3. Dujon, B. *et al. Nature* 369, 371–378 (1994).
4. Feldmann, H. *et al. EMBO J.* 13, 5795-5809 (1994).
5. Galibert, A. *et al. EMBO J.* 15, 2031–2049 (1996).
6. Sharp, P. & Lloyd, A. *Nucleic Acids Res.* 21, 179–183 (1993).
7. Rodriguez-Medina, J. R. & Rymond, B. C. *Mol. Gen. Genet.* 243, 532–539 (1994).
8. Sharp, P. M. & Li, W. H. *Nucleic Acids Res.* 15, 1281–1295 (1987).
9. Olson, M. V. in *The Molecular and Cellular Biology of the Yeast* Saccharomyces, Vol. 1 (ed. Broach J. R., Pringle. J. R. & Jones, E. W.) 1–39 (Cold Spring Harbor Laboratory Press, NY, 1991).
10. Louis, E. J. & Borts, R. *Genetics* 139, 125–136 (1995).
11. Viswanathan, M., Muthukumar, G., Lenard, J. & Cong, Y. S. *Gene* 148, 149–153 (1994).
12. Pasero, P. & Marilley, M. *Mol. Gen. Genet.* 236, 448–452 (1993).
13. Chindamporn, A., Iwaguchi, S., Nakagawa, Y., Homma, M. & Tanaka, K. *J. Gen. Microbiol.* 139, 1409–1415 (1993).
14. Rustchenko, E. P. & Sherman, F. *Yeast* 10, 1157–1171 (1994).
15. Zamb, T. & Petes, T. D. *Cell* 28, 355–364 (1982).
16. McMahon, M. E., Stamenkovich, D. & Petes, T. D. *Nucleic Acids Res.* 12, 8001–8016 (1984).
17. Kim, K. W., Kamerud, J. Q., Livingston, D. M. & Roon, R. J. *J. Biol. Chem.* 263, 11948–11953 (1988).
18. Piper, P. W., Lockheart, A. & Patel, N. *Nucleic Acids Res.* 12, 4083–4096 (1984).
19. Vassarotti, A. *et al. J. Biotechnol.* 41, 131–137 (1995).
20. Riles, L. *et al. Genetics* 134, 81–150 (1993).
21. Scholler, P., Schwarz, S. & Hoheisel, J. D. *Yeast* 11, 659–666 (1995).
22. Scholler, P. *et al. Nucleic Acids Res.* 23, 3842–3849 (1995).
23. Louis, E. J. *Biochemica* 3, 25–26 (1995).

Acknowledgements. We thank J. Warner for help with interpreting the rDNA sequence. This work was supported by the NIH National Center for Human Genome Research and the European Commision; the Belgian Federal Services for Science Policy (D.W.T.C.) and the Service Fédéraux des Affaires Scientifiques, Techniques et Culturelles, Pôles d'attraction Inter-universitaire, the Region de Bruxelles-Capitale, Region Wallonne, the Research Fund of the Katholieke Universiteit Leuven and The Wellcome Trust.

Correspondence and requests for materials should be addressed to M.J. (e-mail: mj@genetics.wustl.edu/)

The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XIII

S. Bowman, C. Churcher, K. Badcock, D. Brown, T. Chillingworth, R. Connor, K Dedman, K. Devlin, S. Gentles, N. Hamlin, S. Hunt, K. Jagels, G. Lye, S. Moule, C. Odell, D. Pearson, M. Rajandream, P. Rice, J. Skelton, S. Walsh, S. Whitehead & B. Barrell

The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK

Systematic sequencing of the genome of *Saccharomyces cerevisiae* has revealed thousands of new predicted genes and allowed analysis of long-range features of chromosomal organization. Generally, genes and predicted genes seem to be distributed evenly throughout the genome, having no overall preference for DNA strand. Apart from the smaller chromosomes, which can have substantially lower gene density in their telomeric regions[1–3], there is a consistent average of one open reading frame (ORF) approximately every two kilobases. However, one of the most surprising findings for a eukaryote with approximately 6,000 genes was the amount of apparent redundancy in its genome. This redundancy occurs both between individual ORFs and over more extensive chromosome regions, which have been duplicated preserving gene order and orientation[4–6]. Here we report the entire nucleotide sequence of chromosome XIII, the sixth-largest *S. cerevisiae* chromosome, and demonstrate that its features and organization are consistent with those observed for other *S. cerevisiae* chromosomes. Analysis revealed 459 ORFs, 284 have not been identified previously. Both intra- and interchromosomal duplications of regions of this chromosome have occurred.

