RESEARCH ARTICLE

Laboratory evolution of new lactate transporter genes in a *jen1*Δ mutant of *Saccharomyces cerevisiae* and their identification as *ADY2* alleles by whole-genome resequencing and transcriptome analysis

Stefan de Kok[1,2], Jurgen F. Nijkamp[2,3], Bart Oud[1,2], Filipa C. Roque[1,2], Dick de Ridder[2,3], Jean-Marc Daran[1,2], Jack T. Pronk[1,2] & Antonius J. A. van Maris[1,2]

[1]Department of Biotechnology, Delft University of Technology, Delft, The Netherlands; [2]Kluyver Centre for Genomics of Industrial Fermentation, Delft, The Netherlands; and [3]The Delft Bioinformatics Lab, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Delft, The Netherlands

Correspondence: Antonius J.A. van Maris, Department of Biotechnology, Delft University of Technology, Julianalaan 67, Delft 2628 BC, The Netherlands. Tel.: +31 15 278 1616; fax: +31 15 278 2355; e-mail: A.J.A.vanMaris@TUDelft.nl

Received 24 October 2011; revised 7 December 2011; accepted 18 January 2012. Final version published online 18 January 2012.
DOI: 10.1111/j.1567-1364.2011.00787.x

Editor: Jens Nielsen

Keywords
organic acids; reverse metabolic engineering; chromosomal rearrangement; Ty1 retrotransposon.

Abstract

Laboratory evolution is a powerful approach in applied and fundamental yeast research, but complete elucidation of the molecular basis of evolved phenotypes remains a challenge. In this study, DNA microarray-based transcriptome analysis and whole-genome resequencing were used to investigate evolution of novel lactate transporters in *Saccharomyces cerevisiae* that can replace Jen1p, the only documented *S. cerevisiae* lactate transporter. To this end, a *jen1*Δ mutant was evolved for growth on lactate in serial batch cultures. Two independent evolution experiments yielded growth on lactate as sole carbon source (0.14 and 0.18 h^{-1}, respectively). Transcriptome analysis did not provide leads, but whole-genome resequencing showed different single-nucleotide changes (C755G/Leu219Val and C655G/Ala252Gly) in the acetate transporter gene *ADY2*. Introduction of these *ADY2* alleles in a *jen1*Δ *ady2*Δ strain enabled growth on lactate (0.14 h^{-1} for Ady2p^{Leu219Val} and 0.12 h^{-1} for Ady2p^{Ala252Gly}), demonstrating that these alleles of *ADY2* encode efficient lactate transporters. Depth of coverage of DNA sequencing, combined with karyotyping, gene deletions and diagnostic PCR, showed that an isochromosome III (*c.* 475 kb) with two additional copies of *ADY2*^{C755G} had been formed via crossover between retrotransposons *YCLWΔ15* and *YCRCΔ6*. The isochromosome formation shows how even short periods of selective pressure can cause substantial karyotype changes.

Introduction

Laboratory evolution (referred to as 'evolutionary engineering' in applied contexts) is a powerful strategy to obtain microorganisms with improved characteristics that are difficult to engineer through rational approaches, especially when the genetic determinants for a phenotype are unknown or complex (Sauer, 2001). Subsequent discovery of the underlying mutation(s) is of key importance, both for functional gene analysis and to enable transfer of the acquired traits to nonevolved industrially relevant strain backgrounds ('reverse metabolic engineering'; Bailey *et al.*, 1996). In addition to the mutation(s) of interest, random mutations can accumulate during evolution. Combined with our incomplete understanding of metabolic and regulatory networks, this makes elucidation of the genetic basis of acquired phenotypes the main challenge of reverse metabolic engineering. During evolution, different types of mutations can occur: (1) small local changes (substitutions, insertions, deletions and duplications) that result in either differential gene expression or changes in the amino acid sequence of the encoded protein and/or (2) rearrangements of larger DNA fragments, potentially resulting in a different gene dosage or expression (Sauer, 2001).

Several genome-wide techniques have been described in literature to analyse mutations that occur during laboratory evolution (for reviews, see Gill, 2003; Bro & Nielsen, 2004; Gresham *et al.*, 2008a). Classical techniques include construction and analysis of plasmid-based genomic libraries and quantitative trait locus mapping via crossing. More recent developments have enabled whole-genome DNA sequencing, DNA hybridization to DNA microarrays, transcriptome analysis using DNA microarrays, proteome, metabolome and fluxome profiling (Gill, 2003; Bro & Nielsen, 2004; Gresham *et al.*, 2008b). Analysis with DNA microarrays was, until recently, the only affordable genome-wide method. Microarray-based transcriptome analysis only indirectly indicates changes on the DNA level. Several factors, such as complex regulatory networks, transcription factor binding, translation efficiency and stability of mRNA and proteins, complicate the interpretation of the exact genetic basis of changes observed at mRNA and protein level. In contrast, comparative genome hybridization using DNA microarrays can identify genetic changes directly (Winzeler *et al.*, 1999; Daran Lapujade *et al.*, 2003; Schacherer *et al.*, 2007). However, limitations in the maximum probe density imply that either (1) the resolution of these DNA microarrays is limited to several base pairs, necessitating sequencing of selected regions or (2) multiple DNA microarrays are required to cover larger (eukaryotic) genomes. Furthermore, DNA microarrays can only be used for a comprehensive analysis when the strains under investigation are congenic to the strain on whose genome sequence the microarray design was based. In contrast, whole-genome DNA sequencing, which because of recent price developments has become an interesting technique for analysis of evolved strains (Nowrousian, 2010; Sboner *et al.*, 2011), has the potential to directly reveal all mutations accumulated in an evolved strain. However, discriminating between random and relevant mutations remains a challenging task. Combining DNA sequencing with other genome-wide techniques that provide insight in the underlying physiology can help to elucidate the genetic basis of a phenotype acquired via evolution.

Transport of carboxylic acids plays a key role in weak organic acid stress (Palmqvist & Hahn-Hägerdal, 2000) and in metabolic engineering strategies for organic acid production with *Saccharomyces cerevisiae* (van Maris *et al.*, 2004a, b; Jol *et al.*, 2010). However, the responsible membrane transporters are often poorly studied and encoded by multiple redundant genes (Casal *et al.*, 2008). The goal of this study was to investigate whether laboratory evolution and subsequent elucidation of the underlying mutations can lead to the identification of alternative lactate transporters in *S. cerevisiae*. In *S. cerevisiae*, Jen1p was previously identified as the only efficient lactate importer by generation of a UV mutant unable to grow on lactate and subsequent functional complementation with a genomic library (Casal *et al.*, 1999). The lactate uptake rate of the resulting *jen1* null mutant was close to the detection limit (Casal *et al.*, 1999). Interestingly, export of lactate in engineered lactate producing *S. cerevisiae* is unaffected by deletion of *JEN1* (our unpublished results). Even though mechanisms for import or export can differ, these combined observations suggest the presence of at least one alternative lactate transporter. To explore the presence of other lactate transporters in *S. cerevisiae*, a *jen1*Δ strain was constructed and evolved for growth on lactate as the sole carbon and energy source. The evolved strains were subjected to a combination of transcriptome analysis and whole-genome DNA (re)sequencing to identify the relevant mutations. The resulting lead genes were tested for their lactate transport activity via knockout studies in the evolved strains and introduction into nonevolved strains.

Materials and methods

Strains and maintenance

The *S. cerevisiae* strains used and constructed in this study (Table 1) are all derived from CEN.PK113-7D (van Dijken *et al.*, 2000; Entian & Kotter, 2007). Stock cultures were grown at 30 °C in shake flasks containing either 100 mL synthetic medium (Verduyn *et al.*, 1992) with 20 g L^{-1} glucose as carbon source or 100 mL complex medium containing 20 g L^{-1} glucose, 10 g L^{-1} Bacto yeast extract and 20 g L^{-1} Bacto peptone. After overnight growth, 20% (v/v) glycerol was added and 1-mL aliquots were stored at −80 °C.

Plasmid and strain construction

The plasmids and primers used in this study are listed in Tables 2 and 3, respectively. Transformations of *S. cerevisiae* were carried out according to the LiAc/ssDNA method (Gietz & Woods, 2002). Gene deletions were performed using the loxP-marker-loxP/Cre recombinase system, using pUG6, pUG66, pUG-*natNT2* and pUG-*hphNT1* as templates for PCR amplification of the knockout cassettes (Gueldener *et al.*, 2002). Selection of knockout mutants was performed on agar plates containing 20 g L^{-1} glucose, 10 g L^{-1} Bacto yeast extract, 20 g l^{-1} Bacto peptone and either 200 mg L^{-1} G418, 10 mg L^{-1} phleomycin, 100 mg L^{-1} nourseothricin or 200 mg L^{-1} hygromycin B. The *natNT2* marker was cloned from pFA6a-natNT2 (Janke *et al.*, 2004) to pUG6 via the *Sac*I and *Bgl*II restriction sites, thereby replacing the *KanMX4* marker, resulting in pUG-*natNT2*. The *JEN1* knockout cassette was amplified

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

Strain	Relevant genotype	Source
CEN.PK113-7D	MATa *URA3 ADY2 JEN1*	P. Kötter, Frankfurt
IMK302	MATa *URA3 ADY2 jen1::loxP-KanMX4-loxP*	This study
IMW004	IMK302 evolved for growth on lactate	This study
IMW005	IMK302 evolved for growth on lactate	This study
IMW032	IMW004 *ady2::loxP-hphNT1-loxP*	This study
IMW040	IMW004 *ady2::loxP-hphNT1-loxP ady2::loxP-natNT2-loxP*	This study
IMW041	IMW004 *ady2::loxP-hphNT1-loxP ady2::loxP-natNT2-loxP ady2::loxP-ble-loxP*	This study
IMW033	IMW005 *ady2::loxP-hphNT1-loxP*	This study
IMK322	MATa *ura3::loxP-hphNT1-loxP ADY2 jen1::loxP-KanMX4-loxP*	This study
IMK338	MATa *ura3::loxP ADY2 jen1::loxP*	This study
IMK341	MATa *ura3::loxP ady2::loxP-hphNT1-loxP jen1::loxP*	This study
IMZ271	IMK341 pRS416 (*CEN6/ARS4*, *URA3*)	This study
IMZ272	IMK341 pUDC011 (*CEN6/ARS4*, *URA3*, P_{ADY2}-*ADY2*-T_{ADY2})	This study
IMZ273	IMK341 pUDC012 (*CEN6/ARS4*, *URA3*, P_{ADY2}-*ADY2*^{C655G}-T_{ADY2})	This study
IMZ276	IMK341 pUDC015 (*CEN6/ARS4*, *URA3*, P_{ADY2}-*ADY2*^{C755G}-T_{ADY2})	This study

