

## **Chapter VII**

### **References**

## 7.1 References

- Abecasis G. R., Noguchi E., Heinzmann A., Traherne J. A., Bhattacharyya S., Leaves N. I., Anderson G. G., Zhang Y., Lench N. J., Carey A., Cardon L. R., Moffatt M. F., and Cookson W. O. (2001). Extent and distribution of linkage disequilibrium in three genomic regions. *Am J Hum Genet* **68**: 191-197.
- Adams M. D., Celniker S. E., Holt R. A., Evans C. A., Gocayne J. D., Amanatides P. G., Scherer S. E., Li P. W., Hoskins R. A., Galle R. F., George R. A., Lewis S. E., Richards S., Ashburner M., Henderson S. N., Sutton G. G., Wortman J. R., Yandell M. D., Zhang Q., Chen L. X., Brandon R. C., Rogers Y. H., Blazej R. G., Champe M., Pfeiffer B. D., Wan K. H., Doyle C., Baxter E. G., Helt G., Nelson C. R., Gabor G. L., Abril J. F., Agbayani A., An H. J., Andrews-Pfannkoch C., Baldwin D., Ballew R. M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E. M., Beeson K. Y., Benos P. V., Berman B. P., Bhandari D., Bolshakov S., Borkova D., Botchan M. R., Bouck J., Brokstein P., Brottier P., Burtis K. C., Busam D. A., Butler H., Cadieu E., Center A., Chandra I., Cherry J. M., Cawley S., Dahlke C., Davenport L. B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A. D., Dew I., Dietz S. M., Dodson K., Doup L. E., Downes M., Dugan-Rocha S., Dunkov B. C., Dunn P., Durbin K. J., Evangelista C. C., Ferraz C., Ferriera S., Fleischmann W., Fosler C., Gabrielian A. E., Garg N. S., Gelbart W. M., Glasser K., Glodek A., Gong F., Gorrell J. H., Gu Z., Guan P., Harris M., Harris N. L., Harvey D., Heiman T. J., Hernandez J. R., Houck J., Hostin D., Houston K. A., Howland T. J., Wei M. H., Ibegwam C., *et al.* (2000). The genome sequence of *Drosophila melanogaster*. *Science* **287**: 2185-95.
- Adams M. D., Kelley J. M., Gocayne J. D., Dubnick M., Polymeropoulos M. H., Xiao H., Merril C. R., Wu A., Olde B., Moreno R. F., and *et al.* (1991). Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* **252**: 1651-6.
- Ahringer J. (1997). Turn to the worm! *Curr Opin Genet Dev* **7**: 410-5.
- Altschul S. F., Gish W., Miller W., Myers E. W., and Lipman D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**: 403-10.
- Altschul S. F., Madden T. L., Schaffer A. A., Zhang J., Zhang Z., Miller W., and Lipman D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389-402.
- Altshuler D., Pollara V. J., Cowles C. R., Van Etten W. J., Baldwin J., Linton L., and Lander E. S. (2000). An SNP map of the human genome generated by reduced representation shotgun sequencing. *Nature* **407**: 513-6.
- Ansari-Lari M. A., Oeltjen J. C., Schwartz S., Zhang Z., Muzny D. M., Lu J., Gorrell J. H., Chinault A. C., Belmont J. W., Miller W., and Gibbs R. A. (1998). Comparative sequence analysis of a gene-rich cluster at human chromosome 12p13 and its syntenic region in mouse chromosome 6. *Genome Res* **8**: 29-40.

- Antequera F., and Bird A. (1993). Number of CpG islands and genes in human and mouse. *Proc Natl Acad Sci U S A* **90**: 11995-9.
- Aparicio S., and Brenner S. (1997). How good a model is the Fugu genome? *Nature* **387**: 140.
- Aparicio S. A. J. R. (2000). How to count...human genes. *Nat. Genet.* **25**: 129-130.
- Apweiler R., Attwood T. K., Bairoch A., Bateman A., Birney E., Biswas M., Bucher P., Cerutti L., Corpet F., Croning M. D., Durbin R., Falquet L., Fleischmann W., Gouzy J., Hermjakob H., Hulo N., Jonassen I., Kahn D., Kanapin A., Karavidopoulou Y., Lopez R., Marx B., Mulder N. J., Oinn T. M., Pagni M., Servant F., Sigrist C. J., and Zdobnov E. M. (2000). InterPro-an integrated documentation resource for protein families, domains and functional sites. *Bioinformatics* **16**: 1145-50.
- Ardlie K. G., Kruglyak L., and Seielstad M. (2002). Patterns of linkage disequilibrium in the human genome. *Nat Rev Genet* **3**: 299-309.
- Arnold C., and Hodgson I. J. (1991). Vectorette PCR: a novel approach to genomic walking. *PCR Methods Appl* **1**: 39-42.
- Bairoch A., and Apweiler R. (2000). The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res* **28**: 45-8.
- Baldi P., and Baisnee P. F. (2000). Sequence analysis by additive scales: DNA structure for sequences and repeats of all lengths. *Bioinformatics* **16**: 865-89.
- Bankier A. T., Weston K. M., and Barrell B. G. (1987). Random cloning and sequencing by the M13/dideoxynucleotide chain termination method. *Methods Enzymol* **155**: 51-93.
- Barbesino G., Tomer Y., Concepcion E., Davies T. F., and Greenberg D. A. (1998). Linkage analysis of candidate genes in autoimmune thyroid disease: 1. Selected immunoregulatory genes. International Consortium for the Genetics of Autoimmune Thyroid Disease. *J Clin Endocrinol Metab* **83**: 1580-4.
- Barbesino G., Tomer Y., Concepcion E. S., Davies T. F., and Greenberg D. A. (1998). Linkage analysis of candidate genes in autoimmune thyroid disease. II. Selected gender-related genes and the X-chromosome. International Consortium for the Genetics of Autoimmune Thyroid Disease. *J Clin Endocrinol Metab* **83**: 3290-5.
- Bargmann C. I. (2001). High-throughput reverse genetics: RNAi screens in *Caenorhabditis elegans*. *Genome Biol* **2**: REVIEWS1005.
- Barnes M. R. (2002). SNP and mutation data on the Web - hidden treasures for uncovering. *Comparative and Functional Genomics* **3**: 67-74.
- Bassett D. E., Jr., Basrai M. A., Connelly C., Hyland K. M., Kitagawa K., Mayer M. L., Morrow D. M., Page A. M., Resto V. A., Skibbens R. V., and Hieter P. (1996). Exploiting the complete yeast genome sequence. *Curr Opin Genet Dev* **6**: 763-6.
- Batzler M. A., Deininger P. L., Hellmann-Blumberg U., Jurka J., Labuda D., Rubin C. M., Schmid C. W., Zietkiewicz E., and Zuckerkandl E. (1996). Standardized nomenclature for Alu repeats. *J Mol Evol* **42**: 3-6.

- Batzoglou S., Pachter L., Mesirov J. P., Berger B., and Lander E. S. (2000). Human and mouse gene structure: comparative analysis and application to exon prediction. *Genome Res* **10**: 950-8.
- Baudin A., Ozier-Kalogeropoulos O., Denouel A., Lacroute F., and Cullin C. (1993). A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **21**: 3329-30.
- Beaudoing E., Freier S., Wyatt J. R., Claverie J. M., and Gautheret D. (2000). Patterns of variant polyadenylation signal usage in human genes. *Genome Res* **10**: 1001-10.
- Bench A. J., Aldred M. A., Humphray S. J., Champion K. M., Gilbert J. G., Asimakopoulos F. A., Deloukas P., Gwilliam R., Bentley D. R., and Green A. R. (1998). A detailed physical and transcriptional map of the region of chromosome 20 that is deleted in myeloproliferative disorders and refinement of the common deleted region. *Genomics* **49**: 351-62.
- Bench A. J., Nacheva E. P., Hood T. L., Holden J. L., French L., Swanton S., Champion K. M., Li J., Whittaker P., Stavrides G., Hunt A. R., Huntly B. J., Campbell L. J., Bentley D. R., Deloukas P., and Green A. R. (2000). Chromosome 20 deletions in myeloid malignancies: reduction of the common deleted region, generation of a PAC/BAC contig and identification of candidate genes. UK Cancer Cytogenetics Group (UKCCG). *Oncogene* **19**: 3902-13.
- Benson D. A., Karsch-Mizrachi I., Lipman D. J., Ostell J., Rapp B. A., and Wheeler D. L. (2002). GenBank. *Nucleic Acids Res* **30**: 17-20.
- Bentley D. R. (2000). The Human Genome Project-an overview. *Med Res Rev* **20**: 189-96.
- Bentley D. R., Deloukas P., Dunham A., French L., Gregory S. G., Humphray S. J., Mungall A. J., Ross M. T., Carter N. P., Dunham I., Scott C. E., Ashcroft K. J., Atkinson A. L., Aubin K., Beare D. M., Bethel G., Brady N., Brook J. C., Burford D. C., Burrill W. D., Burrows C., Butler A. P., Carder C., Catanese J. J., Clee C. M., Clegg S. M., Copley V., Coffey A. J., Cole C. G., Collins J. E., Conquer J. S., Cooper R. A., Culley K. M., Dawson E., Dearden F. L., Durbin R. M., de Jong P. J., Dharmi P. D., Earthrowl M. E., Edwards C. A., Evans R. S., Gillson C. J., Ghori J., Green L., Gwilliam R., Halls K. S., Hammond S., Harper G. L., Heathcott R. W., Holden J. L., Holloway E., Hopkins B. L., Howard P. J., Howell G. R., Huckle E. J., Hughes J., Hunt P. J., Hunt S. E., Izmajlowicz M., Jones C. A., Joseph S. S., Laird G., Langford C. F., Lehvaslaiho M. H., Leversha M. A., McCann O. T., McDonald L. M., McDowall J., Maslen G. L., Mistry D., Moschonas N. K., Neocleous V., Pearson D. M., Phillips K. J., Porter K. M., Prathalingam S. R., Ramsey Y. H., Ranby S. A., Rice C. M., Rogers J., Rogers L. J., Sarafidou T., Scott D. J., Sharp G. J., Shaw-Smith C. J., Smink L. J., Soderlund C., Sotheran E. C., Steingruber H. E., Sulston J. E., Taylor A., Taylor R. G., Thorpe A. A., Tinsley E., Warry G. L., Whittaker A., Whittaker P., Williams S. H., Wilmer T. E., Wooster R., *et al.* (2001). The physical maps for sequencing human chromosomes 1, 6, 9, 10, 13, 20 and X. *Nature* **409**: 942-3.
- Bernardi G. (1989). The isochore organization of the human genome. *Annu Rev Genet* **23**: 637-61.

- Bernardi G., Olofsson B., Filipski J., Zerial M., Salinas J., Cuny G., Meunier-Rotival M., and Rodier F. (1985). The mosaic genome of warm-blooded vertebrates. *Science* **228**: 953-8.
- Bernot A., Heilig R., Clepet C., Smaoui N., Da Silva C., Petit J. L., Devaud C., Chiannilkulchai N., Fizames C., Samson D., Cruaud C., Caloustian C., Gyapay G., Delpech M., and Weissenbach J. (1998). A transcriptional Map of the FMF region. *Genomics* **50**: 147-60.
- Bertina R. M., Koeleman B. P., Koster T., Rosendaal F. R., Dirven R. J., de Ronde H., van der Velden P. A., and Reitsma P. H. (1994). Mutation in blood coagulation factor V associated with resistance to activated protein C. *Nature* **369**: 64-7.
- Bickmore W. A., and Sumner A. T. (1989). Mammalian chromosome banding-an expression of genome organization. *Trends Genet* **5**: 144-8.
- Bihoreau M. T., Gauguier D., Kato N., Hyne G., Lindpaintner K., Rapp J. P., James M. R., and Lathrop G. M. (1997). A linkage map of the rat genome derived from three F2 crosses. *Genome Res* **7**: 434-40.
- Bihoreau M. T., Sebag-Montefiore L., Godfrey R. F., Wallis R. H., Brown J. H., Danoy P. A., Collins S. C., Rouard M., Kaisaki P. J., Lathrop M., and Gauguier D. (2001). A high-resolution consensus linkage map of the rat, integrating radiation hybrid and genetic maps. *Genomics* **75**: 57-69.
- Birchall P. S., Fishpool R. M., and Albertson D. G. (1995). Expression patterns of predicted genes from the *C. elegans* genome sequence visualized by FISH in whole organisms. *Nat Genet* **11**: 314-20.
- Bird A. P. (1986). CpG-rich islands and the function of DNA methylation. *Nature* **321**: 209-13.
- Bird A. P. (1987). CpG islands as gene markers in the vertebrate nucleus. *TIG* **3**: 342-347.
- Blackstock W. P., and Weir M. P. (1999). Proteomics: quantitative and physical mapping of cellular proteins. *Trends Biotechnol* **17**: 121-7.
- Blattner F. R., Plunkett G., 3rd, Bloch C. A., Perna N. T., Burland V., Riley M., Collado-Vides J., Glasner J. D., Rode C. K., Mayhew G. F., Gregor J., Davis N. W., Kirkpatrick H. A., Goeden M. A., Rose D. J., Mau B., and Shao Y. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* **277**: 1453-74.
- Blechs Schmidt K., Schweiger M., Wertz K., Poulson R., Christensen H. M., Rosenthal A., Lehrach H., and Yaspo M. L. (1999). The mouse Aire gene: comparative genomic sequencing, gene organization, and expression. *Genome Res* **9**: 158-66.
- Boehnke M. (2000). A look at linkage disequilibrium. *Nat Genet* **25**: 246-7.
- Boguski M. S., Lowe T. M., and Tolstoshev C. M. (1993). dbEST-database for expressed sequence tags. *Nat Genet* **4**: 332-3.
- Botstein D., White R. L., Skolnick M., and Davis R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* **32**: 314-31.

- Bowden D. W., Sale M., Howard T. D., Qadri A., Spray B. J., Rothschild C. B., Akots G., Rich S. S., and Freedman B. I. (1997). Linkage of genetic markers on human chromosomes 20 and 12 to NIDDM in Caucasian sib pairs with a history of diabetic nephropathy. *Diabetes* **46**: 882-6.
- Braun A., Little D. P., Reuter D., Muller-Mysok B., and Koster H. (1997a). Improved analysis of microsatellites using mass spectrometry. *Genomics* **46**: 18-23.
- Braun A., Little D. P., and Koster H. (1997b). Detecting CFTR gene mutations by using primer oligo base extension and mass spectrometry. *Clin Chem* **43**: 1151-8.
- Brenner S., Elgar G., Sandford R., Macrae A., Venkatesh B., and Aparicio S. (1993). Characterization of the pufferfish (*Fugu*) genome as a compact model vertebrate genome. *Nature* **366**: 265-8.
- Brett D., Hanke J., Lehmann G., Haase S., Delbruck S., Krueger S., Reich J., and Bork P. (2000). EST comparison indicates 38% of human mRNAs contain possible alternative splice forms. *FEBS Lett* **474**: 83-6.
- Brickner A. G., Koop B. F., Aronow B. J., and Wiginton D. A. (1999). Genomic sequence comparison of the human and mouse adenosine deaminase gene regions. *Mamm Genome* **10**: 95-101.
- Brookes A. J. (1999). The essence of SNPs. *Gene* **234**: 177-86.
- Brookes A. J., Lehvaslaiho H., Siegfried M., Boehm J. G., Yuan Y. P., Sarkar C. M., Bork P., and Ortigao F. (2000). HGBASE: a database of SNPs and other variations in and around human genes. *Nucleic Acids Res* **28**: 356-60.
- Brüls T., Gyapay G., Petit J. L., Artiguenave F., Vico V., Qin S., Tin-Wollam A. M., Da Silva C., Muselet D., Mavel D., Pelletier E., Levy M., Fujiyama A., Matsuda F., Wilson R., Rowen L., Hood L., Weissenbach J., Saurin W., and Heilig R. (2001). A physical map of human chromosome 14. *Nature* **409**: 947-8.
- Buckler A. J., Chang D. D., Graw S. L., Brook J. D., Haber D. A., Sharp P. A., and Housman D. E. (1991). Exon amplification: a strategy to isolate mammalian genes based on RNA splicing. *Proc Natl Acad Sci U S A* **88**: 4005-9.
- Burge C., and Karlin S. (1997). Prediction of complete gene structures in human genomic DNA. *J Mol Biol* **268**: 78-94.
- Burke D. T., Carle G. F., and Olson M. V. (1987). Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* **236**: 806-12.
- Burns N., Grimwade B., Ross-Macdonald P. B., Choi E. Y., Finberg K., Roeder G. S., and Snyder M. (1994). Large-scale analysis of gene expression, protein localization, and gene disruption in *Saccharomyces cerevisiae*. *Genes Dev* **8**: 1087-105.
- Burset M., and Guigo R. (1996). Evaluation of gene structure prediction programs. *Genomics* **34**: 353-67.
- Burset M., Seledtsov I. A., and Solovyev V. V. (2000). Analysis of canonical and non-canonical splice sites in mammalian genomes. *Nucleic Acids Res* **28**: 4364-75.

