anti-leukaemia treatments have historically sought to eliminate proliferating cells. But, like normal HSCs, L-HSCs rarely proliferate and so are resistant to such therapies. The identification of L-HSC-specific genetic targets, especially those involved in the enhanced self-renewal seen in L-HSCs, could lead to new treatments that stop the disease at its source.

John E. Dick is in the Division of Cell and Molecular Biology, University Health Network, and the Department of Molecular Genetics and Microbiology, University of Toronto, 620 University Avenue, Toronto, Ontario M5G 2C1, Canada. E-mail: jdick@uhnres.utoronto.ca


15 MAY 2003

Genomics: Yeast rises again
Steven L. Salzberg

What’s in a genome? The short answer is that you can’t really say in detail for any one species until you have the genome sequence of a variety of other species — some closely related, others less so — to compare it with.

While the human and mouse genomes have lately dominated public discussions of genome science, yeast researchers have quietly continued to forge ahead with analysing Saccharomyces cerevisiae and related yeast species. Yeast has long been a favourite of biologists, who use it as a model for investigating the biology of higher organisms. Although yeasts are unicellular, like bacteria, they have a nucleus, putting them in the biological kingdom of euakaryotes (Eukaryota), a group that includes humans. Yeast was the first eukaryote to have its genome completely sequenced, in 1996.

And yeast has long been a part of human commerce and culture, being used in brewing and baking since ancient times. Saccharomyces cerevisiae itself was first isolated in beer in 1837.

On page 241 of this issue, Kellis and colleagues describe the results of sequencing three more yeast species — Saccharomyces paradoxus, S. bayanus and S. mikatae — and comparing them with S. cerevisiae. Two of these species, S. bayanus and S. paradoxus, are used in winemaking, but the third new species were chosen mainly because they are evolutionarily very close to S. cerevisiae, and the implications of the study go well beyond brewing and viniculture. The analyses will produce a substantial revision in our knowledge of the yeast genome, and provide strategic directions for how we might select other sequencing targets to advance understanding of the human genome.

The original S. cerevisiae project was a model of international collaboration among small laboratories: 55% of the genome was sequenced by a network of European labs, with the rest of it being generated at five large (at the time) centres. More than 600 scientists participated in the project, sequencing a cosmid (a 40-kilobase fragment) at a time with the rest of it being generated at five large centres. More than 600 scientists participated in the project, sequencing a cosmid (a 40-kilobase fragment) at a time with the rest of it being generated at five large centres.
species (including S. bayanus), and failing to find matches for 742 genes, estimated that the gene count should be reduced to 5,651. The crux of their argument was the assumption that if a gene is functional, it should be conserved among closely related species (see Fig. 1). Because the Génolevures project covered only 20–40% of each genome, the possibility remained that those ‘missing’ genes might be found in the unsequenced regions.

Kellis and colleagues have now closed the door on that possibility: their sequence data cover 98% of two species and 93% of the third. Based on sequence alignments among the species, they have largely confirmed the results of the earlier study and argue that 503 genes should be deleted from the yeast catalogue, leaving 5,726 genes, of which 43 are newly discovered in their study.

Regulatory sequences, which sit outside genes and turn them on and off, are the key to understanding how a genome fits together. Whether we are comparing human and mouse, or yeast and yeast, we still have to answer the puzzling question of how seemingly huge differences in physical, biochemical or behavioural characteristics can result from sometimes tiny differences in the protein sequences. Regulatory sites occur virtually anywhere in the vast areas between protein-coding regions, and they can be identified because — unlike non-functional regions — they are conserved. The closer two species are, the more regulatory sites they are likely to share. Unfortunately, if the species are close enough, many pieces of non-functional DNA will be conserved merely by chance. A nice solution to this problem is to sequence DNA that has structural similarity to human as these four yeast species share, and even those have different chromosome numbers: chimpanzee (Pan troglodytes) has 24, gibbon (Hylobates concolor) has 25, and macaque (Macaca fuscata) has 21. At the sequence level, we might gain more information from studying more distant mammals, such as cat (Felis catus), pig (Sus scrofa) and dolphin (Tursiops truncatus).

This new study of yeast genomes makes it clear that comparative genomics sequencing has tremendous analytical power: it offers the prospect of enhancing our knowledge of thousands of genes at once as well as providing fresh clues about the function of the vast amount of genomic DNA that does not encode genes. As it has done before, lowly yeast shows us a path towards a better understanding of our own biology.

Steven L. Salzberg is at The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, Maryland 20850, USA.

e-mail: salzberg@tigr.org

Molecular biology
Disruptive influence
Marco Foiani

Recombination is a vital cellular process implicated in DNA metabolism — but it must be tightly controlled. The discovery of a protein that disrupts recombination intermediates sheds light on the control mechanisms.

On pages 305 and 309 of this issue, Krejci and co-workers and Veaute and colleagues describe a biochemical mechanism that controls the genome ‘shuffling’ occurring in dividing cells and in DNA repair. Their findings have implications for how genome stability is maintained, and hence for the development of cancer.

Genome shuffling is referred to as ‘recombination’, and is a cellular process by which extensive tracts of DNA are moved from one part of the genome to another. There are several recombination pathways, some not yet well characterized, which are routinely used by normal cells to repair damaged chromosomes, to assist in DNA synthesis, and even to regulate gene expression. Recombination also occurs during the production of eggs and sperm, in which its function is to mix the genetic information such that each egg and each sperm is genetically different.

Despite its importance, however, recombination can sometimes be harmful: it can generate damaging genomic rearrangements, as well as intermediate structures that cannot be processed normally. Cells need to coordinate recombination with other responses to DNA damage, with progression through the cell-division cycle, and with chromosome replication.

Otherwise, cells invariably become genetically unstable as the protein that disrupts recombination take over the chromosomes. During DNA replication, for instance, the double helix unwinds and separates, and the two strands are used as templates to make another helix. Replication frequently stalls, and a ‘checkpoint’ ensures that the separated DNA (the replication fork) maintains its integrity during these pauses. Without this checkpoint, abnormal replication intermediates form and are processed by unscheduled recombination. In addition, in some inherited human diseases — Werner, Bloom and Rothmund–Thomson syndromes — mutations in enzymes implicated in DNA metabolism (DNA helicases) cause increased recombination, genome instability and a predisposition to cancer.

So cells must have mechanisms to control recombination and to prevent harmful chromosome rearrangements. One possible mechanism in yeast involves the Srs2 protein (a human relative of which has not yet been discovered, but it is surely only a matter of time). Srs2 is another DNA helicase — it can unwind double helices — and it has previously been implicated in DNA replication, in restarting the cell cycle after DNA-damage-