Chromosome XIII of *S. cerevisiae* is 924,430 base pairs long, and contains 459 ORFs. Eight of these are TyA and TyB ORFs from four Ty1 retrotransposons present on chromosome XIII in strain AB972, two of which are located on each arm of the chromosome, and these are excluded from further analyses. The average gene density on this chromosome is one ORF for every 1,997 base pairs of DNA, which correlates well with that observed for other *S. cerevisiae* chromosomes, with 74.2% of DNA on this chromosome contributing to ORFs. An average chromosome XIII ORF is 494 codons long.

Of the 451 *S. cerevisiae* ORFs on chromosome XIII, 167 (37.0%) encode previously identified proteins. A further 281 (62.3%) predicted genes have not been previously sequenced; 121 (26.8%) of these ORFs have similarities to genes for which some biochemical information is available. However, several of this category of ORF have their best protein similarity to a protein of unknown function. A total of 160 ORFs (35.5%) encode predicted proteins that are not significantly similar to proteins of known function. Because of the rapidly advancing progress of other systematic sequencing projects, many of these ORFs have homology to hypothetical proteins both in yeasts and higher organisms. A total of 51 predicted genes have similarity only to predicted proteins of unknown function. Although the majority are most similar to another *S. cerevisiae* hypothetical protein (50.9%), several have their best homology to an ORF identified in systematic sequencing of the yeast *Schizosaccharomyces pombe*[7] (11.8%), or to predicted proteins in the nematode *Caenorhabditis elegans*[8] (17%). Thus they are members of gene families whose function is currently unknown. There were no significant protein sequence similarities for 109 ORFs, of which 11 are thought to be questionable ORFs based on their length, codon adaptation index (CAI) value and positional base preferences.

During the systematic sequencing of other chromosomes, several putative pseudogenes were identified[1,9]. These consisted of ORFs separated by a stop codon or frameshift from upstream or downstream sequences that shared a common homology to a single *S. cerevisiae* ORF. Most of these pseudogenes identified occur close to the telomeres of chromosomes. Three ORFs on chromosome XIII (YMR084W, YMR085W and YMR326C) have been classified as putative pseudogene ORFs. Of these, only YMR326C is located close to one of the chromosome telomeres; all three have strong similarity to sequences found elsewhere in the *S. cerevisiae* genome. These frameshifts have been confirmed by sequencing genomic DNA.

The average intergenic distance between adjacent ORFs depends on their relative orientation. This is certainly the case on chromosome XIII, in which 204 ORFs are arranged in tandem with an average intergenic distance of 450 base pairs. Of these, 110 are divergent and are an average 616 bp apart, and 111 are convergent and an average of 260 bp apart. This is consistent with a greater sequence requirement for the regulation of gene expression from promoter elements than for transcription termination.

Of the 451 ORFs on chromosome XIII, 24 (5.3%) are predicted to contain introns. There seems to be no preference for DNA strand, with 229 genes coded on the Watson strand and 222 on the Crick strand. There is no evidence of any significant clustering of related genes. However, there are several instances in which two very similar ORFs occur close to one another in tandem; for example, YMR169C and YMR170C/ALD2 (aldehyde dehydrogenases), and YMR006C and YMR008C/PLB1 (lysophospholipases).

The longest ORF on chromosome XIII is *HFA1*, (which is homologous to *FAS3*), a putative acetyl-CoA carboxylase that had been sequenced previously[10] (2,123 codons). A total of 39 ORFs on this chromosome are more than 1,000 codons in length. *S. cerevisiae* genes of less than 100 codons with no homology are difficult to detect[11]. On chromosome XIII, 10 ORFs shorter than 100 amino acids in length have been identified. The smallest of these is YMR248C, which is just 55 amino acids long, and may be spliced to a second small ORF immediately upstream. The smallest ORF on this chromosome that encodes a previously characterized protein is *COX7*, which is 59 amino acids long and encodes cytochrome oxidase polypeptide VII (ref.12).

