Chapter VII References

Aach J., Bulyk M. L., Church G. M., Comander J., Derti A., and Shendure J. (2001). Computational comparison of two draft sequences of the human genome. *Nature* **409:** 856-9.

Achaz G., Netter P., and Coissac E. (2001). Study of intrachromosomal duplications among the eukaryote genomes. *Mol Biol Evol* **18:** 2280-8.

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.

Adams M. D., Soares M. B., Kerlavage A. R., Fields C., and Venter J. C. (1993). Rapid cDNA sequencing (expressed sequence tags) from a directionally cloned human infant brain cDNA library. *Nat Genet* **4:** 373-80.

Aissani B., and Bernardi G. (1991). CpG islands, genes and isochores in the genomes of vertebrates. *Gene* **106:** 185-95.

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.

Amadou C., Ribouchon M. T., Mattei M. G., Jenkins N. A., Gilbert D. J., Copeland N. G., Avoustin P., and Pontarotti P. (1995). Localization of new genes and markers to the distal part of the human major histocompatibility complex (MHC) region and comparison with the mouse: new insights into the evolution of mammalian genomes. *Genomics* **26:** 9-20.

Anderson L., and Seilhamer J. (1997). A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis* **18:** 533-7.

Anderson S., Bankier A. T., Barrell B. G., de Bruijn M. H., Coulson A. R., Drouin J., Eperon I. C., Nierlich D. P., Roe B. A., Sanger F., Schreier P. H., Smith A. J., Staden R., and Young I. G. (1981). Sequence and organization of the human mitochondrial genome. *Nature* **290:** 457-65.

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.

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.

Argraves W. S., Tran H., Burgess W. H., and Dickerson K. (1990). Fibulin is an extracellular matrix and plasma glycoprotein with repeated domain structure. *J Cell Biol* **111:** 3155-64.

Ashburner M., Misra S., Roote J., Lewis S. E., Blazej R., Davis T., Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D., Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A., Palazzolo M., Reese M. G., Spradling A., Tsang G., Wan K., Whitelaw K., Celniker S., and et al. (1999). An exploration of the sequence of a 2.9-Mb region of the genome of Drosophila melanogaster: the Adh region. *Genetics* **153:** 179-219.

Attwood T. K., Blythe M. J., Flower D. R., Gaulton A., Mabey J. E., Maudling N., McGregor L., Mitchell A. L., Moulton G., Paine K., and Scordis P. (2002). PRINTS and PRINTS-S shed light on protein ancestry. *Nucleic Acids Res* **30:** 239-41.

Audic S., and Claverie J. M. (1998). Visualizing the competitive recognition of TATA-boxes in vertebrate promoters. *Trends Genet* **14:** 10-1.

Bairoch A., and Apweiler R. (2000). The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res* **28:** 45-8.

Baker W., van den Broek A., Camon E., Hingamp P., Sterk P., Stoesser G., and Tuli M. A. (2000). The EMBL nucleotide sequence database. *Nucleic Acids Res* **28:** 19-23.

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.

Bargmann C. I. (2001). High-throughput reverse genetics: RNAi screens in Caenorhabditis elegans. *Genome Biol* **2**.

Bateman A., Birney E., Durbin R., Eddy S. R., Finn R. D., and Sonnhammer E. L. (1999). Pfam 3.1: 1313 multiple alignments and profile HMMs match the majority of proteins. *Nucleic Acids Res* **27:** 260-2.

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.

Baulande S., Lasnier F., Lucas M., and Pairault J. (2001). Adiponutrin, a transmembrane protein corresponding to a novel dietary- and obesity-linked mRNA specifically expressed in the adipose lineage. *J Biol Chem* **276:** 33336-44.

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.

Bell C. J., Budarf M. L., Nieuwenhuijsen B. W., Barnoski B. L., Buetow K. H., Campbell K., Colbert A. M., Collins J., Daly M., Desjardins P. R., and et al. (1995). Integration of physical, breakpoint and genetic maps of chromosome 22. Localization of 587 yeast artificial chromosomes with 238 mapped markers. *Hum Mol Genet* **4:** 59-69.

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., Cobley 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., Dhami 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. (1993). The isochore organization of the human genome and its evolutionary history--a review. *Gene* **135:** 57-66.

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.

Bichara M., Pinet I., Schumacher S., and Fuchs R. P. (2000). Mechanisms of dinucleotide repeat instability in Escherichia coli. *Genetics* **154:** 533-42.

Bichara M., Schumacher S., and Fuchs R. P. (1995). Genetic instability within monotonous runs of CpG sequences in Escherichia coli. *Genetics* **140:** 897-907.

Bird A., Taggart M., Frommer M., Miller O. J., and Macleod D. (1985). A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. *Cell* **40:** 91-9.

Bird A., Tate P., Nan X., Campoy J., Meehan R., Cross S., Tweedie S., Charlton J., and Macleod D. (1995). Studies of DNA methylation in animals. *J Cell Sci Suppl* **19:** 37-9.

Bird A. P. (1986). CpG-rich islands and the function of DNA methylation. *Nature* **321:** 209-13.

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.

Blechschmidt 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.

Boguski M. S., Lowe T. M., and Tolstoshev C. M. (1993). dbEST--database for "expressed sequence tags". *Nat Genet* **4:** 332-3.

Boguski M. S., and Schuler G. D. (1995). ESTablishing a human transcript map. *Nat Genet* **10:** 369-71.

Bonfield J. K., Smith K., and Staden R. (1995). A new DNA sequence assembly program. *Nucleic Acids Res* **23:** 4992-9.

Bork P., and Koonin E. V. (1998). Predicting functions from protein sequences--where are the bottlenecks? *Nat Genet* **18:** 313-8.

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.

Bouffard G. G., Idol J. R., Braden V. V., Iyer L. M., Cunningham A. F., Weintraub L. A., Touchman J. W., Mohr-Tidwell R. M., Peluso D. C., Fulton R. S., Ueltzen M. S., Weissenbach J., Magness C. L., and Green E. D. (1997). A physical map of human chromosome 7: an integrated YAC contig map with average STS spacing of 79 kb. *Genome Res* **7:** 673-92.

Boyd J. M., Gallo G. J., Elangovan B., Houghton A. B., Malstrom S., Avery B. J., Ebb R. G., Subramanian T., Chittenden T., Lutz R. J., and et al. (1995). Bik, a novel death-inducing protein shares a distinct sequence motif with Bcl-2 family proteins and interacts with viral and cellular survival-promoting proteins. *Oncogene* **11:** 1921-8.

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.

Brocchieri L. (2001). Phylogenetic inferences from molecular sequences: review and critique. *Theor Popul Biol* **59:** 27-40.

Brosius J. (1999). Genomes were forged by massive bombardments with retroelements and retrosequences. *Genetica* **107:** 209-38.

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.

Burge C. B., and Karlin S. (1998). Finding the genes in genomic DNA. *Curr Opin Struct Biol* **8:** 346-54.

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.

Burt D. W., Bruley C., Dunn I. C., Jones C. T., Ramage A., Law A. S., Morrice D. R., Paton I. R., Smith J., Windsor D., Sazanov A., Fries R., and Waddington D. (1999). The dynamics of chromosome evolution in birds and mammals. *Nature* **402:** 411-3.

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.

Casalotti S. O., Pelaia G., Yakovlev A. G., Csikos T., Grayson D. R., and Krueger K. E. (1992). Structure of the rat gene encoding the mitochondrial benzodiazepine receptor. *Gene* **121:** 377-82.

