FEMS Yeast Research, 15, 2015, 1–11

doi: 10.1093/femsyr/fou008
Research Article

RESEARCH ARTICLE

Loss of lager specific genes and subtelomeric regions define two different *Saccharomyces cerevisiae* lineages for *Saccharomyces pastorianus* Group I and II strains

Chandre Monerawela[1], Tharappel C. James[1], Kenneth H. Wolfe[2] and Ursula Bond[1,*]

[1]Moyne Institute, School of Genetics and Microbiology, Trinity College Dublin, College Green, Dublin 2, Ireland and [2]UCD Conway Institute, School of Medicine and Medical Science, University College Dublin, Dublin 4, Ireland

*Corresponding author: Moyne Institute, School of Genetics and Microbiology, Trinity College Dublin, College Green, Dublin 2, Ireland. Tel:+353-1-896-26578; E-mail: ubond@tcd.ie

One sentence summary: We show that *Saccharomyces pastorianus* Group II yeasts are related to stout yeasts while Group I resemble ale yeasts and that the two groups arose by independent hybridization events.

Editor: Isak S. Pretorius

ABSTRACT

Lager yeasts, *Saccharomyces pastorianus*, are interspecies hybrids between *S. cerevisiae* and *S. eubayanus* and are classified into Group I and Group II clades. The genome of the Group II strain, Weihenstephan 34/70, contains eight so-called 'lager-specific' genes that are located in subtelomeric regions. We evaluated the origins of these genes through bioinformatic and PCR analyses of *Saccharomyces* genomes. We determined that four are of *cerevisiae* origin while four originate from *S. eubayanus*. The Group I yeasts contain all four *S. eubayanus* genes but individual strains contain only a subset of the *cerevisiae* genes. We identified *S. cerevisiae* strains that contain all four *cerevisiae* 'lager-specific' genes, and distinct patterns of loss of these genes in other strains. Analysis of the subtelomeric regions uncovered patterns of loss in different *S. cerevisiae* strains. We identify two classes of *S. cerevisiae* strains: ale yeasts (Foster O) and stout yeasts with patterns of 'lager-specific' genes and subtelomeric regions identical to Group I and II *S. pastorianus* yeasts, respectively. These findings lead us to propose that Group I and II *S. pastorianus* strains originate from separate hybridization events involving different *S. cerevisiae* lineages. Using the combined bioinformatic and PCR data, we describe a potential classification map for industrial yeasts.

Key words: lager yeasts; origins; stout yeasts; ale yeasts

INTRODUCTION

Beer is produced by the fermentation of sugars, mostly maltose, into alcohol by the action of yeasts and is generally classified as ale, stout or lager, based on the fermentation method and the type of yeast used in the process (Wunderlich and Back 2009). Stout and ales yeast are traditionally called top fermenters and lager yeasts, bottom fermenters, as the former tend to float to the top of the vessel at the end of the fermentation, while the latter sediment to the bottom of the vessel. Stout is a distinct type of beer. It is dark in colour due to the use of roasted barley or roasted malt in the grist. Ale brewing was established in the Middle Ages, and fermentations are carried out at 20–30°C with ale yeasts that are mostly diploid in nature belonging to the species *Saccharomyces cerevisiae* (Legras *et al.*, 2007). Ale yeasts are classified into four groups based on their degree of

Received: 1 September 2014; **Accepted:** 26 November 2014

flocculation (Gilliland 1955). Lager brewing emerged during the 15th century in Bavaria and requires cold temperatures between 8 and 15°C. Lager fermentations utilize the species *S. pastorianus*, formerly referred to as *S. carlsbergensis*, an allopolyploid interspecies hybrid that arose from the fusion of two distinct *Saccharomyces* species (Nilsson-Tillgren *et al.*, 1981; Dunn and Sherlock 2008; Bond 2009; Libkind *et al.*, 2011; Walther, Hesselbart and Wendland 2014).

Genetic analyses including single chromosome transfer experiments, individual gene sequencing, restriction fragment length polymorphisms and subsequently whole genome sequencing indicate that the parental genomes of *S. pastorianus* originate from an *S. cerevisiae* yeast and the newly discovered *S. eubayanus* (Nilsson-Tillgren *et al.*, 1981; Rainieri *et al.*, 2006; Legras *et al.*, 2007; Dunn and Sherlock 2008; Liti *et al.*, 2009; Nakao *et al.*, 2009; Libkind *et al.*, 2011). *Saccharomyces eubayanus* was first identified in an environmental niche in Patagonia, South America, but recently strains of *S. eubayanus* have been identified in the US and China although not as yet in Europe (Bing *et al.*, 2014; Peris *et al.*, 2014). It is proposed that *S. eubayanus* may have been transported to Europe through old trade routes between Europe and the Far East or during the period of exploration of the New World in the 15th and 16th centuries (Libkind *et al.*, 2011; Bing *et al.*, 2014).

In addition to the parental chromosomes, *S. pastorianus* contains unique lager-specific (LgS) hybrid chromosomes resulting from homeologous recombination between the parental chromosomes at specific sites (Bond *et al.*, 2004; Dunn and Sherlock 2008; Usher and Bond 2009). Furthermore, two distinct types of *S. pastorianus* strains have been identified; Group I, or Saaz-type strains, contain a roughly haploid *S. cerevisiae* DNA content, while Group II strains, the Frohberg type, contain a diploid *S. cerevisiae* content (Dunn and Sherlock 2008; Walther *et al.*, 2014). Whole genome sequencing of the two groups showed that the reference Group I strain CBS1513 (*S. carlsbergensis*) has 47 chromosomes (3n-1 ploidy with 29 different chromosome structures), whereas the reference Group II strain Weihenstephan (WS) 34/70 has approximately 64 chromosomes (4n ploidy with 36 different chromosome structures) (Nakao *et al.*, 2009; Walther *et al.*, 2014). Extensive aneuploidy is observed within each group, with different strains possessing different copy numbers of *S. cerevisiae*, *S. eubayanus* and hybrid chromosomes (Bond *et al.*, 2004; Hewitt *et al.*, 2014; Walther *et al.*, 2014). Group I *S. pastorianus* strains have also undergone substantial loss of heterozygosity and chromosome loss in the *S. cerevisiae* genome complement (Dunn and Sherlock 2008; Hewitt *et al.*, 2014; Walther *et al.*, 2014). The Group I and II differ in fermentation performances, in adaptation to growth at low temperature and in sugar utilization (Gibson *et al.*, 2013; Walther *et al.*, 2014).

There is disagreement about whether *S. pastorianus* originated from a single hybridization between *S. cerevisiae* and *S. eubayanus*, or if the Group I and II strains originated from two independent hybridizations. The major differences in DNA content and chromosome copy number between Groups I and II suggest two independent hybridizations (Dunn and Sherlock 2008). However, three sites of rearrangement between *S. cerevisiae* and *S. eubayanus* DNA, at loci *HSP82*, *XRN1/KEM1* and *MAT*, forming hybrid chromosomes, are shared identically between Group I and II *S. pastorianus* strains (Hewitt *et al.*, 2014; Walther *et al.*, 2014) which suggests a common origin and hence a single hybridization (Walther *et al.*, 2014). To reconcile these observations, one or other of two unlikely processes must have happened: either concerted loss of *S. cerevisiae* DNA occurred in the Group I lineage after a shared hybridization event or three identical interchromosomal rearrangements occurred in parallel in the Group I and Group II lineages after independent hybridization events.

The origin of the parental species contributing to the complex polyploid genomes of the lager yeasts represents a gap in our knowledge of the evolution of these important industrial yeasts. The genome sequence of the prototype Group II *S. pastorianus* WS 34/70 strain identified eight apparent LgS genes based on their absence from any of the *Saccharomyces* genome databases at the time (Nakao *et al.*, 2009). The genes are located in the subtelomeric regions (STRs) of the chromosomes, areas that display high levels of sequence diversity and are the focal points for rapid evolution (Brown, Murray and Verstrepen 2010; Dunn *et al.*, 2012; Bergstrom *et al.*, 2014). We hypothesized that identifying the origins of these LgS genes may provide insight into the ancestry of the genomes of *S. pastorianus* strains. We examined the publically available genome sequences from species in the genus *Saccharomyces* for the presence of the LgS genes. We also compared the STRs of the *S. cerevisiae*-derived chromosomes of *S. pastorianus* to other *S. cerevisiae* genomes to identify the most likely parental source of this portion of the lager yeast genomes. Using the patterns of presence or absence of the LgS genes and STRs, we show that Group I and Group II lager yeasts appear to originate separately from two different *S. cerevisiae* yeasts: Fosters O-like ale yeast and stouts yeast, respectively.