- Cai W. W., Chow C. W., Damani S., Gregory S. G., Marra M., and Bradley A. (2001). An SSLP marker-anchored BAC framework map of the mouse genome. *Nat Genet* **29**: 133-4.
- Carlson C. S., Newman T. L., and Nickerson D. A. (2001). SNPping in the human genome. *Curr Opin Chem Biol* **5**: 78-85.
- Carver E. A., and Stubbs L. (1997). Zooming in on the human-mouse comparative map: genome conservation re-examined on a high-resolution scale. *Genome Res* **7**: 1123-37.
- Centola M., Chen X., Sood R., Deng Z., Aksentijevich I., Blake T., Ricke D. O., Wood G., Zaks N., Richards N., Krizman D., Mansfield E., Apostolou S., Liu J., Shafran N., Vedula A., Hamon M., Cercek A., Kahan T., Gumucio D., Callen D. F., Richards R. I., Moyzis R. K., Kastner D. L., and *et al.* (1998). Construction of an approximately 700-kb transcript map around the familial Mediterranean fever locus on human chromosome 16p13.3. *Genome Res* **8**: 1172-91.
- Chakravarti A. (1999). Population genetics-making sense out of sequence. *Nat Genet* **21**: 56-60.
- Chakravarti A. (2001). To a future of genetic medicine. *Nature* **409**: 822-3.
- Chambon-Pautas C., Cave H., Gerard B., Guidal-Giroux C., Duval M., Vilmer E., and Grandchamp B. (1998). High-resolution allelotyping analysis of childhood B-lineage acute lymphoblastic leukemia. *Leukemia* **12**: 1107-13.
- Chasman D., and Adams R. M. (2001). Predicting the functional consequences of non-synonymous single nucleotide polymorphisms: structure-based assessment of amino acid variation. *J Mol Biol* **307**: 683-706.
- Chen J., Greenblatt I. M., and Dellaporta S. L. (1992). Molecular analysis of Ac transposition and DNA replication. *Genetics* **130**: 665-76.
- Chiaromonte F., Yang S., Elnitski L., Yap V. B., Miller W., and Hardison R. C. (2001). Association between divergence and interspersed repeats in mammalian noncoding genomic DNA. *Proc Natl Acad Sci USA* **98**: 14503-8.
- Chiaromonte F., Yap V. B., and Miller W. (2002). Scoring pairwise genomic sequence alignments. *Pac Symp Biocomput*: 115-26.
- Chumakov I. M., Le Gall I., Billault A., Ougen P., Soularue P., Guillou S., Rigault P., Bui H., De Tand M. F., Barillot E., and *et al.* (1992). Isolation of chromosome 21-specific yeast artificial chromosomes from a total human genome library. *Nat Genet* **1**: 222-5.
- Chumakov I. M., Rigault P., Le Gall I., Bellanne-Chantelot C., Billault A., Guillou S., Soularue P., Guasconi G., Poullier E., Gros I., and *et al.* (1995). A YAC contig map of the human genome. *Nature* **377**: 175-297.
- Church D. M., Banks L. T., Rogers A. C., Graw S. L., Housman D. E., Gusella J. F., and Buckler A. J. (1993). Identification of human chromosome 9 specific genes using exon amplification. *Hum Mol Genet* **2**: 1915-20.
- Claverie J. M. (1997). Computational methods for the identification of genes in vertebrate genomic sequences. *Hum Mol Genet* **6**: 1735-44.

- Claverie J. M. (2001). Gene number. What if there are only 30,000 human genes? *Science* **291**: 1255-7.
- Colgan D. F., and Manley J. L. (1997). Mechanism and regulation of mRNA polyadenylation. *Genes Dev* **11**: 2755-66.
- Collinge J., Sidle K. C., Meads J., Ironside J., and Hill A. F. (1996). Molecular analysis of prion strain variation and the aetiology of 'new variant' CJD. *Nature* **383**: 685-90.
- Collins F., and Galas D. (1993). A new five-year plan for the U.S. Human Genome Project. *Science* **262**: 43-6.
- Collins F. S. (1995). Positional cloning moves from perditional to traditional. *Nat Genet* **9**: 347-50.
- Collins F. S. (2001). Contemplating the end of the beginning. *Genome Res* **11**: 641-3.
- Collins F. S., Brooks L. D., and Chakravarti A. (1998b). A DNA polymorphism discovery resource for research on human genetic variation. *Genome Res* **8**: 1229-31.
- Collins F. S., and Guttmacher A. E. (2001). Genetics moves into the medical mainstream. *Jama* **286**: 2322-4.
- Collins F. S., Guyer M. S., and Charkravarti A. (1997). Variations on a theme: cataloging human DNA sequence variation. *Science* **278**: 1580-1.
- Collins F. S., Patrinos A., Jordan E., Chakravarti A., Gesteland R., and Walters L. (1998a). New goals for the U.S. Human Genome Project: 1998-2003. *Science* **282**: 682-9.
- Collins F. S., and Weissman S. M. (1984). The molecular genetics of human hemoglobin. *Prog Nucleic Acid Res Mol Biol* **31**: 315-462.
- Collins J., and Hohn B. (1978). Cosmids: a type of plasmid gene-cloning vector that is packageable in vitro in bacteriophage lambda heads. *Proc Natl Acad Sci U S A* **75**: 4242-6.
- Collins J., Saari B., and Anderson P. (1987). Activation of a transposable element in the germ line but not the soma of *Caenorhabditis elegans*. *Nature* **328**: 726-8.
- Coulondre C., Miller J. H., Farabaugh P. J., and Gilbert W. (1978). Molecular basis of base substitution hotspots in *Escherichia coli*. *Nature* **274**: 775-80.
- Coulson A., Sulston J., Brenner S., and Karn J. (1986). Toward a physical map of the genome of the nematode *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci.* **83**: 7821-7825.
- Couque N., Chambon-Pautas C., Cave H., Bardet V., Duval M., Vilmer E., and Grandchamp B. (1999). Mapping of chromosome 20 for loss of heterozygosity in childhood ALL reveals a 1,000-kb deletion in one patient. *Leukemia* **13**: 1972-4.
- Cox D. R., Burmeister M., Price E. R., Kim S., and Myers R. M. (1990). Radiation hybrid mapping: a somatic cell genetic method for constructing high-resolution maps of mammalian chromosomes. *Science* **250**: 245-50.

- Crabtree J., Wiltshire T., Brunk B., Zhao S., Schug J., Stoeckert C. J., Jr., and Bucan M. (2001). High-resolution BAC-based map of the central portion of mouse chromosome 5. *Genome Res* **11**: 1746-57.
- Craig J. M., and Bickmore W. A. (1993). Chromosome bands--flavours to savour. *Bioessays* **15**: 349-54.
- Crnogorac-Jurcevic T., Brown J. R., Lehrach H., and Schalkwyk L. C. (1997). *Tetraodon fluviatilis*, a new puffer fish model for genome studies. *Genomics* **41**: 177-84.
- Crollius H. R., Jaillon O., Bernot A., Dasilva C., Bouneau L., Fischer C., Fizames C., Wincker P., Brottier P., Quetier F., Saurin W., and Weissenbach J. (2000a). Estimate of human gene number provided by genome-wide analysis using *Tetraodon nigroviridis* DNA sequence. *Nat Genet* **25**: 235-8.
- Crollius H. R., Jaillon O., Dasilva C., Ozouf-Costaz C., Fizames C., Fischer C., Bouneau L., Billault A., Quetier F., Saurin W., Bernot A., and Weissenbach J. (2000b). Characterization and repeat analysis of the compact genome of the freshwater pufferfish *Tetraodon nigroviridis*. *Genome Res* **10**: 939-49.
- Cross S. H., and Bird A. P. (1995). CpG islands and genes. *Curr Opin Genet Dev* **5**: 309-14.
- Cross S. H., Charlton J. A., Nan X., and Bird A. P. (1994). Purification of CpG islands using a methylated DNA binding column. *Nat Genet* **6**: 236-44.
- Cross S. H., Clark V. H., and Bird A. P. (1999). Isolation of CpG islands from large genomic clones. *Nucleic Acids Res* **27**: 2099-107.
- Cross S. H., Clark V. H., Simmen M. W., Bickmore W. A., Maroon H., Langford C. F., Carter N. P., and Bird A. P. (2000). CpG island libraries from human chromosomes 18 and 22: landmarks for novel genes. *Mamm Genome* **11**: 373-83.
- Davignon J., Gregg R. E., and Sing C. F. (1988). Apolipoprotein E polymorphism and atherosclerosis. *Arteriosclerosis* **8**: 1-21.
- Davuluri R. V., Grosse I., and Zhang M. Q. (2001). Computational identification of promoters and first exons in the human genome. *Nat Genet* **29**: 412-7.
- Dawson E., Abecasis G. R., Bumpstead S., Chen Y., Hunt S., Beare D. M., Pabial J., Dibling T., Tinsley E., Kirby S., Carter D., Papaspyridonos M., Livingstone S., Ganske R., Lohmussaer E., Zernant J., Tonisson N., Remm M., Magi R., Puurand T., Vilo J., Kurg A., Rice K., Deloukas P., Mott R., Metspalu A., Bentley D. R., Cardon L. R., and Dunham I. (2002). A first-generation linkage disequilibrium map of human chromosome 22. *Nature* **416**: 544-8.
- Dawson E., Chen Y., Hunt S., Smink L. J., Hunt A., Rice K., Livingstone S., Bumpstead S., Bruskiwich R., Sham P., Ganske R., Adams M., Kawasaki K., Shimizu N., Minoshima S., Roe B., Bentley D., and Dunham I. (2001). A SNP resource for human chromosome 22: extracting dense clusters of SNPs from the genomic sequence. *Genome Res* **11**: 170-8.
- Deak P., Omar M. M., Saunders R. D., Pal M., Komonyi O., Szidonya J., Maroy P., Zhang Y., Ashburner M., Benos P., Savakis C., Siden-Kiamos I., Louis C., Bolshakov V. N., Kafatos F. C., Madueno E., Modolell J., and Glover D. M.

- (1997). P-element insertion alleles of essential genes on the third chromosome of *Drosophila melanogaster*: correlation of physical and cytogenetic maps in chromosomal region 86E-87F. *Genetics* **147**: 1697-722.
- DeBerardinis R. J., Goodier J. L., Ostertag E. M., and Kazazian H. H., Jr. (1998). Rapid amplification of a retrotransposon subfamily is evolving the mouse genome. *Nat Genet* **20**: 288-90.
- DeBry R. W., and Seldin M. F. (1996). Human/mouse homology relationships. *Genomics* **33**: 337-51.
- Dehal P., Predki P., Olsen A. S., Kobayashi A., Folta P., Lucas S., Land M., Terry A., Ecale Zhou C. L., Rash S., Zhang Q., Gordon L., Kim J., Elkin C., Pollard M. J., Richardson P., Rokhsar D., Uberbacher E., Hawkins T., Branscomb E., and Stubbs L. (2001). Human chromosome 19 and related regions in mouse: conservative and lineage-specific evolution. *Science* **293**: 104-11.
- Deloukas P., Matthews L. H., Ashurst J., Burton J., Gilbert J. G., Jones M., Stavrides G., Almeida J. P., Babbage A. K., Bagguley C. L., Bailey J., Barlow K. F., Bates K. N., Beard L. M., Beare D. M., Beasley O. P., Bird C. P., Blakey S. E., Bridgeman A. M., Brown A. J., Buck D., Burrill W., Butler A. P., Carder C., Carter N. P., Chapman J. C., Clamp M., Clark G., Clark L. N., Clark S. Y., Clee C. M., Clegg S., Copley V. E., Collier R. E., Connor R., Corby N. R., Coulson A., Coville G. J., Deadman R., Dhami P., Dunn M., Ellington A. G., Frankland J. A., Fraser A., French L., Garner P., Grafham D. V., Griffiths C., Griffiths M. N., Gwilliam R., Hall R. E., Hammond S., Harley J. L., Heath P. D., Ho S., Holden J. L., Howden P. J., Huckle E., Hunt A. R., Hunt S. E., Jekosch K., Johnson C. M., Johnson D., Kay M. P., Kimberley A. M., King A., Knights A., Laird G. K., Lawlor S., Lehtvaslaiho M. H., Leversha M., Lloyd C., Lloyd D. M., Lovell J. D., Marsh V. L., Martin S. L., McConnell L. J., McLay K., McMurray A. A., Milne S., Mistry D., Moore M. J., Mullikin J. C., Nickerson T., Oliver K., Parker A., Patel R., Pearce T. A., Peck A. I., Phillimore B. J., Prathalingam S. R., Plumb R. W., Ramsay H., Rice C. M., Ross M. T., Scott C. E., Sehra H. K., Shownkeen R., Sims S., Skuce C. D., *et al.* (2001). The DNA sequence and comparative analysis of human chromosome 20. *Nature* **414**: 865-71.
- Deloukas P., Schuler G. D., Gyapay G., Beasley E. M., Soderlund C., Rodriguez-Tome P., Hui L., Matise T. C., McKusick K. B., Beckmann J. S., Bentolila S., Bihoreau M., Birren B. B., Browne J., Butler A., Castle A. B., Chiannilkulchai N., Clee C., Day P. J., Dehejia A., Dibling T., Drouot N., Duprat S., Fizames C., Bentley D. R., and *et al.* (1998). A physical map of 30,000 human genes. *Science* **282**: 744-6.
- DeRisi J. L., Iyer V. R., and Brown P. O. (1997). Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278**: 680-6.
- Deutsch S., Iseli C., Bucher P., Antonarakis S. E., and Scott H. S. (2001). A cSNP map and database for human chromosome 21. *Genome Res* **11**: 300-7.
- Devaux F., Marc P., and Jacq C. (2001). Transcriptomes, transcription activators and microarrays. *FEBS Lett* **498**: 140-4.
- Devlin B., and Risch N. (1995). A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics* **29**: 311-22.