Table 2. Plasmids used in this study

Plasmid	Characteristic	Reference
pCR-BLUNT II-TOPO	Gateway entry plasmid	Invitrogen
pRS416	Centromeric plasmid, *URA3*	Sikorski & Hieter (1989)
pUG6	PCR template for loxP-*KanMX4*-loxP cassette	Gueldener *et al.* (2002)
pUG66	PCR template for loxP-*ble*-loxP cassette	Gueldener *et al.* (2002)
pSH47	Centromeric plasmid, *URA3*, P_{GAL1}-*Cre*-T_{CYC1}	Gueldener *et al.* (2002)
pUG-*hphNT1*	PCR template for loxP-*hphNT1*-loxP cassette	De Kok *et al.* (2011)
pFA6a-*natNT2*	Plasmid with *natNT2* marker	Janke *et al.* (2004)
pUG-*natNT2*	PCR template for loxP-*natNT2*-loxP cassette	This study
pUD151	Gateway entry clone, P_{ADY2}-*ADY2*-T_{ADY2}	This study
pUD152	Gateway entry clone, P_{ADY2}-*ADY2*^{C655G}-T_{ADY2}	This study
pUD152	Gateway entry clone, P_{ADY2}-*ADY2*^{C755G}-T_{ADY2}	This study
pUDC011	Centromeric plasmid, *URA3*, P_{ADY2}-*ADY2*-T_{ADY2}	This study
pUDC012	Centromeric plasmid, *URA3*, P_{ADY2}-*ADY2*^{C655G}-T_{ADY2}	This study
pUDC015	Centromeric plasmid, *URA3*, P_{ADY2}-*ADY2*^{C755G}-T_{ADY2}	This study

from pUG6 using primers JEN1 KO Fw and JEN1 KO Rv and transformed to CEN.PK113-7D, resulting in strain IMK302. Correct integration of the knockout cassette was confirmed using primer combinations JEN1 Ctrl Fw/KanMX4 Ctrl Rv and KanMX4 Ctrl FW/JEN1 Ctrl Rv. The *URA3* knockout cassette was amplified from pUG-hphNT using primers URA3 KO Fw and URA3 KO Rv and transformed to IMK302, resulting in IMK322. Correct integration of the knockout cassette was confirmed using primer combinations URA3 Ctrl Fw/hphNT1 Ctrl Rv and hphNT1 Ctrl FW/URA3 Ctrl Rv. The *KanMX4* and *hphNT1* markers in IMK322 were removed using the Cre/loxP system with pSH47 (Gueldener *et al.*, 2002), resulting in IMK338. Correct marker removal was confirmed using primer combinations JEN1 Ctrl Fw/JEN1 Ctrl Rv and URA3 Ctrl Fw/URA3 Ctrl Rv.

The first *ADY2* knockout cassette was amplified from pUG-*hphNT1* using primers ADY2 KO Fw A and ADY2 KO Rv B and transformed to IMK338, IMW004 and IMW005, resulting in IMK341, IMW032 and IMW033, respectively. Correct integration of the knockout cassette was confirmed using primer combinations ADY2 Ctrl Fw/hphNT1 Ctrl Rv and hphNT1 Ctrl FW/ADY2 Ctrl Rv. The second *ADY2* knockout cassette was amplified from pUG-*natNT2* using primers ADY2 KO Fw B and ADY2 KO Rv B and transformed to IMW032, resulting in IMW040. Correct integration of the knockout cassette was confirmed using primer combinations ADY2 Ctrl Fw/natNT2 Ctrl Rv and natNT2 Ctrl FW/ADY2 Ctrl Rv. The third *ADY2* knockout cassette was amplified from pUG66 using primers ADY2 KO Fw C and ADY2 KO Rv C and transformed to IMW040, resulting in

Table 3. Primers used in this study

Name	Sequence (5′→3′)
JEN1 KO Fw	TGCACATCATTGTTGAGAAATAGTTTTGGAAGTTGTCTAGTCCTTCTCCCCAGCTGAAGCTTCGTACGC
JEN1 KO Rv	GAAATGCAGTTACATAGAGAAGCGAACACGCCCTAGAGAGCAATGAAAAGGCATAGGCCACTAGTGGATCTG
JEN1 Ctrl Fw	GCGCGGCTTGAAACTATTTCTCC
JEN1 Ctrl Rv	GGCCCATTCAGTGCAAGAACC
ADY2 KO Fw A	CACAGATATAACTAAACAACCACAAAACAACTCATATACAAACAAATAATCAGCTGAAGCTTCGTACGC
ADY2 KO Rv A	CAATAGTTCTCGTTATTAGTAGGTCGTGCTCTTAAAAGATTACCCTTTCAGCATAGGCCACTAGTGGATCTG
ADY2 KO Fw B	ATGTCTGACAAGGAACAAACGAGCGGAAACACAGATTTGGAGAATGCACCCAGCTGAAGCTTCGTACGC
ADY2 KO Rv B	GATGGTAATGGGAATGGACGAGCCAGTACATATGAATTCTGCTTTGTAGCGCATAGGCCACTAGTGGATCTG
ADY2 KO Fw C	GATACTATAGTTCCCATGATAACGACGTTAATGGCGTTGCAGAAGATGAACCAGCTGAAGCTTCGTACGC
ADY2 KO Rv C	CTGCATATGCGTTGTACCAAGCAATGAAAGCAACAACAACTCCCAGGACACGCATAGGCCACTAGTGGATCTG
ADY2 Ctrl Fw	GGGCGGCTATTTCGGTATTG
ADY2 Ctrl Rv	GCAACATCTTCTCCGCTAACC
URA3 KO Fw	TCCTAGTCCTGTTGCTGCCAAGCTATTTAATATCATGCACGAAAAGCAAACAAACTTGTGGGGCACACCCAA TTTAGACACAGCTGAAGCTTGTACGC
URA3 KO Rv	TTCCCAGCCTGCTTTTCTGTAACGTTCACCCTCTACCTTAGCATCCCTTCCCTTTGCAAATAGGCGCGTCAAG TACAAAGTTAGCATGCCACTAGTGGATCTG
URA3 Ctrl Fw	GCTACTGCGCCAATTGATGAC
URA3 Ctrl Rv	CGAGATTCCCGGGTAATAACTG
KanMX4 Ctrl Fw	TCGTATGTGAATGCTGGTCG
KanMX4 Ctrl Rv	CGCACGTCAAGACTGTCAAG
hphNT1 Ctrl Fw	ACGCGGATTTCGGCTCCAAC
hphNT1 Ctrl Rv	AGACGTCGCGGTGAGTTCAG
natNT2 Ctrl Fw	CCACTCTTGACGACACGGCTTAC
natNT2 Ctrl Rv	GTAAGCCGTGTCGTCAAGAGTGG
ble Ctrl Fw	CGAATTGCTTGCAGGCATCTC
ble Ctrl Rv	AATCTCGTGATGGCAGGTTGG
ADY2 inside Fw	GCCTTTGGTGGTACCTTGAATC
ADY2 inside Rv	GCTCTTGTGACACCAAGTCTATTAGC
ADY2p Fw	CACCACCCGGGAGGATCCGATAGGCGTCGTATATAGTCTCTTC
ADY2t Rv	GCACTAGTTCTAGACGCATTTGGATCGTCCAG
P5	TGCACCAGCAAATGAGGCTTC
P6	ATAGTCAATGCCCGTGGTTC

IMW041. Correct integration of the knockout cassette was confirmed using primer combinations ADY2 Ctrl Fw/ble Ctrl Rv and ble Ctrl FW/ADY2 Ctrl Rv. Presence or absence of the *ADY2* gene was confirmed using primers ADY2 inside Fw and ADY2 inside Rv. The *ADY2* gene including promoter and terminator was amplified from genomic DNA of CEN.PK113-7D, IMW004 and IMW005 using primers ADY2p Fw and ADY2t Rv, and the resulting PCR products were cloned into pCR-BLUNT II-TOPO using Gateway Technology (Invitrogen, Carlsbad, NM), resulting in pUD151, pUD153 and pUD152, respectively. The *ADY2* alleles including promoter and terminator were subsequently cloned into pRS416 via the *Bam*HI and *Xba*I sites, resulting in pUDC011, pUDC015 and pUDC012, respectively. pRS416, pUDC011, pUDC012 and pUDC015 were transformed into IMK341 and selected on a synthetic medium without uracil, resulting in IMZ271, IMZ272, IMZ273 and IMZ276.

Molecular biology techniques

PCR amplification was performed using Phusion Hot Start High Fidelity Polymerase (Finnzymes, Espoo, Finland) according to manufacturer's instructions in a Biometra TGradient Thermocycler (Biometra, Gottingen, Germany). DNA fragments were separated on a 1% (w/v) agarose (Sigma, St. Louis, MO) gel in 1xTAE (40 mM Tris–acetate pH 8.0 and 1 mM EDTA). Isolation of fragments from gel was performed with the Zymoclean Gel DNA Recovery kit (Zymo Research, Orange, CA). Restriction endonucleases (New England Biolabs, Beverly, MA) and DNA ligases (Roche, Basel, Switzerland) were used according to manufacturer's instructions. Transformation and amplification of plasmids was performed in *Escherichia coli* One Shot TOP10 competent cells (Invitrogen) according to manufacturer's instructions. Plasmids were isolated from *E. coli* with the Sigma GenElute Plasmid Miniprep Kit (Sigma). DNA constructs were routinely

sequenced by Baseclear BV (Baseclear, Leiden, The Netherlands).