MATERIALS AND METHODS

Strains

Saccharomyces cerevisiae, *S. pastorianus* and *S. eubayanus* strains were acquired from various culture collections including the National Collection of Yeast Cultures (NCYC, Norwich UK), the Collection de Levures d'Interet Biotechnologique (CLIB, Paris France), the Belgium Lager Strains Collection (BLSC), Commonwealth Agricultural Bureaux International (CABI; National Collection of Fungus Cultures, UK), European Saccharomyces cerevisiae Archive for Functional Analysis (EUROSCARF, Frankfurt, Germany), the Portuguese Yeast Culture Collection (PYCC, Lisbon, Portugal), VTT Technical Research Centre of Finland (VTT, Helsinki, Finland), AB Mauri (NSW, Australia), Guinness Brewery (Dublin, Ireland) as well as individually sourced (Tables 2 and 3)

Yeast propagation

Yeast strains were cultured in YEP (3% yeast extract peptone) supplemented with 2% glucose or maltose at 30°C overnight.

Bioinformatic analysis of the LgS genes

DNA sequences of the *S. pastorianus* WS 34/70 LgS genes (Table 1) were compared to the genome sequences of *S. cerevisiae* strains (Table 2) in the *Saccharomyces* Genome Database (SGD, http://www.yeastgenome.org/) using WU-BLAST2 (http://www.yeastgenome.org/cgi-bin/blast-sgd.pl) and to the available genomes of other *Saccharomyces* species in SGD using Fungal-BLAST (http://www.yeastgenome.org/cgi-bin/blast-fungal.pl) using the same search parameters. In total, 42 genomes were included in the analysis. The majority of *S. cerevisiae* genome sequences were directly submitted to NCBI by Fay and colleagues (Washington University, St Louis), while the remainder of the

Table 1. LgS gene chromosome location and genome coordinates.

Gene	Putative gene product	Chromosome location	W34/70 contig	Coordinates
LgS-TYP	Tyrosine permease	*S. c* XIII/Hybrid XIII	ABPO01000020.1[a]	59175–60848
LgS-AMD1	Amidase	*S. c* XIII/Hybrid XIII	ABPO01000303.1[a]	7653–5857
LgS-TRR	Transcriptional regulatory protein	*S. c* XIV	ABPO01000417.1	4560–6947
LgS-Hypo	Hypothetical protein	*S. c* XVI	ABPO01000069.1	26652–25534
LgS-FSY	Fructose symporter	*S. e* IV/II[c]	ABPO01002087.1	375–2087
LgS-AMD2	Amidase	*S. e* V	ABPO01000969.1[b]	5466–6653
			ABPO01003184.1[b]	942–64
LgS-MCT	Monocarbohydrate transporter	*S. e* VI	ABPO01000605.1	9701–8631
LgS-MEL	Melibiase	*S. e* X/Hybrid X	ABPO01000141.1[a]	17859–16444

[a]One contig represents two different chromosomes, [b]LgS-AMD2 ORF in W34/70 on two contigs, [c]Reciprocal translocation between two *S. eubayanus* chromosomes.

Table 2. Detection of lager *S. cerevisiae* subtelomeric sequences and LgS genes in *S. cerevisiae* yeasts.

Strain	Origin	1[b]		2		3		4		5		6		7		8		9		10		11		12		13		TYP	AMD1	14		TRR	15		16		HYPO
		L[a]	R[a]	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	R-13	R-13	L	R	L-14	L	R	L	R	L-16
FostersO	Ale							■																		■		■	■									
FostersB	Ale	■						■																		■		■	■									
CBS 1513	*Sp* Group 1							■	■			■	■									■	■	■	■	■		■	■					■		■	
JAY291	Bioethanol							■												■						■		■	■	■		■					
ZTW1	Bioethanol																									■		■	■	■	■	■					■
CBS7960	Bioethanol	■			■				■											■						■		■	■	■		■					■
EC9-8	Canyon Isolate																									■		■	■								
YJM789	Clinical Isolate																									■		■	■								
Y10	Coconut	■								■										■						■			■						■		
CLIB215	Bakery	■						■																		■		■	■	■		■					■
CLIB324	Bakery	■																								■		■	■	■		■					■
BY4741	Laboratory																									■		■	■	■		■					■
BY4742	Laboratory																									■		■	■	■		■					■
S288C	Laboratory																									■		■	■	■		■					■
Sigma1278b	Laboratory																									■		■	■	■		■					■
W303	Laboratory																									■		■	■	■		■					■
CEN.PK113-7D	Laboratory	■																								■		■	■	■		■					■
FL100	Laboratory																											■									■
T7	Oak																													■		■					
Kyokai no.7	Sake							■																		■		■		■		■					■
UC5	Sake																									■		■		■	■	■					■
Vin13	Wine																									■		■	■	■		■					■
AWRI1631	Wine	■																								■		■	■	■		■					■
M22	Wine	■																								■		■	■	■		■				■	■
T73	Wine	■							■											■						■		■	■	■		■					■
EC1118	Wine																									■		■	■	■		■					
Lalvin QA23	Wine																									■		■	■	■		■					
VL3	Wine	■																								■		■	■	■		■					
YJM269	Wine	■																		■						■			■								
AWRI796	Wine							■																						■		■					■
RM11-1a	Wine	■																												■		■					■
PW5	Wine													■																							
Total Strains Missing Region		12	0	0	1	0	0	7	3	0	1	1	1	1	0	0	0	0	0	5	0	1	1	1	1	1	25	26	25	23	2	23	0	1	1	2	19

■ shaded area; not detected.

Saccharomyces genomes were deposited by various other groups (Cliften *et al.*, 2003; Kellis *et al.*, 2003; Wei *et al.*, 2007; Borneman *et al.*, 2008; Doniger *et al.*, 2008; Argueso *et al.*, 2009; Liti *et al.*, 2009; Novo *et al.*, 2009; Dowell *et al.*, 2010; Akao *et al.*, 2011; Ralser *et al.*, 2012; Liti *et al.*, 2013). Maximum-likelihood phylogenetic trees were built using the program MEGA5.2 (http://www.megasoftware.net) (Tamura *et al.*, 2011). Trees were constructed using the default parameters including the bootstrap method with 500 bootstrap replications.

Bioinformatic analysis of the STRs of the *S. cerevisiae*-like chromosomes in *S. pastorianus*

Query sequences corresponding to the STRs of *S. cerevisiae* chromosomes in the *S. pastorianus* strain WS 34/70 (Nakao *et al.*, 2009) were obtained from NCBI (http://www.ncbi.nlm.nih.gov;ABPO01). STRs were defined as being less than 30 kbp from the ends of the contigs lying closest to the ends of chromosomes. The chromosome coordinates for each STR are listed in Table S1 (Supporting Information). The query sequences were used to search for homologous regions in *S. cerevisiae* strains using WU-BLAST2 and in the *S. pastorianus* strain CBS1513 using NCBI-BLAST using the same search parameters. The identified sequences were aligned using the ClustalW2 server (http://www.ebi.ac.uk/Tools/msa/clustalw2/). A sliding window DNA sequence comparison algorithm was used to generate percentage sequence identity maps directly from the ClustalW2 alignments. Maximum-likelihood phylogenetic trees were built using the program MEGA5.2. Pairwise distances of each *S. cerevisiae* genome to the subtelomeric query sequence were calculated in MEGA 5.2 using the maximum composite likelihood model and the default parameters.

Table 3. Detection of LgS genes and STRs IV-L and XIII-R in *S. cerevisiae* and *S. pastorianus* strains.