- Dib C., Faure S., Fizames C., Samson D., Drouot N., Vignal A., Millasseau P., Marc S., Hazan J., Seboun E., Lathrop M., Gyapay G., Morissette J., and Weissenbach J. (1996). A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* **380**: 152-4.
- Dietrich W. F., Copeland N. G., Gilbert D. J., Miller J. C., Jenkins N. A., and Lander E. S. (1995). Mapping the mouse genome: current status and future prospects. *Proc Natl Acad Sci U S A* **92**: 10849-53.
- Dietrich W. F., Miller J., Steen R., Merchant M. A., Damron-Boles D., Husain Z., Dredge R., Daly M. J., Ingalls K. A., O'Connor T. J., and *et al.* (1996). A comprehensive genetic map of the mouse genome. *Nature* **380**: 149-52.
- Dillon N., and Sabbattini P. (2000). Functional gene expression domains: defining the functional unit of eukaryotic gene regulation. *Bioessays* **22**: 657-65.
- Donis-Keller H., Green P., Helms C., Cartinhour S., Weiffenbach B., Stephens K., Keith T. P., Bowden D. W., Smith D. R., Lander E. S., and *et al.* (1987). A genetic linkage map of the human genome. *Cell* **51**: 319-37.
- Down T. A., and Hubbard T. J. (2002). Computational detection and location of transcription start sites in mammalian genomic DNA. *Genome Res* **12**: 458-61.
- Driever W., Solnica-Krezel L., Schier A. F., Neuhauss S. C., Malicki J., Stemple D. L., Stainier D. Y., Zwartkruis F., Abdelilah S., Rangini Z., Belak J., and Boggs C. (1996). A genetic screen for mutations affecting embryogenesis in zebrafish. *Development* **123**: 37-46.
- Driever W., Stemple D., Schier A., and Solnica-Krezel L. (1994). Zebrafish: genetic tools for studying vertebrate development. *Trends Genet* **10**: 152-9.
- Dunham I. (2000). The gene guessing game. *Yeast* **17**: 218-24.
- Dunham I., Shimizu N., Roe B. A., Chisoe S., Hunt A. R., Collins J. E., Bruskiewich R., Beare D. M., Clamp M., Smink L. J., Ainscough R., Almeida J. P., Babbage A., Bagguley C., Bailey J., Barlow K., Bates K. N., Beasley O., Bird C. P., Blakey S., Bridgeman A. M., Buck D., Burgess J., Burrill W. D., O'Brien K. P., and *et al.* (1999). The DNA sequence of human chromosome 22. *Nature* **402**: 489-95.
- Dunning A. M., Durocher F., Healey C. S., Teare M. D., McBride S. E., Carlomagno F., Xu C. F., Dawson E., Rhodes S., Ueda S., Lai E., Luben R. N., Van Rensburg E. J., Mannermaa A., Kataja V., Rennart G., Dunham I., Purvis I., Easton D., and Ponder B. A. (2000). The extent of linkage disequilibrium in four populations with distinct demographic histories. *Am J Hum Genet* **67**: 1544-54.
- Durbin R., and Thierry-Mieg J. (1994). The ACEDB genome database. *Proceedings of the International Symposium on Computational Methods in Genome Research (1992; Heidelberg, Germany)*.
- Dutt M. J., and Lee K. H. (2000). Proteomic analysis. *Curr Opin Biotechnol* **11**: 176-9.
- Duyk G. M., Kim S. W., Myers R. M., and Cox D. R. (1990). Exon trapping: a genetic screen to identify candidate transcribed sequences in cloned mammalian genomic DNA. *Proc Natl Acad Sci U S A* **87**: 8995-9.

- Elgar G., Clark M., Green A., and Sandford R. (1997). How good a model is the Fugu genome? *Nature* **387**: 140.
- Elgar G., Sandford R., Aparicio S., Macrae A., Venkatesh B., and Brenner S. (1996). Small is beautiful: comparative genomics with the pufferfish (Fugu rubripes). *Trends Genet* **12**: 145-50.
- Elvin P., Slynn G., Black D., Graham A., Butler R., Riley J., Anand R., and Markham A. F. (1990). Isolation of cDNA clones using yeast artificial chromosome probes. *Nucleic Acids Res* **18**: 3913-7.
- Engels W. R., Johnson-Schlitz D. M., Eggleston W. B., and Sved J. (1990). High-frequency P element loss in Drosophila is homolog dependent. *Cell* **62**: 515-25.
- Eppig J. T. (1996). Comparative maps: adding pieces to the mammalian jigsaw puzzle. *Curr Opin Genet Dev* **6**: 723-30.
- Eppig J. T., and Nadeau J. H. (1995). Comparative maps: the mammalian jigsaw puzzle. *Curr Opin Genet Dev* **5**: 709-16.
- Esnault C., Maestre J., and Heidmann T. (2000). Human LINE retrotransposons generate processed pseudogenes. *Nat Genet* **24**: 363-7.
- Ewing B., and Green P. (2000). Analysis of expressed sequence tags indicates 35,000 human genes. *Nat Genet* **25**: 232-4.
- Fei Z., Ono T., and Smith L. M. (1998). MALDI-TOF mass spectrometric typing of single nucleotide polymorphisms with mass-tagged ddNTPs. *Nucleic Acids Res* **26**: 2827-8.
- Fenn J. B., Mann M., Meng C. K., Wong S. F., and Whitehouse C. M. (1989). Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**: 64-71.
- Fickett J. W., and Hatzigeorgiou A. G. (1997). Eukaryotic promoter recognition. *Genome Res* **7**: 861-78.
- Fields C., Adams M. D., White O., and Venter J. C. (1994). How many genes in the human genome? *Nat Genet* **7**: 345-6.
- Fire A., Xu S., Montgomery M. K., Kostas S. A., Driver S. E., and Mello C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**: 806-11.
- Fleischmann R. D., Adams M. D., White O., Clayton R. A., Kirkness E. F., Kerlavage A. R., Bult C. J., Tomb J. F., Dougherty B. A., Merrick J. M., and *et al.* (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**: 496-512.
- Foote S., Vollrath D., Hilton A., and Page D. C. (1992). The human Y chromosome: overlapping DNA clones spanning the euchromatic region. *Science* **258**: 60-6.
- Footz T. K., Brinkman-Mills P., Banting G. S., Maier S. A., Riazi M. A., Bridgland L., Hu S., Birren B., Minoshima S., Shimizu N., Pan H., Nguyen T., Fang F., Fu Y., Ray L., Wu H., Shaull S., Phan S., Yao Z., Chen F., Huan A., Hu P., Wang Q., Loh P., Qi S., Roe B. A., and McDermid H. E. (2001). Analysis of the cat eye syndrome critical region in humans and the region of conserved

- synteny in mice: a search for candidate genes at or near the human chromosome 22 pericentromere. *Genome Res* **11**: 1053-70.
- Fraser P., and Grosveld F. (1998). Locus control regions, chromatin activation and transcription. *Curr Opin Cell Biol* **10**: 361-5.
- Frisse L., Hudson R. R., Bartoszewicz A., Wall J. D., Donfack J., and Di Rienzo A. (2001). Gene conversion and different population histories may explain the contrast between polymorphism and linkage disequilibrium levels. *Am J Hum Genet* **69**: 831-43.
- Fromont-Racine M., Rain J. C., and Legrain P. (1997). Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. *Nat Genet* **16**: 277-82.
- Fu D. J., Tang K., Braun A., Reuter D., Darnhofer-Demar B., Little D. P., O'Donnell M. J., Cantor C. R., and Koster H. (1998). Sequencing exons 5 to 8 of the p53 gene by MALDI-TOF mass spectrometry. *Nat Biotechnol* **16**: 381-4.
- Gabriel S. B., Schaffner S. F., Nguyen H., Moore J. M., Roy J., Blumenstiel B., Higgins J., DeFelice M., Lochner A., Faggart M., Liu-Cordero S. N., Rotimi C., Adeyemo A., Cooper R., Ward R., Lander E. S., Daly M. J., and Altshuler D. (2002). The structure of haplotype blocks in the human genome. *Science* **296**: 2225-9.
- Gardiner K. (1995). Human genome organization. *Curr Opin Genet Dev* **5**: 315-22.
- Gardiner K., and Mural R. J. (1995). Getting the message: identifying transcribed sequences. *Trends Genet* **11**: 77-9.
- Gates M. A., Kim L., Egan E. S., Cardozo T., Sirotkin H. I., Dougan S. T., Lashkari D., Abagyan R., Schier A. F., and Talbot W. S. (1999). A genetic linkage map for zebrafish: comparative analysis and localization of genes and expressed sequences. *Genome Res* **9**: 334-47.
- Gautheret D., Poirot O., Lopez F., Audic S., and Claverie J. M. (1998). Alternate polyadenylation in human mRNAs: a large-scale analysis by EST clustering. *Genome Res* **8**: 524-30.
- Geisler R., Rauch G. J., Baier H., van Bebber F., Brobeta L., Dekens M. P., Finger K., Fricke C., Gates M. A., Geiger H., Geiger-Rudolph S., Gilmour D., Glaser S., Gnugge L., Habeck H., Hingst K., Holley S., Keenan J., Kirn A., Knaut H., Lashkari D., Maderspacher F., Martyn U., Neuhauss S., Haffter P., and *et al.* (1999). A radiation hybrid map of the zebrafish genome. *Nat Genet* **23**: 86-9.
- Ghosh S., Watanabe R. M., Hauser E. R., Valle T., Magnuson V. L., Erdos M. R., Langefeld C. D., Balow J., Jr., Ally D. S., Kohtamaki K., Chines P., Birznieks G., Kaleta H. S., Musick A., Te C., Tannenbaum J., Eldridge W., Shapiro S., Martin C., Witt A., So A., Chang J., Shurtleff B., Porter R., Boehnke M., and *et al.* (1999). Type 2 diabetes: evidence for linkage on chromosome 20 in 716 Finnish affected sib pairs. *Proc Natl Acad Sci U S A* **96**: 2198-203.
- Ghosh S., Watanabe R. M., Valle T. T., Hauser E. R., Magnuson V. L., Langefeld C. D., Ally D. S., Mohlke K. L., Silander K., Kohtamaki K., Chines P., Balow Jr J., Birznieks G., Chang J., Eldridge W., Erdos M. R., Karanjawala Z. E., Knapp J. I., Kudelko K., Martin C., Morales-Mena A., Musick A., Musick T., Pfahl C., Porter R., and Rayman J. B. (2000). The Finland-United States

- investigation of non-insulin-dependent diabetes mellitus genetics (FUSION) study. I. An autosomal genome scan for genes that predispose to type 2 diabetes. *Am J Hum Genet* **67**: 1174-85.
- Gilbert W. (1978). Why genes in pieces? *Nature* **271**: 501.
- Gilley J., Armes N., and Fried M. (1997). Fugu genome is not a good mammalian model. *Nature* **385**: 305-6.
- Gilley J., and Fried M. (1999). Extensive gene order differences within regions of conserved synteny between the Fugu and human genomes: implications for chromosomal evolution and the cloning of disease genes. *Hum Mol Genet* **8**: 1313-20.
- Goffeau A., Barrell B. G., Bussey H., Davis R. W., Dujon B., Feldmann H., Galibert F., Hoheisel J. D., Jacq C., Johnston M., Louis E. J., Mewes H. W., Murakami Y., Philippsen P., Tettelin H., and Oliver S. G. (1996). Life with 6000 genes. *Science* **274**: 546, 563-7.
- Goldstein D. B. (2001). Islands of linkage disequilibrium. *Nat Genet* **29**: 109-11.
- Goss S. J., and Harris H. (1975). New method for mapping genes in human chromosomes. *Nature* **255**: 680-4.
- Göttgens B., Barton L. M., Gilbert J. G., Bench A. J., Sanchez M. J., Bahn S., Mistry S., Grafham D., McMurray A., Vaudin M., Amaya E., Bentley D. R., Green A. R., and Sinclair A. M. (2000). Analysis of vertebrate SCL loci identifies conserved enhancers. *Nat Biotechnol* **18**: 181-6.
- Göttgens B., Gilbert J. G., Barton L. M., Grafham D., Rogers J., Bentley D. R., and Green A. R. (2001). Long-range comparison of human and mouse SCL loci: localized regions of sensitivity to restriction endonucleases correspond precisely with peaks of conserved noncoding sequences. *Genome Res* **11**: 87-97.
- Gough S. C. (2000). The genetics of Graves' disease. *Endocrinol Metab Clin North Am* **29**: 255-66.
- Gray I. C., Campbell D. A., and Spurr N. K. (2000). Single nucleotide polymorphisms as tools in human genetics. *Hum Mol Genet* **9**: 2403-8.
- Green E. D., and Olson M. V. (1990). Chromosomal region of the cystic fibrosis gene in yeast artificial chromosomes: a model for human genome mapping. *Science* **250**: 94-8.
- Green E. D., and Olson M. V. (1990). Systematic screening of yeast artificial-chromosome libraries by use of the polymerase chain reaction. *Proc Natl Acad Sci U S A* **87**: 1213-7.
- Green E. D., Riethman H. C., Dutchik J. E., and Olson M. V. (1991). Detection and characterization of chimeric yeast artificial-chromosome clones. *Genomics* **11**: 658-69.
- Green P. (2002). Whole-genome disassembly. *Proc Natl Acad Sci U S A* **99**: 4143-4.
- Gregory S. G., Howell G. R., and Bentley D. R. (1997). Genome mapping by fluorescent fingerprinting. *Genome Res* **7**: 1162-8.

- Griffin T. J., Hall J. G., Prudent J. R., and Smith L. M. (1999). Direct genetic analysis by matrix-assisted laser desorption/ionization mass spectrometry. *Proc Natl Acad Sci U S A* **96**: 6301-6.
- Griffin T. J., and Smith L. M. (2000). Single-nucleotide polymorphism analysis by MALDI-TOF mass spectrometry. *Trends Biotechnol* **18**: 77-84.
- Griffin T. J., and Smith L. M. (2002). Single-nucleotide polymorphism analysis by MALDI-TOF mass spectrometry. *A TRENDS Guide to Genetic Variation and Genomic medicine*: S10-18.
- Griffin T. J., Tang W., and Smith L. M. (1997). Genetic analysis by peptide nucleic acid affinity MALDI-TOF mass spectrometry. *Nat Biotechnol* **15**: 1368-72.
- Griffiths W. J., Jonsson A. P., Liu S., Rai D. K., and Wang Y. (2001). Electrospray and tandem mass spectrometry in biochemistry. *Biochem J* **355**: 545-61.
- Gu Z., Hillier L., and Kwok P. Y. (1998). Single nucleotide polymorphism hunting in cyberspace. *Hum Mutat* **12**: 221-5.
- Gu Z., Wang H., Nekrutenko A., and Li W. H. (2000). Densities, length proportions, and other distributional features of repetitive sequences in the human genome estimated from 430 megabases of genomic sequence. *Gene* **259**: 81-8.
- Guigo R., Agarwal P., Abril J. F., Burset M., and Fickett J. W. (2000). An assessment of gene prediction accuracy in large DNA sequences. *Genome Res* **10**: 1631-42.
- Gut I. G. (2001). Automation in genotyping of single nucleotide polymorphisms. *Hum Mutat* **17**: 475-92.
- Gyapay G., Morissette J., Vignal A., Dib C., Fizames C., Millasseau P., Marc S., Bernardi G., Lathrop M., and Weissenbach J. (1994). The 1993-94 Genethon human genetic linkage map. *Nat Genet* **7**: 246-339.
- Gyapay G., Schmitt K., Fizames C., Jones H., Vega-Czarny N., Spillett D., Muselet D., Prud'Homme J. F., Dib C., Auffray C., Morissette J., Weissenbach J., and Goodfellow P. N. (1996). A radiation hybrid map of the human genome. *Hum Mol Genet* **5**: 339-46.
- Hacia J. G., and Collins F. S. (1999). Mutational analysis using oligonucleotide microarrays. *J Med Genet* **36**: 730-6.
- Hacia J. G., Sun B., Hunt N., Edgemon K., Mosbrook D., Robbins C., Fodor S. P., Tagle D. A., and Collins F. S. (1998). Strategies for mutational analysis of the large multiexon ATM gene using high-density oligonucleotide arrays. *Genome Res* **8**: 1245-58.
- Haff L. A., and Smirnov I. P. (1997). Single-nucleotide polymorphism identification assays using a thermostable DNA polymerase and delayed extraction MALDI-TOF mass spectrometry. *Genome Res* **7**: 378-88.
- Haffter P., Granato M., Brand M., Mullins M. C., Hammerschmidt M., Kane D. A., Odenthal J., van Eeden F. J., Jiang Y. J., Heisenberg C. P., Kelsh R. N., Furutani-Seiki M., Vogelsang E., Beuchle D., Schach U., Fabian C., and Nusslein-Volhard C. (1996). The identification of genes with unique and essential functions in the development of the zebrafish, *Danio rerio*. *Development* **123**: 1-36.

