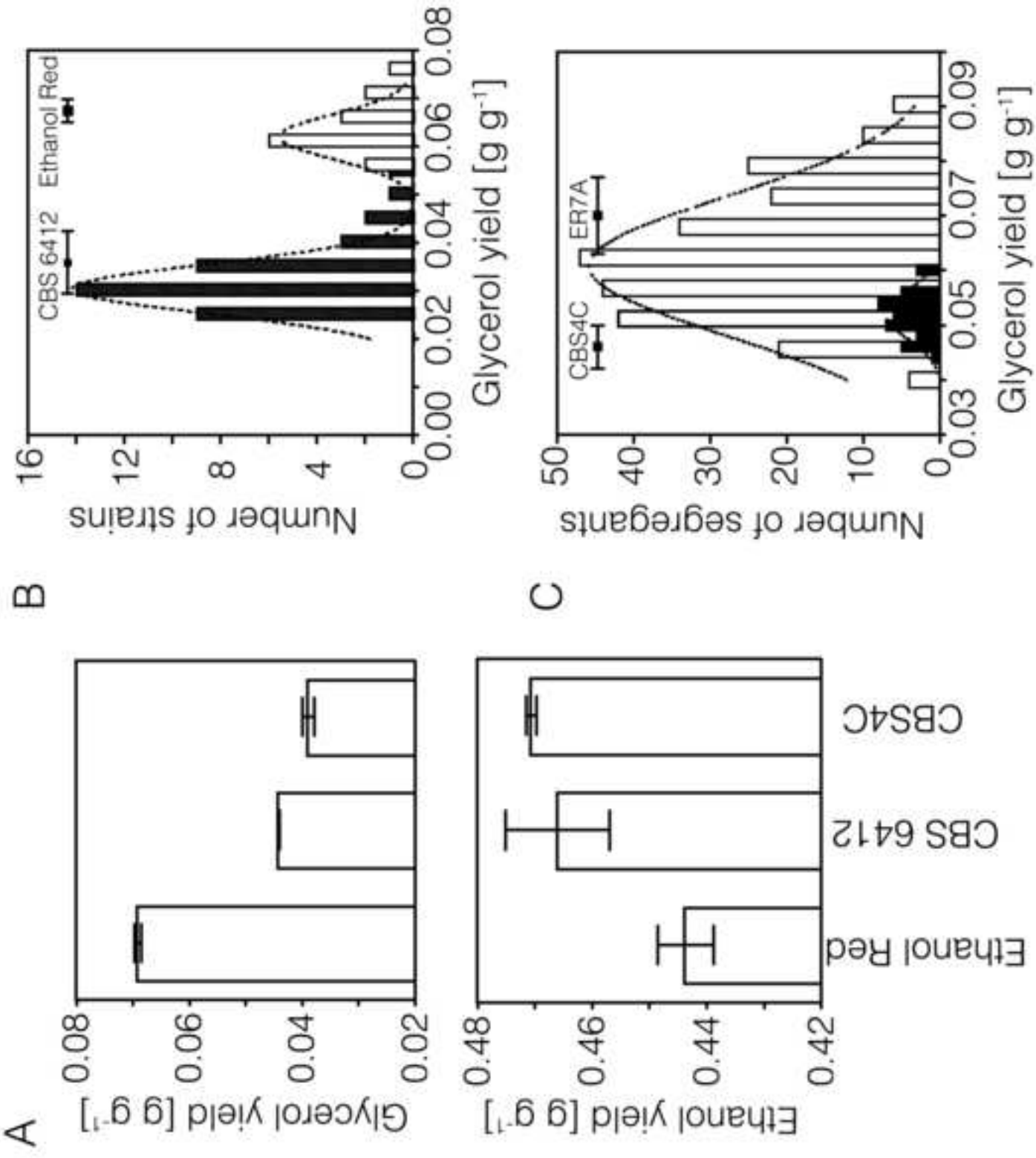

A
Glycerol yield [g g-1]
0.08
0.06
0.04
0.02
Ethanol yield [g g-1]
0.48
0.46
0.44
0.42
Ethanol Red
CBS 6412
CBS4C
B
Number of strains
16
12
8
4
0
CBS 6412
Ethanol Red
0.00
0.02
0.04
0.06
0.08
Glycerol yield [g g-1]
C
Number of segregants
50
40
30
20
10
0
CBS4C
ER7A
0.03
0.05
0.07
0.09
Glycerol yield [g g-1]