Strain	Group	Source	TYP	AMD	R-XIII	TRR	Hypo	L-IV
CMBS	Group2	BLSC						
7012	Group2	Guinness						
6701	Group2	Guinness						
3701	Group2	Guinness						
2001	Group2	Guinness						
2002	Group2	Guinness						
2106	Group2	Guinness						
A15	Group2	VTT						
1511	Stout Ale	NCYC						
1063	Stout Ale	NCYC						
1064	Stout Ale	NCYC						
1164	Stout Ale	Guinness						
ATH	Coconut Water	India						
Malabar	Coconut Water	India						
Rauch Bier	Ale	Beer Bottle						
Schneider Weisse	Ale	Beer Bottle						
87	Distillers	NCYC						
ABMY0497	Distillers	Mauri						
Y1	Industrial	Guinness						
Chimay	Ale	Beer Bottle						
Kloster Hell	Ale	Beer Bottle						
ABMY0154	Distillers	Mauri						
ABMY0514	Distillers	Mauri						
R103	Industrial	Guinness						
2592	Distillers	NCYC						
ABMY0904	Distillers	Mauri						
ABMY1359	Distillers	Mauri						
CBS 1503	Group1	CLIB						
CBS 1513	Group1	CLIB						
CBS 1538	Group1	CLIB						
1075	Ale	NCYC						
1077	Ale	NCYC						
662	Ale	Guinness						
Hens Tooth	Ale	Beer Bottle						
Sierra Nevada	Ale	Beer Bottle						
Coopers Brewery	Ale	Beer Bottle						
1151	Ale	NCYC						
2397	Ale	NCYC						
BY5625	Laboratory	CABI						
K103	Industrial	Guinness						
Y133	Industrial	Guinness						
Y2694	Industrial	Guinness						
S288C	Laboratory	Novagen						
S150	Laboratory	Lab strain						
W303	Laboratory	EUROSCARF						
BY4741	Laboratory	EUROSCARF						

Shaded regions: not detected.

Single nucleotide polymorphism (SNP) analysis

Query sequences corresponding to the non-STRs of *S. cerevisiae*-like chromosomes in the *S. pastorianus* strain WS 34/70 were obtained from NCBI after selecting *S. cerevisiae* contigs greater than 10 000 bp in size and greater than 30 kb from the telomeres, as defined above, from the supplemental information of Nakao *et al.* (2009). The chromosome coordinates for each non-STR are listed in Table S2 (Supporting Information). The query sequences were used to search for homologous regions in *S. cerevisiae* ale strains Fosters B and Fosters O using WU-BLAST2 and in the *S. pastorianus* strain CBS1513 using NCBI-BLAST using the same search parameters. The identified regions were aligned using MEGA5.2. SNPs were manually counted in the sequence alignment ignoring sequencing errors (N bases) and insertions (Ty elements).

DNA extraction and polymerase chain reaction (PCR)

DNA was extracted from the strains listed in Table 3 using the phenol–chloroform method as previously described (James, Campbell and Bond 2002). PCRs were carried out in 25 μl reactions containing 250 μM dNTPs, 0.4 nM of each oligonucleotide primer, 1X *Taq* DNA polymerase buffer (10 mM Tris-HCl, pH 8.3,

50 mM KCl, 1.5 mM $MgCl_2$, 1% Triton, New England Biolabs, MA, USA), 1U *Taq* DNA polymerase (YB *Taq*, Yorkshire Bioscience Ltd, York, UK) and 200 ng template DNA. Samples were initially denatured at 95°C for 3 min followed by 35 cycles of denaturation at 95°C for 30 s and annealing and extension at 72°C for 1 min 45 s. A final extension at 72°C for 7 min was applied. The DNA sequences of primers used in PCR reactions are shown in Table S3 (Supporting Information). Products were electrophoresed on a 1.6% agarose TBE gels and were visualized by staining with ethidium bromide.

Fermentations

Fermentations were carried out in wort with a Brix value of 12. Lager and ale yeast cultures were pitched at cell densities of 1.5 $\times$ 10^7 cells ml^{-1}. Fermentations were carried out at 15 or 22°C under microaerophilic conditions in sterile 250 ml cylindrical glass flasks. Samples (40 ml) were collected in pre-chilled tubes and centrifuged at 4000 rpm for 10 min at 4°C. Samples were collected at intervals during the fermentations and wort brix values were measured using an electronic refractometer. Following fermentation, the cells were pelleted and washed in 10 ml of cold sterile distilled water. The cells were recovered and resuspended in 1 ml of water and frozen at −70°C.

RNA extraction and reverse-transcription PCR (RT-PCR)

RNA was extracted from the fermentation samples using the hot-phenol method as previously described (James *et al.*, 2002). RNA samples (4 μg/reaction) were treated with DNase I (Promega Inc., WI, USA) to remove any contaminating DNA. DNase-treated RNA was reverse transcribed in a mixture containing 10X Reverse Transcriptase Buffer, 0.5 mM dNTP mix, 1 mM reverse primer and 50U Reverse Transcriptase (Applied Biosystems/M-MLV) in a reaction volume of 20 μl. The reactions were carried out 25°C for 10 min, 37°C for 2 h and finally 85°C for 5 min. 2 μl of cDNA was amplified by PCR as previously described (Beggs, James and Bond 2012) using the primers listed in Table S3 (Supporting Information). The products were electrophoresed on a 1.6% agarose TBE gels and stained with ethidium bromide.

RESULTS

'Lager-specific' genes are present in some *S. cerevisiae* strains

Given the increased sequence diversity in the STRs of the *S. pastorianus and S. cerevisiae* genomes (Dunn *et al.*, 2012; Bergstrom *et al.*, 2014) and in particular the presence of LgS genes in STRs, we reasoned that a detailed analysis of these regions of the chromosomes might aid in identifying the most likely parental origins for the two genomes present in the species. Contigs containing the LgS genes were identified from the Group II strain WS 34/70 genome sequence (Nakao *et al.*, 2009) (Table 1). Four of the genes, LgS-TYP, LgS-AMD1, LgS-TRR and LgS-Hypo, are located on *S. cerevisiae*-like chromosomes while the other four genes, LgS-FSY, LgS-AMD2, LgS-MCT and LgS-MEL, are found on *S. eubayanus* chromosomes. With the exception of LgS-Hypo, seven of the genes are located within 30 kb of the telomeres. Bioinformatic analysis of genomes representing the genus *Saccharomyces*, using the eight LgS genes as query sequences, revealed that orthologues are present in many of the genomes with varying percentage identities (70–100%). Further phylogenetic analyses identified the *Saccharomyces* species showing the closest relationship for seven of the LgS genes. While the four *S. cerevisiae* LgS genes LgS-TYP, LgS-AMD1, LgS-TRR and LgS-Hypo clustered most closely with *S. cerevisiae* strains, two different phylogenetic clades were identified within this group (Fig. 1A). LgS-TYP and LgS-AMD1 genes, which are colocated on the right STRs of both the *S. cerevisiae* and the hybrid-type chromosome XIII (STR XIII-R), clustered with the wine yeasts RM11-1a and AWRI796 (Table 2). On the other hand LgS-TRR and LgS-Hypo, which are located on the left STRs of the *S. cerevisiae* chromosome XIV and the hybrid chromosome XVI respectively (Table 1), show highest identity to the Fosters B and O ale yeast strains (Fig. 1A). While homologues of LgS-TYP, LgS-AMD1 and LgS-Hypo were identified in other non-*S. cerevisiae* yeasts, the identified sequences had lower similarity (79–90% compared to 94–100% for the *S. cerevisiae* yeasts).

The three LgS genes, LgS-FSY, LgS-AMD2 and LgS-MEL, cluster most closely with the *S. bayanus* and *S. eubayanus* genomes (Fig. 1B). LgS-FSY and LgS-MEL show 99% identity to *S. eubayanus* gene sequences. The nearest neighbour for LgS-AMD2 is *S. bayanus*, with 92% identity; this gene is not present in the *S. eubayanus* genome. Homologues of the *S. eubayanus* LgS genes can be detected in other *Saccharomyces* yeasts, *albeit* with lower percentage identities (70–87%). The fourth LgS gene, LgS-MCT, of suspected *S. eubayanus* ancestry, was not found in any of the publically available *Saccharomyces* yeast genome sequences. Since horizontal gene transfer in yeast is quite rare and LgS-MCT resides on an *S. eubayanus*-derived chromosome in lager yeasts, it is most likely that this gene is absent from the genome sequence of the studied type strain of *S. eubayanus*. We confirmed the assignment of *S. pastorianus* LgS-MCT, LgS-AMD2, LgS-FSY and LgS-MEL to *S. eubayanus* ancestry by PCR from *S. eubayanus* genomic DNA (Fig. 2A). None of the eight LgS genes could be amplified from the *S. cerevisiae* reference strain S288C (Fig. 2B).

