Chapter 5

Mapping the Human Genome, Cloned Genes, DNA Polymorphisms, and Inherited Disease

Thomas B. Shows, Alan Y. Sakaguchi, and Susan L. Naylor

Department of Human Genetics
Roswell Park Memorial Institute
New York State Department of Health
Buffalo, New York

INTRODUCTION

It is estimated that the human haploid genome is composed of 3×10^9 nucleotides and that only 10% of it consists of coding and regulatory sequences.[14] If a gene is approximately 10^4 nucleotides in length, which includes the coding region and the intervening and flanking sequences, this estimate would predict that there are about $3–10 \times 10^4$ human genes coding for different protein products. Since gene clustering in humans has become evident (for example, the hemoglobin, immunoglobin, and HLA clusters), these estimated gene products may be grouped in from 3000 to 15,000 clusters.[15] Further, based upon genetic and molecular studies of microorganisms, *Drosophila,* and the mouse, there are perhaps 5×10^4 structural genes estimated in humans,[14,106,183] which is in agreement with the number of estimated protein products in the human genome. Mapping the human genome involves partitioning the total number of genes into individual maps representing the 24 different human nuclear chromosomes and linearly ordering them on each chromosome. A similar exercise has mapped the 37 genes encoded in the DNA of the mitochondrial genome.[2]

H. Harris et al. (eds.), *Advances in Human Genetics*
© Plenum Press, New York 1982

As seen in more complete maps of microorganisms and viruses, a precise molecular knowledge of the human gene map is a prerequisite for comprehending each gene's function and mode of expression, and for revealing the genetic history of man in comparison with other animal species. An understanding of the human gene map should promote a genetic knowledge of how genes function individually and as coordinated sets in man. Such information is essential for defining all aspects of normal and abnormal human biology and development.

The ultimate goal for achieving this understanding is to determine the nucleotide sequence of each chromosome and to delineate this sequence into specific genes. While this seemed utopian just a few years ago, the rapid advances in technology portend that construction of a molecular map of man will be possible. This goal has been achieved for human mitochondrial DNA and for the SV40 genome, where the linear order of its seven genes and their nucleotide sequence has been determined.[2,60,205] This reality is approaching for the nuclear gene maps, since a combination of classical Mendelian genetics, somatic cell genetics, and recombinant DNA technologies has currently made it possible to determine the chromosome assignments of a great number of human genes and, for some genes, their nucleotide sequences. Classical family studies are limited because of generation time, size of sibships, and number of genetic markers. Information is usually limited to linkage relationships and the assignment of a few genes to specific chromosomes as a result of abnormal chromosomes, X linkage, or gene dosage. The technology of somatic cell genetics is a culmination of cytogenetics, cell culture, and biochemical genetics, which in concert have contributed to an exponential increase in the number of human genes assigned to specific chromosomes and regions of chromosomes. Even with the limitation of having to deal with gene markers expressed only in cell culture, currently about 85% of the markers that have been mapped were mapped by the parasexual technology of somatic cell genetics using man–rodent somatic cell hybrids that retain only specific numbers and combinations of human chromosomes. Current development for cloning specific human genes and DNA segments by recombinant DNA technology, coupled with the previously mentioned methodology and now including *in situ* mapping, allows one to predict that it will be possible to map the majority of the human genome with few refinements in technology. Examination of human–rodent somatic cell hybrids segregating human chromosomes with a cloned gene probe has recently made it possible to assign to chromosomal regions the

genes for globin(s),[4,71,81,104,122,139,150,228] insulin,[48,85,140,189,191] growth hormone,[70,190] chorionic somatomammotropin,[70,86,190] ACTH,[193] leukocyte interferon,[194] and fibroblast interferon.[194] Although only a relatively small number of human genes have been cloned, rapid technological progress in the identification and isolation of cloned genes will ensure that their number will increase[201,253] For example, it is now possible to sequence polypeptides at the picomolar level.[101] Synthesizing oligonucleotide primers corresponding in sequence to unique stretches of oligopeptides in sequenced proteins provides a tool to "fish out" the corresponding gene from a cDNA or genomic library or from cellular mRNA.[180,201,253,268] In addition, the entire human genome has been cloned as nonspecific DNA segments comprising a genomic library.[137,154] By mapping these cloned DNA segments of known nucleotide sequences, a molecular map of man will be constructed.

As the human gene map (Appendix A) becomes more complete, its usefulness in other areas of human biology will increase proportionally. Since cloned genes and DNA segments have identified what appears to be a high level of polymorphism at the DNA level,[8,275] a very large resource of useful genetic markers has become available. These DNA polymorphisms will be mapped to specific chromosomes and be used in Mendelian gene linkage studies. This will make it possible to construct large numbers of genetic linkage groups, located on specific chromosomal regions, for genetic counseling and prenatal diagnosis of inherited molecular disease. The framework of using these technologies for predicting and recognizing carriers and for diagnosis of serious inherited disease has been formulated.[19,109] Also, knowing the location of genes will aid in their isolation for possibilities of gene transfer, in perhaps altering a specific gene for therapy, and in the isolation of specific genes for cloning and production of a specific therapeutic product.

In this review, the identification, mapping, and significance of these new markers and the beginnings of a molecular map will be presented. This information will be integrated into and compared with the current human gene map. The current status of the human nuclear and mitochondrial gene maps will be given, along with uses of the map, particularly regarding the dissection of molecular disease, genetic counseling, and prenatal diagnosis. Comparative genetic maps of other mammals will be presented and considered for their important human gene mapping predictions regarding genetic and molecular markers and inherited disease. The human genome is rapidly being mapped and its organizational features

are being determined; this information will be very important for understanding human biology and molecular disease.

The Impact of Recombinant DNA on Human Genetics

The application of recombinant DNA techniques to human genetics promises to have an impact unequalled in any other area of human biology. The ability to clone individual genes and the use of restriction endonucleases have allowed genetics to progress from a description of phenotypic traits to the actual genotype of an individual.[19,109] Present technology will allow the determination of changes in DNA sequence that give rise to variations in polymorphic proteins used in genetic linkage studies. Moreover, variations in DNA sequences among individuals can themselves now be used as polymorphic markers in linkage analysis.[69,114,145] The isolation and chromosomal assignment of polymorphic fragments of DNA promises to be a powerful tool in analyzing the Mendelian inheritance of linked genetic loci involved in inherited diseases.[19] Recent advances in chromosome isolation and separation[23] have allowed the construction of chromosome-specific recombinant libraries of DNA fragments.[42,276] Individual clones from these libraries will be useful for detecting restriction-site polymorphisms on individual chromosomes, for a comparison of the molecular organization of different chromosomes, and for analyzing regions of chromosomes whose aberrations are thought to be involved in certain forms of neoplastic disease.[67,143,216]

There is perhaps no better example of the impact of recombinant DNA on human genetics than the analysis of the β-like hemoglobin locus in man. This locus has been localized to the short arm of chromosome 11 using somatic cell hybrids and molecular hybridization methods[46,81,104,228] and also by using chromosomes isolated by flow cytometry.[139] Linkage of the β-like genes had been predicted from genetic studies and from the existence of polypeptide fusion products, termed hemoglobin Lepore and hemoglobin Kenya.[3,94] Recently, the entire locus has been isolated by recombinant DNA techniques and the complete nucleotide sequence determined.[54,68] The β-like genes are arranged on the chromosome in the order in which they are expressed during ontogenic development: $5'$-ϵ-$^{G}\gamma$-$^{A}\gamma$-δ-β-$3'$. The genes are all transcribed from the same DNA strand. To the $5'$ side of ϵ and between $^{A}\gamma$ and δ are two β-like pseudogenes, ψβ-2 and ψβ-1, respectively, that apparently do not code for functional proteins.[68]

The α-like globin gene cluster shares several structural similarities with the β-like globin locus. As with β-like genes, the genes of the α-like cluster are arranged in the order in which they are developmentally expressed: 5′-ζ2-ζ1-α2-α1-3′. In addition, ψal, an α-like pseudogene, is located between ζ1 and α2. Gene mapping studies have localized the α-globin genes to the chromosome 16 short arm.[4,45,71,122] The coding sequences of the α-like genes are interrupted by two introns that occur in identical positions in the β-globin genes, and also in the globin genes of the mouse.[134,141]

The $^{G}\gamma$, $^{A}\gamma$ genes are virtually identical in nucleotide sequence, as is also suggested by restriction mapping analysis for the α1 and α2 genes. These structural similarities appear to be due to a mechanism of gene matching and correction during mammalian evolution, perhaps through unequal crossing over and/or gene conversion.[250] The occurrence of individuals in the human population who possess chromosomes with only one or three α-globin genes would be consonant with the idea that "editing" of globin genes could occur by unequal crossing over.[250]

The occurrence and location of pseudogenes in the β-like and α-like gene clusters of all mammals examined so far provides grist for speculation that they are the result of gene duplication and sequence divergence.[155] None of the pseudogenes examined could encode a functional globin polypeptide product due to small insertions or deletions that would alter the translational reading frame. One of the two mouse α-pseudogenes is unusual since each of the two introns normally present have been excised in a manner conforming to the rules of RNA splicing, and the coding sequences are therefore contiguous.[179,265] Furthermore, both the αψ3 and αψ4 pseudogenes of the mouse are not syntenic with the active mouse α-globin genes on chromosome 11. They are instead respectively located on mouse chromosomes 15 and 17, as shown by blot hybridization experiments using hamster–rodent somatic cell hybrids.[142,204] Several mechanisms have been proposed to explain the dispersed arrangement of the mouse α-like genes, the most interesting of which involves the transposition of host genes by retroviruses.[74,142] Leder *et al.*[142] have suggested that the αψ3 pseudogene in the form of a genomic fragment could have been conveyed to chromosome 15 by a retrovirus. They support this idea with data demonstrating the presence of retroviral-like sequences in the vicinity of the αψ3 pseudogene. Comparative chromosomal and molecular mapping of the α-globin loci in both man and mouse has thus revealed the similarities (e.g., location and number of intervening

sequences) and also the differences (e.g., location of pseudogenes) displayed by these gene clusters.

Chromosomal Mapping of the β-Globin Complex

The human β-globin gene was first mapped to chromosome 11 by Deisseroth *et al.*,[46] using liquid hybridization techniques and somatic cell hybrids. Stringent hybridization conditions were used so that human DNA preferentially annealed to the human mRNA for β-globin used as a probe. After the β-globin gene had been cloned,[153] several laboratories set about to confirm the location of the gene by using somatic cell hybrids containing translocations of chromosome 11, and regionally assigned the β-globin gene to the short arm (Fig. 1). These investigators used what was a relatively new method for detecting single-copy sequences in DNA cut with restriction endonucleases that in essence allows one DNA sequence out of a million to be identified.[18,254] In addition, hybridization *in situ* to metaphase chromosomes[150] and to DNA isolated from chromosomes separated by flow sorting[139] exemplifies the varied methodology used to assign the β-globin genes. Geneticists, armed with these techniques, have begun to confront the task of constructing a molecular map of man.

Globin DNA Polymorphisms and Prenatal Diagnosis

Demonstration of linkage between a polymorphic DNA locus and an inherited genetic disease locus could be used predictively to detect heterozygotes and individuals at risk for developing disease. DNA fragment length polymorphisms have been used successfully in the prenatal diagnosis of sickle cell anemia[69,109,203] and β-thalassemia.[111,145] By following the segregation of the DNA polymorphism in families with these diseases, the occurrence of an affected child was predicted antenatally.[110,145] As the taking of fetal red blood cells for prenatal diagnosis results in an abortion rate of 7%, analysis of amniotic fluid cells using cloned DNA probes provides a safer way of assessing diseases *in utero*. The use of DNA polymorphisms for constructing a genetic linkage map of man and its uses in prenatal and antenatal diagnosis of Mendelian disorders is reviewed in a latter section.

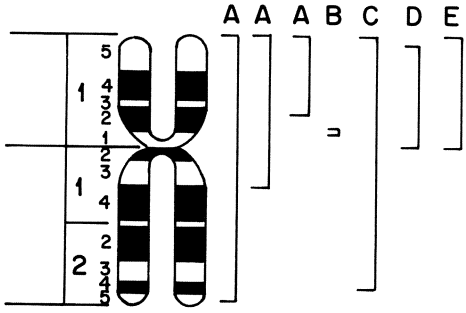

Fig. 1. Regional mapping of the β-globin gene. The original assignment of β-globin by Deisseroth *et al.*[46] was made using liquid hybridization techniques and cell hybrids. Various methodologies using cloned DNA probes have been implemented in the regional location of the β-globin gene (see text). A. Flow sorting (Lebo *et al.*[139]). B. Chromosome specific libraries and somatic cell hybrids (Gusella *et al.*[81]). C. Somatic cell hybrids (Scott *et al.*[228]). D. Somatic cell hybrids (Jeffreys *et al.*[104]). E. *In situ* hybridization (Malcolm *et al.*[150]).

CONSTRUCTING THE MOLECULAR MAP

There is a great deal of optimism that the enormous complexity of the human genome will be tamed and unraveled by a combination of several experimental approaches. Recombinant DNA techniques allowing isolation of specific genes, rapid DNA sequencing to provide a precise nucleotide map, and mapping techniques to determine chromosomal location of genes will complement one another in constructing a molecular map of man. It is no longer fanciful to speculate that soon the nucleotide sequence of large portions of human chromosomes, perhaps even whole chromosomes, will be determined using present DNA sequencing strategies.[165,224] Human geneticists will eventually have molecular information about the 5000 kb of DNA residing in a chromosome band that could provide an understanding of chromosome regions associated with neoplasia and inherited disease.[67,143,260]

Assignment of Cloned Genes

Somatic cell hybrids have been used successfully to chromosomally assign genes from several species of mammals by following the segregation of gene products or phenotypes in hybrids losing chromosomes of

one of the parental cells (see Ref. 240 for details). Mapping with cloned probes overcomes a major drawback to the above method of gene assignment: Gene expression is required to map a protein, but the gene itself can be assigned directly using a cloned DNA probe. Mapping structural gene sequences directly overcomes the uncertainty that a regulatory gene is being assigned in hybrids. Many genes coding for proteins characteristic of differentiated cells are not expressed in hybrid cells.[41] These same genes can be chromosomally assigned using cloned probes.

The methodology for mapping cloned genes using somatic cell hybrids is depicted in Fig. 2. Somatic cell hybrids are constructed by fusing rodent cells possessing a selectable recessive mutation with human leukocytes or fibroblasts.[117] Fused cells are then placed in selective medium to isolate hybrid clones.[146,260] These individual hybrid clones contain different complements of human chromosomes, whose presence can be detected karyotypically and by assaying for gene products assigned to those chromosomes.[237] DNA is isolated from the same samples of hybrid clones. The DNAs are digested with a restriction endonuclease and the fragments are separated by electrophoresis in agarose.[254] The restriction endonuclease chosen will be one that produces at least one human DNA fragment with a unique mobility compared to the mouse DNA fragments hybridizing with the probe. This is usually observed, since there has been considerable divergence in nucleotide sequence in man and mouse in and around homologous genes.[54,262] Of course, if the human gene probe does not hybridize to mouse DNA, this is no longer a consideration. It should be mentioned that human genes can be mapped using probes derived from other species as long as there is sufficient sequence homology to allow detection by molecular hybridization.

To test a probe we usually use EcoRI, BamHI, and HindIII initially to screen for differences in the size of human and mouse DNA fragments, principally because these enzymes are inexpensive and frequently yield simple hybridization patterns on Southern blots.[254] Once a suitable restriction enzyme is chosen, DNAs from a panel of hybrids are analyzed by Southern blotting using the gene probe which has been radioactively labeled with ^{32}P by nick translation[152,209] or by gap filling.[181] The hybrids are chosen so that collectively all of the human autosomes and the X chromosome are represented in a high proportion of the cells. The hybridized blots are exposed to x-ray film, which is developed to reveal specific bands representing fragments of DNA that have hybridized to the probe (Fig. 3). A correlation between the presence of a specific hy-

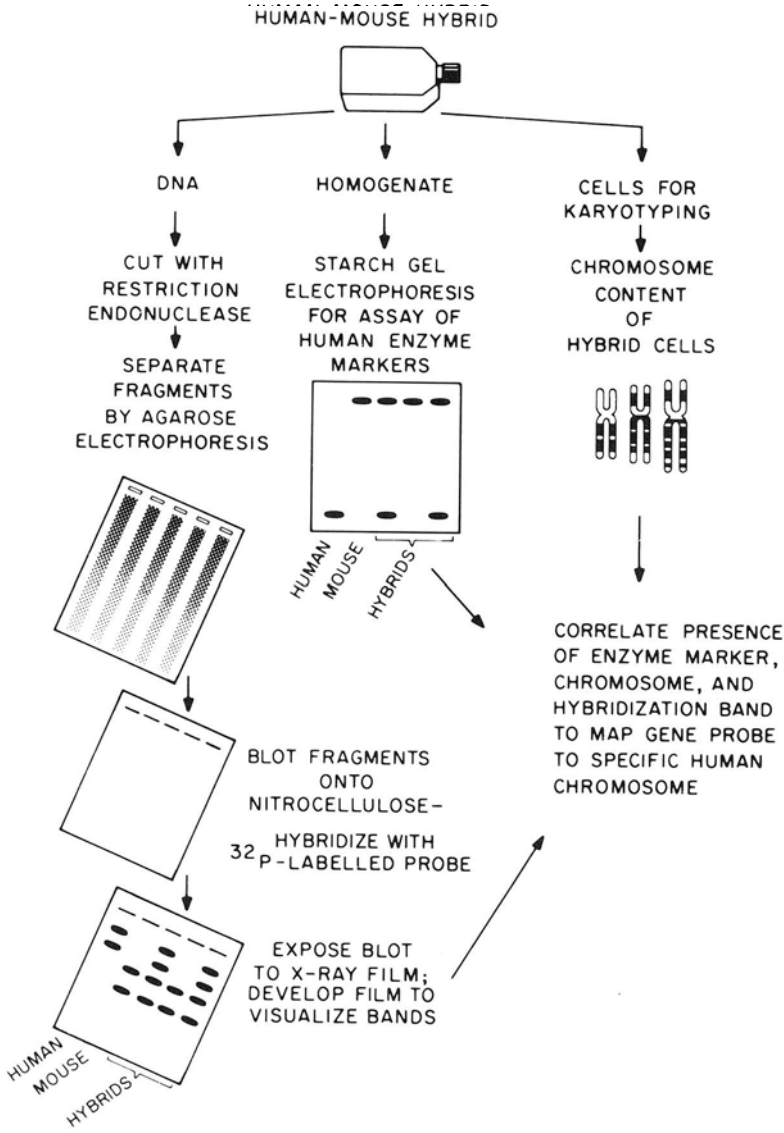

Fig. 2. Mapping of cloned genes in human–rodent somatic cell hybrids by Southern blotting. Clones of cell hybrids are grown and harvested; some of the cells are used for preparing DNA,[176] cell-free homogenates,[229] and for karyotyping.[199] Human enzyme markers previously assigned to each human autosome and the X chromosome are assayed by histochemical staining (see Table VII). Human chromosome content of hybrids is confirmed by direct karyotyping of hybrid cells. DNA from a group of cell hybrids is cleaved with a restriction endonuclease, which gives a unique pattern of human and mouse fragments when hybridized to a radiolabeled gene probe after Southern blotting. The segregation of the human-specific bands on the blots is correlated with the results of histochemical staining and karyotyping to assign the gene to a particular chromosome.

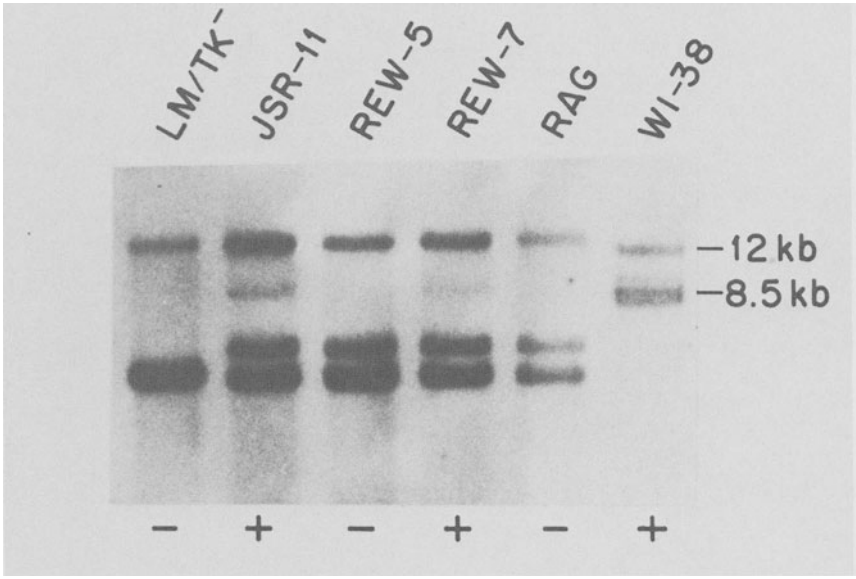

Fig. 3. Hybridization of the rat chymotrypsinogen-B (CTRB) cDNA clone to cell hybrid and parental DNAs. EcoRI-digested human DNA (WI-38) produces bands of 12 and 8.5 kb. The 8.5-kb fragment can be distinguished from bands produced from mouse DNA (RAG and LM/TK$^-$). DNA from the hybrids JSR-11 and REW-7 contained the 8.5-kb human band, while REW-5 DNA did not. The CTRB probe also detects a fragment length polymorphism in mouse DNA (two bands in LM/TK$^-$, three bands in RAG).

bridization band with a single human chromosome assigns the gene to that chromosome. For example, in Table I, it can be seen that the 8.5-kb band specific for chymotrypsinogen B, a serine protease, is seen only when human chromosome 16 is present in hybrid cells tested.[221]

cDNA Probes for Specific Genes

The isolation of a recombinant DNA probe for a specific gene can be accomplished by the synthesis of cDNA from specific messenger RNA coding for the protein of interest.[29,263] Messenger RNA is isolated from a cell type that produces fairly large amounts of the protein. The specific mRNA can be enriched by a combination of techniques, including (a) its abundance in a specific tissue,[43] (b) isolation of specific message from polysomes by antibody precipitation,[76] and (c) size fractionation of the total message.[213] The purity of the mRNA can be tested by translating the message in an *in vitro* system.[196,211] A complementary cDNA is then

TABLE I. Segregation of Chymotrypsinogen B with Chromosomes in Somatic Cell Hybrids[a]

	CTRB	1	2	3	4	5	6	7	8	9'	10	11	12	13	14	15	16	17	18	19	20	21	22	X	17/3	3/17	11/X	X/11	X/3	3/X	X/22	22/X	5/X
TSL-2	−	−	+	−	−	+	+	−	−	−	+	−	+	−	−	−	−	−	+	−	+	+	−	+	+	+							
ICL-15	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	+	−	−	+	+	−	−									
XER-11	+	+	−	+	+	−	+	+	+	+	+	−	+	−	−	+	+	+	+	+	+	+	+	−			+	+					
XER-7	−	+	+	+	+	+	+	+	+	+	+	−	+	+	+	+	−	−	+	+	−	−	−	+			+	−					
EXR-3	+	+	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+	+	+	−	+	+	+	−			+	+					
XER-9	−	−	+	+	+	−	−	−	+	−	+	−	+	−	+	−	−	+	+	−	−	+	−	−			+	−					
EXR-9	+	+	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+	+	+	+	+	+	+	−			+	+					
MAR-2	+	+	+	−	+	−	−	+	−	−	+	−	−	−	−	+	+	+	−	+	+	+	+	+									
DUA-5 BSAgA	−	−	−	+	−	+	−	−	−	−	−	−	−	−	+	−	−	+	+	−	−	+	−	−									
XTR-22	−	−	+	−	+	+	+	−	+	−	+	+	−	+	−	+	−	−	+	+	+	+	+	−					+	−			
REX-12	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−							−	+	
ALR-1BSH6	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	−	−	+	+	+	−	+									+

[a] Karotypes of hybrids were prepared at the same time as DNA for hybridization experiments. The presence of CTRB in compared to the chromosome content of hybrids. Several of the hybrids were constructed from human cells containing translocations. TSL [46,XX,t(3;17)(p21;q13)]; XER [46,X,t(X;11)(q11;p11)]; EXR [46,X,t(X;11)(q22;q13)]; XTR [46,X,t(X;3)(q28;q21)]; REX [46,X,t(X;22)(q22;q13)]; and ALR [46,X,t(X;5)(q22;q35)].

made by using reverse transcriptase.[108,214,267] After synthesizing the homologous strand, the cDNA is inserted into a bacterial plasmid and then used to transform an appropriate bacterial host cell. Bacteria containing the chimeric plasmid are selected by a unique property, often resistance to certain antibiotics.[16]

cDNAs coding for proteins that are not usually expressed in interspecies hybrids have been isolated and used in gene mapping studies. These include β-globin on chromosome 11,[104,228] growth hormone on chromosome 17,[70,190] insulin on chromosome 11,[48,191] and the proopiocortin gene on chromosome 2.[193] Since many of these genes are members of multigene families, one may actually be mapping a cluster of genes. This is true for the interferons and β-globins. The probe used to map growth hormone also hybridizes with the gene coding for chorionic somatomammotropin, which shares greater than 90% homology at the mRNA level. Consequently, both genes were mapped together using the GH probe.[190] Human insulin, however, appears to be a true single-copy gene,[7] as had been suggested from genetic studies and amino acid sequencing data.

A fascinating complex has been isolated and characterized from the pituitary gland. The entire proopiocortin (POC) complex of peptides is contained within one precursor molecule.[172] Processing of the precursor molecule yields adrenocorticotropic hormone (ACTH) and β-lipotropin (βLPH). ACTH is further processed to α-melanocyte stimulating hormone (αMSH) and corticotropin-like intermediate lobe peptide (CLIP), whereas βLPH is processed to β-endophorin and α-lipotropin (αLPH), which in turn is cleaved to produce β-melanocyte stimulating hormone (βMSH). The *POC* gene complex is contained on one 13-kb EcoRI fragment of human DNA located on chromosome 2.[193]

We have mapped chymotrypsinogen B, one of a family of serine proteases that includes elastase and trypsin, to human chromosome 16 (Table I; Ref. 221). With the cDNA probes for trypsin and elastase, it will be interesting to determine if these genes are syntenic with CTRB. One or both genes could reside on another chromosome, a situation observed with prolactin, which shares about 40% homology at the mRNA level with GH and CSH but resides on chromosome 6.[192] It is worth noting that α-haptoglobin, which shares some homology with the serine proteases,[52] has been linked to chromosome 16 by family studies.[149,212] The haptoglobins are thought to have arisen from the serine proteases as a result of gene duplication and subsequent divergence.[126]

The occurrence of gene families and genes with partial homology can complicate mapping studies in hybrids with molecular probes. It is necessary for the investigator to determine by independent studies which band on a Southern blot corresponds to the gene that is to be mapped in hybrids. Otherwise, a related gene or pseudogene may actually be the sequence being mapped.

Representatives of several gene families and single-copy genes have been assigned using cloned probes and somatic cell hybrids (Table II). These genes are involved in several different areas of cellular function, including the translational apparatus ($tRNA_i^{met}$), signal transmission (*IFF, IFL, GH, CSH, POC, PRL, INS*), cellular homeostatis (*CTRB, HBA, HBB*), cell social behavior (*HLA*), cellular architecture (*COL1A2,* histones), and the immunological system (immunoglobulin genes).

In Situ Hybridization

The technique of *in situ* hybridization has been refined to the point where it can now be used to map single-copy genes.[71,84,150] In this method a radiolabeled RNA or DNA probe is hybridized to fixed metaphase chromosomes. The preparations are exposed to autoradiographic emulsion to detect silver grains formed as the result of β-emission from the radiolabeled probe. The grain distribution over the metaphase chromosomes is determined to localize the specific gene being mapped. The preparations are stained to facilitate identification of specific chromosomes.

The sensitivity of the method has been increased by a number of procedures. Malcolm *et al.*[150] used a 4.4-kb probe labeled to a specific activity of 1.7×10^8 dpm/μg to confirm the mapping of the β-globin gene to the short arm of chromosome 11. Probes of even higher specific activity (10^9 dpm/μg) were used along with dextran sulfate to confirm the mapping of the α-globin genes on chromosome 16.[71] In conjunction with larger probes (≥5 kb), dextran sulfate is thought to promote the formation of probe networks and thereby increase the hybridization signal, perhaps 20-fold.[71,84] In spite of this, hybridization efficiencies may be as low as 1–2%; and in an analysis of hybridization to the α-globin locus, Gerhard *et al.*[71] concluded that only 50% of the available sites had been saturated with probe. These results emphasize the necessity of analyzing many metaphase chromosomes to arrive at a statistically significant distribution of silver grains allowing a single-copy assignment, since in most cases only one or a few silver grains will be observed per chromosome. The

TABLE II. Cloned Genes Mapped to Human Chromosomes

Gene locus	Symbol	Chromosome	Reference	Polymorphic	Reference
Albumin	ALB	4q	84a		
Chorionic somatomammotropin (placental lactogen)	CSH	17q	70, 86, 190		200
Chymotrypsinogen B	CTRB	16	221	+	
α_2 Collagen I	COL1A2	7	107		
α-Globin	HBA	16p	4, 45, 71, 122	+	259
β-Globin	HBB	11p	46, 81, 83, 104, 139, 150, 228	+	109, 111, 197
γ-Globin	HBG	11p	81, 104, 139	+	203
δ-Globin	HBD	11p	81, 104, 139, 228		
Growth hormone	GH	17q	70, 86, 190	+	200
Growth hormone-like gene	GHL	17q	190		
Histocompatibility complex β locus	HLA	6p	100, 153	+	13
Histones	H1, H2B, H4	7q	27, 256		
	H2A, H3	7q	27, 256		
Immunoglobulin					
Kappa constant	IGKC	2	159		
Kappa variable	IGKV	2p	151		
Mu, gamma, alpha constant	IGHC	14	39, 251, 252		
Lambda constant	IGLC	22	159		
Insulin	INS	11p	48, 85, 140, 189, 191	+	8, 9, 127, 215, 264
Interferon					
Leukocyte	IFL	9	194		
Fibroblast	IFF	9	194		
Prolactin	PRL	6	192		
Proopiocortin	POC	2	193		
$tRNA_i^{met}$	TRM1	6	177		
	TRM2	6	177		

technique also requires careful attention in the preparation of metaphase chromosomes, which can affect the efficiency of hybridization.