Figure 2

Figure 3



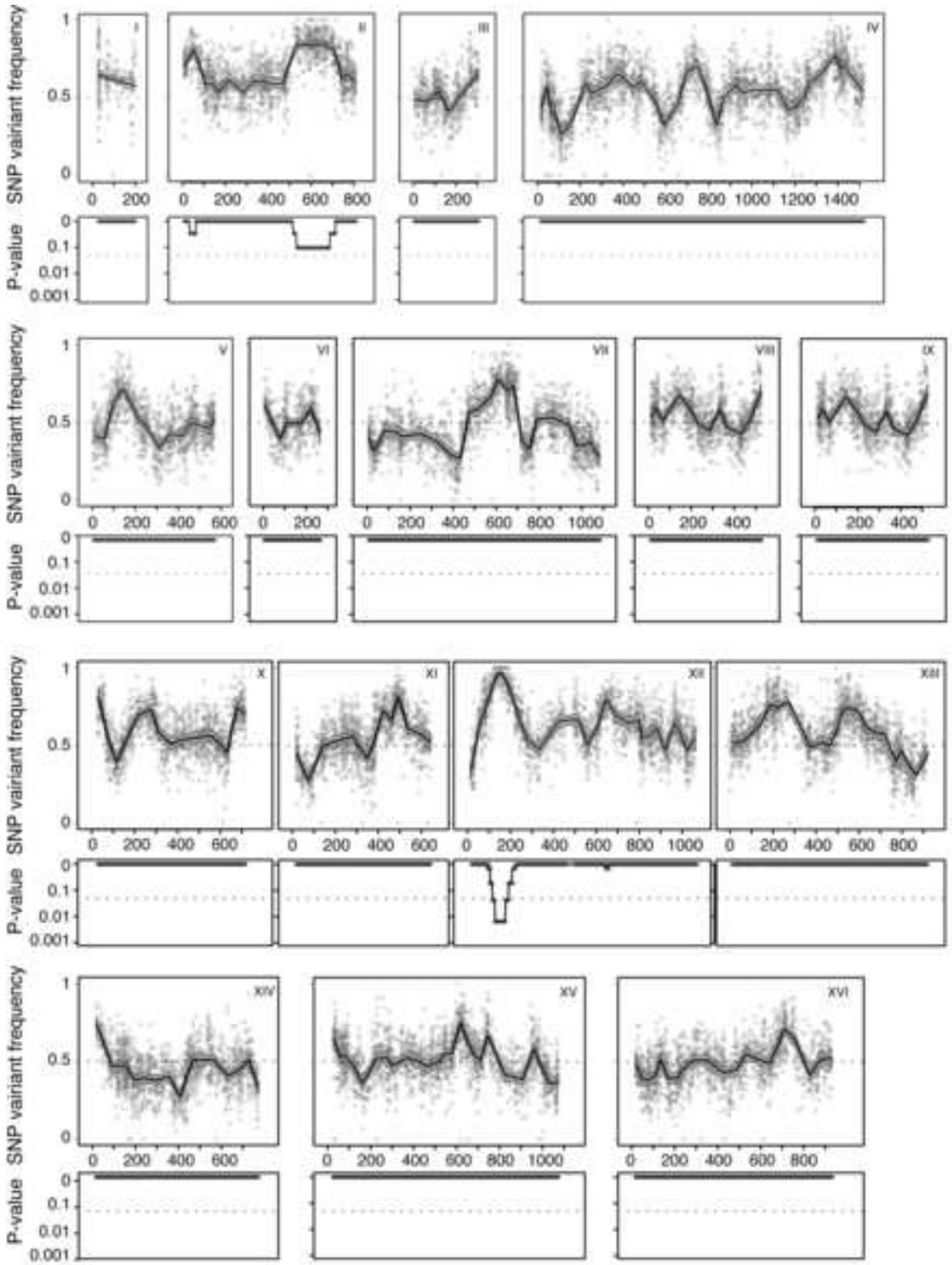
