Chapter 7

Principles and Applications of Genomics and Proteomics in the Analysis of Industrial Yeast Strains

URSULA BOND[1] AND ANDERS BLOMBERG[2]

[1]*Moyne Institute for Preventive Medicine, Microbiology Department, University of Dublin, Trinity College, Dublin 2, Ireland*
(e-mail: ubond@tcd.ie)
[2]*Department of Cell and Molecular Biology, Lundberg Laboratory, Göteborg University, 413 90 Göteborg, Sweden*
(e-mail: anders.blomberg@gmm.gu.se)

7.1 Introduction

Yeasts are presumably the oldest commercially cultured microorganisms and are widely used in the food and beverage industries. *Saccharomyces cerevisiae* belongs to the phylum *Hemiascomycetes*, and at least 1,000 separate strains of *S. cerevisiae* are currently being used in the baking, brewing, distilling and wine-making industries with at least two million tons of yeasts produced per year in the brewing and baking industries alone. Industrial yeast strains can be obtained from a number of repositories, such as the National Collection of Yeasts Culture (http://www.ncyc.co.uk/), The Culture Collection of Yeasts (http://www.chem.sk/yeast/culture_collection_of_yeasts.htm), The European Culture Collections' Organization (ECCO; http://www.eccosite.org/) and the Collection de Levures d'Interêt Biotechnology (http://www.inra.fr/Internet/), to name but a few. The high fermentative capacity of yeasts, together with their ability to withstand the extreme environmental conditions experienced during industrial fermentations, has led to the selection of strains with unique characteristics. In this chapter we focus on the recent use of genomics and proteomics approaches to gain an understanding of the nature of the genomes of industrial strains of yeasts, the expression of genes within these genomes and their final proteome complement.

7.2 DNA Sequencing of Yeast Genomes

To completely understand the molecular and physiological composition of any organism, it is essential to have the complete DNA sequence of its genome. This systematic approach, at a minimum, allows the researcher to determine the number of genes encoding functional proteins and provides opportunities for the theoretical

The Yeast Handbook
Amparo Querol, Graham H. Fleet (Eds.):
Yeasts in Food and Beverages

and experimental analysis of all these genes. In this section we outline the background to the sequencing of the haploid genome of the yeast *S. cerevisiae* and show how this endeavour acts as a paradigm for the analysis of the more complex genomes of the industrial strains of yeasts.

The establishment of the *Saccharomyces* Genome Database (SGD; http://www.yeastgenome.org) has been an important hub in the dissemination of information regarding the yeast genome and allows one to access the DNA sequence information of all open reading frames (ORFs). Additionally, the SGD provides links to other web-based databases such as GenBank (DNA) and GenPept (protein) at the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov), Munich Information Centre for Protein Sequences (http://mips.gsf.de/genre/proj/yeast/), Protein Information Resources (http://pir.georgetown.edu/) and SwissProt (http://us.expasy.org/sprot/), to name but a few. Links to gene expression and other functional genomics databases are also available which provide the possibility for direct comparisons of the gene expression profiles for the entire genome under a variety of environmental conditions and in a variety of genetic backgrounds (see later). Presently, the Yeast Genome Sequencing Project and the subsequent annotation of the sequences into a user-friendly database at the SGD is the prototype for the sequencing and annotation of other genomes.

7.2.1 Sequencing of the Genome of *Saccharomyces cerevisiae*

In order to obtain the entire genome sequence of an organism, it is essential to first create a complete representative genomic library. In the early 1980s, the pioneering work of Burke and Olsen allowed the cloning of large genomic fragments in the order of several 100 kb in length into yeast artificial chromosomes (YACs) (Burke et al. 1987). YAC plasmids contain both centromere and telomere sequences, thus allowing the plasmid and the inserted DNA to replicate like a chromosome. Yeast genomic libraries were also prepared in cosmids which are replicating plasmids containing the cos sites of the bacteriophage λ. Genomic fragments as large as 45 kbp can be cloned into cosmids. By sequencing and examining overlapping sequences, of YAC or cosmid clones, a physical map of the yeast genome was created (Cherry et al. 1997). This map with its linked library of DNA clones provided the starting point for the yeast genome sequencing project.

The DNA sequence of the *S. cerevisiae* genome was completed in 1996 (Goffeau et al. 1996). This was achieved through the cooperation of over 600 scientists from Europe, the USA and Japan using automated robotic machines to sequence DNA from random YAC or cosmid clones. Adding to this was DNA sequence information emerging from a "network" of small to medium-sized yeast laboratories which also led the interpretation and verification of the DNA sequences emanating from the DNA sequencing centres. From this information, it was possible to deduce the organisation of the *Saccharomyces* genome, which was shown to be composed of twelve million base pairs arranged on 16 chromosomes (Mewes et al. 1997). Originally, 6,275 theoretical ORFs containing greater than 100 amino acids were identified in the genome sequence. Following subsequent reanalysis of the data and the inclusion of additional information from the scientific literature, this number

was revised downward as of May 2002 to 6,062, of which 3,966 represented ORFs previously identified by genetic analysis or by the presence of structurally and functionally related orthologues in other species and 2,096 genes of unknown function. The estimate for the number of protein coding genes has more recently been further refined and the current estimate at the SGD (as of January 2005) is 5,798. In addition to ORFs, sequences corresponding to non-protein coding genes such as those coding for ribosomal RNA, transfer RNAs (tRNAs), small nuclear RNAs, small nucleolar RNAs, non-coding RNAs, transposable elements (Ty) and long terminal repeats (LTRs) have been identified in the genome.

7.2.2 Genome Sequencing of Other Yeast Species

Since the initial sequencing of the *S. cerevisiae* genome, the DNA sequences of a number of other members of the *Saccharomyces sensu stricto* group have been completed. A comparative analysis of the genomic sequences of *S. bayanus*, *S. mikatae* and *S. paradoxus*, which are separated from *S. cerevisiae* by approximately five million to twenty million years of evolution, confirms the organisation of ORFs onto 16 chromosomes with an average genome size of approximately 11.5×10^6–12×10^6 bp for this genera (Kellis et al. 2003). The four genomes show a high degree of conservation of synteny with only 1.3% of sites of insertions or deletions falling within protein coding regions. The 32 telomeric and subtelomeric regions of the 16 chromosomes represent the regions of greatest sequence diversity. By using a reading frame conservation (RFC) test, together with manual inspection of dubious ORFs for all initially predicted 6,275 theoretical ORFs from the *S. cerevisiae* genome, Kellis et al. (2003) identified 5,458 ORF orthologs in all four species, thus leading to a re-evaluation of the number of true ORFs in the *S. cerevisiae* genome.The comparative genomic analysis from multiples members of the *Saccharomyces sensu stricto* group, in addition to defining the number of true ORFs in these species, has allowed a comprehensive analysis of the rate of evolution of these genomes and the identification of ORFs unique to each species. Additionally, the genome alignments of all four species has allowed the identification of regulatory elements in the intergenic regions of the genomes. The sequencing of more distantly-related yeasts such as *S. kluyveri* (Cliften et al. 2003; Kellis et al. 2003), *Candida glabrata*, *Kluyveromyces lactis*, *Debaryomyces hansenii*, *Yarrowia lipolytica* (Dujon et al. 2004), *Ashbya gossypii* (Dietrich et al. 2004) and *Kluyveromyces waltii* (Kellis et al. 2004) has added to our understanding of the evolution of the genomes of the *Hemiascomycetes*.

7.3 Whole Genome Approaches to the Characterisation of Industrial Strains of Yeasts

The complete genome sequence is an important resource allowing scientists to examine the physiology and evolution of related organisms. It is also the prerequisite for some of the developed genome-wide techniques that have revolutionised the way biological systems are studied. In this section we outline some of the basics of these techniques aiming at the analyses of the complete complement of transcripts (the transcriptome) and proteins (the proteome) in a cell.

7.3.1 Microarray Technology for Genome and Transcriptome Analysis

Microarray technology grew out of the complete DNA sequencing of the haploid *S. cerevisiae* genome and involves robotic application of DNA, representing each gene in the genome, to glass slides, silicone or nylon membranes. The DNAs are arrayed in an orderly fashion to allow easy identification of genes within the array (Chittur 2004; Epstein and Butow 2000; Gerhold et al. 1999; Hardiman 2004). Two types of DNA microarray chips are currently used. Firstly, DNA sequences (500–5,000-bases long) representing individual ORFs are PCR-amplified using DNA oligos specific to each gene. These DNAs are robotically "spotted" onto the glass slides or nylon membranes. The second method generates arrays of oligonucleotides (20–80-mer oligos) representing each gene. These oligonucleotides can be directly synthesised using photolithographic techniques in situ or by conventional synthesis followed by on-chip immobilisation. A variation on this second approach is to include internal controls on the microarrays in which single nucleotide mismatches of each oligonucleotide are included on the chips to allow quantification of the specificity of hybridisation to a given probe. The prototype oligonucleotide microarrays were developed at Affymetrix, which sells its products under the GeneChip trademark.

The arraying of the whole genome, representing each individual gene, on a single matrix, allows the simultaneous analysis of the complete messenger RNA (mRNA) profile (transcriptome) of an organism in a single experiment. Relative steady-state levels of mRNAs are normally examined temporally or spatially under experimental conditions where perturbations from the normal growing conditions are imposed. These perturbation may include changes in genetic background, environmental changes or pharmacological changes, to name but a few. This is achieved through the principles of nucleic acid hybridisation. Briefly, heat denatured double-stranded DNA or single-stranded DNA on the microarray chips is incubated with a labelled probe. For transcriptome analysis, the probes are prepared by first converting RNA to complementary DNAs (cDNAs) by random priming using the enzyme reverse transcriptase. This cDNA probe represents a "snapshot" of the total pool of mRNAs present in the cell under a specific set of experimental conditions. A similar cDNA pool is prepared from RNA extracted from cells grown under "control" conditions. The cDNA probes are differentially labelled by incorporating the fluorescently tagged nucleotides Cy3 (green; 635 nm; control) and Cy5 (red; 532 nm; experimental), respectively, during the reverse transcription reaction. The differentially labelled cDNAs are then mixed and added to the microarray chip, where they compete for hybridisation to the DNA sequences on the chip. The extent of hybridisation is directly proportional to the amount of cDNA, representing a specific mRNA in the sample, that is complementary to a given DNA sequence on the chip and the degree of sequence homology between the two sequences.

Hybridisations are normally carried out in a solution containing 5X SSC (0.6 M sodium chloride, 0.06 M sodium citrate), 7% sodium dodecyl sulfate (SDS) and 50 mM sodium phosphate, pH 7.0. Blocking reagents such as Denhardt's solution [0.02% poly(vinylpyrrollidone), 0.02% bovine serum albumin, 0.002% Ficoll 400] or commercially available blocking reagents are added to the hybridisation solution to

increase the effective concentration of the probe and to reduce non-specific hybridisation. Hybridisations are normally carried out at 68°C or at a temperature below the melting temperature of the probe. Under these hybridisation conditions the labelled probe will bind to its complementary sequence on the DNA chip. The stringency of hybridisation can be varied by changing the temperature of hybridisation or the composition of the hybridisation solution. Following hybridisation, the filters are washed to remove unhybridised probe. Again the stringency of hybridisation can be adjusted here by varying the washing conditions; high-stringency washing (0.5X SSC, 0.1% SDS; 68°C) will allow only completely identical sequences to hybridise, while low-stringency conditions (2X SSC, 0.1% SDS; 68°C or at a lower temperature) will allow hybridisation between DNAs containing mismatches.

The hybridised DNA can then be measured using a fluorescence scanner. The readout from the fluorescence detector is then analysed using programs such as Genepix Pro or ScanAlyze and is expressed as pixels of green or red fluorescent light per square millimetre for each gene. Following normalisation of the data to correct for different efficiencies of labelling and corrections for size of the spot, subtraction of background and removal of spurious readings, the data can be directly fed into spreadsheet programs. The data are presented as a normalised linear ratio or a normalised $\log_2$ ratio of red-to-green fluorescence. The red-to-green fluorescence ratio gives a direct measurement of the relative proportions of RNA (cDNA) in the starting samples. A $\log_2$ ratio of greater than zero indicates a higher level of RNA (cDNA) or DNA in the experimental sample compared with the control sample, while a $\log_2$ ratio of less than zero indicates the opposite.

The red-to-green fluorescence ratios can then be fed directly into clustering programs such as Cluster (http://rana.lbl.gov/eisensoftware.htm) (Eisen et al. 1998). Clustering programs use hierarchical or *K*-means algorithms as a means of identifying and correlating patterns of gene expression and can be used to group together, into expression classes, genes showing similar gene expression patterns. Clustering programs also allow data from multiple microarray experiments to be analysed simultaneously. The cluster output can be viewed in TreeView (http://rana.lbl.gov/eisensoftware.htm) as a colour-coded graphical representation of expression profiles at a glance (Fig. 7.1).

Microarray technology has been exploited to generate a vast amount of data examining the gene expression patterns of *S. cerevisiae* under a variety of experimental conditions. The majority of these can be accessed through the SGD or directly at the site Yeast Microarray Global viewer (yMGV; www.transcriptome.ens.fr/). The latter site contains data from 1,544 experiments mainly showing gene expression patterns for the haploid yeast *S. cerevisiae*. Comparison of gene expression data generated from different sources is often hampered by differing experimental parameters being examined. However, as mentioned before, programs such as Cluster allow the side-by-side clustering of gene expression patterns from any number of differing sources and will reveal overall similarities and differences in the patterns. The yMGV site provides a graphical representation of gene expression variations for each published genome-wide experiment. Additionally, one can examine the effects of experimental conditions on one or a group of genes and identify groups of genes sharing similar transcription profiles in a defined subset of experiments.