Originally, the usefulness of *in situ* hybridization was limited to repetitive sequences, such as the 18S and 28S ribosomal DNAs that in humans are present at about 300 copies per diploid genome. The rDNA genes were located at the secondary constrictions of chromosomes 13, 14, 15, 21, and 22 in humans as tandem repeats of ten or more genes.[57,89] The 5S RNAs are also a repetitive cluster of genes localized by *in situ* hybridization to the long arm of chromosome 1.[105,257] In *Drosophila, in situ* hybridization has been a powerful tool for localizing genes on salivary gland polytene chromosomes, in which genetic loci have been amplified several hundredfold.[198]

The increasing number of single- and low-copy genes assigned by *in situ* hybridization attests to its usefulness as a gene mapping tool (Table III). It is interesting to point out, however, that except for an undefined DNA segment on chromosome 1,[84] no original assignment of a single- or low-copy gene has been made yet using *in situ* hybridization. It is important currently to use *in situ* hybridization in conjunction with somatic cell hybrids and Southern blotting to map cloned probes. For example, it might be difficult to determine by *in situ* hybridization the chromosomal distribution of related genes that are dispersed in the genome, such as the *GH-CSH-PRL* gene family[190,192] or the mouse α-like globin gene family, which includes two pseudogenes.[142,204] The chromosomal assignments of these gene families have been accomplished using somatic cell hybrids.

TABLE III. Genes Assigned by *in Situ* Hybridization

Chromosome	Gene	Symbol	Reference
1	14.9-kb DNA segment	D1S1	84
	5S RNA	RN5S	105, 257
2	Kappa light chain immunoglobulin	IGKC	151
4	Albumin	ALB	84a
7	Histones	H1, H2B, H4,	27, 256
		H2A, H3	27, 256
11	β-Globin	HBB	150
	Insulin	INS	85
16	α-Globin	HBA	4, 71
17	Growth hormone	GH	86
	Chorionic somatomammotropin	CSH	86
13, 14, 15, 21, 22	Ribosomal RNA	RNR	57, 89

Undefined DNA Segments as Genetic Markers

The function of the vast majority of human DNA is still undefined. Therefore, most of the clones that comprise a recombinant DNA library of the human genome[137] are simply "anonymous." They can be used, however, as genetic markers that when mapped can serve as landmarks to delineate a given region of a human chromosome.

Probes of undetermined coding capacity have been used successfully to define specific segments of human chromosome 11,[82] and the technology to isolate and map undefined probes is currently available.[42,82,112,135] Determining the presence of a human undefined DNA segment by hybridization to cell hybrid chromosomal DNA does not require a gene product to be expressed, which is necessary for current mapping studies with cell hybrids. Thus, not knowing the coding capacity of a single-copy probe does not diminish its utility or importance as a gene marker. Once an undefined probe has been mapped to a specific human chromosome, it becomes a useful marker for genetic studies. As a marker, the undefined probe should be as useful as protein products (e.g., G6PD) expressed in hybrids and serves as a reliable marker for regional mapping and linear ordering of chromosome regions.

Undefined DNA segments can be used to analyze chromosome breaks,[82] and have a great potential as tools to define the extent of chromosome deletions and translocations involved in inherited disease and neoplasia.[67,143,216] For example, specific interstitial chromosome deletions have been detected in some patients with Wilms' tumor (del [11p13])[67] and retinoblastoma (del [13q14]),[143] and the absence of these chromosomal regions could be determined using DNA probes. Furthermore, DNA probes might reveal smaller deletions or rearrangements in affected patients with no cytologically detectable chromosome abnormality, as has been done in the hemoglobinopathies.[63,163,184,185,186,259]

Sources of Undefined DNA Fragments

Complete genomic libraries are most often comprised of a collection of recombinant bacteriophage or cosmid clones containing DNA segments collectively representing an entire genome.[80,154] These libraries are important sources of defined and undefined DNA sequences for mapping and are also valuable for isolating regions of DNA adjacent to structural genes that may function in a regulatory manner.

Undefined DNA probes can be isolated from recombinant DNA libraries prepared from human genomic DNA,[137] from flow sorted chromosomes,[42,276] from somatic cell hybrids containing single or few human chromosomes,[42,82,112] and from cDNA libraries prepared from mRNA.[34] Complete libraries contain the entire human genome cloned into bacteriophage or cosmid vectors carrying 20–50-kb segments of DNA.[80,102,137] To be useful in mapping studies, probes must contain single- or low-copy sequences devoid of repetitive elements which are widely dispersed throughout the human genome.[93,227,261] Techniques for identifying single-copy sequences in recombinant bacteriophage or plasmid clones have been described.[82]

To facilitate the isolation of single-copy clones, libraries of short DNA pieces (1–2 kb) can be constructed.[225] A high proportion of clones from such a library should be suitable single-copy probes without further modification. An alternative approach is the construction of a cDNA library derived from mRNA. Many of the cDNA clones should reflect single- or low-copy gene sequences, although recently the existence of repetitive sequence elements in such clones has been reported.[34] The suitability of any undefined DNA fragment should be evaluated by blot hybridization to human DNA.

Several experimental approaches have been used to enrich for DNA sequences specific for individual chromosomes. Recombinant libraries prepared from cell hybrids containing a single human chromosome effectively isolate that set of human DNA from the rest of the genome, and offer a short-cut to obtaining chromosome-specific markers. Gusella *et al.*[82] constructed a DNA library from a human–hamster hybrid containing only human chromosome 11, and were able to identify and isolate several human-specific DNA probes from this library. Five of these cloned DNA probes could be regionally assigned to different segments of human chromosome 11. In view of the genetic importance of chromosome 11 (Wilms tumor, HBB, insulin), these DNA probes could be important markers in fine structural analysis of this chromosome. A similar approach has been used to construct recombinant libraries specific for the human X chromosome[22,226] and chromosomes 12[42] and 21.[112]

Using a different approach, Wolf *et al.*[273] made a partial library of cellular DNA from an individual polysomic for the X chromosome. Since the X chromosome was more abundant than all the other chromosomes, they were able to isolate several clones specific for the X chromosome. The library was assayed using ^{32}P-labeled plasmid clones to screen

BamHI-digested DNA from a mouse–human hybrid containing the X chromosome, and also one lacking the X. Since there are many chromosome anomalies producing polysomy, this method is applicable for the isolation of sequences specific for several chromosomes.

A technically difficult, if not expensive, approach utilizes chromosomes separated by flow sorting[23] to construct DNA libraries.[42,273] The majority of human chromosomes can be isolated in high purity by flow sorting,[23,42] and this brings within the realm of feasibility DNA libraries for each of the human chromosomes. Recombinant DNA libraries specific for chromosomes 22 and X have been reported using this methodology,[42,273] and undoubtedly more chromosomes will be cloned in the future.

Chromosome-specific DNA libraries are a short-cut to obtaining markers for a given chromosome, yet even with these probes it is a good idea to confirm their chromosomal location by mapping studies. This applies especially to clones that might belong to a dispersed family of related genes residing on several different chromosomes. The strategy for chromosomal assignment of undefined DNA segments is virtually identical to that for mapping cloned genes (Fig. 2). As with cloned genes, deletion chromosomes and translocations can be utilized for regional mapping of undefined DNA segments. An example of the mapping of an undefined DNA segment in hybrid cells is shown in Figure 4. HS-3, which detects a 9-kb EcoRI fragment, identified a DNA segment on chromosome 3. Using hybrid cells made with a human fibroblast containing chromosome 3 translocations, D3S1 (*D*NA; chromosome *3*; *S*egment *1*) was localized to the 3p21 → q21 region.[175]

Several undefined DNA segments have already been assigned to

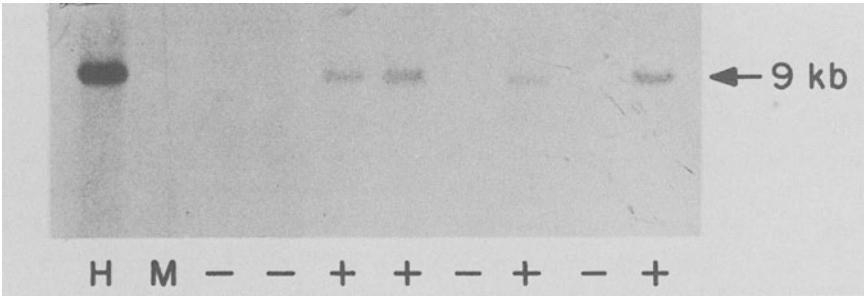

Fig. 4. Hybridization of a cloned, undefined DNA segment to human and hybrid DNA. The HS-3 probe detects a 9.0-kb band in human DNA digested with EcoRI but does not hybridize to mouse DNA under moderately stringent conditions. The human band was detected only in hybrids containing human chromosome 3.

TABLE IV. Chromosomal Localization of Undefined DNA Segments[a]

DNA segment[b]	Name[c]	Chromosome	References
D1S1	H3	1 (p36)	84
D1Z1	λHS3	1 (q12)	77
D6S1	B120	6	222
D10S1	P3-10	10	113
D11S2		11 (pter→p13)	82
D11S3		11 (p13→p1208)	82
D11S4		11 (p1208→p11)	82
D11S5		11 (p11→q13)	82
D11S6		11	82
D11S7		11	82
D11S8		11	82
D11S9		11	82
D11S10		11	82
D11S11		11	82
D12S1	L2.2	12	202
D14S2	L1.48	14	202
D16S2	L1.19	16	202
D19S1	3-10	19	
D20S1	18BA	20	22
D20S2	46EA	20	22
D22S1	3-2	22	114
DXS1	8EA	X (pter→q1)	22
DXS2	33HA	X (pter→q1)	22
DXS3	19HA	X (q1→q21)	22
DXS4	12EA	X (q21→q24)	22
DXS5	14BA	X (q21→q24)	22
DXS6	32EA	X (q2→qter)	22
DXS7	L1.28	X	202

[a] Undefined segments which have not *yet* been found to be polymorphic.
[b] The gene nomenclature for undefined DNA segments is *D*NA, chromosome number, *S*egment (or *Z* for repetitive), and sequential number.
[c] Designation of the laboratory that cloned the DNA fragment.

human chromosomes (Table IV), and this number will grow rapidly as more of the genome is dissected.

DNA Polymorphisms as Genetic Markers

Botstein *et al.*[19] proposed that undefined fragments of DNA that recognize restriction-site polymorphisms in human DNA could be used

as genetic markers to study the Mendelian inheritance of human genetic diseases. They reasoned that with a sufficient number of polymorphic DNA markers spaced at 20-cM intervals in the genome any genetic disease could be linked to one of them. By analyzing pedigrees in which polymorphic traits and inherited disease are segregating, it might be possible to map the disease loci by demonstrating linkage to a particular DNA polymorphism. This could be accomplished without any knowledge of the biochemical defects involved in a particular disease. Successful application of this system of linkage analysis would represent a monumental methodological advance, since only 2–3% of all Mendelian disorders have been chromosomally assigned.

Use of restriction endonucleases (>200 have now been identified[210]) and the detection of single-copy DNA sequences by Southern blotting, combined with recombinant DNA techniques, form the experimental basis for the construction of a genetic linkage map of man using a collection of random, polymorphic DNA fragments (see Ref. 19 for an extensive discussion).

Wyman and White[275] described the first highly polymorphic locus, *D14S1*, in human DNA using an arbitrary restriction fragment of DNA. At least eight alleles were observed, and >75% of individuals examined were heterozygotes. The DNA fragment length polymorphisms displayed Mendelian inheritance through a three-generation pedigree. Restriction enzyme analysis suggested that the polymorphism at this locus was due to DNA rearrangements rather than base-pair substitutions or modifications. This highly polymorphic locus is located on the q21 → qter region of human chromosome 14.[49]

DNA fragment length polymorphism due to DNA rearrangement approximately 700 bp upstream from the single human insulin gene has been extensively analyzed.[8,9] Restriction mapping and DNA sequencing studies indicate that the polymorphisms are due to insertions of DNA of two general size classes, 0–600 and 1600–2000 bp. No correlation was apparent that could distinguish normal individuals from individuals with insulin dependent (Type I) or non-insulin-dependent (Type II) diabetes, although such an association with Type II diabetes has been claimed.[215] Based upon analysis of two cloned allelic insulin genes, one of which possessed a 100-bp single-copy insert, Bell *et al.*[8] concluded that the smaller class of polymorphic fragments probably was not due to a transposable genetic element. Mechanisms other than those involving transposable elements have been suggested to account for highly polymorphic

loci in human DNA; for example, those observed in the α-globin gene complex.[90]

DNA Polymorphisms and Prenatal Diagnosis

In their original studies, Kan and Dozy[109] reported a polymorphism for an HpaI restriction enzyme site, about 5 kb from the 3′ end of the human β-globin gene. They observed that 87% of the β^S-containing chromosomes in their survey of American blacks lacked the HpaI site and the gene was present in a 13-kb fragment of DNA. In contrast, β^A-containing chromosomes were largely associated with a 7.6-kb or, less often, with a 7.0-kb fragment after HpaI digestion of DNA. Another group later found in a study of American blacks in the Eastern United States that the HpaI site was absent in 60% of β^S chromosomes and 8% of β^A chromosomes.[203] Prenatal diagnosis of sickle cell anemia using this linkage disequilibrium of the 13-kb fragment and β^S gene would be exact in only 36% of fetuses at risk (AS × AS parents), and sickle cell anemia could be excluded in an additional 24% of at-risk pregnancies where the β^S gene is segregating with the 13-kb fragment in only one parent.[203] By combining the HpaI polymorphism with a HindIII polymorphism located in the $^G\gamma$ and $^A\gamma$ intervening sequences, an exact diagnosis of sickle cell anemia prenatally can be increased to 80%.

More recently, Geever *et al.*[69] have demonstrated that the sickle cell gene can be detected directly using the restriction enzyme DdeI. The β^S mutation converts the codon for glutamic acid (GAG) to that for valine (GTG) at amino acid position 6, and changes a restriction endonuclease recognition site for DdeI (CTNAG), where N = any nucleotide. Thus the DdeI site is absent in the β^S gene and can be detected by Southern blotting. The method of detecting the β^S gene should therefore be applicable for all couples at risk and does not require family studies. Geever *et al.*[69] point out, however, that heterozygous HbC and nondeletion type β-thalassemia in a heterozygote for HbS, usually less severe disorders than sickle cell anemia, would give identical blot hybridization patterns as sickle cell trait; in these instances, family studies would be informative in distinguishing the disorders.

It is also possible to use DNA polymorphisms in prenatal diagnosis of β-thalassemia, in which β-globin synthesis is very low or absent, using a $^G\gamma$-globin gene probe[145] or with a β-globin gene probe.[111] Kan *et al.*[111] observed that the β^o-thalassemia gene in Sardinians was associated with

a 9.3-kb BamHI fragment, while in normal individuals the β-globin genes reside in a 22-kb BamHI fragment. This polymorphism was due to variation in a BamHI site on the 3′ side of the β-globin gene, and allows the presence of a β-thalassemia gene to be ruled out in Sardinians if a 22-kb fragment is present in the DNA cleaved with BamHI. This is in contrast to persons of West African origin, in whom the presence of a 13-kb HpaI fragment suggests the presence of a β^S gene.[109]

Thus, DNA polymorphic variants can be associated with the abnormal gene (13 kb; β^S) or the normal gene (22 kb; β^o-thalassemia) and depend upon the variant on the chromosome in which the mutation arose. In Sardinians, the β^o-thalassemia mutation arose on a chromosome carrying a 9.3-kb BamHI variant, while the sickle mutation in West Africans arose on a chromosome with the 13-kb variant.[111] DNA polymorphisms are useful, thus, not only in prenatal diagnosis, but also for anthropological studies.

Clearly, the power of prenatal diagnosis could be increased enormously if polymorphic DNA loci linked to genetic disease, such as cystic fibrosis, muscular dystrophy, or Huntington's disease, could be isolated. The methodologies reviewed here and being used in several laboratories should result in rapid progress toward this goal.

Identification of DNA Polymorphisms

Based upon the number of sequence polymorphisms detected in the β-globin cluster of man, it has been estimated that the human haploid genome would contain 3×10^7 sequence variants, or one every 100 bp.[103] Since there are 24 human chromosomes per haploid genome, there would be on the average approximately 10^6 sequence polymorphisms per chromosome. For example, Jeffreys[103] surveyed 60 individuals using eight restriction enzymes for polymorphism in the β-globin complex. He found three polymorphic fragments, one of which was found at a frequency of 0.23 in the population, and two of which were less common. Likewise, Kan and Dozy and colleagues have found two polymorphisms near the β-globin gene: One is segregated with the sickle cell trait in a West African population,[109] and the other can be used to monitor β-thalassemia in Sardinian populations.[111]

Only a fraction of single-base changes will create or destroy a restriction enzyme site. Even if only 10% of such sequence polymorphisms can be recognized by a restriction enzyme, there would still be 10^5 possible

polymorphisms per chromosome that could be used as markers. If the average chromosome contains 10^8 bp, a sequence polymorphism would occur for each kb of DNA. Hence, a good chance exists that a given undefined DNA fragment will identify a DNA polymorphism.

In order to determine whether a DNA marker is sufficiently polymorphic to be used in population studies, the DNA from cells of 10–20 individuals is tested for restriction site polymorphisms. This number of individuals should reveal the common polymorphisms. For screening purposes, DNA can be isolated from placentas, fibroblasts, and/or leukocytes. The restriction enzymes used for screening include six-base cutters (EcoRI, BamHI, and HindIII) and four-base cutters (TaqI and MspI). White and co-workers[5,225] reported that DNA polymorphisms were detected frequently with the latter two restriction enzymes. They proposed that the CpG dinucleotide in the recognition sequences for TaqI and MspI (5′TCGA3′ and 5′CCGG3′, respectively) is a mutational hot spot in human DNA. One to two percent of the cytosine residues in the CpG doublet in animal cells are methylated. 5-Me-C can be deaminated to thymine, and in bacteria this can produce base substitutions.[32]

DNA polymorphisms give rise to fragment length variations that can be visualized by Southern blotting.[8,275] The polymorphisms can arise from single-base-pair substitutions that create or destroy a restriction enzyme recognition site, from additions or deletions of DNA, or from other DNA rearrangements, such as inversions. Such DNA fragment length variants should segregate in a codominant Mendelian fashion, and a person could be homozygous or heterozygous, depending upon the variant carried by each chromosome homolog. Figure 5 shows a simple example of a DNA polymorphism detected in three individuals, one of whom is a heterozygote. As elegantly demonstrated by Kan and co-workers[109] with the sickle hemoglobin gene, this approach can form the basis for a genetic analysis of pedigrees in which an inherited disease is segregating.[30,207] A major advantage of using random DNA polymorphisms in genetic linkage studies is that the probes need not encode a known product, as was required in studies involving the β-hemoglobin locus.

The search for DNA polymorphisms can be facilitated by the somatic cell hybrid strategy. By mapping both specific and undefined probes, one can produce a map with many markers, each of which may be polymorphic either within the gene loci or in a neighboring area. By defining more markers, there is a much greater chance of detecting linkage to an inherited disease. In addition, our laboratory has mapped the genetic lo-

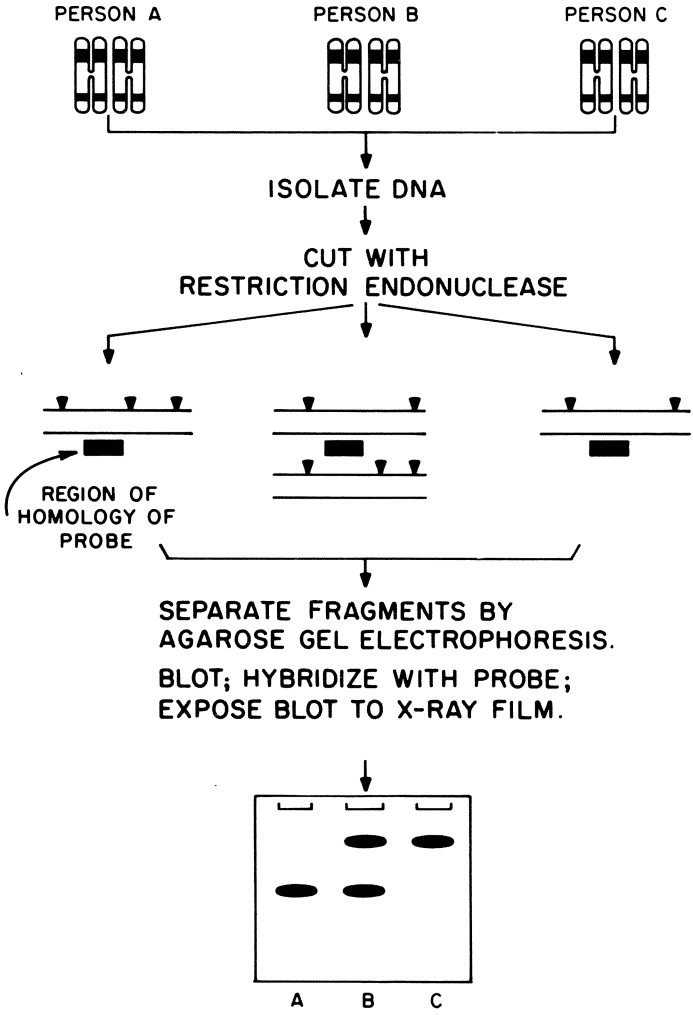

Fig. 5. Detection of DNA polymorphisms in humans. DNA is isolated from individuals, cleaved with a restriction endonuclease, and the fragments are blotted using the Southern procedure.[254] The blot is hybridized to a radiolabeled probe which detects a specific site in DNA (■). The normal distribution of restriction enzyme sites (▼) is shown for the DNA of person A. Person C possesses DNA which has lost one of the restriction sites due to mutation (e.g., single base substitution, or DNA rearrangement). Therefore the probe recognizes only a fast migrating band in homozygote person A, and a slow migrating band in homozygote person C. In contrast, person B is a heterozygote, since both fast and slow bands are detected, arising from the different chromosome homologs.

cation of 18 inherited metabolic diseases, and the chromosomal location of some 50 metabolic diseases (not including X-linked disease) are known (see section on Mapping of Disease). Identification of a polymorphic DNA sequence that occurs frequently in the population and that maps in the same region as an already mapped inherited metabolic disease locus would suggest the feasibility of using family studies to demonstrate the closeness of linkage between the two loci. Some of these diseases can be detected prenatally; however, in those instances in which the genotype of an individual is ambiguous, polymorphic DNA markers would greatly facilitate the reliability of the testing procedure. Table V lists undefined DNA polymorphic loci that have been mapped to date.

Using somatic cell hybrids, we recently localized an undefined genomic DNA fragment, D3S1, to human chromosome 3.[175] D3S1 detects a HindIII polymorphism in human DNA (Fig. 6). In our initial studies DNA from 15 individuals was surveyed: nine individuals had only a 26-kb band, while six had both a 26-kb and a 15-kb band. To characterize the polymorphism, somatic cell hybrids were constructed using human cells from an individual whose DNA gave the two-band pattern. The 15-kb band and the 26-kb band segregated independently in the hybrids, suggesting that they are allelic, i.e., reside on different chromosome homologs.

As an adjunct to mapping studies, any given cloned gene is also used to screen human DNAs for polymorphism. As already mentioned, CTRB

TABLE V. Chromosomal Location of Polymorphic Undefined DNA Segments

DNA segment[a]	Name[b]	Chromosome	References
D2S1	L2.3	2	202
D3S1	HS-3	3 (p21→q21)	175
D5S1	L1.7	5	202
D11S1		11 (p11→q13)	82
D14S1	1016-18 pAW101	14 (q21→qter)	49, 275
D16S1	L1.36	16	202
D17S1	p12-2	17	5
DXS8	RB6	X (q)	42
DXS9	RC8	X (p)	42

[a] The gene nomenclature for undefined DNA segments is *D*NA, chromosome number, *S*egment, and sequential number.
[b] Designation of the laboratory that cloned the DNA fragment.

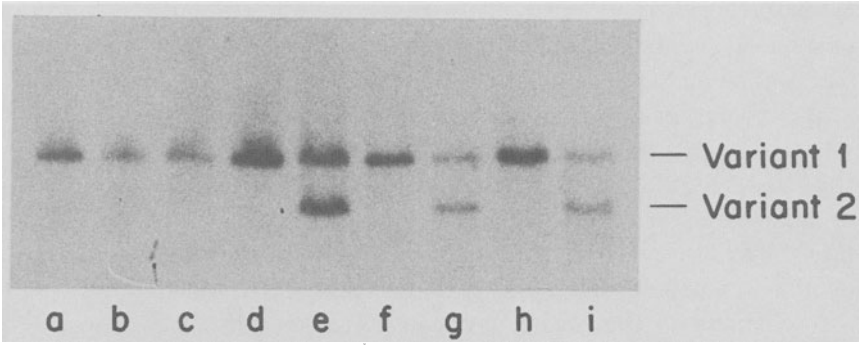

Fig. 6. Detection of a restriction fragment length polymorphism located on chromosome 3. Human DNA digested with HindIII yielded two patterns after hybridization with HS-3. Variant 1 is 26 kb, while variant 2 is 15 kb. Studies with somatic cell hybrids suggest that these are independent alleles.

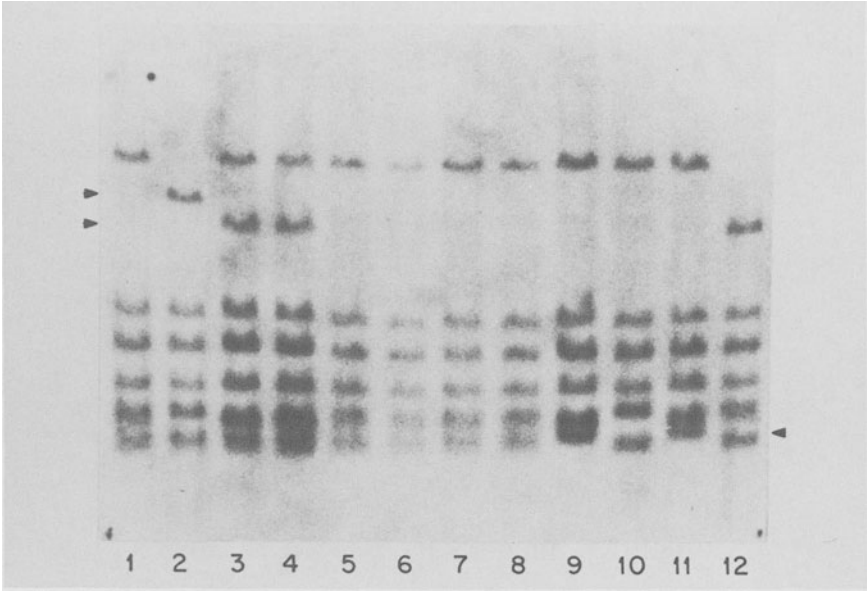

Fig. 7. Polymorphism of chymotrypsinogen B (CTRB) in human DNA. DNA was isolated from 12 unrelated individuals and cleaved with TaqI. The fragments were analyzed by Southern blotting using a ^{32}P-labeled rat CTRB probe. Blots were hybridized as described by Owerbach *et al.*[189] TaqI recognizes the sequence 5′TCGA3′. Arrows indicate fragment length variants detected by the CTRB probe.

was assigned to human chromosome 16 using somatic cell hybrids. Following the experimental paradigm of White and co-workers,[5,225] we screened 12 human DNAs digested with TaqI for CTRB polymorphism. Since TaqI is a four-base cutter (5′TCGA3′), a complex pattern was observed, with several fragment length polymorphisms apparent (Fig. 7). It is interesting to note that with EcoRI, the enzyme used in the mapping studies, CTRB recognizes a polymorphism in the two mouse cell lines used in the construction of hybrids (LM/TK$^-$ and RAG; see Fig. 3). Other cloned genes that detect DNA polymorphisms are listed in Table II.

CONSTRUCTING THE GENETIC MAP

The genetic map is composed of markers that represent expressed gene products, functions, and phenotypes. This is distinguished from the molecular map, which is being constructed by restriction mapping and sequencing an ever-increasing number of the genes themselves and undefined DNA fragments. For determining the genetic map, a rather large variety of markers has been described (Table VI). These markers, which represent a majority of cellular functions, are diverse and range from defined gene products to biochemically undefined syndromes, as can be observed in Table VI and Appendix A. In addition to their diversity, there are a large number of markers that are available for mapping studies, over 600 that have been or that will be mapped in the next year, and some additional 2000 inherited traits that will be more difficult to map.[161] For example, roughly 100 additional markers that have become available recently are the external membrane proteins and cytoplasmic proteins detected by 2D electrophoresis[33,160,188] as well as surface antigens detected by monoclonal antibodies.[75] Routinely, new classes of markers suitable for mapping become available as new techniques for testing are described. In many of these instances, additional polymorphisms will be identified for genetic family studies.

The three principal strategies for determining close linkage and assignment of genes to chromosomes center around different methodologies being used in family studies, somatic cell genetics, and *in situ* hybridization. These mapping procedures generally apply to contructing both genetic and molecular maps. The success of these rapidly developing and increasingly diverse methodologies in combination with classical genetic

TABLE VI. Classes of Markers (with Examples) Used in Linkage and Chromosome Assignment Studies

Classical family studies	Somatic cell genetics		*In situ* hybridization and flow sorting
Blood group antigens (ABO, RH)	Activators (plasminogen activator)	Mitochondrial markers (nuclear superoxide dismutase-2)	Cloned genes (*GH, INS, HBA*)
Cancer-associated (retinoblastoma)	Cell culture markers (polykarocytosis promoter, temperature-sensitive markers, transport markers)	Modifiers (adenosine deaminase complexing protein)	DNA polymorphisms
Chromosomal markers (uncoiler gene, deletions)	Cell receptors (epidermal growth factor receptor)	Nutritional auxotrophs (Gly complementing gene)	DNA segments (undefined)
Clotting factors (Factor XII)	Chromosomal proteins (histones)	Organelle markers (lysosomal enzymes)	rRNA
Cloned genes (insulin)	Cloned genes (*HBB, POC, GH*)	rRNA	
Complement components	Cytoplasmic proteins (2D gel markers)	tRNA	
Developmental markers (fetal Hb)	DNA polymorphism	Repetitive DNA sequences (DXZ1, satellite DNA)	
Diseases (defect unknown) (Prader–Willi syndrome)	DNA segments (undefined)	Ribosomal proteins	
DNA polymorphism	Drug sensitivity (ouabain)	Species-specific antigens	
DNA segments (undefined)	Enzymes (lactate dehydrogenase A)	Structural proteins (collagen)	
Enzymes (adenosine deaminase, esterase D)	Gene expression marker (TAT regulator gene)	Toxin sensitivity (diphtheria toxin)	
Histocompatibility markers (HLA)	Hormones (*PRL, CSH*)	Viral associated markers (human coronavirus sensitivity)	
Metabolic disease (defect known) (phenylketonuria)	Immunoglobulins		
Morphological traits (dermatoglyphics)	Membrane proteins (surface antigens)		
Mitochondrial markers (nuclear glutamic oxaloacetic transaminase-M)			
Serum proteins (albumin, haptoglobin)			
Sex-linked traits (G6PD-deficiency, color-blindness)			

approaches for human gene mapping is evidenced by the exponentially increasing genetic map of the human genome (Appendix A). Construction of a genetic map has proceeded rapidly over the last 10 years, stimulated by six biannual International Human Gene Mapping Workshops.[95–100] The evolution of the gene map is presented in those publications along with methodology and marker detection for defining the map. The gene map presented here is refined from those publications. Only selected aspects of contructing the map are described here. Additional features are found in the Workshop reports or numerous reviews.[37,162,207,218,234,235,240]

Close Linkage

There are several procedures to determine if genes are closely linked or adjacent to each other. If adjacent (not close enough to be linked) and on the same chromosome, they are said to be syntenic.[206]

Family Studies

When two loci are sufficiently close together on a chromosome such that the segregation of their alleles in family members is not random, these loci are said to be closely linked. The map distance between two loci is determined by the frequency of recombination between the loci, which is directly proportional to the distance between the loci. Polymorphic loci for these studies are identified in Table V and Appendix A. Establishing linkage of loci in man has encompassed a highly sophisticated statistical and computer technology taking into account the many parameters that can affect the segregation of traits. For a complete review of linkage analysis in man and polymorphic markers used for this analysis, see Conneally and Rivas.[30]

Unequal Crossing Over and Amino Acid Sequencing

The now classical example of unequal crossing over is the observation of hemoglobin Lepore involving the β and δ hemoglobin loci.[3] In family studies, recombination has not been observed for the β and δ loci, indicating close linkage.[168] Because of nonhomologous pairing, crossing over has occurred within nonhomologous but adjacent β and δ loci, resulting

in a new fusion protein—Hb Lepore. Subsequent amino acid sequencing of Hb Lepore revealed that the protein was composed of part of the β and δ peptides.[87] This finding demonstrates that the two genes are adjacent.