Laboratory evolution of strains IMW004 and IMW005

Glycerol stocks of IMK302 were used to inoculate two 500-mL shake flasks containing 100 mL of synthetic medium (Verduyn *et al.*, 1992) with 5 g L^{-1} L-lactic acid (Fluka 09578) and 25 mM MES buffer. The pH of the medium was set to 5.0 with 2 M KOH and 2 M H_2SO_4 prior to autoclaving (121 °C, 20 min). When growth was observed, a small aliquot (0.1–2 mL) was transferred to a new flask. After 10 transfers, the culture was plated on agar plates containing 20 g L^{-1} glucose, 10 g L^{-1} Bacto yeast extract and 20 g L^{-1} Bacto peptone to obtain single colonies. Per evolution line, four colonies were tested for growth on lactate. One colony per evolution line was stocked and named as IMW004 and IMW005.

Cultivation and media

Shake-flask precultures for characterization experiments were inoculated with glycerol stocks and grown in synthetic medium (Verduyn *et al.*, 1992) with 20 g L^{-1} ethanol as carbon source.

Characterization of growth on lactate in shake flasks with an initial pH of 5.0 was performed in synthetic medium (Verduyn *et al.*, 1992) with 5 g L^{-1} lactate and 25 mM MES buffer, as described earlier. Cells growing exponentially on ethanol were used to inoculate 500-mL shake flasks containing 100 mL medium and incubated in an Innova incubator shaker (New Brunswick Scientific, Edison, NJ) at 200 rpm and 30 °C.

Aerobic batch cultures for transcriptome analysis were carried out at 30 °C in 2-L laboratory fermentors (Applikon, Schiedam, The Netherlands) with a working volume of 1 L. Synthetic medium (Verduyn *et al.*, 1992) was supplemented with 20 g L^{-1} lactic acid (Fluka 09578) as sole carbon source. The pH of this medium was set to 5.0 before autoclaving (120 °C). Antifoam Emulsion C (Sigma) was autoclaved separately (120 °C) as a 20% (w/v) solution and added to a final concentration of 0.2 g L^{-1}. The culture pH was maintained at 5.0 by automatic addition of 2 M KOH and 2 M H_2SO_4. Cultures were stirred at 800 rpm and sparged with 500 mL air per min. The precultures were grown on ethanol as described earlier. Six hours before the inoculation of the batch cultures, 25 mM lactate was added to the preculture to induce lactate metabolism. Samples for RNA extraction were taken during the exponential growth phase at a cell dry weight concentration of 3–4 g L^{-1} and a lactate concentration of 5–7 g L^{-1}. Sequential batch cultures for the characterization of growth rate and biomass yield were run in the same setup as described earlier, but with 10 g L^{-1} lactate to prevent possible nutrient limitations above 5 g L^{-1} cell dry weight. The first batch was inoculated with a preculture growing exponentially on ethanol, as described earlier. To increase reproducibility of the results, a sequential batch reactor (SBR) was used and the third cycle was sampled (Abbott *et al.*, 2009).

Growth assays on agar plates

Growth on ethanol, lactate and pyruvate was tested by spotting 10 μL of serial dilutions of a culture growing exponentially on ethanol, on synthetic medium agarose (1% w/v) plates supplemented with 25 mM MES and either 5 g L^{-1} ethanol, lactate or pyruvate. Before spotting, the cultures were first washed 2 times and diluted to the appropriate cell concentration with synthetic medium (Verduyn *et al.*, 1992). Pictures were taken after 6 days incubation at 30 °C.

Analytical methods

Optical density at 660 nm was measured with a Libra S11 spectrophotometer (Biochrom, Cambridge, UK). Culture dry weights were determined via filtration of appropriate sample volumes (10–20 mL) over dry preweighed nitrocellulose filters (Gelman Laboratory, Ann Arbor, MI) with a pore size of 0.45 μm. After removal of the medium, the filters were washed twice with demineralized water, dried in a microwave oven for 20 min at 350 W and weighed. Culture supernatants were obtained after the centrifugation of the broth. Supernatants and media were analysed via HPLC using an Aminex HPX-87H ion exchange column operated at 60 °C with 5 mM H_2SO_4 as mobile phase at a flow rate of 0.6 mL min^{-1}. Off-gas was first cooled in a condenser (2 °C) and dried with a Perma Pure Dryer (Permapure, Toms River, NJ). CO_2 concentrations in the off-gas were measured with a NGA 2000 Rosemount gas analyzer (Rosemount Analytical, Orrville, OH).

Microarrays and analysis

Sampling of cells from the batch cultures and total RNA extraction was performed as described previously (Abbott *et al.*, 2007). Probe preparation and hybridization to Affymetrix Genechip® microarrays were performed according to Affymetrix's instructions. The one-cycle eukaryotic target labelling assay was used, starting with 15 μg of total RNA. The quality of total RNA, cRNA and fragmented cRNA was checked using the Agilent BioAnalyzer 2100 (Agilent Technologies, Amstelveen, the

Netherlands). Results for each strain were obtained from two independent culture replicates. The Significance Analysis of Microarrays (SAM version 1.12; Tusher *et al.*, 2001) add-in to Microsoft Excel was used for comparison of replicate array experiments using a fold-change threshold of two and an expected false discovery rate of 5%. Transcript data have been deposited in the Genome Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE31634.

Whole-genome sequencing and analysis

DNA of the reference strain CEN.PK113-7D was prepared as described previously (Burke *et al.*, 2000). A library of 200-bp fragments was created and sequenced paired-end using the Illumina Solexa system, generating *c.* 56 million 36-bp paired reads. These reads were mapped to the genome of *S. cerevisiae* strain S288C (SGD project) using the Burrows–Wheeler Alignment tool (BWA; Li & Durbin, 2009). Subsequently, a consensus sequence was generated using SAMtools' pileup2fq (Li *et al.*, 2009) with default parameters, except for the minimum depth and root-mean-square mapping quality, which were set to 10 and 25, respectively. Consensus bases not passing the filtering thresholds, mostly in repetitive regions, were replaced by 'N'. This consensus was used as CEN.PK113-7D reference genome.

Genomic DNA from IMW004 and IMW005 was isolated using the Qiagen 100/G kit (Qiagen, Hilden, Germany). A library of 200-bp genomic fragments was created and sequenced paired-end (50-bp reads) using an Illumina HiSeq 2000 sequencer by Baseclear BV (Baseclear). The individual reads were mapped onto the genome of CEN.PK113-7D, using BWA (Li & Durbin, 2009) and further processed using SAMtools (Li *et al.*, 2009). Single-nucleotide variations and small insertions and deletions were extracted from the mapping using SAMtools' varFilter. Default settings were used, except that the maximum read depth was set to 300× (−D300) and the minimum small insertions and deletions score for single-nucleotide variations filtering was set to 60 (−G60). To minimize false positive mutation calls, custom Perl scripts and Microsoft Excel were used for further mutation filtering. First, mutation calls containing ambiguous bases in either reference or mapping consensus were filtered out. Second, only the single-nucleotide variations with a quality of at least 20 and small insertions and deletions with a quality of at least 50 were kept. Variant quality is defined as the Phred-scaled probability that the mutation call is incorrect (Ewing & Green, 1998). Third, mutations with a depth of coverage < 10× were discarded. Fourth, insertion and deletion mutation calls were only kept when at least 85% of the reads spanning the location confirmed the insertion or deletion. Fifth, small insertions and deletions that were close to an 'N' in the reference were removed, because this would complicate correct alignments and introduce false positive mutation calls. For small insertions and deletions, a window size of 40 bp was used, and for single-nucleotide variations, a window size of 5 bp. Copy number variation was analysed using CNV-seq (Xie & Tammi, 2009).

Pulsed-field gel electrophoresis and southern blotting

Chromosomes were separated using the CHEF yeast genomic DNA Plug Kit (170-3593; Bio-Rad, Richmond, CA) according to manufacturer's instructions and subsequently transferred onto Hybond-N^+ nylon membranes (RPN303; Amersham Biosciences, Piscataway, NJ). Southern blotting, signal generation and signal detection were performed using the Gene Images AlkPhos Kit, CPD Star detection reagent and Hyperfilm ECL (RPN 3680, RPN3682 and 28-9068; Amersham Biosciences). The DNA probe used for southern blotting was amplified from CEN.PK113-7D genomic DNA with primers ADY2 inside Fw and ADY2 inside Rv (Table 3).

Results

Laboratory evolution of Jen1p-independent growth on lactate

To test whether laboratory evolution and subsequent analysis of the underlying mutations is a powerful strategy to identify mutants that can transport lactate in *S. cerevisiae*, a *jen1*Δ knockout strain was constructed and evolved for growth on lactate. Hitherto, Jen1p is the only efficient importer of lactate in *S. cerevisiae* described in literature (Casal *et al.*, 1999). Consistent with earlier reports, *S. cerevisiae* IMK302 (*jen1*Δ) did not show growth on plates with lactate after 6 days, whereas the reference strain CEN.PK113-7D (*JEN1*) grew normally (Fig. 1). In shake-flask cultures with 5 g L^{-1} lactate at pH 5, the same results were obtained [$\mu < 0.001$ h^{-1} for IMK302 (*jen1*Δ)], but after prolonged incubation (5–15 days), growth was observed in two independent experiments. To select for faster growth, small aliquots of the shake-flask cultures were transferred to fresh medium for nine consecutive times. After *c.* 100 generations in 10 shake-flask cultures, single-colony isolates were obtained by plating on nonselective agar plates containing glucose, yeast extract and peptone. Per evolution line, four single-colony isolates were tested for growth on lactate and yielded essentially the same growth rates as the mixed population from which they were isolated. One single-