- Hansen R. S., Wijmenga C., Luo P., Stanek A. M., Canfield T. K., Weemaes C. M., and Gartner S. M. (1999). The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. *Proc Natl Acad Sci U S A* **96**: 14412-7.
- Hardison R. C., Oeltjen J., and Miller W. (1997). Long human-mouse sequence alignments reveal novel regulatory elements: a reason to sequence the mouse genome. *Genome Res* **7**: 959-66.
- Harrington C. A., Rosenow C., and Retief J. (2000). Monitoring gene expression using DNA microarrays. *Curr Opin Microbiol* **3**: 285-91.
- Harvey D. J. (2001). Identification of protein-bound carbohydrates by mass spectrometry. *Proteomics* **1**: 311-28.
- Hattori M., Fujiyama A., Taylor T. D., Watanabe H., Yada T., Park H. S., Toyoda A., Ishii K., Totoki Y., Choi D. K., Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D., Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W., Rosenthal A., Kudoh J., Schibuya K., Kawasaki K., Asakawa S., Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S. E., Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brant P., Scharfe M., Schon O., Desario A., Reichelt J., Kauer G., Blocker H., Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E., Haaf T., Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F., Lehrach H., Reinhardt R., and Yaspo M. L. (2000). The DNA sequence of human chromosome 21. *Nature* **405**: 311-9.
- Hieter P., and Boguski M. (1997). Functional genomics: it's all how you read it. *Science* **278**: 601-2.
- Hill W. G., and Robertson A. (1968). Linkage-disequilibrium in finite populations. *Theor. Appl. Genet.* **38**: 226-31.
- Hillier L. D., Lennon G., Becker M., Bonaldo M. F., Chiapelli B., Chisoe S., Dietrich N., DuBuque T., Favello A., Gish W., Hawkins M., Hultman M., Kucaba T., Lacy M., Le M., Le N., Mardis E., Moore B., Morris M., Parsons J., Prange C., Rifkin L., Rohlfing T., Schellenberg K., Marra M., and *et al.* (1996). Generation and analysis of 280,000 human expressed sequence tags. *Genome Res* **6**: 807-28.
- Hogenesch J. B., Ching K. A., Batalov S., Su A. I., Walker J. R., Zhou Y., Kay S. A., Schultz P. G., and Cooke M. P. (2001). A comparison of the Celera and Ensembl predicted gene sets reveals little overlap in novel genes. *Cell* **106**: 413-5.
- Holloway J. W., Beghe B., Turner S., Hinks L. J., Day I. N., and Howell W. M. (1999). Comparison of three methods for single nucleotide polymorphism typing for DNA bank studies: sequence-specific oligonucleotide probe hybridisation, TaqMan liquid phase hybridisation, and microplate array diagonal gel electrophoresis (MADGE). *Hum Mutat* **14**: 340-7.
- Holmquist G. P. (1992). Chromosome bands, their chromatin flavors, and their functional features. *Am J Hum Genet* **51**: 17-37.

- Horton R., Niblett D., Milne S., Palmer S., Tubby B., Trowsdale J., and Beck S. (1998). Large-scale sequence comparisons reveal unusually high levels of variation in the HLA-DQB1 locus in the class II region of the human MHC. *J Mol Biol* **282**: 71-97.
- Houck C. M., Rinehart F. P., and Schmid C. W. (1979). A ubiquitous family of repeated DNA sequences in the human genome. *J Mol Biol* **132**: 289-306.
- Hubbard T., Barker D., Birney E., Cameron G., Chen Y., Clark L., Cox T., Cuff J., Curwen V., Down T., Durbin R., Eyraas E., Gilbert J., Hammond M., Huminiecki L., Kasprzyk A., Lehvaslaiho H., Lijnzaad P., Melsopp C., Mongin E., Pettett R., Pocock M., Potter S., Rust A., Schmidt E., Searle S., Slater G., Smith J., Spooner W., Stabenau A., Stalker J., Stupka E., Ureta-Vidal A., Vastrik I., and Clamp M. (2002). The Ensembl genome database project. *Nucleic Acids Res* **30**: 38-41.
- Hudson T. J., Engelstein M., Lee M. K., Ho E. C., Rubenfield M. J., Adams C. P., Housman D. E., and Dracopoli N. C. (1992). Isolation and chromosomal assignment of 100 highly informative human simple sequence repeat polymorphisms. *Genomics* **13**: 622-9.
- Hudson T. J., Stein L. D., Gerety S. S., Ma J., Castle A. B., Silva J., Slonim D. K., Baptista R., Kruglyak L., Xu S. H., and *et al.* (1995). An STS-based map of the human genome. *Science* **270**: 1945-54.
- Hughes T. R., Marton M. J., Jones A. R., Roberts C. J., Stoughton R., Armour C. D., Bennett H. A., Coffey E., Dai H., He Y. D., Kidd M. J., King A. M., Meyer M. R., Slade D., Lum P. Y., Stepaniants S. B., Shoemaker D. D., Gachotte D., Chakraborty K., Simon J., Bard M., and Friend S. H. (2000). Functional discovery via a compendium of expression profiles. *Cell* **102**: 109-26.
- Hugot J. P., Chamaillard M., Zouali H., Lesage S., Cezard J. P., Belaiche J., Almer S., Tysk C., O'Morain C. A., Gassull M., Binder V., Finkel Y., Cortot A., Modigliani R., Laurent-Puig P., Gower-Rousseau C., Macry J., Colombel J. F., Sahbatou M., and Thomas G. (2001). Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* **411**: 599-603.
- Humphray S. J., Knaggs S. J., and Ragoussis I. (2001). Contiguation of bacterial clones. *Methods Mol Biol* **175**: 69-108.
- International Human Genome Mapping Consortium (IHGMC): McPherson J. D., Marra M., Hillier L., Waterston R. H., Chinwalla A., Wallis J., Sekhon M., Wylie K., Mardis E. R., Wilson R. K., Fulton R., Kucaba T. A., Wagner-McPherson C., Barbazuk W. B., Gregory S. G., Humphray S. J., French L., Evans R. S., Bethel G., Whittaker A., Holden J. L., McCann O. T., Dunham A., Soderlund C., Scott C. E., Bentley D. R., Schuler G., Chen H. C., Jang W., Green E. D., Idol J. R., Maduro V. V., Montgomery K. T., Lee E., Miller A., Emerling S., Kucherlapati, Gibbs R., Scherer S., Gorrell J. H., Sodergren E., Clerc-Blankenburg K., Tabor P., Naylor S., Garcia D., de Jong P. J., Catanese J. J., Nowak N., Osoegawa K., Qin S., Rowen L., Madan A., Dors M., Hood L., Trask B., Friedman C., Massa H., Cheung V. G., Kirsch I. R., Reid T., Yonescu R., Weissenbach J., Bruls T., Heilig R., Branscomb E., Olsen A., Doggett N., Cheng J. F., Hawkins T., Myers R. M., Shang J., Ramirez L., Schmutz J., Velasquez O., Dixon K., Stone N. E., Cox D. R., Haussler D., Kent W. J., Furey T., Rogic S., Kennedy S., Jones S., Rosenthal A., Wen G.,

- Schilhabel M., Gloeckner G., Nyakatura G., Siebert R., Schlegelberger B., Korenberg J., Chen X. N., Fujiyama A., Hattori M., Toyoda A., Yada T., Park H. S., Sakaki Y., Shimizu N., *et al.* (2001). A physical map of the human genome. *Nature* **409**: 934-41.
- International Human Genome Sequencing Consortium (IHGSC): Lander E. S., Linton L. M., Birren B., Nusbaum C., Zody M. C., Baldwin J., Devon K., Dewar K., Doyle M., FitzHugh W., Funke R., Gage D., Harris K., Heaford A., Howland J., Kann L., Lehoczky J., LeVine R., McEwan P., McKernan K., Meldrim J., Mesirov J. P., Miranda C., Morris W., Naylor J., Raymond C., Rosetti M., Santos R., Sheridan A., Sougnez C., Stange-Thomann N., Stojanovic N., Subramanian A., Wyman D., Rogers J., Sulston J., Ainscough R., Beck S., Bentley D., Burton J., Clee C., Carter N., Coulson A., Deadman R., Deloukas P., Dunham A., Dunham I., Durbin R., French L., Grafham D., Gregory S., Hubbard T., Humphray S., Hunt A., Jones M., Lloyd C., McMurray A., Matthews L., Mercer S., Milne S., Mullikin J. C., Mungall A., Plumb R., Ross M., Shownkeen R., Sims S., Waterston R. H., Wilson R. K., Hillier L. W., McPherson J. D., Marra M. A., Mardis E. R., Fulton L. A., Chinwalla A. T., Pepin K. H., Gish W. R., Chissoe S. L., Wendl M. C., Delehaunty K. D., Miner T. L., Delehaunty A., Kramer J. B., Cook L. L., Fulton R. S., Johnson D. L., Minx P. J., Clifton S. W., Hawkins T., Branscomb E., Predki P., Richardson P., Wenning S., Slezak T., Doggett N., Cheng J. F., Olsen A., Lucas S., Elkin C., Uberbacher E., Frazier M., *et al.* (2001). Initial sequencing and analysis of the human genome. *Nature* **409**: 860-921.
- International SNP Map Working Group (ISNPMWG): Sachidanandam R., Weissman D., Schmidt S. C., Kakol J. M., Stein L. D., Marth G., Sherry S., Mullikin J. C., Mortimore B. J., Willey D. L., Hunt S. E., Cole C. G., Coggill P. C., Rice C. M., Ning Z., Rogers J., Bentley D. R., Kwok P. Y., Mardis E. R., Yeh R. T., Schultz B., Cook L., Davenport R., Dante M., Fulton L., Hillier L., Waterston R. H., McPherson J. D., Gilman B., Schaffner S., Van Etten W. J., Reich D., Higgins J., Daly M. J., Blumenstiel B., Baldwin J., Stange-Thomann N., Zody M. C., Linton L., Lander E. S., and Altshuler D. (2001). A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409**: 928-33.
- Ioannou P. A., Amemiya C. T., Garnes J., Kroisel P. M., Shizuya H., Chen C., Batzer M. A., and de Jong P. J. (1994). A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nat Genet* **6**: 84-9.
- Irizarry K., Kustanovich V., Li C., Brown N., Nelson S., Wong W., and Lee C. J. (2000). Genome-wide analysis of single-nucleotide polymorphisms in human expressed sequences. *Nat Genet* **26**: 233-6.
- Isaksson A., Landegren U., Syvanen A. C., Bork P., Stein C., Ortigao F., and Brookes A. J. (2000). Discovery, scoring and utilization of human single nucleotide polymorphisms: a multidisciplinary problem. *Eur J Hum Genet* **8**: 154-6.
- Jackson I. J. (2001). Mouse genomics: making sense of the sequence. *Curr Biol* **11**: R311-4.
- Jackson P. E., Scholl P. F., and Groopman J. D. (2000). Mass spectrometry for genotyping: an emerging tool for molecular medicine. *Molecular Medicine Today* **6**: 271-76.

- Jacq C., Miller J. R., and Brownlee G. G. (1977). A pseudogene structure in 5S DNA of *Xenopus laevis*. *Cell* **12**: 109-20.
- James M. R., and Lindpaintner K. (1997). Why map the rat? *Trends Genet* **13**: 171-3.
- Jareborg N., Birney E., and Durbin R. (1999). Comparative analysis of noncoding regions of 77 orthologous mouse and human gene pairs. *Genome Res* **9**: 815-24.
- Jeffreys A. J., Kauppi L., and Neumann R. (2001). Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. *Nat Genet* **29**: 217-22.
- Jeffreys A. J., Wilson V., and Thein S. L. (1985). Hypervariable minisatellite regions in human DNA. *Nature* **314**: 67-73.
- Jenkins S., and Gibson N. (2002). High-throughput SNP genotyping. *Comparative and Functional Genomics* **3**: 57-66.
- Ji L., Malecki M., Warram J. H., Yang Y., Rich S. S., and Krolewski A. S. (1997). New susceptibility locus for NIDDM is localized to human chromosome 20q. *Diabetes* **46**: 876-81.
- Ji Y., Eichler E. E., Schwartz S., and Nicholls R. D. (2000). Structure of chromosomal duplicons and their role in mediating human genomic disorders. *Genome Res* **10**: 597-610.
- Jordan E., and Collins F. S. (1996). A march of genetic maps. *Nature* **380**: 111-2.
- Jorde L. B. (2000). Linkage disequilibrium and the search for complex disease genes. *Genome Res* **10**: 1435-44.
- Jurka J. (2000). Repbase update: a database and an electronic journal of repetitive elements. *Trends Genet* **16**: 418-20.
- Jurka J., and Milosavljevic A. (1991). Reconstruction and analysis of human Alu genes. *J Mol Evol* **32**: 105-21.
- Jurka J., Zietkiewicz E., and Labuda D. (1995). Ubiquitous mammalian-wide interspersed repeats (MIRs) are molecular fossils from the mesozoic era. *Nucleic Acids Res* **23**: 170-5.
- Kan Z., Rouchka E. C., Gish W. R., and States D. J. (2001). Gene structure prediction and alternative splicing analysis using genomically aligned ESTs. *Genome Res* **11**: 889-900.
- Karas M., and Hillenkamp F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* **60**: 2299-301.
- Kass D. H. (2001). Impact of SINEs and LINEs on the Mammalian Genome. *Current Genomics* **2**: 199-219.
- Kelley S. (2000). Getting started with Acedb. *Brief Bioinform* **1**: 131-7.
- Kelly P. D., Chu F., Woods I. G., Ngo-Hazelett P., Cardozo T., Huang H., Kimm F., Liao L., Yan Y. L., Zhou Y., Johnson S. L., Abagyan R., Schier A. F., Postlethwait J. H., and Talbot W. S. (2000). Genetic linkage mapping of zebrafish genes and ESTs. *Genome Res* **10**: 558-67.
- Kent W. J. (2002). BLAT-the BLAST-like alignment tool. *Genome Res* **12**: 656-64.