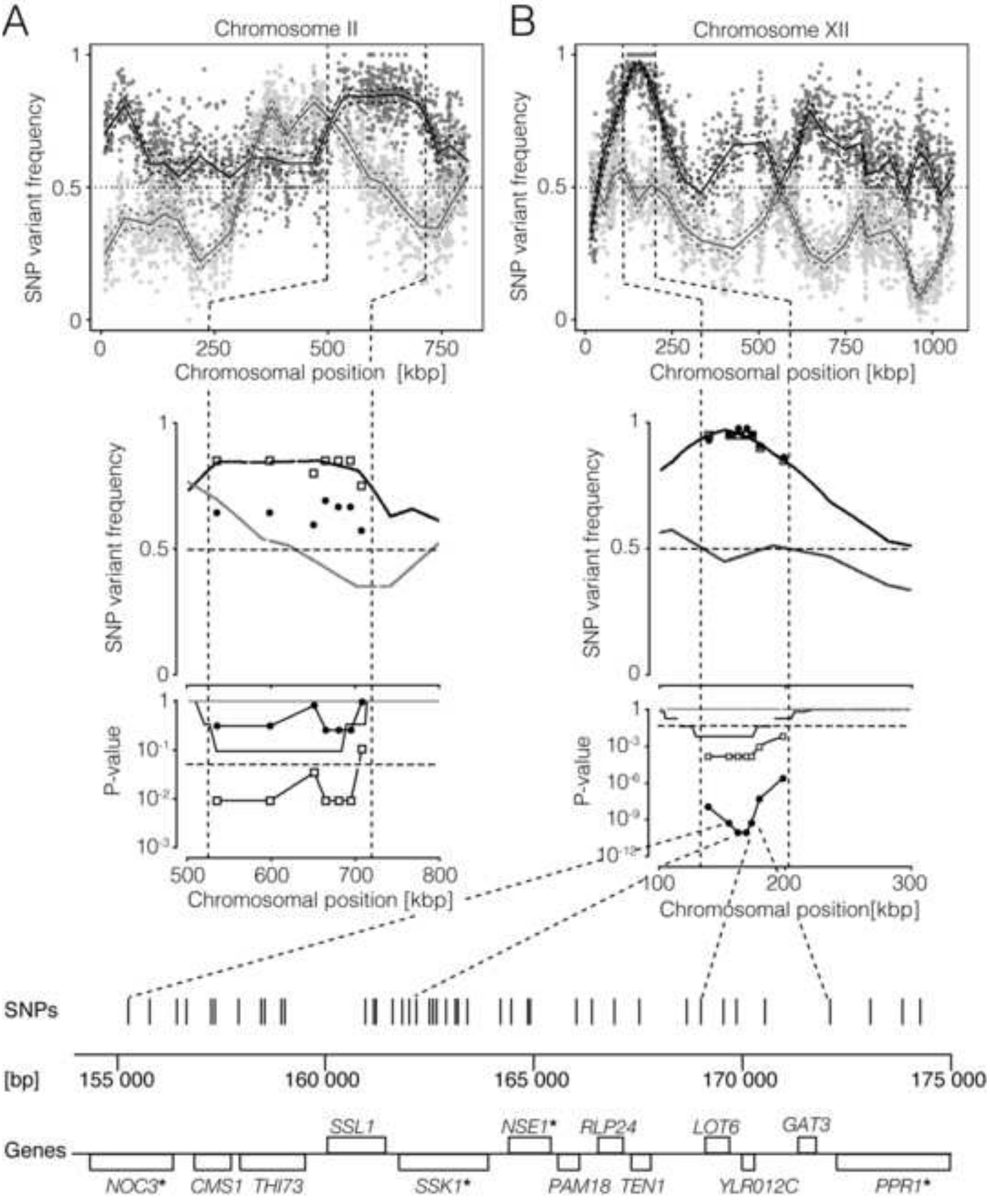