Castells A., Ino Y., Louis D. N., Ramesh V., Gusella J. F., and Rustgi A. K. (1999). Mapping of a target region of allelic loss to a 0.5-cM interval on chromosome 22q13 in human colorectal cancer. *Gastroenterology* **117:** 831-7.

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.

Chambers D. M., Rouleau G. A., and Abbott C. M. (2001). Comparative genomic analysis of genes encoding translation elongation factor 1B(alpha) in human and mouse shows EEF1B1 to be a recent retrotransposition event. *Genomics* **77:** 145-8.

Chan M. F., Liang G., and Jones P. A. (2000). Relationship between transcription and DNA methylation. *Curr Top Microbiol Immunol* **249:** 75-86.

Chang Y. J., McCabe R. T., Rennert H., Budarf M. L., Sayegh R., Emanuel B. S., Skolnick P., and Strauss J. F., 3rd (1992). The human "peripheral-type" benzodiazepine receptor: regional mapping of the gene and characterization of the receptor expressed from cDNA. *DNA Cell Biol* **11:** 471-80.

Cheung V. G., Nowak N., Jang W., Kirsch I. R., Zhao S., Chen X. N., Furey T. S., Kim U. J., Kuo W. L., Olivier M., Conroy J., Kasprzyk A., Massa H., Yonescu R., Sait S., Thoreen C., Snijders A., Lemyre E., Bailey J. A., Bruzel A., Burrill W. D., Clegg S. M., Collins S., Dhami P., Friedman C., Han C. S., Herrick S., Lee J., Ligon A. H., Lowry S., Morley M., Narasimhan S., Osoegawa K., Peng Z., Plajzer-Frick I., Quade B. J., Scott D., Sirotkin K., Thorpe A. A., Gray J. W., Hudson J., Pinkel D., Ried T., Rowen L., Shen-Ong G. L., Strausberg R. L., Birney E., Callen D. F., Cheng J. F., Cox D. R., Doggett N. A., Carter N. P., Eichler E. E., Haussler D., Korenberg J. R., Morton C. C., Albertson D., Schuler G., de Jong P. J., and Trask B. J. (2001). Integration of cytogenetic landmarks into the draft sequence of the human genome. *Nature* **409:** 953-8.

Chittenden T., Flemington C., Houghton A. B., Ebb R. G., Gallo G. J., Elangovan B., Chinnadurai G., and Lutz R. J. (1995). A conserved domain in Bak, distinct from BH1 and BH2, mediates cell death and protein binding functions. *Embo J* **14:** 5589-96.

Chomczynski P., and Sacchi N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate- phenol-chloroform extraction. *Anal Biochem* **162:** 156-9.

Chumakov I., Rigault P., Guillou S., Ougen P., Billaut A., Guasconi G., Gervy P., LeGall I., Soularue P., Grinas L., and et al. (1992). Continuum of overlapping clones spanning the entire human chromosome 21q. *Nature* **359:** 380-7.

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.

Cleves A. E. (1997). Protein transports: the nonclassical ins and outs. *Curr Biol* **7:** R318-20.

Coldwell M. J., Mitchell S. A., Stoneley M., MacFarlane M., and Willis A. E. (2000). Initiation of Apaf-1 translation by internal ribosome entry. *Oncogene* **19:** 899-905.

Colgan D. F., and Manley J. L. (1997). Mechanism and regulation of mRNA polyadenylation. *Genes Dev* **11:** 2755-66.

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.

Collins J. E., Cole C. G., Smink L. J., Garrett C. L., Leversha M. A., Soderlund C. A., Maslen G. L., Everett L. A., Rice K. M., Coffey A. J., and et al. (1995). A high-density YAC contig map of human chromosome 22. *Nature* **377:** 367-79.

Cooper D. N., and Krawczak M. (1989). Cytosine methylation and the fate of CpG dinucleotides in vertebrate genomes. *Hum Genet* **83:** 181-8.

Cooper P. R., Smilinich N. J., Day C. D., Nowak N. J., Reid L. H., Pearsall R. S., Reece M., Prawitt D., Landers J., Housman D. E., Winterpacht A., Zabel B. U., Pelletier J., Weissman B. E., Shows T. B., and Higgins M. J. (1998). Divergently transcribed overlapping genes expressed in liver and kidney and located in the 11p15.5 imprinted domain. *Genomics* **49:** 38-51.

Corpet F., Servant F., Gouzy J., and Kahn D. (2000). ProDom and ProDom-CG: tools for protein domain analysis and whole genome comparisons. *Nucleic Acids Res* **28:** 267-9.

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. (1996). The Caenorhabditis elegans genome project. C. elegans Genome Consortium. *Biochem Soc Trans* **24:** 289-91.

Cousin H., Gaultier A., Bleux C., Darribere T., and Alfandari D. (2000). PACSIN2 is a regulator of the metalloprotease/disintegrin ADAM13. *Dev Biol* **227:** 197-210.

Cox D. R. (1992). Radiation hybrid mapping. *Cytogenet Cell Genet* **59:** 80-1.

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.

Craig J. M., and Bickmore W. A. (1993). Chromosome bands--flavours to savour. *Bioessays* **15:** 349-54.

Creighton T. E. (1993). Proteins. Second edition. W. H. Freeman and Company.

Crick F. (1957). On protein synthesis. *Symp. Soc. Exp. Biol*. **12**: 138-163.

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. (2000). 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., 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.

Das M., Burge C. B., Park E., Colinas J., and Pelletier J. (2001). Assessment of the total number of human transcription units. *Genomics* **77:** 71-8.

de Souza S. J., Camargo A. A., Briones M. R., Costa F. F., Nagai M. A., Verjovski-Almeida S., Zago M. A., Andrade L. E., Carrer H., El-Dorry H. F., Espreafico E. M., Habr-Gama A., Giannella-Neto D., Goldman G. H., Gruber A., Hackel C., Kimura E. T., Maciel R. M., Marie S. K., Martins E. A., Nobrega M. P., Paco-Larson M. L., Pardini M. I., Pereira G. G., Pesquero J. B., Rodrigues V., Rogatto S. R., da Silva I. D., Sogayar M. C., de Fatima Sonati M., Tajara E. H., Valentini S. R., Acencio M., Alberto F. L., Amaral M. E., Aneas I., Bengtson M. H., Carraro D. M., Carvalho A. F., Carvalho L. H., Cerutti J. M., Correa M. L., Costa M. C., Curcio C., Gushiken T., Ho P. L., Kimura E., Leite L. C., Maia G., Majumder P., Marins M., Matsukuma A., Melo A. S., Mestriner C. A., Miracca E. C., Miranda D. C., Nascimento A. N., Nobrega F. G., Ojopi E. P., Pandolfi J. R., Pessoa L. G., Rahal P., Rainho C. A., da Ros N., de Sa R. G., Sales M. M., da Silva N. P., Silva T. C., da Silva W., Jr., Simao D. F., Sousa J. F., Stecconi D., Tsukumo F., Valente V., Zalcbeg H., Brentani R. R., Reis F. L., Dias-Neto E., and Simpson A. J. (2000). Identification of human chromosome 22

transcribed sequences with ORF expressed sequence tags. *Proc Natl Acad Sci U S A* **97:** 12690-3.

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.

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.

den Dunnen J. T., van Neck J. W., Cremers F. P., Lubsen N. H., and Schoenmakers J. G. (1989). Nucleotide sequence of the rat gamma-crystallin gene region and comparison with an orthologous human region. *Gene* **78:** 201-13.

Deng A. Y., and Rapp J. P. (1994). Evaluation of the angiotensin II receptor AT1B gene as a candidate gene for blood pressure. *J Hypertens* **12:** 1001-6.