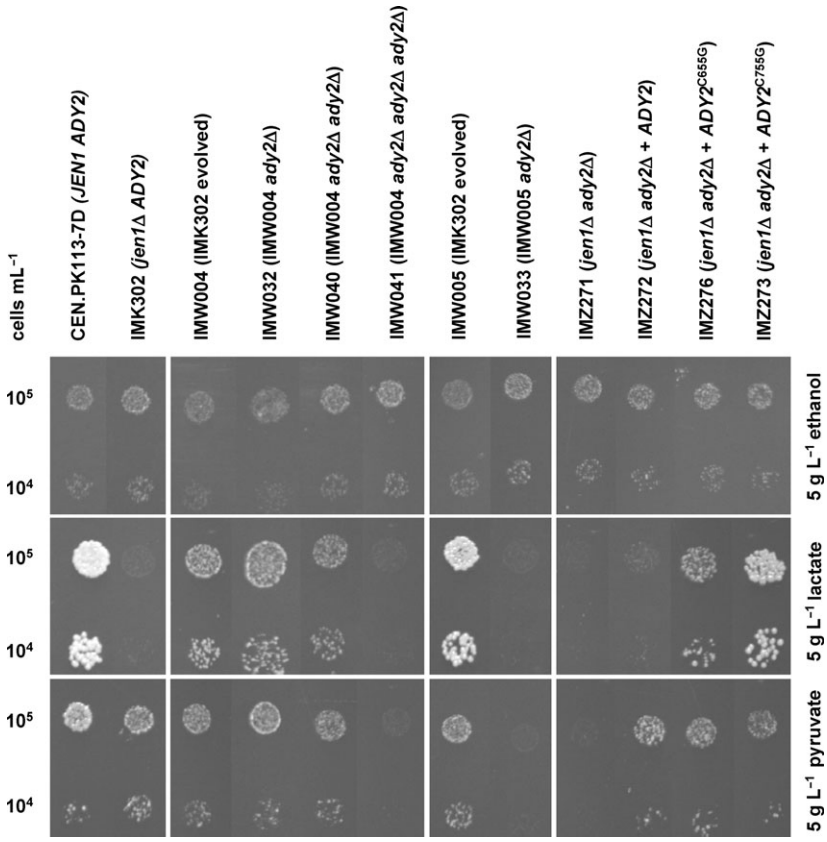

Fig. 1. Growth of *Saccharomyces cerevisiae* strains with different combinations of *JEN1* and *ADY2* alleles on ethanol, lactate and pyruvate. 10 μL of a suspension with the indicated cell concentration was spotted onto synthetic medium agarose (1% w/v) plates at pH 5 with 5 g L^{-1} ethanol (upper panel), 5 g L^{-1} lactate (middle panel) or 5 g L^{-1} pyruvate (lower panel) as the sole carbon source. Pictures were taken after 6 days of incubation at 30 °C.

colony isolate from each evolution experiment was chosen for further characterization. These resulting evolved strains were named IMW004 and IMW005 and grew at maximum specific growth rates of 0.14 and 0.18 h^{-1}, respectively, on 20 g L^{-1} lactate in aerobic bioreactor batch cultures (pH 5.0).

Mutation analysis through whole-genome DNA sequencing and transcriptome analysis

Two techniques were used to study the molecular basis of Jen1p-independent growth on lactate in the evolved strains IMW004 and IMW005: transcriptome analysis with DNA microarrays and whole-genome DNA resequencing. In view of highly reproducible growth conditions and the ability to control the specific growth rate, chemostat cultures are normally the preferred cultivation type for comparative transcriptome analysis (Daran-Lapujade *et al.*, 2008). However, chemostat cultures growing on lactate demonstrated persistent oscillations that precluded steady-state analysis (data not shown). Therefore, cell samples for transcriptome analysis were obtained from aerobic bioreactor batch cultures grown on 20 g L^{-1} lactate and at pH 5.0. To minimize differences in nutrient concentration and growth conditions, all cultures were sampled at a biomass dry weight concentration between 3 and 4 g L^{-1}, corresponding to 5–7 g L^{-1} residual lactate. The average deviation of the mean of all genes with mRNA expression levels above background (in this case 12) was between 16% and 21%, which was only slightly higher than chemostat-based transcriptome analysis (10–15%; Basso *et al.*, 2011). When comparing the transcriptome data of CEN.PK113-7D (*JEN1*) with IMW004 and IMW005 (both *jen1*Δ evolved), only a very small number of genes (11 for IMW004 and 18 for IMW005) were expressed at different levels based on the statistical criteria applied in this study (absolute fold difference $\geq$ 2; false discovery rate 5%, see Materials and methods; Fig. 2). In addition to *JEN1* (deleted in the evolved strain) and the introduced *KANMX4* dominant marker gene, the general amino acid permease *GAP1* was the only gene that was differentially expressed in both evolution lines IMW004 and IMW005. However, the transcript levels of *GAP1* were only 2.6- and 2.0-fold higher in IMW004 and IMW005, respectively, than in the reference strain CEN.PK113-7D.

In parallel, genomic DNA of IMW004 and IMW005 (both *jen1*Δ evolved) was sequenced using Illumina technology, yielding *c.* 12.8 million 50-bp reads for IMW004 and *c.* 13.7 million reads for IMW005. The average depth

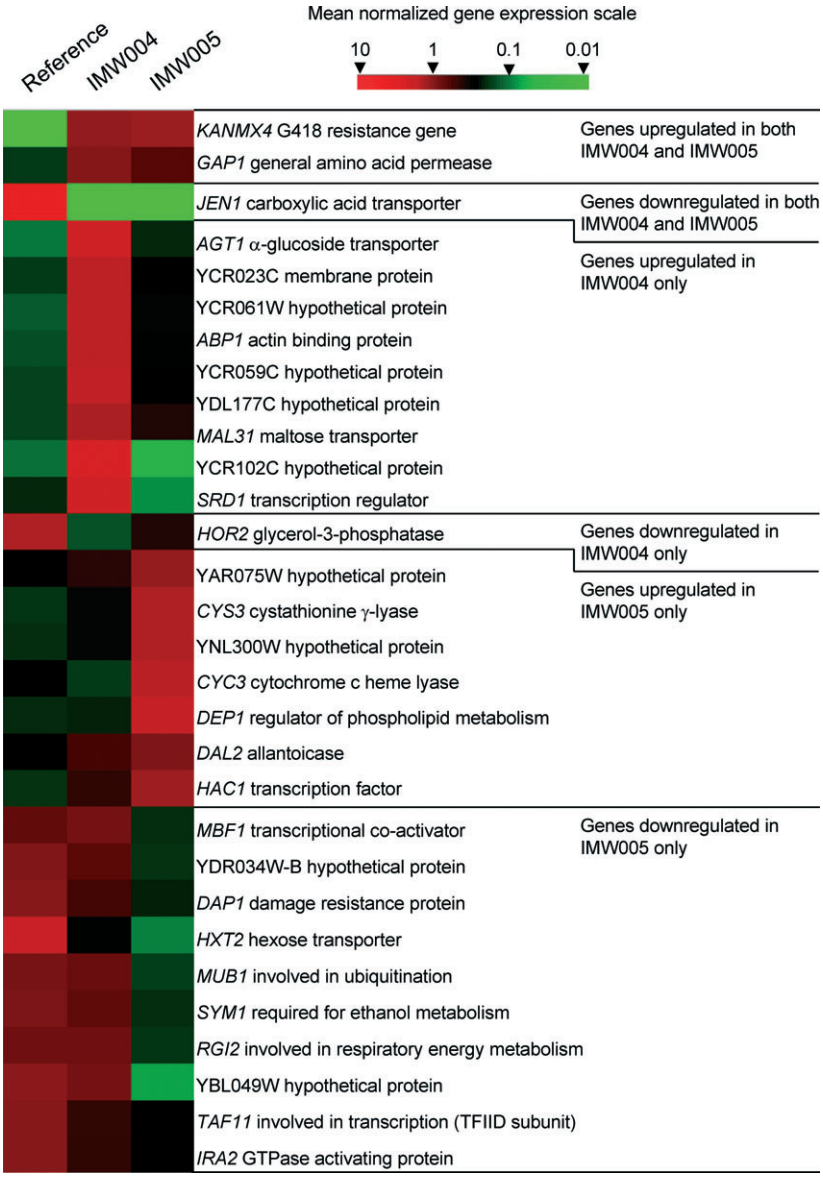

Fig. 2. Differentially expressed genes in IMW004 and IMW005 (both *jen1*Δ evolved) compared with the reference strain CEN. PK113-7D (*JEN1*) during exponential growth on 20 g L^{-1} lactate at pH 5.0. Expression data were compared using a fold-change threshold of two and an estimated false discovery rate of 5%.

of coverage was 51.5-fold for IMW004 and 56.6-fold for IMW005. The sequencing reads were mapped against the genome sequence of the reference strain CEN.PK113-7D, which was constructed by mapping CEN.PK113-7D sequencing reads to the S288C sequence. This resulted in 20 and 16 single-nucleotide variations for IMW004 and IMW005, respectively. No insertions or deletions were found. The set of relevant mutations was reduced to 7 for IMW004 and 10 for IMW005 (Table 4) by eliminating single-nucleotide variations outside coding regions, because no links were found with mRNA expression levels of adjacent genes. Nucleotide changes that only result in different codon usage for the same amino acid might affect translation efficiency and were therefore not discarded. Interestingly, several of the remaining single-nucleotide variations were either identical in both IMW004 and IMW005 or different mutations had occurred in the same genes (Table 4). Such mutations might (1) present genes that are relevant for the Jen1p-independent growth on lactate, (2) have accumulated during the construction of the *jen1*Δ strain and have been present before the start of the evolution experiment or (3) represent errors in the reference sequence.