Analysis of STRs of *S. cerevisiae* chromosomes in *S. pastorianus*

We expanded the analysis to the STRs of all 16 *S. cerevisiae* chromosomes in *S. pastorianus* WS 34/70. The regions examined lie closer to the telomeres than the LgS genes (Table 1 and Table S1, Supporting Information). The 32 *S. pastorianus* WS 34/70 STRs of *S. cerevisiae* ancestry were used as query sequences in pairwise alignments to the genomes of a diverse range of *S. cerevisiae* strains (Table 2). Averaging the nucleotide divergence level (number of base substitutions per site) for all 32 STRs identifies the Fosters O and B ale strains as the nearest neighbours of the *S. cerevisiae* component of the WS 34/70 strain (Fig. 3). Average divergences of 0.0010 and 0.0022 substitutions per site were observed for Fosters B and Fosters O, respectively. Several wine yeasts (Lalvin QA23, Vin13, EC1118, AWRI1631, AWRI796 and VL3) show divergence <0.005, while the divergence of the other *S. cerevisiae* strains ranges from 0.005 to 0.020. Of the 32 phylogenetic trees, the lager yeast clustered with the ale yeasts at 11 STRs (data not shown). For 17 other STRs, various yeast strains, including the ales, clustered with the lager yeast, indicating no discrimination between yeast strains. For the remaining four *S. pastorianus* STRs, IV-L, IX-R, XII-R and XIII-R), the ale yeasts did not cluster near *S. pastorianus* (data not shown).

The analysis of each STR separately uncovered several *Saccharomyces* strains in which the lager yeast STRs were absent (Table 2). The most commonly absent STRs in the strains are STRs XIII-R (25/31), XIV-L (23/31), I-L (12/31) and IV-L (7/31) (Table 2). STRs VII-L, X-L, XVI-L, II-R, IV-R, V-R and XVI-R also appeared to be absent in some strains (Table 2) but omissions in genome

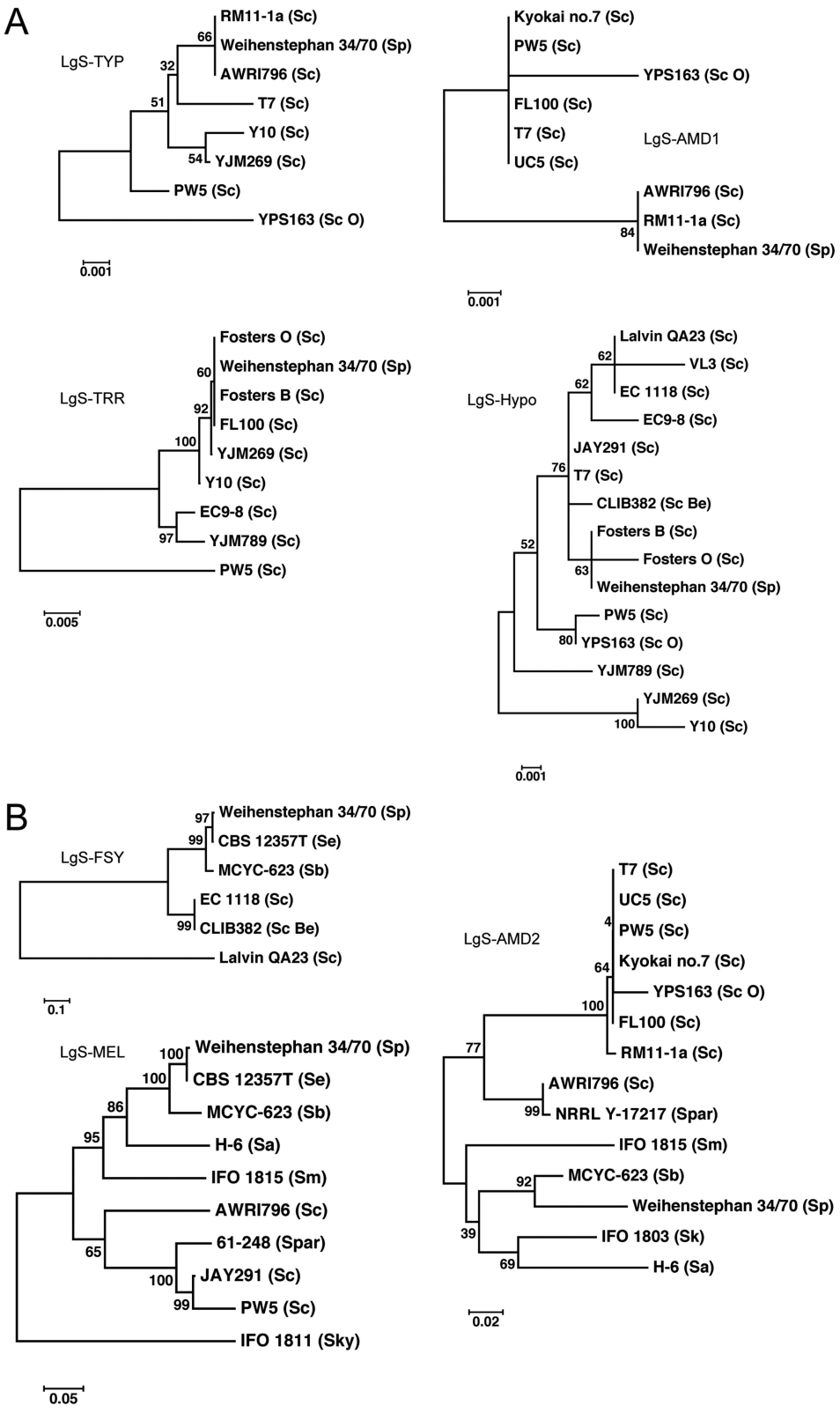

Figure 1. Phylogenetic trees for the *S. pastorianus* 'lager specific' genes. *Saccharomyces cerevisiae* genes (**A**), *S. eubayanus* genes (**B**). The names of the genes are shown beside each panel. In total, 42 genomes were analysed. The clades clustering closest to Weihenstephan 34/70 are shown only. Bootstrap confidence values are shown at branch points. Scale bars, bases substitutions per site. *Saccharomyces cerevisiae* (Sc) strains are described in Table 2 with two exceptions; Sc O, from oak tree and Sc Be, a beer isolate. *S. pastorianus* (Sp), *S. eubayanus* (Se), *S. bayanus* (Sb), *S. paradoxus* (Spar), *S. mikatae* (Sm) *S. kudriavzevii* (Sk) *S. arboricola* (Sa) *S. kluyveri* (Sky).

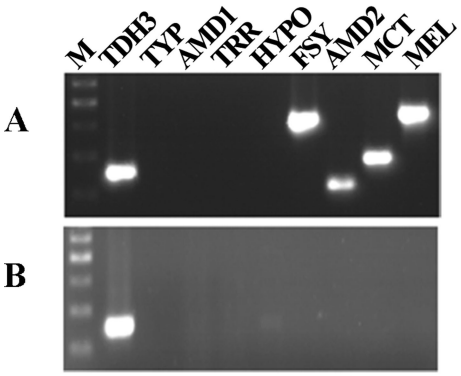

Figure 2. Detection of LgS genes by PCR in (**A**) *S. eubayanus* PYCC6148 and (**B**) *S. cerevisiae* S288C. The names of the LgS genes are shown above the lanes. M, molecular weight marker Q-step ladder 1 showing the range 200–600 bp. TDH3 was used as a positive control for PCR reactions.

sequence assembly cannot be completely ruled out. Analysis of the Group I *S. pastorianus* strain CBS1513 revealed that its reported genome sequence (Walther *et al.*, 2014) lacks 12 *S. cerevisiae*-derived STRs that are present in the Group II strain WS 34/70 (STRs IV-L, IV-R, VI-L, VI-R, XI-L, XI-R, XII-L, XII-R, XIII-L, XIII-R, XV-R and XVI-R; Table 2).