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.

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., 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.

Ding D. Q., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T., and Hiraoka Y. (2000). Large-scale screening of intracellular protein localization in living fission yeast cells by the use of a GFP-fusion genomic DNA library. *Genes Cells* **5:** 169-90.

Doggett N. A., Goodwin L. A., Tesmer J. G., Meincke L. J., Bruce D. C., Clark L. M., Altherr M. R., Ford A. A., Chi H. C., Marrone B. L., and et al. (1995). An integrated physical map of human chromosome 16. *Nature* **377:** 335-65.

Dongre A. R., Eng J. K., and Yates J. R., 3rd (1997). Emerging tandem-mass-spectrometry techniques for the rapid identification of proteins. *Trends Biotechnol* **15:** 418-25.

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.

Dulbecco R. (1986). A turning point in cancer research: sequencing the human genome. *Science* **231:** 1055-6.

Dunham I., Shimizu N., Roe B. A., Chissoe 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.

Duret L., Mouchiroud D., and Gautier C. (1995). Statistical analysis of vertebrate sequences reveals that long genes are scarce in GC-rich isochores. *J Mol Evol* **40:** 308-17.

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.

Eisenhaber F., Persson B., and Argos P. (1995). Protein structure prediction: recognition of primary, secondary, and tertiary structural features from amino acid sequence. *Crit Rev Biochem Mol Biol* **30:** 1-94.

Emanuelsson O., Nielsen H., Brunak S., and von Heijne G. (2000). Predicting subcellular localization of proteins based on their N- terminal amino acid sequence. *J Mol Biol* **300:** 1005-16.

Epp T. A., Wang R., Sole M. J., and Liew C. C. (1995). Concerted evolution of mammalian cardiac myosin heavy chain genes. *J Mol Evol* **41:** 284-92.

Eppig J. T., and Nadeau J. H. (1995). Comparative maps: the mammalian jigsaw puzzle. *Curr Opin Genet Dev* **5:** 709-16.

Etzold T., Ulyanov A., and Argos P. (1996). SRS: information retrieval system for molecular biology data banks. *Methods Enzymol* **266:** 114-28.

Falany C. N., Xie X., Wang J., Ferrer J., and Falany J. L. (2000). Molecular cloning and expression of novel sulphotransferase-like cDNAs from human and rat brain. *Biochem J* **346 Pt 3:** 857-64.

Favello A., Hillier L., and Wilson R. K. (1995). Genomic DNA sequencing methods. *Methods Cell Biol* **48:** 551-69.

Feinberg A. P., and Vogelstein B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* **132:** 6-13.

Fichant G. A., and Burks C. (1991). Identifying potential tRNA genes in genomic DNA sequences. *J Mol Biol* **220:** 659-71.

Fitch W. M. (1970). Distinguishing homologous from analogous proteins. *Syst Zool* **19:** 99-113.

Fodor S. P., Read J. L., Pirrung M. C., Stryer L., Lu A. T., and Solas D. (1991). Light-directed, spatially addressable parallel chemical synthesis. *Science* **251:** 767-73.

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.

Foy C., Newton V., Wellesley D., Harris R., and Read A. P. (1990). Assignment of the locus for Waardenburg syndrome type I to human chromosome 2q37 and possible homology to the Splotch mouse. *Am J Hum Genet* **46:** 1017-23.

Frazer K. A., Sheehan J. B., Stokowski R. P., Chen X., Hosseini R., Cheng J. F., Fodor S. P., Cox D. R., and Patil N. (2001). Evolutionarily conserved sequences on human chromosome 21. *Genome Res* **11:** 1651-9.

Galtier N., Gouy M., and Gautier C. (1996). SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Comput Appl Biosci* **12:** 543-8.

Garcia-Mata R., Bebok Z., Sorscher E. J., and Sztul E. S. (1999). Characterization and dynamics of aggresome formation by a cytosolic GFP- chimera. *J Cell Biol* **146:** 1239-54.

Gardiner K. (1995). Human genome organization. *Curr Opin Genet Dev* **5:** 315-22.

Gardiner K. (1996). Base composition and gene distribution: critical patterns in mammalian genome organization. *Trends Genet* **12:** 519-24.

Gardiner K., and Mural R. J. (1995). Getting the message: identifying transcribed sequences. *Trends Genet* **11:** 77-9.

Gardiner-Garden M., and Frommer M. (1987). CpG islands in vertebrate genomes. *J Mol Biol* **196:** 261-82.

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.

Gelfand M. S. (1995). Prediction of function in DNA sequence analysis. *J Comput Biol* **2:** 87-115.

Gemmill R. M., Chumakov I., Scott P., Waggoner B., Rigault P., Cypser J., Chen Q., Weissenbach J., Gardiner K., Wang H., and et al. (1995). A second-generation YAC contig map of human chromosome 3. *Nature* **377:** 299-319.

Gianfrancesco F., Esposito T., Ruini L., Houlgatte R., Nagaraja R., D'Esposito M., Rocchi M., Auffray C., Schlessinger D., D'Urso M., and Forabosco A. (1997). Mapping of 59 EST gene markers in 31 intervals spanning the human X chromosome. *Gene* **187:** 179-84.

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.

Goss S. J., and Harris H. (1975). New method for mapping genes in human chromosomes. *Nature* **255:** 680-4.

Gottgens 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.

Goulding M. D., Chalepakis G., Deutsch U., Erselius J. R., and Gruss P. (1991). Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. *Embo J* **10:** 1135-47.

Graves J. A. (1996). Mammals that break the rules: genetics of marsupials and monotremes. *Annu Rev Genet* **30:** 233-60.

Gray N. K., and Wickens M. (1998). Control of translation initiation in animals. *Annu Rev Cell Dev Biol* **14:** 399-458.

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., Riethman H. C., Dutchik J. E., and Olson M. V. (1991). Detection and characterization of chimeric yeast artificial-chromosome clones. *Genomics* **11:** 658-69.

Gregory S. G., Howell G. R., and Bentley D. R. (1997). Genome mapping by fluorescent fingerprinting. *Genome Res* **7:** 1162-8.

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.

Gumucio D. L., Wiebauer K., Caldwell R. M., Samuelson L. C., and Meisler M. H. (1988). Concerted evolution of human amylase genes. *Mol Cell Biol* **8:** 1197-205.

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.

Gygi S. P., Rochon Y., Franza B. R., and Aebersold R. (1999). Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol* **19:** 1720-30.

Haig D. (1999). A brief history of human autosomes. *Philos Trans R Soc Lond B Biol Sci* **354:** 1447-70.

Han J., Sabbatini P., and White E. (1996). Induction of apoptosis by human Nbk/Bik, a BH3-containing protein that interacts with E1B 19K. *Mol Cell Biol* **16:** 5857-64.

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.

Heller R. A., Schena M., Chai A., Shalon D., Bedilion T., Gilmore J., Woolley D. E., and Davis R. W. (1997). Discovery and analysis of inflammatory disease-related genes using cDNA microarrays. *Proc Natl Acad Sci U S A* **94:** 2150-5.

Hess G. F., Drong R. F., Weiland K. L., Slightom J. L., Sclafani R. A., and Hollingsworth R. E. (1998). A human homolog of the yeast CDC7 gene is overexpressed in some tumors and transformed cell lines. *Gene* **211:** 133-40.

Hieter P., and Boguski M. (1997). Functional genomics: it's all how you read it. *Science* **278:** 601-2.

Hillier L. D., Lennon G., Becker M., Bonaldo M. F., Chiapelli B., Chissoe 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.