Rearrangement of chromosome III

Rearrangement of larger DNA fragments, including deletions and duplications, is a well-known phenomenon dur-

Table 4. Mutation calls inside genes in IMW004 and IMW005 identified via whole-genome sequencing

Gene	Description	Nucleotide change	Amino acid change
Genes mutated in both IMW004 and IMW005			
YAT1	Mitochondrial carnitine acetyltransferase	C75A	None
ADY2	Acetate transporter	C655G (IMW005)	Leu219Val (IMW005)
ADY2	Acetate transporter	C755G (IMW004)	Ala252Gly (IMW004)
BSC1	Hypothetical protein	T553A	Thr185Ser
FLO11	Glycoprotein (flocculin)	G1716A (IMW005)	None
FLO11	Glycoprotein (flocculin)	C1743T (IMW005)	None
FLO11	Glycoprotein (flocculin)	G2361A	None
DDC1	DNA damage checkpoint protein	G1513C	Ala505Pro
Genes mutated only in IMW004			
YNL179C	Hypothetical protein	T156A	Leu52Phe
CAF120	Part of transcriptional regulatory complex	G1689A	None
Genes mutated only in IMW005			
FIT1	Cell wall mannoprotein	A1067G	Val356Ala
PAU2	Member of seripauperin multigene family	T318C	None
YHR028W-A	Hypothetical protein	T217C	Ser73Phe

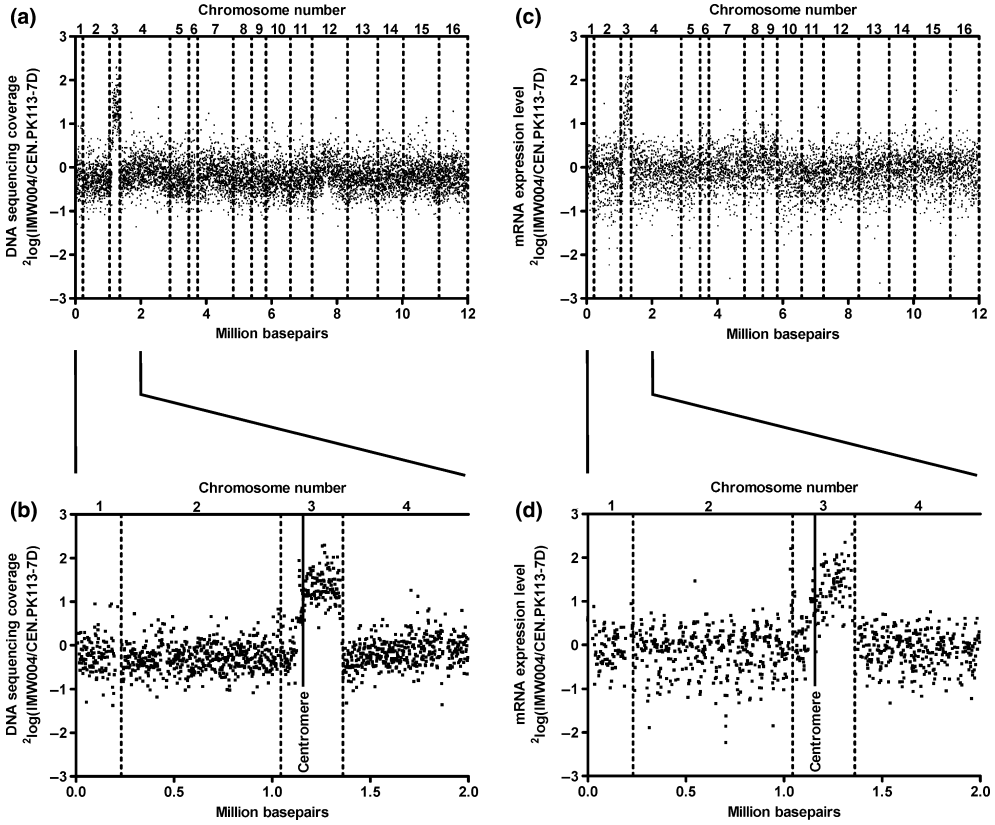

Fig. 3. Comparison of depth of coverage of DNA sequencing (a, b) and mRNA expression levels (c, d) of CEN.PK113-7D (*JEN1*) and IMW004 (*jen1*Δ evolved). Values were plotted according to their chromosomal location. Chromosome numbers are indicated above the panels. Depth of coverage of DNA sequencing was averaged and plotted per 1500 bp. For mRNA expression levels, only genes with an expression level above 12 were considered.

ing (yeast) evolution (Wolfe & Shields, 1997; Dunham *et al.*, 2002; Wisselink *et al.*, 2010; Dhar *et al.*, 2011). To analyse these larger structural variations, the genome-wide mRNA expression levels were plotted according to their chromosomal location. Interestingly, both the mRNA expression levels and depth of coverage of DNA sequenc-

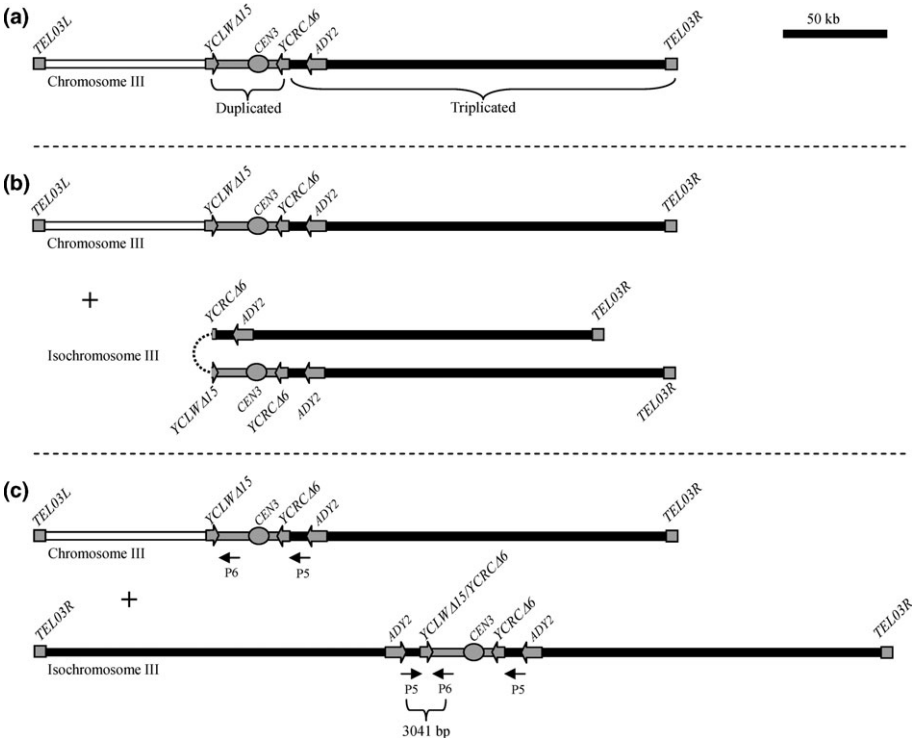

Fig. 4. Analysis of chromosomal rearrangements in IMW004. (a) Copy number estimates on chromosome III based on depth of coverage of DNA sequencing and mRNA expression levels. (b) Crossover between homologous regions of *YCLWΔ15* and *YCRCΔ6*, (c) resulting in the formation in an additional pseudopalindromic isochromosome III, containing two extra copies of *ADY2*^{C755G}. Indicated are the binding sites of primers P5 and P6, which were used to confirm that crossover took place as proposed. Duplicated parts of chromosome III are indicated in grey; triplicated regions in black. Relevant features are indicated, but not drawn to scale.

ing showed ^{2}log ratios between 1 and 2 for a group of genes on chromosome III, when comparing IMW004 and CEN.PK113-7D (Fig. 3). Detailed investigation of the depth of coverage of DNA sequencing indicated a duplication of chromosome III between 83 and 126 kb, including the centromere, and a triplication between 126 kb to the end of the right arm (Fig. 4a). These observations led to the hypothesis that, in addition to chromosome III, an additional pseudopalindromic isochromosome III was formed containing the centromere and two copies of the right arm of chromosome III. Karyotyping confirmed the formation of an additional chromosome with a size of *c.* 475 kb (Fig. 5), which corresponds well to the proposed rearrangement (Fig. 4).

To exactly determine the chromosomal rearrangement, the DNA sequences 5 kb upstream and downstream of the observed breakpoints (i.e. 78–88 and 121–131 kb) were aligned with each other. The highest homology was observed between the Ty1-type long terminal repeat retrotransposons *YCLWΔ15* and *YCRCΔ6*, with 96% homology over a region of 176 bp. To verify whether crossover took place between *YCLWΔ15* and *YCRCΔ6* as proposed in Fig. 4, primer combinations were designed (P5 and P6) that should give a 3041-bp PCR product if crossover had indeed taken place between *YCLWΔ15* and *YCRCΔ6*. In line with the proposed mechanism of rearrangement, the expected 3-kb PCR product was obtained from IMW004, but not from CEN.PK113-7D genomic DNA. Sequencing of this DNA fragment showed that crossover took place between the last 98 homologous base pairs of *YCLWΔ15* and *YCRCΔ6*.

Verification of the identified mutations in the acetate transporter *ADY2*

The *ADY2* gene, which encodes an acetate transporter (Paiva *et al.*, 2004), contained mutations that resulted in amino acid changes in both independently evolved strains (Table 4) and is located on the triplicated region of chromosome III of IMW004. To analyse whether the different copies of *ADY2* in IMW004 all contained the same mutation, the individual sequencing reads mapping to *ADY2* were analysed. At all positions, a constant signal was observed, showing that IMW004 carried three identical copies of *ADY2*^{C755G}. *ADY2* was not amongst the genes whose mRNA levels were significantly different from

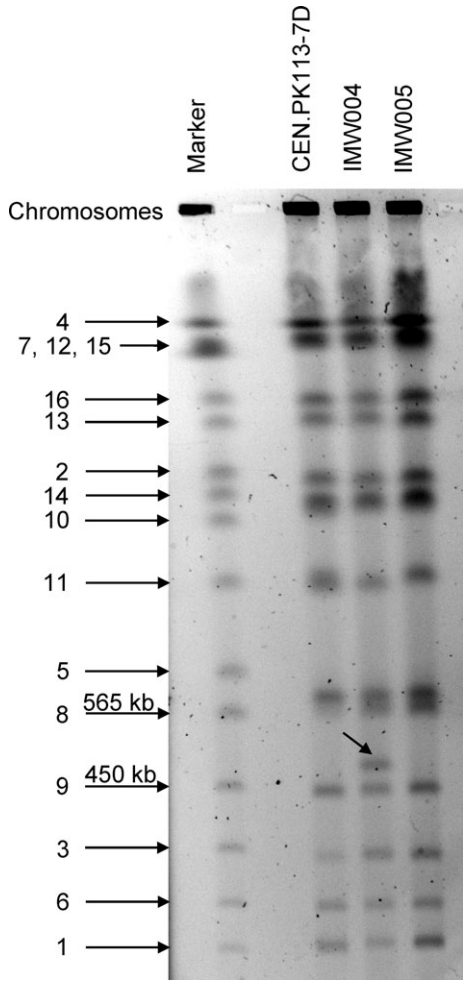

Fig. 5. Pulsed-field gel electrophoresis (PFGE) analysis of the reference strain CEN.PK113-7D and the evolved strains IMW004 and IMW005. Chromosome numbers and sizes are indicated on the left side. YNN295 (Bio-Rad #170-3605) was used as marker. The additional chromosome (c. 475 kb) present in IMW004 is indicated by the diagonal arrow.

those in the reference strain (1.4-fold in IMW004 and 0.6-fold IMW005).