Figure 4





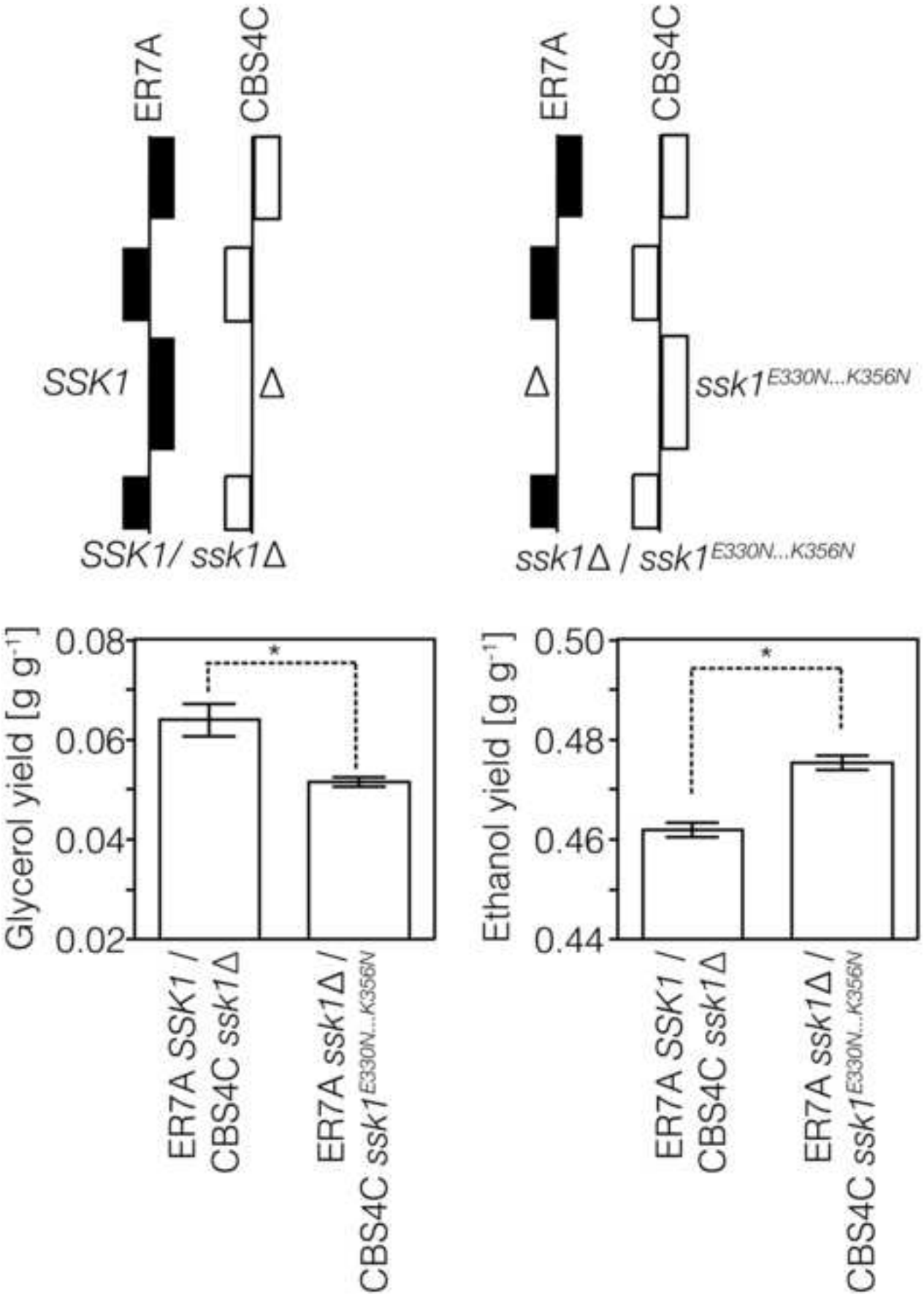


Figure 6



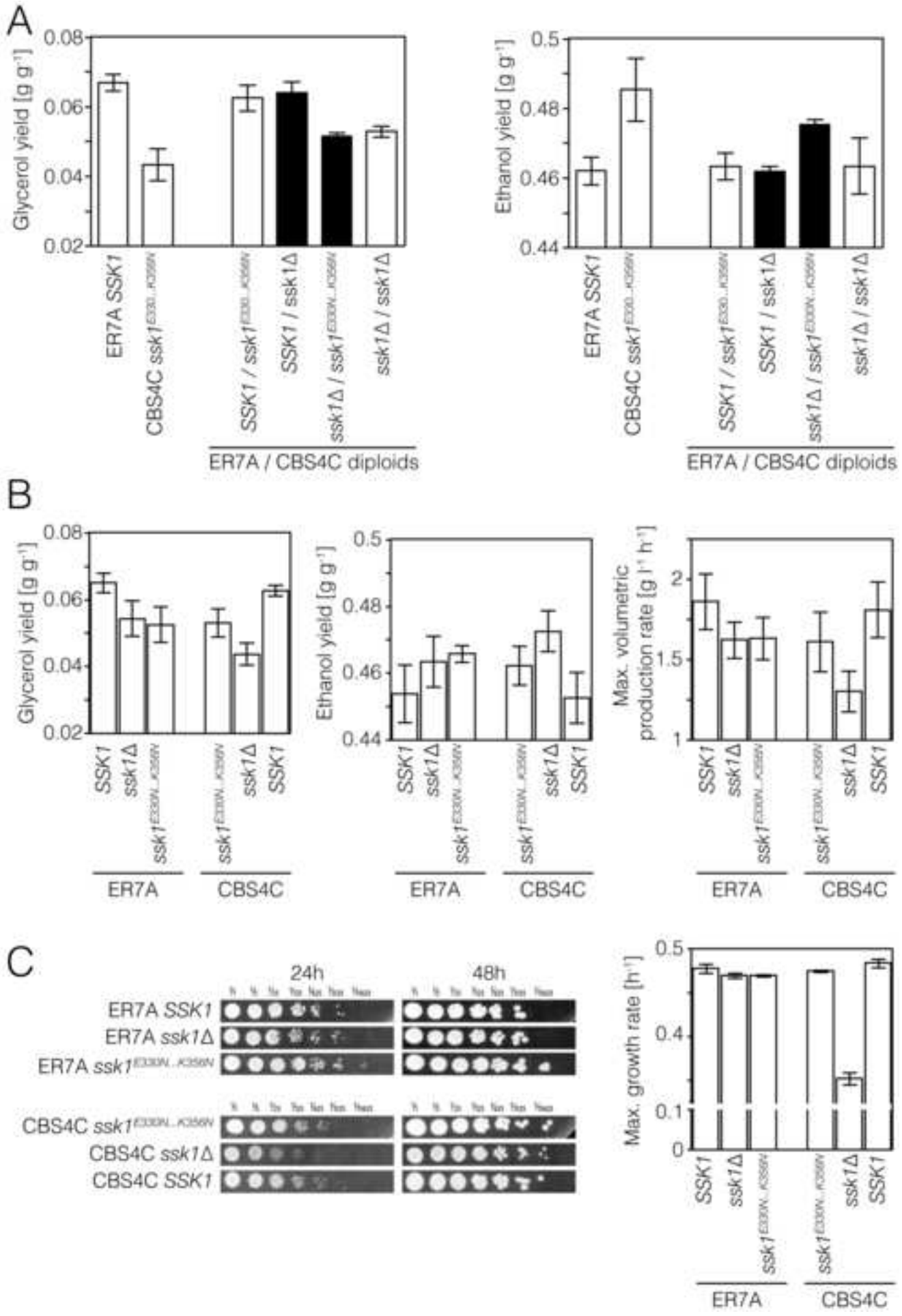


Figure 7



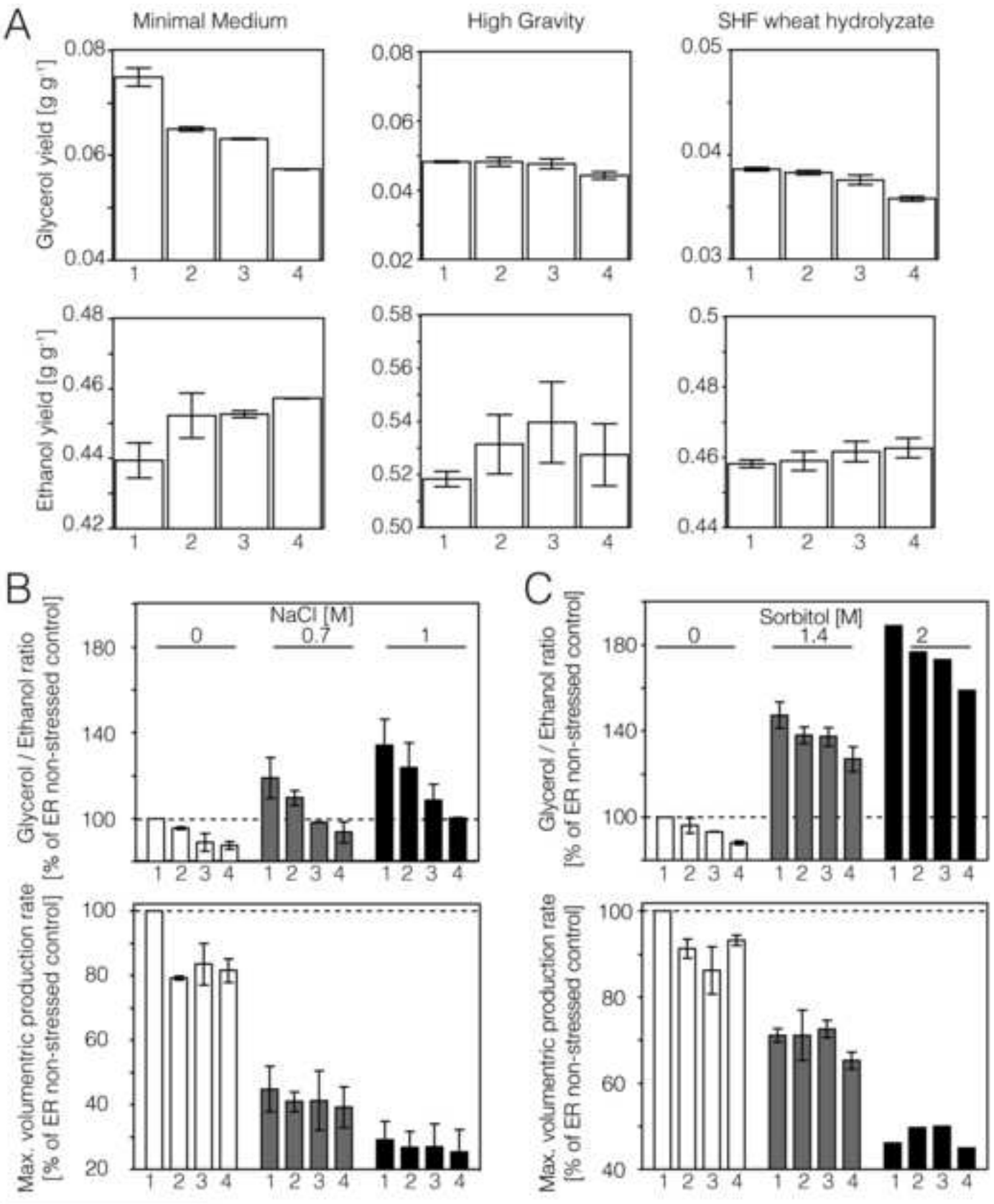


**Table 1.** *Saccharomyces cerevisiae* strains used in this work.

| **Strain** | **Genotype** | **Source** |
|---|---|---|
| CBS6412 | *Diploid, ssk1$^{E330N...K356N}$/ssk1$^{E330N...K356N}$* | CBS-KNAW |
| Ethanol Red | *Diploid, SSK1/SSK1* | Fermentis, S. I. Lesaffre |
| ER7A | *Segregant 7A of Ethanol Red, Matα* | This study |
| CBS4C | *Segregant 4C of CBS6412, Mat**a*** | This study |
| ER7A *ssk1*Δ | *ER7A, ssk1*Δ *(marker removed)* | This study |
| CBS4C *ssk1*Δ | *CBS4C, ssk1*Δ *(marker removed)* | This study |