We also analysed the *S. cerevisiae* genomes for the presence or absence of the four *S. cerevisiae*-like LgS genes. The pattern of absence of the four *S. cerevisiae* LgS genes differed from strain to strain (Table 2). Just a single strain, the wine yeast PW5, contained DNA sequences corresponding to all four genes, despite displaying a high divergence level (0.017). LgS-AMD1 sequences were identified in 7 yeast strains, LgS-TYP in 6, LgS-TRR in 9, and LgS-Hypo in 13. The strains Fosters O and B strains, which showed the lowest divergence level, lack both LgS-TYP and LgS-AMD1. Notably, all four *S. cerevisiae* LgS genes are absent from bakery and laboratory strains, with the exception of laboratory strain FL100, which contains both LgS-AMD1 and LgS-TRR. The four *S. cerevisiae* LgS genes are also missing in a subset of wine strains (Table 2).

PCR detection of 'Lager-specific' genes and STRs in brewery, industrial and laboratory *S. cerevisiae*

Since the presence or absence of the *S. pastorianus* STRs and LgS genes appears to be good discriminators of *S. cerevisiae* strains, yeasts from different industrial sources as well as several laboratory strains were examined by PCR analysis for the presence or absence of the LgS genes and for two of the STRs, namely STR IV-L and STR XIII-R. The region STR IV-L was chosen because it is absent in the Fosters O and B ale strains. PCR analysis was also carried out on DNA extracted from several Group I and II lager yeasts and the patterns were compared to those produced by the industrial yeasts (Fig. 4 and Table 3).

The analysis revealed a close relationship between Group II *S. pastorianus* lager yeasts and *S. cerevisiae* strains used for stout production. PCR products for all six regions were amplified from (a) all Group II lager yeasts, (b) four yeast strains classified as stout yeasts (Fig. 4) and (c) from two yeast strains isolated from a fermented alcoholic beverage, Toddy, which is produced from coconut water in rural south west India (Table 3). Yeasts isolated from two German bottled beers (Rauch Bier and Schneider Weisse) showed the same pattern with the exception that they lacked STR IV-L (Table 3).

Group I lager yeasts (CBS 1503, 1513 and 1538) lack DNA corresponding to STR IV-L and STR XIII-R, LgS-AMD1 and LgS-TYP (Fig. 4) as do strains 1075, 1077 (NCYC) and 662 (Guinness), whose fermentation products are classified as ales. The same pattern was observed in several yeast strains isolated from bottled craft beers (Table 3). This pattern is the same as that identified in Fosters O and B ale yeasts by bioinformatic analysis (Table 2). The Group I strain CBS1538 also lacks DNA corresponding to the LgS-Hypo gene (Fig. 4). Two other ale yeasts (NCYC 1151 and NCYC 2397) show a similar pattern as other ale yeasts with the exception that these strains contain the STR IV-L (Table 3).

Several yeast strains, designated as distillers or industrial yeasts, cluster together in lacking LgS-TYP. In some of these strains, either STR IV-L, LgS-TRR or LgS-Hypo are additionally absent (Table 3). Interestingly, yeast from bottles of the Chimay and Kloster Hell beers cluster with this group. Finally, all laboratory yeasts examined here, with one exception, cluster together

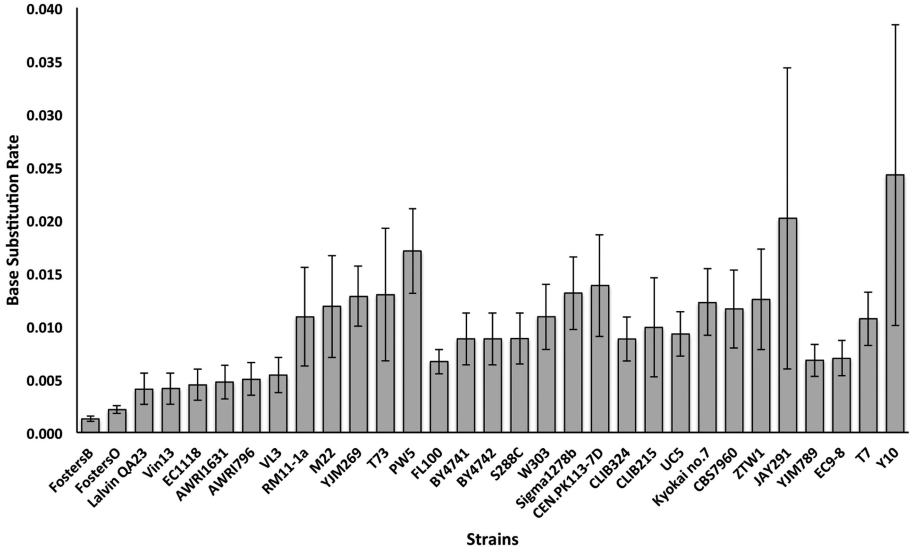

Figure 3. Base substitution rates in the STRs of *S. cerevisiae* strains. The 32 STRs of *S. cerevisiae* chromosomes in *S. pastorianus* Weihenstephan 34/70 were used as query sequences to identify homologous regions in 31 *S. cerevisiae* strains (Table 2). Pairwise distances of each *S. cerevisiae* genome to the subtelomeric query sequence were calculated in MEGA 5.2. Error bars show the standard error of the mean for all 32 STRs.

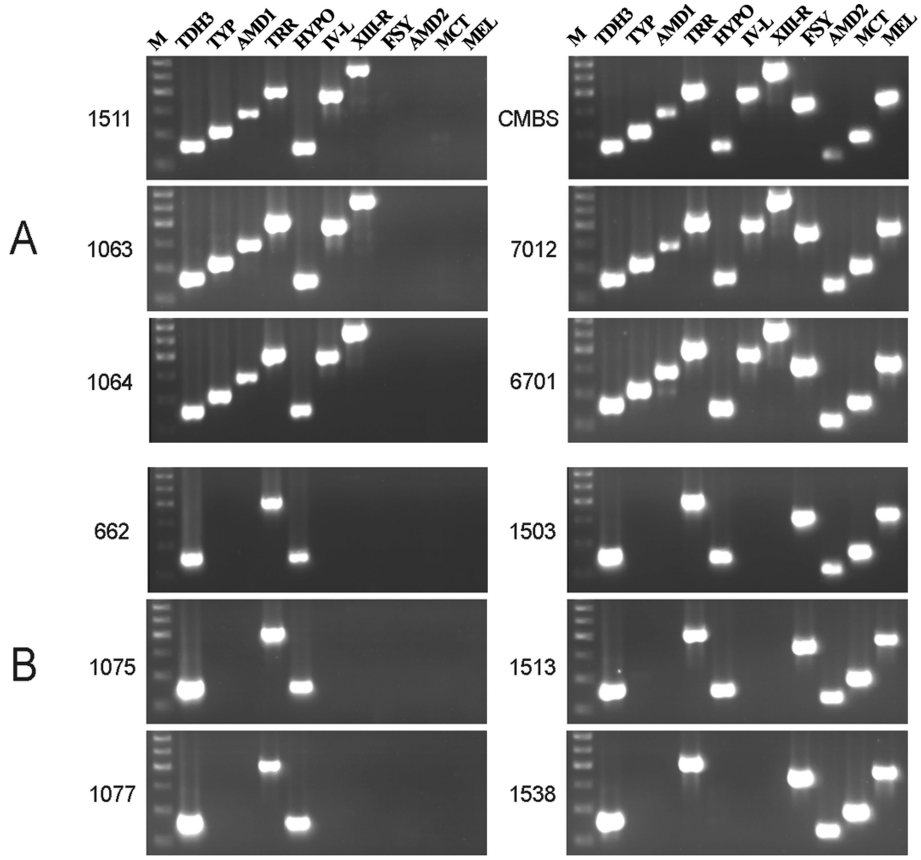

Figure 4. Detection of the LgS genes and the STRs of chromosomes IV-L and XIII-R by PCR in stout (1511, 1063 and 1064) and Group 2 lager yeasts (CMBS, 7012 and 6701) (**A**) and ale (662 and 1075, 1077) and Group 1 lager yeasts (CBS 1503, CBS 1513 and CBS 1538) (**B**). The STRs and LgS gene names are shown above the lanes. Lane M, molecular weight marker Q-step ladder 1 showing the range 200–700 bp. TDH3 was used as a positive control for PCR reactions.

in lacking all four LgS genes as well as STR XIII-R but contain STR IV-L. Thus, *S. cerevisiae* strains of different origins can be grouped together based on the presence or absence of these six markers, and furthermore Group I and II lager yeasts can be clearly distinguished from each other.