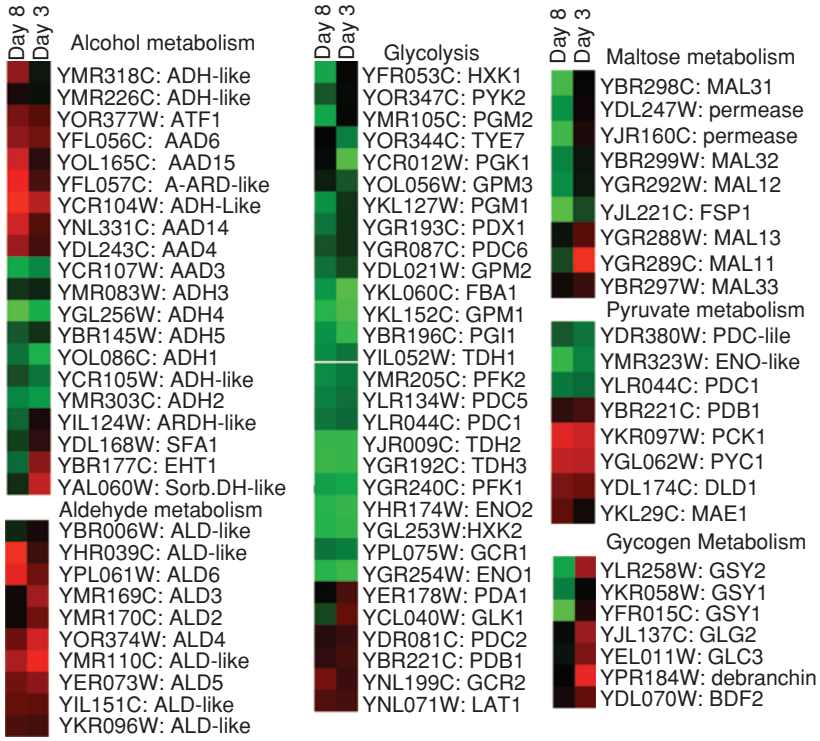

Fig. 7.1. TreeView display showing expression profiles of genes encoding proteins involved in carbohydrate metabolism. Log_2 ratios of transcript levels on days 3 and 8 of a brewing fermentation with a lager strain of a yeast relative to the levels on day 1 were obtained from a microarray analysis (James et al. 2003). Genes were first grouped together by cellular process and then clustered to reveal transcripts showing similar levels of expression (log_2 ratios) using the program Cluster. The range and intensity of colours from *red* to *green* represents a continuum of highest levels of induction (*red*) or repression (*green*)

One of the most useful datasets is that of gene expression patterns of *S. cerevisiae* under a variety of environmental conditions experienced by yeasts, such as heat and cold shock, amino acid starvation, nitrogen depletion, and during the exponential and stationary growth phase (Gasch et al. 2000).

7.3.2 Technologies for Proteome Analysis

The mRNA expression changes estimated by microarray analysis should ideally reflect the change in the amount of protein under the same experimental conditions. However, in many instances this is at its best a good approximation and rather large discrepancies are revealed. Additionally, transcriptome analysis does not reflect the complex myriad of post-translational features, like protein modifications

(e.g. phosphorylation, N-terminal acetylation or ribosylation), protein association with cofactors (e.g. NADH or zinc), protein complex formation (e.g. the ribosome contains almost 100 components), protein localisation (e.g. into mitochondria or the nucleus) and protein degradation (e.g. ubiquitination and breakdown via the proteasome), that eventually define the active proteome component of the cell. Ideally all these levels of complexity should be examined for a full understanding of protein activity in the cell and hence the system under study. This, in short, defines the great challenge in proteomics. In addition, the quantitative range of proteins in the cell is huge; recent estimates indicate that protein abundance in yeasts covers roughly 5 orders of magnitude (Ghaemmaghami et al. 2003), which of course adds to the technological challenge. Proteomics on *S. cerevisiae* currently includes a plethora of techniques, where many of the methodologies rely on the fact that the full genome sequence is available. In the analysis of industrial yeast strains where the full genome has not yet been released, the most fundamental proteomics approach with electrophoretic separation of cell extracts, subsequent image analysis and protein identification has so far been applied. However, the application of some large-scale techniques in proteomics, presently only within reach for the analysis of laboratory *S. cerevisiae* strains, will most likely in the near future also be applied for the analysis of industrial strains (Sect. 7.6).

7.3.2.1 Two-Dimensional Polyacrylamide Gel Electrophoresis

The standard methodology in proteomics has been the combination of two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) separation and quantification with mass spectrometry (MS) based identification of resolved proteins. Separating/arraying all proteins in a yeast cell provides a substantial experimental challenge. The technique applied must deal with the great qualitative and quantitative complexity and should be able to do so with a reasonably high throughput of samples. The technique most frequently used to obtain high-resolution separation of proteins is 2D-PAGE, which resolves proteins in two consecutive steps that separate on the basis of independent protein-specific properties, i.e. the isoelectric point (pI) and the molecular weight (M_r). The principle is simple and elegant and when first applied in the mid-1970s it was a major breakthrough in the molecular global analysis of biological samples (Klose 1975; O'Farrell 1975). The major procedure has not changed much over the years; however, substantial refinements in chemistry and hardware have made the currently applied technique vastly superior to the initial setup.

Isoelectric focusing of native proteins is well established and has been extensively used. To increase the resolving power, separation under denaturing conditions by adding high amounts of urea (about 9 M) and 1–4% of a non-ionic detergent was developed. The denaturing capacity is increased even further by inclusion of thiourea in addition to urea as a chaotroph, and the use of zwitterionic amphiphilic compounds (e.g. 3-[3-(Cholamidopropyl)dimethylammonio]-1-propanesulfonate, CHAPS, or SB 3–10) (Rabilloud et al. 1997). This procedure is particularly useful for the resolution of problematic proteins (Rabilloud 1998), but has become the standard procedure in many laboratories.

Separation based on different protein pIs was initially achieved by the inclusion of carrier ampholytes, molecules that in an electric field will generate a pH gradient in which proteins migrate until they reach the pH where their net charge is zero. However, when applying a wide-range ampholyte mixture, such as one with pH 3–10, the separation will not cover this wide a pH range at the end of the run; the final gradient will cover not more than about pH 4–6.5 (O'Farrell et al. 1977). The consequence is, of course, that proteins with pI values outside the produced pH range of the gradient will not focus. One solution to this problem is to pursue isoelectric focusing for shorter times, leading to a gradient with greater pH coverage (O'Farrell et al. 1977). This technique is then a non-equilibrium one, and proteins will, in general, not reach their true pI value. These non-equilibrium pH gradient gel electrophoresis (NEPHGE) gels have been applied to the separation of yeast samples, especially in the past (Bataillé et al. 1988; Boucherie et al. 1995; Brousse et al. 1985; Iida and Yahara 1984; Shin et al. 1987). The main problem with the NEPHGE procedure is reproducibility and standardisation.

The technical solution to at least some of the problems mentioned before for ampholytes came with the invention and application of immobilines (Bjellqvist et al. 1982; Görg et al. 1988, 1999). Immobilines produce a pH gradient that is generated by the covalent anchoring of the immobilines to the polymer matrix. The consequence is that the gradient will be stable and not collapse during prolonged running time. In addition, shallow (zooming in on narrower pH regions) or very broad gradients can be formed, and even the use of sigmoid-shaped non-linear pH 3–10 gradients for optimal protein pattern spread has been applied for the analysis of yeasts (Muller et al. 1999; Norbeck and Blomberg 1997, 2000). In addition, with the development of more alkaline pH gradients (up to pH 12.5) the 2D system can resolve very alkaline proteins like histones and many ribosomal proteins (Wildgruber et al. 2002). The use of shallower gradients, where better physical separation is obtained between proteins with closely matching pI values, can be very useful, and will minimise the shielding effect on low-abundance proteins from abundant neighbours (Wildgruber et al. 2000).

For the 2D separation there is presently no good system for the parallel separation of proteins in the range 1–600 kDa, which is the size range provided by the yeast proteome. Size differences can be dealt with but only if we apply slightly different techniques for different size classes: i.e. altering the concentration of the polyacrylamide matrix will enable better separation of small proteins (high total percentage of acrylamide, %T) or large proteins (low %T) (Garrels 1979). However, choosing either of these extremes will of course compromise the resolution of the opposite size class. In practice this means that more gels have to be run and fewer samples will be analysed. Most groups have thus adopted some intermediate concentration (10–12%T) as the standard, thus accepting minor losses in resolution of extremes in the process.

7.3.2.2 Post-Separation Analyses – Image Analysis

Visualisation of 2D-PAGE-separated proteins can be performed by different means, and when dealing with the industrial-scale visualisation this is achieved by the use of

some sensitive stain. Silver deposits at protein spots can detect low nanogram quantities of protein; however, different silver staining protocols are more or less sensitive and some are more useful for quantitative purposes (Blomberg 2002; Rabilloud 1992). The signal-response curve for silver staining is linear for all protocols over a rather narrow dynamic range, and proper quantification over a wider range can only be achieved by the use of calibration strips. A recent development of great use in protein quantitation is the use of fluorescent dyes. The currently most frequently used dye is Sypro Ruby, which has been documented to generate a linear response over some orders of magnitude. This staining procedure is, however, a bit less sensitive than silver staining (Rabilloud et al. 2001; Blomberg et al., unpublished data). In Fig. 7.2a the 2D pattern of lager yeast strain CMBS33 from Sypro Ruby staining can be observed from cells grown under laboratory conditions in synthetic defined medium containing glucose. Recently dual labelling with covalent linkage of fluorescent dyes prior to electrophoretic separation was also applied, i.e. fluorescence-based multiplexed proteomics (Patton and Beechem 2002). Different samples are labelled with dyes with different spectral properties, mixed and separated by 2D-PAGE. The resulting gel is scanned in different spectral windows and the individual quantities estimated. The main advantage of this methodology is that between-gel variation is cancelled, which makes sample comparison and spot matching more straightforward. However, if more than two or three samples are to be analysed, which is usually the case in proteome analysis if statistics are to be applied (which it should), the problem still persists with gel-pattern matching. Any of these means for

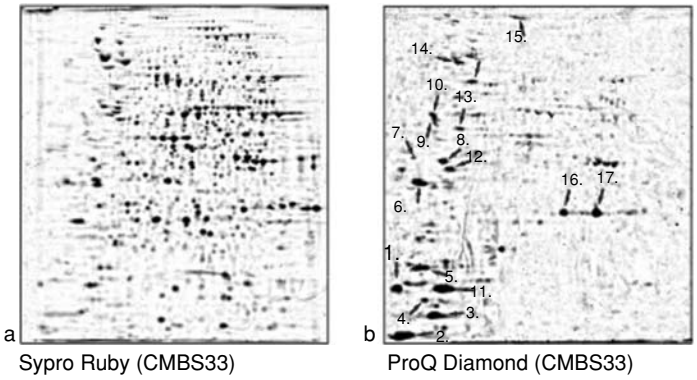

Fig. 7.2. Differential staining with Sypro Ruby and ProQ diamond of 2D-separated proteins from a lager yeast strain. The lager strain CMBS33 was grown in synthetic defined medium with glucose as the carbon and energy source and harvested in the mid-exponential growth phase. (**a**) Total proteins were separated using a wide-range non-linear pH 3–10 gradient in the first dimension and were visualised by the fluorescent dye Sypro Ruby. (**b**) The same gel as in **a** was subsequently stained with the phosphoprotein specific stain ProQ diamond to visualise proteins in the pattern that were modified by phosphorylation (about 25 proteins displayed significant phosphorylation specific staining). The acidic side in the first dimension is to the *left*. (Adopted from R. Caesar, E. Pettersson, J. Gustafsson, A. Hashemi, A. Blomberg, unpublished results)

visualisation can be utilised for protein expression analysis of samples cultivated on an industrial scale. However, the most precise quantitative method is not really applicable to large-scale fermentations – isotopic labelling of proteins during cultivation by addition of a radioactive amino acid, usually ^{35}S-methionine. After separation the gels are dried and exposed to image plates that have a higher sensitivity and a wider linear range of response compared with those of ordinary X-ray film (Blomberg 2002).

Protein stains have also been developed to specifically capture protein features. One good example of this is the recently developed stain ProQ diamond, where gels are fluorescently stained and imaged to reveal phosphorylation levels using this fluorescent phosphosensor dye (Schulenberg et al. 2003). The initial staining step is subsequently followed by staining and imaging to reveal general protein expression levels using a total protein indicator such as Sypro Ruby. ProQ diamond specific proteins can subsequently be identified by peptide mass fingerprinting. Figure 7.2b displays the phosphorylated proteins indicated by ProQ diamond staining in the lager strain CMBS33. Some of the phopsphorylated proteins were identified earlier as phosphoproteins in *S. cerevisiae*, while others are novel phosphorylation targets (R. Caesar, E. Pettersson, J. Gustafsson, A. Hashemi, A. Blomberg, unpublished results).

Generated images can be analysed for quantitative changes in protein spot volume. For this task there are currently a number of different commercial software packages. All of them are based on some procedure for background subtraction, smoothing and subsequent spot identification and quantification (Blomberg 2002). Different software packages apply different strategies in the analysis of spot volumes. Some use mathematical modelling based on spot width and peak height, while others identify the outer borders of the spot and sum all the pixel values; however, the quantitative results are usually more or less similar, at least for well-resolved spots. The final analysis is to compare the quantitative values of individual protein spots between different samples. This matching process is supposed to be automated in many of the software packages, but still much time is spent in manual editing and matching of different images. Currently procedures are being developed that hopefully will lead to more accurate automatic handling of large collections of image data; these are based on image warping prior to matching (Gustafsson et al. 2002; Veeser et al. 2001).

7.3.2.3 Protein Identification: a Historical Survey

The 2D map obtained is of rather limited biological value if resolved protein spots are not identified. This was initially performed by co-migration with purified preparations of enzymes. Applying this procedure, Calvin McLaughlin and colleagues were able to produce the first annotated 2D map of *S. cerevisiae*, which in 1978 contained identity for 17 proteins (Elliott and McLaughlin 1978). The proteins identified were mostly very abundant and involved in glycolysis, like hexokinase, triosephosphate isomerase and enolase.

Another early procedure for the identification of proteins in the 2D patterns was by the use of antibodies. Proteins were separated and subsequently blotted onto

filters where polyclonal antibodies were used for identification. This approach was first attempted for yeasts by the Bordeaux group lead by Helian Boucherie. Antibodies were produced against purified preparations of enolase and glyceraldehyde-3-phosphate dehydrogenase, and these antibodies were used for protein identification on 2D gels (Brousse et al. 1985) and were also applied in the first analysis of some wine yeasts (Brousse et al. 1985). The antibodies detected both Eno1p and Eno2p as well as Tdh2p and Tdh3p (as we know them today). This exemplifies a weak point when using polyclonal antibodies; they are prone to cross-reactivity. The future use of antibodies in protein characterisation would most likely be found in the large scale analysis of diverse protein modifications or in the analysis of low-abundance proteins.

The advent of molecular biology techniques provided novel avenues for protein identification since genes could be cloned and overexpressed by the use of multicopy plasmids. This strategy was explored for the identification of the location of *PGI* and *PGM* encoded polypeptides in the 2D pattern (Bataillé et al. 1987), as greatly enlarged spots. A variation on this theme would be the 2D analysis of a gene knockout, where the gene product in question is scored as a missing 2D spot in the mutant.