- Kent W. J., Sugnet C. W., Furey T. S., Roskin K. M., Pringle T. H., Zahler A. M., and Haussler A. D. (2002). The Human Genome Browser at UCSC. *Genome Res* **12**: 996-1006.
- Kessler M. M., Beckendorf R. C., Westhafer M. A., and Nordstrom J. L. (1986). Requirement of A-A-U-A-A-A and adjacent downstream sequences for SV40 early polyadenylation. *Nucleic Acids Res* **14**: 4939-52.
- Kim J., Gordon L., Dehal P., Badri H., Christensen M., Groza M., Ha C., Hammond S., Vargas M., Wehri E., Wagner M., Olsen A., and Stubbs L. (2001). Homology-driven assembly of a sequence-ready mouse BAC contig map spanning regions related to the 46-Mb gene-rich euchromatic segments of human chromosome 19. *Genomics* **74**: 129-41.
- Kim U. J., Shizuya H., de Jong P. J., Birren B., and Simon M. I. (1992). Stable propagation of cosmid sized human DNA inserts in an F factor based vector. *Nucleic Acids Res* **20**: 1083-5.
- Kirpekar F., Nordhoff E., Larsen L. K., Kristiansen K., Roepstorff P., and Hillenkamp F. (1998). DNA sequence analysis by MALDI mass spectrometry. *Nucleic Acids Res* **26**: 2554-9.
- Kong A., Gudbjartsson D. F., Sainz J., Jonsdottir G. M., Gudjonsson S. A., Richardsson B., Sigurdardottir S., Barnard J., Hallbeck B., Masson G., Shlien A., Palsson S. T., Frigge M. L., Thorgeirsson T. E., Gulcher J. R., and Stefansson K. (2002). A high-resolution recombination map of the human genome. *Nat Genet* **31**: 241-7.
- Koop B. F., and Hood L. (1994). Striking sequence similarity over almost 100 kilobases of human and mouse T-cell receptor DNA. *Nat Genet* **7**: 48-53.
- Korn B., Sedlacek Z., Manca A., Kioschis P., Konecki D., Lehrach H., and Poustka A. (1992). A strategy for the selection of transcribed sequences in the Xq28 region. *Hum Mol Genet* **1**: 235-42.
- Köster H., Tang K., Fu D. J., Braun A., van den Boom D., Smith C. L., Cotter R. J., and Cantor C. R. (1996). A strategy for rapid and efficient DNA sequencing by mass spectrometry. *Nat Biotechnol* **14**: 1123-8.
- Kozak M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res* **15**: 8125-48.
- Kristensen V. N., Kelefiotis D., Kristensen T., and Borresen-Dale A. L. (2001). High-throughput methods for detection of genetic variation. *Biotechniques* **30**: 318-22, 324, 326 passim.
- Kruglyak L. (1997). The use of a genetic map of biallelic markers in linkage studies. *Nat Genet* **17**: 21-4.
- Kruglyak L. (1999). Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat Genet* **22**: 139-44.
- Kruglyak L., and Nickerson D. A. (2001). Variation is the spice of life. *Nat Genet* **27**: 234-6.
- Kwitek A. E., Tonellato P. J., Chen D., Gullings-Handley J., Cheng Y. S., Twigger S., Scheetz T. E., Casavant T. L., Stoll M., Nobrega M. A., Shiozawa M., Soares M. B., Sheffield V. C., and Jacob H. J. (2001). Automated construction of

- high-density comparative maps between rat, human, and mouse. *Genome Res* **11**: 1935-43.
- Kwok P. Y. (1998). Genotyping by mass spectroscopy takes flight. *Nat. Biotech.* **16**: 1314-15.
- Kwok P. Y., and Gu Z. (1999). Single nucleotide polymorphism libraries: why and how are we building them? *Mol Med Today* **5**: 538-43.
- Laken S. J., Jackson P. E., Kinzler K. W., Vogelstein B., Strickland P. T., Groopman J. D., and Friesen M. D. (1998). Genotyping by mass spectrometric analysis of short DNA fragments. *Nat Biotechnol* **16**: 1352-6.
- Lamerdin J. E., Montgomery M. A., Stilwagen S. A., Scheidecker L. K., Tebbs R. S., Brookman K. W., Thompson L. H., and Carrano A. V. (1995). Genomic sequence comparison of the human and mouse XRCC1 DNA repair gene regions. *Genomics* **25**: 547-54.
- Lamerdin J. E., Stilwagen S. A., Ramirez M. H., Stubbs L., and Carrano A. V. (1996). Sequence analysis of the ERCC2 gene regions in human, mouse, and hamster reveals three linked genes. *Genomics* **34**: 399-409.
- Landegren U., Nilsson M., and Kwok P. Y. (1998). Reading bits of genetic information: methods for single-nucleotide polymorphism analysis. *Genome Res* **8**: 769-76.
- Lander E. S. (1996). The new genomics: global views of biology. *Science* **274**: 536-9.
- Larsen F., Gundersen G., Lopez R., and Prydz H. (1992). CpG islands as gene markers in the human genome. *Genomics* **13**: 1095-107.
- Lee K. H. (2001). Proteomics: a technology-driven and technology-limited discovery science. *Trends Biotechnol* **19**: 217-22.
- Lee P. S., and Lee K. H. (2000). Genomic analysis. *Curr Opin Biotechnol* **11**: 171-5.
- Lee T. I., and Young R. A. (2000). Transcription of eukaryotic protein-coding genes. *Annu Rev Genet* **34**: 77-137.
- Legrain P., and Selig L. (2000). Genome-wide protein interaction maps using two-hybrid systems. *FEBS Lett* **480**: 32-6.
- Leushner J. (2001). MALDI TOF mass spectrometry: an emerging platform for genomics and diagnostics. *Expert Rev Mol Diagn* **1**: 11-8.
- Leushner J., and Chiu N. H. (2000). Automated mass spectrometry: a revolutionary technology for clinical diagnostics. *Mol Diagn* **5**: 341-8.
- Lewin B. (1994). *Genes V*: 657-660.
- Lewontin R. C. (1964). The interaction of selection and linkage. I. General considerations; heterotic models. *Genetics* **49**: 49-67.
- Li L., Krantz I. D., Deng Y., Genin A., Banta A. B., Collins C. C., Qi M., Trask B. J., Kuo W. L., Cochran J., Costa T., Pierpont M. E., Rand E. B., Piccoli D. A., Hood L., and Spinner N. B. (1997). Alagille syndrome is caused by mutations in human Jagged1, which encodes a ligand for Notch1. *Nat Genet* **16**: 243-51.
- Li Q., Harju S., and Peterson K. R. (1999). Locus control regions: coming of age at a decade plus. *Trends Genet* **15**: 403-8.

- Li W. H., Ellsworth, D.L., Krushkal, J., Chang, B.H., and Hewett-Emmett, D (1996). Rates of nucleotide substitution in primates and rodents and the generation-time effect hypothesis. *Mol. Phylogenet. Evol.* **5**: 182-187.
- Liang F., Holt I., Pertea G., Karamycheva S., Salzberg S. L., and Quackenbush J. (2000). Gene index analysis of the human genome estimates approximately 120,000 genes. *Nat Genet* **25**: 239-40.
- Liang P., and Pardee A. B. (1992). Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* **257**: 967-71.
- Liang P., and Pardee A. B. (1995). Recent advances in differential display. *Curr Opin Immunol* **7**: 274-80.
- Lieberman A. P., and Fischbeck K. H. (2000). Triplet repeat expansion in neuromuscular disease. *Muscle Nerve* **23**: 843-50.
- Lindsay S., and Bird A. P. (1987). Use of restriction enzymes to detect potential gene sequences in mammalian DNA. *Nature* **327**: 336-8.
- Litt M., and Luty J. A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am J Hum Genet* **44**: 397-401.
- Livak K. J. (1999). Allelic discrimination using fluorogenic probes and the 5' nuclease assay. *Genet Anal* **14**: 143-9.
- Livak K. J., Flood S. J., Marmaro J., Giusti W., and Deetz K. (1995). Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridisation. *PCR Methods Appl.* **4**: 357-62.
- Logsdon J. M., Jr., Tyshenko M. G., Dixon C., J D. J., Walker V. K., and Palmer J. D. (1995). Seven newly discovered intron positions in the triose-phosphate isomerase gene: evidence for the introns-late theory. *Proc Natl Acad Sci U S A* **92**: 8507-11.
- Long M., Rosenberg C., and Gilbert W. (1995). Intron phase correlations and the evolution of the intron/exon structure of genes. *Proc Natl Acad Sci U S A* **92**: 12495-9.
- Loots G. G., Locksley R. M., Blankespoor C. M., Wang Z. E., Miller W., Rubin E. M., and Frazer K. A. (2000). Identification of a coordinate regulator of interleukins 4, 13, and 5 by cross-species sequence comparisons. *Science* **288**: 136-40.
- Lovett M., Kere J., and Hinton L. M. (1991). Direct selection: a method for the isolation of cDNAs encoded by large genomic regions. *Proc Natl Acad Sci U S A* **88**: 9628-32.
- Lundwall A. (1996). The structure of the semenogelin gene locus--nucleotide sequence of the intergenic and the flanking DNA. *Eur J Biochem* **235**: 466-70.
- Lyamichev V., Mast A. L., Hall J. G., Prudent J. R., Kaiser M. W., Takova T., Kwiatkowski R. W., Sander T. J., de Arruda M., Arco D. A., Neri B. P., and Brow M. A. (1999). Polymorphism identification and quantitative detection of genomic DNA by invasive cleavage of oligonucleotide probes. *Nat Biotechnol* **17**: 292-6.

- Maestre J., Tchenio T., Dhellin O., and Heidmann T. (1995). mRNA retroposition in human cells: processed pseudogene formation. *Embo J* **14**: 6333-8.
- Maillet I., Lagniel G., Perrot M., Boucherie H., and Labarre J. (1996). Rapid identification of yeast proteins on two-dimensional gels. *J Biol Chem* **271**: 10263-70.
- Makalowski W. (2000). Genomic scrap yard: how genomes utilize all that junk. *Gene* **259**: 61-7.
- Makalowski W., and Boguski M. S. (1998). Evolutionary parameters of the transcribed mammalian genome: an analysis of 2,820 orthologous rodent and human sequences. *Proc Natl Acad Sci U S A* **95**: 9407-12.
- Makalowski W., Zhang J., and Boguski M. S. (1996). Comparative analysis of 1196 orthologous mouse and human full-length mRNA and protein sequences. *Genome Res* **6**: 846-57.
- Mallon A. M., Platzer M., Bate R., Gloeckner G., Botcherby M. R., Nordsiek G., Strivens M. A., Kioschis P., Dangel A., Cunningham D., Straw R. N., Weston P., Gilbert M., Fernando S., Goodall K., Hunter G., Greystrom J. S., Clarke D., Kimberley C., Goerdes M., Blechschmidt K., Rump A., Hinzmann B., Mundy C. R., Miller W., Poustka A., Herman G. E., Rhodes M., Denny P., Rosenthal A., and Brown S. D. (2000). Comparative genome sequence analysis of the Bpa/Str region in mouse and Man. *Genome Res* **10**: 758-75.
- Marra M., Hillier L., Kucaba T., Allen M., Barstead R., Beck C., Blistain A., Bonaldo M., Bowers Y., Bowles L., Cardenas M., Chamberlain A., Chappell J., Clifton S., Favello A., Geisel S., Gibbons M., Harvey N., Hill F., Jackson Y., Kohn S., Lennon G., Mardis E., Martin J., Waterston R., and *et al.* (1999). An encyclopedia of mouse genes. *Nat Genet* **21**: 191-4.
- Marra M. A., Kucaba T. A., Dietrich N. L., Green E. D., Brownstein B., Wilson R. K., McDonald K. M., Hillier L. W., McPherson J. D., and Waterston R. H. (1997). High throughput fingerprint analysis of large-insert clones. *Genome Res* **7**: 1072-84.
- Marras S. A., Kramer F. R., and Tyagi S. (1999). Multiplex detection of single-nucleotide variations using molecular beacons. *Genet Anal* **14**: 151-6.
- Marshall E. (1999). Drug firms to create public database of genetic mutations. *Science* **284**: 406-7.
- Marth G., Yeh R., Minton M., Donaldson R., Li Q., Duan S., Davenport R., Miller R. D., and Kwok P. Y. (2001). Single-nucleotide polymorphisms in the public domain: how useful are they? *Nat Genet* **27**: 371-2.
- Martin D. B., and Nelson P. S. (2001). From genomics to proteomics: techniques and applications in cancer research. *Trends Cell Biol* **11**: S60-5.
- Martindale D. W., Wilson M. D., Wang D., Burke R. D., Chen X., Duronio V., and Koop B. F. (2000). Comparative genomic sequence analysis of the Williams syndrome region (LIMK1-RFC2) of human chromosome 7q11.23. *Mamm Genome* **11**: 890-8.

- Martzen M. R., McCraith S. M., Spinelli S. L., Torres F. M., Fields S., Grayhack E. J., and Phizicky E. M. (1999). A biochemical genomics approach for identifying genes by the activity of their products. *Science* **286**: 1153-5.
- Masood E. (1999). As consortium plans free SNP map of human genome. *Nature* **398**: 545-6.
- Mattick J. S. (2001). Non-coding RNAs: the architects of eukaryotic complexity. *EMBO Rep* **2**: 986-91.
- Matz M. V., and Lukyanov S. A. (1998). Different strategies of differential display: areas of application. *Nucleic Acids Res* **26**: 5537-43.
- Mayor C., Brudno M., Schwartz J. R., Poliakov A., Rubin E. M., Frazer K. A., Pachter L. S., and Dubchak I. (2000). VISTA : visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics* **16**: 1046-7.
- McCarthy L. C., Bihoreau M. T., Kiguwa S. L., Browne J., Watanabe T. K., Hishigaki H., Tsuji A., Kiel S., Webber C., Davis M. E., Knights C., Smith A., Critcher R., Huxtall P., Hudson J. R., Jr., Ono T., Hayashi H., Takagi T., Nakamura Y., Tanigami A., Goodfellow P. N., Lathrop G. M., and James M. R. (2000). A whole-genome radiation hybrid panel and framework map of the rat genome. *Mamm Genome* **11**: 791-5.
- Meisler M. H. (2001). Evolutionarily conserved noncoding DNA in the human genome: how much and what for? *Genome Res* **11**: 1617-8.
- Mighell A. J., Markham A. F., and Robinson P. A. (1997). Alu sequences. *FEBS Lett* **417**: 1-5.
- Mighell A. J., Smith N. R., Robinson P. A., and Markham A. F. (2000). Vertebrate pseudogenes. *FEBS Lett* **468**: 109-14.
- Miki R., Kadota K., Bono H., Mizuno Y., Tomaru Y., Carninci P., Itoh M., Shibata K., Kawai J., Konno H., Watanabe S., Sato K., Tokusumi Y., Kikuchi N., Ishii Y., Hamaguchi Y., Nishizuka I., Goto H., Nitanda H., Satomi S., Yoshiki A., Kusakabe M., DeRisi J. L., Eisen M. B., Iyer V. R., Brown P. O., Muramatsu M., Shimada H., Okazaki Y., and Hayashizaki Y. (2001). Delineating developmental and metabolic pathways in vivo by expression profiling using the RIKEN set of 18,816 full-length enriched mouse cDNA arrays. *Proc Natl Acad Sci U S A* **98**: 2199-204.
- Miles C., Elgar G., Coles E., Kleinjan D. J., van Heyningen V., and Hastie N. (1998). Complete sequencing of the Fugu WAGR region from WT1 to PAX6: dramatic compaction and conservation of synteny with human chromosome 11p13. *Proc Natl Acad Sci U S A* **95**: 13068-72.
- Miller R. D., and Kwok P. Y. (2001). The birth and death of human single-nucleotide polymorphisms: new experimental evidence and implications for human history and medicine. *Hum Mol Genet* **10**: 2195-8.
- Miller W. (2001). Comparison of genomic DNA sequences: solved and unsolved problems. *Bioinformatics* **17**: 391-7.
- Mills A. A., and Bradley A. (2001). From mouse to man: generating megabase chromosome rearrangements. *Trends Genet* **17**: 331-9.