Hirsch T., Decaudin D., Susin S. A., Marchetti P., Larochette N., Resche-Rigon M., and Kroemer G. (1998). PK11195, a ligand of the mitochondrial benzodiazepine receptor, facilitates the induction of apoptosis and reverses Bcl-2-mediated cytoprotection. *Exp Cell Res* **241:** 426-34.

Hodgson C. P., and Fisk R. Z. (1987). Hybridization probe size control: optimized 'oligolabelling'. *Nucleic Acids Res* **15:** 6295.

Hofmann K., Bucher P., Falquet L., and Bairoch A. (1999). The PROSITE database, its status in 1999. *Nucleic Acids Res* **27:** 215-9.

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.

Hoja M. R., Wahlestedt C., and Hoog C. (2000). A visual intracellular classification strategy for uncharacterized human proteins. *Exp Cell Res* **259:** 239-46.

Hood L., Rowen L., and Koop B. F. (1995). Human and mouse T-cell receptor loci: genomics, evolution, diversity, and serendipity. *Ann N Y Acad Sci* **758:** 390-412.

Houlgatte R., Mariage-Samson R., Duprat S., Tessier A., Bentolila S., Lamy B., and Auffray C. (1995). The Genexpress Index: a resource for gene discovery and the genic map of the human genome. *Genome Res* **5:** 272-304.

Huang X. Q., Hardison R. C., and Miller W. (1990). A space-efficient algorithm for local similarities. *Comput Appl Biosci* **6:** 373-81.

Huang S. H., Yang A. Y. and Holcenberg J. (1993). Amplification of gene ends from gene libraries by polymerase chain reaction with single-sided specificity. *Methods in Molecular Biology, PCR Protocols: Current Methods and Applications* (ed. White B. A.) 357-363 (Humana Press, Totowa, New Jersey).

Hubbard T., and Birney E. (2000). Open annotation offers a democratic solution to genome sequencing. *Nature* **403:** 825.

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 D. C. (2000). MIRs as agents of mammalian gene evolution. *Trends Genet* **16:** 60-2.

Hurst L. D., and Eyre-Walker A. (2000). Evolutionary genomics: reading the bands. *Bioessays* **22:** 105-7.

Huynen M. A., and Bork P. (1998). Measuring genome evolution. *Proc Natl Acad Sci U S A* **95:** 5849-56.

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.

Jackson R. J., and Kaminski A. (1995). Internal initiation of translation in eukaryotes: the picornavirus paradigm and beyond. *Rna* **1:** 985-1000.

Jang W., Hua A., Spilson S. V., Miller W., Roe B. A., and Meisler M. H. (1999). Comparative sequence of human and mouse BAC clones from the mnd2 region of chromosome 2p13. *Genome Res* **9:** 53-61.

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.

Jones C. T., Morrice D. R., Paton I. R., and Burt D. W. (1997). Gene homologs on human chromosome 15q21-q26 and a chicken microchromosome identify a new conserved segment. *Mamm Genome* **8:** 436-40.

Jurka J., and Kapitonov V. V. (1999). Sectorial mutagenesis by transposable elements. *Genetica* **107:** 239-48.

Kamb A., Wang C., Thomas A., DeHoff B. S., Norris F. H., Richardson K., Rine J., Skolnick M. H., and Rosteck P. R., Jr. (1995). Software trapping: a strategy for finding genes in large genomic regions. *Comput Biomed Res* **28:** 140-53.

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. A., 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., Toyo-oka 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.

Keusch J. J., Manzella S. M., Nyame K. A., Cummings R. D., and Baenziger J. U. (2000). Cloning of Gb3 synthase, the key enzyme in globo-series glycosphingolipid synthesis, predicts a family of alpha 1, 4- glycosyltransferases conserved in plants, insects, and mammals. *J Biol Chem* **275:** 25315-21.

Khan A. S., Wilcox A. S., Polymeropoulos M. H., Hopkins J. A., Stevens T. J., Robinson M., Orpana A. K., and Sikela J. M. (1992). Single pass sequencing and physical and genetic mapping of human brain cDNAs. *Nat Genet* **2:** 180-5.

Khorana H. G., Buchi H., Ghosh H., Gupta N., Jacob T. M., Kossel H., Morgan R., Narang S. A., Ohtsuka E., and Wells R. D. (1966). Polynucleotide synthesis and the genetic code. *Cold Spring Harb Symp Quant Biol* **31:** 39-49.

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.

Kojima Y., Fukumoto S., Furukawa K., Okajima T., Wiels J., Yokoyama K., Suzuki Y., Urano T., and Ohta M. (2000). Molecular cloning of globotriaosylceramide/CD77 synthase, a glycosyltransferase that initiates the synthesis of globo series glycosphingolipids. *J Biol Chem* **275:** 15152-6.

Koop B. F. (1995). Human and rodent DNA sequence comparisons: a mosaic model of genomic evolution. *Trends Genet* **11:** 367-71.

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.

Kozak M. (1980). Evaluation of the "scanning model" for initiation of protein synthesis in eucaryotes. *Cell* **22:** 7-8.

Kozak M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res* **15:** 8125-48.

Kozak M. (1999). Initiation of translation in prokaryotes and eukaryotes. *Gene* **234:** 187-208.

Kozak M. (2000). Do the 5'untranslated domains of human cDNAs challenge the rules for initiation of translation (or is it vice versa)? *Genomics* **70:** 396-406.

Krauter K., Montgomery K., Yoon S. J., LeBlanc-Straceski J., Renault B., Marondel I., Herdman V., Cupelli L., Banks A., Lieman J., and et al. (1995). A second-generation YAC contig map of human chromosome 12. *Nature* **377:** 321-33.

Kremer E. J., et al. (1991). Mapping of DNA instability ath the fragil X to a trinucleotide repeat sequence p(CCG)n. *Science* **252**, 1711-4.

Krichevsky A. M., Metzer E., and Rosen H. (1999). Translational control of specific genes during differentiation of HL-60 cells. *J Biol Chem* **274:** 14295-305.

Kruglyak S., Durrett R. T., Schug M. D., and Aquadro C. F. (1998). Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. *Proc Natl Acad Sci U S A* **95:** 10774-8.

Kulp D., Haussler D., Reese M. G., and Eeckman F. H. (1996). A generalized hidden Markov model for the recognition of human genes in DNA. *Proc Int Conf Intell Syst Mol Biol* **4:** 134-42.

Kuster B., and Mann M. (1998). Identifying proteins and post-translational modifications by mass spectrometry. *Curr Opin Struct Biol* **8:** 393-400.

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.

Lander E. S. (1996). The new genomics: global views of biology. *Science* **274:** 536-9.

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.

Larsen F., Gundersen G., Lopez R., and Prydz H. (1992). CpG islands as gene markers in the human genome. *Genomics* **13:** 1095-107.

Lee L. G., Connell C. R., Woo S. L., Cheng R. D., McArdle B. F., Fuller C. W., Halloran N. D., and Wilson R. K. (1992). DNA sequencing with dye-labeled terminators and T7 DNA polymerase: effect of dyes and dNTPs on incorporation of dye-terminators and probability analysis of termination fragments. *Nucleic Acids Res* **20:** 2471-83.

Lennon G., Auffray C., Polymeropoulos M., and Soares M. B. (1996). The I.M.A.G.E. Consortium: an integrated molecular analysis of genomes and their expression. *Genomics* **33:** 151-2.

Lewin B. (1997). Genes VI. Oxford University Press and Cell Press.

Li W. (2001). Delineating relative homogeneous G+C domains in DNA sequences. *Gene* **276:** 57-72.