A cheap, rapid, yet efficient way of identifying proteins on a laboratory scale is based on the determination of their amino acid ratios (Garrels et al. 1994; Maillet et al. 1996). Specific double-labelling with ^{3}H and ^{14}C or ^{35}S-labelled amino acids, chosen among those that are specifically incorporated into proteins without interconversion, allowed an accurate measurement of different amino acid ratios for roughly 200 proteins. This double-labelling technique is particularly suited for the analysis of proteins that are only transiently expressed; these proteins will be difficult to analyse by other means since in many instances they will be present at very low levels.

The first technique for direct identification of resolved proteins was microsequencing. In principle, microsequencing works extremely well in most cases; however, a number of problems have also been encountered, e.g. proteins frequently carry modifications in their N-terminus, which blocked and stopped the sequencing reaction. Thus, novel procedures for generating material to be analysed in the microsequencer were developed. The most appealing and experimentally straightforward method was in-gel digestion with a specific protease, mostly trypsin, and subsequent elution and fractionation of the peptides generated (Norbeck and Blomberg 1995; Rosenfeld et al. 1992). The in-gel digestion turned out to be a major breakthrough in the analysis of proteins and is still the main method in use today for the generation of peptide cocktails from 2D-resolved proteins. A number of laboratories have successfully adopted this strategy of microsequencing of isolated peptides and used it for the annotation of their yeast map (Garrels et al. 1997; Norbeck and Blomberg 1995, 1997; Sanchez et al. 1996). Microsequencing is a very powerful technique when it comes to stringent identification of proteins, in particular for closely related proteins (Norbeck and Blomberg 1995). However, this technique also has a number of drawbacks: (1) it is rather expensive to run and maintain (e.g. chemicals have to be ultrapure), (2) the analysis time for each amino acid is about 30–45 min, at least for the somewhat older microsequencers, which results in few protein identifications per day, and (3) rather large amounts of protein are required (1–10 pmol). This

excludes the use of microsequencing in any type of really large scale programme when linked to the characterisation of hundreds of proteins resolved by 2D-PAGE.

7.3.2.4 Protein Identification: Mass Spectrometry

Currently the methodology of choice for large-scale protein identification is based on the use of mass spectrometry. Molecules to be analysed by mass spectrometry are ionised in the gas phase and are subsequently resolved in relation to their mass-to-charge ratio. Key in the protein analyses by mass spectrometers was to get large molecules like proteins and peptides into the gas phase as ions. Much effort was spent in the past to develop non-destructive "soft" ionisation methods that are compatible with studies of proteins and peptides. Two different ionisation methods were eventually developed that proved versatile, non-destructive and robust in the ionisation of a wide spectrum of peptides and proteins.

In matrix-assisted laser desorption ionisation (MALDI) a large excess of matrix material is mixed with the analyte molecule and a small volume of the mixture is placed on a metal target where it is allowed to dry and form crystals (Mann et al. 2001). Nanosecond laser pulses from nitrogen lasers are then used to irradiate the crystals under vacuum. The matrix is a small organic molecule like the commonly used compound α-cyano-4-hydroxycinnamic acid, and it is believed that the matrix molecules absorb the laser energy, resulting in desorption, and throw out of a small volume of matrix and embedded analytes. The matrix thus serves the purpose of propelling the non-volatile proteins and peptides into the gas phase. The ionisation process of a peptide competes with ionisation of all other peptides in the mixture, and this suppression phenomenon results in a context-dependent signal intensity for a particular peptide (Larsson et al. 1997). This phenomenon makes MALDI-MS non-quantitative and also explains why some peptides never "fly" in the analysis of complex peptide mixtures.

In the other ionisation method, electrospray ionisation (ESI), a low flow rate of liquid (microlitres per minute or less) is pumped through a very narrow and high electric potential needle. This procedure leads to a dispersion of the liquid into micrometre-sized charged droplets, which is called an electrospray (Mann et al. 2001). When these small droplets evaporate charge is transferred to the analyte molecules. Electrospray is regarded as a very soft ionisation method since it rarely fragments the analyte ions. ESI works well on many types of macromolecules and usually results in multiply charged larger molecules. Since the separation and analysis in the mass spectrometer is based on mass over charge, multiple charges produced by ESI bring even very large proteins into the mass/charge analytical range of most mass spectrometers. This ionisation method can be directly coupled in-line with a high-performance liquid chromatography system. This system design is rather robust since sample cleanup, separation and concentration is performed in a single in-line chromatographic step. Many of the very potent applications of ESI-MS in proteomics are currently based on in-line separation of peptides by 1D or 2D chromatography systems.

Mass spectrometers measure the mass-to-charge ratio of ions. This can be achieved by separation based on time-of-flight (TOF-MS), quadrupole electric fields

generated by metal rods (quadrupole MS), or selective ejection of ions from a 3D trapping field (ion-trap MS) (Mann et al. 2001). When structural information is to be gained, e.g. in peptide sequencing, two different steps of *m*/*z* analysis are performed in tandem MS (MS/MS) with some kind of fragmentation procedure, in a collision-induced dissociation (CID) chamber, placed in between. The MS/MS analysis can in principle be performed by employing the same *m*/*z* separation principle twice (e.g. TOF-TOF) or by combining two different ones (e.g. quadrupole TOF). Both ionisation procedures described, MALDI and ESI, can be coupled to any of the *m*/*z* separation methods; however, some combinations appear to be a better match and are more frequently found in current commercial instruments (e.g. MALDI-TOF; ESI ion trap).

In the TOF-type mass spectrometers ions are accelerated to a fixed amount of kinetic energy by a strong electric field, applied some 100–500 ns after the laser pulse, and the ions travel down a flight tube. This procedure leads to molecules with low *m*/*z* values having higher velocities and getting to the detector before the ones with high *m*/*z* values. The mass accuracy of MALDI-TOF instruments is roughly 10 ppm (0.01 Da for a 1-kDa peptide). The quadrupole is a mass filter that is composed of four metallic rods to which an oscillating electric field is applied. This filter lets through only ions with a certain *m*/*z* value, with the other ones not reaching the detector. The mass-over-charge spectrum is obtained by changing the amplitude of the electric field and recording the ions that reach the detector. Most peptide sequence analyses have been performed on triple quadrupole instruments where the design is divided into three sections. Two sections that provide *m*/*z* filters and one central quadrupole section that contains the ions during fragmentation thus constitute the CID chamber. Quadrupole mass spectrometers are capable of unit mass resolution and mass accuracy of 100–500 ppm. In ion traps the ions are physically trapped in a 3D electric field. Ion traps capture the continuous beam of ions up to the limit of the maximum number of ions that can be introduced into the trap without distorting the electrical field. After capture, the ions are subjected to additional electrical fields that lead to one ion after the other being ejected from the trap. The ejected ions are detected and this produces the *m*/*z* spectrum. In cases where MS/MS analysis is to be performed with the ion trap, all except the desired ion are first ejected. After that the remaining ion is fragmented by collision to a gas and the fragments are analysed by the same procedure as before. This construction allows for multiple rounds of analysis and fragmentations, with the result that a large number of MS/MS analysis can be performed (MS^n). Ion traps are compact, robust and very versatile instruments for which the operation can be highly automated. The mass accuracy of the ion traps is slightly less than for the quadrupole instruments, which is usually not a great problem in the analysis of yeast proteins because of the relatively small genome (only roughly 6,000 proteins).

Algorithms for protein identification based on MS data has been developed and refined (Mann et al. 2001). Originally protein identification was almost exclusively performed using MALDI-TOF-MS data, where the peptide mass fingerprint generated was compared with the theoretically predicted masses of peptides for each entry in the protein database. Trypsin digestion has most frequently been used (cleavage after arginine and lysine residues), but in principle any type of specific protease

could be used in the generation of peptides. The first step in the procedure is that visualised proteins are physically cut out from preparative gels. These gel pieces are destained and washed, dehydrated and then rehydrated with trypsin solution. With the growing number of sequences in the databases, more sophisticated algorithms have been developed that generate lists of protein hits with corresponding probability scores. However, as a rule of thumb for an unambiguous identification, when high mass accuracy in the range 10–50 ppm is achieved, at least five peptide mass MALDI-MS data need to match that should cover at least 15% of the length of the protein. MS/MS data can also be used for database searches. Since sequence information is revealed in the tandem mass spectra, these searches provide higher confidence and greater discrimination. Several alternative algorithms exist, but here it suffices to mention that experimental spectra are matched against calculated fragmentation spectra for all peptides in the database. The power of the MS/MS approach truly comes into play when mixtures of proteins (or whole proteomes) are to be analysed (Washburn et al. 2001), since single peptide fragmentations can be sufficient to identify a particular protein in the sample.

7.3.2.5 Yeast Proteome Databases

Currently the most impressive and complete global analysis of the yeast proteome includes the analysis of 400 proteins corresponding to 279 different genes for the laboratory strain S288c during exponential growth in synthetic defined medium (Perrot et al. 1999). In this analysis, a combination of different techniques was used in the identification of the protein spots; genetic alterations (gene overexpression or deletion) 121 proteins, amino acid composition 114 proteins, and MS 221 proteins. Many of the most abundant proteins in the 2D pattern are involved in energy metabolism, of which proteins encoded by 48 genes have been identified. A large proportion of the proteins identified are involved in biosynthesis of either small molecules (65 different genes) or macromolecules (79 genes). Some of the metabolic pathways are almost completely covered, like purine synthesis, which was also used for the first complete study of all its components (Denis et al. 1998). However, only a small proportion (five proteins) of the subunits in the ribosome were separated and identified. A good proportion of the aminoacyl-tRNA synthetases have been identified, 13 out of the 20 believed to be present in the genome. In addition, a high number of chaperones are present; these are mostly rather dominant and are represented by a large variety of proteins and different families in yeasts. Also identified are components involved in the cellular response to growth perturbations (stress); osmotic adaptation (four proteins) and detoxification (six proteins). This data can be explored at the Bordeaux 2D database (YPM; http://www.ibgc.u-bordeaux2.fr/YPM/).

7.4 Genome Constitution of Industrial Strains of Yeasts

The genomic analysis of industrial strains of yeasts has lagged behind that of the prototype haploid *S. cerevisiae* species. This is partly due to the complex nature of their genomes. Industrial strains of yeasts, in particular those involved in the beer and wine industries, all belong to the *Saccharomyces sensu stricto* group. Currently,

this taxonomic group includes seven yeast species, *S. cerevisiae*, *S. paradoxus*, *S. pastorianus*, *S. cariocanus*, *S. kudriavzevii*, *S. bayanus* and *S. mikatae*.

Yeasts isolated from wine fermentation in which no starting inoculum is used tend to belong to the species *S. cerevisiae* or *S. bayanus*, with the majority being *S. cerevisiae* (Mortimer 2000). A small number of wine yeasts appear to be interspecies hybrids of *S. cerevisiae* and *S. bayanus*, while one particular cider yeast strain (CID1) contains DNA from three separate species (de Barros Lopes et al. 2002; Mortimer 2000). The majority (approximately 70%) of wine yeasts are prototrophic, heterozygous and homothallic and most possess diploid genomes (Mortimer 2000).

Industrial beer fermentations in modern-day breweries use two uniquely different strains of yeasts, the top-fermenting ale strains and the bottom-fermenting lager strains, the names "top" and "bottom" reflecting whether the yeast rises to the top of the fermentation or settles to the bottom at the end of fermentation (Kielland-Brandt et al. 1995; Kodama et al. 2005). Top-fermenting yeasts produce ales and closely resemble *S. cerevisiae*, while the bottom-fermenting lager strains appear to have arisen from a fusion between two yeast species, one closely resembling *S. cerevisiae*, while the other appears similar to *S. bayanus* and/or *S. pastorianus* CBS1503 (*S. monacensis*) (Kodama et al. 2005). The lager yeasts have been grouped with *S. pastorianus* (Vaughan-Martini and Martini 1987) but are commonly referred to as *S. carlsbergensis*.

The lager yeast genome is believed to have undergone a genome duplication following the species fusion, leading to an allotetraploid strain containing varying numbers of *S. cerevisiae* and non-*S. cerevisiae* chromosomes. Generally lager yeast genomes are considered to be aneuploid in nature, possessing unequal numbers of chromosomes, e.g. a strain could have one copy of chromosome III and five copies of chromosome IV. In addition, recombination between the two homeologous sets of chromosomes has occurred, generating mosaic chromosomes (Kodama et al. 2005). These complex genomes appears to confer unique selective characteristics to these strains, e.g. the ability of lager yeasts to grow at low temperatures (7–13°C) and to withstand high osmotic pressure, high hydrostatic pressure, and high ethanol and CO_2 concentrations.

7.4.1 Classical Approaches to the Analysis of Industrial Yeast Genomes

7.4.1.1 Single Chromosome Transfer of the Lager Yeast Genome

The poor sporulation ability of lager yeast strains has hampered classical genetic analysis of these polyploid strains. However, preceding the development of molecular approaches for genome analysis, studies using the technique of single chromosome transfer from lager strains to well-defined laboratory *S. cerevisiae* strains helped to elucidate the polyploid and mosaic nature of lager yeast genomes (Casey 1986; Kielland-Brandt et al. 1995; Nilsson-Tillgren et al. 1981). Using this experimental approach three chromosome types in the lager strains have been defined: (1) *S. cerevisiae* chromosomes capable of recombining with *S. cerevisiae* chromosomes, (2) non-*S. cerevisiae* chromosomes showing no recombination and (3) mosaic chromosomes where regions of the lager yeast chromosome can recombine with an

S. cerevisiae chromosome (Kielland-Brandt et al. 1995). The conclusions reached by these seminal experiments defined the complex makeup of the lager yeast genomes and have been borne out and refined by subsequent analysis using molecular and whole genome approaches (see later).