Analysis of non-telomeric chromosomal regions in Group I and II lager yeasts

The data thus far indicate that Group I and II lager yeasts can be distinguished by the pattern of loss of STRs and LgS genes. The Fosters O and B ale yeasts show the same LgS gene and STR loss pattern as Group I lager yeasts, yet pairwise distance analysis also reveals high sequence identity between Fosters O/B and Group II lager yeasts (Fig. 3). Genomic sequences for three Group I, including the Carlsberg strain CBS 1513 (Table 3) and two Group II lager yeasts, have recently been published (Nakao *et al.*, 2009; Hewitt *et al.*, 2014; Walther *et al.*, 2014). To tease out the relationship between Group I and Group II lager yeasts and to explore the possibility that the Group I and II lager yeasts arose from a single hybridization event and thus share a common ancestor, we compared the *S. cerevisiae* genomes of CBS1513, WS 34/70 and the two ale yeasts Fosters O and B for SNPs at 10 arbitrarily chosen non-telomeric regions (Table 4). The regions surrounding the three recombination sites on hybrid chromosomes III, VII and XVI that are conserved between Group I and Group II strains were also examined as these represent points of observed identity between the two groups. An average of 3.3 SNPs kbp^{-1} was observed between the Group I and Group II lager yeast strains from an analysis of over 126 000 bp. Based on the estimated single nucleotide mutation rate of 1.67×10^{-10} per base per generation in *S. cerevisiae* (Zhu *et al.*, 2014), the rate of SNP accumulation is over 10-fold higher than that expected from divergence following a single hybridization event occurring some 500 years ago. Comparison of SNPs between the Group I and Group II strains and the Fosters O and B ales revealed similar SNP accumulation rates (Table 4). CBS1513 was slightly more similar to both Fosters O and B ale strains than was the Group II WS 34/70 strain, but the differences were not statistically significant. Thus, Group I and II lager yeasts are indistinguishable from Fosters O and B ale yeasts and from each other at a SNP level and contain a much higher degree of diversity than expected from a single hybridization event occurring about 500 years ago. Surprisingly, analysis of the DNA sequences surrounding the shared recombination site on the hybrid chromosome III revealed a significantly higher nucleotide divergence between Group I and Group II lager yeasts than at the other chromosomal regions examined. This divergence is accounted for by the fact that the Group I strain contains an (α) mating-type cassette at the MAT locus while the Group II strain contains an (a) cassette (Table 4).

Expression of LgS genes during fermentation

Finally, we asked if the LgS genes are transcriptionally active and thus have the potential to encode for functional gene products. Gene-specific transcripts were detected by RT-PCR from RNA

Table 4. SNP divergence in non-STRs among Fosters ales and the *S. cerevisiae*-derived chromosomes of Group I and II *S. pastorianus* genomes.

Chromosome	Region	Query length (bp)	FostersB vs WS34/70	FostersO vs WS34/70	FostersB vs CBS1513	FostersO vs CBS1513	WS34/70 vs CBS1513
I	YAL033W-YAL029C	9564	3.14	2.4	2.51	1.88	1.05
II	YBR046C-YBR053C	10243	6.44	6.44	2.15	3.61	4.88
IV	YDR004W-YDR010C	10210	5.48	4.51	5.11	3.73	5.78
V	YER048C-YER052C	10990	1.55	3.91	0.94	3.38	0.91
VII	YGL114W-YGL108C	10612	1.04	2.17	2.55	4.43	3.3
VIII	YHR046C-YHR050W	11899	5.21	8.32	3.78	3.02	7.73
IX	YIL027C-YIL021W	10189	5.69	4.32	5.1	3.53	1.57
XIII	YMR041C-YMR048W	10982	9.29	9.11	8.1	8.47	3.73
XIV	YNL239W-YNL229C	17571	1.2	2.45	4.44	5.12	4.27
XV	YOL045W-YOL039W	9917	3.43	2.52	2.72	2.22	3.33
VII*	YGL173C	4111	1.22	3.65	1.7	4.14	1.7
XVI*	YPL242C-YPL240C	4502	2.44	4.44	2.44	3.55	1.56
III*^	YCR039C-YCR042C	5782	3.67	5.03	83.16	84.52	81.46
Average SNP kbp^{-1}			3.84	4.52	3.46	3.92	3.32

* hybrid chromosome, ^ not included in average SNP kbp^{-1}.

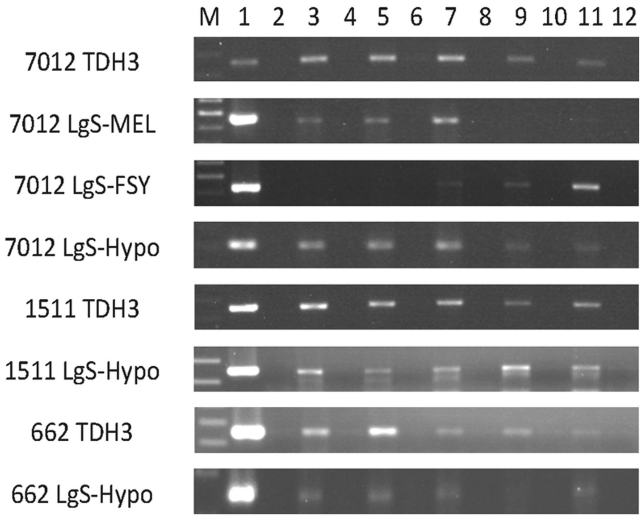

Figure 5. Detection of LgS gene transcripts by RT-PCR in lager and ale yeasts under fermentation conditions. The yeast strains and LgS gene names are shown to the left of the panels. Lane M, molecular weight marker Q-step ladder 1 showing the range 200–300 bp for TDH3 and LgS-Hypo and 300–400 bp for LgS-FSY and LgS-MEL; lane 1, PCR products from genomic DNA; 2, as lane 1 but PCR minus genomic DNA; lanes 3, 5, 7, 9, 11, RT-PCR reactions using RNA isolated from days 1, 2, 3, 5 and 7 of fermentation respectively. Lanes 4, 6, 8, 10 and 12, RT-PCR reactions from RNA isolated from days 1, 2, 3, 5 and 7 but without reverse transcriptase.

isolated at different stages of fermentations carried out with the Group II lager yeast, 7012 (Fig. 5). Of the eight LgS genes, transcripts were only detected for LgS-MEL, LgS-FSY and LgS-Hypo. Transcripts for LgS-MEL were detected on days 1–3 of fermentation and not in the latter stages (days 5–7), while conversely LgS-FSY transcripts are detected on days 3–7 and not in the early stages (Fig. 5). LgS-Hypo appears to be constitutively expressed throughout the fermentation (Fig. 5). The same pattern of expression was also observed for LgS-Hypo in fermentations carried out with either stout or ale yeasts (Fig. 5, 1511 and 662, respectively). As expected, transcripts for LgS MEL and LgS-FSY are not detected in the ale/stout fermentations, as the *S. eubayanus* genes are not present in these strains (data not shown).

DISCUSSION

The recent availability of whole genome sequences for Group I and II lager yeasts and for hundreds of *Saccharomyces* species affords an opportunity to re-evaluate the origins of the lager yeast genomes and to uncover minute but specific differences between yeasts that have evolved as part of industrial domestication. The first insight into unique characteristics of lager yeasts was the identification of a set of LgS genes in the genome of the prototypic Group II lager yeast, WS 34/70 (Nakao *et al.*, 2009). The genes were designated 'lager-specific' as they were absent from the genomes of the reference strains *S. cerevisiae* S288C and *S. bayanus* var. *uvaurm* (CBS7001; *S. bayanus*), the most likely parental genomes of the lager yeasts at the time of publication (Nakao *et al.*, 2009).