Liang P., and Pardee A. B. (1992). Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* **257:** 967-71.

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.

Liu C. C., Simonsen C. C., and Levinson A. D. (1984). Initiation of translation at internal AUG codons in mammalian cells. *Nature* **309:** 82-5.

Liu J., and Rost B. (2001). Comparing function and structure between entire proteomes. *Protein Sci* **10:** 1970-9.

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.

Lund J., Chen F., Hua A., Roe B., Budarf M., Emanuel B. S., and Reeves R. H. (2000). Comparative sequence analysis of 634 kb of the mouse chromosome 16 region of conserved synteny with the human velocardiofacial syndrome region on chromosome 22q11.2. *Genomics* **63:** 374-83.

Lundin L. G. (1993). Evolution of the vertebrate genome as reflected in paralogous chromosomal regions in man and the house mouse. *Genomics* **16:** 1-19.

Lupas A. (1996). Coiled coils: new structures and new functions. *Trends Biochem Sci* **21:** 375-82.

Lutz-Freyermuth C., Query C. C., and Keene J. D. (1990). Quantitative determination that one of two potential RNA-binding domains of the A protein component of the U1 small

nuclear ribonucleoprotein complex binds with high affinity to stem-loop II of U1 RNA. *Proc Natl Acad Sci U S A* **87:** 6393-7.

Maglott D. R., Katz K. S., Sicotte H., and Pruitt K. D. (2000). NCBI's LocusLink and RefSeq. *Nucleic Acids Res* **28:** 126-8.

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.

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.

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.

Matsuo K., Clay O., Takahashi T., Silke J., and Schaffner W. (1993). Evidence for erosion of mouse CpG islands during mammalian evolution. *Somat Cell Mol Genet* **19:** 543-55.

Matsuura T., Yamagata T., Burgess D. L., Rasmussen A., Grewal R. P., Watase K., Khajavi M., McCall A. E., Davis C. F., Zu L., Achari M., Pulst S. M., Alonso E., Noebels J. L., Nelson D. L., Zoghbi H. Y., and Ashizawa T. (2000). Large expansion of the ATTCT pentanucleotide repeat in spinocerebellar ataxia type 10. *Nat Genet* **26:** 191-4.

Maxam A. M., and Gilbert W. (1977). A new method for sequencing DNA. *Proc Natl Acad Sci U S A* **74:** 560-4.

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., Asakawa S., et al. (2001). A physical map of the human genome. *Nature* **409:** 934-41.

Meisler M. H. (2001). Evolutionarily conserved noncoding DNA in the human genome: how much and what for? *Genome Res* **11:** 1617-8.

Merilainen J., Lehto V. P., and Wasenius V. M. (1997). FAP52, a novel, SH3 domain-containing focal adhesion protein. *J Biol Chem* **272:** 23278-84.

Merkulov G. V., and Boeke J. D. (1998). Libraries of green fluorescent protein fusions generated by transposition in vitro. *Gene* **222:** 213-22.

Mironov A. A., Fickett J. W., and Gelfand M. S. (1999). Frequent alternative splicing of human genes. *Genome Res* **9:** 1288-93.

Molnar A., and Georgopoulos K. (1994). The Ikaros gene encodes a family of functionally diverse zinc finger DNA-binding proteins. *Mol Cell Biol* **14:** 8292-303.

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.

Montzka K. A., and Steitz J. A. (1988). Additional low-abundance human small nuclear ribonucleoproteins: U11, U12, etc. *Proc Natl Acad Sci U S A* **85:** 8885-9.

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.

Mott R. (1997). EST_GENOME: a program to align spliced DNA sequences to unspliced genomic DNA. *Comput Appl Biosci* **13:** 477-8.

Mukherjee S., and Das S. K. (1989). Subcellular distribution of "peripheral type" binding sites for [3H]Ro5- 4864 in guinea pig lung. Localization to the mitochondrial inner membrane. *J Biol Chem* **264:** 16713-8.

Muller D., Thieke K., Burgin A., Dickmanns A., and Eilers M. (2000). Cyclin E-mediated elimination of p27 requires its interaction with the nuclear pore-associated protein mNPAP60. *Embo J* **19:** 2168-80.

Muller S., Stanyon R., O'Brien P. C., Ferguson-Smith M. A., Plesker R., and Wienberg J. (1999). Defining the ancestral karyotype of all primates by multidirectional chromosome painting between tree shrews, lemurs and humans. *Chromosoma* **108:** 393-400.

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., Lamble 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., Heathcott 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.

Mungall A. J., Humphray S. J., Ranby S. A., Edwards C. A., Heathcott R. W., Clee C. M., Holloway E., Peck A. I., Harrison P., Green L. D., Butler A. P., Langford C. F., William R. G., Huckle E. J., Baron L., Smith A., Leversha M. A., Ramsey Y. H., Clegg S. M., Rice C. M., Maslen G. L., Hunt S. E., Scott C. E., Soderlund C. A., Dunham I., and et al. (1997). From long range mapping to sequence-ready contigs on human chromosome 6. *DNA Seq* **8:** 151-4.

Murakami K., Yubisui T., Takeshita M., and Miyata T. (1989). The NH2-terminal structures of human and rat liver microsomal NADH- cytochrome b5 reductases. *J Biochem (Tokyo)* **105:** 312-7.

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.

Nadeau J. H., and Sankoff D. (1998). Counting on comparative maps. *Trends Genet* **14:** 495-501.

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.

Nagaraja R., MacMillan S., Kere J., Jones C., Griffin S., Schmatz M., Terrell J., Shomaker M., Jermak C., Hott C., Masisi M., Mumm S., Srivastava A., Pilia G., Featherstone T., Mazzarella R., Kesterson S., McCauley B., Railey B., Burough F., Nowotny V., D'Urso M., States D., Brownstein B., and Schlessinger D. (1997). X chromosome map at 75-kb STS resolution, revealing extremes of recombination and GC content. *Genome Res* **7:** 210-22.

Nagase T., Seki N., Tanaka A., Ishikawa K., and Nomura N. (1995). Prediction of the coding sequences of unidentified human genes. IV. The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by analysis of cDNA clones from human cell line KG-1. *DNA Res* **2:** 167-74, 199-210.

Nakai A., and Ishikawa T. (2000). A nuclear localization signal is essential for stress-induced dimer-to- trimer transition of heat shock transcription factor 3. *J Biol Chem* **275:** 34665-71.

Nakai K., and Horton P. (1999). PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem Sci* **24:** 34-6.

Nielsen H., Engelbrecht J., Brunak S., and von Heijne G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* **10:** 1-6.

Nirenberg M., Caskey T., Marshall R., Brimacombe R., Kellogg D., Doctor B., Hatfield D., Levin J., Rottman F., Pestka S., Wilcox M., and Anderson F. (1966). The RNA code and protein synthesis. *Cold Spring Harb Symp Quant Biol* **31:** 11-24.

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.

Oakey R. J., Watson M. L., and Seldin M. F. (1992). Construction of a physical map on mouse and human chromosome 1: comparison of 13 Mb of mouse and 11 Mb of human DNA. *Hum Mol Genet* **1:** 613-20.

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.

O'Farrell P. H. (1975). High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* **250:** 4007-21.

Okada N., and Hamada M. (1997). The 3' ends of tRNA-derived SINEs originated from the 3' ends of LINEs: a new example from the bovine genome. *J Mol Evol* **44:** S52-6.

Okubo K., Hori N., Matoba R., Niiyama T., Fukushima A., Kojima Y., and Matsubara K. (1992). Large scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression. *Nat Genet* **2:** 173-9.