- Mironov A. A., Fickett J. W., and Gelfand M. S. (1999). Frequent alternative splicing of human genes. *Genome Res* **9**: 1288-93.
- Moffatt M. F., Traherne J. A., Abecasis G. R., and Cookson W. O. (2000). Single nucleotide polymorphism and linkage disequilibrium within the TCR alpha/delta locus. *Hum Mol Genet* **9**: 1011-9.
- Monaco A. P., Neve R. L., Colletti-Feener C., Bertelson C. J., Kurnit D. M., and Kunkel L. M. (1986). Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. *Nature* **323**: 646-50.
- Montgomery K. T., Lee E., Miller A., Lau S., Shim C., Decker J., Chiu D., Emerling S., Sekhon M., Kim R., Lenz J., Han J., Ioshikhes I., Renault B., Marondel I., Yoon S. J., Song K., Murty V. V., Scherer S., Yonescu R., Kirsch I. R., Ried T., McPherson J., Gibbs R., and Kucherlapati R. (2001). A high-resolution map of human chromosome 12. *Nature* **409**: 945-6.
- Moore M. J., and Sharp P. A. (1993). Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. *Nature* **365**: 364-8.
- Morgan J. G., Dolganov G. M., Robbins S. E., Hinton L. M., and Lovett M. (1992). The selective isolation of novel cDNAs encoded by the regions surrounding the human interleukin 4 and 5 genes. *Nucleic Acids Res* **20**: 5173-9.
- Mori I., Benian G. M., Moerman D. G., and Waterston R. H. (1988). Transposable element Tc1 of *Caenorhabditis elegans* recognizes specific target sequences for integration. *Proc Natl Acad Sci U S A* **85**: 861-4.
- Morton N. E. (1991). Parameters of the human genome. *Proc Natl Acad Sci U S A* **88**: 7474-6.
- Mullikin J. C., Hunt S. E., Cole C. G., Mortimore B. J., Rice C. M., Burton J., Matthews L. H., Pavitt R., Plumb R. W., Sims S. K., Ainscough R. M., Attwood J., Bailey J. M., Barlow K., Bruskiewich R. M., Butcher P. N., Carter N. P., Chen Y., Clee C. M., Coggill P. C., Davies J., Davies R. M., Dawson E., Francis M. D., Joy A. A., Lambie R. G., Langford C. F., Macarthy J., Mall V., Moreland A., Overton-Larty E. K., Ross M. T., Smith L. C., Steward C. A., Sulston J. E., Tinsley E. J., Turney K. J., Willey D. L., Wilson G. D., McMurray A. A., Dunham I., Rogers J., and Bentley D. R. (2000). An SNP map of human chromosome 22. *Nature* **407**: 516-20.
- Mungall A. J., Edwards C. A., Ranby S. A., Humphray S. J., Heathcote R. W., Clee C. M., East C. L., Holloway E., Butler A. P., Langford C. F., Gwilliam R., Rice K. M., Maslen G. L., Carter N. P., Ross M. T., Deloukas P., Bentley D. R., and Dunham I. (1996). Physical mapping of chromosome 6: a strategy for the rapid generation of sequence-ready contigs. *DNA Seq* **7**: 47-9.
- Murakami K., and Takagi T. (1998). Gene recognition by combination of several gene-finding programs. *Bioinformatics* **14**: 665-75.
- Mural R. J., Adams M. D., Myers E. W., Smith H. O., Miklos G. L., Wides R., Halpern A., Li P. W., Sutton G. G., Nadeau J., Salzberg S. L., Holt R. A., Kodira C. D., Lu F., Chen L., Deng Z., Evangelista C. C., Gan W., Heiman T. J., Li J., Li Z., Merkulov G. V., Milshina N. V., Naik A. K., Qi R., Shue B. C., Wang A., Wang J., Wang X., Yan X., Ye J., Yooseph S., Zhao Q., Zheng L., Zhu S. C., Biddick K., Bolanos R., Delcher A. L., Dew I. M., Fasulo D.,

- Flanigan M. J., Huson D. H., Kravitz S. A., Miller J. R., Mobarry C. M., Reinert K., Remington K. A., Zhang Q., Zheng X. H., Nusskern D. R., Lai Z., Lei Y., Zhong W., Yao A., Guan P., Ji R. R., Gu Z., Wang Z. Y., Zhong F., Xiao C., Chiang C. C., Yandell M., Wortman J. R., Amanatides P. G., Hladun S. L., Pratts E. C., Johnson J. E., Dodson K. L., Woodford K. J., Evans C. A., Gropman B., Rusch D. B., Venter E., Wang M., Smith T. J., Houck J. T., Tompkins D. E., Haynes C., Jacob D., Chin S. H., Allen D. R., Dahlke C. E., Sanders R., Li K., Liu X., Levitsky A. A., Majoros W. H., Chen Q., Xia A. C., Lopez J. R., Donnelly M. T., Newman M. H., Glodek A., Kraft C. L., Nodell M., Ali F., An H. J., Baldwin-Pitts D., Beeson K. Y., Cai S., *et al.* (2002). A comparison of whole-genome shotgun-derived mouse chromosome 16 and the human genome. *Science* **296**: 1661-71.
- Murphy W. J., Stanyon R., and O'Brien S. J. (2001). Evolution of mammalian genome organization inferred from comparative gene mapping. *Genome Biol* **2**: REVIEWS0005.
- Murray J. C., Buetow K. H., Weber J. L., Ludwigsen S., Scherpbier-Heddema T., Manion F., Quillen J., Sheffield V. C., Sunden S., Duyk G. M., and *et al.* (1994). A comprehensive human linkage map with centimorgan density. Cooperative Human Linkage Center (CHLC). *Science* **265**: 2049-54.
- Myers E. W., Sutton G. G., Smith H. O., Adams M. D., and Venter J. C. (2002). On the sequencing and assembly of the human genome. *Proc Natl Acad Sci U S A* **99**: 4145-6.
- Nagaraja R., Kere J., MacMillan S., Masisi M. J., Johnson D., Molini B. J., Halley G. R., Wein K., Trusgnich M., Eble B., and *et al.* (1994). Characterization of four human YAC libraries for clone size, chimerism and X chromosome sequence representation. *Nucleic Acids Res* **22**: 3406-11.
- Nakamura Y., Leppert M., O'Connell P., Wolff R., Holm T., Culver M., Martin C., Fujimoto E., Hoff M., Kumlin E., and *et al.* (1987). Variable number of tandem repeat (VNTR) markers for human gene mapping. *Science* **235**: 1616-22.
- Ng P. C., and Henikoff S. (2002). Accounting for human polymorphisms predicted to affect protein function. *Genome Res* **12**: 436-46.
- Ning Z., Cox A. J., and Mullikin J. C. (2001). SSAHA: a fast search method for large DNA databases. *Genome Res* **11**: 1725-9.
- Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayasi Y., Sato S., Nagase T., Seki N., Ishikawa K., and Tabata S. (1994). Prediction of the coding sequences of unidentified human genes. I. The coding sequences of 40 new genes (KIAA0001-KIAA0040) deduced by analysis of randomly sampled cDNA clones from human immature myeloid cell line KG-1. *DNA Res* **1**: 27-35.
- Noordewier M. O., and Warren P. V. (2001). Gene expression microarrays and the integration of biological knowledge. *Trends Biotechnol* **19**: 412-5.
- Nordborg M., and Tavaré S. (2002). Linkage disequilibrium: what history has to tell us. *Trends Genet* **18**: 83-90.

- Nowotny P., Kwon J. M., and Goate A. M. (2001). SNP analysis to dissect human traits. *Curr Opin Neurobiol* **11**: 637-41.
- Nusbaum C., Slonim D. K., Harris K. L., Birren B. W., Steen R. G., Stein L. D., Miller J., Dietrich W. F., Nahf R., Wang V., Merport O., Castle A. B., Husain Z., Farino G., Gray D., Anderson M. O., Devine R., Horton L. T., Jr., Ye W., Wu X., Kouyoumjian V., Zemsteva I. S., Wu Y., Collymore A. J., Courtney D. F., and *et al.* (1999). A YAC-based physical map of the mouse genome. *Nat Genet* **22**: 388-93.
- O'Brien S. J., Menotti-Raymond M., Murphy W. J., Nash W. G., Wienberg J., Stanyon R., Copeland N. G., Jenkins N. A., Womack J. E., and Marshall Graves J. A. (1999). The promise of comparative genomics in mammals. *Science* **286**: 458-62, 479-81.
- Oeltjen J. C., Malley T. M., Muzny D. M., Miller W., Gibbs R. A., and Belmont J. W. (1997). Large-scale comparative sequence analysis of the human and murine Bruton's tyrosine kinase loci reveals conserved regulatory domains. *Genome Res* **7**: 315-29.
- Ogura Y., Bonen D. K., Inohara N., Nicolae D. L., Chen F. F., Ramos R., Britton H., Moran T., Karaliuskas R., Duerr R. H., Achkar J. P., Brant S. R., Bayless T. M., Kirschner B. S., Hanauer S. B., Nunez G., and Cho J. H. (2001). A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* **411**: 603-6.
- Olivier M., Bustos V. I., Levy M. R., Smick G. A., Moreno I., Bushard J. M., Almendras A. A., Sheppard K., Zierten D. L., Aggarwal A., Carlson C. S., Foster B. D., Vo N., Kelly L., Liu X., and Cox D. R. (2001). Complex high-resolution linkage disequilibrium and haplotype patterns of single-nucleotide polymorphisms in 2.5 Mb of sequence on human chromosome 21. *Genomics* **78**: 64-72.
- Olson M., Hood L., Cantor C., and Botstein D. (1989). A common language for physical mapping of the human genome. *Science* **245**: 1434-5.
- Olson M. V., Dutchik J. E., Graham M. Y., Brodeur G. M., Helms C., Frank M., MacCollin M., Scheinman R., and Frank T. (1986). Random-clone strategy for genomic restriction mapping in yeast. *Proc Natl Acad Sci U S A* **83**: 7826-30.
- Osoegawa K., Tateno M., Woon P. Y., Frengen E., Mammoser A. G., Catanese J. J., Hayashizaki Y., and de Jong P. J. (2000). Bacterial artificial chromosome libraries for mouse sequencing and functional analysis. *Genome Res* **10**: 116-28.
- Pandey A., and Mann M. (2000). Proteomics to study genes and genomes. *Nature* **405**: 837-46.
- Paracchini S., Arredi B., Chalk R., and Tyler-Smith C. (2002). Hierarchical high-throughput SNP genotyping of the human Y chromosome using MALDI-TOF mass spectrometry. *Nucleic Acids Res* **30**: e27.
- Parimoo S., Patanjali S. R., Shukla H., Chaplin D. D., and Weissman S. M. (1991). cDNA selection: efficient PCR approach for the selection of cDNAs encoded in large chromosomal DNA fragments. *Proc Natl Acad Sci U S A* **88**: 9623-7.

- Pastinen T., Kurg A., Metspalu A., Peltonen L., and Syvanen A. C. (1997). Minisequencing: a specific tool for DNA analysis and diagnostics on oligonucleotide arrays. *Genome Res* **7**: 606-14.
- Pastinen T., Raitio M., Lindroos K., Tainola P., Peltonen L., and Syvanen A. C. (2000). A system for specific, high-throughput genotyping by allele-specific primer extension on microarrays. *Genome Res* **10**: 1031-42.
- Patil N., Berno A. J., Hinds D. A., Barrett W. A., Doshi J. M., Hacker C. R., Kautzer C. R., Lee D. H., Marjoribanks C., McDonough D. P., Nguyen B. T., Norris M. C., Sheehan J. B., Shen N., Stern D., Stokowski R. P., Thomas D. J., Trulson M. O., Vyas K. R., Frazer K. A., Fodor S. P., and Cox D. R. (2001). Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. *Science* **294**: 1719-23.
- Paule M. R., and White R. J. (2000). Survey and summary: transcription by RNA polymerases I and III. *Nucleic Acids Res* **28**: 1283-98.
- Pearce S. H., Vaidya B., Imrie H., Perros P., Kelly W. F., Toft A. D., McCarthy M. I., Young E. T., and Kendall-Taylor P. (1999). Further evidence for a susceptibility locus on chromosome 20q13.11 in families with dominant transmission of Graves disease. *Am J Hum Genet* **65**: 1462-5.
- Pennacchio L. A., and Rubin E. M. (2001). Genomic strategies to identify mammalian regulatory sequences. *Nat Rev Genet* **2**: 100-9.
- Persidis A. (1999). Bioinformatics. *Nat Biotechnol* **17**: 828-30.
- Persidis A. (2000). Data mining in biotechnology. *Nat Biotechnol* **18**: 237-8.
- Peters J., Church D., Zuberi A. R., and Selley R. (1999). Mouse chromosome 2. *Mamm Genome* **10**: 941.
- Pfost D. R., Boyce-Jacino M. T., and Grant D. M. (2000). A SNPshot: pharmacogenetics and the future of drug therapy. *Trends Biotechnol* **18**: 334-8.
- Picoult-Newberg L., Ideker T. E., Pohl M. G., Taylor S. L., Donaldson M. A., Nickerson D. A., and Boyce-Jacino M. (1999). Mining SNPs from EST databases. *Genome Res* **9**: 167-74.
- Pletcher M. T., Wiltshire T., Cabin D. E., Villanueva M., and Reeves R. H. (2001). Use of comparative physical and sequence mapping to annotate mouse chromosome 16 and human chromosome 21. *Genomics* **74**: 45-54.
- Ponger L., Duret L., and Mouchiroud D. (2001). Determinants of CpG islands: expression in early embryo and isochore structure. *Genome Res* **11**: 1854-60.
- Prak E. T., and Kazazian H. H., Jr. (2000). Mobile elements and the human genome. *Nat Rev Genet* **1**: 134-44.
- Price J. A., Brewer C. S., Howard T. D., Fossey S. C., Sale M. M., Ji L., Krolewski A. S., and Bowden D. W. (1999). A physical map of the 20q12-q13.1 region associated with type 2 diabetes. *Genomics* **62**: 208-15.
- Pritchard J. K., and Przeworski M. (2001). Linkage disequilibrium in humans: models and data. *Am J Hum Genet* **69**: 1-14.