Since the LgS genes represent points of diversity in *S. pastorianus*, we re-evaluated the origins of these genes through a bioinformatic analysis of *Saccharomyces* genomes to determine the most likely origin of the parent genomes. Our results indicate that four of the genes are of *cerevisiae* origin while the other four originate from *S. eubayanus*. The *cerevisiae* LgS genes are good discriminators of *S. cerevisiae* strains and are indeed absent from the reference strain S288C.

An expanded analysis of the STRs of the Group II WS 34/70 *S. cerevisiae* chromosomes uncovered additional differences between *S. cerevisiae* strains. The chromosomal regions most commonly lost in *S. cerevisiae* strains are STR XIII-R and XIV-L. Loss of DNA sequences at STR XIII-R has also been reported in several other *S. cerevisiae* strains (Bergstrom *et al.*, 2014). Three of the LgS genes (LgS-TYP, LgS-AMD1 and LgS-TRR) lie in these STRs. In the majority of cases, the loss of LgS-AMD1 and LgS-TRR is correlated with the loss of STR XIII-R and XIV-L, respectively; however, both bioinformatic and PCR analysis indicate that LgS-TYP, which lies more distal from the telomere than LgS-AMD1 and STR XIII-R sequences, can be independently lost (Table 4, Supporting Information). In strain Y10, LgS-TYP was present while STR XIII-R was absent. The pattern of presence or absence of the LgS in *S. cerevisiae* strains explains the phylogenetic trees derived for these genes (Fig. 1A). Both HYPO and TRR cluster most closely with the Fosters O and B ales, confirming an ale origin for the *cerevisiae* component of the lager yeast genome (Legras *et al.*, 2007; Dunn and Sherlock 2008), while AMD1 and TYP1 fail

to do so as the region where these genes reside (STR XIII-L) is absent in both ale strains.

The STR analysis uncovered new regions of the *S. pastorianus* Group II genome not previously identified as being absent in *S. cerevisiae* strains. For example, STR I-L is missing in several wine yeasts, in two bakery yeasts and also distinguishes Fosters O and B ale yeasts. It is possible that these regions may be present in the strains but missing from the assembled genome sequences, but the absence of these regions from multiple strains suggests that they are truly missing.

Loss of LgS-Hypo in 19 *S. cerevisiae* strains is not correlated with the loss of the STR XVI-L, which is only absent in a single yeast strain, Y10. This was not unexpected as LgS-Hypo resides approximately 56 kb from the telomere and therefore based on our definition and that of others (Brown *et al.*, 2010) is considered non-subtelomeric. However, at this distance LgS-Hypo may still be influenced by telomere dynamics. Examination of chromosomal regions surrounding LgS-Hypo reveals the presence of a Ty element, which may have been responsible for the loss of this gene from various *S. cerevisiae* yeasts. Interestingly, of all of the *S. cerevisiae* LgS genes, LgS-Hypo is the only one that is transcribed during fermentation conditions, at levels that can be detected by RT-PCR, indicating that it may encode for a functional protein. The putative function of Hypo is unknown, however, Hypo shares a domain motif and 27% amino acid identity with saccharopine dehydrogenase from the bacterial species *Alteromonas macleodii*. Saccharopine, an intermediate metabolite in lysine catabolism, was first discovered in brewer's yeasts (Darling and Larsen 1961). The Hypo protein also contains two putative transmembrane domains at the N-terminus. Further analysis of these domains may provide insight into its cellular role.

The pattern of loss of the four *S. cerevisiae* LgS genes and STR XIII-R and STR IV-L clearly distinguishes between the Group I and Group II lager yeasts and furthermore separates 'ale' yeasts (including stouts) into Group I or II-like clades. The six DNA sequences are found in all Group II lager yeasts examined, in strains of yeasts designated as stout and in natural yeasts used for the fermentation of Toddy from coconut water. Additionally, five of the six DNA sequences are also present in yeasts isolated from two different beer bottles, a Weisse bier and a Rauch Bier. Stout is mainly brewed in the British Isles and may have evolved in the unique fermentation conditions used for stout production, which is distinctive in the use of roasted malt barley. Interestingly, Rauch (smoked) bier also uses a roasting method in its fermentation process (Pavsler and Buiatti 2009).

Group I lager yeasts and Group I-like ale yeasts are typified by the absence of STR XIII-R and associated LgS-TYP and LgS-AMD1 as well as STR IV-L. This category includes several yeast strains isolated from bottled beers. Based on the combined data from bioinformatic and PCR analysis, we hypothesize that the *S. cerevisiae* genome in Group I yeasts originates from Fosters O-like ale yeasts while the Group II genome originates from the stout yeast lineage exemplified by the prototype strain, NCYC 1511. Fosters O is distinguished from Fosters B by the presence of STR I-L which is also present in Group I lager yeasts. It will be interesting to see how our grouping of ale yeasts relates to a previous classification of these yeasts based on flocculation properties (Gilliland 1955), particularly as the FLO genes, which encode for flocculation proteins, are located in STRs.

While Group I and II lager yeasts are clearly distinguishable by types and numbers of chromosomes, the two groups share three common breakpoints at which homeologous recombination between the *S. cerevisiae* and *S. eubayanus* parental chromosomes has generated hybrid chromosomes (Hewitt *et al.*, 2014; Walther *et al.*, 2014). This observation has led others to propose a common parental origin for the *S. cerevisiae* genome for the two groups (Walther *et al.*, 2014). However, that proposal also necessitates the massive and specific loss of almost one entire copy of the *S. cerevisiae* component of the hybrid genome in an early Group I strain. The alternative hypothesis, that the three shared rearrangements are convergent between Group I and II strains, seems more probable, particularly since one of the three breakpoints is in the MAT locus and may have been caused by attempted mating-type switching (Nakao *et al.*, 2009). We have previously shown that homeologous recombination at breakpoints can be induced by exposure of yeasts to stressful environmental conditions (James *et al.*, 2008). Thus, breakpoints may represent natural fragile sites that are prone to recombination and therefore may arise independently in the two groups. Indeed, genome and PCR analyses have identified at least two Group I lager strains that do not share these common breakpoints (Monerawela and Bond, unpublished and Hewitt *et al.*, 2014). Our analysis also revealed that the Group I strain (CBS 1531) contains an (α) cassette at the MAT locus on the hybrid chromosome III while the Group II strain WS34/70 contains an (a) cassette, indicating further genome diversity between the two groups. Furthermore, the SNP analysis of non- STRs of the Group I and II strains revealed a much higher divergence rate than that expected from a single hybridization event that occurred sometime during the past 500 years. We observed a similar average SNP accumulation rate between Group I, Group II lager yeasts and the two ale yeasts Fosters O and B. This low level of sequence diversity between the *S. cerevisiae* genomes confirms the recent data from Bergstrom *et al.* (2014), who demonstrated that genetic distances, as measured by SNPs, are extremely low in *S. cerevisiae* strains. Instead, major genetic diversity in *S. cerevisiae* arises from loss of function and copy number variations in STRs (Bergstrom *et al.*, 2014), confirming our findings here. While the accumulated data support the hypothesis that at least two independent hybridization events, involving different *S. cerevisiae* brewing strains, contributed to the generation of the current lager yeasts, we cannot, at present, rule out the possibility that multiple rounds of hybridizations may have occurred.

Finally, based on the pattern of the presence or absence of the four LgS genes and two STRs, XIII-R and IV-L, we propose a broad grouping of industrial and laboratory yeasts (Table S4, Supporting Information). Stout, Toddy and Group II yeasts contain all six markers and thus represent a group of strains with genomes that represent the most complete gene repertoire of the species *S. cerevisiae*. Closely related to this group are the wine yeasts RM11-1a, AWR1796 and a yeast strain isolated from oak trees (T7). Distiller and other industrial yeasts are grouped based on the absence of LgS-TYP and additionally in some strains STR IV-L. Two 'ale' yeast strains recovered from two different bottled craft ales (Belgian Chimay and German Kloster Hell) also share LgS and STR features of the distiller strains as do two sake yeasts and strain FL100. The latter is classified as a laboratory yeast but based on the present data perhaps should be reclassified. The majority of ale and all Group I lager yeasts lack LgS-TYP, LgS-AMD1, STR XIII-R and STR IV-L sequences. Finally, laboratory yeasts and the wine yeasts, in addition to lacking STR XIII-R, LgS-AMD1, LgS-TYP, STR XIV-L and LgS-TRR also lack LgS-Hypo. These groupings may provide insight into the state of evolution and domestication of *S. cerevisiae* strains.