Olski T. M., Noegel A. A., and Korenbaum E. (2001). Parvin, a 42 kDa focal adhesion protein, related to the alpha-actinin superfamily. *J Cell Sci* **114:** 525-38.

Olson J. M., Ciliax B. J., Mancini W. R., and Young A. B. (1988). Presence of peripheral-type benzodiazepine binding sites on human erythrocyte membranes. *Eur J Pharmacol* **152:** 47-53.

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.

Orengo C. A., Todd A. E., and Thornton J. M. (1999). From protein structure to function. *Curr Opin Struct Biol* **9:** 374-82.

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.

Ozols J., Carr S. A., and Strittmatter P. (1984). Identification of the NH2-terminal blocking group of NADH-cytochrome b5 reductase as myristic acid and the complete amino acid sequence of the membrane-binding domain. *J Biol Chem* **259:** 13349-54.

Pan T. C., Kluge M., Zhang R. Z., Mayer U., Timpl R., and Chu M. L. (1993). Sequence of extracellular mouse protein BM-90/fibulin and its calcium- dependent binding to other basement-membrane ligands. *Eur J Biochem* **215:** 733-40.

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.

Parola A. L., Stump D. G., Pepperl D. J., Krueger K. E., Regan J. W., and Laird H. E., 2nd (1991). Cloning and expression of a pharmacologically unique bovine peripheral- type benzodiazepine receptor isoquinoline binding protein. *J Biol Chem* **266:** 14082-7.

Pepperkok R., Simpson J. C., and Wiemann S. (2001). Being in the right location at the right time. *Genome Biol* **2**.

Peri S., and Pandey A. (2001). A reassessment of the translation initiation codon in vertebrates. *Trends Genet* **17:** 685-7.

Persson B., Zigler J. S., Jr., and Jornvall H. (1994). A super-family of medium-chain dehydrogenases/reductases (MDR). Sub- lines including zeta-crystallin, alcohol and polyol dehydrogenases, quinone oxidoreductase enoyl reductases, VAT-1 and other proteins. *Eur J Biochem* **226:** 15-22.

Piatigorsky J., and Wistow G. (1991). The recruitment of crystallins: new functions precede gene duplication. *Science* **252:** 1078-9.

Pichon B., Mercan D., Pouillon V., Christophe-Hobertus C., and Christophe D. (2000). A method for the large-scale cloning of nuclear proteins and nuclear targeting sequences on a functional basis. *Anal Biochem* **284:** 231-9.

Pietrini G., Carrera P., and Borgese N. (1988). Two transcripts encode rat cytochrome b5 reductase. *Proc Natl Acad Sci U S A* **85:** 7246-50.

Pletcher M. T., Roe B. A., Chen F., Do T., Do A., Malaj E., and Reeves R. H. (2000). Chromosome evolution: the junction of mammalian chromosomes in the formation of mouse chromosome 10. *Genome Res* **10:** 1463-7.

Prober J. M., Trainor G. L., Dam R. J., Hobbs F. W., Robertson C. W., Zagursky R. J., Cocuzza A. J., Jensen M. A., and Baumeister K. (1987). A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science* **238:** 336-41.

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.

Qualmann B., and Kelly R. B. (2000). Syndapin isoforms participate in receptor-mediated endocytosis and actin organization. *J Cell Biol* **148:** 1047-62.

Quandt K., Frech K., Karas H., Wingender E., and Werner T. (1995). MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res* **23:** 4878-84.

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.

Reichhardt T. (1999). It's sink or swim as a tidal wave of data approaches. *Nature* **399:** 517-20.

Rettenberger G., Klett C., Zechner U., Bruch J., Just W., Vogel W., and Hameister H. (1995). ZOO-FISH analysis: cat and human karyotypes closely resemble the putative ancestral mammalian karyotype. *Chromosome Res* **3:** 479-86.

Rhodes M., Straw R., Fernando S., Evans A., Lacey T., Dearlove A., Greystrong 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.

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.

Riond J., Mattei M. G., Kaghad M., Dumont X., Guillemot J. C., Le Fur G., Caput D., and Ferrara P. (1991). Molecular cloning and chromosomal localization of a human peripheral- type benzodiazepine receptor. *Eur J Biochem* **195:** 305-11.

Ritter B., Modregger J., Paulsson M., and Plomann M. (1999). PACSIN 2, a novel member of the PACSIN family of cytoplasmic adapter proteins. *FEBS Lett* **454:** 356-62.

Rivas E., Klein R. J., Jones T. A., and Eddy S. R. (2001). Computational identification of noncoding RNAs in E. coli by comparative genomics. *Curr Biol* **11:** 1369-73.

Roest Crollius H., Jaillon O., Bernot A., Dasilva C., Bouneau L., Fischer C., Fizames C., Wincker P., Brottier P., Quetier F., Saurin W., and Weissenbach J. (2000). Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence. *Nat Genet* **25:** 235-8.

Rolls M. M., Stein P. A., Taylor S. S., Ha E., McKeon F., and Rapoport T. A. (1999). A visual screen of a GFP-fusion library identifies a new type of nuclear envelope membrane protein. *J Cell Biol* **146:** 29-44.

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.

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.

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., Della Valle G., and Bernardi G. (1992). The highest gene concentrations in the human genome are in telomeric bands of metaphase chromosomes. *Proc Natl Acad Sci U S A* **89:** 4913-7.

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.

Sadusky T. J., Kemp T. J., Simon M., Carey N., and Coulton G. R. (2001). Identification of Serhl, a new member of the serine hydrolase family induced by passive stretch of skeletal muscle in vivo. *Genomics* **73:** 38-49.

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.

Saitou N., and Nei M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4:** 406-25.

Salamov A. A., and Solovyev V. V. (2000). Ab initio gene finding in Drosophila genomic DNA. *Genome Res* **10:** 516-22.

Sander C., and Schneider R. (1991). Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins* **9:** 56-68.

Sanger F., Coulson A. R., Hong G. F., Hill D. F., and Petersen G. B. (1982). Nucleotide sequence of bacteriophage lambda DNA. *J Mol Biol* **162:** 729-73.

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.

Sargent M. G., and Bennett M. F. (1986). Identification of a specific membrane-particle-associated DNA sequence in Bacillus subtilis. *J Bacteriol* **166:** 38-43.

Sawin K. E., and Nurse P. (1996). Identification of fission yeast nuclear markers using random polypeptide fusions with green fluorescent protein. *Proc Natl Acad Sci U S A* **93:** 15146-51.

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.

Scherthan H., Cremer T., Arnason U., Weier H. U., Lima-de-Faria A., and Fronicke L. (1994). Comparative chromosome painting discloses homologous segments in distantly related mammals. *Nat Genet* **6:** 342-7.

Schuler G. D. (1997). Sequence mapping by electronic PCR. *Genome Res* **7:** 541-50.

Schuler G. D. (1998). Sequence alignment and database searching. *Methods Biochem Anal* **39:** 145-71.

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., Chiannilkulchai 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. (1996a). A gene map of the human genome. *Science* **274:** 540-6.

Schuler G. D., Epstein J. A., Ohkawa H., and Kans J. A. (1996b). Entrez: molecular biology database and retrieval system. *Methods Enzymol* **266:** 141-62.

Schultz J., Copley R. R., Doerks T., Ponting C. P., and Bork P. (2000). SMART: a web-based tool for the study of genetically mobile domains. *Nucleic Acids Res* **28:** 231-4.