- Pruitt K. D., and Maglott D. R. (2001). RefSeq and LocusLink: NCBI gene-centered resources. *Nucleic Acids Res* **29**: 137-40.
- Przeworski M., Hudson R. R., and Di Rienzo A. (2000). Adjusting the focus on human variation. *Trends Genet* **16**: 296-302.
- Qiu Y., Cavelier L., Chiu S., Yang X., Rubin E., and Cheng J. F. (2001). Human and mouse ABCA1 comparative sequencing and transgenesis studies revealing novel regulatory sequences. *Genomics* **73**: 66-76.
- Ranade K., Chang M. S., Ting C. T., Pei D., Hsiao C. F., Olivier M., Pesich R., Hebert J., Chen Y. D., Dzau V. J., Curb D., Olshen R., Risch N., Cox D. R., and Botstein D. (2001). High-throughput genotyping with single nucleotide polymorphisms. *Genome Res* **11**: 1262-8.
- Rapp J. P. (2000). Genetic analysis of inherited hypertension in the rat. *Physiol Rev* **80**: 135-72.
- Reese M. G., Hartzell G., Harris N. L., Ohler U., Abril J. F., and Lewis S. E. (2000). Genome annotation assessment in *Drosophila melanogaster*. *Genome Res* **10**: 483-501.
- Reich D. E., Cargill M., Bolk S., Ireland J., Sabeti P. C., Richter D. J., Lavery T., Kouyoumjian R., Farhadian S. F., Ward R., and Lander E. S. (2001). Linkage disequilibrium in the human genome. *Nature* **411**: 199-204.
- Reichhardt T. (1999). It's sink or swim as a tidal wave of data approaches. *Nature* **399**: 517-20.
- Remm M., and Metspalu A. (2002). High-density genotyping and linkage disequilibrium in the human genome using chromosome 22 as a model. *Curr Opin Chem Biol* **6**: 24-30.
- Ren B., Robert F., Wyrick J. J., Aparicio O., Jennings E. G., Simon I., Zeitlinger J., Schreiber J., Hannett N., Kanin E., Volkert T. L., Wilson C. J., Bell S. P., and Young R. A. (2000). Genome-wide location and function of DNA binding proteins. *Science* **290**: 2306-9.
- Rhodes M., Straw R., Fernando S., Evans A., Lacey T., Dearlove A., Greystrom J., Walker J., Watson P., Weston P., Kelly M., Taylor D., Gibson K., Mundy C., Bourgade F., Poirier C., Simon D., Brunialti A. L., Montagutelli X., Gu'enet J. L., Haynes A., and Brown S. D. (1998). A high-resolution microsatellite map of the mouse genome. *Genome Res* **8**: 531-42.
- Rich S. S. (1990). Mapping genes in diabetes. Genetic epidemiological perspective. *Diabetes* **39**: 1315-9.
- Riley J., Butler R., Ogilvie D., Finniear R., Jenner D., Powell S., Anand R., Smith J. C., and Markham A. F. (1990). A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucleic Acids Res* **18**: 2887-90.
- Rioux J. D., Daly M. J., Silverberg M. S., Lindblad K., Steinhart H., Cohen Z., Delmonte T., Kocher K., Miller K., Guschwan S., Kulbokas E. J., O'Leary S., Winchester E., Dewar K., Green T., Stone V., Chow C., Cohen A., Langelier D., Lapointe G., Gaudet D., Faith J., Branco N., Bull S. B., McLeod R. S., Griffiths A. M., Bitton A., Greenberg G. R., Lander E. S., Siminovitch K. A.,

- and Hudson T. J. (2001). Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. *Nat Genet* **29**: 223-8.
- Risch N., and Merikangas K. (1996). The future of genetic studies of complex human diseases. *Science* **273**: 1516-7.
- Risch N. J. (2000). Searching for genetic determinants in the new millennium. *Nature* **405**: 847-56.
- Rogic S., Mackworth A. K., and Ouellette F. B. (2001). Evaluation of gene-finding programs on mammalian sequences. *Genome Res* **11**: 817-32.
- Rogozin I. B., Mayorov V. I., Lavrentieva M. V., Milanese L., and Adkison L. R. (2000). Prediction and phylogenetic analysis of mammalian short interspersed elements (SINEs). *Brief Bioinform* **1**: 260-74.
- Rommens J. M., Iannuzzi M. C., Kerem B., Drumm M. L., Melmer G., Dean M., Rozmahel R., Cole J. L., Kennedy D., Hidaka N., and *et al.* (1989). Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science* **245**: 1059-65.
- Roses A. D. (2000). Pharmacogenetics and the practice of medicine. *Nature* **405**: 857-65.
- Ross P., Hall L., Smirnov I., and Haff L. (1998). High level multiplex genotyping by MALDI-TOF mass spectrometry. *Nat Biotechnol* **16**: 1347-51.
- Ross-MacDonald P., Coelho P. S., Roemer T., Agarwal S., Kumar A., Jansen R., Cheung K. H., Sheehan A., Symoniatis D., Umansky L., Heidtman M., Nelson F. K., Iwasaki H., Hager K., Gerstein M., Miller P., Roeder G. S., and Snyder M. (1999). Large-scale analysis of the yeast genome by transposon tagging and gene disruption. *Nature* **402**: 413-8.
- Rothstein R. (1991). Targeting, disruption, replacement, and allele rescue: integrative DNA transformation in yeast. *Methods Enzymol* **194**: 281-301.
- Rozen S., and Skaletsky H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* **132**: 365-86.
- Rubin G. M. (1996). Around the genomes: the Drosophila genome project. *Genome Res* **6**: 71-9.
- Rushforth A. M., Saari B., and Anderson P. (1993). Site-selected insertion of the transposon Tc1 into a Caenorhabditis elegans myosin light chain gene. *Mol Cell Biol* **13**: 902-10.
- Saccone S., Caccio S., Kusuda J., Andreozzi L., and Bernardi G. (1996). Identification of the gene-richest bands in human chromosomes. *Gene* **174**: 85-94.
- Saccone S., De Sario A., Wiegant J., Raap A. K., Della Valle G., and Bernardi G. (1993). Correlations between isochores and chromosomal bands in the human genome. *Proc Natl Acad Sci U S A* **90**: 11929-33.
- Saiki R. K., Gelfand D. H., Stoffel S., Scharf S. J., Higuchi R., Horn G. T., Mullis K. B., and Erlich H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**: 487-91.

- Saiki R. K., Scharf S., Faloona F., Mullis K. B., Horn G. T., Erlich H. A., and Arnheim N. (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**: 1350-4.
- Salamov A. A., and Solovyev V. V. (2000). Ab initio gene finding in Drosophila genomic DNA. *Genome Res* **10**: 516-22.
- Sanger F., Nicklen S., and Coulson A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**: 5463-7.
- Sassaman D. M., Dombroski B. A., Moran J. V., Kimberland M. L., Naas T. P., DeBerardinis R. J., Gabriel A., Swergold G. D., and Kazazian H. H., Jr. (1997). Many human L1 elements are capable of retrotransposition. *Nat Genet* **16**: 37-43.
- Sauer S., Lechner D., Berlin K., Lehrach H., Escary J. L., Fox N., and Gut I. G. (2000). A novel procedure for efficient genotyping of single nucleotide polymorphisms. *Nucleic Acids Res* **28**: E13.
- Schalkwyk L. C., Cusack B., Dunkel I., Hopp M., Kramer M., Palczewski S., Piefke J., Scheel S., Weiher M., Wenske G., Lehrach H., and Himmelbauer H. (2001). Advanced integrated mouse YAC map including BAC framework. *Genome Res* **11**: 2142-50.
- Scheetz T. E., Raymond M. R., Nishimura D. Y., McClain A., Roberts C., Birkett C., Gardiner J., Zhang J., Butters N., Sun C., Kwitek-Black A., Jacob H., Casavant T. L., Soares M. B., and Sheffield V. C. (2001). Generation of a high-density rat EST map. *Genome Res* **11**: 497-502.
- Scherf M., Klingenhoff A., Frech K., Quandt K., Schneider R., Grote K., Frisch M., Gailus-Durner V., Seidel A., Brack-Werner R., and Werner T. (2001). First pass annotation of promoters on human chromosome 22. *Genome Res* **11**: 333-40.
- Scherf M., Klingenhoff A., and Werner T. (2000). Highly specific localization of promoter regions in large genomic sequences by PromoterInspector: a novel context analysis approach. *J Mol Biol* **297**: 599-606.
- Schmid A., Schindelholz B., and Zinn K. (2002). Combinatorial RNAi: a method for evaluating the functions of gene families in Drosophila. *Trends Neurosci* **25**: 71-4.
- Schuler G. D., Boguski M. S., Stewart E. A., Stein L. D., Gyapay G., Rice K., White R. E., Rodriguez-Tome P., Aggarwal A., Bajorek E., Bentolila S., Birren B. B., Butler A., Castle A. B., Chiannikulchai N., Chu A., Clee C., Cowles S., Day P. J., Dibling T., Drouot N., Dunham I., Duprat S., East C., Hudson T. J., and *et al.* (1996). A gene map of the human genome. *Science* **274**: 540-6.
- Schwartz S., Zhang Z., Frazer K. A., Smit A., Riemer C., Bouck J., Gibbs R., Hardison R., and Miller W. (2000). PipMaker-a web server for aligning two genomic DNA sequences. *Genome Res* **10**: 577-86.
- Sharp P. A., and Burge C. B. (1997). Classification of introns: U2-type or U12-type. *Cell* **91**: 875-9.

- Shehee W. R., Loeb D. D., Adey N. B., Burton F. H., Casavant N. C., Cole P., Davies C. J., McGraw R. A., Schichman S. A., Severynse D. M., and *et al.* (1989). Nucleotide sequence of the BALB/c mouse beta-globin complex. *J Mol Biol* **205**: 41-62.
- Shen L. X., Basilion J. P., and Stanton V. P., Jr. (1999). Single-nucleotide polymorphisms can cause different structural folds of mRNA. *Proc Natl Acad Sci U S A* **96**: 7871-6.
- Sherry S. T., Ward M. H., Kholodov M., Baker J., Phan L., Smigielski E. M., and Sirotkin K. (2001). dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* **29**: 308-11.
- Shimoda N., Knapik E. W., Ziniti J., Sim C., Yamada E., Kaplan S., Jackson D., de Sauvage F., Jacob H., and Fishman M. C. (1999). Zebrafish genetic map with 2000 microsatellite markers. *Genomics* **58**: 219-32.
- Shizuya H., Birren B., Kim U. J., Mancino V., Slepak T., Tachiiri Y., and Simon M. (1992). Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc Natl Acad Sci U S A* **89**: 8794-7.
- Shoemaker D. D., Lashkari D. A., Morris D., Mittmann M., and Davis R. W. (1996). Quantitative phenotypic analysis of yeast deletion mutants using a highly parallel molecular bar-coding strategy. *Nat Genet* **14**: 450-6.
- Shoemaker D. D., Schadt E. E., Armour C. D., He Y. D., Garrett-Engle P., McDonagh P. D., Loerch P. M., Leonardson A., Lum P. Y., Cavet G., Wu L. F., Altschuler S. J., Edwards S., King J., Tsang J. S., Schimmack G., Schelter J. M., Koch J., Ziman M., Marton M. J., Li B., Cundiff P., Ward T., Castle J., Krolewski M., Meyer M. R., Mao M., Burchard J., Kidd M. J., Dai H., Phillips J. W., Linsley P. S., Stoughton R., Scherer S., and Boguski M. S. (2001). Experimental annotation of the human genome using microarray technology. *Nature* **409**: 922-7.
- Shyamala V., and Ames G. F. (1989). Genome walking by single-specific-primer polymerase chain reaction: SSP-PCR. *Gene* **84**: 1-8.
- Shyamala V., and Ames G. F. (1993). Genome walking by single specific primer-polymerase chain reaction. *Methods Enzymol* **217**: 436-46.
- Sidow A., and Thomas W. K. (1994). A molecular evolutionary framework for eukaryotic model organisms. *Curr Biol* **4**: 596-603.
- Simpson A. J. G., de Souza S. J., Camargo A. A., and Brentani R. R. (2001). Definition of the gene content of the human genome: the need for deep experimental verification. *Comparative and Functional Genomics* **2**: 169-175.
- Smink L. (2000). Genome studies of human chromosome 22q13.3. PhD thesis.
- Smit A. F. (1996). The origin of interspersed repeats in the human genome. *Curr Opin Genet Dev* **6**: 743-8.
- Smit A. F. (1999). Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Curr Opin Genet Dev* **9**: 657-63.
- Smit A. F., and Riggs A. D. (1995). MIRs are classic, tRNA-derived SINEs that amplified before the mammalian radiation. *Nucleic Acids Res* **23**: 98-102.

- Smit A. F., and Riggs A. D. (1996). Tiggers and DNA transposon fossils in the human genome. *Proc Natl Acad Sci U S A* **93**: 1443-8.
- Smith S. F., Snell P., Gruetzner F., Bench A. J., Haaf T., Metcalfe J. A., Green A. R., and Elgar G. (2002). Analyses of the extent of shared synteny and conserved gene orders between the genome of *Fugu rubripes* and human 20q. *Genome Res* **12**: 776-84.
- Smith V., Chou K. N., Lashkari D., Botstein D., and Brown P. O. (1996). Functional analysis of the genes of yeast chromosome V by genetic footprinting. *Science* **274**: 2069-74.
- Soderlund C., Humphray S., Dunham A., and French L. (2000). Contigs built with fingerprints, markers, and FPC V4.7. *Genome Res* **10**: 1772-87.
- Solovyev V. V., Salamov A. A., and Lawrence C. B. (1994). Predicting internal exons by oligonucleotide composition and discriminant analysis of spliceable open reading frames. *Nucleic Acids Res* **22**: 5156-63.
- Sonnhammer E. L., and Durbin R. (1994). A workbench for large-scale sequence homology analysis. *Comput Appl Biosci* **10**: 301-7.
- Sonnhammer E. L., and Durbin R. (1995). A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. *Gene* **167**: GC1-10.
- Spradling A. C., Stern D. M., Kiss I., Roote J., Lavery T., and Rubin G. M. (1995). Gene disruptions using P transposable elements: an integral component of the *Drosophila* genome project. *Proc Natl Acad Sci U S A* **92**: 10824-30.
- Stemers F. J., Ferguson J. A., and Walt D. R. (2000). Screening unlabeled DNA targets with randomly ordered fiber-optic gene arrays. *Nat Biotechnol* **18**: 91-4.
- Stoesser G., Baker W., van den Broek A., Camon E., Garcia-Pastor M., Kanz C., Kulikova T., Leinonen R., Lin Q., Lombard V., Lopez R., Redaschi N., Stoehr P., Tuli M. A., Tzouvara K., and Vaughan R. (2002). The EMBL Nucleotide Sequence Database. *Nucleic Acids Res* **30**: 21-6.
- Stone D. L., Slavotinek A., Bouffard G. G., Banerjee-Basu S., Baxevanis A. D., Barr M., and Biesecker L. G. (2000). Mutation of a gene encoding a putative chaperonin causes McKusick-Kaufman syndrome. *Nat Genet* **25**: 79-82.
- Strausberg R. L., Feingold E. A., Klausner R. D., and Collins F. S. (1999). The mammalian gene collection. *Science* **286**: 455-7.
- Stringer C. B., and Andrews P. (1988). Genetic and fossil evidence for the origin of modern humans. *Science* **239**: 1263-8.
- Stumpf M. P. (2002). Haplotype diversity and the block structure of linkage disequilibrium. *Trends Genet* **18**: 226-8.
- Sturtevant J. (2000). Applications of differential-display reverse transcription-PCR to molecular pathogenesis and medical mycology. *Clin Microbiol Rev* **13**: 408-27.
- Sulston J. E., Mallet F., Staden R., Durbin R., Horsnell T., and Coulson A. (1980). Software for genome mapping by fingerprinting techniques. *CABIOS* **4**: 125-132.

- Summers T. J., Thomas J. W., Lee-Lin S. Q., Maduro V. V., Idol J. R., and Green E. D. (2001). Comparative physical mapping of targeted regions of the rat genome. *Mamm Genome* **12**: 508-12.
- Sun X., Ding H., Hung K., and Guo B. (2000). A new MALDI-TOF based mini-sequencing assay for genotyping of SNPS. *Nucleic Acids Res* **28**: E68.
- Sunyaev S., Ramensky V., Koch I., Lathe W., 3rd, Kondrashov A. S., and Bork P. (2001). Prediction of deleterious human alleles. *Hum Mol Genet* **10**: 591-7.
- Syvänen A. C. (2001). Accessing genetic variation: genotyping single nucleotide polymorphisms. *Nat Rev Genet* **2**: 930-42.
- Taillon-Miller P., Gu Z., Li Q., Hillier L., and Kwok P. Y. (1998). Overlapping genomic sequences: a treasure trove of single-nucleotide polymorphisms. *Genome Res* **8**: 748-54.
- Talbot W. S., and Hopkins N. (2000). Zebrafish mutations and functional analysis of the vertebrate genome. *Genes Dev* **14**: 755-62.
- Tang K., Fu D. J., Julien D., Braun A., Cantor C. R., and Koster H. (1999). Chip-based genotyping by mass spectrometry. *Proc Natl Acad Sci U S A* **96**: 10016-20.
- Tateno Y., Imanishi T., Miyazaki S., Fukami-Kobayashi K., Saitou N., Sugawara H., and Gojobori T. (2002). DNA Data Bank of Japan (DDBJ) for genome scale research in life science. *Nucleic Acids Res* **30**: 27-30.
- Tautz D. (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res* **17**: 6463-71.
- Taylor J. G., Choi E. H., Foster C. B., and Chanock S. J. (2001). Using genetic variation to study human disease. *Trends Mol Med* **7**: 507-12.
- The *C. elegans* Sequencing Consortium (1998). Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**: 2012-8.
- The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium: Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L. M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M. F., Brownstein M. J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D. A., Kamiya M., Lee N. H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schonbach C., Seya T., Shibata Y., Storch K. F., Suzuki H., Toyooka K., Wang K. H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., and Hayashizaki Y. (2001). Functional annotation of a full-length mouse cDNA collection. *Nature* **409**: 685-90.