Chromosome XIII encodes 21 predicted tRNA genes, of which six are spliced. In addition to the four Ty1 retrotransposons, several long terminal repeat (LTR) sequences are present, providing evidence of previous trans-

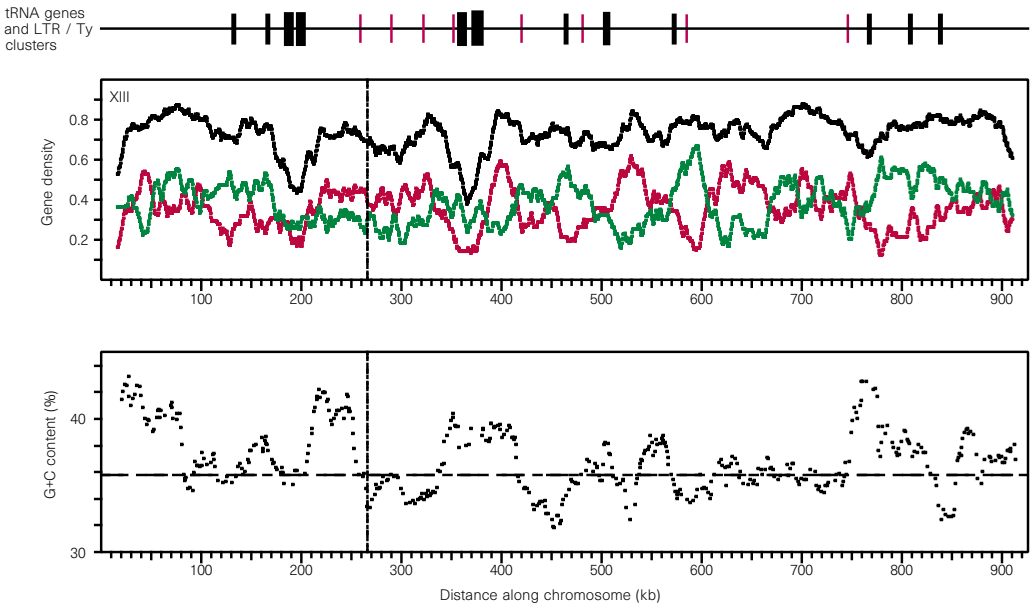

Figure 1 Overall molecular architecture of chromosome XIII. Variation in gene density (top) and base composition (bottom) along the sequence-based map of chromosome XIII (scale in kilobases from left telomere). Veritical broken lines indicate the position of the centromere. Gene density is expressed as the probability for each nucleotide to be part of an ORF. It was calculated using sliding windows of 30 kb (steps of 0.5 kb) for the Watson strand alone (red line), the Crick strand alone (green line), and the sum of both strands (black line). Percentage of G+C was calculated from the silent positions of codons using a sliding window of 13 consecutive ORFs (horizontal broken line indicates average percentage G+C% at silent positions of codons as 35.8). Top line, positions of tRNA genes, solo LTR or Ty elements (thin small vertical lines), or clusters of them (thick small vertical lines) along the chromosome map.

position events on this chromosome. Of the 30 LTRs (whole or partial) identified on this chromosome, 29 are in close proximity to tRNA genes. There are 24 delta elements, with eight of these flanking the four retrotransposons and 16 solo elements. A further five LTRs resemble tau elements, and there is evidence of a single sigma element.

The telomeres of chromosome XIII are highly similar to the telomeres of several other *S. cerevisiae* chromosomes. Adjacent to the terminal $C_{1-3}A$ telomeric repeat at the left telomere is a Y′ element, which is separated from a core X element by the subtelomeric repeats STR-D, STR-C, STR-B and STR-A[13]. The right telomere of chromosome XIII conforms to this structure but does not contain a Y′ element. The right telomere shares a region of 4 kb in which the sequence is almost identical to that found at the right telomere of chromosome XV (EMBL database, accession no. SC23472). The centromere of chromosome XIII is located between bases 268,031 and 268,150; it conforms to the consensus sequence derived from the centromeres of other *S. cerevisiae* chromosomes[14].