7.4.1.2 Electrophoretic Analysis of Lager and Wine Strains of Yeasts

The characterisation of the structure and composition of industrial yeast chromosomes has been greatly aided by the development of pulsed-field gel electrophoresis. Classical single (continuous) field gel electrophoresis on agarose gels in tris(hydroymethyl)aminomethane acetate buffers has an upper limit of separation of approximately 30–50 kb. Pulsed-field gel electrophoresis, in which the direction and the duration of the current are varied (Coulson et al. 1988), relieves this upper restriction and allows the separation of DNA fragments in the size range from 10 kb to 10 Mb. The earliest models of pulsed-field gel electrophoresis employed a single field inversion of the electric current: field-inversion gel electrophoresis (FIGE) (Carle et al. 1986). FIGE periodically inverts the polarity of the electrodes subjecting the DNA to a 180° reorientation. The more complex pulsed-field systems such as contour-clamped homogeneous electric field (CHEF) (Chu 1990; Chu et al. 1986) transverse alternating field electrophoresis (TAFE) (Gardiner et al. 1986) and rotating gel electrophoresis (RGE) (Anand and Southern 1990; Gemmill 1991; Southern et al. 1987) all subject the DNA to smaller changes in orientation, generally between 96 and 120°. This results in the DNA moving forward in a zigzag manner allowing for separation over a wider range of sizes.

The first yeast chromosome fingerprint of the haploid strain of *S. cerevisiae*, using FIGE (Carle and Olson 1985), confirmed the presence of 16 chromosome bands. Separation of industrial yeast chromosomes on CHEF or TAFE gels revealed a wide heterogeneity in the chromosome banding pattern; however, in each case chromosomes equivalent in size to the *S. cerevisiae* chromosomes are apparent. (Casey 1996). Strains of lager yeasts vary in the number of chromosome bands present and in subtle electrophoretic mobilities of individual chromosomes. For example, chromosome fingerprints of a number of yeast cultures from the Stroh Brewery Culture Collection distinguished two unique fingerprints of type "Tuborg" and type "Carlsberg" (Casey 1996). These two strain types differ in the presence or absence of the small chromosomes I and X. By combining the techniques of pulsed-field gel electrophoresis and Southern blotting, the identity of individual chromosome bands was deduced. In a seminal paper by Casey (1986) the presence of three types of chromosome X in *S. carlsbergensis* was revealed using a DNA probe specific to chromosome X. Likewise, by varying the stringency of the hybridisation and washing conditions, the *S. cerevisiae* like, the non-*S. cerevisiae* homeologues and mosaic chromosomes, present in the lager strains, can be distinguished.

Electrophoretic karyotyping of wine yeasts from different sources reveals that strains differ in the number, electrophoretic mobilities and band intensities of chromosomes. Up to 21 chromosome bands have been identified in certain wine yeast strains (Castrejon et al. 2004). Hybridisation of chromosomes separated by pulsed-field gel electrophoresis, using chromosome-specific DNA probes, has allowed the

identification of individual chromosomes of wine yeast (Puig et al. 2000). Furthermore, karyotyping has revealed that the genomes of wine yeasts are subject to genetic rearrangements by undergoing recombination between homologous chromosomes and between paralogous or repeated DNA sequences (Codon et al. 1998; Infante et al. 2003). The recombination events, resulting in the generation of mosaic chromosomes in lager and wine yeasts, appear to occur during mitosis rather than being meiotic in nature (Puig et al. 2000). The resulting complex genomes of lager and wine yeasts lead to poor mating ability, poor sporulation and spore viability and thus contributes to the genetic isolation of these strains.

7.4.2 Comparative Genomic Microarray Analysis

7.4.2.1 Comparative Competitive Genomic Hybridisation of Lager Strains of Yeast

While pulsed-field gel electrophoresis allows the identification of the number of individual chromosomes and chromosome heterogeneity between yeast strains, it cannot reveal any information on the actual gene makeup of these industrial yeast strains. The development of whole genome microarray technology has allowed researchers to address this issue and to ask pertinent questions regarding the biological consequences of the hybrid nature of many industrial strains of yeasts. Such questions include what is the nature of the mosaic chromosomes in industrial strains of yeasts, how many copies of *S. cerevisiae* and non-*S. cerevisiae* genes are present in the genome and what effect does the presence of multiple non-identical genes have on the final transcriptome and proteome content of the yeast?

The technique of comparative competitive genomic hybridisation (CCGH) (Watanabe et al. 2004) to microarrayed *S. cerevisiae* genes on DNA chips can be used to determine the relative copy number of *S. cerevisiae* like genes, at single-gene resolution, in industrial strains of yeasts. In this case, the genomic DNA populations from two different yeast strains are differentially labelled by random priming with Cy3- and Cy5-tagged nucleotides by PCR amplification using enzymes such as Taq polymerase and hybridisations are carried out as described in Sect. 7.3.1. In this experiment, the differentially labelled DNA samples compete for hybridisation to their complementary sequence on the DNA chips. As with microarrays using two differentially labelled cDNA pools (Sect. 7.3.1), a ratio of hybridisation, reflected in red-to-green fluorescence ratios for each gene, is obtained. The ratio of hybridisation for the two competing DNA samples will depend on (1) the degree of sequence homology between the labelled DNA probes and the *S. cerevisiae* DNA arrayed on the chip and (2) the relative abundance (copy number) of the DNA sequence in the labelled DNA sample. In the simplest example, in which differentially labelled DNAs from two haploid strains of *S. cerevisiae* are competitively hybridised to *S. cerevisiae* DNA chips, one expects the ratio of red-to-green fluorescence to be 1.0 for all genes except for individual genes which may differ in the two strains. For polyploid and, in particular, aneuploid strains of yeasts, depending on the gene copy number and the homology between the reactants, the ratio of hybridisation (red-to-green fluorescence ratio) will be greater or less

than 1. This deviation from the control mean ratio of hybridisation is indicative of a higher or a lower copy number, respectively. The caveat in using CCGH analysis for industrial strains of yeasts is that only genes with extensive homology to *S. cerevisiae* will be detected.

Using this approach, Bond et al. (2004), Kodama et al. (2005) and Infante et al. (2003) have examined the copy number of *S. cerevisiae* like genes in lager and wine yeasts. A number of interesting findings have emerged from these studies. In one of the lager yeast studies (Bond et al. 2004), competitive hybridisations were carried out for two lager strains of yeasts and the yeasts were compared with a haploid *S. cerevisiae* strain. When the ratio of hybridisation for each individual gene was arranged according to the *S. cerevisiae* gene order, it was observed that the ratios of hybridisation for genes on a given chromosome were remarkably similar. However, in eight of the 16 chromosomes, distinct "jumps" in the ratios of hybridisation occurred at discrete loci along the chromosome. Examples of two of these jumps in the ratios of hybridisations are shown in Fig. 7.3. For chromosome III (Fig. 7.3a), a distinctive jump occurs at the MAT locus, between YCR039C and YCR040C. In lager strain 6701, genes to the left of the MAT locus show a ratio of hybridisation of 0.6, while for strain CMBS, the average ratio of hybridisation is 1.0. The genes to the right of the MAT locus in both strains show an average ratio of hybridisation of 1.7. It is interesting to note that the majority of the jump locations are conserved between the two lager strains of yeasts (they are definitely different isolates) and also between a third lager strain (Bond and James, unpublished data). However, each strain also possessed its own unique jumps; for strain CMBS a unique jump is observed on chromosome VIII after YHR165C and on chromosome XIII after YMR302C, while strain 6701 has a unique jump on chromosome X after YJR009C. The simplest explanation of the ratios of hybridisation data is that the jump locations represent regions where the homeologous chromosomes have undergone inter-and/or intra-chromosomal translocations. In fact, previous studies have used CCGH to identify genomic rearrangements such as translocations, deletions and amplifications in *S. cerevisiae* strains during experimental evolution conditions (Dunham et al. 2002). While CCGH analysis can reveal relative differences in copy number between two yeast species and/or strains, it does not allow an absolute quantitative determination of gene copy number. However, the relationship between the ratios of hybridisation and the gene copy number can be determined by combining CCGH analysis with quantitative real-time PCR. The technique of real-time-PCR allows the determination of the degree of amplification of known DNA quantities. By comparing the rate of PCR amplification of ORFs displaying unique ratios of hybridisation to the rate of PCR amplification of a known single copy gene, Bond et al. (2004) have shown that there is a direct correlation between the ratio of hybridisation and gene copy number for *S. cerevisiae* like genes in lager strains of yeasts.

An analysis of chromosome XVI in two lager strains of yeasts reveals the presence of four distinct regions with unique ratios of hybridisation (Fig. 7.3b). The region immediately following the left telomere up to the ORF YPL242C shows a ratio of 1.8. This is followed by a long stretch up to YPR159W, where a ratio of 1.0–1.2 is observed. Surprisingly, the region from YPR160C to YPR190C, encompassing 30 genes, many of which are essential in the haploid *S. cerevisiae* species,

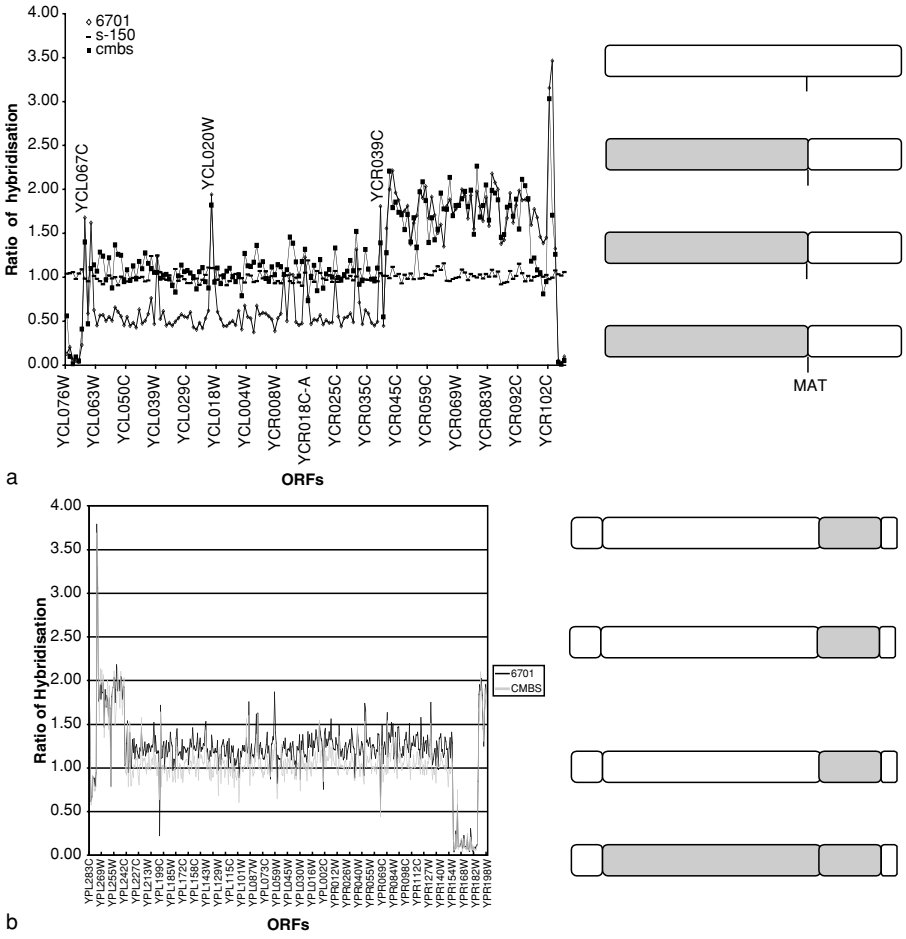

Fig. 7.3. Comparative competitive genomic hybridisation for the determination of the copy number of *S. cerevisiae* like genes in industrial yeast strains. (**a**) Ratio of hybridisation along chromosome III as determined by comparative competitive genomic hybridisation (*CCGH*) for the lager strains 6701 (*diamonds*) and CMBS (*squares*) relative to the haploid strain S-150B (*bars*). Schematic representation of chromosome III in strain 6701 as determined by quantitative-real-time PCR and CCGH. On the basis of these analyses the minimum copy number is one *S. cerevisiae* like chromosome III (*open box*) and three mosaic chromosome IIIs, where the genes to the left of the MAT locus are non *S. cerevisiae* like (*grey*) and those to the right of the MAT locus are *S. cerevisiae* like (*open boxes*). (**b**) Ratio of hybridisations along chromosome XVI for lager strains 6701 (*black line*) and CMBS (*grey line*) relative to the haploid strain S-150. Schematic representation of the minimum copy number and types of chromosome XVI in these strains as determined by CCGH and quantitative real-time PCR. *Open boxes S. cerevsiae* like, *grey boxes* non-*S. cerevisiae* like

shows a ratio of 0.2. From YPR190C to the telomere, the ratio once again rises to 1.8. The low hybridisation signal in the region YPR160C–YPR190C suggests that the *S. cerevisiae* genes in this region are absent in the lager strains or have significantly diverged from the *S. cerevisiae* sequence and most likely the region only contains non-*S. cerevisiae* gene equivalents. Southern blotting analysis of chromosome XVI, separated on CHEF gels, indicates that the latter is the case as a very weak hybridisation signal can be detected with DNA probes prepared from *S. cerevisiae* ORFs in this region (Usher and Bond, unpublished data).

Using the relationship between the ratio of hybridisation and the gene copy number one can deduce that the minimum chromosome XVI complement in these lager strains consists of three *S. cerevisiae* like chromosomes containing the non-*S. cerevisiae* region YPR160C–YPR 190C and one non-*S. cerevisiae* chromosome containing the telomeres and subtelomeric regions of an *S. cerevisiae* like chromosome (Fig. 7.3b). In a CCGH study with a number of lager strains of yeasts, Kodama et al. (2005) showed that the mosaic makeup of chromosome XVI is conserved amongst the lager strains although some variations do exist. These chromosomes most likely arose from recombination events between homeologous chromosomes. It is interesting to note that in the majority of cases, these recombination sites (as defined by the jumps in the ratios of hybridisation) are located at known sites of high genetic recombination (as in the case of the MAT locus on chromosome III) or at sites containing clusters of Ty or tRNA genes. However, the lager-specific jump sites, such as that observed at YHR165W in strain CMBS, do not contain any of these elements. These unique recombination sites may truly represent examples of adaptive evolution conferring specific selective advantages to that particular strain. Using CCGH analysis, Kodama et al. (2005) further showed that *S. pastorianus* (CBS 1538), *S. carlsbergensis* (IFO11023) and *S. monacensis* (CBS1503) lack certain *S. cerevisiae* like chromosomes. Remarkably, the *S. pastorianus* strain CBS1538 appears to lack *S. cerevisiae* like chromosomes II, III, IV, VI, VIII, XII, XV and XVI.