Gene Transfer

The ability to transfer small amounts of DNA to recipient somatic cells in culture facilitates structural mapping of genes closely linked on chromosomes. Chromosome-mediated gene transfer (CMGT) using isolating metaphase chromosomes[158,240] and isolated interphase DNA for DNA-mediated gene transfer (DMGT)[271] can be used in transferring a selectable gene and closely linked genes to cultured somatic cells with a growth requirement. With either technique, CMGT or DMGT, only a small portion of donor genetic material containing at least the selectable gene is ultimately stably retained in recipient cells with a growth requirement.[272] The closer the two genes are, the greater is the likelihood that they will both be transferred and stably retained in a recipient cell. The selectable marker thymidine kinase (TK) is very close to galactokinase (GALK) on human chromosome 17 in the q21 → 22 region.[56] When CMGT is used to transfer and select for *TK*, a number of recipient cells also retain the close marker GALK.[274] Since the band on chromosome 17 probably contains less than 5000 kb,[219] the *TK* and *GALK* genes are estimated to be no more than 5000 kb apart. Mouse ($HPRT^-$) recipient cells isolated after CMGT were found to contain human donor HPRT and a cytologically detectable fragment of the X chromosome in different recipient cells, an unusual occurrence, since donor genetic material usually has not been physically detectable after CMGT.[119,167] These X chromosome fragments were used to regionally localize the X-linked markers *HPRT, G6PD,* and *PGK*. These studies give a semiquantitative measure of physical distances between linked genes and allow gene order to be determined. The transfer of mammalian genes has been reviewed.[240]

A different method of determining close synteny and gene order of human genes has been reported by Goss and Harris,[78,79] and utilizes x-rays to induce chromosome breaks in human cells before fusion with mouse cells. This reduces the likelihood that any two syntenic or linked genes will be cotransferred, the probability being inversely proportional to the distance separating the genes. When irradiated human cells are

fused with mouse cells having a growth requirement and cell hybrids are isolated in selection medium, different sized fragments encoding the human selectable marker are retained. The cotransfer of syntenic genes gives results for linear order in good agreement with those obtained by linkage analysis.

Nucleotide Sequences and Restriction Mapping

Although the linkage of the members of the β-globin family was predicted earlier by fusion proteins (δ-β Lepore and $^{A}\gamma$-β Kenya) and family studies, the actual molecular relationships between the complex were determined by the analysis of cloned genes.[62,68] In 65 kb isolated as a set of overlapping DNA clones, the five β-like polypeptides were determined to be arranged 5′ε-$^{G}\gamma$-$^{A}\gamma$-δ-β3′.[68] In addition, two β-like pseudogenes were found: one between $^{A}\gamma$ and δ and the other 5′ to the ε gene.

Assignment of Genes to Chromosomes and Regions of Chromosomes

There are three major types of strategies, each with specific methodologies, for assigning genes to chromosomes and regions of chromosomes. Family studies represent the classical Mendelian way; somatic cell hybridization represents the parasexual approach; and Southern hybridization combined with cell hybrids and *in situ* hybridization represent the principal molecular approaches. A few biochemical–molecular strategies represent other specialized procedures. These are the principal technologies which have been employed to generate the human nuclear map in Appendix A.

Family Studies

Assignment of genes to specific chromosomes, using family study data, requires the segregation of traits with an identifiable chromosome. For X chromosomal inheritance, X-linked genes are assigned to the X by the characteristic father to daughter or mother to son pattern of inheritance over several generations. For autosomal loci it is necessary for a trait to segregate with a chromosomal variant in a pedigree. For example, a familial chromosome variation on chromosome 1, termed "uncoiler" and designated *1qh,* was used in family studies to show the first chro-

mosome assignment of an autosomal locus, the Duffy blood group.[51] Familial translocations were also used for chromosome assignment and regional location, as shown for the assignment of the α-haptoglobin locus (*HP*) on chromosome 16.[212] A fragile site on chromosome 16 was used to regionally map *HP* to the 16q22 region of the long arm.[149] This strategy does have drawbacks, since there is a need for dominant inheritance, and allelic variation of both the chromosome variation and the trait are necessary.

Gene dosage has also been used in family studies to assign genes to chromosomes. This involves quantitatively measuring a gene product in individuals with a chromosome deletion, or individuals with monosomies or trisomies. Deletion mapping has been employed to map red cell acid phosphatase (*ACP1*) to 2p[58] and adenylate kinase-1 (*AK1*) to 9q34.[59] Three copies of a gene in trisomies have also been employed to map a few genes, for example, superoxide dismutase-1 to the long arm of 21 in trisomy 21.[249]

Cell Hybridization for Gene Mapping

Development of the somatic cell hybrid strategy for mapping human genes has resulted in an exponential increase in the number of human genes assigned to specific chromosomes and regions of chromosomes. The mapping of defined DNA sequences using cloned DNA combined with cell hybrid strategy or isolated metaphase chromosomes allows one to predict the rapid development of a molecular human gene map as well. The somatic cell hybrid gene mapping methodology involves the fusion of human and rodent cells, creating proliferating hybrid somatic cells that contain reduced numbers of human chromosomes.[217,234,269] As described in Fig. 8, human cells (usually fibroblasts or leukocytes) are fused to mouse or Chinese hamster cells that possess selectable markers. Application of a selection system kills the rodent cells; the human cells employed either do not divide or divide slowly; and hybrid cells grow if the human gene that complements the rodent growth deficiency is retained. The most important fact for human genetic studies is the preferential loss of human chromosomes,[269] resulting in man–rodent cell hybrids that are characterized by different numbers and combinations of human chromosomes. Any human gene or product that is expressed in cultured cells (Table VI) and that can be distinguished from a rodent homologous gene or product can be assigned to a specific human chromosome and syntenic

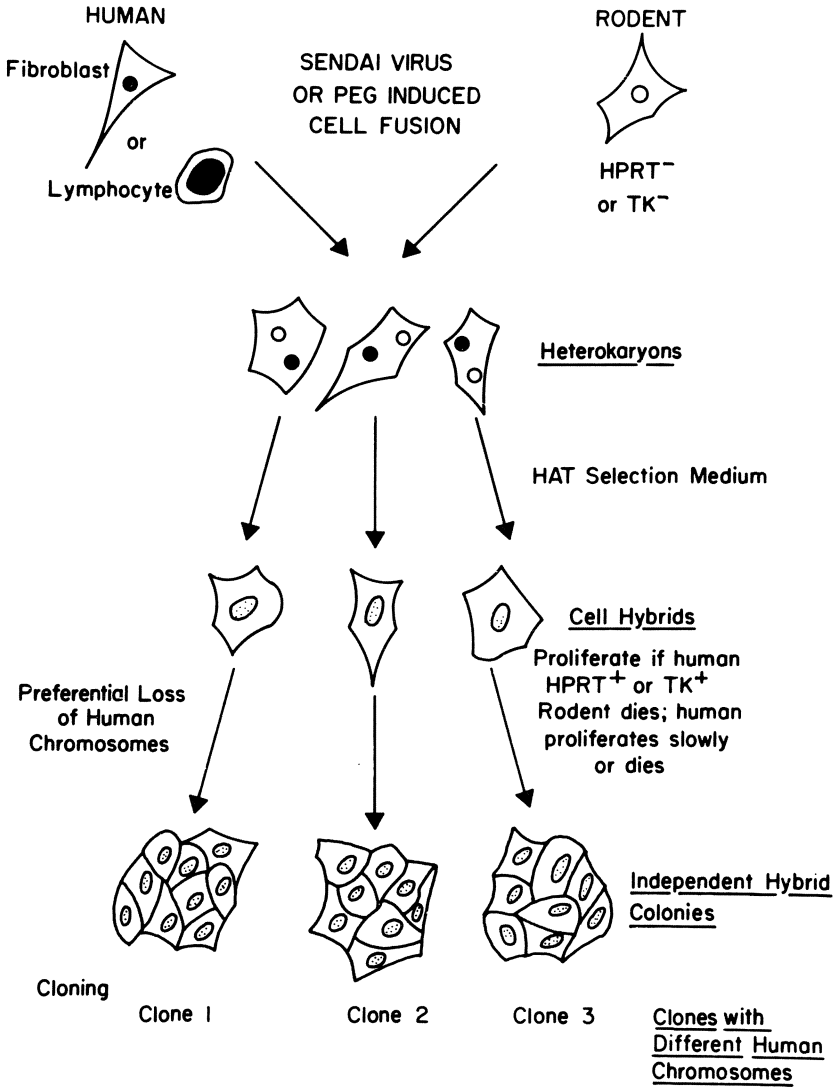

Fig. 8. Somatic cell hybridization scheme featuring the human gene mapping strategy. After cell fusion, human–mouse somatic cell hybrid clones are selected by the HAT selection system. These cell hybrids lose human chromosomes, making it possible to isolate clones with different numbers and combinations of human chromosomes. This partial human genome has made it possible to map human genes. (This figure is taken from Ref.[235] with permission of *Excerpta Medica*.)

group.[234,240] This is accomplished by correlating the presence of a human gene product with the retention of a specific human chromosome.[217,234] Using cells with human chromosome rearrangements "segregating" in cell hybrids permits regional chromosome assignments when a gene marker and a chromosome region are concordantly retained or lost.[234,235,237] This capability for mapping human genes parasexually is the result of a very timely culmination of three technologies: growth and selection of mammalian cells in culture, gene marker identification such as isozyme analyses, and chromosome identification and banding techniques to distinguish chromosomes within and between species. Several reviews have described in detail the technologies necessary to isolate man–rodent cell hybrids and to assign human genes and their products to specific chromosomes.[37,218,234] Chromosome and regional assignment is further optimized by assembling chromosome assignment panels of cell hybrids[36] and, likewise, regional assignment panels. The chromosome assignment panels consist of a minimum number of cell hydrids, each of which has a unique combination of human chromosomes. In such a panel, only a single chromosome unambiguously correlates with the distribution of the new marker in the hybrid panel. Thus, a gene is assigned to a specific chromosome. In order for a panel to be successful, a gene marker must be tested on samples correlated on the same cell passage with karyotype and isozyme marker analyses. A logical extension to this approach is to construct and isolate regional assignment panels where each cell hybrid in a set contains a different region of a specific chromosome. Twenty-four such panels of cell hybrids would contribute to the linear ordering of genes and to fine structural mapping.[234] The chromosome rearrangements necessary for these panels are being obtained from the human population as translocations or deletions, often involving selectable markers, or from human cultured cells that have been irradiated.[78,79] (A variety of chromosomal rearrangements in cultured cells can be obtained from the Mutant Human Cell Repository in Camden, N.J.)

Isolation of Cell Hybrids with Single Human Chromosomes. Obviously, a panel consisting of 24 cell hybrids, each with a single and different human chromosome, would be ideal for human gene mapping. If any human gene or product were expressed in a single human chromosome hybrid, then the marker would be encoded on this chromosome. Several such cell hybrids exist, but it is often necessary for the chromosome retained to encode a selectable marker, such as *HPRT* on the X or *TK* on chromosome 17. Some single-chromosome cell hybrids con-

tain a translocated chromosome involving an autosomal segment and a segment with a selectable marker, such as an X/15 or 3/17 translocation.[174,220] Isolation of single chromosome hybrids is usually generated in three ways: (a) subcloning cell hybrids containing few human chromosomes, no selectable marker involved; (b) subcloning and counterselecting hybrids with selectable markers; and (c) generating microcells with few human chromosomes for fusion to rodent cells.

Microcells are chromosomes which have been enclosed in nuclear and plasma membrane fragments and can be isolated by centrifugation of micronucleated cells in the presence of cytochalasin B.[55] Since microcells contain as few as a single chromosome, they can be used to transfer small numbers of chromosomes by fusion to recipient cells.[64] Microcells have been used to transfer mouse chromosomes containing the human HPRT gene to HPRT-deficient recipient cells.[65] Microcell hybrids were selected on HAT medium and contained from one to five donor mouse chromosomes. These results demonstrate the importance of the microcell methodology for transferring small numbers of intact chromosomes for chromosome assignments and rearranged chromosomes for regional assignments.[65,167] This topic has been reviewed.[240]

Simple Isozyme Markers for Identifying Each Human Chromosome in Cell Hybrids. Using the cell hybrid protocol, it is necessary to recognize the cosegregation of a specific human chromosome, with its biochemical or molecular markers, and the gene to be mapped. This requires identifying on the same cell passage each chromosome cytogenetically as well as the gene markers assigned to it. Karyotypic examination adds strength by demonstrating a specific human chromosome and whether broken chromosomes are present; if appropriate and well-known gene markers are chosen, a minimum of cytogenetic analysis may be necessary. This is acceptable, since stable markers have now been assigned to each chromosome by several laboratories and human chromosome dynamics in cell hybrids is not better understood, or if time is a factor and cytogenetic capabilities are not available.

We have found it convenient and efficient to use isozyme markers to rapidly map genes using cell hybrids.[234] If a catalog of isozyme markers whose genes are encoded on each chromosome is utilized, a large number of cell hybrids can be tested very rapidly for the human chromosomes retained. If isozyme markers encoded on each arm of a chromosome are well chosen, then breakage that may occur can be detected. Since chromosome analysis is time-consuming, we have chosen the isozymes listed

in Table VII for rapid mapping before confirming the results with chromosome analyses. If hybrid cells homogenized at 80 × 10^6 cells/ml in 0.05 M Tris, pH 7.5, are examined by starch-gel electrophoresis and specific histochemical staining,[88,229] as many as 50 hybrids can be examined for their human chromosome content in a week's time.[234] Table VII lists simple and repeatable isozyme markers for each chromosome,

TABLE VII. Enzyme Markers for Assigning Genes to Human Chromosomes

Chromosome	Arm	Enzyme marker[a]
1	p	Enolase-1 (ENO1)[164]
	p	Phosphoglucomutase-1 (PGM1)[242]
	q	Peptidase-C (PEPC)[144,245]
	q	Fumarate hydratase (FH)[266]
2	p	Malate dehydrogenase-1 (MDH1)[230]
	q	Isocitrate dehydrogenase-1 (IDH1)[230]
3	p	Aminoacylase-1 (ACY1)[174]
4	?	Peptidase-S (PEPS)[245]
	p?	Phosphoglucomutase-2 (PGM2)[242]
5	q	β-Hexosaminidase-B (HEXB)[129]
6	p	Glyoxalase I (GLO)[10,123]
	q	Superoxide dimutase-2 (SOD2)[21]
	q	Malic enzyme-1 (ME1)[243]
7	q	β-Glucuronidase (GUSB)[131]
8	p	Glutathione reductase (GSR)[187]
9	p	Aconitase-1 (ACO1)[247]
	q	Adenylate kinase-1 (AK1)[61]
10	q	Glutamic-oxaloacetic transaminase-1 (GOT1)[232]
11	p	Lactate dehydrogenase-A (LDHA)[239]
12	p	Triosephosphate isomerase-1 (TPI1)[125]
	q	Peptidase-B (PEPB)[144]
13	q	Esterase-D (ESD)[233]
14	q	Nucleoside phosphorylase (NP)[208]
15	q	Mannose phosphate isomerase (MPI)[129]
	q	Pyruvate kinase-M2 (PKM2)[231]
16	q	Adenosine phosphoribosyl transferase (APRT)[171,233]
17	q	Galactokinase (GALK)[174]
18	q	Peptidase-A (PEPA)[144,245]
19	?	Glucose phosphate isomerase (GPI)[50]
	?	Peptidase-D (PEPD)[157]
20	q	Adenosine deaminase (ADA)[120]
21	q	Superoxide dismutase-1 (SOD1)[21]
22	q	Aconitase-2 (ACO2)[247]
X	q	Glucose-6-phosphate dehydrogenase (G6PD)[241]
	q	Phosphoglycerate kinase (PGK)[28]

[a] Reference is to methodology used in detecting an enzyme. The assignment of the genes coding for these enzymes is given in the Human Gene Mapping Workshops.[95–100]

which apply to both man–mouse and man–Chinese hamster hybrids. All isozymes are examined by starch-gel electrophoresis, which required only one electrophoretic technique to be developed. By testing as few as nine starch gels made with different buffers and sliced in various horizontal or vertical cuts after electrophoresis, a total of 35 isozymes are scorable with different reaction mixtures specific for each isozyme (see Table VII footnote). A total of 0.4 ml of the cell homogenate is required for a single round of testing of all isozymes in Table VII. Thus, if one is interested in mapping a gene expressed in cell hybrids, this can be accomplished quite rapidly.

In Situ Hybridization and Flow Sorting for Assignment of Genes

The recent advances in *in situ* hybridization techniques has allowed the mapping of single-copy DNA sequences by this method. As indicated in previous sections, several genes have been mapped using this technique (see Table III). Application of recent high-resolution banding techniques may extend the power of *in situ* hybridization.[277]

The fluorescence-activated cell sorter has been successfully used to sort chromosomes.[23] Most chromosomes can be isolated to approximately 90% purity,[23,42] although it has been reported that chromosomes 10, 11, and 12 are difficult to resolve.[42] DNA has been isolated from sorted chromosomes and made into single-chromosome libraries.[42,276] In addition, DNA can be isolated from the fractionated chromosomes and then analyzed for the presence of a specific gene by restriction endonuclease digestion and Southern blotting. For example, confirmation of the location of *HBB* and *INS* to the short arm of chromosome 11 has been done using an X/11 translocation that could be distinguished from the normal chromosome 11 by flow sorting.[139,140] As more laboratories become competent in the art of separating chromosomes, flow sorting will become a significant means of mapping genes.

DISSECTION AND MAPPING OF HUMAN DISEASE

A genetic understanding of disease is gained by knowing where the affected gene is located and its relationship to other and adjacent genes. If a structural gene's function, expression, or product is dependent on additional genes, such as processing genes, temporal genes, or architec-

tural genes, then a mutation in any of these steps in expression could possibly result in a phenotype similar to one involving a structural gene defect. Mutations of one gene that affect the expression of another gene identify loci that function in the final realization of a structural gene product. Thus, if mutant genes involved in a sequence of physiological steps can be chromosomally located, it can be investigated whether diseases affecting the same tissue, pathway, or organelle or that result in similar pathology are the result of mutations to genes that are closely linked, clustered on the same chromosome, or under coordinate control. In situations where genes are closely linked and have very similar amino acid sequences and mRNA homology, such as growth hormone and chorionic somatomammotropin,[156,178,190] but which are expressed in different tissues, it is important to know if mutations in one gene or flanking sequence affect the other gene. Such information will lead to a better understanding of the different numbers and types of genes that are involved in molecular disease and the organization and expression of the human genome as well.

Mapping Inherited Disease

Many inherited diseases are known to have a defect in a specific enzyme, protein, or other biochemical property. Only under certain conditions can a disease of unknown etiology or even a structural gene mutant be assigned to a specific chromosome by family studies (above). However, the gene can be mapped using cell hybrids or *in situ* hybridization if a cloned probe is available or the gene's product is expressed in cultured cells. Since many inherited diseases are the result of a structural gene defect, mapping the normal structural gene identifies the chromosomal location of the disorder. Many of the inborn errors associated with a specific enzyme or molecular defect have been mapped (Table VIII). For example, we have utilized the cell hybrid strategy to map the structural genes of argininosuccinic aciduria (*ASL*),[173] galactose-4-epimerase deficiency (*GALE*),[11] G_{M1}-gangliosidosis (*GLB1*),[247] Tay–Sachs disease (*HEXA*),[128,129] mannosidosis (*MANB*),[26] Sandhoff–Jatzkewitz disease (*HEXB*),[128,130] metachromatic leukodystrophy (*ARSA*),[47] Wolman's disease (*LIPA*),[121] β-glucuronidase deficiency (*GUSB*),[131] Maroteaux–Lamy syndrome (*ARSB*),[47] growth hormone deficiency (*GH*),[190] and several others just to illustrate the straightforward and rapid approach to mapping

structural genes expressed in cell hybrids and associated with disease (Table VIII).

Disorders have been described and mapped (Table VIII) in amino acid, carbohydrate, lipid, mucopolysaccharide, and nucleic acid metabolism. Affected genes functioning in pathways such as the heme, glycolytic, and the citric acid cycle have been mapped. Genes expressed in various tissues or organelles such as the red cell and serum or the lysosome and involved in disease have been mapped. Disorders in immunodeficiency, endocrinopathies, hypercholesterolemia, and various diseases of unknown biochemical defect have also been mapped. To the extent that some viruses cause disease, several viral phenotypes have been mapped.

Although Table VIII represents a modest list of mapped diseases, it represents a remarkable achievement in a short period of time and allows one to predict that with the advent of cloned genes and DNA polymorphisms and through family and cell hybrid studies, such a list will rapidly increase in number and become much more diverse as well.

DNA Polymorphisms for Mapping Disease

While Table VIII compiles a sizable number of diseases that are now mapped, there are some 2000 inherited traits,[161] most rare, that have not been mapped; without the appropriately mapped polymorphic marker segregating in affected families, it would be difficult to do so. Previously the potentially large numbers of DNA polymorphisms were discussed and a strategy presented as to how they could be used to map virtually all disease. If a mutant gene is closely linked to one or several polymorphic loci, then in family studies it becomes possible to predict if an individual carries the altered allele. Additionally, if a chromosomally assigned polymorphism is found to be closely linked to a specific disease in family studies, then the altered gene is assigned to a specific chromosome. It appears to be possible to identify and assign DNA polymorphisms with several alleles at roughly equal intervals of, say, 20-cM on all human chromosomes. Soon, several markers in a catalog of assigned DNA polymorphisms (i.e., Table V) will be available for testing in each family with an inherited disorder, and a determination of close linkage can be ascertained for one or more polymorphisms. Once close linkage and assignment have been made and a genotype determined, then it will be

TABLE VIII. Mapping of Inherited Disease

Disorder		Locus	Chromosome assignment
A. Disorders in metabolism	*Enzyme affected*		
Amino acid metabolism:			
Argininosuccinic aciduria	Argininosuccinate lyase	*ASL*	7pter→q22
Citrullinemia	Argininosuccinate synthetase	*ASS*	9q34→qter
OTC deficiency	Ornithine transcarbamylase	*OTC*	X
Phenylketonuria	Phenylalanine hydroxylase	*PKU*	1
Tetrahydropteroyl-glutamate methyl-transferase deficiency	5-Methyltetrahydropterol-L-glutamate:L-homocysteine *S*-methyltransferase	*MTR*	1
Carbohydrate metabolism:			
Fucosidosis	α-Fucosidase	*FUCA*	1p34→p32
Galactokinase deficiency	Galactokinase	*GALK*	17q21→q22
Galactose-4-epimerase deficiency	Galactose-4-epimerase	*GALE*	1pter→p32
Galactosemia	Galactose-1-phosphate uridylyltransferase	*GALT*	9p
Mannosidosis	α-Mannosidase-B	*MANB*	19pter→q13
Pompe's disease	α-Glucosidase	*GAA*	17q21→q25
Lipid metabolism:			
Fabry's disease	α-Galactosidase	*GLA*	Xq21→q24
GM_1 gangliosidosis	β-Galactosidase	*GLB1*	3p21→q21
GM_2 gangliosidosis (Tay–Sachs)	β-Hexosaminidase-A	*HEXA*	15q22→q23
GM_2 gangliosidosis (Sandhoff's)	β-Hexosaminidase-A and -B	*HEXB*	5q11→q13
Lecithin:cholesterol acyl transferase deficiency	Lecithin:cholesterol acyltransferase	*LCAT*	16cen→q22
Metachromatic leukodystrophy	Arylsulfatase-A	*ARSA*	22q
Steroid sulfatase deficiency (ichthyosis)	Steroid sulfatase	*STS*	Xpter→p22
Wolman's disease	Acid lipase	*LIPA*	10
Mucopolysaccharide metabolism:			
β-Glucuronidase deficiency (MPS VII)	β-Glucuronidase	*GUSB*	7cen→q22
Hunter's disease	Iduronate sulfatase	*IDS*	X

Maroteaux–Lamy syndrome	Arylsulfatase-B	*ARSB*	5
Nucleic acid metabolism:			
Hyperuricemia	Adenine phosphoribosyl transferase	*APRT*	16q12→q22
Lesch–Nyhan syndrome	Hypoxanthine phosphoribosyl transferase	*HPRT*	Xq26→q28
Other enzyme deficiencies:			
Acid phosphatase deficiency (lysosome storage disease)	Acid phosphatase-2	*ACP2*	11p12→cen
Glutathione peroxidase deficiency	Glutathione peroxidase-1	*GPX1*	3p13→q12
B. Erythrocyte abnormalities	*Enzyme affected in glycolytic pathway*		
Hemolytic anemia:			
Glucosephosphate isomerase deficiency	Glucosephosphate isomerase	*GPI*	19pter→q13
G6PD deficiency	Glucose-6-phosphate dehydrogenase	*G6PD*	Xq28
Hexokinase deficiency	Hexokinase-1	*HK1*	10
Phosphogluconate dehydrogenase deficiency	6-Phosphogluconate dehydrogenase	*PGD*	1pter→p34
Phosphoglycerate kinase deficiency	Phosphoglycerate kinase	*PGK*	Xq13
Triosephosphate isomerase deficiency	Triosephosphate isomerase-1	*TPI1*	12p13
Heme and hemoglobin:	*Protein affected*		
Acute intermittent porphyria	Uroporphyrinogen I synthase	*UPS*	11q23→qter
Hemoglobinopathies (sickle cell and thalassemias)	α Chain	*HBA*	16pter→p11
	β Chain	*HBB*	11p1208→1205
Methemoglobinema	NADH-diaphorase	*DIA1*	22q13→qter
Blood group:	*Gene involved*		
Rh incompatibility	Rhesus blood group	*Rh*	1p36→p32
Other:	*Enzyme affected*		
Elliptocytosis		*EL1*	1p
Glutathione reductase deficiency	Glutathione reductase	*GSR*	8p21
Hemolytic disease	Adenylate kinase	*AK1*	9q34
C. Serum abnormalities	*Protein affected*		
Analbuminemia	Albumin	*ALB*	4q11→q13
Antithrombin III deficiency	Antithrombin III	*AT3*	1q23→q25
Hemophilia	Antihemophilic globulin A (factor VIII)	*HEMA*	Xq

(*Continued*)

TABLE VIII. (Continued)

Disorder		Locus	Chromosome assignment
D. Immune deficiency diseases	*Defect*		
C2 Deficiency	Complement component-2	*C2*	6p23→p21
C4 Deficiency	Complement component-4	*C4*	6p23→p21
Nucleoside phosphorylase	Nucleoside phosphorylase	*NP*	14q12→q20
Severe combined immunodeficiency	Adenosine deaminase	*ADA*	20q13.2→qter
E. Endocrinopathies	*Protein defect*		
Congenital adrenal hyperplasia	21-Hydroxylase deficiency	*CAH*	6p23→p21
Growth hormone deficiency	Growth hormone	*GH*	17q22→q24
Hyperinsulinemia	Insulin	*INS*	11p13→p15
F. Diseases (Biochemical defect unknown)			
Adrenoleukodystrophy		*ALD*	Xq28→qter
Cataract, zonular pulverulent		*CAE*	1
Color-blindness (deutan)		*CBD*	Xq
Color-blindness (protan)		*CBP*	Xq
Chronic granulomatous disease		*CGD*	X
Charcot–Marie–Tooth disease (slow conduction type)		*CMT1*	1
Dentinogenesis imperfecta		*DGI*	4
DiGeorge syndrome		*DGS*	22
Hypercholesterolemia		*HC*	6
Muscular dystrophy, Becker type		*MDB*	X
Muscular dystrophy, Duchenne type		*MDD*	Xp21
Nail-patella syndrome		*NPS1*	9q
Ocular albinism		*OA*	X
Prader–Willi syndrome		*PWS*	15q11

Retinoschisis	*RS*	X
Spinal cerebellar ataxia	*SCA1*	6
G. Malignant disorders		
Metastasis (fibronectin deficiency?)	*FN*	8, 11
Retinoblastoma–1	*RB1*	13q
Wilms' tumor—aniridia/ambiguous genitalia/mental retardation	*WAGR*	11p13
H. Virus properties (possibly involved in disease)		
Coronavirus sensitivity	*HCVS*	15q11→qter
Echo 11 sensitivity	*E11S*	19q
Epstein–Barr virus integration site	*EBV*	14
Herpes virus sensitivity	*HV1S*	3 or 11
Poliovirus sensitivity	*PVS*	19

possible to provide adequate genetic counseling and prenatal diagnosis for predicting affected individuals.

Most of the disorders in Table VIII are recessively inherited and the heterozygote is often difficult to detect. This can cast doubt on prenatal diagnosis and genetic counseling. Mapping a fragment length DNA polymorphism near affected genes, listed in Table VIII, will aid in the detection of these carriers. Fragment length polymorphisms also will be important for mapping of fairly common autosomal dominant diseases, such as Huntington's disease, neurofibromatosis, and polycystic kidney disease, which have yet to be assigned to a specific linkage group or chromosome.