Comparison of the positions of genes on chromosome XIII with their corresponding genetically mapped loci[15] shows that, as for most *S. cerevisiae* chromosomes, the two are generally in agreement. Several genes are incorrectly positioned on the genetic map in comparison with the sequence-derived map; for example, *van1* is much closer to the left telomere of chromosome XIII than expected. Small inversions are evident, for example between *UPF1* and *ADH3*, but no gross discrepancies have occurred as observed for chromosome XI (ref. 16). *MEL6* has been mapped to this chromosome, but in the *S. cerevisiae* strain S288C, from which AB972 was derived, this locus is missing. Also, the locus *SUP8* has been only tentatively assigned to the tRNA(Tyr) at base 837928 as a second tRNA(Tyr) is present on the left arm of the chromosome.

Chromosome XIII contains both intra- and interchromosomal duplications of both single ORFs and more extensive regions. The largest intrachromosomal duplication consists of two regions approximately 40 kb to 50 kb in length containing six homologous genes in the same order and orientation with respect to each other; the duplication has occurred between the left and right arms of the chromosome. This encompasses the region from base 32,334 to 73,917 and 790,207 to 840,147. Smaller instances of tandem gene duplication have also occurred within chromosome XIII; for example, YML125C and YML124C (*TUB3*) are similar to YML087C and YML085C (*TUB1*) on the short arm of the chromosome.

The largest interchromosomal duplication observed has occurred between a 200 kb region on the right arm of chromosome XIII and an equivalent region on the left arm of chromosome XI (in the opposite orientation). On chromosome XIII the region spans bases 303,238 to 502,733 and is bounded by ORFs YMR016C and YMR118C, on chromosome XI the coordinates of this repeat occur between bases 179,672 and 357,489. Of a total of approximately 90 genes, 17 show significant similarity and are in the same order and orientation on each chromosome. Further evidence of large scale duplications exists, including two separate regions on both chromosome IV and XVI.

Chromosome XIII has been analysed for variations in base composition, as described in ref. 9. The percentage G+C content in the third position of each codon varies throughout the length of chromosome XIII (Fig. 1), as observed through long-range analysis of other *S. cerevisiae* chromosomes. Detailed analysis at the level of individual ORFs (Fig. 2) shows that, as for chromosome IX, regions of high third-position G+C are shorter and contain fewer ORFs that regions of chromosome III that are rich in G+C. Comparison with chromosome IX shows that regions of high G+C are slightly less evident on chromosome XIII. This is also seen when plotting total G+C content for these two chromosomes: peaks of high G+C are generally lower in magnitude for chromosome XIII (results not shown).

Local areas of high G+C composition have also been observed in several of the intergenic regions of this chromosome, for example regions at approximately 119, 306, 541 and 653 kb are G+C rich. It is possible that these intergenic regions of high G+C may contain previously unidentified, small *S. cerevisiae* genes. These areas do not appear to correlate with sequences of high coding potential. However, this does not exclude the