To investigate the physiological relevance of the mutation (C755G; Ala252Gly) and the apparent triplication of *ADY2* in IMW004, knockout strains were constructed. Diagnostic PCR on genomic DNA of the single, double and triple knockout strains (IMW032, IMW040 and IMW041) only showed absence of *ADY2* in the triple knockout strain (IMW041), thereby confirming that IMW004 contained three copies of $ADY2^{C755G}$. Only the triple knockout strain was unable to grow on lactate (Fig. 1), confirming that $Ady2p^{Ala252Gly}$ was essential for lactate import in the evolved strain IMW004.

To analyse the effect of $ADY2^{C755G}$ gene dosage on the growth rate on lactate, aerobic shake-flask experiments were performed with 5 g L^{-1} lactate at pH 5. IMW004 (3 copies of $ADY2^{C755G}$) grew at a growth rate of 0.10 ± 0.01 h^{-1}; IMW032 (2 copies) at 0.09 ± 0.00 h^{-1}, IMW040 (1 copy) at 0.07 ± 0.00 h^{-1} and IMW041 (0 copies) did not grow ($\mu < 0.001$ h^{-1}).

To analyse whether the identified mutation (C655G; Leu219Val) in *ADY2* could also fully explain the acquired phenotype in IMW005, a knockout of $ADY2^{C655G}$ was constructed. IMW033 (IMW005 *ady2*Δ) was not able to grow on lactate, thereby confirming the crucial role of $Ady2p^{Leu219Val}$ in lactate import in IMW005 (Fig. 1).

Reverse metabolic engineering of mutated *ADY2* alleles

To verify whether mutation of *ADY2* alone is sufficient to allow for growth on lactate, the mutated alleles were introduced into a nonevolved background (*jen1*Δ *ady2*Δ). IMZ271 (*jen1*Δ *ady2*Δ) and IMZ272 (*jen1*Δ *ady2*Δ + *ADY2*) did not show growth on agar plates with 5 g L^{-1} lactate after 6 days of incubation, whereas IMZ273 (*jen1*Δ *ady2*Δ + $ADY2^{C755G}$) and IMZ276 (*jen1*Δ *ady2*Δ + $ADY2^{C655G}$) did grow on these plates (Fig. 1). This confirmed that these alleles can confer the ability to transport lactate. IMZ271 (*jen1*Δ *ady2*Δ) could still grow on acetate (data not shown), which prevented kinetic analysis of acetate transport mediated by the mutated Ady2p alleles. Interestingly, IMZ271 (*jen1*Δ *ady2*Δ) could not grow on pyruvate, whereas IMZ272 (*jen1*Δ *ady2*Δ + *ADY2)*, IMZ273 (*jen1*Δ *ady2*Δ + $ADY2^{C755G}$) and IMZ276 (*jen1*Δ *ady2*Δ + $ADY2^{C655G}$) could, showing that both the nonmutated and the mutated Ady2p alleles can function as efficient pyruvate importers in this strain background (Fig. 1), in contrast to previous reports on other *S. cerevisiae* strains (Akita *et al.*, 2000).

Subsequently, two strains expressing the native *ADY2* allele were compared. IMK302 (*jen1*Δ *ADY2*), which contains the native *ADY2* gene on chromosome III, did not grow on lactate ($\mu < 0.001$ h^{-1}) in shake-flask cultures with 5 g L^{-1} lactate at pH 5. The fact that IMZ272 (*jen1*Δ *ady2*Δ + *ADY2*), containing the native *ADY2* under control of its own promoter on a centromeric vector, grew very slowly on lactate (μ_{max} around 0.02 h^{-1}) demonstrates that even nonmutated Ady2p can transport lactate, although at a very low rate. The observed differences in growth may be due to small differences in expression level between chromosomal (IMK302) and centromeric (IMZ272) expression. This might raise the lactate consumption rate in IMZ272 marginally above the flux required for maintenance metabolism and thus allow for very slow growth (Pirt, 1975; Boender *et al.*, 2009).

To further characterize the mutated alleles of *ADY2*, the reverse engineered strains IMZ273 (*jen1*Δ *ady2*Δ

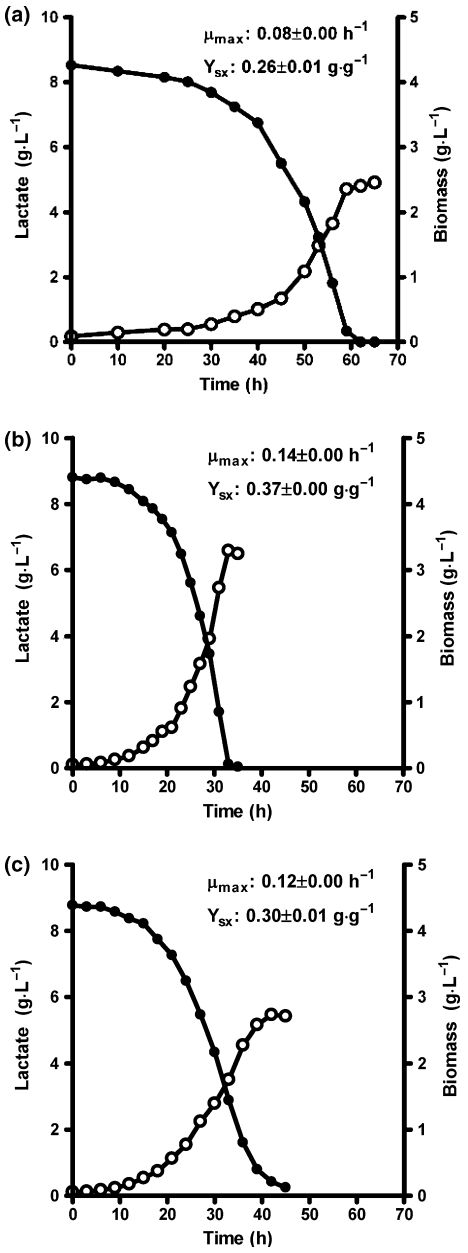

Fig. 6. Growth of CEN.PK 113-7D (*JEN1 ADY2*, a), IMZ273 (*jen1 ady2* + *ADY2*^{C655G}, b) and IMZ276 (*jen1 ady2* + *ADY2*^{C755G}, c) during the third cycle of an aerobic sequential batch fermentation on lactate at pH 5.0. The results shown are from one representative experiment. Averages and mean deviations for the maximum specific growth rates (μ_{max}) and biomass yields (Y_{sx}) were obtained from independent duplicate experiments.

+ *ADY2*^{C755G}) and IMZ276 (*jen1*Δ *ady2*Δ + *ADY2*^{C655G}) were tested in aerobic sequential batch fermentations with 10 g L^{-1} lactate at pH 5.0. The reference strain CEN.PK113-7D (*JEN1 ADY2*) was used for comparison, because in this strain, the large majority of lactate transport proceeds via Jen1p and not via Ady2p (Fig. 1). Both IMZ273 (*jen1*Δ *ady2*Δ + *ADY2*^{C655G}) and IMZ276 (*jen1*Δ *ady2*Δ + *ADY2*^{C755G}) grew rapidly on lactate at maximum specific growth rates of 0.14 ± 0.00 and 0.12 ± 0.00 h^{-1}, respectively. In cultures of strain IMZ276, the Ady2p^{Ala252Gly}-dependent lactate consumption rate decreased towards the end of the fermentation, indicating a low affinity for lactate (Fig. 6). The biomass yield on lactate of IMZ273 (0.37 ± 0.00 g g^{-1}; *jen1*Δ *ady2*Δ + *ADY2*^{C655G}) was considerably higher than those of IMZ276 (0.30 ± 0.01 g g^{-1}; *jen1*Δ *ady2*Δ + *ADY2*^{C755G}) and CEN.PK113-7D (0.26 ± 0.01 g g^{-1}; *JEN1 ADY2*).

Discussion

Ady2p as lactate transporter

No protein structure of Ady2p or close homologues is available, and therefore, the effect of the mutations on the structure and function of Ady2p remains uncertain. However, the observation that single-nucleotide mutations turned the acetate transporter Ady2p into an efficient lactate transporter provides valuable information for future structure-function analysis. In both evolution lines, the identified mutations were located in putative transmembrane domains and resulted in amino acids that are one carbon atom smaller than the original amino acid (alanine to glycine in IMW004 and leucine to valine in IMW005). As lactate ($C_3H_6O_3$) is structurally similar, but one carbon atom larger than acetate ($C_2H_4O_2$), this suggests that extra space is created in the translocation site of Ady2p.