Genetics of Metabolic Pathways

Table VIII lists the chromosome assignments of structural and other genes associated with inherited disease divided into various functional categories. None of the genes in lipid metabolism have been assigned to the same chromosome. Even those genes functioning in the same pathway, such as *GLA, GLB1, HEXA, HEXB,* and *ARSA,* are not located on the same chromosome. The same nonlinkage phenomenon is observed for disease of carbohydrate and mucopolysaccharide metabolism. The acid hydrolases *GLA, GLB1, HEXA, HEXB, ARSA, LIPA, FUCA, MANB, GAA, GUSB,* and *ARSB* are found in the lysosome, and several are fatal when deficient. Close linkage and assignment to the same chromosome would suggest that genetic organization and subcellular organelle localization were associated; however, Table VIII demonstrates lack of clustering for these genes. Although some of the genes are located on the same chromosome, they are not closely linked. Nonlinkage was also observed for amino acid disorders, nucleic acid metabolism, and other disorders with single-enzyme defects (Table VIII). Tissue-specific enzymes, such as the red cell glycolytic enzymes or serum proteins, do not appear to be linked. However, at the molecular level the genome is beginning to show some organizational characteristics. It is becoming apparent that many genes are members of multigene families. For example, certain families of genes appear to have arisen by tandem duplication, and often these genes are still adjacent. The β-globin and α-globin gene families have been described in previous sections. Other genes which have been duplicated and remain together are 18S and 28S rDNA,[270] 5S RNA,[105,257] histones,[116] interferons (several leukocyte and at least one

fibroblast type),[194] $tRNA_i^{met}$,[177] *GH-CSH-GHL*,[190] immunoglobulin gene clusters,[15] and the histocompatibility complex.[1] Other structurally related genes are located in different parts of the genome, such as the prolactin gene located on chromosome 6 separated from the structurally very similar *GH-CSH-GHL* complex on chromosome 17.[190,192] The regulation of related genes may become apparent at the molecular level when both structural and flanking regions have been sequenced. Common flanking sequences involved in control of gene expression might precede each related gene or those functioning in a metabolic pathway.

Genetic Dissection of Disease

Somatic Genetic Dissection

Mapping inherited disease when the altered gene is known is straightforward; however, identifying and mapping genes for a disorder of unknown biochemical etiology often requires several somatic cell genetic methodologies.

Genetic Complementation Studies. I-cell disease (ICD), or mucolipidosis II, will be used as an example of this approach. This disorder is a fatal childhood disease with a complex phenotype of lysosomal enzyme deficiencies in skin fibroblasts and elevated enzyme activities in culture media.[166] In addition, a majority of lysosomal enzymes in fibroblasts from affected individuals is altered electrophoretically.[24,91] In an effort to identify and map the gene(s) involved in ICD, ICD fibroblasts were fused with mouse cells and cell hybrids isolated.[25] In these hybrids the abnormal enzyme phenotypes observed in ICD cells for β-galactosidase, β-hexosaminidase-A and -B, α-galactosidase, arylsulfatase-A and -B, α-mannosidase, α-fucosidase, β-glucuronidase, and acid phosphatase were corrected in ICD × mouse hybrids[25]; that is, the normal human electrophoretic patterns were recovered. The multiple enzyme defects in ICD fibroblasts would suggest the mutation of a gene product important in the expression of a large number of lysosomal enzymes rather than structural gene mutations of all the individual lysosomal enzymes. Post-translational modification of the lysosomal enzymes is suggested from the hybrid data and from the restoration of normal electrophoretic phenotypes of the enzymes after neuraminidase treatment.[24] These results demonstrate that the mouse genome corrected the affected human enzyme

expression in cell hybrids and suggest that the ICD mutation is a processing defect common to the final expression of multiple lysosomal enzymes.

The heterogeneity of enzyme activities and electrophoretic mobilities observed in ICD fibroblasts suggested genetic heterogeneity. Genetic complementation studies to identify possible complementation groups were performed. Many different ICD fibroblasts were fused in 2 × 2 combinations, and multinucleated cells were isolated after a gravity sedimentation procedure.[115] Analysis of these cells for lysosomal enzyme activities and electrophoretic mobilities revealed at least two and possibly three complementation groups within I-cell disease.[248] Thus, at least two mutations, probably involving different genes, result in similar phenotypes. Using the same strategy, several genes have been identified that when altered result in sialidase deficiency.[92] In this way genetic heterogeneity and specific genes can be identified for mapping, genetic counseling, and medical treatment.

Cell Hybrid Studies. Often by fusing enzyme-deficient cells with rodent cells, characteristics of the disorder can be determined. In the above example using I-cell disease, the altered gene was determined to be involved in the processing of lysosomal enzymes by the recovery of normal expression in somatic cell hybrids. Thus, the structural genes of the affected lysosomal enzymes were not affected.

When Tay–Sachs disease (TSD) and Sandhoff–Jatzkewitz disease (SJD) fibroblasts were fused with mouse cells, different results were obtained. TSD fibroblasts, deficient in β-hexosaminidase α subunit (HEXA) known to be encoded on chromosome 15, were fused to mouse cells; the cell hybrids retaining chromosome 15 did not reexpress HEXB.[244] This result was also observed when SJD fibroblasts, deficient in the β-hexosaminidase β subunit (HEXB) encoded on chromosome 5, were fused to mouse cells.[244] If the enzyme deficiency is a structural gene defect, then this observation is expected. If an enzyme deficiency results from the involvement of another gene on a different chromosome, then this can be investigated in cell hybrids. In cell hybrids, if the chromosome carrying the known structural gene chromosome is retained and the affected enzyme continues to be deficient while all other human chromosomes segregate independently, then the mutant gene is most likely a structural gene defect, as is the case of TSD and SJD. However, when human chromosomes segregate and the affected enzyme is recovered, then a nonstructural gene may be responsible for the deficiency.

Dissection Using Recombinant DNA

Genetic analysis and understanding of many Mendelian disorders is hampered when no information is available about the precise molecular defects involved. This was true for some of the hemoglobinopathies until the genes of the α- and β-globin clusters could be isolated and analyzed by recombinant DNA techniques.[54,155]

Those disorders in which a deficiency of a gene product occurs, because of either suspected structural or regulatory mutation, are amenable to analysis by gene cloning methods. Clearly, once a gene is isolated, a comparison can be made between the normal and the defective gene by DNA sequencing. Knowing the actual DNA sequence of a gene might suggest modalities for prenatal diagnosis based upon restriction endonucleases. A clear example of this is the detection of the β^S gene using DdeI.[69] Furthermore, a structural gene probe allows neighboring regions of DNA to be tested for DNA polymorphisms.[8,109,207]

Some of the most important disorders, such as Tay–Sachs disease, involve enzymes coded for by rare messenger RNAs. The structural genes for these diseases would be difficult, therefore, to isolate. However, with several recent experimental strategies, it appears likely that genes coding for rare mRNAs will be isolated.[35,101,180,201,253,268] It would seem especially appropriate to apply cloning methods to isolate genes for lysosomal enzymes involved in many heritable, often fatal, disorders. Molecular analysis of these genes would help to clarify models for their regulation[195] and provide potentially useful probes for diagnostic purposes.

THE HUMAN GENE MAPS

The human nuclear genome is composed of 24 gene maps, one for each of the different chromosomes. The gene maps are presented in Appendix A. Each chromosome has at least three genes assigned to it, except the Y chromosome. Three chromosomes, 1, 6, and X, have more than 30 genes assigned to each. Most chromosomes encode a gene that is polymorphic in the human population or that is a useful marker in genetic studies. The assigned genes in Appendix A provide many reliable and useful markers for human genetic studies, and the usefulness of the map will increase in direct proportion to the number of markers mapped. As the number of polymorphic markers assigned increases, more inherited diseases linked to these polymorphic markers by family studies will be

assigned to specific chromosomes. Such genetic information concerning an inherited disease linked to polymorphic and mapped genes will be of considerable importance for determining the genotype in affected and heterozygous individuals, for diagnosing both pre- and post-natally, for genetic counseling, and for possible therapeutic treatment.

It is estimated that less than 2% of all human genes have been identified and that less than 1% have been assigned to specific chromosomes. More than 375 genes have been assigned and are listed by chromosome in Appendix A and alphabetically in Appendix B. Approximately 100 additional traits that have X-linked patterns of inheritance are not included in Appendices A and B because they are rare and have been catalogued elsewhere.[161] Although only a small fraction of the total number of human genes have been identified to date, the majority of these have been mapped within the last five years.

In contrast, the entire human mitochondrial genome has been mapped.[2] There are 37 genes encoded in a circular DNA molecule with several unique features that distinguish it from nuclear chromosomes. It will be necessary to understand the total human genome composed of nuclear and mitochondrial maps and how each works in relation to each other in order to understand human biology.

Markers for Mapping the Human Genome

The paucity of markers for mapping the human genome just 15 years ago has been transformed into a galaxy of markers that now promises to encompass the entire genome (Tables II–VIII). For previous Mendelian genetic studies it was necessary to rely on a relatively small number of polymorphic enzyme phenotypes, fairly common diseases, blood group antigens, morphological traits, and sex-linked phenotypes.[207] Using these markers and chromosomal abnormalities, it was possible to assign only genes responsible for phenotypes known to be X-linked or associated with a particular chromosomal rearrangement to specific chromosomes; the latter represented a very small number. An equally small number of genes were demonstrated to be closely linked by statistical studies and occasional molecular studies showing duplication of related genes and therefore close linkage. For gene mapping and gene linkage studies in families, it is necessary to have a large number of polymorphic loci to identify each linkage group or chromosome and, additionally, to have a

sufficient number of markers segregating in families to determine linkage for the trait in question. This source of markers does not exist now, but the identification of DNA polymorphisms (above) shows considerable promise, along with current reliable markers, for providing a large supply of the necessary polymorphic markers to adequately study inherited traits and molecular disease.

With the advent of parasexual genetics embodied in man–rodent somatic cell hybrids (Fig. 8), the usable gene markers for mapping human genes increased exponentially as a result of the more than 80 million years of evolutionary divergence between man and mouse or Chinese hamster.[234] Virtually every gene product that can be identified in cultured somatic cells can be distinguished as being human or rodent because of structural or other differences. As a result of the large number of markers, and the cell hybrid system to dissect and reconstruct the human genome, approximately 375 genes and DNA markers additional to those linked or assigned by Mendelian genetics have been mapped. Gene markers that have been used for human mapping studies represent a wide variety of gene functions (Table VI), depending on whether a gene is identified by somatic cell hybridization, biochemical–molecular, or recombinant DNA methodologies, for example. A large number of markers have been described, since man is the most observed and studied species. Many of the gene markers studied in families (Table VI) can be utilized in somatic cell hybrids. The notable exceptions include blood groups, various diseases, and morphological traits; yet many unique markers are expressed in cultured somatic cells and cell hybrids.

The regional assignments of the β-globin locus in 1979[81,104,228] and the insulin,[189] growth hormone,[190] and chorionic somatomammotropin[190] genes in 1980 were the result of combining restriction endonuclease, recombinant DNA, and somatic cell hybrid methodologies. This heralded a new era in human gene mapping and made possible an enormous reservoir of gene markers. In fact, by joining the methodologies of recombinant DNA and somatic cell hybridization with the isolation of single human chromosomes and family studies for identifying gene linkages, it will be possible to map virtually the majority of the human genome.

Currently the most common genes assigned to human chromosomes are those that code for enzymes or structural polypeptides.[234] Mapping these genes has been accomplished principally using the cell hybrid methodology, although these markers can be mapped using other procedures.

Gene linkage studies using family studies are usually restricted to markers that comprise the blood groups, isozymes, serum proteins, clotting factors, sex-linked traits, and diseases. To expand this capability, it is often necessary to combine several genetic methodologies to make an assignment; for example, blood groups not recognizable in cultured cells but known to be linked to enzyme markers or DNA polymorphisms from family studies can be assigned to chromosomes when the enzyme gene or DNA polymorphism is expressed and mapped in cell hybrids. Thus, most types of gene markers recognized in man have been mapped to a specific chromosome, demonstrating the efficiency and versatility of chromosome assignment techniques.

The 24 Nuclear Gene Maps

Chromosome assignments, regional locations, and linkage information of human genes are presented in Appendix A and are derived from six Human Gene Mapping Workshops.[95–100] The original reference for each gene assignment and regional localization has been cited in the Workshop reports. The genes are arranged by chromosome, listed alphabetically, and the proposed linear gene order given for each chromosome map. The polymorphic loci are indicated, representing variable regions of each chromosome and markers for Mendelian gene linkage and population studies. The most consistent smallest regional assignment is given based on studies using rearranged chromosomes or *in situ* hybridization. Whether the assignment is confirmed, provisional (P), tentative (T), or inconsistent (I) is indicated, but this rating does not necessarily relate to the status of the regional location on a particular chromosome. A confirmed assignment is indicated when two or more independent reports assign a gene to the same chromosome and region; a provisional assignment acknowledges an assignment from a single report; a tentative assignment represents results based on inference or tenuous data; an inconsistent assignment occurs when assignment results disagree. A separate dictionary (Appendix B) lists the genes alphabetically with their recommended symbol.

For the first time, all gene symbols have been standardized in accordance with gene nomenclature guidelines proposed by the Gene Mapping Workshop Nomenclature Committee[246] and accepted by a majority

vote in plenary session at the Human Gene Mapping Workshop 6.[100] It provides a uniform system for human gene nomenclature and is referred to as the International System for Human Gene Nomenclature (1979), or ISGN (1979). Amendments to this system have been published in Human Gene Mapping 6.[100] In order to adopt this system, it was necessary to change some gene symbols. However, most changes are minor and do not significantly alter the sense of previous versions of the symbols. Where major changes are made, the previously used symbol and additional symbols which have been used for a gene in the recent past are given.[238,238a,246]

Chromosomal Benchmarks

There are those genes that have proven to be particularly useful and important in human genetic studies, and serve as benchmarks for each chromosome. Those markers that have been particularly useful in human genetics will be identified and briefly discussed. The greatest number of genes have been assigned to chromosomes 1, 6, and the X. Other large chromosomes, namely 2, 3, 4, and 5, have fewer assigned genes than chromosomes 1, 6, and X and illustrate the discrepancy in the observed number of mapped genes and a chromosome's size, where size apparently reflects the relative DNA content. This unresolved observation is apparent for several chromosomes with a disproportionately low number of markers in relation to their size.

Chromosome 1. A variety of different gene markers have been assigned to the largest of human chromosomes. They include viral, blood group, several different types of disease, blood clotting, enzyme, DNA segment, and molecular markers. Many of the markers are variant (although some are rare) in the population, such as *AMY1, AMY2, AT3, CAE, Do, EL1, ENO1, FUCA, Fy, GALE, GDH, PEPC, PGD, PGM1, PKU1, Rh, RP1, SC,* and *UMPK*. With this array of markers segregating in families, extensive genetic studies at the Mendelian, somatic cell, and molecular levels illustrate the usefulness of an extensive map of each chromosome for use in molecular organization of the genome, genetic counseling, prenatal diagnosis, and possible gene therapy. Polymorphic markers exist at several locations on both long and short arms (see Appendix A), providing useful markers for gene linkage studies, that are likely to be segregating in families with an unknown inherited disease.

The short arm is the most extensively mapped and encodes useful and easily assayable markers for family studies, comprising *AMY, GDH, PGD, PGM1, Rh,* and *UMPK,* with *Fy* and *PEPC* being good markers for the long arm. A rather extensive and accurate gene order has been determined (Appendix A). For cultured somatic cells, *AK2, DIS1, ENO1, FH, FUCA, GALE, PEPC, PGD, PGM1,* and *UMPK* are excellent markers. These include markers on both arms and those which are polymorphic, a very useful characteristic for somatic cell genetic and cell hybrid studies. The extensive genetic information on this chromosome has allowed a probable gene order that will provide excellent benchmarks for genetic studies. A detailed account of the chromosome 1 map has been reported by Cook and Hamerton.[31]

Chromosome 2. In contrast, this second largest of human chromosomes possesses only four polymorphic markers: red cell acid phosphatase (*ACP1*) on the distal portion of the short arm; *D2S1,* a DNA polymorphism; *IGKC,* a κ-chain marker, and the Kidd blood group (*Jk*). Many genes on this chromosome are not useful in Mendelian genetic studies; however, cloned gene probes have been utilized for mapping *POC* (proopiocortin), *IGKC,* and *IGKV* (immunoglobulin κ-chain constant and variable regions). A gene order is developing for this chromosome through regional mapping. The *ACP1, IDH1, MDH1,* and *POC* markers are excellent for cell genetic studies.

Chromosome 3. The genes *ACY1* (aminoacylase-1), *GLB1* (β-galactosidase-1), *D3S1* (DNA restriction polymorphism), and *GPX1* (glutathione peroxidase-1) are the most useful on this chromosome for mapping studies in both family and cell genetic studies because of their reliable and relatively simple assay procedures. *GPX1* is polymorphic, and *GLB1* is the lysosomal enzyme deficient in G_{M1} gangliosidosis. *D3S1* is polymorphic and is located in the same chromosome region as *ACY1* and *GLB1*.[174,175] The transferrin (*TF*) locus has been suggested to be located on this chromosome by comparative gene mapping studies of aminoacylase and β-galactosidase in the mouse.[174] It is of interest that the transferrin receptor gene is also located on this chromosome.

Chromosome 4. *ALB* (albumin), *GC* (group-specific component), and *MNS* (blood groups) on the long arm in that order, and *PGM2* (phosphoglucomutase-2) and *PLG* (plasminogen), probably on the short arm, are all polymorphic and have been excellent markers for gene linkage studies. In cell genetic studies *PGM2* and *PEPS* are used to identify this chromosome. Although only a few genes have been assigned to this chro-

mosome relative to its size, there are numerous polymorphic loci available for family studies.

Chromosome 5. β-Hexosaminidase-B (*HEXB*) is deficient in Sandhoff–Jatzkewitz disease, and arylsulfatase-B (*ARSB*) is deficient in Maroteaux-Lamy syndrome. These are rare inherited disorders and are the few markers on this chromosome that are variant and useful in family studies. Markers for broad genetic studies have not been identified for this chromosome, and the linear order is tentative at this stage. The diphtheria toxin sensitivity locus is a good marker in somatic cell studies, and its function can be utilized in a selection system against chromosome 5. An undefined DNA fragment, *D5S1,* which promises to be a useful molecular marker, has been located on this chromosome and is polymorphic.

Chromosome 6. With the exception of chromosome 1, more markers have been assigned to this chromosome than to any other autosome. However, they are concentrated on the short arm (p) of the chromosome. The predominant group of markers on 6p is the major histocompatibility complex, which includes the *HLA* loci, the complement components, glyoxalase I (*GLO1*), and the properdin factor B (*BF*). These markers are highly polymorphic, which increases the possibility of finding heterozygosity in families segregating rare Mendelian traits; hence the preponderance of markers on the short arm. Rare traits that have been linked, or provisionally linked, to the *HLA* loci by family studies are congenital adrenal hyperplasia (*CAH*), and hypercholesterolemia (*HC*). This degree of polymorphism has allowed linear ordering of the genes by recombination studies in families and chromosome localization by chromosomal rearrangements in cell hybrid studies. Other polymorphic loci on chromosome 6 are phosphoglucomutase-3 (*PGM3*) and the P blood group. Reliable markers for somatic cell genetic studies are *GLO* and *HLA* loci on the short arm, and *PGM3, ME1* (malic enzyme-1), and *SOD2* (superoxide dismutase-2) on the long arm. A DNA segment, *D6S1,* and prolactin were assigned using cloned probes and the recombinant DNA strategy.[192,222] A further account of the chromosome 6 genes has been detailed.[100]

Chromosome 7. This chromosome encodes genes with diverse functions representative of many of the different activities of a cell. Markers encoded on 7 are involved in amino acid metabolism, nucleic acid metabolism, blood groups, structural proteins, membrane proteins, chromosomal proteins, hormones, mitochondria, surface antigens, the lysosome, and developmental markers. However, there is no polymorphic

marker on this chromosome, although other markers associated with rarer diseases are on this chromosome, such as mucopolysaccharidosis VII, a β-glucuronidase deficiency. For cell hybrid studies, *GUSB* (β-glucuronidase) and *ASL* (argininosuccinate lyase) are reliable markers that identify this chromosome.

Chromosome 8. Although chromosomes 8, 9, 10, 11, and 12 are similar in size, few informative markers have been assigned to chromosome 8, whereas the other chromosomes have 3–4 times the number of markers assigned. This chromosome does carry the polymorphic marker glutathione reductase (*GSR*), which is informative in family and cell genetic studies. A fibronectin marker (*FNS*) appears to control the retention of *FN* on the cell surface, which may be important in metastasizing cells, some of which do not express *FN*.[187] Another gene on chromosome 11 is involved also in the fibronectin phenotype.[118]

Chromosome 9. The important ABO blood group marker is located at the q terminus of this chromosome. For family studies, this marker and adenylate kinase-1 (*AK1*) are the most useful polymorphic markers on the long arm, while ACO1 (aconitase-1) and *GALT* (galactose-1-phosphate uridylyltransferase) are polymorphic in the population and encoded on the short arm. It is of evolutionary interest that both fibroblast and leukocyte interferon structural genes are located on this chromosome.[194] These and the *ACO1* and *ASS* (argininosuccinate synthetase) markers are very useful for cell culture studies.

Chromosome 10. Although this chromosome displays interesting markers that have been used in somatic cell genetic studies, only *GOT1* (glutamic-oxaloacetic transaminase, soluble form) has been used for linkage studies. Reliable and simple to assay markers that are used to identify the chromosome in cell genetic studies are *GOT1, HK1* (hexokinase-1), *LIPA* (acid lipase-A), and *PP* (pyrophosphatase). For cell genetic studies, there is a curious nonviral gene called *FUSE* involved in the important property of fusing cells.

Chromosome 11. The short arm of this chromosome is being extensively studied at the molecular level, since the hemoglobin ε, γ, δ, and β chains have been localized to the p1208 → p1205 region. The use of recombinant DNA technology (above) has determined the nucleotide sequence of DNA segments within this region and the linear order to be *HBE-HBG2-HBG1-HBD-HBB* reading from 5′ to 3′. In this same region of DNA, several restriction enzyme site polymorphisms have been

found.[109,145,203] These non-alpha-hemoglobin loci have been utilized extensively in family linkage studies over the last 25 years. The DNA polymorphisms will add a new dimension to these studies, including prenatal diagnosis. On the same arm as the non-alpha loci, the insulin structural gene has been assigned using recombinant DNA and somatic cell technology.[48,85,140,189,191] *INS* also appears to be near the *LDHA* (lactate dehydrogenase-A) and *ACP2* (acid phosphatase-2) loci.[191] A DNA polymorphism near the *INS* gene has been described by Bell *et al.*[8] Use of these DNA and protein polymorphisms should prove to be very informative for gene linkage studies and for prenatal diagnostic studies of chromosome 11 markers. Isolation of a large number of cloned DNA segments of chromosome 11 from a genomic library obtained from a cell hybrid retaining only human 11 has provided important markers for the linear order and sequence of the molecular map.[82] A deletion in 11p13 occurs in some patients with Wilms' tumor.[67]

Chromosome 12. Although genetic variants have been described for several of the loci encoded on this chromosome, for example, peptidase-B (*PEPB*) and lactate dehydrogenase-B (*LDHB*), none are polymorphic or have played an extensive role in family gene linkage studies. On this chromosome are such loci as *LDHB, PEPB,* and *D12S1,* which are excellent markers for cultured cells.

Chromosome 13. Few genes have been assigned to this acrocentric chromosome and, of these, esterase-D (*ESD*) is the only polymorphic marker in populations. *ESD* is also a useful marker for somatic cell studies. A deletion in 13q14 occurs in some retinoblastoma patients. Ribosomal DNA genes (*RNR*) have been located on the five acrocentric chromosomes, of which 13 is one. These rDNA genes are redundant on each acrocentric chromosome and possess, within a nontranscribed spacer region, a DNA polymorphism which has proven to be useful in understanding the organization of rDNA genes on the telomeres of these human chromosomes.[127,174]

Chromosome 14. On this acrocentric chromosome resides the rDNA cluster of repeated genes on the short arm and the nucleoside phosphorylase (*NP*) gene on the long arm. Recently the immunoglobulin heavy-chain markers[39,251,252] and a DNA polymorphism (*D14S1*) have been assigned,[49] both of which should be excellent markers for population and family studies. The α_1-antitrypsin (*PI*) marker has been assigned to this chromosome and is very useful in population studies. The reliable

NP marker has been used extensively to identify this chromosome in cultured cells. Although several genes are now assigned to this chromosome, their linear order is not well defined.

Chromosome 15. The rDNA cluster of genes is encoded on the short arm, with several useful markers on the long arm. Of these, *HEXA*, altered in Tay–Sachs disease, and β_2-microglobulin are currently the most important markers for investigating molecular disease on the long arm of this chromosome. The Prader–Willi syndrome has been assigned also to the long arm (q11). An interesting biological marker determining human coronavirus 229E sensitivity in human–rodent hybrids has been located on the long arm of chromosome 15.[220] Heavily used cell genetic markers for this chromosome include *HEXA*, *MPI* (mannose phosphate isomerase), and *PKM2* [pyruvate kinase (M2)], all on the long arm.

Chromosome 16. The important hemoglobin α locus (*HBA*), the haptoglobin locus (*HP*), and *GOT2* (glutamic oxaloacetic transaminase, mitochondrial form) are polymorphic loci and have been used extensively in human family studies. The cloned chymotrypsinogen B gene (*CTRB*) was recently found to recognize a DNA polymorphism and should be useful in family studies.[221] In cell culture studies, the *APRT* locus is most frequently used and separate selection systems are available to select for and against this gene and chromosome. Excellent markers are available for both arms of this chromosome.

Chromosome 17. This chromosome encodes the significant genes thymidine kinase (*TK*), growth hormone (*GH*), and chorionic somatomammotropin hormone (*CSH*) and the polymorphic *GAA* (acid α-glucosidase) locus. The closely related and linked *GH* and *CSH* genes arose by gene duplication, and these loci have been regionally localized by *in situ* hybridization using cloned probes.[86] In cell genetic studies the *TK* gene can be selected for and against,[146] and the closely syntenic locus *GALK* (galactokinase) is widely used to identify 17.

Chromosome 18. Surprisingly, few genes have been assigned to this chromosome, with peptidase-A being the most useful marker in some populations and in somatic genetic studies.

Chromosome 19. Compared to chromosome 18, this smaller chromosome encodes a relatively large number of genes. For Mendelian genetic studies, *C3* (complement component-3) and proline dipeptidase (*PEPD*) are very useful markers. For cell genetic studies, *GPI* (glucose phosphate isomerase) and lysosomal α-mannosidase-B (*MANB*) associated with mannosidosis are important markers. It is interesting to note

that a gene controlling sensitivity of human cells to poliovirus infection is encoded on this chromosome.

Chromosome 20. One of the most useful polymorphic markers for population studies is red cell adenosine deaminase (*ADA*), which is encoded on 20. This marker also identifies this chromosome in cell genetic studies. Deficiency of this gene is associated with one form of severe combined immunodeficiency disease (SCID). The multiple endocrine neoplasia-2 (*MEN2*) disorder and three undefined DNA segments have been assigned to this chromosome recently.

Chromosome 21. One polymorphic locus, *SOD1* (superoxide dismutase-1, soluble form), has been assigned to this chromosome. An interferon receptor (*IFRC*) and *SOD1* are frequently used markers on this chromosome for cell genetic studies.

Chromosome 22. This small acrocentric chromosome also encodes an rDNA gene cluster on its short arm. Although no polymorphic markers have been assigned to this chromosome, arylsulfatase-A (*ARSA*), associated with metachromatic leukodystrophy, and NADH diaphorase (*DIA1*, cytochrome b_5 reductase) are encoded on this chromosome; each is variant in the population, but not at the level of a polymorphism. Both markers can be recognized in cell genetic studies. Noteworthy new assignments are markers for the immunoglobulin λ-chain constant and variable regions (*IGLC, IGLV*) and the DiGeorge syndrome (*DGS*).

X Chromosome. There are over 100 X-linked traits listed in the McKusick catalog.[161] Of these, the most used polymorphic markers in family studies are glucose-6-phosphate dehydrogenase (*G6PD*), the Xg blood group, and hypoxanthine phosphoribosyl transferase (*HPRT*). At the somatic cell level, *G6PD, HPRT,* phosphoglycerate kinase (*PGK*), and α-galactosidase (*GLA*) have been the most significant in mapping studies. The *HPRT* locus can be selected for and against in cultured cells using the HAT/8-azaguanine counterselection system,[260] which has made it possible to select for the X-chromosome and, particularly, a region on the X encoding the *HPRT* locus.[236,237] The X chromosome encodes a wide variety of genes involved in disease, such as clotting factors, muscular dystrophy, color-blindness, enzymes in purine metabolism, immunodeficiences, and disorders of the eye, nervous system, and skin. Other markers are Y histocompatibility antigens, blood groups, and surface antigens. At the molecular level, several DNA segments have been mapped, which will contribute to a molecular description of this chromosome and its genes altered in disease.

The Y Chromosome. The male-determining Y chromosome is known to encode the Y histocompatibility antigen.

It is clear that the number of genes assigned to each chromosome is not proportional to chromosomal size; this relationship does not follow a Poisson distribution.[231] However, all chromosomes encode reliable markers for somatic cell genetic studies except the Y, and 84% of human chromosomes encode a polymorphic marker suitable for gene linkage and population studies. Most chromosomes encode markers associated with rare genetic variation that also could be used in genetic studies. Although the number of genes mapped is small compared with the total, those that have been mapped represent an extremely useful set of benchmarks for human family, population, and somatic cell genetic studies.

Unassigned Linkage Groups

Currently there are 11 linkage groups that are not assigned to a specific chromosome (Appendix C). Closeness of the linkage and the linkage methods are detailed in Human Gene Mapping 6.[100] The gene loci code for markers that are recognized generally in family studies and have not been associated with a specific chromosome using abnormal chromosomes or are not expressed in cultured cells for cell hybrid mapping.

Chromosome assignment of these unassigned markers must await linkage of at least one of the markers in the linkage groups to an already assigned marker or to a known chromosomal rearrangement. Assignment of these linkage groups represents an immediate addition of several loci to a chromosome. Such assignments would bring valuable linkage data to the gene map.