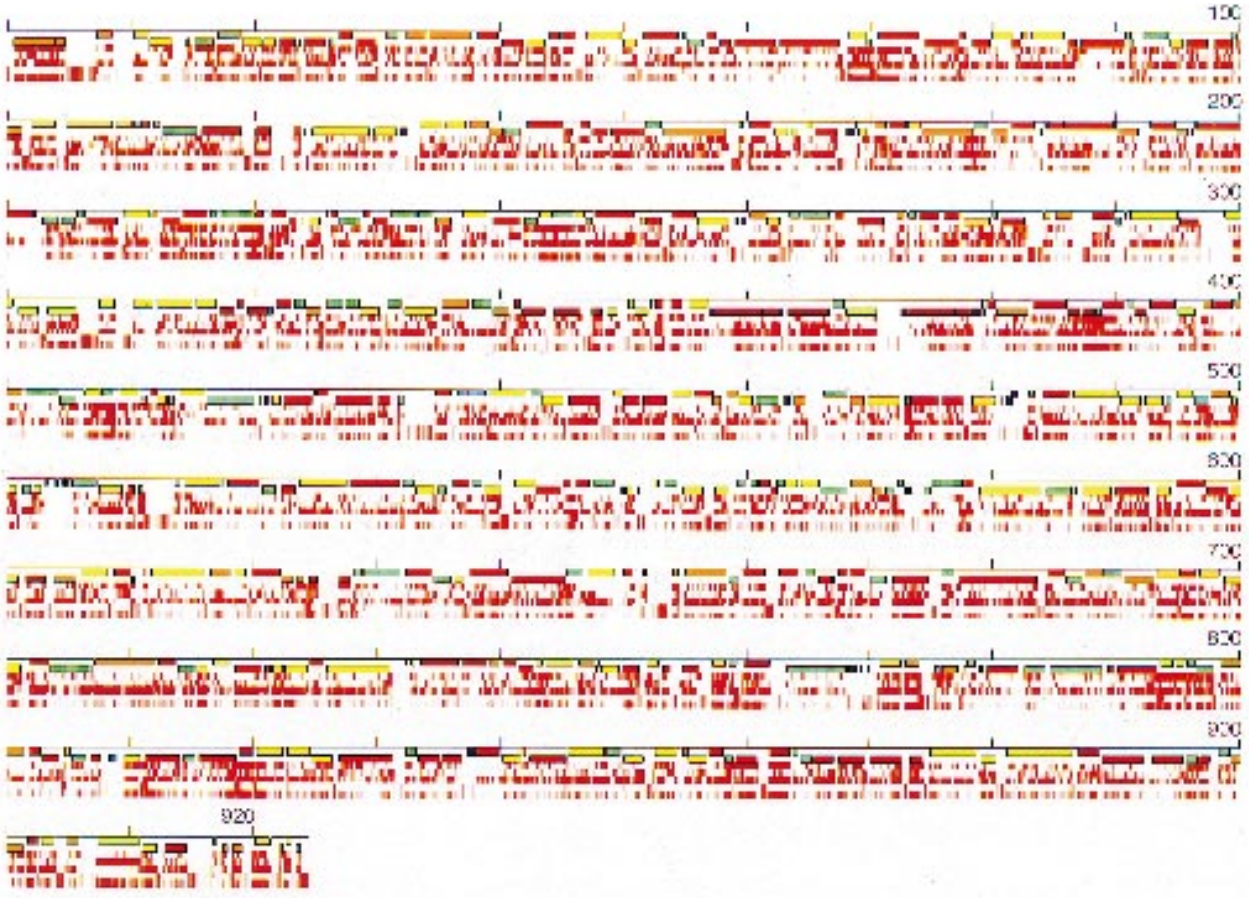

Figure 2 G+C composition of chromosome XIII. Chromosome XIII and its features are drawn to scale. The top graduated line represents the chromosome split into segments of 100 kb, with ORFs indicated below this as coloured boxes. ORFs located on the Watson strand are shown above those on the Crick strand. ORFs encoding previously identified genes are shown in red, those with similarities to hypothetical proteins in orange, and those with no significant similarities in green; pseudogene ORFs are shown in blue. tRNA genes and transposon-derived ORFS are shown in white boxes, with LTRs shown in dark blue (delta), turquoise (tau) and pink (sigma). Below this, variations in G+C composition (calculated using a sliding window of 200 bases) are shown as bars, with gradations of red varying from 35% to 45% G+C content (P. R., unpublished data). Areas of less than 35% are shown in white, with those over 45% red. Five bars of G+C variation are shown, the lowest bar shows total G+C content; above this a second shows G+C content in intergenic regions alone. The G+C composition in each of the three bases of each codon are shown above this, with the central of the five bars representing first position G+C, the next representing second position G+C, and the top bar showing third position G+C.

possibility that coding sequences exist in these areas. No similarities to other *S. cerevisiae* ORFs were found using the program BLASTX, indicating that these areas rich in G+C are not pseudogene remnants. Other possibilities for these G+C peaks in intergenic regions have not yet been fully investigated.

Few ORFs show a high incidence of high G+C content in the second position of their codons. The most pronounced example on chromosome XIII is YMR317W, which encodes a protein rich in serine and threonine residues, and is located close to the right telomere of the chromosome.

Analysis of other *S. cerevisiae* chromosomes has demonstrated that, in general, areas high in G+C content correlate with high gene density, and that regions around chromosome centromeres and telomeres appear to be G+C poor[4–6,16]. This is also the case for chromosome X111. However, the significance of these areas of high and low G+C content is not yet clear.