- The Sanger Institute and Washington University Genome Sequencing Center (1998). Toward a complete human genome sequence. *Genome Res* **8**: 1097-108.
- Thiery J. P., Macaya G., and Bernardi G. (1976). An analysis of eukaryotic genomes by density gradient centrifugation. *J Mol Biol* **108**: 219-35.
- Thomas A., and Skolnick M. H. (1994). A probabilistic model for detecting coding regions in DNA sequences. *IMA J Math Appl Med Biol* **11**: 149-60.
- Thomas J. W., Summers T. J., Lee-Lin S. Q., Maduro V. V., Idol J. R., Mastrian S. D., Ryan J. F., Jamison D. C., and Green E. D. (2000). Comparative genome mapping in the sequence-based era: early experience with human chromosome 7. *Genome Res* **10**: 624-33.
- Thompson J. D., Higgins D. G., and Gibson T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673-80.
- Thornton J. M. (2001). From genome to function. *Science* **292**: 2095-7.
- Tilford C. A., Kuroda-Kawaguchi T., Skaletsky H., Rozen S., Brown L. G., Rosenberg M., McPherson J. D., Wylie K., Sekhon M., Kucaba T. A., Waterston R. H., and Page D. C. (2001). A physical map of the human Y chromosome. *Nature* **409**: 943-5.
- Tilghman S. M., Curtis P. J., Tiemeier D. C., Leder P., and Weissmann C. (1978a). The intervening sequence of a mouse beta-globin gene is transcribed within the 15S beta-globin mRNA precursor. *Proc Natl Acad Sci U S A* **75**: 1309-13.
- Tilghman S. M., Tiemeier D. C., Seidman J. G., Peterlin B. M., Sullivan M., Maizel J. V., and Leder P. (1978b). Intervening sequence of DNA identified in the structural portion of a mouse beta-globin gene. *Proc Natl Acad Sci U S A* **75**: 725-9.
- Tomer Y., Barbesino G., Greenberg D. A., Concepcion E., and Davies T. F. (1998). A new Graves disease-susceptibility locus maps to chromosome 20q11.2. International Consortium for the Genetics of Autoimmune Thyroid Disease. *Am J Hum Genet* **63**: 1749-56.
- Tomer Y., Barbesino G., Greenberg D. A., Concepcion E., and Davies T. F. (1999). Mapping the major susceptibility loci for familial Graves' and Hashimoto's diseases: evidence for genetic heterogeneity and gene interactions. *J Clin Endocrinol Metab* **84**: 4656-64.
- Tomer Y., and Davies T. F. (1997). The genetic susceptibility to Graves' disease. *Baillieres Clin Endocrinol Metab* **11**: 431-50.
- Touchman J. W., Bouffard G. G., Weintraub L. A., Idol J. R., Wang L., Robbins C. M., Nussbaum J. C., Lovett M., and Green E. D. (1997). 2006 expressed-sequence tags derived from human chromosome 7-enriched cDNA libraries. *Genome Res* **7**: 281-92.
- Touchman J. W., Dehejia A., Chiba-Falek O., Cabin D. E., Schwartz J. R., Orrison B. M., Polymeropoulos M. H., and Nussbaum R. L. (2001). Human and mouse alpha-synuclein genes: comparative genomic sequence analysis and identification of a novel gene regulatory element. *Genome Res* **11**: 78-86.

- Trofatter J. A., Long K. R., Murrell J. R., Stotler C. J., Gusella J. F., and Buckler A. J. (1995). An expression-independent catalog of genes from human chromosome 22. *Genome Res* **5**: 214-24.
- Trofatter J. A., MacCollin M. M., Rutter J. L., Murrell J. R., Duyao M. P., Parry D. M., Eldridge R., Kley N., Menon A. G., Pulaski K., and *et al.* (1993). A novel moesin-, ezrin-, radixin-like gene is a candidate for the neurofibromatosis 2 tumor suppressor. *Cell* **72**: 791-800.
- Tyagi S., Bratu D. P., and Kramer F. R. (1998). Multicolor molecular beacons for allele discrimination. *Nat Biotechnol* **16**: 49-53.
- Tyagi S., and Kramer F. R. (1996). Molecular beacons: probes that fluoresce upon hybridization. *Nat Biotechnol* **14**: 303-8.
- Uberbacher E. C., Xu Y., and Mural R. J. (1996). Discovering and understanding genes in human DNA sequence using GRAIL. *Methods Enzymol* **266**: 259-81.
- Van Etten W. J., Steen R. G., Nguyen H., Castle A. B., Slonim D. K., Ge B., Nusbaum C., Schuler G. D., Lander E. S., and Hudson T. J. (1999). Radiation hybrid map of the mouse genome. *Nat Genet* **22**: 384-7.
- Vanin E. F. (1985). Processed pseudogenes: characteristics and evolution. *Annu Rev Genet* **19**: 253-72.
- Velculescu V. E., Zhang L., Vogelstein B., and Kinzler K. W. (1995). Serial analysis of gene expression. *Science* **270**: 484-7.
- Velculescu V. E., Zhang L., Zhou W., Vogelstein J., Basrai M. A., Bassett D. E., Jr., Hieter P., Vogelstein B., and Kinzler K. W. (1997). Characterization of the yeast transcriptome. *Cell* **88**: 243-51.
- Venter J. C., Adams M. D., Myers E. W., Li P. W., Mural R. J., Sutton G. G., Smith H. O., Yandell M., Evans C. A., Holt R. A., Gocayne J. D., Amanatides P., Ballew R. M., Huson D. H., Wortman J. R., Zhang Q., Kodira C. D., Zheng X. H., Chen L., Skupski M., Subramanian G., Thomas P. D., Zhang J., Gabor Miklos G. L., Nelson C., Broder S., Clark A. G., Nadeau J., McKusick V. A., Zinder N., Levine A. J., Roberts R. J., Simon M., Slayman C., Hunkapiller M., Bolanos R., Delcher A., Dew I., Fasulo D., Flanigan M., Florea L., Halpern A., Hannenhalli S., Kravitz S., Levy S., Mobarry C., Reinert K., Remington K., Abu-Threideh J., Beasley E., Biddick K., Bonazzi V., Brandon R., Cargill M., Chandramouliswaran I., Charlab R., Chaturvedi K., Deng Z., Di Francesco V., Dunn P., Eilbeck K., Evangelista C., Gabrielian A. E., Gan W., Ge W., Gong F., Gu Z., Guan P., Heiman T. J., Higgins M. E., Ji R. R., Ke Z., Ketchum K. A., Lai Z., Lei Y., Li Z., Li J., Liang Y., Lin X., Lu F., Merkulov G. V., Milshina N., Moore H. M., Naik A. K., Narayan V. A., Neelam B., Nusskern D., Rusch D. B., Salzberg S., Shao W., Shue B., Sun J., Wang Z., Wang A., Wang X., Wang J., Wei M., Wides R., Xiao C., Yan C., *et al.* (2001). The sequence of the human genome. *Science* **291**: 1304-51.
- Wagener R., Kobbe B., and Paulsson M. (1998). Genomic organisation, alternative splicing and primary structure of human matrillin-4. *FEBS Lett* **438**: 165-70.
- Wall J. D. (2001). Insights from linked single nucleotide polymorphisms: what we can learn from linkage disequilibrium. *Curr Opin Genet Dev* **11**: 647-51.

- Wall J. D., and Przeworski M. (2000). When did the human population size start increasing? *Genetics* **155**: 1865-74.
- Walt D. R. (2000). Techview: molecular biology. Bead-based fiber-optic arrays. *Science* **287**: 451-2.
- Walter M. A., Spillett D. J., Thomas P., Weissenbach J., and Goodfellow P. N. (1994). A method for constructing radiation hybrid maps of whole genomes. *Nat Genet* **7**: 22-8.
- Wang D. G., Fan J. B., Siao C. J., Berno A., Young P., Sapolsky R., Ghandour G., Perkins N., Winchester E., Spencer J., Kruglyak L., Stein L., Hsie L., Topaloglou T., Hubbell E., Robinson E., Mittmann M., Morris M. S., Shen N., Kilburn D., Rioux J., Nusbaum C., Rozen S., Hudson T. J., Lander E. S., and *et al.* (1998a). Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* **280**: 1077-82.
- Wang P. W., Eisenbart J. D., Espinosa R., 3rd, Davis E. M., Larson R. A., and Le Beau M. M. (2000). Refinement of the smallest commonly deleted segment of chromosome 20 in malignant myeloid diseases and development of a PAC-based physical and transcription map. *Genomics* **67**: 28-39.
- Wang P. W., Iannantuoni K., Davis E. M., Espinosa R., 3rd, Stoffel M., and Le Beau M. M. (1998b). Refinement of the commonly deleted segment in myeloid leukemias with a del(20q). *Genes Chromosomes Cancer* **21**: 75-81.
- Wang Z., and Moulton J. (2001). SNPs, protein structure, and disease. *Hum Mutat* **17**: 263-70.
- Wasserman W. W., Palumbo M., Thompson W., Fickett J. W., and Lawrence C. E. (2000). Human-mouse genome comparisons to locate regulatory sites. *Nat Genet* **26**: 225-8.
- Watanabe T. K., Bihoreau M. T., McCarthy L. C., Kiguwa S. L., Hishigaki H., Tsuji A., Browne J., Yamasaki Y., Mizoguchi-Miyakita A., Oga K., Ono T., Okuno S., Kanemoto N., Takahashi E., Tomita K., Hayashi H., Adachi M., Webber C., Davis M., Kiel S., Knights C., Smith A., Critcher R., Miller J., James M. R., and *et al.* (1999). A radiation hybrid map of the rat genome containing 5,255 markers. *Nat Genet* **22**: 27-36.
- Waterston R., and Sulston J. E. (1998). The Human Genome Project: reaching the finish line. *Science* **282**: 53-4.
- Waterston R. H., Lander E. S., and Sulston J. E. (2002). On the sequencing of the human genome. *Proc Natl Acad Sci U S A* **99**: 3712-6.
- Watson J. D. (1990). The human genome project: past, present, and future. *Science* **248**: 44-9.
- Weber J. L. (1990). Human DNA polymorphisms and methods of analysis. *Curr Opin Biotechnol* **1**: 166-71.
- Weber J. L., and May P. E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet* **44**: 388-96.
- Weiss K. M., and Clark A. G. (2002). Linkage disequilibrium and the mapping of complex human traits. *Trends Genet* **18**: 19-24.

- Weissenbach J., Gyapay G., Dib C., Vignal A., Morissette J., Millasseau P., Vaysseix G., and Lathrop M. (1992). A second-generation linkage map of the human genome. *Nature* **359**: 794-801.
- Wenderfer S. E., Slack J. P., McCluskey T. S., and Monaco J. J. (2000). Identification of 40 genes on a 1-Mb contig around the IL-4 cytokine family gene cluster on mouse chromosome 11. *Genomics* **63**: 354-73.
- Wheelan S. J., Boguski M. S., Duret L., and Makalowski W. (1999). Human and nematode orthologs-lessons from the analysis of 1800 human genes and the proteome of *Caenorhabditis elegans*. *Gene* **238**: 163-70.
- Wheeler D. L., Church D. M., Lash A. E., Leipe D. D., Madden T. L., Pontius J. U., Schuler G. D., Schriml L. M., Tatusova T. A., Wagner L., and Rapp B. A. (2002). Database resources of the National Center for Biotechnology Information: 2002 update. *Nucleic Acids Res* **30**: 13-6.
- White K. P. (2001). Functional genomics and the study of development, variation and evolution. *Nat Rev Genet* **2**: 528-37.
- Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S., Ansorge W., Bocher M., Blocker H., Bauersachs S., Blum H., Lauber J., Dusterhoft A., Beyer A., Kohrer K., Strack N., Mewes H. W., Ottenwalder B., Obermaier B., Tampe J., Heubner D., Wambutt R., Korn B., Klein M., and Poustka A. (2001). Toward a catalog of human genes and proteins: sequencing and analysis of 500 novel complete protein coding human cDNAs. *Genome Res* **11**: 422-35.
- Wiginton D. A., Kaplan D. J., States J. C., Akeson A. L., Perme C. M., Bilyk I. J., Vaughn A. J., Lattier D. L., and Hutton J. J. (1986). Complete sequence and structure of the gene for human adenosine deaminase. *Biochemistry* **25**: 8234-44.
- Wilcox A. S., Khan A. S., Hopkins J. A., and Sikela J. M. (1991). Use of 3' untranslated sequences of human cDNAs for rapid chromosome assignment and conversion to STSS: implications for an expression map of the genome. *Nucleic Acids Res* **19**: 1837-43.
- Wilson M. D., Riemer C., Martindale D. W., Schnupf P., Boright A. P., Cheung T. L., Hardy D. M., Schwartz S., Scherer S. W., Tsui L. C., Miller W., and Koop B. F. (2001). Comparative analysis of the gene-dense ACHE/TFR2 region on human chromosome 7q22 with the orthologous region on mouse chromosome 5. *Nucleic Acids Res* **29**: 1352-65.
- Woods I. G., Kelly P. D., Chu F., Ngo-Hazelett P., Yan Y. L., Huang H., Postlethwait J. H., and Talbot W. S. (2000). A comparative map of the zebrafish genome. *Genome Res* **10**: 1903-14.
- Yaspo M. L. (2001). Taking a functional genomics approach in molecular medicine. *Trends Mol Med* **7**: 494-501.
- Yu A., Zhao C., Fan Y., Jang W., Mungall A. J., Deloukas P., Olsen A., Doggett N. A., Ghebranious N., Broman K. W., and Weber J. L. (2001). Comparison of human genetic and sequence-based physical maps. *Nature* **409**: 951-3.

- Zambrowicz B. P., Friedrich G. A., Buxton E. C., Lilleberg S. L., Person C., and Sands A. T. (1998). Disruption and sequence identification of 2,000 genes in mouse embryonic stem cells. *Nature* **392**: 608-11.
- Zamore P. D. (2001). RNA interference: listening to the sound of silence. *Nat Struct Biol* **8**: 746-50.
- Zdobnov E. M., and Apweiler R. (2001). InterProScan--an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* **17**: 847-8.
- Zdobnov E. M., Lopez R., Apweiler R., and Etzold T. (2002). The EBI SRS server--recent developments. *Bioinformatics* **18**: 368-73.
- Zhao S., Shatsman S., Ayodeji B., Geer K., Tsegaye G., Krol M., Gebregeorgis E., Shvartsbeyn A., Russell D., Overton L., Jiang L., Dimitrov G., Tran K., Shetty J., Malek J. A., Feldblyum T., Nierman W. C., and Fraser C. M. (2001). Mouse BAC ends quality assessment and sequence analyses. *Genome Res* **11**: 1736-45.
- Zhou B., Westaway S. K., Levinson B., Johnson M. A., Gitschier J., and Hayflick S. J. (2001). A novel pantothenate kinase gene (PANK2) is defective in Hallervorden-Spatz syndrome. *Nat Genet* **28**: 345-9.
- Zouali H., Hani E. H., Philippi A., Vionnet N., Beckmann J. S., Demenais F., and Froguel P. (1997). A susceptibility locus for early-onset non-insulin dependent (type 2) diabetes mellitus maps to chromosome 20q, proximal to the phosphoenolpyruvate carboxykinase gene. *Hum Mol Genet* **6**: 1401-8.
- Zwaal R. R., Broeks A., van Meurs J., Groenen J. T., and Plasterk R. H. (1993). Target-selected gene inactivation in *Caenorhabditis elegans* by using a frozen transposon insertion mutant bank. *Proc Natl Acad Sci U S A* **90**: 7431-5.