7.4.2.2 CCGH Analysis of Wine Yeast Genomes

Like lager yeasts, wine yeast strains show properties of aneuploidy, polyploidy and hybrid chromosomes. For example, a subgroup of wine yeasts, isolated from the flor vellum of aging sherry wines, have been shown to contain genetic heterogeneity as identified by electrophoretic karyotyping (Infante et al. 2003). Flor vellum is a biofilm that develops on the surface of sherry wines after the fermentative processes is complete. CCGH analysis of two flor yeast strains, *S. cerevisiae* var. *beticus* 11.3 and *S. cerevisiae* var. *montuliensis* 1.28, confirmed the aneuploid and hybrid nature of their chromosomes. By comparing the ratios of hybridisation for the two flor yeast strains, Infante et al. (2003) revealed that certain chromosomes, such as chromosome IV, show identical ratios of hybridisation (and therefore gene copy numbers), while chromosomes I, III, VI, X and XI show ratios of hybridisations 6 times that observed for chromosome IV. The aneuploid nature of the chromosomes is revealed by plotting the log ratios of the hybridisations between the two strains for each ORF as a function of their position on the *S. cerevisiae* chromosomes. As with the lager strains, the points at which the ratios of the hybridisations

change coincide with positions of Ty elements, LTRs and tRNA gene clusters. A number of other breakpoints correspond to positions of known meiosis-induced double-strand breaks (Gerton et al. 2000). Copy number difference between the two flor yeasts involved 38% of *S. cerevisiae* ORFs present in 116 regions of the genome.

Thus, the picture emerging for industrial yeast strains is one of aneuploidy arising from recombination events, mostly between homeologous chromosomes but also including non-homologous interchromosomal recombination events, resulting from the selective pressures experienced by these strains. This hypothesis, put forward by Infante et al. (2003), is supported by results from a number of model systems set up to examine genome evolution. In a study by Delneri et al. (2003), interspecies crosses between two different *S. mikatae* strains and an *S. cerevisiae* strain engineered to contain chromosomes that are collinear with the *S. mikatae* strain resulted in progeny whose genomes displayed widespread aneuploidy. The authors suggest that aneuploidy imposes genetic isolation on the strains, leading to stable genetic properties. In another experiment to examine the effects of selective pressures on the evolutions of genomes, Dunham et al. (2002) grew strains of *S. cerevisiae* for 100–500 generations in glucose-limited chemostats. Again using CCGH analysis, these authors found that these selective pressures resulted in the emergence of aneuploid strains showing gross chromosomal rearrangements such as amplifications, deletions and translocations. Experiments have also shown that spores of *S. cerevisiae*, *S. bayanus* (*S. uvarum*) and *S. pastorianus* are capable of mating to produce diploids (Delneri et al. 2003). However, upon sporulation, the spores produced are sterile and rarely produce meiotic offspring, leading to the genetic isolation of these strains and the maintenance of stable polyploid genomes.

That such chromosomal rearrangements result from selective pressures in nature was elegantly shown by Perez-Ortin et al. (2002) in the analysis of a wine yeast strain T73. In this strain, a reciprocal translocation between chromosomes VIII and XVI was observed. This genome rearrangement confers sulfite resistance to the strain by generating a fusion between the 5′ upstream region of the *SSU1* gene, encoded on chromosome XVI, and the promoter region of the *ECM34* gene from chromosome VIII. The *SSU1* gene mediates sulfite efflux in *S. cerevisiae* (Park and Bakalinsky 2000). Currently no known function has been ascribed to the *ECM34* gene. The resultant chimera (SSUI-R) results in higher expression levels of the *SSU1* gene, thus conferring sulfite resistance on this strain. The translocation event seems to be mediated by a short stretch of sequence homology between the two regions: The SSU1-R allele contains four repeats of a 76-bp sequence which is identical to a single copy of a 77-bp sequence in the *ECM34* gene. It is interesting to note that the site of translocation on chromosome VIII at YHL043W co-localises with a recombination site in the lager yeasts identified by CCGH (Bond et al. 2004), suggesting that this region may represent a particular "hotspot" of recombination common to wine and lager yeasts. However, an analysis of recombination sites identified in flor yeasts and lager yeasts shows few or no sites in common other than sites in subtelomeric regions. This lack of similarity may reflect the different adaptive pressures on these yeast strains owing to the different fermentation conditions encountered in beer making and wine making. It is interesting to note that none of the translocation sites

identified by sequencing in *S. bayanus*, *S. mikatae* and *S. paradoxus* strains (Kellis et al. 2003) have been identified by CCGH in the lager strains, suggesting that none of these species are the true parent of the hybrid lager strains. Alternatively, the recombination sites identified by CCGH may be *S. cerevisiae* specific. It is also possible that the strains have undergone further genomic rearrangements so that the original translocations have been lost or altered.

7.4.3 Comparative Proteomics to Reveal the Hybrid Constitution of Industrial Strains

The molecular details of the mixed constitution of industrial yeast strains can also be analysed by the application of proteomics.

7.4.3.1 Comparative Proteomics of Lager Strains

Boucherie and colleagues (Joubert et al. 2000) used 2D-PAGE in a comparative analysis of proteomes of different lager yeast isolates from various breweries. In agreement with the hypothesis that the genome of lager yeast is a hybrid from at least two different genomes, it was observed that many of the more abundant lager yeast 2D spots seemed to be duplicated (compared with what was found for the laboratory *S. cerevisiae* strain). The authors mentioned in their report that many of these duplicated spots exhibited the same relative abundance while displaying slight differences in their pI and/or M_r. For three of the duplicated non-*S. cerevisiae* spots, their identity as sequence variants of Pdc1p, Eno1/2p and Fba1p was confirmed by the use of microsequencing. Differential labelling of proteins from different strains and subsequent sample mixing followed by 2D-PAGE analysis allowed reliable scoring of the number of co-migrating proteins on a global scale. Independent analysis by two different labelling methods yielded essentially the same result: roughly 85% of the proteins in the lager strain K11 (patent no. FR 2 750 703-A1) co-migrated with proteins from *S. cerevisiae* (based on the analysis of about 300–500 proteins, respectively, in the two independent studies). Comparative proteome analysis via co-migration on 2D gels was also applied earlier to the analysis of strain variants of *S cerevisiae*, where one isolate from fish intestine was related to a laboratory strain (Andlid et al. 1999); for all 984 matched proteins in that study it was apparent that in total 98% of the proteins co-migrated, leaving only 16 strain-specific proteins in the fish isolate. Thus, in the comparison between *S. cerevisiae* and the lager beer strains a much lower similarity was seen compared with what was found between *S. cerevisiae* strains. The overlap to the 2D pattern of *S. cerevisiae* (Joubert et al. 2000) for the lager strain made possible the release of the first 2D map of an industrial yeast encompassing 185 identified proteins. The lager reference 2D pattern was later extended by MALDI-MS analysis to include an additional 30 lager-specific protein spots (Joubert et al. 2001).

The 2D pattern of the lager strain K11 was also compared with that of other type strains/species in an attempt to identify the non-*S. cerevisiae* genetic component of lager strains. It was reported that the proteome of lager brewing yeasts and of the type strains of *S. carlsbergensis*, *S. monacensis* and *S. pastorianus* could be

interpreted as the superimposition of two elementary patterns. One of them originates from proteins encoded by an *S. cerevisiae* like genome and the other apparently from the *S. pastorianus* NRRL Y-1551 strain. Surprisingly it was found that the two different type strains of *S. pastorianus* displayed rather different 2D patterns, despite supposedly being identical isolates. The reason for this discrepancy was not known, and complicates the final interpretation of their result. However, this work constitutes an important example for the use of 2D analysis for the comparative proteomics of yeast strains using spot position (x and y dimensions) as an indicator of protein identity.

More recent analysis of different lager strains adopted the procedure of liquid chromatography (LC) MS/MS analysis of resolved protein spots. In this work (R. Caesar, E. Pettersson, J. Gustafsson, A. Hashemi, M. Rudemo, A. Blomberg, unpublished results) MS/MS analysis of non-*S. cerevisiae* spots in the 2D gel and subsequent comparison to the complete proteomes of the yeasts within the *Saccharomyces sensu stricto* group revealed a phylogenetic tree with great similarity to the one obtained using gene sequence data. In this comparison the non-*S. cerevisiae* proteins in the total protein 2D pattern of the lager strains are clearly most similar to those of *S. bayanus* (Fig. 7.4).

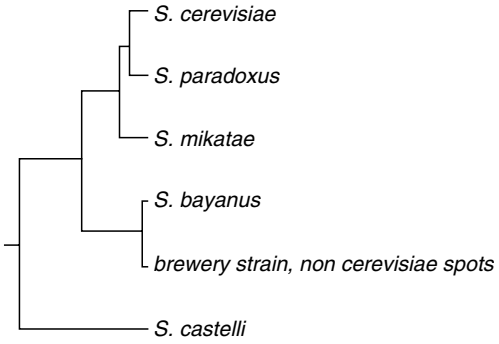

Fig. 7.4. Phylogenetic tree of the non-*Saccharomyces cerevisiae* proteins in the lager strain CMBS33. Seven dominant spots on a CMBS33 2D gel lacking corresponding spots on the gels from the laboratory strain BY4742 were identified as Eft1p, Shm2p, Ilv5p, Pdc1p, Met6p, Pgk1p and Pyk1p using liquid chromatography tandem mass spectrometry (MS/MS). The SEQUEST algorithm was subsequently used to correlate MS/MS spectra from these spots to theoretical mass data derived from a FASTA database containing the amino acid sequences of the identified proteins and their homologues in *S. paradoxus*, *S. mikatae*, *S. bayanus* and *S. castelli*. For each protein the peptide sequences identified from any of the species were compiled to a sequence for the CMBS33 form of the protein. In total, the peptides identified contained 1,237 amino acids covering 32% of the seven proteins. Multiple alignment of the regions covered by the MS/MS analysis were performed between the five *Saccharomyces* species and the brewery strain using ClustalW and a hierarchical tree based on the result was constructed. The brewery strain non *S. cerevisiae* spots differed with one amino acid from *S. bayanus*, with 36 amino acids from *S. cerevisiae*, with 34 amino acids from *S. paradoxus*, with 32 amino acids from *S. mikatae* and with 134 amino acids from *S. castelli*

7.4.3.2 Comparative Proteomics of Ale Strains

The first protein 2D map of an ale-fermenting yeast was recently presented (Kobi et al. 2004). In this analysis 205 spots corresponding to 133 different proteins were identified by MALDI-TOF analysis in the A38 ale strain. Comparison of the proteome of this ale strain with a lager brewing yeast and the *S. cerevisiae* strain S288c confirmed that ale strains are much closer to S288c than lager strains. A zoom-in on a central part of the 2D patterns of three different lager strains, *S. cerevisiae* S288c, and two different ale strains clearly displays these differences and similarities (Fig. 7.5). The A38 ale strain exhibited a very similar pattern to the *S. cerevisiae*

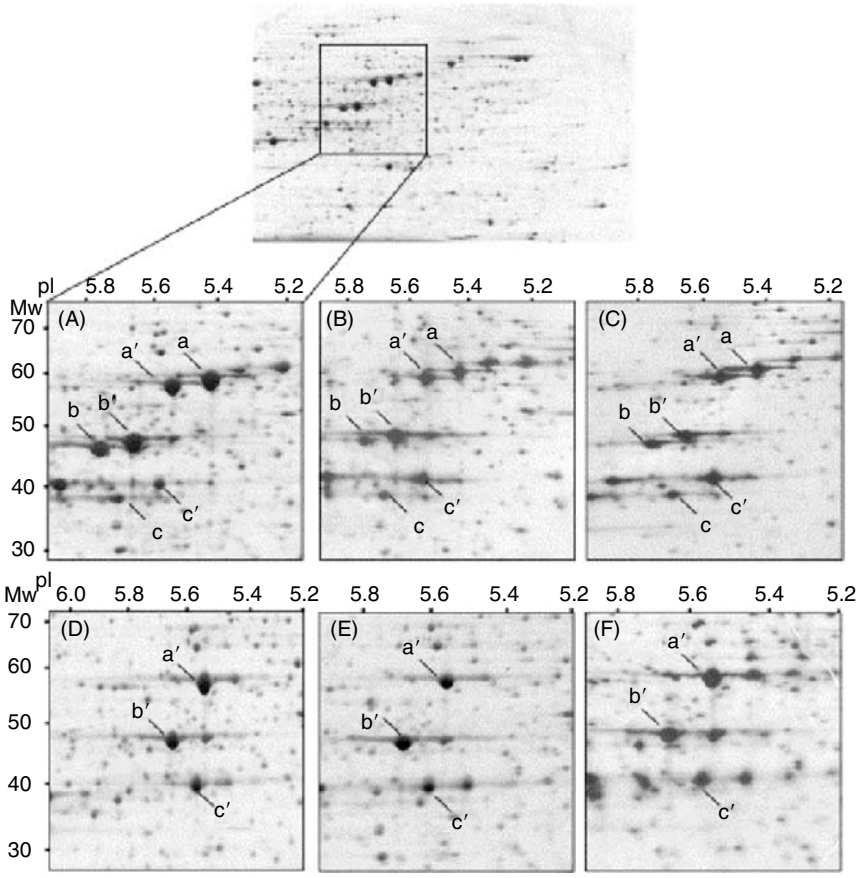

Fig. 7.5. The 2D patterns of ale strains are similar to the one found in laboratory strains. 2D polyacrylamide gel electrophoresis analysis of ale (**e**, **f**), lager (**a–c**) and laboratory (**d**) strains. The protein spots indicated are Pdc1p (*a* and *a*'), Eno2p (*b* and *b*') and Fba1p (*c* and *c*'). 2D images are displayed with the acidic side to the *right*. (Adopted from Kobi et al. 2004 with permission from Elsevier)

strain, and this observation could also be extended to other ale strains. However, this does not mean that the 2D patterns of these ale strains were identical to the laboratory strain S288C. On the contrary, looking at the entire gel for these strains, some discrepancies appear, e.g. the alcohol dehydrogenase isoform Adh4p is present in the ale-brewing strains but not in the laboratory strain under these growth conditions. This is interesting since this zinc-dependent Adh4 protein is reported to be regulated by zinc deficiency (Yuan 2000), which could indicate that the industrial ale strain might exhibit alterations in zinc metabolism. The Adh4p isoform was also found expressed under normal laboratory growth conditions in an industrial baker's yeast strain (Nilsson et al., unpublished results).