Schulz R. A., and Butler B. A. (1989). Overlapping genes of Drosophila melanogaster: organization of the z600- gonadal-Eip28/29 gene cluster. *Genes Dev* **3:** 232-42.

Schwartz F., and Ota T. (1997). The 239AB gene on chromosome 22: a novel member of an ancient gene family. *Gene* **194:** 57-62.

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.

Sealey P. G., Whittaker P. A., and Southern E. M. (1985). Removal of repeated sequences from hybridisation probes. *Nucleic Acids Res* **13:** 1905-22.

Seed B., and Aruffo A. (1987). Molecular cloning of the CD2 antigen, the T-cell erythrocyte receptor, by a rapid immunoselection procedure. *Proc Natl Acad Sci U S A* **84:** 3365-9.

Sehgal A., Briggs J., Rinehart-Kim J., Basso J., and Bos T. J. (2000). The chicken c-Jun 5' untranslated region directs translation by internal initiation. *Oncogene* **19:** 2836-45.

Senior K. (1999). Fingerprinting disease with protein chip arrays. *Mol Med Today* **5:** 326-7.

Shaikh T. H., Kurahashi H., Saitta S. C., O'Hare A. M., Hu P., Roe B. A., Driscoll D. A., McDonald-McGinn D. M., Zackai E. H., Budarf M. L., and Emanuel B. S. (2000). Chromosome 22-specific low copy repeats and the 22q11.2 deletion syndrome: genomic organization and deletion endpoint analysis. *Hum Mol Genet* **9:** 489-501.

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.

Shirabe K., Yubisui T., Borgese N., Tang C. Y., Hultquist D. E., and Takeshita M. (1992). Enzymatic instability of NADH-cytochrome b5 reductase as a cause of hereditary methemoglobinemia type I (red cell type). *J Biol Chem* **267:** 20416-21.

Shirabe K., Yubisui T., Nishino T., and Takeshita M. (1991). Role of cysteine residues in human NADH-cytochrome b5 reductase studied by site-directed mutagenesis. Cys-273 and Cys-283 are located close to the NADH-binding site but are not catalytically essential. *J Biol Chem* **266:** 7531-6.

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-Engele 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.

Simpson J. C., Wellenreuther R., Poustka A., Pepperkok R., and Wiemann S. (2000). Systematic subcellular localization of novel proteins identified by large-scale cDNA sequencing. *EMBO Rep* **1:** 287-92.

Sinden R. R. (1999). Biological implications of the DNA structures associated with disease-causing triplet repeats. *Am J Hum Genet* **64:** 346-53.

Slusher L. B., Gillman E. C., Martin N. C., and Hopper A. K. (1991). mRNA leader length and initiation codon context determine alternative AUG selection for the yeast gene MOD5. *Proc Natl Acad Sci U S A* **88:** 9789-93.

Smink L. J. (2000). Genome studies of human chromosome 22q13.31. PhD Thesis, Open University.

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.

Smith L. M., Kaiser R. J., Sanders J. Z., and Hood L. E. (1987). The synthesis and use of fluorescent oligonucleotides in DNA sequence analysis. *Methods Enzymol* **155:** 260-301.

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., Longden I., and Mott R. (1997). FPC: a system for building contigs from restriction fingerprinted clones. *Comput Appl Biosci* **13:** 523-35.

Solovyev V., and Salamov A. (1997). The Gene-Finder computer tools for analysis of human and model organisms genome sequences. *Proc Int Conf Intell Syst Mol Biol* **5:** 294-302.

Solovyev V. V., Salamov A. A., and Lawrence C. B. (1994). The prediction of human exons by oligonucleotide composition and discriminant analysis of spliceable open reading frames. *Proc Int Conf Intell Syst Mol Biol* **2:** 354-62.

Somma M. P., Gambino I., and Lavia P. (1991). Transcription factors binding to the mouse HTF9 housekeeping promoter differ between cell types. *Nucleic Acids Res* **19:** 4451-8.

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., Laverty 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.

Sprengel R., Werner P., Seeburg P. H., Mukhin A. G., Santi M. R., Grayson D. R., Guidotti A., and Krueger K. E. (1989). Molecular cloning and expression of cDNA encoding a peripheral-type benzodiazepine receptor. *J Biol Chem* **264:** 20415-21.

Steffensen R., Carlier K., Wiels J., Levery S. B., Stroud M., Cedergren B., Nilsson Sojka B., Bennett E. P., Jersild C., and Clausen H. (2000). Cloning and expression of the histo-blood group Pk UDP-galactose: Ga1beta-4G1cbeta1-cer alpha1, 4-galactosyltransferase. Molecular genetic basis of the p phenotype. *J Biol Chem* **275:** 16723-9.

Stein L. (2001). Genome annotation: from sequence to biology. *Nat Rev Genet* **2:** 493-503.

Stein L. D., and Thierry-Mieg J. (1998). Scriptable access to the Caenorhabditis elegans genome sequence and other ACEDB databases. *Genome Res* **8:** 1308-15.

Stephens R. M., and Schneider T. D. (1992). Features of spliceosome evolution and function inferred from an analysis of the information at human splice sites. *J Mol Biol* **228:** 1124-36.

Strachan T., and Read A. P. (1999). Human molecular genetics 2. Second edition. BIOS Scientific Publishers Ltd.

Strausberg R. L., Feingold E. A., Klausner R. D., and Collins F. S. (1999). The mammalian gene collection. *Science* **286:** 455-7.

Strittmatter P., Kittler J. M., and Coghill J. E. (1992). Characterization of the role of lysine 110 of NADH-cytochrome b5 reductase in the binding and oxidation of NADH by site-directed mutagenesis. *J Biol Chem* **267:** 20164-7.

Stryer L. (1988).Biochemistry. Third edition. W.H. Freeman and Company.

Sved J., and Bird A. (1990). The expected equilibrium of the CpG dinucleotide in vertebrate genomes under a mutation model. *Proc Natl Acad Sci U S A* **87:** 4692-6.

Tagle D. A., Koop B. F., Goodman M., Slightom J. L., Hess D. L., and Jones R. T. (1988). Embryonic epsilon and gamma globin genes of a prosimian primate (Galago crassicaudatus). Nucleotide and amino acid sequences, developmental regulation and phylogenetic footprints. *J Mol Biol* **203:** 439-55.

Taketani S., Kohno H., Okuda M., Furukawa T., and Tokunaga R. (1994). Induction of peripheral-type benzodiazepine receptors during differentiation of mouse erythroleukemia cells. A possible involvement of these receptors in heme biosynthesis. *J Biol Chem* **269:** 7527-31.

Tamura M., Yubisui T., Takeshita M., Kawabata S., Miyata T., and Iwanaga S. (1987). Structural comparison of bovine erythrocyte, brain, and liver NADH- cytochrome b5 reductase by HPLC mapping. *J Biochem (Tokyo)* **101:** 1147-59.

Tassabehji M., Read A. P., Newton V. E., Harris R., Balling R., Gruss P., and Strachan T. (1992). Waardenburg's syndrome patients have mutations in the human homologue of the Pax-3 paired box gene. *Nature* **355:** 635-6.

Tatusov R. L., Koonin E. V., and Lipman D. J. (1997). A genomic perspective on protein families. *Science* **278:** 631-7.

Tatusov R. L., Mushegian A. R., Bork P., Brown N. P., Hayes W. S., Borodovsky M., Rudd K. E., and Koonin E. V. (1996). Metabolism and evolution of Haemophilus influenzae deduced from a whole- genome comparison with Escherichia coli. *Curr Biol* **6:** 279-91.

Tautz D. (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res* **17:** 6463-71.