The Mitochondrial Gene Map

The human mitochondrial (mt) genome has been sequenced in its entirety through a collaborative effort of many researchers at the MRC Laboratory at Cambridge[2] (Fig. 9). The 16,529 base pairs sequenced revealed several features unique to mitochondrial DNA. Human mt DNA has been streamlined to a minimal amount of DNA, unlike DNA from mitochondria of other organisms, such as yeast.[17] Contained on one circular molecule are 37 genes including two for rRNAs, 22 for tRNAs, and 13 for proteins. Very few or no noncoding sequences were found between each of these transcribed regions. Both rRNA genes are very small, and

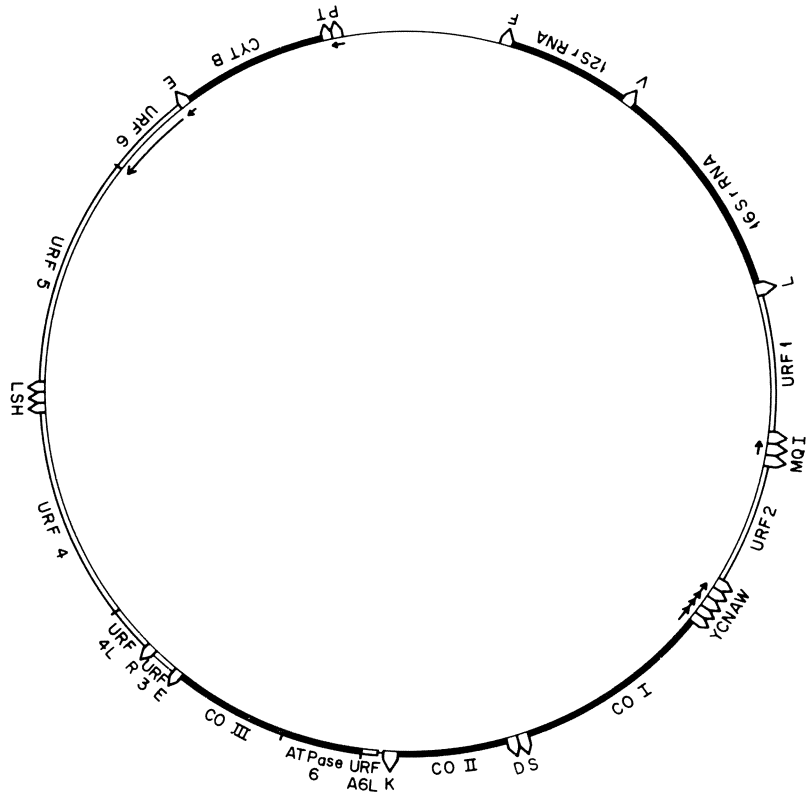

Fig. 9. The gene map of human mitochondrial DNA. The 16,529 base pairs of mt DNA are represented diagrammatically. There is an economy of material; very few noncoding sequences were found. ⌂ indicates locations of the 22 tRNA molecules, denoted by the one-letter symbol for each amino acid. Co I, II, III indicate the three subunits of cytochrome C oxidase; cyt b is cytochrome b. URF 1, 2, 3, 4L, 4, 5, 6, A6L are the eight unidentified reading frames. Also shown are the loci for 12S and 16S rRNA and ATPase 6. Except for the noted regions, transcription occurs clockwise. (Adapted from Anderson *et al.*[2])

5S rRNA found in cellular ribosomes is lacking in the mitochondrial ribosomes. In contrast to the approximately 3000 species of tRNAs found in the nucleus, only 22 tRNAs were found in mt DNA, although these 22 tRNAs were capable of reading all codons in the translation of mRNA. The sets of two codons in the mRNA that signal for the same amino acid are recognized by one anticodon on the tRNA that displays G:U wobble. Even more unusual is the single tRNA that functions in the recognition of four codons by having U in the first position of the anticodon.

Messenger RNA was found for 13 proteins, eight of which code for unidentified polypeptides.[2,169] The genes for these mRNAs have been reduced, so that there are no untranslated leader and trailing regions as found in nuclear genes. In addition, a complete stop codon is not present in some transcribed regions but is formed by polyadenylation. Only one promoter was found for each strand, indicating that differential expression of message does not occur at the level of transcription. Other mechanisms such as stability of the message may control the expression of various transcripts. Curiously, human genomic mt DNA, unlike other species, does not contain intervening sequences in the coding regions.[2,17]

The genetic code for human mt DNA is different from that found for other genomes. Specifically, UGA codes for tryptophan (not stop), AGA and AGG code for stop (not arginine), AUA for methionine (not isoleucine), and AUA and AUU both could initiate translation in addition to AUG.

The DNA sequence of the mitochondrial genome has indicated that perhaps different evolutionary constraints have operated on it compared to the nuclear genome. This is even more striking in view of the fact that both genomes have occupied the same environment, i.e., the mammalian cell, for millions of years.

COMPARATIVE GENETICS OF THE HUMAN GENOME

Ohno[182] postulated that the genes located on the X chromosome would be retained as a conserved linkage group throughout evolution. Because of the selection pressure imposed on these genes by X inactivation, this particular group of genes would remain together for reasons of dosage compensation. An overwhelming amount of evidence now supports this hypothesis.[95–100] Several genes (*G6PD, HPRT, GLA, PGK*) are linked in many species of mammals, ranging from *Mus muculus* to the primates (Table IX).

A few years ago another observation was made: autosomal linkage groups are also conserved over large spans of time.[66,132,133] This fact was not so surprising in primates since their chromosomal banding patterns are very similar.[44,53] Through systematic analysis of chromosomal rearrangements, the phylogeny of the primates has been traced and mimics the phylogenetic tree produced by classifying various physical charac-

TABLE IX. Comparison of Human Chromosomal Gene Maps with Those of Other Species[a]

Chromosome 1 / Marker	Chimpanzee (PTR 2)	Gorilla (GGO 1)	Orangutan (PPY 1)	Rhesus monkey (MML 1)	African green monkey (CAE 6 + CAE 1)	Baboon (PPA 1)	*Cebus capuchan* (CCA 15)	*Microcebus murinus* (MMU 3)	Sheep	Cow	Cat	Rat	Chinese hamster	Mouse
ENO1	+	+	+	+	+[6]	+	+	+	+	+	+[C2]	+[5]	+[2]	+[4]
PGD	+	+	+	+	+[6]	+	+	+	+	+		+[5]	+[2]	+[4]
GDH														+[4]
FUCA	+				+[6]	+								+[4]
AK2	+				+[6]	+						+[5]		+[4]
PGM1	+	+	+	+	+[6]	+	+	+		−	+[C2]	+[5]	+[2]	
GUK		+	+	+	+[6]	+	+							
AMY1														+[3]
AMY2														+[3]
PEPC	+	+	+		+[1]	+	+[U]							−[1]
FH	+	+	+	+	+[1]		+[U]							
RN5S	+	+	+		+[1]	+								

(*Continued*)

TABLE IX. (Continued)

Chromosome 2

Marker	Chimpanzee (PTR 12 + PTR 13)	Gorilla (GGO 11 + GGO 12)	Orangutan (PPY 12 + PPY 11)	Rhesus monkey (MML 9 + MML 15)	Cat	Mouse
ACP1	$+^{12}$		$+^{12}$		$+^{A3}$	$-^{12}$
MDH1	$+^{12}$	$+^{11}$	$+^{12}$	$+^{9}$	$+^{A3}$	
GLAT	$+^{12}$					
IDH1	$+^{13}$	$+^{12}$	$+^{11}$	$+^{15}$	$-^{C1}$	$-^{1}$
GALT	$+^{13}$					

Chromosome 6

Marker	Chimpanzee (PTR 5)	Gorilla (GGO 5)	Orangutan (PPY 5)	Rhesus monkey (MML 2)	Baboon	*Cebus capuchan* (CCA 3)	*Microcebus murinus* (MMU 6)	Sheep	Cow	Dog	Cat	Mouse
MHC	+	+		+	+			+		+		$+^{17}$
GLO1	+										$+^{B2}$	$+^{17}$
PGM3	+	+		+		+		+	+	+	$+^{B2}$	$+^{9}$
ME1	+				+	+	+	+	+		$+^{B2}$	$+^{9}$
SOD2	+	+	+	+					+		$+^{B2}$	$+^{9}$

Chromosome 12

Marker	Chimpanzee (PTR 10)	Gorilla (GGO 10)	Orangutan (PPY 9)	Rhesus monkey (MML 12)	African green monkey (CAE 13)	*Cebus capuchan* (CCA 10)	*Microcebus murinus* (MMU 7)	Sheep	Cow	Pig	Cat	Mouse	Rabbit
LDHB	+	+	+	+	+	+	+	+	+	−	+[B4]		+[4]
TPI1	+	+	+	+	+	+	+	+	+		+[B4]	+[6]	+[4]
GAPD	+	+	+	+	+						+[B4]	+[6]	
PEPB	+	+	+	+	+	+	+	+	+	−	+[B4]	+[10]	
ENO2	+						+						
CS				+	+		+					+[10]	

Chromosome 15

Marker	Chimpanzee (PTR 16)	Gorilla (GGO 15)	Orangutan (PPY 16)	Rhesus monkey (MML 7)	African green monkey	Baboon	*Cebus capuchan*	*Microcebus murinus*	Sheep	Cow	Pig	Cat	Mouse
RNR	−		+										
SORD													+[2]
IDH2													−[7]
B2M													+[2]
MPI	+		+	+								+[B3]	+[9]
PKM2	+	+	+	+	+	+	+	+	−	−	+	+[B3]	+[9]
HEXA	+	+	+	+	+	+	+	+	−	−	+	+[B3]	

(*Continued*)

TABLE IX. (Continued)

Chromosome 17

Marker						
TK1	Mouse	Chinese hamster	African green monkey	Rhesus monkey	Orangutan (PPY 19)	Gorilla (GGO 19)	Chimpanzee (PTR 19)
GALK	$+^{11}$ $+^{11}$	$+^{7}$ $+^{7}$	+ +	+	+	+ +	+ +

X Chromosome

Marker																		
PGK	Kangaroo	Muntjac	Rabbit	Mouse	Chinese hamster	Rat	Horse	Cat	Dog	Pig	Cow	Sheep	*Microcebus murinus*	African green monkey	Rhesus monkey	Orangutan	Gorilla	Chimpanzee
GLA	+		+	+	+	+				+	+	+	+	+			+	+
G6PD			+	+	+	+				+	+	+	+	+	+	+	+	+
HPRT	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
HEMA									+									
TFM				+														
OTC				+														

[a] Superscript numbers refer to the chromosome to which the linkage group is assigned. U: Syntenic, but unassigned to a specific chromosome. Minus sign: Not syntenic with other marker(s). Data compiled from Ref. 44, 53, and 95–100.

teristics. As the evolutionary distance between species increases, a concomitant increase in the number of chromosomal rearrangements is observed.[44] As expected, groups of genes are also linked in several primate species (Table IX). The rearrangements found in the linkage groups are a reflection of chromosomal rearrangements which have taken place in the speciation process.[44,53] For example, the long and short arms of human chromosome 2 have two separate chromosomal counterparts in other primate species.[53]

Several autosomal linkage groups have remained together after even longer periods of evolution evident in present-day mammals as common linkage groups. For example, the short arm of human chromosome 1 contains markers (*ENO1, PGD, AK2*) that are also linked in such divergent species as mouse, monkey, sheep, and rat (Table IX).

Many other markers which are syntenic in man are also syntenic in other species. The most extensive genetic studies have been with *Mus musculus*. As seen in Table X, 14 autosomal syntenic groups of man show at least partial homology in mouse. Often a human syntenic group will be divided into smaller groups located on different chromosomes in other species. For example, the human chromosome 10 markers *PP, HK1, GOTS, LIPA* are located on two chromosomes in mouse: *PP*(*Pyp-1*) and *HK1*(*Hk-1*) on mouse chromosome 10 and *GOTS*(*Got-1*) and *LIPA*(*Lip-1*) on mouse chromosome 19. These data suggest that in humans, *PP* is located closer to *HK1* and *LIPA* is near *GOTS*.[121] Other species also show this conservation of linkage groups; for example, the retention of the *LDHB-PEPB* gene synteny is seen in many species, except pigs and mice (see Table IX). Likewise, *PKM2* and *MPI* are syntenic in several species, but a splitting of these two genes has occurred in the ruminatory ungulates (Table IX).

Rigorous criteria for establishing homology of gene products between species has been established in the Gene Mapping Workshops[100] so that only those genes that are truly homologous will be compared. These criteria include similar amino acid sequence, immunological cross-reactivity, formation of heteropolymers of subunits between species, similar tissue distribution, similar developmental induction time, similar response to specific inhibitors, similar pleiotropic effects, and molecular hybridization to the same probe. These criteria become very important when the comparative maps are used to predict the location of human genes.

Comparison of gene maps at the molecular level also gives an indication of the divergence between species.[6,54] In most species there is a

TABLE X. Comparison of Mouse and Human Autosomal Gene Maps[a]

Human chromosome	Locus[b]	Mouse chromosome
1p	*ENO* (*Eno-1*)	4
	GDH (*Gpd-1*)	4
	PGD (*Pgd*)	4
	FUCA (*Fuca*)	4
	AK2 (*Ak-2*)	4
	PGM1 (*Pgm-2*)	4
	AMY1 (*Amy-1*)	3
	AMY2 (*Amy-2*)	3
1q	*PEPC* (*Dip-2*)	1
3	*ACY1* (*Acy-1*)	9
	GLB1 (*Bgs*)	9
	TF (*Trf*)	9
4	*PGM2* (*Pgm-1*)	5
	PEPS (*Pep-7*)	5
	ALB (*Alb-1*)	5
6	*HLA* (*H-2*)	17
	C4 (*Ss*)	17
	GLO1 (*Glo-1*)	17
	PGM3 (*Pgm-3*)	9
	ME1 (*Mod-1*)	9
	SOD2 (*Sod-2*)	9
7	*GUSB* (*Gus*)	5
	MDH2 (*Mor-1*)	5
	ASL (*Asl*)	5
10	*PP* (*Pyp-1*)	10
	HK1 (*Hk-1*)	10
	GOT1 (*Got-1*)	19
	LIPA (*Lip-1*)	19
11p	*LDHA* (*Ldh-1*)	7
	HBB (*Hbb*)	7
	CAT (*Cs-1*)	2
	ACP2 (*Acp-2*)	2
12	*LDHB* (*Ldh-2*)	-
	TPI (*Tpi*)	6
	GAPD (*Gapd*)	6
	PEPB (*Pep-2*)	10
	CS (*Cis*)	10
15	*B2M* (*Bam*)	2
	SORD (*Sdh-1*)	2
	MPI (*Mpi-1*)	9
	PKM2 (*Pk-3*)	9
	IDH2 (*Idh-2*)	7
16	*CTRB* (*Prt-2*)	8
	GOT2 (*Got-2*)	8

(*Continued*)

TABLE X. (Continued)

Human chromosome	Locus[b]	Mouse chromosome
	APRT (*Aprt*)	8
	HBA (*Hba*)	11
	HPA (*Hpa*)	11
17	*TK1* (*Tk-1*)	11
	GALK (*Glk*)	11
19	*GPI* (*Gpi-1*)	7
	PEPD (*Pep-4*)	7
21	*SOD1* (*Sod-1*)	16
	IFRC (*Ifrc*)	16
	PRGS (*Prgs*)	16
22	*DIA1* (*Dia-1*)	15
	ARSA (*As-1*)	15

[a] Compiled from Ref. 44, 53, and 95–100.
[b] The human gene nomenclature is followed by the mouse nomenclature in parentheses.

single insulin protein consisting of the A and B chains.[7] It is presumed that there is a single gene which codes for insulin, and, in fact, the DNA sequence of human insulin shows only one gene.[7] However, in mouse, rat, and at least two species of fish, there are two forms of the insulin molecule.[148] The two nonallelic rat preproinsulin genes have been isolated and sequenced by Lomedico *et al.*[148] The similarity of the two rat insulin genes had led the authors to hypothesize that this gene duplication is a recent one.[148]

Likewise, the evolution of the primate β-globin gene cluster has been studied by recombinant DNA methods. In humans the β-globin cluster consists of the sequence $5'\epsilon^{G}\gamma^{A}\gamma\delta\beta 3'$; in addition, two pseudogenes for β-globin were found.[68] The restriction maps for gorilla and the yellow baboon are indistinguishable from that of man, although no pseudogenes were found.[6] A new-world monkey, the owl monkey, has only one γ-globin gene, suggesting that the γ-globin gene duplication occurred ~20–40 million years ago.[6] Even more distantly related to man is the brown lemur, which has only the ε, γ, and β genes and a fusion product of ε and β.

In comparing the genome of one species to another, the unexpected does occur. Now that molecular probes are available for many genes, often a nontranscribed duplicate gene or pseudogene is discovered in the

genome which has partial homology to a given probe. For example, in the human β-globin complex there are also two β pseudogenes.[68] From this and other data, it seemed as though pseudogenes were located in a cluster with their transcribed counterparts. In humans, a pseudogene was found near the α-globin complex, as expected.[155] However, when the mouse was analyzed for the location of the two α-globin pseudogenes (α3 and α4), it was found that unlike the transcribed genes on chromosome 11, the mouse α3 was on chromosome 15 and α4 on chromosome 17.[142,204] This observation points out that interspecies comparison using molecular hybridization must be analyzed and interpreted very carefully.

The reason why certain linkage groups are retained through long periods of evolution while others are not is not yet clear. Whether accidental or because of some functional grouping, the identification of conserved linkage groups will (i) give clues about the organization of the genome, (ii) suggest mechanisms in evolution, and (iii) may predict the location of human genes.

The comparison of gene linkage groups between species may lead to the prediction of the location of genes. For example, we have mapped the genes for β-galactosidase and aminoacylase-1 to the p21 → q21 region of chromosome 3.[174] In addition, in collaboration with P. Lalley,[100] we assigned the gene for Acy-1 to chromosome 9 in the mouse. The mouse β-galactosidase gene has also been mapped to chromosome 9. Between these two genes lies the gene for transferrin in the mouse.[174] We would predict, therefore, that transferrin also is located on chromosome 3 in humans. Since we have mapped the DNA polymorphism D3S1 to the same region of chromosome 3, we are currently testing families for linkage of D3S1 and transferrin.

Caution must be exercised in predicting gene location, as all genes used in the comparison must be homologous. Even when these requirements are met, it is always possible that a translocation may have occurred near these genes. With these precautions in mind, however, comparative mapping should be a useful tool for mapping the human genome.

MERGING THE GENETIC AND MOLECULAR MAPS

More than 60 cloned DNA segments have been assigned in humans (see Table XI). Cloned DNA segments have been mapped to all but a few chromosomes, and polymorphic loci (at the DNA level) are located on

TABLE XI. The Human Molecular Map[a]

Chromosome	DNA marker
1	D1S1[b]; D1Z1[c]
2	POC[c]; IGKC[c]; IGKV[b]; D2S1[c]
3	D3S1[c]; D3S2[c]
4	ALB[b]
5	D5S1[c]
6	PRL[c]; TRM1[c]; TRM2[c]; HLA[c]; D6S1[c]
7	COL1A2[c]; H1, H2B, H4, H2A, H3[b]
8	
9	IFF[c]; IFL[c]
10	D1OS1[c]
11	HBB[b,c,d,e]; HBD[c]; HBG[c]; INS[b,c,e]; D11S1–D11S11[f]
12	D12S1[c]
13	RNR[b,c]
14	IGHC[c]; D14S1[c]; D14S2[c]; RNR[b,c]
15	RNR[b,c]
16	CTRB[c]; HBA[b,c,d]; D16S1[c]; D16S2[c]
17	CSH[b,c]; GH[c]; GHL[c]; D17S1[c]
18	
19	D19S1[c]
20	D20S1, D20S2, D20S3[f]
21	RNR[b,c]
22	IGLC[c]; IGLV[c]; D22S1[c]; RNR[b,c]
X	DXS1–DXS7[f]; DXS8[e]; DXS9[e]; DXZ1[c]

[a] Sequenced or potential for being sequenced.
[b] *In situ* hybridization.
[c] Somatic cell hybrids with restriction endonucleases.
[d] Liquid hybridization.
[e] Flow sorting.
[f] Library made from hybrid cell.

most chromosomes (Tables IV and V). Thus, each mapped DNA segment can be sequenced to form a portion of the molecular map. Eventually the genetic map will be superimposed on the nucleotide backbone constituting the molecular map.

It is estimated that 1 cM, a unit of genetic recombination, represents 1×10^6 bp.[206] This is enough genetic material to code for perhaps several hundred genes. Thus, Mendelian genetics and cytogenetics describe the human genome at the 100-gene level. In contrast, the molecular geneticist generates a high-resolution map measured in kilobases. As the human genome contains 33 M, it is apparent that the genetic map and the molecular map describe gene clusters by converging along the chromosome from two different levels of resolution.

The genetic map as produced by sexual and parasexual studies and other methods has provided hundreds of chromosomal markers, some of which are useful for family studies. The molecular map should complement and extend the usefulness of the genetic map in its application to human genetic disease. The use of DNA fragment length polymorphisms for diagnosis of human disease (Table V) was the result of dissection of genes at the molecular level. The current genetic map will be required to chromosomally assign and assemble a useful library of DNA polymorphisms to be used predictively in genetic counseling.[19] As already described, previously assigned chromosomal markers and somatic cell hybrid techniques will expedite the mapping of undefined DNA polymorphisms. To test for Mendelian segregation of DNA polymorphisms and their usefulness as markers, pedigrees that have been well characterized for polymorphic traits will be required.[19,30] It seemed necessary and inevitable for the genetic and molecular approaches to meld, since it has been difficult with family studies to assign to linkage groups diseases such as cystic fibrosis or Huntington's disease. In fact, only a very small minority of inherited diseases have been mapped.

The methods used for determining the genetic map often reveal very little about the molecular nature of the gene involved in the phenotype being observed. Consequently, little information can be obtained from the genetic map about the molecular mechanisms of expression of disease. For example, until the β-globin and α-globin gene clusters could be analyzed at the molecular level, it was unknown that certain hemoglobinopathies were due to actual gene deletions[184,185,186] or that a form of β^+-thalassemia is due to a base substitution in one of the introns of the β-globin gene.[255]

Several features of the human genome could not have been predicted from the Mendelian genetic map, including astonishing discoveries concerning gene structure. The idea that the amino acid sequence of a polypeptide is always colinear with the nucleotide sequence encoding it was firmly laid to rest in 1977 with the discovery of splicing and intervening sequences.[12,20,73] It is now known that the coding sequence of some genes contains introns,[12] whereas others do not.[138] Speculation concerning the role of introns in gene structure, function, and evolution has been made[38,72] and tested experimentally.[148,223,258] Another unexpected finding was the occurrence of pseudogenes found associated with[68] and dissociated from[142,204] their homologous gene cluster.

The molecular map seems to exhibit more fluidity than the "stable"

genetic map. The occurrence of highly polymorphic regions of DNA, perhaps associated with transposable genetic elements, indicates that the mammalian genome is more dynamic than originally thought.[8,275]

CONCLUSIONS

The techniques of DNA cloning and sequencing will eventually make it possible to determine the entire sequence of the human genome. With the rapid advance of technology and the associated concepts, the achievement of this goal is now merely an exercise based on sheer effort that should eventually yield the entire sequence. Interpreting the meaning of the total sequence is a much more difficult question and one to which we may not have the answers until long after the sequence has been determined. For example, knowing the total sequence might still not answer how the genome functions as a unit. In the meantime, the genetic and molecular maps are rapidly being described at an exponential rate. In fact, the molecular map will soon overtake the genetic map in numbers of markers assigned when considering the exceedingly large numbers of undefined DNA segments that are available for mapping. Although not crucial, it may be some time before we can delineate these DNA segments into coding regions that define a function. What is important is that this reservoir of markers will include enough DNA polymorphisms to enable the human geneticist to assign inherited disease to linkage groups and specific chromosome locations. Thus, the human gene map has reached the point, with enough critical information, where it can be used for its predictive value in counseling and diagnosis, for identifying disease loci, and on a larger scale for understanding human biology. As more specific gene loci are sequenced, it is anticipated that the map will yield answers to models of gene regulation which will enhance disease therapy and treatment.[40]

It should be realized that utilizing the human genome, as in gene transfer strategies, has enormous potential in several areas of biology, such as in repair of genetic disease, in food production, in producing therapeutic agents, and in energy production and industrial processes.[240] This concept is not biologically unrealistic, since transfer of DNA is evolutionarily very old, occurring naturally in diverse organisms. Obviously, many obstacles must be surmounted before successful gene ther-

apy becomes a reality in humans[170]; the most formidable is whether the transferred gene can be expressed normally.

When comparing the gene maps of man and several species, it is apparent that understanding the human genome allows us to comprehend our biological past. At the same time, understanding our genome will allow us to look to the future for comprehending both normal and abnormal biology and for treating human disease.

APPENDIX A. Chromosome Assignment of Human Gene Markers[a]

Chromosome	Gene symbol[b]	Gene marker	Polymorphic[c]	E.C. No.	Smallest region[d]	Status[e]
1	*A12M1*	Adenovirus-12 chromosome modification site 1C			q42→q43	P
	A12M2	Adenovirus-12 chromosome modification site 1A			p36	P
	A12M3	Adenovirus-12 chromosome modification site 1B			q21	P
	AK2	Adenylate kinase-2		2.7.4.3	pter→p32	
	AMY1	α-Amylase (salivary)		3.2.1.1	p22.1→q11	
	AMY2	α-Amylase (pancreatic)	+	3.2.1.1	p22.1→q11	
	AT3	Antithrombin III			q23→q25	P
	CAE	Cataract, zonular pulverulent (Fy-linked)				
	CMT1	Charcot–Marie–Tooth disease (slow conduction type)				
	D1S1	DNA segment			p36	P
	D1Z1	DNA satellite 3			q12	P
	Do	Dombrock blood group	+			T
	EL1	Elliptocytosis (Rh-linked)			p	
	EL2	Elliptocytosis (not Rh-linked)				T
	ENO1	Enolase-1		4.2.1.11	p36	
	FH	Fumarate hydratase		4.2.1.2	q42→qter	
	FUCA	α-L-Fucosidase	+	3.2.1.51	p34→p32	
	Fy	Duffy blood group	+		pter→q21 or q25 or q32→qter	
	GALE	UDPGAL-4-epimerase		5.1.3.2	pter→p32	
	GBA	β-Glucosidase, acid		3.2.1.45	p11→qter	P
	GDH	Glucose dehydrogenase	+	1.1.1.47	pter→p21	
	GUK1	Guanylate kinase-1		2.7.4.8	q32→q42	
	GUK2	Guanylate kinase-2		2.7.4.8		
	MTR	Tetrahydropteroylglutamate methyltransferase		2.1.1.13		P

(*Continued*)

APPENDIX A. (Continued)

Chromosome	Gene symbol[b]	Gene marker	Polymorphic[c]	E.C. No.	Smallest region[d]	Status[e]
	PEPC	Peptidase-C	+	3.4.11.* or 3.4.13.*	q25 or q42	
	PFKM	Phosphofructokinase, M subunit		2.7.1.11	p32.1→q32	P
	PGD	Phosphogluconate dehydrogenase	+	1.1.1.44	pter→p34	
	PGM1	Phosphoglucomutase-1	+	2.7.5.1	p22.1	
	PKU1	Phenylketonuria				I
	Rd	Radin blood group			p34→p22.1	P
	Rh	Rhesus blood group	+		p36→p32	
	RN5S	5S RNA			q42→q43	
	RP1	Retinitis pigmentosa-1				T
	Sc	Scianna blood group	+		p34→p32	
	SDH	Succinate dehydrogenase		1.3.99.1	p22.1→qter	P
	UGP1	UPD glucose pyrophosphorylase-1		2.7.7.9	q21→q22	
	UMPK	Uridine monophosphate kinase	+	2.7.4.*	p32	
Probable gene order: pter-(*ENO1,A12M2,D1S1,PGD,GDH,GALE*)-*EL1-Rh-FUCA*-(*Sc,Rd,AK2*)-*UMPK-PGM1*-(*AMY1,AMY2*)-cen-(*D1Z1,Fy,CAE,A12M3,UGP1*)-*AT3-GUK-PEPC*-(*A12M1,RN5S,FH*)-qter						
2	*ACP1*	Acid phosphatase-1	+	3.1.3.2	p23 or p25	
	ADCP2	Adenosine deaminase complexing protein-2				
	AHH	Arylhydrocarbon hydroxylase		1.14.14.1		P
	D2S1	DNA segment	+			P
	GLAT	Galactose enzyme activator			p22→p11	P
	IDH1	Isocitrate dehydrogenase (soluble)		1.1.1.42	q32→qter	
	IF1	Interferon-1			p23→qter	P
	IGK	Immunoglobulin κ chain, region not specified				
	IGKC	Immunoglobulin κ chain, constant region	+			P
	IGKV	Immunoglobulin κ chain, variable region			p13→cen	P
	Jk	Kidd blood group	+			

	MDH1	Malate dehydrogenase, NAD (soluble)		1.1.1.37	p23	
	POC	Proopiocortin (adrenocorticotropin/β-lipotropin)				P
	RACH	Regulator of acetylcholinesterase				P
	RPE	Ribulose-5-phosphate 3-epimerase		5.1.3.1		P
	UGP2	UDG glucose pyrophosphorylase-2		2.7.7.9		P
Probable gene order: *pter-ACP1-MDH1-GLAT-IGKV*-cen-*IDH1*-qter						
3	*ACY1*	Aminoacylase-1		3.5.1.14	p21	
	AF8T	AF8T temperature sensitivity complementing				P
	D3S1	DNA segment	+			P
	D3S2	DNA segment	+			P
	GLB1	β-Galactosidase-1		3.2.1.23	p21→q21	
	GPX1	Glutathione peroxidase-1	+	1.11.1.9	p13→q12	
(or 11)	*Hv1S*	Herpes simplex virus type 1 sensitivity				I
	TFRC	Transferrin receptor				P
Probable gene order: pter-*ACY1*-(*GLB*,cen,*GPX1*)-qter						
4	*ALB*	Albumin	+		q11→q13	
	DGI	Dentinogenesis imperfecta				P
	GC	Group-specific protein	+		q11→q13	
	MN	MN blood group	+		q28→q31	
	PEPS	Peptidase-S		3.4.11.* or 3.4.13.*	p12→q12	
	PGM2	Phosphoglucomutase-2	+	2.7.5.1	p14→q12	
	PLG	Plasminogen	+		pter→q22 or q25→qter	P
	PPAT	Phosphoribosyl pyrophosphate amidotransferase		2.4.2.14	pter→q21	P
	QDPR	Quinoid dehydrogenase reductase		1.6.5.1		P
	Sf	Stoltzfus blood group				P
	Ss	Ss blood group	+		q28→q31	
	TYS	Sclerotylosis				P
Probable gene order: pter-(*PGM2*,*PEPS*,cen)-(*ALB*,*GC*)-(*MN*,*Ss*)-qter						

(*Continued*)

APPENDIX A. (Continued)