| ER7A x CBS4C | *Hybrid diploid ER7A x CBS4C* | This study |
| ER7A x CBS4C ssk1Δ | *Hybrid diploid ER7A x CBS4C ssk1*Δ | This study |
| ER7A ssk1Δ x CBS4C | *Hybrid diploid ER7A ssk1*Δ *x CBS4C* | This study |
| ER7A ssk1Δ x CBS4C ssk1Δ | *Hybrid diploid ER7A ssk1*Δ *x CBS4C ssk1*Δ | This study |
| ER7A *ssk1$^{E330N...K356N}$* | *ER7A ssk1$^{E330N...K356N}$ (SSK1 allele exchanged)* | This study |
| CBS4C *SSK1* | *CBS4C SSK1 (SSK1 allele exchanged)* | This study |
| HG5 | Ethanol Red *ssk1*Δ*/ssk1*Δ | This study |
| HG7 | Ethanol Red *ssk1$^{E330N...K356N}$/ssk1*Δ | This study |
| HG8 | Ethanol Red *ssk1$^{E330N...K356N}$/ ssk1$^{E330N...K356N}$* | This study |

***SSK1*** **alleles:** *SSK1 = SSK1 allele of Ethanol Red, ssk1$^{E330N...K356N}$ = SSK1 allele of CBS6412, ssk1*Δ *= deletion of SSK1*

**Highlights**

- Low glycerol yield in yeast fermentation was analyzed as a complex genetic trait
- Pooled-segregant whole-genome sequence analysis revealed one major locus
- A truncated, partially mistranslated *SSK1* allele was identified as causative gene
- The *SSK1* allele reduced glycerol/ethanol ratio more than *ssk1*Δ in industrial strain
- The *SSK1* allele caused less side-effects on growth, volumetric productivity and osmotolerance than the deletion of *SSK1*.