7.4.3.3 Comparative Proteomics of Baker's Yeast Strains

Four industrial baker's yeast strains with various fermentative capacities were compared with a laboratory *S. cerevisiae* strain during fed-batch cultivation. After 2D-PAGE analysis the strains were grouped according to differences in protein expression and 2D spot positional differences (Nilsson et al., unpublished results). It was found that the baker's yeast strains were more similar to each other than to the laboratory strain. However, overall the 2D patterns were rather similar among all strains and the most outstanding difference was found for the protein spots corresponding to alcohol dehydrogenase I and II, where all strains examined exhibited large variations. In particular it was noticed that the baker's yeast strains contained several alcohol dehydrogenase spots not present in the laboratory strain. Subsequent MALDI-MS analysis of these alcohol dehydrogenase spots revealed a rather large apparent variability in the alcohol dehydrogenase protein sequence, and some of the sequences seemed to be sequence hybrids between Adh1p and Adh2p.

7.4.4 Genomic Sequencing of Industrial Strains of Yeast

The ultimate characterisation of industrial yeast strains will require the complete genome sequence at a nucleotide level; however, this has been hampered by the polyploid nature of their genomes which requires sequencing and analysis of both the *S. cerevisiae* and the non-*S. cerevisiae* components of the genome. The first description of the sequencing of an industrial lager strains was reported by Suntory, Japan (Nakao et al. 2003). This group sequenced the lager strain Weihenstephan 34/70 (*S. pastorianus*) using a random shotgun approach. Two types of contigs covering 23.2 Mbp (95% coverage) were identified, displaying either 98% identity or 85% identity to the *S. cerevisiae* DNA sequence (Kodama et al. 2005; Nakao et al. 2003). The sequence data confirmed the findings of classical genetics of the presence of three types of chromosomes in the lager strains, *S. cerevisiae* like, non-*S. cerevisiae* like and mosaic chromosomes, arising from recombination events between homeologous chromosomes and identified 37 chromosomal varieties. In general, with the exception of a number of translocations and inversions, the gene synteny is identical to that of *S. cerevisiae*. The non-*S. cerevisiae* like DNA sequences most closely resemble those of *S. bayanus* (*uvarum*); however, the sequences are not identical. This may be due to the difference in strains of *S. bayanus* as only a single *S. bayanus* sequence

is currently available. Interestingly, this study identified 20 ORFs with no significant identities to *S. cerevisiae* ORFs. Furthermore, the recombination sites on the mosaic chromosomes, identified by CCGH analysis (Bond et al. 2004), are confirmed by the sequence analysis, although the Weihenstephan strain contains a number of unique mosaic chromosomes that are not present in the 6701 or CMBS strains characterised by CCGH analysis and vice versa (Fig. 7.6).

7.5 Analyses of the Industrial Process

The novel tools for genome-wide analysis of the yeast transcriptome and proteome are not only instrumental for a better description of industrial strains but also allow a more comprehensive analysis of the industrial-scale fermentation process. The environmental conditions experienced by yeasts in industrial settings are quite unique in their complexity and this complexity is rarely experienced by strains propagated under laboratory growth conditions. For example, lager and ale yeasts during industrial fermentation simultaneously experience a wide array of conditions such as anaerobiosis, high alcohol concentrations, high hydrostatic pressure and high cell density. Likewise, wine yeasts are exposed to a unique set of conditions such as high sugar concentrations, anaerobiosis, low pH and high ethanol concentrations during oenological fermentations. One might consider all of these conditions as imposing severe stress on the yeasts, and therefore it is of particular interest to examine the gene expression profiles in order to understand the stress responses of these yeasts

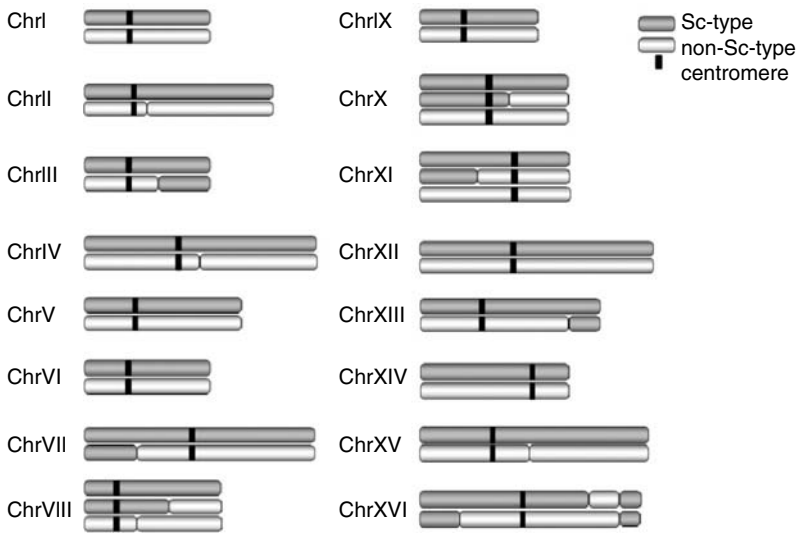

Fig. 7.6. Genome composition of the lager strain Weihenstephan 34/70 (*S. pastorianus*) as deduced by DNA sequencing. (Adopted from Kodama et al. 2005 with permission from Springer-Verlag)

to their environment and to examine the expression patterns of genes involved in metabolic processes, in particular those associated with carbohydrate metabolism.

A number of transcriptome datasets are currently available profiling the gene expression patterns in lager yeasts (Dawes et al. 2002; Higgins et al. 2003a, b; James et al. 2002, 2003; Olesen et al. 2002; Panoutsopoulou et al. 2001; Pugh et al. 2002) and wine yeasts (Backhus et al. 2001; Cavalieri et al. 2000; Erasmus et al. 2003; Rossignol et al. 2003) under fermentation conditions. It should be pointed out that in all transcriptome studies outlined in this section, it is only possible to examine gene expression of *S. cerevisiae* like genes in either lager yeasts or wine yeasts. The contribution of the non-*S. cerevisiae* genes to the overall gene expression patterns remains unexplored at this stage. Likewise the effects of multiple copies of similar but non-identical genes on the overall gene expression pattern are currently unknown, as is whether control of gene expression is conserved between the *S. cerevisiae* and non-*S. cerevisiae* genes. Currently these datasets have not been compiled into a format (website) that allows direct comparison of the different datasets; however, some comparisons have been carried out (James and Bond, unpublished data). While direct comparisons of the data are hampered by the varying experimental conditions used (parameters such as time points sampled during the fermentation, starting cell densities, industrial growth media and temperature used all vary between the datasets), an overall picture of the physiological state of the yeast under fermentation conditions can be deduced.

7.5.1 Beer Production

7.5.1.1 Gene Expression Patterns in Lager Yeasts Under Fermentation Conditions

Among the common gene expression themes that emerge from the analysis of lager yeast fermentations is the co-ordinate upregulation of genes affected by anaerobiosis and those required for ergosterol and fatty acid metabolism. The latter set of genes are required for continued cell membrane biosynthesis, a process requiring the presence of molecular oxygen. Thus, brewers generally actively aerate the wort at the start of the fermentation. This small quantity of initially added O_2 is sufficient to allow continued ergosterol biosynthesis during the fermentation. The co-ordinate upregulation of genes required for fatty acid and sterol metabolism is further verified by the upregulation of a number of genes involved in peroxisomal metabolism such as the PEX genes *PEX*, *5*, *10*, *14*, *21*, *11* and *18*. Likewise, the PAU gene family, initially identified as genes responding to anaerobic growth conditions, but more recently suggested to be sterol carriers (Wilcox et al. 2002), are co-ordinately upregulated as fermentation proceeds (James et al. 2003). High levels of ergosterol may be essential to protect yeast from ethanol stress (Alexandre et al. 1994). Another interesting finding from the transcriptome analysis was the co-ordinate upregulation during large-scale fermentation of genes involved in aldehyde metabolism (Fig. 7.1).

Of the genes that show decreased transcript levels on days 3 and 8, with respect to day 1, of fermentation, the most abundant classes are those required for protein and amino acid biosynthesis. These account for more than 30% of the downregulated genes. The co-ordinate downregulation of these genes most likely reflects the

low level of general cell metabolism occurring during industrial fermentation, where cells may only undergo one to two cell divisions. The co-ordinate downregulation of protein synthesis genes has also been observed in stationary-phase yeast cultures (Gray et al. 2004). The majority of genes involved in glycolysis are similarly downregulated as fermentation proceeds (Fig. 7.1). These genes are generally regulated by catabolite repression (glucose repression) and their downregulation may reflect continued levels of useable carbohydrates even at day 8 of fermentation.

7.5.1.2 Application of Proteomics to Analyse Ale Fermentation

Kobi et al. (2004) followed proteome changes during the fermentation process of the ale strain A38 in a 10-hl pilot device, for the first, second and third generation. The A38 strain was initially grown in aerobic conditions with saccharose as the sole carbon source before pitching in wort under anaerobic conditions for the first fermentation. To mimic the industrial process, the yeasts were harvested at the end of the fermentation and re-inoculated in the fresh wort for a second and then third generation. In particular, proteome analysis and comparison was performed during each generation (at the start and after roughly 200 h). A comparison between the beginning and the end of the first generation showed that 50 of the 85 differentially expressed proteins were repressed, mostly glycolytic enzymes, proteins involved in acetylcoenzyme A formation, proteins of the tricarboxylic acid cycle, and proteins involved in respiration. It is clear that most of the changes reveal an adaptation to anaerobic conditions. In addition, most of the proteins induced at later times in the fermentation process were protein fragments belonging to either proteins in carbon metabolism, or protein or amino acid biosynthesis pathways. It was suggested that intracellular proteolysis influenced the regulation of these proteins during the industrial fermentation process.

Many fewer changes in protein abundance were scored between the beginning and the end of the third generation, and the observed changes exhibited lower levels of response. Among the proteins that displayed at least a twofold change were proteins involved in methionine biosynthesis (e.g. Sah1p, Met6p and Met3p) as well as some involved in carbon metabolism (e.g. Fba1p, Adh1p and Ald6p); all the example proteins indicated displayed induction except Ald6p, which was repressed.

7.5.1.3 Stress Responses During Beer Fermentation

Considering the extreme environmental conditions experienced by yeasts during the fermentation process, one might expect the induction of stress genes in these cells. Stress-responsive genes are generally regulated by transcription factors such as the heat shock factor, Hsf1 and the Msn2/Msn4 proteins, which bind to heat shock elements (HSE) and stress responsive elements (STREs), respectively. Msn2/Msn4p are zinc-finger proteins that are activated by a number of stress conditions encountered by yeasts such as entry into the stationary phase, carbon source starvation and osmotic stress. Surprisingly, transcriptome analysis during fermentation indicates that genes under the control of Hsf1 and Msn2/Msn4 are downregulated as fermentation proceeds (James et al. 2003). These results are consistent with previous

data showing that Hsp104p levels were repressed during fermentation (Gray et al. 2004; Brosnan et al 2000). The lack of expression for stress-responsive genes appears to be unique to lager strains of yeasts. A comparative analysis of genes encoding heat shock proteins (HSP genes) in a haploid laboratory strain of *S. cerevisiae* and the lager strain 6701, grown under identical fermentation conditions, revealed that a subset of the HSP genes (*HSP104*, *HSP30*, *HSP26* and *HSP12*) are in fact highly induced late in fermentation (on day 8) exclusively in the haploid laboratory strain (James et al. 2002; James and Bond, unpublished results), a response not encountered for the lager strain. These results therefore suggest that the transcriptional regulation of HSE-/STRE-regulated HSP genes is significantly different in lager and laboratory strains of yeasts.

The regulation of genes responding to oxidative stress has also been extensively examined in lager strains of yeasts during fermentation (Higgins et al. 2003a; James et al. 2003). These genes are generally under the control of the transcription factor Yap1p. Expression levels of these genes are elevated on days 1 and 3 of fermentation most likely in response to the production of O_2 free radicals. The expression patterns of these genes differ in the various reported datasets (Higgins et al. 2003a; James et al. 2003; Olesen et al. 2002). This most likely reflects the different regimes for oxygenation of the media at the beginning of fermentation. A number of oxidative responsive genes are also induced late in fermentation; these include the genes *LYS7*, *SOD1*, *TSA1* and *MXR1* (James et al. 2002, 2003). The *MXR1* gene encodes the protein methionine sulfoxide reductase, which reverses the oxidation of methionine residues and which has been shown to be required for the reduction of dimethyl sulfoxide to dimethyl sulfide, which is a thioester of major importance for the aroma and flavour of beer (Hansen et al. 2002).

Kobi et al. (2004) reported that a number of stress proteins exhibited changes in ale yeasts during the first-generation fermentation, like Hsp26p and Ssa4p. However, the only stress proteins that also increased during the third generation of fermentation were Kar2p and Ssa1p (both these proteins encode chaperones that are required for protein folding). Yeast strains must cope with various stresses during the fermentation process; however, there is clearly no strong stress response in the ale strain.

7.5.2 Wine Fermentation

7.5.2.1 Gene Expression Patterns in Wine Strains Under Fermentation Conditions

The conditions for wine and lager fermentations differ in a number of important aspects, such as starting cell concentrations, growth media (wort, rich in maltose for beer, and grape juice, rich in fructose and glucose for wine), pH, temperature and length of incubation. However, yeasts in both fermentations share some common environmental conditions, such as anaerobiosis and high ethanol concentrations at the end of the fermentation. In general, carbohydrates do not become limiting during wine fermentations. Instead cells enter the stationary phase owing to limiting nitrogen concentrations and/or attainment of maximal cell density. Surprisingly, the expression patterns are remarkably similar in both lager and wine fermentations;

however, unique expression patterns are also observed under both conditions. As observed in lager strains of yeasts, there is a co-ordinate downregulation of genes involved in protein and amino acid biosynthesis and upregulation of PAU genes and genes encoding proteins involved in aldehyde metabolism. Significant differences between the expression patterns observed in the wine and lager fermentations include the downregulation of genes for ergosterol biosynthesis in wine fermentations and the upregulation of genes required for glycolysis. The former results from the inclusion of ergosterol in the culture medium, while the latter most likely reflects the concentrations of carbohydrates at different times during the wine and lager fermentations. One interesting similarity is the upregulation of genes encoding aryl alcohol dehydrogenases. The contribution of these genes to total alcohol synthesis is presently unknown. It has also been suggested that these genes are induced in response to oxidative stress (Rossignol et al. 2003).

Transciptome analysis during wine fermentation also revealed major changes in gene expression patterns as a result of nitrogen depletion as the fermentation proceeded and confirmed that growth arrest (entry into the stationary phase) is a direct consequence of nitrogen depletion (Rossignol et al. 2003). The major changes in the gene expression pattern associated with nitrogen depletion were the induction of genes required for the metabolism of poor alternative nitrogen sources such as genes required for proline, allantoin and urea utilisation, nitrogen permeases and genes encoding proteins required for the management of glutamate pools. All of these genes are under the control of the TOR pathway.

