A Gene Map of the Human Genome

G. D. Schuler,* M. S. Boguski, E. A. Stewart, L. D. Stein, G. Gyapay, K. Rice, R. E. White, P. Rodriguez-Tomé, A. Aggarwal, E. Bajorek, S. Bentolila, B. B. Birren, A. Butler, A. B. Castle, N. Chiannilkulchai, A. Chu, C. Clee, S. Cowles, P. J. R. Day, T. Dibling, N. Drouot, I. Dunham, S. Duprat, C. East, C. Edwards, J.-B. Fan, N. Fang, C. Fizames, C. Garrett, L. Green, D. Hadley, M. Harris, P. Harrison, S. Brady, A. Hicks, E. Holloway, L. Hui, S. Hussain, C. Louis-Dit-Sully, J. Ma, A. MacGilvery, C. Mader, A. Maratukulam, T. C. Matise, K. B. McKusick, J. Morissette, A. Mungall, D. Muselet, H. C. Nusbaum, D. C. Page, A. Peck, S. Perkins, M. Piercy, F. Qin, J. Quackenbush, S. Ranby, T. Reif, S. Rozen, C. Sanders, X. She, J. Silva, D. K. Slonim, C. Soderlund, W.-L. Sun, P. Tabar, T. Thangarajah, N. Vega-Czarny, D. Vollrath, S. Voyticky, T. Wilmer, X. Wu, M. D. Adams, C. Auffray, N. A. R. Walter, R. Brandon, A. Dehejia, P. N. Goodfellow, R. Houlgatte, J. R. Hudson Jr., S. E. Ide, K. R. Iorio, W. Y. Lee, N. Seki, T. Nagase, K. Ishikawa, N. Nomura, C. Phillips, M. H. Polymeropoulos, M. Sandusky, K. Schmitt, R. Berry, K. Swanson, R. Torres, J. C. Venter, J. M. Sikela, J. S. Beckmann, J. Weissenbach, R. M. Myers, D. R. Cox, M. R. James, D. Bentley, P. Deloukas, E. S. Lander, T. J. Hudson

The human genome is thought to harbor 50,000 to 100,000 genes, of which about half have been sampled to date in the form of expressed sequence tags. An international consortium was organized to develop and map gene-based sequence tagged site markers on a set of two radiation hybrid panels and a yeast artificial chromosome library. More than 16,000 human genes have been mapped relative to a framework map that contains about 1000 polymorphic genetic markers. The gene map unifies the existing genetic and physical maps with the nucleotide and protein sequence databases in a fashion that should speed the discovery of genes underlying inherited human disease. The integrated resource is available through a site on the World Wide Web at http://www.ncbi.nlm.nih.gov/SCIENCE96/.

Central to the description of an organism's genome is a comprehensive catalog of the sequence and location of all its genes. Gene maps are now available for those organisms whose complete genomic sequence has been determined, including 141 viruses, 51 organelles, two eubacteria, one archeon, and one eukaryote (the yeast, *Saccharomyces cerevisiae*) (*1*). Such a map of the human genome should become available by 2005, as a result of the efforts by the Human Genome Project to determine the complete 3 billion nucleotides of the human DNA sequence and develop suitable computer and laboratory tools for recognizing genes.

In view of the tremendous value of a human gene map for biomedical research, it is not reasonable to wait until the complete sequence is available to begin preparing such a map. There are compelling reasons for constructing a series of increasingly comprehensive gene maps and cross-referencing them to the human genetic map. A key application is the positional cloning (Fig. 1) of disease-causing genes. Genetic mapping of affected families with polymorphic markers that span the genome permits localization of the disease gene to a candidate region, often in the range of 2 to 5 megabases (Mb). Such intervals are physically mapped with overlapping DNA clones, which usually serve as substrates to identify genes ("transcripts") in the region. Subsequently, the genes are scrutinized for the presence of sequence mutations in affected individuals. Regional transcript mapping by current methods, which is difficult and time-consuming, would be supplanted by the availability of a comprehensive, whole-genome gene map. Such a resource would accelerate gene searches for simple Mendelian traits and is essential in the case of complex (polygenic) traits, for which limited genetic resolution will necessitate sifting through multimegabase regions. The availability of an expanding gene inventory for any candidate region is predicted to make the "positional candidate" approach the predominant method for cloning human disease genes