Chromosome	Gene symbol[b]	Gene marker	Polymorphic[c]	E.C. No.	Smallest region[d]	Status[e]
5	*ARSB*	Arysulfatase-B		3.1.6.1		
	AVRR	Antiviral state repressor regulator			p	P
	D5S1	DNA segment	+			P
	DTS	Diphtheria toxin sensitivity			q15→qter	
	HEXB	Hexosaminidase-B (β subunit)		3.2.1.30	q11→q13	
	IF2	Interferon-2			p	P
	LARS	Leucyl-tRNA synthetase		6.1.1.4		P
Probable gene order: cen-*HEXB-DTS*-qter						
6	*ADCP1*	Adenosine deaminase complexing protein-1				P
	BEVI	Baboon M7 virus infection				
	BF	Properdin factor B (glycine-rich glycoprotein)	+		p23→p2105	
	C2	Complement component-2	+		p23→p2105	P
	C4F	Complement component-4F (Rogers)	+		p23→p2105	
	C4S	Complement component-4S (Chido)	+		p23→p2105	
	CAH	Congenital adrenal hyperplasia III (21-hydroxylase deficiency)			p23→p2105	
	D6S1	DNA segment				P
	GLO	Glyoxalase I (lactoyl-glutathione lyase)	+	4.4.1.5	p23→p2100	
	HAF	Clotting factor XII (Hageman)				P
	HC	Hypercholesterolemia				P
	HLA-A	HLA-A	+		p23→p2105	
	HLA-B	HLA-B	+		p23→p2105	
	HLA-C	HLA-C	+		p23→p2105	
	HLA-D	HLA-D	+		p23→p2105	
	HLA-DR	HLA-D related	+		p23→p2105	
	IS	Immune suppression				P
	ME1	Malic enzyme (soluble)		1.1.1.40	q12→q15	

	MRBC	Monkey red blood cell receptor				P
	P	P blood group	+			P
	PGM3	Phosphoglucomutase-3	+	2.7.5.1	q12→qter	
	PLA	Plasminogen activator				P
	PRL	Prolactin				P
	S5	Surface antigen (chromosome 6)				P
	SCA1	Spinal cerebellar ataxia				P
	SOD2	Superoxide dismutase (mitochondrial)		1.15.1.1	q21	
	TRM1	$tRNA_i^{met}$				P
	TRM2	$tRNA_i^{met}$				P

Probable gene order: pter-*HLA A-HLA C-HLA B-(BF,C2,C4F,C4S,CAH)-HLA D-HLA DR-GLO*-cen-*ME1-PGM3-SOD2*-qter

7	*ASL*	Argininosuccinate lyase		4.3.2.1	pter→q22	P
	BLVR	Biliverdin reductase		1.3.1.24	p14-cen	P
	COL1A1	Collagen, type I, α1				P
	COL1A2	Collagen, type I, α2				
	COL3A1	Collagen, type III, α1				P
	EGFR	Epidermal growth factor receptor			p13→p11	
	GCF1	Growth control factor-1				P
	GUSB	β-Glucuronidase		3.2.1.31	cen→q22	
	H1	H1 Histone			q22 or q32→q36	
	H2A	H2A Histone			q22 or q32→q36	
	H2B	H2B Histone			q22 or q32→q36	
	H3	H3 Histone			q22 or q32→q36	
	H4	H4 Histone			q22 or q32→q36	
	HADH	Hydroxyacyl-CoA dehydrogenase		1.1.1.35		P
	MDH2	Malate dehydrogenase, NAD (mitochondrial)		1.1.1.37	p22→q22	
	NHCP	Non-histone chromosome protein				P
	NM	Neutrophil migration			q22→qter	P
	PSP	Phosphoserine phosphatase				P
	S6	Surface antigen (chromosome 7)-1 (MW 165,000)			pter→p11	P

(*Continued*)

APPENDIX A. (Continued)

Chromosome	Gene symbol[b]	Gene marker	Polymorphic[c]	E.C. No.	Smallest region[d]	Status[e]
	S7	Surface antigen (chromosome 7)-2				P
	UP	Uridine phosphorylase		2.4.2.3		
Probable gene order: pter-(*S6,BLVR,EGFR*)-cen-*GUSB*-(*NM,H1,H2A,H2B,H3,H4*)-qter						
8						
	FNS	Fibronectin surface control				P
(or 16)	*GPT*	Glutamic-pyruvic transaminase (alanine aminotransferase)		2.6.1.2	8 or 16pter→16p11	I
	GSR	Glutathione reductase	+	1.6.4.2	p211	
(or 12)	*SPH1*	Spherocytosis-1			8p11 or 12p13	I
Probable gene order: pter-*GSR*-cen						
9	*ABO*	ABO blood group	+		q34	
	ACO1	Aconitase (soluble)	+	4.2.1.3	p22→p13	
	AK1	Adenylate kinase-1	+	2.7.4.3	q34	
	AK3	Adenylate kinase-3		2.7.4.10	p24→p13	
	ASS	Argininosuccinate synthetase		6.3.4.5	q34→qter	P
	DNCM	DNA associated with cytoplasmic membrane			ph	P
	FPGS	Folylpolyglutamate synthetase				
	GALT	Galactose-1-phosphate uridylyltransferase	+	2.7.7.12	pter→p13	
	IFF	Interferon, fibroblast β type			p24→qter	
	IFL	Interferon, leukocyte α type				P
	NPSI	Nail-patella syndrome type 1			q	
	ORM	Orosomucoid			q	P
	WS1	Waardenburg syndrome, type I				T
Probable gene order: pter-*GALT*-(*AK3,ACO1*)-cen-(*ABO, NPS1,AK1,ASS*)-qter						

10	*ADK*	Adenosine kinase		2.7.1.20	q11→q24	
(18, P)	*CGH*	Chorionic gonadotropin				P
	D10S1	DNA segment				P
	FUSE	Polykaryocytosis promoter				P
	GOT1	Glutamic-oxaloacetic transaminase (soluble)	+	2.6.1.1	q24→q25	
	GSAS	Glutamate-γ-semialdehyde synthetase				P
	HK1	Hexokinase-1		2.7.1.1		
	LIPA	Lipase-A				
	M130	External membrane protein-130 (MW 130,000)				P
	PFKF	Phosphofructokinase, F subunit		2.7.1.11		P
	PP	Pyrophosphatase (inorganic)		3.6.1.1	p15→q24	
Probable gene order: cen-*ADK*-*GOT1*-qter						
11	*ACP2*	Acid phosphatase-2		3.1.3.2	p12→cen	
	BVIX	BALB virus induction xenotropic (induction of xenotropic oncornavirus)				P
	CAT	Catalase			p13	
	CPSD	Cathepsin D		3.4.23.5	pter→q12	P
	D11S1	DNA segment	+		pter→p13	P
	D11S2	DNA segment			p13→p1208	P
	D11S3	DNA segment			p1208→p11	P
	D11S4	DNA segment			p11→q13	P
	D11S5	DNA segment				P
	D11S6	DNA segment				P
	D11S7	DNA segment				P
	D11S8	DNA segment				P
	D11S9	DNA segment				P
	D11S10	DNA segment				P
	D11S11	DNA segment				P
	ESA4	Esterase-A_4		3.1.1.1	cen→q22	
	FCP	Hemoglobin F cell production				

(*Continued*)

APPENDIX A. (Continued)

Chromosome	Gene symbol[b]	Gene marker	Polymorphic[c]	E.C. No.	Smallest region[d]	Status[e]
	FN	Fibronectin (large, external, transformation sensitive protein)				
	GST1	Glutathione *S*-transferase-1				P
	HBB	Hemoglobin β	+		p1208→p1205	
	HBD	Hemoglobin δ			p1208→p1205	
	HBE	Hemoglobin ϵ			p1208→p1205	
	HBG1	Hemoglobin γ^A	+		p1208→p1205	
	HBG2	Hemoglobin γ^G	+		p1208→p1205	
(or 3)	*HV1S*	Herpes simplex virus type 1 sensitivity				I
	INS	Insulin			p15.5→p13	
	LDHA	Lactate dehydrogenase-A		1.1.1.27	p1208→p1203	
	MIC1	Antigen identified by monoclonal antibody W6/34				P
	MIC4	Antigen identified by monoclonal antibody F10.44.2				P
	S1	Lethal antigen-1			pter→p13	P
	S2	Lethal antigen-2			q13→qter	P
	S3	Lethal antigen-3			pter→p13	P
	S4	Species antigen			p	P
	UPS	Uroporphyrinogen I synthase		4.3.1.8	q23→qter	
	WAGR	Wilms' tumor—aniridia, genitourinary abnormalities, and mental retardation triad			p13	
Probable gene order: pter-(*S1,S3,INS,D11S1*)-(*CAT,WAGR,D11S2*)-(*LDHA,HBE,HBG2,HBG1,HBD,HBB,D11S3,ACP2*)-cen-(*ESA4,S2,UPS*)-qter						
12	*BCT1*	Branched chain aminotransferase-1		2.6.1.26	pter→q12	
	CS	Citrate synthase		4.1.3.7	p11→qter	
	D12S1	DNA segment				P

	ENO2	Enolase-2		4.2.1.11	p11→qter	
	GAPD	Glyceraldehyde-3-phosphate dehydrogenase		1.2.1.12	p13	
	GPD1	Glycerol-3-phosphate dehydrogenase		1.1.1.8		P
	KAR	α-Keto acid (aromatic) reductase			p	P
	LDHB	Lactate dehydrogenase-B		1.1.1.27	p12.4→p12.1	
	MIC3	Antigen identified by monoclonal antibody 602-29				P
	PEPB	Peptidase-B		3.4.11.* or 3.4.13.*	q21	
	S8	Surface antigen (chromosome 12)-1				P
	SHMT	Serine hydroxymethyl transferase		2.1.2.1	q12→q14	
(or 8)	*SPH1*	Spherocytosis-1			8p11 or 12p13	I
	TPI1	Triosephosphate isomerase-1		5.3.1.1	p13	
	TPI2	Triosephosphate isomerase-2		5.3.1.1		
Probable gene order: pter-(*GAPD,TPI*)-*LDHB*-cen-*SHMT-PEPB*-qter						
13	*ESD*	Esterase-D	+	3.1.1.1	q14	
	RB1	Retinoblastoma-1			q	
(and 14, 15, 21, 22)	*RNR*	Ribosomal RNA			p12	
Probable gene order: pter-*RNR*-cen-*RB1-ESD*-qter						
14	*CKBB*	Creatine kinase BB isozyme		2.7.3.2		
	D14S1	DNA segment	+		q21→qter	P
	D14S2	DNA segment				P
	EBV	Epstein–Barr virus				P
	ESAT	Esterase activator				P
	IGHA1	Immunoglobulin α^1 heavy chain				
	IGHA2	Immunoglobulin α^2 heavy chain				
	IGHD	Immunoblobulin δ heavy chain				
	IGHE	Immunoglobulin ϵ heavy chain				
	IGHF	Immunoglobulin heavy chain flanking region				

(*Continued*)

APPENDIX A. (Continued)

Chromosome	Gene symbol[b]	Gene marker	Polymorphic[c]	E.C. No.	Smallest region[d]	Status[e]
	IGHG1	Immunoglobulin γ^1 heavy chain	+			
	IGHG2	Immunoglobulin γ^2 heavy chain	+			
	IGHG3	Immunoblogulin γ^3 heavy chain	+			
	IGHG4	Immunoglobulin γ^4 heavy chain	+			
	IGHM	Immunoglobulin μ heavy chain				
	IGHV	Immunoglobulin heavy chain, variable region (chain not specified)				
	LCH	Lentil agglutinin binding				P
	M195	External membrane protein-195 (MW 195,000)				P
	NP	Nucleoside phosphorylase		2.4.2.1	q12→q20	
	PFGS	Phosphoribosylformylglycinamide synthetase (formylglycinamide ribotide aminotransferase)		6.3.5.3		P
	PGFT	Phosphoribosylglycinamide formyltransferase		2.1.2.2		P
	PI	α_1-Antitrypsin (α_1-protease inhibitor)	+			
(and 13, 15, 21, 22)	*RNR*	Ribosomal RNA			p12	
	WARS	Tryptophanyl-tRNA synthetase		6.1.1.2	q21→qter	
Probable gene order: pter-*RNR*-cen-*NP*-(*D14S1*,*WARS*)-qter						
15	*B2M*	β_2-Microglobulin			q22→qter	
	BVIN	BALB virus induction N-tropic (induction of N-tropic oncornavirus)				P
	GANC	α-Glucosidase (neutral)-C	+	3.2.1.20		P
	HCVS	Human coronavirus sensitivity			q11→qter	P
	HEXA	Hexosaminidase-A (α subunit)	+	3.2.1.30	q22→q23	
	IDH2	Isocitrate dehydrogenase (mitochondrial)		1.1.1.42	q21→qter	

	MANA	α-Mannosidase-A		3.2.1.24	q11→qter	P
	MPI	Mannose phosphate isomerase		5.3.1.8	q22→qter	
	PKM2	Pyruvate kinase (M2)		2.7.1.40	q22→qter	
	PWS	Prader–Willi syndrome			q11	
(and 13, 14, 21, 22)	*RNR*	Ribosomal RNA			p12	
	SORD	Sorbitol dehydrogenase		1.1.1.14	pter→q21	P

Probable gene order: pter-*RNR*-cen-*PWS*-(*MANA,IDH2,HEXA,MPI,PKM2,B2M*)-qter

16	*APRT*	Adenine phosphoribosyltransferase		2.4.2.7	q12→q22	
	AVR	Regulator of antiviral state				P
	CTH	Cystathionase		4.4.1.1		P
	CTRB	Chymotrypsinogen B	+			P
	D16S1	DNA segment	+			P
	D16S2	DNA segment				P
	DIA4	Diaphorase (NADH/NADPH)			q12→q21	
	ESB3	Esterase-B3				P
	GCF2	Growth control factor-2				P
	GOT2	Glutamic-oxaloacetic transaminase (mitochondrial)	+	2.6.1.1		
(or 8)	*GPT*	Glutamic-pyruvic transaminase	+	2.6.1.2	8 or 16pter→16p11	I
	HAGH	Hydroxyacyl glutathione hydrolase		3.1.2.6		P
	HBA	Hemoglobin α	+		pter→p11	
	HBZ	Hemoglobin ζ			pter→p12	
	HP	Haptoglobin	+		cen→q22	
	IFR	Interferon production regulator				P
	LCAT	Lecithin-cholesterol acyltransferase		2.3.1.43	cen→q22	
	LIPB	Lipase-B				P
	PGP	Phosphoglycollate phosphatase	+	3.1.3.18	p13→p12	
	TK2	Thymidine kinase (mitochondrial)		2.7.1.21		P

Probable gene order: pter-(*HBZ,HBA,PGP*)-cen-(*HP,LCAT,DIA4,APRT*)-qter

(*Continued*)

APPENDIX A. (Continued)

Chromosome	Gene symbol[b]	Gene marker	Polymorphic[c]	E.C. No.	Smallest region[d]	Status[e]
17	*A12M4*	Adenovirus-12 chromosome modification site 17			q21→q22	P
	COLM	Collagen marker				P
	CSH	Chorionic somatomammotropin hormone			q22→q24	
	D17S1	DNA segment				P
	GAA	α-Glucosidase (acid)	+	3.2.1.20	q21→q25	
	GALK	Galactokinase		2.7.1.6	q21→q22	
	GH	Growth hormone			q22→q24	
	GHL	Growth hormone-like				P
	S9	Surface antigen (chromosome 17)-1				P
	TK1	Thymidine kinase (soluble)		2.7.1.21	q21→q22	
Probable gene order: cen-(*TK1*,*GALK*,*A12M4*)-(*GAA*,*GH*,*CSH*)-qter						
18 (and 10)	*CGH*	Chorionic gonadotropin				P
	PEPA	Peptidase-A	+	3.4.11.* or 3.4.13.*	q23	
Probable gene order: cen-*PEPA*-qter						
19	*BCT2*	Branched chain aminotransferase-2		2.6.1.26		P
	C3	Complement component-3	+			P
	DNL	Lysosomal DNAse			pter→q13	P
	D19S1	DNA segment				P
	E11S	ECHO 11 virus sensitivity			q	P
	GPI	Glucose phosphate isomerase		5.3.1.9	pter→q13	
	GUSM	Mouse β-glucuronidase modifier				P
	MANB	α-Mannosidase-B (lysosomal)		3.2.1.24	pter→q13	
	M7V1	Baboon virus replication				P
	PEPD	Peptidase-D (proline dipeptidase)	+	3.4.13.9	pter→q13	
	PVS	Poliovirus sensitivity				
	RDRC	RD114 virus receptor				P

20	*ADA*	Adenosine deaminase	+	3.5.4.4	q13.2→qter	
	D20S1	DNA segment				P
	D20S2	DNA segment				P
	D20S3	DNA segment				P
	DCE	Desmosterol-to-cholesterol enzyme				P
	ITPA	Inosine triphosphatase (nucleoside triphosphate pyrophosphatase)		3.6.1.19	p	
	MEN2	Multiple endocrine neoplasia-2			p12.2	P
Probable gene order: pter-*MEN2*-cen-*ADA*-qter						
21	*AABT*	β-Amino acid transport				P
	HTOR	5-Hydroxytryptamine oxygenase regulator				P
	IFRC	Interferon receptor			q21→qter	
	PAIS	Phosphoribosylaminoimidazole synthetase				P
	PFKL	Phosphofructokinase, liver type		2.7.1.11		P
	PRGS	Phosphoribosylglycinamide synthetase		6.3.4.13		
(and 13, 14, 15, 22)	*RNR*	Ribosomal RNA			p12	
	S14	Surface antigen (chromosome 21)				P
	SOD1	Superoxide dismutase (soluble)	+	1.15.1.1	q22.1	
Probable gene order: pter-*RNR*-cen-*SOD1*-*IFRC*-qter						
22	*ACO2*	Aconitase (mitochondrial)		4.2.1.3	q11→q13	
	ARSA	Arylsulfatase-A		3.1.6.1	q13.31→qter	
	DGS	DiGeorge syndrome			q11	P
	DIA1	Diaphorase (NADH) (cytochrome b_5 reductase)		1.6.2.2	q13.31→qter	P
	D22S1	DNA segment				P
	GLB2	β-Galactosidase-2			q11→qter	P
	IGLC	Immunoglobulin λ chain, constant region				P
	IGLV	Immunoglobulin λ chain, variable region				T
	NAGA	α-*N*-Acetyl-galactosaminidase		3.2.1.49	q13	P
(and 13, 14, 15, 21)	*RNR*	Ribosomal RNA			p12	
	S13	Surface antigen (chromosome 22)				P
Probable gene order: pter-*RNR*-cen-*DGS*-*ACO2*-*NAGA*-(*ARSA*,*DIA1*)-qter						

(*Continued*)

APPENDIX A. (Continued)

Chromosome	Gene symbol[b]	Gene marker	Polymorphic[c]	E.C. No.	Smallest region[d]	Status[e]
X	*ALD*	Adrenoleukodystrophy			q28→qter	
	BA2R	BALB/c 3T3 ts2 temperature sensitivity complementing			q13→q27	P
	C1HR	C1AGOH temperature sensitivity complementing				P
	CBD	Color-blindness (deutan)			q	
	CBP	Color-blindness (protan)			q	
	CGD	Chronic granulomatous disease				
	DHTR	Dihydrotestosterone receptor			p11→q11	
	DXS1	DNA segment			pter→q1	P
	DXS2	DNA segment			pter→q1	P
	DXS3	DNA segment			q1→q21	P
	DXS4	DNA segment			q21→q24	P
	DXS5	DNA segment			q21→q24	P
	DXS6	DNA segment			q2→qter	P
	DXS7	DNA segment				P
	DXS8	DNA segment	+		q	P
	DXS9	DNA segment	+		p	P
	DXZ1	DNA segment (repetitive)			q23→q25	P
	GLA	α-Galactosidase		3.2.1.22	q21→q24	
	G6PD	Glucose-6-phosphate dehydrogenase	+	1.1.1.49	q28	
	HEMA	Antihemophilic globulin A (factor VIII, hemophilia A)			q	
	HPRT	Hypoxanthine phosphoribosyl transferase		2.4.2.8	q26→q28	
	HYB	Y histocompatibility antigen, regulator			pter→p22.3	P
	HYC	Y histocompatibility antigen, receptor				
	MAOA	Monoamine oxidase		1.4.3.4		P
	MDB	Muscular dystrophy, Becker type				
	MDD	Muscular dystrophy, Duchenne type			p21	

MIC2	Antigen identified by monoclonal antibody 12E7				P
OA	Ocular albinism				
OTC	Ornithine transcarbamylase		2.1.3.3		
PGK	Phosphoglycerate kinase		2.7.2.3	q13	
PRPS	Phosphoribosyl pyrophosphate synthetase		2.7.6.1	q21→q27	
RNN1	Nuclear RNA-1				P
RS	Retinoschisis				
S10	Surface antigen (X-linked)-1			q26→q28	P
S11	Surface antigen (X-linked)-2				P
S12	Surface antigen (X-linked)-3			q26→qter	P
STS	Steroid sulfatase (microsomal)			pter→p22.3	
TATR	Tyrosine aminotransferase regulator				
Xg	Xg blood group	+		pter→p22.3	
Xk	Kell blood group precursor				
XM	α_2-Macroglobulin				
XPAC	Fast kinetic complementation DNA repair in Xeroderma pigmentosum, group A				P

Probable gene order: pter-(*Xg,STS,HYB*)-*MDD*-cen-(*PGK,DXS3*)-(*GLA,DXS4,DXS5*)-(*DXZ1,PRPS,DXS6*)-(*HPRT,S10,S12*)-(*G6PD-CBD-CBP-HEMA-ALD*)-qter

Y	*HYA*	Y histocompatibility antigen

[a] Compiled from the Human Gene Mapping Workshop 6 and previous workshops.[95–100]
[b] Gene nomenclature follows accepted guidelines.[238,246]
[c] Refers to being polymorphic in at least one major ethnic group.
[d] Chromosome terminology follows the Paris Conference,[199] and smallest region refers to the smallest, most consistent region.
[e] A provisional assignment is listed as P, a tentative assignment as T, and an inconsistent assignment as I. Assignments without qualifications have been confirmed by two or more independent investigators.

APPENDIX B. Alphabetical Listing of Human Genes Assigned to Chromosomes

Gene marker	Gene symbol[a]	Chromosome assignment[b]	Status[c]
ABO blood group	*ABO*	9	
Acid phosphatase-1	*ACP1*	2	
Acid phosphatase-2	*ACP2*	11	
Aconitase (mitochondrial)	*ACO2*	22	
Aconitase (soluble)	*ACO1*	9	
Adenine phosphoribosyltransferase	*APRT*	16	
Adenosine deaminase	*ADA*	20	
Adenosine deaminase complexing protein-1	*ADCP1*	6	P
Adenosine deaminase complexing protein-2	*ADCP2*	2	
Adenosine kinase	*ADK*	10	
Adenovirus-12 chromosome modification site 1A	*A12M2*	1	P
Adenovirus-12 chromosome modification site 1B	*A12M3*	1	P
Adenovirus-12 chromosome modification site 1C	*A12M1*	1	P
Adenovirus-12 chromosome modification site 17	*A12M4*	17	P
Adenylate kinase-1	*AK1*	9	
Adenylate kinase-2	*AK2*	1	
Adenylate kinase-3	*AK3*	9	
Adrenoleukodystrophy	*ALD*	X	
AF8 temperature sensitivity complementing	*AF8T*	3	P
Albumin	*ALB*	4	
β-Amino acid transport	*AABT*	21	P
Aminoacylase-1	*ACY1*	3	
α-Amylase (pancreatic)	*AMY2*	1	
α-Amylase (salivary)	*AMY1*	1	
Antigen identified by monoclonal antibody W6/34	*MIC1*	11	P
Antigen identified by monoclonal antibody 12E7	*MIC2*	X	P
Antigen identified by monoclonal antibody 602–29	*MIC3*	12	P
Antigen identified by monoclonal antibody F10.44.2	*MIC4*	11	P
Antihemophilic globulin A (factor VIII, hemophilia A)	*HEMA*	X	
Antithrombin III	*AT3*	1	P
α_1-Antitrypsin (α_1-protease inhibitor)	*PI*	14	
Antiviral state repressor regulator	*AVRR*	5	P
Argininosuccinate lyase	*ASL*	7	P
Argininosuccinate synthetase	*ASS*	9	P
Arylhydrocarbon hydroxylase	*AHH*	2	P
Arylsulfatase-A	*ARSA*	22	
Arylsulfatase-B	*ARSB*	5	
Baboon M7 virus infection	*BEVI*	6	
Baboon virus replication	*M7VI*	19	P
BALB virus induction N-tropic (induction of N-tropic oncornavirus)	*BVIN*	15	P
BALB virus induction xenotropic (induction of xenotropic oncornavirus)	*BVIX*	11	P

(*Continued*)

APPENDIX B. (Continued)

Gene marker	Gene symbol[a]	Chromosome assignment[b]	Status[c]
BALB/c 3T3 ts2 temperature sensitivity complementing	*BA2R*	X	P
Biliverdin reductase	*BLVR*	7	P
Branched-chain aminotransferase-1	*BCT1*	12	
Branched-chain aminotransferase-2	*BCT2*	19	P
C1AGOH temperature sensitivity complementing	*C1HR*	X	P
Catalase	*CAT*	11	
Cataract, zonular pulverulent (*Fy*-linked)	*CAE*	1	
Cathepsin D	*CPSD*	11	P
Charcot–Marie–Tooth disease (slow conduction type)	*CMT1*	1	
Chorionic gonadotropin	*CGH*	10 & 18	P
Chorionic somatomammotropin hormone	*CSH*	17	
Chronic granulomatous disease	*CGD*	X	
Chymotrypsinogen B	*CTRB*	16	P
Citrate synthase	*CS*	12	
Clotting factor XII (Hageman)	*HAF*	6	P
Collagen, type I, α1	*COL1A1*	7	P
Collagen, type I, α2	*COL1A2*	7	
Collagen, type III, α1	*COL3A1*	7	P
Collagen marker	*COLM*	17	P
Color-blindness (deutan)	*CBD*	X	
Color-blindness (protan)	*CBP*	X	
Complement component-2	*C2*	6	P
Complement component-3	*C3*	19	P
Complement component-4F	*C4F*	6	
Complement component-4S	*C4S*	6	
Congenital adrenal hyperplasia III (21-hydroxylase deficiency)	*CAH*	6	
Creatine kinase BB isozyme	*CKBB*	14	
Cystathionase	*CTH*	16	P
Dentinogenesis imperfecta	*DGI*	4	P
Desmosterol-to-cholesterol enzyme	*DCE*	20	P
Diaphorase (NADH) (cytochrome b_5 reductase)	*DIA1*	22	P
Diaphorase (NADH/NADPH)	*DIA4*	16	
DiGeorge syndrome	*DGS*	22	P
Dihydrotestosterone receptor	*DHTR*	X	
Diphtheria toxin sensitivity	*DTS*	5	
DNA associated with cytoplasmic membrane	*DNCM*	9	P
DNA satellite 3	*D1Z1*	1	P
DNA segment on chromosome 1	*D1S1*	1	P
DNA segment on chromosome 2	*D2S1*	2	P
DNA segment on chromosome 3	*D3S1*	3	P
DNA segment on chromosome 3	*D3S2*	3	P

(*Continued*)

APPENDIX B. (Continued)

Gene marker	Gene symbol[a]	Chromosome assignment[b]	Status[c]
DNA segment on chromosome 5	*D5S1*	5	P
DNA segment on chromosome 6	*D6S1*	6	P
DNA segment on chromosome 10	*D10S1*	10	P
DNA segment on chromosome 11	*D11S1*	11	P
DNA segment on chromosome 11	*D11S2*	11	P
DNA segment on chromosome 11	*D11S3*	11	P
DNA segment on chromosome 11	*D11S4*	11	P
DNA segment on chromosome 11	*D11S5*	11	P
DNA segment on chromosome 11	*D11S6*	11	P
DNA segment on chromosome 11	*D11S7*	11	P
DNA segment on chromosome 11	*D11S8*	11	P
DNA segment on chromosome 11	*D11S9*	11	P
DNA segment on chromosome 11	*D11S10*	11	P
DNA segment on chromosome 11	*D11S11*	11	P
DNA segment on chromosome 12	*D12S1*	12	P
DNA segment on chromosome 14	*D14S1*	14	P
DNA segment on chromosome 14	*D14S2*	14	P
DNA segment on chromosome 16	*D16S1*	16	P
DNA segment on chromosome 16	*D16S2*	16	P
DNA segment on chromosome 17	*D17S1*	17	P
DNA segment on chromosome 19	*D19S1*	19	P
DNA segment on chromosome 20	*D20S1*	20	P
DNA segment on chromosome 20	*D20S2*	20	P
DNA segment on chromosome 20	*D20S3*	20	P
DNA segment on chromosome 22	*D22S1*	22	P
DNA segment on chromosome X	*DXS1*	X	P
DNA segment on chromosome X	*DXS2*	X	P
DNA segment on chromosome X	*DXS3*	X	P
DNA segment on chromosome X	*DXS4*	X	P
DNA segment on chromosome X	*DXS5*	X	P
DNA segment on chromosome X	*DXS6*	X	P
DNA segment on chromosome X	*DXS7*	X	P
DNA segment on chromosome X	*DXS8*	X	P
DNA segment on chromosome X	*DXS9*	X	P
DNA segment (repetitive) X	*DXZ1*	X	P
Dombrock blood group	*Do*	1	T
Duffy blood group	*Fy*	1	
ECHO 11 virus sensitivity	*E11S*	19	P
Elliptocytosis (Rh-linked)	*EL1*	1	
Elliptocytosis (not Rh-linked)	*EL2*	1	T
Enolase-1	*ENO1*	1	
Enolase-2	*ENO2*	12	
Epidermal growth factor receptor	*EGFR*	7	
Epstein–Barr virus	*EBV*	14	P

(*Continued*)

APPENDIX B. (Continued)

Gene marker	Gene symbol[a]	Chromosome assignment[b]	Status[c]
Esterase-A_4	*ESA4*	11	
Esterase activator	*ESAT*	14	P
Esterase-B3	*ESB3*	16	P
Esterase-D	*ESD*	13	
External membrane protein-130 (MW 130,000)	*M130*	10	P
External membrane protein-195 (MW 195,000)	*M195*	14	P
Fast kinetic complementation DNA repair in Xeroderma pigmentosum, group A	*XPAC*	X	P
Fibronectin	*FN*	11	
Fibronectin surface control	*FNS*	8	P
Folylpolyglutamate synthetase	*FPGS*	9	
α-L-Fucosidase	*FUCA*	1	
Fumarate hydratase	*FH*	1	
Galactokinase	*GALK*	17	
Galactose enzyme activator	*GLAT*	2	P
Galactose-1-phosphate uridylyltransferase	*GALT*	9	
α-Galactosidase	*GLA*	X	
β-Galactosidase-1	*GLB1*	3	
β-Galactosidase-2	*GLB2*	22	P
Glucose dehydrogenase	*GDH*	1	
Glucose phosphate isomerase	*GPI*	19	
Glucose-6-phosphate dehydrogenase	*G6PD*	X	
α-Glucosidase, acid	*GAA*	17	
β-Glucosidase, acid	*GBA*	1	P
α-Glucosidase (neutral)-C	*GANC*	15	P
β-Glucuronidase	*GUSB*	7	
Glutamate-γ-semialdehyde synthetase	*GSAS*	10	P
Glutamic-oxaloacetic transaminase (mitochondrial)	*GOT2*	16	
Glutamic-oxaloacetic transaminase (soluble)	*GOT1*	10	
Glutamic-pyruvic transaminase (alanine aminotransferase)	*GPT*	8 or 16	I
Glutathione peroxidase-1	*GPX1*	3	
Glutathione reductase	*GSR*	8	
Glutathione S-transferase-1	*GST1*	11	P
Glyceraldehyde-3-phosphate dehydrogenase	*GAPD*	12	
Glycerol-3-phosphate dehydrogenase	*GPD1*	12	P
Glyoxalase I (lactoyl-glutathione lyase)	*GLO*	6	
Group-specific protein	*GC*	4	
Growth control factor-1	*GCF1*	7	P
Growth control factor-2	*GCF2*	16	P
Growth hormone	*GH*	17	
Growth hormone-like	*GHL*	17	P
Guanylate kinase-1	*GUK1*	1	
Guanylate kinase-2	*GUK2*	1	