Two of the genes located on chromosome XIII have human homologues that have been implicated in certain forms of cancer. *MLH1* encodes a DNA mismatch repair protein that is very similar to a human protein defective in some forms of hereditary non-polyposis colon cancer[17]. In *S. cerevisiae,* null mutants of this gene are viable, but show an elevated rate of spontaneous mutations and an increased instability of simple repeat sequences[18,19]. A second gene, *SGS1*, is homologous to the human *BLM* gene, which is involved in Bloom's syndrome[20]. Both the human and yeast genes encode DNA helicases. Mutations in the human gene confer a predisposition to many types of cancer and also cause other clinical defects[20]. In *S. cerevisiae*, null mutants are viable, but again show genomic instability[21,22]. Future analysis of these and other *S. cerevisiae* genes should assist the understanding of the molecular mechanisms underlying many human diseases.

Chromosome XIII is the largest *S. cerevisiae* chromosome to be sequenced by a single laboratory. Analysis shows that its features are typical of large *S. cerevisiae* chromosomes. □

Methods

Sequencing. Chromosome XIII was sequenced using a cosmid-based strategy that uses DNA isolated from the S288C-derived strain AB972. Methods used for sequence generation and assembly have been described in detail elsewhere[8,23]. Cosmids were chosen from the map generated by L. Riles and M. Olson (personal communication), giving rise to two large contigs with a central gap. This was filled using lambda clones L-6543 and L-6223. The two chromosome telomeres, pEL161H (left) and pEL175H (right), were provided as plasmid clones[24] by E. Louis. Two further lambda clones, L-4987 and L-7056, were sequenced to complete the left end of the chromosome. Sequence from pEL161H overlapped L-4987, so no further gap filling was required at the left telomere of chromosome XIII. However, at the right telomere a gap between cosmid clone 9924 and plasmid pEL175H remained, and

was filled using long-range polymerase chain reaction (PCR) from *S. cerevisiae* genomic DNA. The sequence generated for this chromosome extends into the $C_{1-3}A$ telomeric repeat sequences on both chromosome arms, although the exact number of these repeats has not been determined. Sequencing was considered to be finished when each base had been sequenced on both strands and all ambiguities had been resolved.

Analysis. For each completed clone, a consensus of the nucleotide sequence was generated in the Staden sequence assembly package XBAP[25], flanked by short regions of sequence overlapping neighbouring clones. This sequence was analysed primarily within the DIANA (Display and Analyse) package (T. Horsnell and B. B., unpublished), a sequence editor with a graphical interface. ORFs equal to or greater than 100 codons in length were marked and trimmed to their first methionine. Each ORF was screened against the SWIR database, a non-redundant compilation of the protein databases Swiss-Prot[26], TrEMBL[27] and WormPep, using the program FASTA[28] with limited optimization. The consensus sequence for each clone was screened against SWIR using BLASTX[29], and EMBL/EMNEW using BLASTN[29], to detect small ORFs less than 100 amino acids in length, other genome features, and local similarity. Some features were specifically identified; Prosite[30] amino-acid motifs (regular expression searching), transposon LTRs (GCG Wordsearch/Segments) and tRNAs (tRNA scan). Individual annotated clones were submitted to the EMBL database within days of being finished. The complete chromosomal sequence was built from overlapping clones and also submitted to the EMBL database as a single record (accession no. SCCHRXIII, Z271257).

Received 26 July 1996; accepted 11 March 1997.