7.5.2.2 Stress Responses During Wine Yeast Fermentation

One of the major differences in gene expression patterns between wine and lager yeast fermentations is the upregulation of HSP genes in wine yeast (Rossignol et al. 2003). As mentioned before, this group of genes appear to be actively repressed during lager yeast fermentations. The expression patterns of HSP genes in wine yeasts most closely resembles that observed in haploid *S. cerevisiae* strains grown under fermentation conditions (James et al. 2002). The HSP genes, such as *HSP30*, *HSP26* and *HSP104*, which are under the control of STREs are particularly induced during wine fermentations. Such genes have previously been shown to be induced following diauxic shift and upon entry into the stationary phase. Both conditions see a shift from usage of fermentable carbohydrates to non-fermentable carbon sources such as acetate and ethanol. It is possible that conditions at diauxic shift are radically different in wine and lager production. Alternatively, the differential expression of HSP genes in wine and lager yeasts raises the possibilities that chromatin structure and/or the presence of homeologous non-*S. cerevisiae* genes contribute to these gene expression patterns.

In addition to HSP genes, 58% of genes previously defined as being regulated as part of the common environmental response (CER) and the environmental stress response (ESR) (Gasch et al. 2000) are upregulated during wine fermentations (Rossignol et al. 2003). Additionally, genes responsive to ethanol stress are also induced. Many of these genes are involved in cell wall biogenesis, suggesting that cell wall alterations may help yeasts cope with ethanol stress. It is interesting to note

that ten genes involved in cell wall biogenesis are also induced late in lager yeast fermentations (James et al. 2003).

A wild-type *S. cerevisiae* wine strain isolated from the natural must of spontaneous grape fermentation was analysed for proteome changes during semiaerobic growth conditions (Trabalzini et al. 2003). In particular the response of this wine strain to the exhaustion of glucose from the medium was of interest since this is believed to be the main reason for unwanted stuck fermentation during vinification. When glucose was depleted from the medium a large number of changes in protein abundance were apparent. Roughly 50 proteins displayed decreased amounts at the end of fermentation, usually in the range twofold to threefold repression and, interestingly, some of these proteins are currently functionally not characterised (e.g. Ybr025cp and Yir035cp). In addition, a large number of protein spots with low molecular weight increased during the later phases of fermentation. Identification of these spots indicated that they belonged to protein fragments of larger proteins, which indicated proteolytic breakdown as part of the cellular response. However, also well-known stress responsive proteins like Sod1p, Tsa1p and Ctt1p exhibited increased expression when glucose was depleted, indicating a clear stress response in this wine yeast.

7.5.2.3 Application of Proteomics to Understand Factors That Affect Wine Haze

The clarity of white wine is highly important for the winemaker. Bottles showing haziness, likely to be rejected by the consumer, result from the aggregation of grape proteins naturally present in wine. To prevent haze formation, winemakers usually lower the concentration of wine proteins through the use of bentonite. Unfortunately bentonite also removes wine aroma components, hence lowering wine quality. Thus, alternative methods of protein stabilisation are being investigated by the wine industry.

Mannoproteins from yeast cell walls are known to be released into the extracellular medium during yeast growth, in particular during the stationary phase (Dupin et al. 2000). One of the proteins released, invertase, was shown to lower haze formation, probably by competing with grape-derived proteins for some unknown factor(s) in wine that is required to form large highly light scattering protein aggregates that are responsible for the haze. To better characterise the protein components of wine a 2002 vintage Sauvignon Blanc wine was analysed by LC-MS/MS (Kwon 2004). Wine proteins were concentrated by the use of a 5-kDa cellulose membrane centrifugal filter tube, followed by salting-out precipitation in saturated ammonium sulfate aqueous solutions and subsequent centrifugation to isolate the protein pellets. Proteins were separated by SDS-PAGE and Coomassie visualised bands were cut out for trypsination and peptide elution. The LC-MS/MS analyses resulted in the identification of 12 different cell wall or plasma membrane associated proteins from *S. cerevisiae*: e.g. Gas1p, a GPI-anchored 1,3-glucosyltransferase, Pho3p, a periplasmic acid phosphatase, Suc4p, invertase, Bgl2p, *endo*-1,3 glucanase and Yju1p (Cwp1p), reported as a structural component of the cell wall. Many of these released yeast proteins could have an influence on the wine haze formation. It was proposed that MS identification of proteins could be used as a quality indicator of wine.

7.5.3 Industrial Production of High-Quality Baker's Yeast

Two aspects of industrial baker's yeast production have been investigated using proteomics. One important feature of baker's yeast performance is high fermentative activity even after long-term cold storage. To identify molecular effects during this storage regime on a baker's yeast strain, 2D analysis was performed before and after storage for 26 days at 4°C (Nilsson et al. 2001a). It was clear from this analysis that cold storage resulted in large changes in the protein content. However, even after this long period of storage the original state of the culture, e.g. exponential growth in the respiro-fermentative state or in the transition phase to respiration, could still be distinguished by analysis of the 2D pattern. The fermentative activity after storage was also shown to be dependent on the initial state, and thus it was concluded that the initial state at harvest was important for the long-term performance of this industrial yeast strain. The baker's yeast cells experience different types of starvation regimes during industrial production. To investigate the effects on these industrial strains from either nitrogen or carbon limited starvation, a differentially starved baker's yeast strain was analysed for protein changes (Nilsson et al. 2001b). It was found that for certain glycolytic enzymes a significant change in the amount of protein could be observed during these starvation regimes, e.g. Pdc1p decreased during nitrogen starvation, while Adh1p was downregulated during carbon starvation. However, the recorded changes in the protein levels did not correlate with the observed changes in fermentative capacity.

7.6 Future Perspectives

The complete genome sequence of the first industrial yeast strain, the lager yeast Weihenstephan 34/70 (*S. pastorianus*), will be instrumental in our analysis of these hybrid genomes. It will not only potentially indicate the full compliment of genes in these cells, but will also provide the basis for a number of novel technological possibilities in the analysis of industrial yeast strains. The genome sequence will soon be opened up for expression analysis of not only the *S. cerevisiae* part of the genome but will also put the non *S. cerevisae* part within experimental reach; not before long the first microarrays with a full complement of genes from industrial yeast will be available. In addition, it will make possible a more complete and precise proteome analysis by providing a better template for database searches using MS data. It will also make possible non-gel-based proteomics approaches that are based on whole cell trypsination prior to high-resolution LC-MS/MS analysis (Washburn et al. 2001). Hopefully these novel possibilities will lead to the publication of studies where combined transcriptome and proteome analyses are presented, which are missing at present. Of great general interest will be the analysis of differences in the regulatory components of these mixed genomes where industrial strains might have evolved/been selected to respond quite differently to the well-studied laboratory *S. cerevisiae* strains (see previous sections on the stress response). Hopefully the genome sequence and the novel experimental possibilities will also lead to the development of dedicated databases where transciptome and proteome data from industrial yeast strains can be presented, compared and analysed. The interplay between

the hybrid proteomes will be another important avenue for future studies, since the genome sequence puts within experimental resolution a view of the physical protein interaction network, by either two-hybrid analysis or co-immunoprecipitation (e.g. tandem affinity purification tagging), in particular regarding the mixed composition of protein complexes. These studies will also be of fundamental importance in our understanding of the functionality of hybrid protein complexes, where different components compete in the formation of complexes and play different functional roles in their final activity.

Many challenges remain and will require further advancements both in the tools of bioinformatics as well as in analytical techniques. One important aspect is the finite gene number for industrial strains, a number that even for the well-studied laboratory *S. cerevisiae* strain, as mentioned earlier, has changed over the years and most likely will continue to change. In this respect the mixed genomes of some industrial yeast strains will provide an even greater challenge in gene identification. However, maybe the greatest obstacle in our strive for a more detailed functional analysis of individual genes in the genetic background of industrial strains will be the difficulty in generating gene deletions in these mixed genomes of several similar and identical gene copies. This problem will certainly ensure that the evolutionary link to *S. cerevisae* will persist not only in the industrial production line but also in future functional analysis of non-*S. cerevisae* genes in the potent laboratory test-bed *S. cerevisiae*.

References

Alexandre H, Rousseaux I, Charpentier C (1994) Ethanol adaptation mechanisms in *Saccharomyces cerevisiae*. Biotechnol Appl Biochem 20:173–183

Anand R, Southern EM (1990) Pulsed field gel electrophoresis. In: Rickwood D, Hames BD (eds) Gel electrophoresis of nucleic acids: a practical approach. IRL Press at Oxford University Press, New York, pp 101–123

Andlid T, Blomberg L, Gustafsson L, Blomberg A (1999) Characterisation of *Saccharomyces cerevisiae* CBS 7764 isolated from rainbow trout intestine. Syst Appl Microbiol 22:145–155

Backhus LE, DeRisi J, Bisson LF (2001) Functional genomic analysis of a commercial wine strain of *Saccharomyces cerevisiae* under differing nitrogen conditions. FEMS Yeast Res 1:111–125; Erratum (2003) FEMS Yeast Res 4:123

Bataillé N, Peypouquet M-F, Boucherie H (1987) Identification of polypeptides of the carbon metabolism machinery on the two-dimensional map of *Saccharomyces cerevisiae*. Location of 23 additional polypeptides. Yeast 3:11–21

Bataillé N, Thoraval D, Boucherie H (1988) Two-dimensional gel analysis of yeast proteins: Application to the study of changes in the levels of major polypeptides of *Saccharomyces cerevisiae* depending on the fermentable or nonfermentable nature of carbon source. Electrophoresis 9:774–780

Bjellqvist B, Ek K, Righetti PG, Gianazza E, Görg A, Westermeier R, Postel WJ (1982) J Biochem Biophys Methods 6:317–339

Blomberg A (2002) Use of two-dimensional gels in yeast proteomics. Methods Enzymol 350:559–584

Bond U, Neal C, Donnelly D, James TC (2004) Aneuploidy and copy number breakpoints in the genome of lager yeasts mapped by microarray hybridisation. Curr Genet 45:360–370

Boucherie H, Bataille N, Fitch IT, Perrot M, Tuite MF (1995) Differential synthesis of glyceraldehyde-3-phosphate dehydrogenase polypeptides in stressed yeast cells. FEMS Microbiol Lett 125:127–134

Brosnan MP, Donnelly D, James TC, Bond U (2000) The stress response is repressed during fermentation in brewery strains of yeast. J Appl Microbiol 88:746–755

Brousse M, Bataillé N, Boucherie H (1985) Identification of glycolytic enzyme polypeptides on the two-dimensional map of *Saccharomyces cerevisiae* and application to the study of some wine yeasts. Appl Environ Microbiol 50:951–957

Burke DT, Carle GF, Olson MV (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science 236:806–812

Carle GF, Olson MV (1985) An electrophoretic karyotype for yeast. Proc Natl Acad Sci USA 82:3756–3760

Carle GF, Frank M, Olson MV (1986) Electrophoretic separation of large DNA molecules by periodic inversion of the electric field. Science 232:65–68

Casey GP (1986) Molecular and genetic analysis of chromosomes X in *Saccharomyces carlsbergensis*. Carlsberg Res Commun 51:343–362

Casey GP (1996) Practical applications of pulsed field electrophoresis and yeast chromosome fingerprinting in brewing QA and R&D. Techn Q 33:1–10

Castrejon F, Martinez-Force E, Benitez T, Codon A (2004) Genetic analysis of apomictic wine yeasts. Curr Genet 45:187–196

Cavalieri D, Townsend JP, Hartl DL (2000) Manifold anomalies in gene expression in a vineyard isolate of *Saccharomyces cerevisiae* revealed by DNA microarray analysis. Proc Natl Acad Sci USA 97:12369–12374

Cherry JM et al (1997) Genetic and physical maps of *Saccharomyces cerevisiae*. Nature 387:67–73

Chittur SV (2004) DNA microarrays: tools for the 21st century. Comb Chem High Throughput Screen 7:531–537

Chu G (1990) Pulsed-field electrophoresis: theory and practice. In: Birren B, Lai E (eds) Methods: a companion to methods of enzymology. Pulsed-field electrophoresis. Academic, San Diego, pp 129–142

Chu G, Vollrath D, Davis RW (1986) Separation of large DNA molecules by contour-clamped homogeneous electric fields. Science 234:1582–1585

Cliften P, Sudarsanam P, Desikan A, Fulton L, Fulton B, Majors J, Waterston R, Cohen BA, Johnston M (2003) Finding functional features in *Saccharomyces* genomes by phylogenetic footprinting. Science 301:71–76

Codon A, Benitez T, Korhola M (1998) Chromosomal polymorphism and adaptation to specific industrial environments of *Saccharomyces* strains. Appl Microbiol Biotechnol 49:154–168

Coulson A, Waterston R, Kiff J, Sulston J, Kohara Y (1988) Genome linking with yeast artificial chromosomes. Nature 335:184–186

Dawes IW et al (2002) Brewing yeast, gene expression and beer quality. Am Soc Brew Chem Abstr 14

de Barros Lopes M, Bellon JR, Shirley NJ, Ganter PF (2002) Evidence for multiple interspecific hybridization in *Saccharomyces sensu stricto species*. FEMS Yeast Res 1:323–331

Delneri D, Colson I, Grammenoudi S, Roberts I, Louis E, Oliver S (2003) Engineering evolution to study speciation in yeasts. Nature 422:68–72

Denis V, Boucherie H, Monribot C, Daignan-Fornier B (1998) Role of Myb-like protein Bas1p in *Saccharomyces cerevisiae*: a proteome analysis. Mol Microbiol 30:557–566

Dietrich FS et al (2004) The Ashbya gossypii genome as a tool for mapping the ancient *Saccharomyces cerevisiae* genome. Science 304:304–307