According to the literature, Jen1p catalyses an electroneutral lactate-proton symport, which is energetically equivalent to diffusion of the undissociated acid (Cassio *et al.*, 1987; Casal *et al.*, 1999). Complete aerobic dissimilation of 1 lactate then yields 7 ATP: 6 from oxidative phosphorylation at an *in vivo* P/O ratio of 1.0 (Van Gulik & Heijnen, 1995; Bakker *et al.*, 2001) and 1 from substrate level phosphorylation in the tricarboxylic-acid cycle. To generate and maintain a proton-motive force, the plasma membrane H^+-ATPase in yeast exports (only) a single proton for each ATP that it hydrolyses (van Leeuwen *et al.*, 1992; Weusthuis *et al.*, 1993). Lactate transport mechanisms that result in outward translocation of a proton and/or inward translocation of a negative charge can save ATP equivalents that would otherwise be used for the build-up or maintenance of the proton-motive force. These savings can therefore not exceed the *in vivo* activity of the plasma membrane H^+-ATPase. In this way, lactate import via a lactate anion uniport or via a lactate-proton antiport mechanism could increase the apparent ATP yield from lactate dissimilation to 8 or 9 mol mol^{-1}, corresponding to increases

in the biomass yield on lactate of 14% and 29%, respectively. The reverse engineered strain IMZ273 ($Ady2p^{Leu219Val}$) displayed a 24 ± 5% higher biomass yield on lactate than IMZ276 ($Ady2p^{Ala252Gly}$) and a 41 ± 3% higher biomass yield than the reference strain CEN.PK113-7D (Fig. 6), in which lactate import primarily proceeds via Jen1p (Fig. 1). Taking into account that an additional *c.* 5% of the biomass yield increase can be explained owing to higher specific growth rates of the evolved strains, which causes a lower impact of maintenance-energy requirements (Pirt, 1975; Boender *et al.*, 2009), these biomass yield values might be consistent with a different mode of energy coupling of lactate transport. Further analysis of the energy coupling of lactate transport by these Ady2p isoforms, involving *in vitro* studies, will provide valuable insights into the energetics of organic acid transport in yeast and its impact on growth energetics.

Chromosomal rearrangement

Chromosomal rearrangement is a well-known phenomenon during (yeast) evolution (Wolfe & Shields, 1997; Dhar *et al.*, 2011) and occurs often by recombination of Ty1-type retrotransposons (reviewed by Garfinkel, 2005). The occurrence of a *c.* 475-kb isocytochrome III in strain IMW004 shows how even a short (*c.* 100 generations) period of selective growth can select for a massive chromosomal rearrangement, in this case resulting in duplication of around 25 genes and triplication of about 100 genes. In addition to enhancing performance under the conditions that led to their selective advantage, such rearrangements may be 'hopeful monsters' that open up novel evolutionary paths via specialization of duplicated genes (Theißen, 2009).

In-depth analysis of the sequencing reads showed that all three copies of *ADY2* in IMW004 contained the same nucleotide change (C755G), which suggests that the mutation in *ADY2* occurred before the triplication event. This indicates that first a gain of function mutation in *ADY2* provided the cells with a selective advantage over the parental strain. The subsequent retrotransposon-mediated triplication of $ADY2^{C755G}$ apparently gave sufficient evolutionary advantage to compensate for the burden of an additional chromosome. In line with this hypothesis, a strain derived from IMW004 containing only one copy of $ADY2^{C755G}$ (IMW040) showed a 30% lower growth rate on lactate. The observation that the second and third copy of $ADY2^{C755G}$ did not lead to a two- and threefold higher growth rates can be explained by the saturation of transcription, possibly due to a limited availability of transcription factors that regulate *ADY2* expression, such as Cat8p, Adr1p and Snf1p (Casal *et al.*, 2008). Indeed, the *ADY2* mRNA expression level of IMW004 (three $ADY2^{C755G}$ copies) was only 40% higher than that of CEN.PK113-7D (one *ADY2* copy). Simulation of the evolution experiment, based on the observed cell densities and the relatively small difference in growth rate between strains expressing one ($0.07\ h^{-1}$) or three ($0.10\ h^{-1}$) $ADY2^{C755G}$ copies, indicated that the formation of the isochromosome III occurred early in the evolution experiment before the first transfer. Accordingly, no further increase in the growth rate on lactate was observed in the remaining nine shake flasks of the evolution experiment.

Functional analysis through laboratory evolution and reverse engineering

Laboratory evolution combined with molecular analysis of the evolved strains enabled the successful identification in *S. cerevisiae* of genes encoding transporter mutants that can import lactate. Although this strategy can only be used in cases where the gene function can be linked to a selective advantage, it can be very useful in situations where only a low residual enzyme activity is observed for which the genetic determinants are unknown, such as for glucose transport or disaccharide hydrolysis in mutant strains of *S. cerevisiae* in which all known genes for these activities have been deleted (Wieczorke *et al.*, 1999; De Kok *et al.*, 2011). In addition, this approach can contribute genes and proteins with new, interesting properties for metabolic engineering. Elucidation of the genetic basis of the acquired phenotype is a crucial step in laboratory evolution-based functional analysis, which in this study was approached by a combination of transcriptome analysis and whole-genome resequencing. Subsequent verification via knockout studies and reverse metabolic engineering showed that the relevant changes could be identified via whole-genome resequencing, but not via transcriptome analysis. In yeast, only few studies have been performed to identify the relevant genetic changes and reverse engineer them into nonevolved strain background. When studies on both prokaryotic and eukaryotic laboratory evolved strains are considered, a clear picture emerges that, in many studies, transcriptome analysis did not allow rapid identification of relevant genetic changes that had been selected for in laboratory evolution experiments (Ferea *et al.*, 1999; Riehle *et al.*, 2003; van Maris *et al.*, 2004c; Fong *et al.*, 2005, 2006; Hu & Wood, 2010), even when combined with metabolome analysis (Vijayendran *et al.*, 2008) or proteome analysis (Lewis *et al.*, 2010). In contrast, studies in which the genetic basis for the acquired phenotype was elucidated and reconstructed via reverse metabolic engineering analysed

the DNA level via whole-genome resequencing or hybridization of genomic DNA to microarrays (Herring *et al.*, 2006; Gresham *et al.*, 2008b; Conrad *et al.*, 2009; Atsumi *et al.*, 2010; Lee & Palsson, 2010; Tremblay *et al.*, 2010; Hong *et al.*, 2011). These observations support the view that analysis of the genome sequence rather than the transcriptome is the most promising first-line analytical approach for elucidating relevant changes that occurred during evolution. Transcriptome analysis can then be used as a follow-up in cases where whole-genome resequencing does not generate obvious and productive leads. In this respect, mRNA sequencing is an interesting development that allows simultaneous measurement of mRNA levels and identification of mutations inside coding regions (Wang *et al.*, 2009; Oshlack *et al.*, 2010). Independent of the technique used, analysing multiple independently evolved strains strongly facilitates dissection of relevant and random genetic changes.

Acknowledgements

We thank Marinka Almering, Mark Bisschops, Lizanne Bosman, Daniel Gonzalez Ramos, Marit Hebly and Duygu Yilmaz for experimental assistance. This work was financially supported by Tate & Lyle Ingredients Americas Inc. The Kluyver Centre for Genomics of Industrial Fermentations is supported by the Netherlands Genomics Initiative.

References

Abbott DA, Knijnenburg TA, De Poorter LMI, Reinders MJT, Pronk JT & Van Maris AJA (2007) Generic and specific transcriptional responses to different weak organic acids in anaerobic chemostat cultures of *Saccharomyces cerevisiae*. *FEMS Yeast Res* **7**: 819–833.

Abbott DA, Suir E, Duong GH, De Hulster E, Pronk JT & van Maris AJA (2009) Catalase overexpression reduces lactic acid-induced oxidative stress in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* **75**: 2320–2325.

Akita O, Nishimori C, Shimamoto T, Fujii T & Iefuji H (2000) Transport of pyruvate in *Saccharomyces cerevisiae* and cloning of the gene encoded pyruvate permease. *Biosci Biotechnol Biochem* **64**: 980–984.

Atsumi S, Wu TY, Machado IMP, Huang WC, Chen PY, Pellegrini M & Liao JC (2010) Evolution, genomic analysis, and reconstruction of isobutanol tolerance in *Escherichia coli*. *Mol Syst Biol* **6**: 449.

Bailey JE, Sburlati A, Hatzimanikatis V, Lee K, Renner WA & Tsai PS (1996) Inverse metabolic engineering: a strategy for directed genetic engineering of useful phenotypes. *Biotechnol Bioeng* **52**: 109–121.

Bakker BM, Overkamp KM, Maris AJA, Kötter P, Luttik MAH, Dijken JP & Pronk JT (2001) Stoichiometry and compartmentation of NADH metabolism in *Saccharomyces cerevisiae*. *FEMS Microbiol Rev* **25**: 15–37.

Basso TO, de Kok S, Dario M *et al.* (2011) Engineering topology and kinetics of sucrose metabolism in *Saccharomyces cerevisiae* for improved ethanol yield. *Metab Eng* **13**: 694–703.

Boender LGM, de Hulster EAF, van Maris AJA, Daran-Lapujade PAS & Pronk JT (2009) Quantitative physiology of *Saccharomyces cerevisiae* at near-zero specific growth rates. *Appl Environ Microbiol* **75**: 5607–5614.

Bro C & Nielsen J (2004) Impact of 'ome' analyses on inverse metabolic engineering. *Metab Eng* **6**: 204–211.

Burke D, Dawson D & Stearns T (2000) *Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual*. CSHL Press, New York, USA.

Casal M, Paiva S, Andrade RP, Gancedo C & Leao C (1999) The lactate-proton symport of *Saccharomyces cerevisiae* is encoded by *JEN1*. *J Bacteriol* **181**: 2620–2623.

Casal M, Paiva S, Queirós O & Soares Silva I (2008) Transport of carboxylic acids in yeasts. *FEMS Microbiol Rev* **32**: 974–994.

Cassio F, Leao C & Van Uden N (1987) Transport of lactate and other short-chain monocarboxylates in the yeast *Saccharomyces cerevisiae*. *Appl Environ Microbiol* **53**: 509–513.

Conrad TM, Joyce AR, Applebee MK, Barrett CL, Xie B, Gao Y & Palsson B (2009) Whole-genome resequencing of *Escherichia coli* K-12 MG1655 undergoing short-term laboratory evolution in lactate minimal media reveals flexible selection of adaptive mutations. *Genome Biol* **10**: R118.

Daran Lapujade P, Daran JM, Kötter P, Petit T, Piper MDW & Pronk JT (2003) Comparative genotyping of the *Saccharomyces cerevisiae* laboratory strains S288C and CEN. PK113 7D using oligonucleotide microarrays. *FEMS Yeast Res* **4**: 259–269.