1. Bussey, H. *et al. Proc. Natl Acad. Sci. USA* 92, 3809–3813 (1995).
2. Oliver, S. G. *et al. Nature* 357, 38–46 (1992).
3. Murakami, Y. *et al. Nature Genet.* 10, 261–268 (1995).
4. Feldmann, H. *et al. EMBO J.* 13, 5795–5809 (1994).
5. Johnston, M. *et al. Science* 265, 2077–2082 (1994).
6. Galibert, F. *et al. EMBO J.* 15, 2031–2049 (1996).
7. http://www.sanger.ac.uk/yeast/pombe.html
8. Wilson, R. *et al. Nature* 368, 32–38 (1994).
9. Churcher, C. *et al*, *Nature* (this issue).
10. Kearsey, S.E. *DNA Sequence* 4, 69–70 (1993)
11. Termier, M. & Kalogeropoulos, A. *Yeast*, 12, 369–384 (1996)
12. Calder, K. M. & McEwen, J. E. *Nucleic Acids Res.* 18, 1632 (1990).
13. Louis, E. J., Naumova, E. S., Lee, A., Naumov, G. & Haber, J. E. *Genetics* 136, 789–802 (1994).
14. Fitzgerald–Hayes, M. *Yeast* 3, 187–200 (1987).
15. Mortimer, R.K., Cherry, J.M., Dietrich, F.M., Riles, L., Olson, M.S. & Botstein, D. http://genome.www.stanford.edu/saccdb/edition12.html (1995).
16. Dujon, B. *et al. Nature* 369, 371–378 (1994).
17. Papadopoulos, N. *et al. Science* 263, 1625–1629 (1994).
18. Strand, M., Earley, M.C., Crouse, G.F. & Petes, T.D. *Proc. Natl Acad. Sci. USA* 92, 10418–10421 (1995).
19. Prolla, T. A., Christie, D. M. & Liskay, R. M. *Mol. Cell. Biol.* 14, 407–415 (1994).
20. Ellis, N.A. *et al. Cell* 83, 655–666 (1995).
21. Gangloff, S., McDonald, J. P., Bendixen, C., Arthur, L. & Rothstein, R. *Mol. Cell. Biol.* 14, 8391–8398 (1994).
22. Watt, P.M., Louis, E.J., Borts, R.H. & Hickson, I.D. *Cell* 81, 253–260 (1995).
23. Smith, V. *et al. Methods Enzmol.* 218, 173–187 (1993).
24. Louis, E. J. *Biochemica* 3, 25–26 (1995).
25. Dear, S. & Staden, R. *Nucleic Acids Res.* 19, 3907–3911 (1991).
26. Bairoch, A. *Nucleic Acids Res.* 19, 2247–2249 (1991).
27. Bairoch, A. & Apweiler, R. *Nucleic Acids Res.* 24, 21–25 (1996).
28. Pearson, W.R. & Lipman, D.J. *Proc. Natl Acad. Sci. USA* 85, 2444–2448 (1988).
29. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D. J. *J. Mol. Biol.* 215 403–410 (1990).
30. Bairoch, A., Butcher, P. & Hoffman, K. *Nucleic Acids Res.* 24, 189–196 (1995).

Acknowledgements. We thank A. Fraser for cosmid DNA preparation; M. Jones and the subcloning group for library preparation; the staff in the gel-pouring and media kitchens for their help; the computer support and software development groups and R. Staden for software support; we thank L. Riles, M. Olsen and E. Louis for gifts of cosmid, lambda and plasmid clones; B. Dujon for providing Fig. 1 using unpublished software developed in collaboration with C. Marck, and D. Harris, J. Sulston and K. Plucknett for critical reading of the manuscript. This work was funded by the Wellcome Trust.

Correspondence and requests for materials should be addressed to B.B. (e-mail: barrell@sanger.ac.uk).
Clone accession numbers and other information can be found on http://www.sanger.ac.uk/yeast/home.html.

The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XIV and its evolutionary implications