(*Continued*)

APPENDIX B. (Continued)

Gene marker	Gene symbol[a]	Chromosome assignment[b]	Status[c]
H1 Histone	*H1*	7	
H2A Histone	*H2A*	7	
H2B Histone	*H2B*	7	
H3 Histone	*H3*	7	
H4 Histone	*H4*	7	
Haptoglobin	*HP*	16	
Hemoglobin α	*HBA*	16	
Hemoglobin β	*HBB*	11	
Hemoglobin δ	*HBD*	11	
Hemoglobin ε	*HBE*	11	
Hemoglobin γ^A	*HBG1*	11	
Hemoglobin γ^G	*HBG2*	11	
Hemoglobin ζ	*HBZ*	16	
Hemoglobin F cell production	*FCP*	11	
Herpes simplex virus type 1 sensitivity	*HV1S*	3 or 11	I
Hexokinase-1	*HK1*	10	
Hexosaminidase-A (α subunit)	*HEXA*	15	
Hexosaminidase-B (β subunit)	*HEXB*	5	
HLA-A	*HLA-A*	6	
HLA-B	*HLA-B*	6	
HLA-C	*HLA-C*	6	
HLA-D	*HLA-D*	6	
HLA-D related	*HLA-DR*	6	
Human coronavirus sensitivity	*HCVS*	15	P
Hydroxyacyl glutathione	*HAGH*	16	P
Hydroxyacyl-CoA dehydrogenase	*HADH*	7	P
5-Hydroxytryptamine oxygenase regulator	*HTOR*	21	P
Hypercholesterolemia	*HC*	6	P
Hypoxanthine phosphoribosyl transferase	*HPRT*	X	
Immune suppression	*IS*	6	P
Immunoglobulin α^1 heavy chain	*IGHA1*	14	
Immunoglobulin α^2 heavy chain	*IGHA2*	14	
Immunoglobulin δ heavy chain	*IGHD*	14	
Immunoglobulin ε heavy chain	*IGHE*	14	
Immunoglobulin heavy-chain flanking region	*IGHF*	14	
Immunoglobulin γ^1 heavy chain	*IGHG1*	14	
Immunoglobulin γ^2 heavy chain	*IGHG2*	14	
Immunoglobulin γ^3 heavy chain	*IGHG3*	14	
Immunoglobulin γ^4 heavy chain	*IGHG4*	14	
Immunoglobulin μ heavy chain	*IGHM*	14	
Immunoglobulin heavy chain, variable region (chain not specified)	*IGHV*	14	
Immunoglobulin κ chain, constant region	*IGKC*	2	P
Immunoglobulin κ chain, variable region	*IGKV*	2	P

(*Continued*)

APPENDIX B. (Continued)

Gene marker	Gene symbol[a]	Chromosome assignment[b]	Status[c]
Immunoglobulin λ chain, constant region	*IGLC*	22	P
Immunoglobulin λ chain, variable region	*IGLV*	22	P
Inosine triphosphatase (nucleoside triphosphate pyrophosphatase)	*ITPA*	20	
Insulin	*INS*	11	
Interferon-1	*IF1*	2	P
Interferon-2	*IF2*	5	P
Interferon, fibroblast β type	*IFF*	9	
Interferon, leukocyte α type	*IFL*	9	P
Interferon production regulator	*IFR*	16	P
Interferon receptor	*IFRC*	21	
Isocitrate dehydrogenase (mitochondrial)	*IDH2*	15	
Isocitrate dehydrogenase (soluble)	*IDH1*	2	
Kell blood group precursor	*Xk*	X	
α-Keto acid (aromatic) reductase	*KAR*	12	P
Kidd blood group	*Jk*	2	
Lactate dehydrogenase-A	*LDHA*	11	
Lactate dehydrogenase-B	*LDHB*	12	
Lecithin-cholesterol acyltransferase	*LCAT*	16	
Lentil agglutinin binding	*LCH*	14	P
Lethal antigen-1	*S1*	11	P
Lethal antigen-2	*S2*	11	P
Lethal antigen-3	*S3*	11	P
Leucyl-tRNA synthetase	*LARS*	5	P
Lipase-A	*LIPA*	10	
Lipase-B	*LIPB*	16	P
Lysosomal DNAse	*DNL*	19	P
α_2-Macroglobulin	*XM*	X	
Malate dehydrogenase, NAD (mitochondrial)	*MDH2*	7	
Malate dehydrogenase, NAD (soluble)	*MDH1*	2	
Malic enzyme (soluble)	*ME1*	6	
Mannose phosphate isomerase	*MPI*	15	
α-Mannosidase-A	*MANA*	15	P
α-Mannosidase-B (lysosomal)	*MANB*	19	
β_2-Microglobulin	*B2M*	15	
MN blood group	*MN*	4	
Monoamine oxidase	*MAOA*	X	P
Monkey red blood cell receptor	*MRBC*	6	P
Mouse β-glucuronidase modifier	*GUSM*	19	P
Multiple endocrine neoplasia-2	*MEN2*	20	P
Muscular dystrophy, Becker type	*MDB*	X	
Muscular dystrophy, Duchenne type	*MDD*	X	

(*Continued*)

APPENDIX B. (Continued)

Gene marker	Gene symbol[a]	Chromosome assignment[b]	Status[c]
α-*N*-acetyl-galactosaminidase	*NAGA*	22	P
Nail-patella syndrome type 1	*NPS1*	9	
Neutrophil migration	*NM*	7	P
Non-histone chromosome protein	*NHCP*	7	P
Nuclear RNA-1	*RNN1*	X	P
Nucleoside phosphorylase	*NP*	14	
Ocular albinism	*OA*	X	
Ornithine transcarbamylase	*OTC*	X	
Orosomucoid	*ORM*	9	P
P blood group	*P*	6	P
Peptidase-A	*PEPA*	18	
Peptidase-B	*PEPB*	12	
Peptidase-C	*PEPC*	1	
Peptidase-D (proline dipeptidase)	*PEPD*	19	
Peptidase-S	*PEPS*	4	
Phenylketonuria	*PKU1*	1	I
Phosphofructokinase, F subunit	*PFKF*	10	P
Phosphofructokinase, liver type	*PFKL*	21	P
Phosphofructokinase, M subunit	*PFKM*	1	P
Phosphoglucomutase-1	*PGM1*	1	
Phosphoglucomutase-2	*PGM2*	4	
Phosphoglucomutase-3	*PGM3*	6	
Phosphogluconate dehydrogenase	*PGD*	1	
Phosphoglycerate kinase	*PGK*	X	
Phosphoglycollate phosphatase	*PGP*	16	
Phosphoribosyl pyrophosphate amidotransferase	*PPAT*	4	P
Phosphoribosyl pyrophosphate synthetase	*PRPS*	X	
Phosphoribosylaminoimidazole synthetase	*PAIS*	21	P
Phosphoribosylformylglycinamide synthetase (formylglycinamide ribotide aminotransferase)	*PFGS*	14	P
Phosphoribosylglycinamide formyltransferase	*PGFT*	14	P
Phosphoribosylglycinamide synthetase	*PRGS*	21	
Phosphoserine phosphatase	*PSP*	7	P
Plasminogen	*PLG*	4	P
Plasminogen activator	*PLA*	6	P
Poliovirus sensitivity	*PVS*	19	
Polykaryocytosis promoter	*FUSE*	10	P
Prader–Willi syndrome	*PWS*	15	
Prolactin	*PRL*	6	P
Proopiocortin (adrenocorticotropin/β-lipotropin)	*POC*	2	P
Properdin factor B (glycine-rich-β-glycoprotein)	*BF*	6	
Pyrophosphatase (inorganic)	*PP*	10	
Pyruvate kinase (M2)	*PKM2*	15	

(*Continued*)

APPENDIX B. (Continued)

Gene marker	Gene symbol[a]	Chromosome assignment[b]	Status[c]
Quinoid dehydrogenase reductase	*QDPR*	*4*	*P*
Radin blood group	*Rd*	1	P
RD114 virus receptor	*RDRC*	19	P
Regulator of acetylcholinesterase	*RACH*	2	P
Regulator of antiviral state	*AVR*	16	P
Retinitis pigmentosa-1	*RP1*	1	T
Retinoblastoma-1	*RB1*	13	
Retinoschisis	*RS*	X	
Rhesus blood group	*Rh*	1	
Ribosomal RNA	*RNR*	13, 14, 15, 21, 22	
Ribulose-5-phosphate 3-epimerase	*RPE*	2	P
RNA, 5S	*RN5S*	1	
Scianna blood group	*Sc*	1	
Sclerotylosis	*TYS*	4	P
Serine hydroxymethyl transferase	*SHMT*	12	
Sorbitol dehydrogenase	*SORD*	15	P
Species antigen	*S4*	11	P
Spherocytosis-1	*SPH1*	8 or 12	I
Spinal cerebellar ataxia	*SCA1*	6	P
Ss blood group	*Ss*	4	
Steroid sulfatase (microsomal)	*STS*	X	
Stoltzfus blood group	*Sf*	4	P
Succinate dehydrogenase	*SDH*	1	P
Superoxide dismutase (mitochondrial)	*SOD2*	6	
Superoxide dismutase (soluble)	*SOD1*	21	
Surface antigen (chromosome 6)	*S5*	6	P
Surface antigen (chromosome 7)-1 (MW 165,000)	*S6*	7	P
Surface antigen (chromosome 7)-2	*S7*	7	P
Surface antigen (chromosome 12)-1	*S8*	12	P
Surface antigen (chromosome 17)-1	*S9*	17	P
Surface antigen (chromosome 22)	*S13*	22	P
Surface antigen (chromosome 21)	*S14*	21	P
Surface antigen (X-linked)-1	*S10*	X	P
Surface antigen (X-linked)-2	*S11*	X	P
Surface antigen (X-linked)-3	*S12*	X	P
Tetrahydropteroylglutamate methyltransferase	*MTR*	1	P
Thymidine kinase (mitochondrial)	*TK2*	16	P
Thymidine kinase (soluble)	*TK1*	17	
Transferrin receptor	*TFRC*	3	P
Triosephosphate isomerase$_1$	*TPI1*	12	
Triosephosphate isomerase$_2$	*TPI2*	12	
$tRNA_i^{met}$	*TRM1*	6	P

(Continued)

APPENDIX B. (Continued)

Gene marker	Gene symbol[a]	Chromosome assignment[b]	Status[c]
$tRNA_i^{met}$	*TRM2*	6	P
Tryptophanyl-tRNA synthetase	*WARS*	14	
Tyrosine aminotransferase regulator	*TATR*	X	
UDP glucose pyrophosphorylase-1	*UGP1*	1	
UDP glucose pyrophosphorylase-2	*UGP2*	2	P
UDPGAL-4-epimerase	*GALE*	1	
Uridine monophosphate kinase	*UMPK*	1	
Uridine phosphorylase	*UP*	7	
Uroporphyrinogen I synthase	*UPS*	11	
Waardenburg syndrome, type I	*WS1*	9	T
Wilms' tumor—aniridia, genitourinary abnormalities, and mental retardation triad	*WAGR*	11	
Xg blood group	*Xg*	X	
Y histocompatibility antigen	*HYA*	Y	
Y histocompatibility antigen, regulator	*HYB*	X	P
Y histocompatibility antigen, receptor	*HYC*	X	

[a] The gene symbols follow nomenclature guidelines.[238,246]
[b] The assignments have been compiled from the Human Gene Mapping Workships.[95–100]
[c] Status refers to whether an assignment is confirmed (blank), provisional (P), tentative (T), or inconsistent (I) (see text).

APPENDIX C. Linkage Groups Not Assigned to Chromosomes[a]

Gene symbol	Gene marker	E.C. No.	Linked genes
ATN	Albinism, tyrosinase-negative		*CLA*
C6	Complement component-6		*C7*
C7	Complement component-7		*C6*
CCAT	Congenital cataract		*Ii*
CHE1	Cholinesterase (serum)-1	3.1.1.8	*TF*
CLA	Cerebellar ataxia (autosomal recessive)		*ATN*
DB	Double-band salivary protein		*GL, PA, PM, PR, PS*
DM	Myotonic dystrophy		*Lu, Se*
EBR3	Epidermolysis bullosa progressiva		*HOAC*
EBS1	Epidermolysis bullosa simplex (Ogna)		*GPT*
GL	Parotid salivary glycoprotein		*DB, PA, PM, PR, PS*
H	H antigen		*Se*
HHG	Hypergonadotropic hypogonadism		*MSS*

(Continued)

APPENDIX C. (Continued)

Gene symbol	Gene marker	E.C. No.	Linked genes
HOAC	Hypoacusis-2 (recessive)		*EBR3*
Ii	Ii blood group		*CCAT*
K	Kell blood group		*PTC*
Le	Lewis blood group		*C3*
Lu	Lutheran blood group		*DM, Se*
MSS	Marinesco–Sjøgren syndrome		*HHG*
PA	Parotid acidic salivary protein		*DB, GL, PM, PR, PS*
PM	Parotid middle band protein		*DB, GL, PA, PR, PS*
PR	Proline-rich salivary protein		*DB, GL, PA, PM, PS*
PS	Parotid (protein) size variant		*DB, GL, PA, PM, PR*
PTC	Phenylthiocarbamide tasting		*K*
Se	ABH secretion		*DM, H, Lu*
TF	Transferrin		*CHE1*

[a] Compiled from Human Gene Mapping Workshop 6.[100] Gene nomenclature follows the system approved by the Workshop.[246]

ACKNOWLEDGMENTS. Supported by Grants GM20454 (National Institute of General Medical Sciences), HDO5196 (National Institute of Child Health and Human Development), CD62 (American Cancer Society) and MOD-692 and -485 (March of Dimes). The authors are indebted to C. Young, L. Haley, M. Byers, R. Eddy, M. Henry, and S. Shows for their excellent assistance and dedicated efforts. The help and collaboration[238] of Dr. Phyllis McAlpine in constructing the gene map is greatly appreciated.

REFERENCES

1. Amos, D. B., and Kostyu, D. D., 1980, HLA—A central immunological agency of man, in: *Advances in Human Genetics, Vol. 10* (H. Harris and K. Hirschhorn, eds.), pp. 137–208, Plenum Press, New York.
2. Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, E., Schreier, P. H., Smith, A. J. H., Staden, R., and Young, I. G., 1981, Sequence and organization of the human mitochondrial genome, *Nature* **290**:457–465.
3. Baglioni, C., 1962, The fusion of two peptide chains in hemoglobin Lepore and its interpretation as a genetic deletion, *Proc. Nat. Acad. Sci.* **48**:1880–1886.
4. Barg, R., Barton, R., Caine, A., Clements, R. L., Ferguson-Smith, M. A., Malcolm, S., Morrison, H., and Murphy, C. S., 1982, Regional localization of human α-globin gene to the short arm of chromosome 16 (16p12→pter) using both somatic cell hybrids and *in situ* hybridization, *Cytogenet. Cell Genet,* in press.
5. Barker, D., and White, R., 1982, More base pair change polymorphisms at sites containing CpG, *Cytogenet. Cell Genet,* in press.

6. Barrie, P. A., Jeffreys, A. J., and Scott, A. F., 1981. Evolution of the β-globin gene cluster in man and the primates, *J. Mol. Biol.* **149**:319–336.
7. Bell, G. I., Pictet, R. L., Rutter, W. J., Cordell, B., Tischer, E., and Goodman, H., 1980, Sequence of the human insulin gene, *Nature* **284**:26–32.
8. Bell, G. I., Karem, J. H., and Rutter, W. J., 1981, A polymorphic DNA region adjacent to the 5′ end of the human insulin gene, *Proc. Nat. Acad. Sci. USA* **78**:5759–5763.
9. Bell, G. I., Selby, M. J., and Rutter, W. J., 1982, Sequence of a highly polymorphic DNA segment in the 5′ flanking region of the human insulin gene, *Nature* **295**:31–35.
10. Bender, K., and Grzeschik, K.-H., 1976, Assignment of the genes for the human glyoxalase I to chromosome 6 and for human esterase D to chromosome 13, *Cytogenet. Cell Genet.* **16**:93–96.
11. Benn, P. A., Shows, T. B., D'Ancona, G. G., Croce, C. M., Orkwiszewski, K. G., and Mellman, W. J., 1979, Assignment of a gene for uridine diphosphate galactose-4-epimerase to human chromosome 1 by somatic cell hybridization, *Cytogenet. Cell Genet.* **24**:138–142.
12. Berget, S. M., Moore, C., and Sharp, P. A., 1977, Spliced segments at the 5′ terminus of adenovirus 2 late mRNA, *Proc. Nat. Acad. Sci.* **74**:3171–3175.
13. Biro, P. A., Pereira, D., Sood, A. K., de Martinville, B., Francke, U., and Weissman, S. M., 1981, The structure of the human major histocompatibility locus, *J. Supramolec. Struct.*, in press.
14. Bishop, J. O., 1974, The gene numbers game, *Cell* **2**:81–86.
15. Bodmer, W. F., 1981, Gene clusters, genome organization, and complex phenotypes, *Am. J. Hum. Genet.* **33**:664–682.
16. Bolivar, F., 1978, Construction and characterization of new cloning vehicles: III. Derivatives of plasmid pBR322 carrying unique EcoRI sites for selection of EcoRI generated recombinant DNA molecules, *Gene* **4**:121–136.
17. Borst, P., and Grivell, L. A., 1981, One gene's intron is another gene's exon, *Nature* **289**:439–440.
18. Botchan, M., Topp, W., and Sambrook, J., 1976, The arrangement of simian virus 40 sequences in the DNA of transformed cells, *Cell* **9**:269–287.
19. 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**:314–331.
20. Brack, C., and Tonegawa, S., 1977, Variable and constant parts of the immunoglobulin light chain gene of a mouse myeloma cell are 1250 nontranslated bases apart, *Proc. Nat. Acad. Sci. USA* **74**:5652–5656.
21. Brewer, G. J., 1967, Achromatic regions of tetrazolium stained starch gels: Inherited electrophoretic variation, *Am. J. Hum. Genet.* **19**:674–680.
22. Bruns, G., Gusella, J., Keys, C., Housman, D., and Gerald, P., 1982, Isolation of "unique sequence" DNA segments from the human X, *Cytogenet. Cell Genet.*, in press.
23. Carrano, A. V., Gray, J. W., Langlois, R. G., Burkhart-Schultz, K. J., and Van Dilla, M. A., 1979, Measurement and purification of human chromosomes by flow cytometry and sorting, *Proc. Nat. Acad. Sci. USA* **76**:1382–1384.
24. Champion, M. J., and Shows, T. B., 1977, Electrophoretic abnormalities of lysosomal enzymes in mucolipidosis fibroblast lines, *Am. J. Hum. Genet.* **29**:149–163.
25. Champion, M. J., and Shows, T. B., 1977, Correction of human mucolipidosis II enzyme abnormalities in somatic cell hybrids, *Nature* **270**:64–66.
26. Champion, M. J., and Shows, T. B., 1977, Mannosidosis: Assignment of the lysosomal α-mannosidase B gene to chromosome 19 in man, *Proc. Nat. Acad. Sci.* **74**:2968–2972.

27. Chandler, M. E., Kedes, L. H., Coh, R. H., and Yunis, J. J., 1979, Genes coding for histone proteins in man are located on the distal end of the long arm of chromosome 7, *Science* **205**:908–910.
28. Chen, S. H., Malcom, L. A., Yosida, A., and Giblett, E. R., 1971, Phosphoglycerate kinase: An X-linked polymorphism in man, *Am. J. Hum. Genet.* **23**:87–91.
29. Collins, J., 1977, Gene cloning with small plasmids, *Curr. Top. Microbiol. Immunol.* **78**:121–170.
30. Conneally, P. M., and Rivas, M. L., 1980, Linkage analysis in man, in: *Advances in Human Genetics, Vol. 10,* (H. Harris and K. Hirschhorn, eds.), pp. 209–266, Plenum Press, New York.
31. Cook, P. J. L., and Hamerton, J. L., 1979, Report of the committee on the genetic constitution of chromosome 1, *Cytogenet. Cell Genet.* **25**:9–20.
32. Coulondre, C., Miller, J. H., Farabaugh, P. J., and Gilbert, W., 1978, Molecular basis of base substitution hotspots in *Escherichia coli, Nature* **274**:775–780.
33. Cox, D. R., Francke, U., and Epstein, C. J., 1981, Assignment of genes to the human X chromosome by the two-dimensional electrophoresis analysis of total cell proteins from rodent–human somatic cell hybrids, *Am. J. Hum. Genet.* **33**:495–512.
34. Crampton, J. M., Davies, K. E., and Knapp, T. B., 1981, The occurrence of families of repetitive sequences in a library of cloned cDNA from human lymphocytes, *Nucleic Acids Res.* **9**:3821–3834.
35. Crea, R., Kraszewski, A., Hirose, T., and Itakura, K., 1978, Chemical synthesis of genes for human insulin, *Proc. Nat. Acad. USA Sci.* **75**:5765–5769.
36. Creagan, R. P., and Ruddle, F. H., 1975, The clone panel: A systematic approach to gene mapping using interspecific somatic cell hybrids, *Cytogenet. Cell Genet.* **14**:282–286.
37. Creagan, R. P., and Ruddle, F. H., 1977, New approaches to human gene mapping by somatic cell genetics, in: *Molecular Structure of Human Chromosomes* (J. Yunis, ed.), pp. 89–142, Academic Press, New York.
38. Crick, F., 1979, Split genes and RNA splicing, *Science* **204**:264–271.
39. Croce, C. M., Shander, M., Martinis, J., Cecurel, L., D'Ancona, G., Dolby, T. W., and Koprowski, H., 1979, Chromosomal location of the genes for human immunoglobulin heavy chains, *Proc. Nat. Acad. Sci. USA* **76**:3416–3419.
40. Davidson, E. H., and Britten, R. J., 1979, Regulation of gene expression: possible role of repetitive sequences, *Science* **204**:1052–1059.
41. Davidson, R. L., 1974, Gene expression in somatic cell hybrids, *Ann. Rev. Genet.* **8**:195–218.
42. Davies, K. E., Young, B. D., Ellis, R. G., Hill, M. E., and Williamson, R. E., 1981, Cloning of a representative genomic library of the human X chromosome after sorting by flow cytometry, *Nature* **293**:374–376.
43. Deeley, R. G., Ucell, D. S., Burns, A. T. H., Gordon, J. I., and Goldberger, R. F., 1977, Kinetics of avian vitellogenin messenger RNA induction, *J. Biol. Chem.* **252**:7913–7915.
44. de Grouchy, J., Turleau, C., and Finaz, C., 1978, Chromosomal phylogeny of the primates, *Ann. Rev. Genet.* **12**:289–328.
45. Deisseroth, A., Nienhuis, A., Turner, P., Velez, R., Anderson, W. F., Ruddle, F., Lawrence, J., Creagen, R., and Kucherlapati, R., 1977, Localization of the human alpha-globin structural gene to chromosome 16 in somatic cell hybrids by molecular hybridization assay, *Cell* **12**:205–218.
46. Deisseroth, A., Nienhuis, A., Lawrence, J., Giles, R., Turner, P., and Ruddle, F., 1978, Chromosomal localization of human beta-globin gene on human chromosome 11 in somatic cell hybrids, *Proc. Nat. Acad. Sci. USA* **75**:1456–1460.

47. DeLuca, C., Brown, J. A., and Shows, T. B., 1979, Lysosomal arylsulfatase deficiencies in humans: Chromosome assignments for arylsulfatase A and B, *Proc. Nat. Acad. Sci. USA* **76**:1957–1961.
48. de Martinville, B., Leary, J., Ullrich, A., and Francke, U., 1982, The human insulin gene maps on the short arm of chromosome 11, *Cytogenet. Cell Genet.*, in press.
49. de Martinville, B., Wyman, A., White, R., and Francke, U., in press, Assignment of the first highly polymorphic DNA marker locus to a human chromosome region, *Cytogenet. Cell Genet.*
50. Detter, J. C., Ways, P. O., Giblett, E. R., Baughan, M. A., Hophenson, D. A., Povey, S., and Harris, A., 1968, Inherited variations in human phosphohexose isomerase, *Ann. Hum. Genet.* **31**:329–338.
51. Donahue, R. P., Bias, W. B., Renwick, J. H., and McKusick, V. A., 1968, Probable assignment of the Duffy blood group locus to chromosome 1 in man, *Proc. Nat. Acad. Sci. USA* **66**:949–952.
52. Doolittle, R. F., 1981, Similar amino acid sequences: Chance or common ancestry?, *Science* **214**:149–159.
53. Dutrillaux, B., 1979, Chromosomal evolution in primates: Tentative phylogeny from *Microcebus murinus* (prosimian) to man, *Hum. Genet.* **48**:251–314.
54. Efstratiadis, A., Posakony, J. W., Maniatis, T., Lawn, R. A., O'Connell, C., Spritz, R. A., DeRiel, J. K., Forget, B. G., Weissman, S. M., Slightom, J. L., Bleichl, A. E., Smithies, O., Baralle, F. E., Shoulders, C. C., and Proudfoot, N. J., 1980, The structure and evolution of the human β-globin gene family, *Cell* **21**:653–668.
55. Ege, T., and Ringertz, N. R., 1974, Preparation of microcells by enucleation of micronucleate cells, *Exp. Cell Res.* **87**:378–382.
56. Elsevier, S. M., Kucherlapati, R. S., Nichols, E. A., Creagen, R. P., Giles, R. E., Ruddle, F. H., Willecke, K., and McDougall, J. K., 1974, Assignment of the gene for galactokinase to human chromosome 17 and its regional localization to band q21-22, *Nature* **251**:633–636.
57. Evans, H. J., Buckland, R. A., and Pardue, M. L., 1974, Location of the genes coding for 18S and 28S ribosomal RNA in the human genome, *Chromosoma* **48**:405–426.
58. Ferguson-Smith, M. A., Newman, B. F., Ellis, P. M., Thomson, M. G., and Riley, I. D., 1973, Assignment by deletion of human red cell and phosphatase gene locus to the short arm of chromosome 2, *Nature New Biol.* **243**:271–272.
59. Ferguson-Smith, M. A., Aitken, D. A., Truleau, C., and de Grouchy, J., 1976, Localization of the human ABO:NP-1AK-1 linkage group by regional assignment of Ak-1 to 9q34, *Hum. Genet.* **34**:35–43.
60. Fiers, W., Contreras, R., Haegeman, G., Rogiers, R., Van de Voorde, A., Van Heuverswyn, H., Van Herreweghe, J., Volckaert, G., and Ysebaert, M., 1978, Complete nucleotide sequence of SV40 DNA, *Nature* **273**:113–120.
61. Fildes, R. A., and Harris, H., 1966, Genetically determined variation of adenylate kinase in man, *Nature* **209**:261–263.
62. Flavell, R. A., Kooter, J. M., and de Boer, E., 1978, Analysis of the β-δ-globin gene loci in normal and Hb Lepore DNA: Direct determination of gene linkage and intergene distance, *Cell* **15**:25–41.
63. Flavell, R. A., Bernards, R., Kooter, J. M., and de Boer, E., 1979, The structure of the human β-globin gene in β-thalassaemia, *Nucleic Acids Res.* **6**:2749–2760.
64. Fournier, R. E. K., and Ruddle, F. H., 1977, Microcell-mediated transfer of murine chromosomes into mouse, Chinese hamster and human somatic cells, *Proc. Nat. Acad. Sci. USA* **74**:319–323.
65. Fournier, R. E. K., and Ruddle, F. H., 1977, Stable association of the human trans-

genome and host murine chromosomes demonstrated with trispecific microcell hybrids, *Proc. Nat. Acad. Sci. USA* **74**:3937–3941.

66. Francke, U., Lalley, P. A., Moss, W., Ivy, J., and Minna, J. D., 1977, Gene mapping in *Mus musculus* by interspecific cell hybridization: Assignment of the genes for tripeptidase-1 to chromosome 10, dipeptidase-2 to chromosome 18, acid phosphatase-1 to chromosome 12, and adenylate kinase-1 to chromosome 2, *Cytogenet. Cell Genet.* **19**:47–84.
67. Francke, U., Holmes, L. B., Atkins, L., and Riccardi, V. M., 1979, Aniridia-Wilms tumor association: Evidence for specific deletion of 11p13, *Cytogenet. Cell Genet.* **24**:185–192.
68. Fritsch, E. F., Lawn, R. M., and Maniatis, T., 1980, Molecular cloning and characterization of the human β-like globin gene cluster, *Cell* **19**:959–972.
69. Geever, R. F., Wilson, L. B., Nallaseth, F. S., Milner, P. F., Bittner, M., and Wilson, J. T., 1981, Direct identification of sickle cell anemia by blot hybridization, *Proc. Nat. Acad. Sci. USA* **78**:5081–5085.
70. George, D. L., Phillips, J. A. III, Francke, U., and Seeburg, P. H., 1981, The genes for growth hormone and chorionic somatomammotropin are on the long arm of human chromosome 17 in region q21→qter, *Hum. Genet.* **57**:138–141.
71. Gerhard, D. S., Kawasaki, E. S., Bancroft, F. C., and Szabo, P., 1981, Localization of a unique gene by direct hybridization *in situ, Proc. Nat. Acad. Sci. USA* **78**:3755–3759.
72. Gilbert, W., 1978, Why genes in pieces?, *Nature* **271**:501.
73. Glover, D. M., and Hogness, D. S., 1977, A novel arrangement of the 18S and 28S sequences in a repeating unit of *Drosophila melanogaster* rDNA, *Cell* **10**:167–176.
74. Goff, S. P., Gilboa, E., Witte, O. N., and Baltimore, D., 1980, Structure of the Abelson murine leukemia virus genome and the homologous cellular gene: Studies with cloned viral DNA, *Cell* **22**:777–785.
75. Goodfellow, P., Banting, G., Lenz, R., Povey, S., and McMichael, A., 1980, A human X-linked antigen defined by a monoclonal antibody, *Som. Cell Genet.* **6**:777–788.
76. Gordon, J. T., Burns, A. T. H., Christmann, J. L., and Deeley, R. C., 1978, Cloning of a double stranded cDNA that codes for a portion of chicken preproalbumin, *J. Biol. Chem.* **253**:8629–8639.
77. Gosden, J. R., Laurie, S. S., and Cooke, J. H., 1981, A cloned repeated DNA sequence in human chromosome heteromorphisms, *Cytogenet. Cell Genet.* **29**:32–39.
78. Goss, S. J., and Harris, H., 1977, Gene transfer by means of cell fusion I, *J. Cell Sci.* **25**:17–37.
79. Goss, S. J., and Harris, H., 1977, Gene transfer by means of cell fusion II, *J. Cell Sci.* **25**:39–58.
80. Grosveld, F. G., Dahl, H.-H. M., de Boer, E., and Flavell, R. A., 1981, Isolation of β-globin-related genes from a human cosmid library, *Gene* **13**:227–237.
81. Gusella, J., Varsanyi-Breiner, A., Kao, F.-T., Jones, C., Puck, T. T., Keys, C., Orkin, S., and Housman, D., 1979, Precise localization of human beta-globin gene complex on chromosome 11, *Proc. Nat. Acad. Sci. USA* **76**:5239–5243.
82. Gusella, J. F., Keys, C., Varsanyi-Breiner, A., Kao, F.-T., Jones, C., Puck, T. T., and Housman, D., 1980. Isolation and localization of DNA segments from specific chromosomes, *Proc. Nat. Acad. Sci. USA* **77**:2829–2833.
83. Haigh, L. S., Anderson, W. F., and Francke, U., 1979, Regional mapping of the β-globin gene (HBB) to 11p, *Cytogenet. Cell Genet.* **25**:162.
84. Harper, M. E., and Saunders, G. F., 1981, Localization of single copy DNA sequences on G-banded human chromosomes by *in situ* hybridization, *Chromosoma* **83**:431–439.