Daran-Lapujade P, Daran JM, van Maris AJA, de Winde JH & Pronk JT (2008) Chemostat-based micro-array analysis in baker's yeast. *Adv Microb Physiol*, **54**: 257–311, 414–417.

De Kok S, Yilmaz D, Suir E, Pronk JT, Daran JM & Van Maris AJA (2011) Increasing free-energy (ATP) conservation in maltose-grown *Saccharomyces cerevisiae* by expression of a heterologous maltose phosphorylase. *Metab Eng* **13**: 518–526.

Dhar R, Sägesser R, Weikert C, Yuan J & Wagner A (2011) Adaptation of *Saccharomyces cerevisiae* to saline stress through laboratory evolution. *J Evol Biol* **24**: 1135–1153.

Dunham MJ, Badrane H, Ferea T, Adams J, Brown PO, Rosenzweig F & Botstein D (2002) Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae*. *P Natl Acad Sci USA* **99**: 16144–16149.

Entian K & Kotter P (2007) Yeast genetic strain and plasmid collections. *Yeast Gene Analysis* (Stansfield I & StarkMJR, eds). Academic Press Ltd., San Diego. *Methods Microbiol* **36**: 629–666.

Ewing B & Green P (1998) Base-calling of automated sequencer traces using *Phred*. II. Error probabilities. *Genome Res* **8**: 186–194.

Ferea TL, Botstein D, Brown PO & Rosenzweig RF (1999) Systematic changes in gene expression patterns following adaptive evolution in yeast. *P Natl Acad Sci USA* **96**: 9721–9726.

Fong SS, Joyce AR & Palsson BØ (2005) Parallel adaptive evolution cultures of *Escherichia coli* lead to convergent growth phenotypes with different gene expression states. *Genome Res* **15**: 1365–1372.

Fong SS, Nanchen A, Palsson BO & Sauer U (2006) Latent pathway activation and increased pathway capacity enable *Escherichia coli* adaptation to loss of key metabolic enzymes. *J Biol Chem* **281**: 8024–8033.

Garfinkel D (2005) Genome evolution mediated by Ty elements in *Saccharomyces*. *Cytogenet Genome Res* **110**: 63–69.

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

Gill RT (2003) Enabling inverse metabolic engineering through genomics. *Curr Opin Biotechnol* **14**: 484–490.

Gresham D, Dunham MJ & Botstein D (2008a) Comparing whole genomes using DNA microarrays. *Nat Rev Genet* **9**: 291–302.

Gresham D, Desai MM, Tucker CM, Jenq HT, Pai DA, Ward A, DeSevo CG, Botstein D & Dunham MJ (2008b) The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. *PLoS Genet* **4**: e1000303.

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

Herring CD, Raghunathan A, Honisch C, Patel T, Applebee MK, Joyce AR, Albert TJ, Blattner FR, Van den Boom D & Cantor CR (2006) Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nat Genet* **38**: 1406–1412.

Hong KK, Vongsangnak W, Vemuri GN & Nielsen J (2011) Unravelling evolutionary strategies of yeast for improving galactose utilization through integrated systems level analysis. *P Natl Acad Sci USA* **108**: 12179–12184.

Hu H & Wood TK (2010) An evolved *Escherichia coli* strain for producing hydrogen and ethanol from glycerol. *Biochem Biophys Res Commun* **391**: 1033–1038.

Janke C, Magiera MM, Rathfelder N, Taxis C, Reber S, Maekawa H, Moreno Borchart A, Doenges G, Schwob E & Schiebel E (2004) A versatile toolbox for PCR based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast* **21**: 947–962.

Jol SJ, Kümmel A, Hatzimanikatis V, Beard DA & Heinemann M (2010) Thermodynamic calculations for biochemical transport and reaction processes in metabolic networks. *Biophys J* **99**: 3139–3144.

Lee DH & Palsson BO (2010) Adaptive evolution of *Escherichia coli* K-12 MG1655 during growth on a nonnative carbon source, L-1,2-propanediol. *Appl Environ Microbiol* **76**: 4158–4168.

Lewis NE, Hixson KK, Conrad TM, Lerman JA, Charusanti P, Polpitiya AD, Adkins JN, Schramm G, Purvine SO & Lopez-Ferrer D (2010) Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models. *Mol Syst Biol* **6**: 390.

Li H & Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**: 1754–1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G & Durbin R (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* **25**: 2078–2079.

Nowrousian M (2010) Next-generation sequencing techniques for eukaryotic microorganisms: sequencing-based solutions to biological problems. *Eukaryot Cell* **9**: 1300–1310.

Oshlack A, Robinson MD & Young MD (2010) From RNA-seq reads to differential expression results. *Genome Biol* **11**: 220.

Paiva S, Devaux F, Barbosa S, Jacq C & Casal M (2004) Ady2p is essential for the acetate permease activity in the yeast *Saccharomyces cerevisiae*. *Yeast* **21**: 201–210.

Palmqvist E & Hahn-Hägerdal B (2000) Fermentation of lignocellulosic hydrolysates. II: inhibitors and mechanisms of inhibition. *Bioresour Technol* **74**: 25–33.

Pirt SJ (1975) *Principles of Microbe and Cell Cultivation*. John Wiley & Sons, Inc, New York, USA.

Riehle MM, Bennett AF, Lenski RE & Long AD (2003) Evolutionary changes in heat-inducible gene expression in lines of *Escherichia coli* adapted to high temperature. *Physiol Genomics* **14**: 47–58.

Sauer U (2001) Evolutionary engineering of industrially important microbial phenotypes. *Adv Biochem Eng Biotechnol* **73**: 129–169.

Sboner A, Mu XJ, Greenbaum D, Auerbach RK & Gerstein MB (2011) The real cost of sequencing: higher than you think! *Genome Biol* **12**: 125.

Schacherer J, Ruderfer DM, Gresham D, Dolinski K, Botstein D & Kruglyak L (2007) Genome-wide analysis of nucleotide-level variation in commonly used *Saccharomyces cerevisiae* strains. *PLoS ONE* **2**: e322.

SGD project *Saccharomyces* genome database. Available at: http://www.yeastgenome.org/, accessed 21 July 2010.

Sikorski RS & Hieter P (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**: 19–27.

Theißen G (2009) Saltational evolution: hopeful monsters are here to stay. *Theory Biosci* **128**: 43–51.

Tremblay PL, Summers ZM, Glaven RH, Nevin KP, Zengler K, Barrett CL, Qiu Y, Palsson BO & Lovley DR (2010) A c-type cytochrome and a transcriptional regulator responsible for enhanced extracellular electron transfer in *Geobacter sulfurreducens* revealed by adaptive evolution. *Environ Microbiol* **13**: 13–23.

Tusher VG, Tibshirani R & Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. *P Natl Acad Sci USA* **98**: 5116–5121.

van Dijken J, Bauer J, Brambilla L, Duboc P, Francois J, Gancedo C, Giuseppin M, Heijnen J, Hoare M & Lange H (2000) An interlaboratory comparison of physiological and genetic properties of four *Saccharomyces cerevisiae* strains. *Enzyme Microb Technol* **26**: 706–714.

Van Gulik W & Heijnen J (1995) A metabolic network stoichiometry analysis of microbial growth and product formation. *Biotechnol Bioeng* **48**: 681–698.

van Leeuwen CC, Weusthuis RA, Postma E, van den Broek PJ & van Dijken JP (1992) Maltose/proton co-transport in *Saccharomyces cerevisiae*. Comparative study with cells and plasma membrane vesicles. *Biochem J* **284**: 441–445.

van Maris AJA, Konings WN, van Dijken JP & Pronk JT (2004a) Microbial export of lactic and 3-hydroxypropanoic acid: implications for industrial fermentation processes. *Metab Eng* **6**: 245–255.

van Maris AJA, Winkler AA, Porro D, van Dijken JP & Pronk JT (2004b) Homofermentative lactate production cannot sustain anaerobic growth of engineered *Saccharomyces cerevisiae*: possible consequence of energy-dependent lactate export. *Appl Environ Microbiol* **70**: 2898–2905.

van Maris AJA, Geertman JMA, Vermeulen A, Groothuizen MK, Winkler AA, Piper MDW, van Dijken JP & Pronk JT (2004c) Directed evolution of pyruvate decarboxylase-negative *Saccharomyces cerevisiae*, yielding a C_2-independent, glucose-tolerant, and pyruvate-hyperproducing yeast. *Appl Environ Microbiol* **70**: 159–166.

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

Vijayendran C, Barsch A, Friehs K, Niehaus K, Becker A & Flaschel E (2008) Perceiving molecular evolution processes in *Escherichia coli* by comprehensive metabolite and gene expression profiling. *Genome Biol* **9**: R72.

Wang Z, Gerstein M & Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* **10**: 57–63.

Weusthuis RA, Adams H, Scheffers WA & van Dijken J (1993) Energetics and kinetics of maltose transport in *Saccharomyces cerevisiae*: a continuous culture study. *Appl Environ Microbiol* **59**: 3102–3109.

Wieczorke R, Krampe S, Weierstall T, Freidel K, Hollenberg CP & Boles E (1999) Concurrent knock-out of at least 20 transporter genes is required to block uptake of hexoses in *Saccharomyces cerevisiae*. *FEBS Lett* **464**: 123–128.

Winzeler E, Lee B, McCusker J & Davis R (1999) Whole genome genetic-typing in yeast using high-density oligonucleotide arrays. *Parasitology* **118**: 73–80.

Wisselink HW, Cipollina C, Oud B, Crimi B, Heijnen J, Pronk J & van Maris A (2010) Metabolome, transcriptome and metabolic flux analysis of arabinose fermentation by engineered *Saccharomyces cerevisiae*. *Metab Eng* **12**: 537–551.

Wolfe KH & Shields DC (1997) Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* **387**: 708–713.

Xie C & Tammi MT (2009) CNV-seq, a new method to detect copy number variation using high-throughput sequencing. *BMC Bioinformatics* **10**: 80.