SUPPLEMENTARY DATA

Supplementary data is available at FEMSYR online.

FUNDING

This research was supported by a 1592 PhD studentship grant to CM from Trinity College Dublin and a European Research Council Advanced Grant no. 268893 to KW.

Conflict of interest statement. None declared.

REFERENCES

Akao T, Yashiro I, Hosoyama A, *et al.* Whole-genome sequencing of sake yeast *Saccharomyces cerevisiae* Kyokai no. 7. *DNA Res* 2011;**18**:423–34.

Argueso JL, Carazzolle MF, Mieczkowski PA, *et al.* Genome structure of a *Saccharomyces cerevisiae* strain widely used in bioethanol production. *Genome Res* 2009;**19**:2258–70.

Beggs S, James TC, Bond U. The PolyA tail length of yeast histone mRNAs varies during the cell cycle and is influenced by Sen1p and Rrp6p. *Nucleic Acids Res* 2012;**40**:2700–11.

Bergstrom A, Simpson JT, Salinas F, *et al.* A high-definition view of functional genetic variation from natural yeast genomes. *Mol Biol Evol* 2014;**31**:872–88.

Bing J, Han PJ, Liu WQ, *et al.* Evidence for a Far East Asian origin of lager beer yeast. *Curr Biol* 2014;**24**:R380–1.

Bond U. The genomes of lager yeasts. *Adv Appl Microbiol* 2009;**69**:159–82.

Bond U, Neal C, Donnelly D, *et al.* Aneuploidy and copy number breakpoints in the genome of lager yeasts mapped by microarray hybridisation. *Curr Genet* 2004;**45**:360–70.

Borneman AR, Forgan AH, Pretorius IS, *et al.* Comparative genome analysis of a *Saccharomyces cerevisiae* wine strain. *FEMS Yeast Res* 2008;**8**:1185–95.

Brown CA, Murray AW, Verstrepen KJ. Rapid expansion and functional divergence of subtelomeric gene families in yeasts. *Curr Biol* 2010;**20**:895–903.

Cliften P, Sudarsanam P, Desikan A, *et al.* Finding functional features in Saccharomyces genomes by phylogenetic footprinting. *Science* 2003;**301**:71–6.

Darling S, Larsen PO. Saccharopine, a new amino acid in Baker's and Brewer's yeast: I. Isolation and properties. *Acta Chem Scand* 1961;**15**:743–9.

Doniger SW, Kim HS, Swain D, *et al.* A catalog of neutral and deleterious polymorphism in yeast. *PLoS Genet* 2008;**4**:e1000183.

Dowell RD, Ryan O, Jansen A, *et al.* Genotype to phenotype: a complex problem. *Science* 2010;**328**:469.

Dunn B, Sherlock G. Reconstruction of the genome origins and evolution of the hybrid lager yeast Saccharomyces pastorianus. *Genome Res* 2008;**18**:1610–23.

Dunn B, Richter C, Kvitek DJ, *et al.* Analysis of the *Saccharomyces cerevisiae* pan-genome reveals a pool of copy number variants distributed in diverse yeast strains from differing industrial environments. *Genome Res* 2012;**22**:908–24.

Gibson BR, Storgards E, Krogerus K, *et al.* Comparative physiology and fermentation performance of Saaz and Frohberg lager yeast strains and the parental species *Saccharomyces eubayanus*. *Yeast* 2013;**30**:255–66.

Gilliland RB. The flocculation of brewing yeasts. *Brewers Guild J* 1955;**5**:246–68.

Hewitt SK, Donaldson IJ, Lovell SC, *et al.* Sequencing and characterisation of rearrangements in three S. pastorianus strains reveals the presence of chimeric genes and gives evidence of breakpoint reuse. *PLoS One* 2014;**9**:e92203.

James TC, Campbell SG, Bond UM. Comparative analysis of global gene expression in lager and laboratory yeast strains grown in wort. *Proc IEEE* 2002;**90**:1887–99.

James TC, Usher J, Campbell S, *et al.* Lager yeasts possess dynamic genomes that undergo rearrangements and gene amplification in response to stress. *Curr Genet* 2008;**53**:139–52.

Kellis M, Patterson N, Endrizzi M, *et al.* Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* 2003;**423**:241–54.

Legras JL, Merdinoglu D, Cornuet JM, *et al.* Bread, beer and wine: *Saccharomyces cerevisiae* diversity reflects human history. *Mol Ecol* 2007;**16**:2091–102.

Libkind D, Hittinger CT, Valerio E, *et al.* Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. *P Natl Acad Sci USA* 2011;**108**:14539–44.

Liti G, Carter DM, Moses AM, *et al.* Population genomics of domestic and wild yeasts. *Nature* 2009;**458**:337–41.

Liti G, Nguyen Ba AN, Blythe M, *et al.* High quality de novo sequencing and assembly of the Saccharomyces arboricolus genome. *BMC Genomics* 2013;**14**:69.

Nakao Y, Kanamori T, Itoh T, *et al.* Genome sequence of the lager brewing yeast, an interspecies hybrid. *DNA Res* 2009;**16**:115–29.

Nilsson-Tillgren T, Gjermansen C, Kielland-Brandt MC, *et al.* Genetic differences between Saccharomyces carlsbergensis and S. cerevisiae. Analysis of chromosome III by single chromosome transfer. *Carlsberg Res Commun* 1981;**46**:65–76.

Novo M, Bigey F, Beyne E, *et al.* Eukaryote-to-eukaryote gene transfer events revealed by the genome sequence of the wine yeast *Saccharomyces cerevisiae* EC1118. *P Natl Acad Sci USA* 2009;**106**:16333–8.

Pavsler A, Buiatti S. Non-lager Beer. In: Preedy VR (ed). *Beer in Health and Disease Prevention*. San Diego: Academic Press, 2009, 17–30.

Peris D, Sylvester K, Libkind D, *et al.* Population structure and reticulate evolution of *Saccharomyces eubayanus* and its lager-brewing hybrids. *Mol Ecol* 2014;**23**:2031–45.

Rainieri S, Kodama Y, Kaneko Y, *et al.* Pure and mixed genetic lines of *Saccharomyces bayanus* and *Saccharomyces pastorianus* and their contribution to the lager brewing strain genome. *Appl Environ Microb* 2006;**72**:3968–74.

Ralser M, Kuhl H, Ralser M, *et al.* The *Saccharomyces cerevisiae* W303-K6001 cross-platform genome sequence: insights into ancestry and physiology of a laboratory mutt. *Open Biol* 2012;**2**:120093.

Tamura K, Peterson D, Peterson N, *et al.* MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 2011;**28**:2731–9.

Usher J, Bond U. Recombination between homoeologous chromosomes of lager yeasts leads to loss of function of the hybrid GPH1 gene. *Appl Environ Microb* 2009;**75**:4573–9.

Walther A, Hesselbart A, Wendland J. Genome sequence of *Saccharomyces carlsbergensis*, the world's first pure culture lager yeast. *G3 (Bethesda)* 2014;**4**:783–93.

Wei W, McCusker JH, Hyman RW, *et al.* Genome sequencing and comparative analysis of *Saccharomyces cerevisiae* strain YJM789. *P Natl Acad Sci USA* 2007;**104**:12825–30.

Wunderlich S, Back W. Overview of manufacturing beer: Ingredients, processes, and quality criteria. In: Preedy VR (ed). *Beer in Health and Disease Prevention*. San Diego: Academic Press, 2009, 1–16.

Zhu YO, Siegal ML, Hall DW, *et al.* Precise estimates of mutation rate and spectrum in yeast. *P Natl Acad Sci USA* 2014;**111**:E2310–8.