Dujon B et al. (2004) Genome evolution in yeasts. Nature 430:35–44
Dunham MJ, Badrane H, Ferea T, Adams J, Brown PO, Rosenzweig F, Botstein D (2002) Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 99:16144–16149
Dupin IV, McKinnon BM, Ryan C, Boulay M, Markides AJ, Jones GP, Williams PJ, Waters EJ (2000) *Saccharomyces cerevisiae* mannoproteins that protect wine from protein haze: their release during fermentation and lees contact and a proposal for their mechanism of action. J Agric Food Chem 48:3098–3105
Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 95:14863–14868
Elliott SG, McLaughlin CS (1978) Rate of macromolecular synthesis through the cell cycle of the yeast *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 75:4384–4388
Epstein CB, Butow RA (2000) Microarray technology-enhanced versatility, persistent challenge. Curr Opin Biotechnol 11:36–41
Erasmus DJ, van der Merwe GK, van Vuuren HJ (2003) Genome-wide expression analyses: metabolic adaptation of *Saccharomyces cerevisiae* to high sugar stress. FEMS Yeast Res 3:375–399
Gardiner K, Laas W, Patterson DS (1986) Fractionation of large mammalian DNA restriction fragments using vertical pulsed-field gradient gel electrophoresis. Somatic Cell Mol Genet 12:185–195
Garrels JI (1979) Two-dimensional gel electrophoresis and computer analysis of proteins synthesized by clonal cell lines. J Biol Chem 254:7961–7977
Garrels JI, Futcher B, Kobayashi R, Latter GI, Schwender B, Volpe T, Warner JR, McLaughlin CS (1994) Protein identifications for a *Saccharomyces cerevisiae* protein database. Electrophoresis 15:1–21
Garrels JI et al (1997) Proteome studies of Saccharomyces cerevisiae: identification and characterization of abundant proteins. Electrophoresis 18:1347–1360
Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO (2000) Genomic expression programs in the response of yeast cells to environmental changes. Mol Cell Biol 11:4241–4257
Gemmill RM (1991) Pulsed field gel electrophoresis. In: Chrambach A, Dunn MJ, Radola BJ (eds) Advances of electrophoresis. VCH, Weinheim, Germany, pp 1–48
Gerhold D, Rushmore T, Caskey CT (1999) DNA chips: promising toys have become powerful tools. Trends Biochem Sci 24:168–173
Gerton JL, DeRisi J, Shroff R, Lichten M, Brown PO, Petes TD (2000) Global mapping of meiotic recombination hotspots and coldspots in the yeast *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 97:11383–11390
Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS (2003) Global analysis of protein expression in yeast. Nature 425:737–741
Goffeau A et al (1996) Life with 6,000 genes. Science 274:546
Görg A, Postel W, Günther S (1988) The current state of two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis 9:531–546
Görg A, Obermaier C, Boguth G, Weiss W (1999) Recent developments in two-dimensional gel electrophoresis with immobilized pH gradients: wide H gradients up to pH 12, longer separation distances and simplified procedures. Electrophoresis 20:712–717
Gray JV, Petsko GA, Johnston GC, Ringe D, Singer RA, Werner-Washburne M (2004) "Sleeping beauty": quiescence in *Saccharomyces cerevisiae*. Microbiol Mol Biol Rev 68:187–206
Gustafsson JS, Blomberg A, Rudemo M (2002) Warping two-dimensional electrophoresis gel images to correct for geometric distortions of the spot pattern. Electrophoresis 23:1731–1744

Hansen J, Bruun SV, Bech LM, Gjermansen C (2002) The level of MXR1 gene expression in brewing yeast during beer fermentation is a major determinant for the concentration of dimethyl sulfide in beer. FEMS Yeast Res 2:137–149

Hardiman G (2004) Microarray platforms – comparisons and contrasts. Pharmacogenomics 5:487–502

Higgins VJ, Beckhouse AG, Oliver AD, Rogers PJ, Dawes IW (2003a) Yeast genome-wide expression analysis identifies a strong ergosterol and oxidative stress response during the initial stages of an industrial lager fermentation. Appl Environm Microbiol 69:4777–4787

Higgins VJ, Rogers PJ, Dawes IW (2003b) Application of genome-wide expression analysis to identify molecular markers useful in monitoring industrial fermentations. Appl Environ Microbiol 69:7535–7540

Iida H, Yahara I (1984) Specific early-G_1 blocks accompanied with stringent response in *Saccharomyces cerevisiae* lead to growth arrest in resting state similar to the G_0 of higher eucaryotes. J Cell Biol 98:1185–1193

Infante JJ, Dombek KM, Rebordinos L, Cantoral JM, Young ET (2003) Genome-wide amplifications caused by chromosomal rearrangements play a major role in the adaptive evolution of natural yeast. Genetics 165:1745–1759

James T, Campbell S, Bond U (2002) Comparative analysis of global gene expression in lager and laboratory yeast strains grown in wort. Proc IEEE 90:1887–1899

James TC, Campbell S, Donnelly D, Bond U (2003) Transcription profile of brewery yeast under fermentation conditions. J Appl Microbiol 94:432–448

Joubert R, Brignon P, Lehmann C, Monribot C, Gendre F, Boucherie H (2000) Two-dimensional gel analysis of the proteome of lager brewing yeasts. Yeast 16:511–522

Joubert R, Strub JM, Zugmeyer S, Kobi D, Carte N, Van Dorsselaer A, Boucherie H, Jaquet-Guffreund L (2001) Identification by mass spectrometry of two-dimensional gel electrophoresis-separated proteins extracted from lager brewing yeast. Electrophoresis 22:2969–2982

Kellis M, Patterson N, Endrizzi M, Birren B, Lander ES (2003) Sequencing and comparison of yeast species to identify genes and regulatory elements. Nature 423:241–254

Kellis M, Birren BW, Lander ES (2004) Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. Nature 428:617–624

Kielland-Brandt MC, Nilsson-Tillgren T, Gjermansen C, Holmberg S, Pedersen MB (1995) Genetics of brewing yeasts. In: Wheals AE, Rose AH, Harrison JS (eds) The yeasts. Academic, London, pp 223–253

Klose J (1975) Protein mapping by combined isoelectric focussing and electrophoresis in mouse tissues. a novel approach to testing for induced point mutations in mammals. Hum Genet 26:231–243

Kobi D, Zugmeyer S, Potier S, Jaquet-Gutfreund L (2004) Two-dimensional protein map of an "ale"-brewing yeast strain: proteome dynamics during fermentation. FEMS Yeast Res 5:213–230

Kodama Y, Kielland-Brandt MC, Hansen J (2005) Lager brewing yeast. In: Sunnerhagen P, Piskur J (eds) Comparative fungal genomics. Springer, Berlin Heidelberg New York (in press)

Kwon SW (2004) Profiling of soluble proteins in wine by nano-high-performance liquid chromatography/tandem mass spectrometry. J Agric Food Chem 52:7258–7263

Larsson T, Norbeck J, Karlsson H, Karlsson K-A, Blomberg A (1997) Identifications of two-dimensional gel electrophoresis resolved yeast proteins by matrix-assisted laser desorption ionisation mass spectrometry. Electrophoresis 18:418–423

Maillet I, Lagniel G, Perrot M, Boucherie H, LaBarre J (1996) Rapid identifications of yeast proteins on two dimensional gels. J Biol Chem 271:10263–10270

Mann M, Hendrickson RC, Pandey A (2001) Analysis of proteins and proteomes by mass spectrometry. Annu Rev Biochem 70:437–473

Mewes HW et al (1997) Overview of the yeast genome. Nature 387:7–65; Erratum (1997) Nature 387:737

Mortimer RK (2000) Evolution and variation of the yeast (*Saccharomyces*) genome. Genome Res 10:403–409; Erratum (2000) Genome Res 10:891

Muller EH, Richards EJ, Norbeck J, Byrne KL, Karlsson K-A, Pretorius GHJ, Meacock PA, Blomberg A, Hohmann S (1999) Thiamine repression and pyruvate decarboxylase autoregulation independently control the expression of the *Saccharomyces cerevisiae* PDC5 gene. FEBS Lett 449:245–250

Nakao Y, Kodama Y, Nakamura N, Takehito I, Hattori M, Shiba T, Ashikari T (2003) Whole genome sequence of a lager brewing yeast. Proceedings of the 29th European brewery convention congress, Dublin, pp 524–530

Nilsson A, Norbeck J, Oelz R, Blomberg A, Gustafsson L (2001a) Fermentative capacity after cold storage of baker's yeast is dependent on the initial physiological state but not correlated to the levels of glycolytic enzymes. Int J Food Microbiol 71:111–124

Nilsson A, Pahlman IL, Jovall PA, Blomberg A, Larsson C, Gustafsson L (2001b) The catabolic capacity of *Saccharomyces cerevisiae* is preserved to a higher extent during carbon compared to nitrogen starvation. Yeast 18:1371–1381

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

Norbeck J, Blomberg A (1995) Gene linkage of two-dimensional polyacrylamide gel electrophoresis resolved proteins from isogene families in *Saccharomyces cerevisiae* by microsequencing of in-gel trypsin generated peptides. Electrophoresis 16:149–156

Norbeck J, Blomberg A (1997) Two-dimensional electrophoretic separation of yeast proteins using a non-linear wide range (pH 3–10) immobilized pH gradient in the first dimension; reproducability and evidence for isoelectric focusing of alkaline (pI >7) proteins. Yeast 13:529–539

Norbeck J, Blomberg A (2000) The level of cAMP dependent protein kinase A activity strongly affects osmotolerance and osmo instigated gene expression changes in *Saccharomyces cerevisiae*. Yeast 16:121–137

Olesen K, Felding T, Gjermansen C, Holmberg S, Pertersen J, Kielland-Brandt MC (2002) The dynamics of the *Saccharomyces carlsbergensis* brewing yeast transcriptome during a production scale lager beer fermentation. FEMS Yeast Res 2:563–573

O'Farrell PH (1975) High resolution two-dimensional electrophoresis of proteins. J Biol Chem 250:4007–4021

O'Farrell PZ, Goodman HM, O'Farrell PH (1977) High resolution two-dimensional electrophoresis of basic as well as acidic proteins. Cell 12:1133–1142

Panoutsopoulou K, Wu J, Hayes A, Butler P, Oliver SG (2001) Yeast transcriptome analysis during the brewing process. Yeast 18: supplement abstract 16–30 from the 20th international conference on yeast genetics and molecular biology, Prague, Czech Republic, 26–31 August

Park H, Bakalinsky AT (2000) SSU1 mediates sulphite efflux in *Saccharomyces cerevisiae*. Yeast 16:881–888

Patton WF, Beechem JM (2002) Rainbow's end: the quest for multiplexed fluorescence quantitative analysis in proteomics. Curr Opin Chem Biol 6:63–69

Perez-Ortin JE, Querol A, Puig S, Barrio E (2002) Molecular characterization of a chromosomal rearrangement involved in the adaptive evolution of yeast strains. Genome Res 12:1533–1539

Perrot M, Sagliocco F, Mini T, Monribot C, Scheider U, Schevchenko A, Mann M, Jenö P, Boucherie H (1999) Two-dimensional gel protein database of *Saccharomyces cerevisiae* (update 1999). Electrophoresis 20:2280–2298

Pugh T et al (2002) Global analysis of yeast gene expression during a brewing fermentation. Annual meeting of the American Society of Brewing Chemists, Tuscon, AZ, O-2

Puig S, Querol A, Barrio E, Perez-Ortin JE (2000) Mitotic recombination and genetic changes in *Saccharomyces cerevisiae* during wine fermentation. Appl Environ Microbiol 66:2057–2061

Rabilloud T (1992) A comparison between low background silver diammine and silver nitrate protein stains. Electrophoresis 13:429–439

Rabilloud T (1998) Use of thiourea to increase the solubility of membrane proteins in two-dimensional electrophoresis. Electrophoresis 19:758–760

Rabilloud T, Adessi C, Giraudel A, Lunardi J (1997) Improvements of the solubilisation of proteins in two-dimensional electrophoresis with immobilised pH gradients. Electrophoresis 18:307–316

Rabilloud T, Strub JM, Luche S, van Dorsselaer A, Lunardi J (2001) A comparison between Sypro Ruby and ruthenium II tris(bathophenanthroline disulfonate) as fluorescent stains for protein detection in gels. Proteomics 1:699–704

Rosenfeld J, Capdevielle J, Guillemot JC, Ferrara P (1992) In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. Anal Biochem 203:173–179

Rossignol T, Dulau L, Julien A, Blondin B (2003) Genome-wide monitoring of wine yeast gene expression during alcoholic fermentation. Yeast 20:1369–1385

Sanchez J-C, Golaz O, Frutiger S, Schaller D, Appel R, Bairoch A, Hughes G, Hochstrasser D (1996) The yeast SWISS-2DPAGE database. Electrophoresis 17:556–565

Schulenberg B, Aggeler R, Beechem JM, Capaldi RA, Patton WF (2003) Analysis of steady-state protein phosphorylation in mitochondria using a novel fluorescent phosphosensor dye. J Biol Chem 278:27251–27255

Shin D-Y, Matsumoto K, Iida H, Uno I, Ishikawa T (1987) Heat shock response of *Saccharomyces cerevisiae* mutants altered in cyclic AMP-dependent protein phosphorylation. Mol Cell Biol 7:244–250

Southern EM, Anand R, Brown WRA, Fletcher DS (1987) A model for the separation of large DNA molecules by crossed field gel electrophoresis. Nucleic Acids Res 15:5925–5943

Trabalzini L et al (2003) Proteomic response to physiological fermentation stresses in a wild-type wine strain of *Saccharomyces cerevisiae*. Biochem J 370:35–46

Vaughan-Martini A, Martini A (1987) Three newly delimited species of *Saccharomyces sensu stricto*. Antonie van Leeuwenhoek 53:77–84

Veeser S, Dunn MJ, Yang GZ (2001) Multiresolution image registration for two-dimensional gel electrophoresis. Proteomics 1:856–870

Washburn MP, Wolters D, Yates JR III (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol 19:242–247

Watanabe M, Takimoto T, Akiba Y, Watanabe T (2004) A new approach to species determination for yeast strains: DNA microarray-based comparative genomic hybridization using a yeast DNA microarray with 6000 genes. Br Poult Sci 45:133–138

Wilcox LJ, Balderes DA, Wharton B, Tinkelenberg AH, Rao G, Sturley SL (2002) Transcriptional profiling identifies two members of the ATP-binding cassette transporter superfamily required for sterol uptake in yeast. J Biol Chem 277:32466–32472

Wildgruber R, Harder A, Obermaier C, Boguth G, Weiss W, Fey SJ, Larsen PM, Görg A (2000) Towards higher resolution: two-dimensional electrophoresis of *Saccharomyces cerevisiae* proteins using overlapping narrow immobilized pH gradients. Electrophoresis 21:2610–2616

Wildgruber R, Reil G, Drews O, Parlar H, Gorg A (2002) Web-based two-dimensional database of *Saccharomyces cerevisiae* proteins using immobilized pH gradients from pH 6 to pH 12 and matrix-assisted laser desorption/ionization-time of flight mass spectrometry. Proteomics 2:727–732

Yuan DS (2000) Zinc-regulated genes in *Saccharomyces cerevisiae* revealed by transposon tagging. Genetics 156:45–58