G. D. Schuler and M. S. Boguski, National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, 8600 Rockville Pike, Bethesda, MD 20894, USA. E. A. Stewart, A. Aggarwal, E. Bajorek, A. Chu, S. Cowles, J.-B. Fan, N. Fang, D. Hadley, M. Harris, S. Brady, S. Hussain, C. Mader, A. Maratukulam, K. B. McKusick, S. Perkins, M. Piercy, F. Qin, J. Quackenbush, T. Reif, C. Sanders, X. She, W.-L. Sun, P. Tabar, D. Vollrath, S. Voyticky, R. M. Myers, D. R. Cox, Department of Genetics, Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305, USA. L. D. Stein, B. B. Birren, A. B. Castle, L. Hui, J. Ma, H. C. Nusbaum, D. C. Page, S. Rozen, J. Silva, D. K. Slonim, X. Wu, Whitehead Institute for Biomedical Research, Massachusetts Institute of Technology Center for Genome Research, 9 Cambridge Center, Cambridge, MA 02142, USA. G. Gyapay, S. Bentolila, N. Chiannilkulchai, N. Drouot, S. Duprat, C. Fizames, D. Muselet, N. Vega-Czarny, J. S. Beckmann, J. Weissenbach, Généthon, CNRS URA 1922, 1 rue de l'Internationale, 91000 Evry, France. K. Rice, A. Butler, C. Clee, T. Dibling, I. Dunham, C. East, C. Edwards, C. Garrett, L. Green, P. Harrison, A. Hicks, E. Holloway, A. MacGilvery, A. Mungall, A. Peck, S. Ranby, C. Soderlund, T. Wilmer, D. Bentley, P. Deloukas, The Sanger Centre, Hinxton Hall, Hinxton, Cambridge CB10 1SA, UK. R. E. White, P. J. R. Day, C. Louis-Dit-Sully, T. Thangarajah, M. R. James, Wellcome Trust Centre for Human Genetics, Nuffield Department of Clinical Medicine, University of Oxford, Windmill Road, Oxford OX3 7BN, UK. P. Rodriguez-Tomé, European Molecular Biology Laboratory Outstation, Hinxton, The European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK. T. C. Matise, Laboratory of Statistical Genetics, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA. J. Morissette, Centre de Recherche du Centre Hospitalier de l'Universite Laval, 2705 Boulevard Laurier, Ste-Foy, Quebec G1V 4G2, Canada. M. D. Adams, R. Brandon, C. Phillips, M. Sandusky, J. C. Venter, The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850, USA. C. Auffray and R. Houlgatte, Genexpress, CNRS UPR 420, 7-19 rue, Guy Moquet-Batiment G, 94801 Villejuif, France. N. A. R. Walter, K. R. Iorio, R. Berry, J. M. Sikela, Department of Pharmacology and Molecular Biology Program, University of Colorado Health Sciences Center, 4200 E. Ninth Avenue, Denver, CO 80262, USA. A. Dehejia, S. E. Ide, M. H. Polymeropoulos, R. Torres, Laboratory of Genetic Disease Research, National Center for Human Genome Research, National Institutes of Health, Bethesda, MD 20892, USA. P. N. Goodfellow and K. Schmitt, Department of Genetics, Cambridge University, Tennis Court Road, Cambridge CB2 3EH, UK. J. R. Hudson Jr., W. Y. Lee, K. Swanson, Research Genetics, 2130 S. Memorial Parkway, Huntsville, AL 35801, USA. N. Seki, T. Nagase, K. Ishikawa, N. Nomura, Kazusa DNA Research Institute, 1532-3 Yana, Kisarazu, Chiba 292, Japan. E. S. Lander, Whitehead Institute for Biomedical Research, Massachusetts Institute of Technology Center for Genome Research, 9 Cambridge Center, Cambridge, MA 02142, USA, and Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. T. J. Hudson, Whitehead Institute for Biomedical Research, Massachusetts Institute of Technology Center for Genome Research, 9 Cambridge Center, Cambridge, MA 02142, USA, Departments of Medicine and Human Genetics and Montreal General Hospital Research Institute, McGill University, Montreal H3G 1A4, Canada.

*To whom correspondence should be addressed.