Thanassi D. G., and Hultgren S. J. (2000). Multiple pathways allow protein secretion across the bacterial outer membrane. *Curr Opin Cell Biol* **12:** 420-30.

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.

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.

Tomita M. (2001). Whole-cell simulation: a grand challenge of the 21st century. *Trends Biotechnol* **19:** 205-10.

Toth G., Gaspari Z., and Jurka J. (2000). Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Res* **10:** 967-81.

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.

Tran H., Mattei M., Godyna S., and Argraves W. S. (1997). Human fibulin-1D: molecular cloning, expression and similarity with S1- 5 protein, a new member of the fibulin gene family. *Matrix Biol* **15:** 479-93.

Trichet V., Ruault M., Roizes G., and De Sario A. (2000). Characterization of the human tubulin tyrosine ligase-like 1 gene (TTLL1) mapping to 22q13.1. *Gene* **257:** 109-17.

Trichet V., Shkolny D., Dunham I., Beare D., and McDermid H. E. (1999). Mapping and complex expression pattern of the human NPAP60L nucleoporin gene. *Cytogenet Cell Genet* **85:** 221-6.

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* **75:** 826.

Tsai D. E., Kenan D. J., and Keene J. D. (1992). In vitro selection of an RNA epitope immunologically cross-reactive with a peptide. *Proc Natl Acad Sci U S A* **89:** 8864-8.

Uberbacher E. C., and Mural R. J. (1991). Locating protein-coding regions in human DNA sequences by a multiple sensor-neural network approach. *Proc Natl Acad Sci U S A* **88:** 11261-5.

Vagner S., Gensac M. C., Maret A., Bayard F., Amalric F., Prats H., and Prats A. C. (1995). Alternative translation of human fibroblast growth factor 2 mRNA occurs by internal entry of ribosomes. *Mol Cell Biol* **15:** 35-44.

Valdes J. M., Tagle D. A., and Collins F. S. (1994). Island rescue PCR: a rapid and efficient method for isolating transcribed sequences from yeast artificial chromosomes and cosmids. *Proc Natl Acad Sci U S A* **91:** 5377-81.

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.

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.

Verkerk A. J., et al. (1991). Identification of a gene (FMR-1) containing a CGG repeat conicident with a breakpoint cluster region exhibiting langth variation in fragil X syndrome. *Cell* **65**: 904-14.

Verma S., Budarf M. L., Emanuel B. S., and Chinnadurai G. (2000). Structural analysis of the human pro-apoptotic gene Bik: chromosomal localization, genomic organization and localization of promoter sequences. *Gene* **254:** 157-62.

Vulpe C., Levinson B., Whitney S., Packman S., and Gitschier J. (1993). Isolation of a candidate gene for Menkes disease and evidence that it encodes a copper-transporting ATPase. *Nat Genet* **3:** 7-13.

Walter G., Bussow K., Cahill D., Lueking A., and Lehrach H. (2000). Protein arrays for gene expression and molecular interaction screening. *Curr Opin Microbiol* **3:** 298-302.

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.

Watson J. D. (1990). The human genome project: past, present, and future. *Science* **248:** 44-9.

Weber J. L., Polymeropoulos M. H., May P. E., Kwitek A. E., Xiao H., McPherson J. D., and Wasmuth J. J. (1991). Mapping of human chromosome 5 microsatellite DNA polymorphisms. *Genomics* **11:** 695-700.

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.

Weterman M. A., Wilbrink M., Dijkhuizen T., van den Berg E., and Geurts van Kessel A. (1996). Fine mapping of the 1q21 breakpoint of the papillary renal cell carcinoma-associated (X;1) translocation. *Hum Genet* **98:** 16-21.

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.

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.

Wingender E., Chen X., Hehl R., Karas H., Liebich I., Matys V., Meinhardt T., Pruss M., Reuter I., and Schacherer F. (2000). TRANSFAC: an integrated system for gene expression regulation. *Nucleic Acids Res* **28:** 316-9.

Wolfe K. H., Sharp P. M., and Li W. H. (1989). Mutation rates differ among regions of the mammalian genome. *Nature* **337:** 283-5.

Wolfsberg T. G., and Landsman D. (1997). A comparison of expressed sequence tags (ESTs) to human genomic sequences. *Nucleic Acids Res* **25:** 1626-32.

Wu X. R., and Sun T. T. (1993). Molecular cloning of a 47 kDa tissue-specific and differentiation- dependent urothelial cell surface glycoprotein. *J Cell Sci* **106:** 31-43.

Xiao Z., Jiang X., Beckett M. L., and Wright G. L., Jr. (2000). Generation of a baculovirus recombinant prostate-specific membrane antigen and its use in the development of a novel protein biochip quantitative immunoassay. *Protein Expr Purif* **19:** 12-21.

Xu Y., Mural R. J., and Uberbacher E. C. (1995). Correcting sequencing errors in DNA coding regions using a dynamic programming approach. *Comput Appl Biosci* **11:** 117-24.

Yang S., Tutton S., Pierce E., and Yoon K. (2001). Specific double-stranded RNA interference in undifferentiated mouse embryonic stem cells. *Mol Cell Biol* **21:** 7807-16.

Yeh R. F., Lim L. P., and Burge C. B. (2001). Computational inference of homologous gene structures in the human genome. *Genome Res* **11:** 803-16.

Yu S. et al. (1991). Fragile X genotype characterised by an unstable region of DNA. *Science* **252**: 1179-81 (1991).

Yuasa T., Yoshiki T., Tanaka T., Kim C. J., Isono T., and Okada Y. (1998). Expression of uroplakin Ib and uroplakin III genes in tissues and peripheral blood of patients with transitional cell carcinoma. *Jpn J Cancer Res* **89:** 879-82.

Yubisui T., Miyata T., Iwanaga S., Tamura M., Yoshida S., Takeshita M., and Nakajima H. (1984). Amino acid sequence of NADH-cytochrome b5 reductase of human erythrocytes. *J Biochem (Tokyo)* **96:** 579-82.

Yubisui T., Naitoh Y., Zenno S., Tamura M., Takeshita M., and Sakaki Y. (1987). Molecular cloning of cDNAs of human liver and placenta NADH-cytochrome b5 reductase. *Proc Natl Acad Sci U S A* **84:** 3609-13.

Zenno S., Hattori M., Misumi Y., Yubisui T., and Sakaki Y. (1990). Molecular cloning of a cDNA encoding rat NADH-cytochrome b5 reductase and the corresponding gene. *J Biochem (Tokyo)* **107:** 810-6.

Zhang C., Yu Y., Zhang S., Liu M., Xing G., Wei H., Bi J., Liu X., Zhou G., Dong C., Hu Z., Zhang Y., Luo L., Wu C., Zhao S., and He F. (2000). Characterization, chromosomal assignment, and tissue expression of a novel human gene belonging to the ARF GAP family. *Genomics* **63:** 400-8.

Zhuo D., Zhao W. D., Wright F. A., Yang H. Y., Wang J. P., Sears R., Baer T., Kwon D. H., Gordon D., Gibbs S., Dai D., Yang Q., Spitzner J., Krahe R., Stredney D., Stutz A., and Yuan B. (2001). Assembly, annotation, and integration of UNIGENE clusters into the human genome draft. *Genome Res* **11:** 904-18.

Ziauddin J., and Sabatini D. M. (2001). Microarrays of cells expressing defined cDNAs. *Nature* **411:** 107-10.

Zoubak S., Clay O., and Bernardi G. (1996). The gene distribution of the human genome. *Gene* **174:** 95-102.

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.