P. Philippsen[1,2], K. Kleine[3], R. Pöhlmann[1], A. Düsterhöft[4], K. Hamberg[2], J. H. Hegemann[2], B. Obermaier[5,6], L. A. Urrestarazu[7], R. Aert[8], K. Albermann[3], R. Altmann[1], B. André[7], V. Baladron[9], J. P. G. Ballesta[10], A.-M. Bécam[11], J. Beinhauer[2], J. Boskovic[10], M. J. Buitrago[9], F. Bussereau[12], F. Coster[13], M. Crouzet[14], M. D'Angelo[15], F. Dal Pero[15], A. De Antoni[15], F. Del Rey[9], F. Doignon[14], H. Domdey[5], E. Dubois[16], T. Fiedler[2], U. Fleig[2], M. Floeth[4], C. Fritz[4], C. Gaillardin[17], J. M. Garcia-Cantalejo[10], N. N Glansdorff[16], A. Goffeau[13], U. Gueldener[2], C. Herbert[11], K. Heumann[3], D. Heuss-Neitzel[4], H. Hilbert[4], K. Hinni[1], I. Iraqui Houssaini[7], M. Jacquet[12], A. Jimenez[10], J.-L. Jonniaux[13], L. Karpfinger[3], G. Lanfranchi[15], A. Lepingle[17], H. Levesque[17], R. Lyck[2], M. Maftahi[17], L. Mallet[12], K. C. T. Maurer[18], F. Messenguy[16], H. W. Mewes[3], D. Möstl[4], F. Nasr[11], J.-M. Nicaud[17], R. K. Niedenthal[2], D. Pandolfo[15], A. Piérard[16], E. Piravandi[5], R. J. Planta[18], T. M. Pohl[19], B. Purnelle[13], C. Rebischung[1], M. Remacha[10], J. L. Revuelta[9], M. Rinke[5], J. E. Saiz[9], F. Sartorello[15], B. Scherens[16], M. Sen-Gupta[2], A. Soler-Mira[10], J. H. M. Urbanus[18], G. Valle[15], L. Van Dyck[13], P. Verhasselt[8], F. Vierendeels[16], S. Vissers[7], M. Voet[8], G. Volckaert[8], A. Wach[1], R. Wambutt[20], H. Wedler[20], A. Zollner[3] & J. Hani[3]

[1]Institute for Applied Microbiology, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland
[2]Justus-Liebig-Universität Giessen, Institut für Mikro- und Molekularbiologie, Frankfurter Strasse 107, D-35392 Giessen, Germany
[3]Martinsrieder Institut für Protein Sequenzen, Max-Planck-Institut für Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, Germany
[4]QIAGEN GmbH, Max-Volmer-Strasse 4, D-40724 Hilden, Germany
[5]Laboratorium für Molekulare Biologie, Genzentrum der LMU München, Feodor-Lynen-Strasse 25, D-811377 München, Germany
[6]MediGene GmbH, Lochhamer Strasse 11, D-82152 Martinsried, Germany
[7]Université Libre de Bruxelles, Physiologie Cellulaire et Génétique des Levures, Boulevard du Triompe CP244, B-1050 Bruxelles, Belgium
[8]Katholieke Universiteit Leuven, Laboratory of Gene Technology, Willem de Croylaan 42, B-3001 Leuven, Belgium
[9]Departamento de Microbiologia y Genética, Universidad de Salamanca, Avenida del Campo Charro s/n, E-37007 Salamanca, Spain
[10]Centro de Biologia Molecular, CSIC & UAM, Cantoblanco, E-28049 Madrid, Spain
[11]Centre de Génétique Moléculaire, Laboratoire propre du CNRS associé à l'Université Pierre et Marie Curie, F-91198 Gif-sur-Yvette, France
[12]Université Paris-Sud, Institut de Génétique et Microbiologie, Laboratoire Information Génétique et Développement, Bât. 400, F-91405 Orsay Cedex, France
[13]Unité de Biochimie Physiologique, Université Catholique de Louvain, Place Croix du Sud 2/20, B-1348 Louvain-la-Neuve, Belgium
[14]LBMS, Université de Bordeaux 2, UPR CNRS 9026, BP 64, 146 rue Léo Saignat, F-33076 Bordeaux Cedex, France
[15]Department of Biology, CRIBI Biotechnology Centre, University of Padova, via Trieste, 75, I-35121 Padova, Italy
[16]CERIA-COOVI, Avenne E. Gryson 1, B-1070 Brussels, Belgium
[17]Institut National Agronomique Paris-Grignon, Laboratoire de Génétique Moléculaire et Cellulaire, Centre de Biotechnologies Agro-Industrielles, F-78850 Thiverval-Grignon, France
[18]Department of Biochemistry and Molecular Biology, IMBW, BioCentrum Amsterdam, Vrije Universiteit de Boelelaan 1083, NL-1081 HV Amsterdam, Netherlands
[19]GATC-Gesellschaft für Analyse-Technik und Consulting mbH, Fritz-Arnold-Strasse 23, D-78467 Konstanz, Germany
[20]AGON GmbH, Glienicker Weg 185, D-12489 Berlin, Germany