84a. Harper, M. E., and Saunders, G. F., 1981, Chromosomal localization of human insulin gene, placental lactogen-growth hormone genes, and other single copy genes by *in situ* hybridization, *Am. J. Hum. Genet.* **33:**105A.
85. Harper, M. E., Ullrich, A., and Saunders, G. F., 1981, Localization of the insulin gene to the distal end of short arm of chromosome 11, *Proc. Nat. Acad. Sci. USA* **78:**4458–4460.
86. Harper, M. E., Barrera-Saldana, H. A., and Saunders G. F., in press, Chromosomal localization of the human placental lactogen-growth hormone gene cluster to 17q22–24, *Am. J. Hum. Genet.*
87. Harris, H., 1975, *The Principles of Human Genetics,* 2nd ed., pp. 104–109, North-Holland/American Elsevier, Amsterdam/New York.
88. Harris, H., and Hopkinson, D. A., 1976, *Handbook of Enzyme Electrophoresis in Human Genetics,* North-Holland, Amsterdam.
89. Henderson, A. S., Warburton, D., and Atwood, K. C., 1972, Location of ribosomal DNA in the human chromosome complement, *Proc. Nat. Acad. Sci. USA* **69:**3394–3398.
90. Higgs, D. R., Goodbourn, S. E. Y., Wainscoat, J. S., Clegg, J. B., and Weatherall, D. J., 1981, Highly variable regions of DNA flank the human α globin genes, *Nucleic Acids Res.* **9:**4213–4224.
91. Honey, N. K., Miller, A. L., and Shows, T. B., 1981, The mucolipidoses: Identification by abnormal electrophoretic patterns of lysosomal hydrolases, *Am. J. Med. Genet.* **9:**239–253.
92. Hoogeveen, A. T., Verheijen, F. W., d'Azzo, A., and Galjaard, H., 1980, Genetic heterogeneity in human neuraminidase deficiency, *Nature* **285:**500–502.
93. 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:**298–306.
94. Huisman, T. H. J., Wrightstone, R. N., Wilson, J. B., Schroeder, W. A., and Kendall, A. G., 1972, Haemoglobin Kenya, a product of fusion of γ and β polypeptide chains, *Arch. Biochem. Biophys.* **153:**850–853.
95. Human Gene Mapping 1, 1974, *Cytogenet. Cell Genet.* **13:**1–216.
96. Human Gene Mapping 2, 1975, *Cytogenet. Cell Genet.* **14:**163–480.
97. Human Gene Mapping 3, 1976, *Cytogenet. Cell Genet.* **16:**1–452.
98. Human Gene Mapping 4, 1978, *Cytogenet. Cell Genet.* **22:**1–730.
99. Human Gene Mapping 5, 1979, *Cytogenet. Cell Genet.* **25:**1–236.
100. Human Gene Mapping 6, 1982, *Cytogenet. Cell Genet.,* in press.
101. Hunkapiller, M. W., and Hood, L. E., 1980, New protein sequenator with increased sensitivity, *Science* **207:**523–525.
102. Ish-Horowicz, D., and Burke, J. F., 1981, Rapid and efficient cosmid cloning, *Nucleic Acids Res.* **9:**2989–2998.
103. Jeffreys, A. J., 1979, DNA sequence variants in the ^{G}gamma-, ^{A}gamma-, delta- and beta-globin genes of man, *Cell* **18:**1–10.
104. Jeffreys, A. J., Craig, I. W., and Francke, U., 1979, Localization of the G-gamma, A-gamma, delta-, and beta-globin genes on the short arm of human chromosome 11, *Nature* **281:**606–608.
105. Johnson, L. D., Henderson, A. S., and Atwood, K. C., 1974, Location of the genes for 5S RNA in the human chromosome complement, *Cytogenet. Cell Genet.* **13:**103:105.
106. Judd, B. H., She, M. W., and Kaufman, T. C., 1972, The anatomy and function of a segment of the X chromosome of *Drosophila melanogaster, Genetics* **71:**139–156.
107. Junien, C., Weil, D., Myers, J. C., Van Cong, N., Chu, M.-L., Foubert, C., Gross, M. S., Prochop, D. J., Kaplan, J. C., and Ramirez, F., 1982, Assignment of the human pro α2(I) collagen structural gene to chromosome 7 by molecular hybridization, *Cytogenet. Cell Genet.,* in press.

108. Kacian, D. L., Spiegelman, S., Bank, A., Terada, M., Metafora, S., Dow, L., and Marks, P. A., 1972, *In vitro* synthesis of DNA components of human genes for globins, *Nature New Biol.* **235**:167.
109. Kan, Y. W., and Dozy, A. M., 1978, Polymorphism of DNA sequence adjacent to human α-globin structural gene: Relationship to sickle mutation, *Proc. Nat. Acad. Sci. USA* **75**:5631–5635.
110. Kan, Y. W., and Dozy, A. M., 1978, Antenatal diagnosis of sickle-cell anemia by DNA. analysis of amniotic-fluid cells, *Lancet* **II**:910–911.
111. Kan, Y. W., Lee, K. Y., Furbetta, M., Angius, A., and Cao, A., 1980, Polymorphism of DNA sequence in the β-globin gene region, *New Eng. J. Med.* **302**:185–188.
112. Kao, F.-T., Jones, C., Puck, T. T., Law, M. L., Gusella, J., Keys, C., Varsanyi-Breiner, A., and Housman, D., 1980, Human DNA segments as genetic markers, *Am. J. Hum. Genet.* **32**:44A.
113. Kao, F.-T., Hartz, J. A., and Law, M. L., 1981, DNA segment marker for human chromosome 10, *J. Cell Biol.* **91**:386a.
114. Kao, F.-T., Hartz, J. A., Law, M. L., and Davidson, J. N., 1982, Isolation, characterization and chromosomal localization of DNA segments from a human genomic library in: Abstract book, 6th International Congress of Human Genetics, Israel.
115. Karig Hohmann, L., and Shows, T. B., 1979, Complementation of genetic disease: A velocity sedimentation procedure for the enrichment of heterokaryons, *Somat. Cell Genet.* **5**:1013–1029.
116. Kedes, L. H., 1976, Histone messengers and histone genes, *Cell* **8**:321–331.
117. Klebe, R. J., Chen, T. R., and Ruddle, F. H., 1970, Controlled production of proliferating somatic cell hybrids, *J. Cell Biol.* **45**:74–82.
118. Klinger, H. P., and Ruoslahti, E., 1980, Human chromosome 11 is syntenic with human specific fibronectin production in human × mouse cell hybrids, *Cytogenet. Cell Genet.* **28**:271–279.
119. Klobutcher, L. A., and Ruddle, F. H., 1979. Phenotype stabilization and integration of transferred material in chromosome mediated gene transfer, *Nature* **280**:657–660.
120. Koch, G., and Shows, T. B., 1978, A gene on human chromosome 6 functions in assembly of tissue-specific adenosine deaminase isozymes, *Proc. Nat. Acad. Sci. USA* **75**:3876–3880.
121. Koch, G., Lalley, P. A., McAvoy, M., and Shows, T. B., 1981, Assignment of *LIPA*, associated with human acid lipase deficiency, to human chromosome 10 and comparative assignment to mouse chromosome 19, *Somat. Cell Genet.* **7**:345–358.
122. Koeffler, H. P., Sparkes, R. S., Stang, H., Mohandas, T., and Sparkes, M. C., 1982, Regional assignment of the human genes for α-globin and phosphoglycolate phosphatase to 16p, *Cytogenet. Cell Genet.*, in press.
123. Kompf, J., Bissbort, S., and Ritler, H., 1975, Red cell glyoxylase I (EC 4.4.1.5): Formal genetics and linkage relations, *Human Genetik* **28**:249–251.
124. Krystal, M., D'Eustachio, P., Ruddle, F. H., and Arnheim, N., 1981, Human nucleolus organizers on nonhomologous chromosomes can share the same ribosomal gene variants, *Proc. Nat. Acad. Sci. USA* **78**:5744–5748.
125. Kucherlapati, R. S., Creagan, R. P., Nichols, E. A., Borgaonkai, D. S., and Ruddle, F. H., 1975, Synteny relationships of four human genes: Mannose phosphate isomerase and pyruvate kinase-3 and triose phosphate isomerase to lactate dehydrogenase-B, *Cytogenet. Cell Genet.* **14**:194–199.
126. Kurosky, A., Barnett, D. R., Lee, T.-H., Touchstone, B., Hay, R. E., Arnott, M. S., Bowman, B. H., and Fitch, W. M., 1980, Covalent structure of human haptoglobin: A serine protease homolog, *Proc. Nat. Acad. Sci. USA* **77**:3388–3392.
127. Kwok, S. C. M., Chan, S. J., Rubenstein, A. H., Poucher, R., and Steiner, D. F.,

1981, Loss of a restriction endonuclease cleavage site in the gene of a structurally abnormal human insulin, *Biochem. Biophys. Res. Commun.* **98:**844–849.

128. Lalley, P. A., and Shows, T. B., 1976, Human hexosaminidase A phenotype requires genes assigned to chromosome 5 and 15 for expression, *Cytogenet. Cell Genet.* **16:**192–196.
129. Lalley, P. A., Rattazzi, M. C., and Shows, T. B., 1974, Human β-D-N acetylhexosaminidase A and B: Expression and linkage relationships in somatic cell hybrids, *Proc. Nat. Acad. Sci. USA* **71:**1569–1573.
130. Lalley, P. A., Brown, J. A., and Shows, T. B., 1976, Assignment of the hexosaminidase B gene to chromosome 5 and its segregation after diphtheria toxin selection employing an X;5 translocation; and the linkage of hexosaminidase A, mannosephosphate isomerase, and pyruvate kinase (M2) genes using man–rodent cell hybrids, *Cytogenet. Cell Genet.* **16:**188–191.
131. Lalley, P. A., Brown, J. A., Eddy, R. L., Haley, L. L., Byers, M. G., Goggin, A., and Shows, T. B., 1977, Human β-glucuronidase: Assignment of the structural gene to chromosome 7 using somatic cell hybrids, *Biochem. Genet.* **15:**367–382.
132. Lalley, P. A., Francke, U., and Minna, J. D., 1978, Homologous genes for enolase, phosphogluconate dehydrogenase, phosphoglucomutase, and adenylate kinase are syntenic on mouse chromosome 4 and human chromosome 1p, *Proc. Nat. Acad. Sci. USA* **75:**2382–2386.
133. Lalley, P. A., Minna, J. D., and Francke, U., 1978. Conservation of autosomal gene synteny groups in mouse and man, *Nature* **274:**160–163.
134. Lauer, J., Shen, C.-K. J., and Maniatis, T., 1980, The chromosomal arrangement of human α-like globin genes: Sequence homology and α-globin gene deletions, *Cell* **20:**119–130.
135. Law, M. L., Kao, F.-T., Patterson, D., and Davidson, J. N., 1980, Isolation of recombinant clones containing DNA segments from human chromosomes 12 and 21, *J. Cell Biol.* **87:**109a.
136. Law, M. L., Kao, F.-T., and Davidson, J. N., 1982, Isolation and regional assignment of DNA fragments from human chromosome 12, in: Abstract book, 6th International Congress of Human Genetics, Israel.
137. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G., and Maniatis, T., 1978, The isolation and characterization of linked δ- and β-globin genes from a cloned library of human DNA, *Cell* **15:**1157–1174.
138. Lawn, R. M., Gross, M., Hauck, C. M., Franke, A. E., Gray, P. V., and Goeddel, D., 1981, DNA sequence of a major leukocyte interferon gene, *Proc. Nat. Acad. Sci. USA* **78:**5435–5439.
139. Lebo, R. V., Carrano, A. V., Burkhart-Schultz, K. J., Dozy, A. M., and Kan, Y. W., 1979, Assignment of human beta-, gamma, and deltaglobin genes to the short arm of chromosome 11 by chromosome sorting and DNA restriction enzyme analysis, *Proc. Nat. Acad. Sci. USA* **76:**5804–5808.
140. Lebo, R. V., Yu, L. C., Cordell, B., Cheung, M. C., Chang, J. C., Carrano, A. V., Goodman, H. M., and Kan, Y. W., 1981, Subchromosomal assignment of the human insulin gene to the terminal portion of the short arm of chromosome 11 by chromosome sorting and DNA restriction enzyme analysis, *Clin. Res.* **29:**114A.
141. Leder, P., Hansen, J. N., Konkel, D., Leder, A., Nishioka, Y., and Talkington, C., 1980, Mouse globin system: A functional and evolutionary analysis, *Science* **209:**1336–1392.
142. Leder, A., Swan, D., Ruddle, F. H., D'Eustachio, P., and Leder, P., 1981, Dispersion of α-like globin genes of the mouse to three different chromosomes, *Nature* **293:**196–200.

143. Lele, K. P., Penrose, L. S., and Stallard, H. B., 1963, Chromosome deletion in a case of retinoblastoma, *Ann. Hum. Genet.* **27**:171–174.
144. Lewis, W. H. P., and Harris, H., 1967, Human red cell peptidases, *Nature* **215**:315–355.
145. Little, P. F. R., Annison, G., Darling, S., Williamson, R., Camba, L., and Modell, B., 1980, Model for antenatal diagnosis of beta-thalassemia and other monogenic disorders by molecular analysis of linked DNA polymorphisms, *Nature* **285**:144–147.
146. Littlefield, J. W., 1964, Selection of hybrids from matings of fibroblasts *in vitro* and their presumed recombinants, *Science* **145**:709–710.
147. Lobban, P. E., and Kaiser, A. D., 1973, Enzymatic end-to-end joining of DNA molecules, *J. Mol. Biol.* **78**:453–471.
148. Lomedico, P., Rosenthal, N., Efstratiadis, A., Gilbert, W., Kolodner, R., and Tizard, R., 1979, The structure and evolution of the two nonallelic rat preproinsulin genes, *Cell* **18**:545–558.
149. Magenis, R. E., Hecht, F., and Lourien, E. W., 1970, Heritable fragile site on chromosome 16: Probable localization of haptoglobin locus in man, *Science* **170**:85–87.
150. Malcolm, S., Barton, P., Murphy, C., and Ferguson-Smith, M., 1981, Chromosomal localization of a single copy gene by *in situ* hybridization—human β-globin genes on the short arm of chromosome 11, *Ann. Hum. Genet.* **45**:135–141.
151. Malcolm, S., Barton, P., Bentley, D. L., Ferguson-Smith, M. A., Murphy, C. S., and Rabbitts, T. H., 1982, Assignment of a V_K locus for immunoglobulin light chains to the short arm of chromosome 2 (2 cen → p 13) by *in situ* hybridization using a cRNA probe of H_K101λCh4A, *Cytogenet. Cell Genet.*, in press.
152. Maniatis, T., Jeffrey, A., and Kleid, D. G., 1975, Nucleotide sequence of the rightward operator of phage λ, *Proc. Nat. Acad. Sci. USA* **72**:1184–1188.
153. Maniatis, T., Kee, S. G., Efstratiadis, A., and Kafatos, F. C., 1976, Amplification and characterization of a β-globin gene synthesized *in vitro, Cell* **8**:163–182.
154. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K., and Efstratiadis, A., 1978, The isolation of structural genes from libraries of eucaryotic DNA. *Cell* **15**:687–701.
155. Maniatis, T., Fritsch, E. F., Lauer, J., and Lawn, R. M., 1980, The molecular genetics of human hemoglobins, *Ann. Rev. Genet.* **14**:145–178.
156. Martial, J. A., Hallewell, R. A., Baxter, J. D., and Goodman, H. M., 1979, Human growth hormone: Complementary DNA cloning and expression in bacteria, *Science* **205**:602–605.
157. McAlpine, P. J., Mohandas, T., Ray, M., Wang, H., and Hamerton, J. L., 1976, Assignment of the peptidase D gene locus (PEPD) to chromosome 19 in man, *Cytogenet. Cell Genet.* **16**:204–205.
158. McBride, O. W., and Ozer, H. L., 1973, Transfer of genetic information by purified chromosomes, *Proc. Nat. Acad. Sci. USA* **70**:1258–1262.
159. McBride, O. W., Swan, D., Leder, P., Hieter, P., and Hallis, G., 1982, Chromosomal location of human immunoglobulin light chain constant region genes, *Cytogenet. Cell Genet.*, in press.
160. McConkey, E. H., Taylor, B. J., and Phan, D., 1979, Human heterozygosity: A new estimate, *Proc. Nat. Acad. Sci. USA* **76**:6500–6504.
161. McKusick, V. A., 1978, *Mendelian Inheritance in Man: Catalogs of Autosomal Dominant, Autosomal Recessive, and X-Linked Phenotypes,* 5th ed., Johns Hopkins University Press, Baltimore.
162. McKusick, V. A., 1980, The anatomy of the human genome, *J Heredity* **71**:370–391.
163. Mears, J. G., Ramirez, F., Leibowitz, D., Nakamura, F., Bloom, A., Konotey-Ahulu, F., and Bank, A., 1978, Changes in restricted human cellular DNA fragments containing

globin gene sequences in thalassemias and related disorders, *Proc. Nat. Acad. Sci. USA* **75**:1222–1226.

164. Meera Khan, P., Doppert, B. A., Hayemeijer, A., and Westerveld, A., 1973, The human loci for phosphopyruvate hydratase and guanylate kinase are syntenic with the PGD-PGM1 linkage group in man–Chinese hamster somatic cell hybrids, *Cytogenet. Cell Genet.* **13**:130–131.
165. Messing, J., Crea, R., and Seeburg, P. H., 1981, A system for shotgun DNA sequencing, *Nucleic Acids Res.* **9**:309–321.
166. Miller, A. L., Freeze, H. H., and Kress, B. C., 1981, I-cell disease, in: *Lysosomes and Lysosomal Storage Diseases* (J. W. Callahan and J. A. Lowden, eds.), pp. 271–287. Raven Press, New York.
167. Miller, C. L., and Ruddle, F. H., 1978, Cotransfer of human X-linked markers into murine somatic cells via isolated metaphase chromosomes, *Proc. Nat. Acad. Sci. USA* **75**:3346–3350.
168. Mishur, M. K., and Nance, W. E., 1969, Further evidence for close linkage of the Hb^{β} and Hb^{δ} loci in man, *J. Med. Genet.* **6**:190–192.
169. Montaya, J., Ojala, D., and Attardi, G., 1981, Distinctive features of the 5′-terminal sequences of the human mitochondrial mRNA, *Nature* **290**:465–474.
170. Morrow, J. F., 1976, The prospects for gene therapy in humans, *Ann. N.Y. Acad. Sci.* **265**:13–21.
171. Mowbray, S., Watson, B., and Harris, H., 1972, A search for electrophoretic variants of human adenine phosphoribosyl transferase, *Ann. Hum. Genet. Lond.* **36**:153–161.
172. Nakaniski, S., Inoue, A., Keta, T., Nakamura, M., Chang, A. C. Y., Cohen, S. N., and Shosaku, N., 1979, Nucleotide sequence of cloned cDNA for bovine corticotropin-β-lipotropin precursor, *Nature* **278**:423–427.
173. Naylor, S. L., Klebe, R. J., and Shows, T. B., 1978, Argininosuccinic aciduria: Assignment of the argininosuccinate lyase gene to the pter→q22 region of human chromosome 7 by bioautography, *Proc. Nat. Acad. Sci. USA* **75**:6159–6162.
174. Naylor, S. L., Elliott, R. W., Brown, J. A., and Shows, T. B., 1982, Mapping of aminoacylase-1 and β-galactosidase-A to homologous regions of human chromosome 3 and mouse chromosome 9 suggests location of additional genes, *Am. J. Hum. Genet.*, in press.
175. Naylor, S. L., Sakaguchi, A. Y., Gusella, J. F., Housman, D., and Shows, T. B., 1982, Mapping of an arbitrary restriction polymorphism to human chromosome 3, *Cytogenet. Cell Genet.*, in press.
176. Naylor, S. L., Sakaguchi, A. Y., Schmickel, R. D., Gutai, M. G., and Shows, T. B., Submitted, Organization of rDNA spacer fragment variants among human acrocentric chromosomes in somatic cell hybrids.
177. Naylor, S. L., Sakaguchi, A. Y., Shows, T. B., Zasloff, M., and Holmes, M., in preparation, Human chromosome 6 contains two nonallelic $tRNA_i^{met}$ genes.
178. Niall, H. D., Hogan, M. L., Sauer, R., Rosenblum, Y., and Greenwood, F. E., 1971, Sequences of pituitary and placental lactogenic and growth hormones, *Proc. Nat. Acad. Sci USA* **68**:866–970.
179. Nishioka, Y., Leder, A., and Leder, P., 1980, Unusual α-globin-like gene that has clearly lost both globin intervening sequences, *Proc. Nat. Acad. Sci. USA* **77**:2806–2809.
180. Noyes, B. E., Mevarech, M., Stein, R., and Agarwal, K. L., 1979, Detection and partial sequence analysis of gastrin mRNA by using an oligodeoxynucleotide probe, *Proc. Nat. Acad. Sci. USA* **76**:1770–1774.
181. O'Farrell, P. H., Kutter, E., and Nakanishi, M., 1980, A restriction map of the bacteriophage T4 genome, *Molec. Gen. Genet.* **179**:421–435.

182. Ohno, S., 1973, Ancient linkage groups and frozen accidents, *Nature* **244**:259–262.
183. Ohta, T., and Kumura, M., 1971, Functional organization of genetic material as a product of molecular evolution, *Nature* **233**:118–119.
184. Orkin, S. H., Alter, B. P., Altay, C., Mahoney, M. J., Lazarus, H., Hobbins, J. C., and Nathan, D. G., 1978, Application of endonuclease mapping to the analysis and prenatal diagnosis of thalassemias caused by globin-gene deletion, *New Eng. J. Med.* **299**:166–172.
185. Orkin, S. H., Old, J. M., Weatherall, D. J., and Nathan, D. G., 1979, Partial deletion of β-globin gene DNA in certain patients with β^o-thalassemia, *Proc. Nat. Acad. Sci. USA* **76**:2400–2404.
186. Orkin, S. H., Kolodner, R., Michelson, A., and Husson, R., 1980, Cloning and direct examination of a structurally abnormal human β^o-thalassemia globin gene, *Proc. Nat. Acad. Sci. USA* **77**:3558–3562.
187. Owerbach, D., Doyle, D., and Shows, T. B., 1978, Genetics of the large, external, transformation-sensitive (LETS) protein: Assignment of a gene coding for expression of LETS to human chromosome 8, *Prog. Nat. Acad. Sci. USA* **75**:5640–5644.
188. Owerbach, D., Doyle, D., and Shows, T. B., 1979, Genetics of the cell surface: Assignment of genes coding for external membrane proteins to human chromosomes 10 and 14, *Som. Cell Genet.* **5**:281–301.
189. Owerbach, D., Bell, G. I., Rutter, W. J., and Shows, T. B., 1980, The insulin gene is located on chromosome 11 in humans, *Nature* **286**:82–84.
190. Owerbach, D., Rutter, W. J., Martial, J. A., Baxter, J. D., and Shows, T. B., 1980, Genes for growth hormone, chorionic somatomammotropin, and growth hormone-like gene on chromosome 17 in humans, *Science* **209**:289–292.
191. Owerbach, D., Bell, G. I., Rutter, W. J., Brown, J. A., and Shows, T. B., 1981, The insulin gene is located on the short arm of chromosome 11 in humans, *Diabetes* **20**:267–270.
192. Owerbach, D., Rutter, W. J., Cooke, N. E., Martial, J. A., and Shows, T. B., 1981, The prolactin gene is located on chromosome 6 in humans, *Science* **212**:815–816.
193. Owerbach, D., Rutter, W. J., Roberts, J. L., Whitfeld, P., Shine, J., Seeburg, P. H., and Shows, T. B., 1981, The proopiocortin (adrenocorticotropin/β-lipotropin) gene is located on chromosome 2 in humans, *Somat. Cell Genet.* **7**:359–369.
194. Owerbach, D., Rutter, W. J., Shows, T. B., Gray, P., Goeddel, D. V., and Lawn, R. M., 1981, The leukocyte and fibroblast interferon genes are located on human chromosome 9, *Proc. Nat. Acad. Sci. USA* **78**:3123–3127.
195. Paigen, K., 1979, Acid hydrolases as models of genetic control, *Ann. Rev. Genet.* **13**:417–466.
196. Palmeter, R. D., 1973, Ovalbumin messenger ribonucleic acid translation, *J. Biol. Chem.* **248**:2095–2106.
197. Panny, S. R., Scott, A. F., Smith, K. D., Phillips, J. A., III, Kazazian, H. H., Jr., Talbot, C. C., Jr., and Boehm, C. D., 1981, Population heterogeneity of the HpaI restriction site associated with the beta globin gene: Implications for prenatal diagnosis, *Am. J. Hum. Genet.* **33**:25–35.
198. Pardue, M. L., Bonner, J. J., Lengyel, J., and Spradling, A. S., 1977, *In situ* hybridization for the study of chromosome structure and function, in: *Molecular Human Cytogenetics* (R. S. Sparkes and D. E. Comings, eds.), pp. 217–232, Academic Press, New York.
199. Paris Conference, 1971, Standardization in human genetics, in: *Birth Defects: Original Article Series VIII,* Vol. 7, The National Foundation, New York.

200. Parks, J. S., Herd, J. E., and Wurzel, J. M., 1981, Human growth hormone (HGH) deficiency and polymorphism within the HGH and human placental lactogen (HPL) gene cluster, *Pediat. Res.* **15**:513.
201. Parnes, J. R., Velan, B., Felsenfeld, A., Ramanathan, L., Ferrini, U., Appella, E., and Seidman, J. G., 1981, Mouse β_2-microglobulin cDNA clones: A screening procedure for cDNA clones corresponding to rare mRNAs, *Proc. Nat. Acad. Sci. USA* **78**:2253–2257.
202. Pearson, P. L., Bakher, E., and Havell, R. A., 1982, Considerations in designing an efficient strategy for localizing unique sequence DNA fragments to human chromosomes, *Cytogenet. Cell Genet.*, in press.
203. Phillips, J. A., III, Panny, S. R., Kazazian, H. H., Jr., Boehm, C. D., Scott, A. F., and Smith, K. D., 1980, Prenatal diagnosis of sickle cell anemia by restriction endonuclease analysis: HindIII polymorphisms in γ-globin genes extend test applicability, *Proc. Nat. Acad. Sci. USA* **77**:2853–2856.
204. Popp, R. A., Lalley, P. A., Whitney, J. B., III, and Anderson, W. F., 1981, Mouse α-globin genes and α-like pseudo genes are not syntenic, *Proc. Nat. Acad. Sci. USA* **78**:6362–6366.
205. Reddy, V. B., Thimmappaya, B., Dhar, R., Subramanian, K. N., Zain, B. S., Pan, J., Ghosh, P. K., Celma, M. L., and Weissman, S. M., 1978, The genome of simian virus 40, *Science* **200**494–502.
206. Renwick, J. H., 1969, Progress in mapping human autosomes, *Br. Med. Bull.* **25**:65–73.
207. Renwick, J. H., 1971, The mapping of human chromosomes, *Ann. Rev. Genet.* **5**:81–120.
208. Riccuiti, F., and Ruddle, F. H., 1973, Assignment of nucleoside phosphorylase to D-14 and localization of X-linked loci in man by somatic cell genetics, *Nature New Biol.* **241**:180–182.
209. Rigby, P. W. J., Dieckmann, M., Rhodes, C., and Berg, P., 1977, Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I, *J. Mol. Biol.* **113**:237–251.
210. Roberts, R. J., 1980, Restriction and modification enzymes and their recognition sequences, *Nucleic Acids Res.* **8**:r63–r80.
211. Roberts, B. E., and Paterson, B. M., 1973, Efficient transcription of tobacco mosaic virus RNA and rabbit globin 9S RNA in a cell-free system from commercial wheat germ, *Proc. Nat. Acad. Sci. USA* **70**:2330–2334.
212. Robson, E. B., Polani, P. E., Dart, S. J., Jacobs, P. A., and Renwick, J. H., 1969, Probable assignment of the alpha locus of haptoglobin to chromosome 16 in man, *Nature* **223**:1163–1165.
213. Rosen, J. M., Woo, S. L., Holder, J. W., Mears, A. R., and O'Malley, B. W., 1975, Preparation and preliminary characteristics of purified ovalbumin messenger RNA from the hen oviduct, *Biochemistry* **14**:69–78.
214. Ross, J., Aviv, H., Scolnick, E., and Leder, P., 1972, *In vitro* synthesis of DNA complementary to purified rabbit globin on RNA, *Proc. Nat. Acad. Sci. USA* **69**:264–268.
215. Rotwein, P., Chyn, R., Chirgwin, J., Cordell, B., Goodman, H. M., and Permutt, M. A., 1981, Polymorphism in the 5′-flanking region of the human insulin gene and its possible relation to type 2 diabetes, *Science* **213**:1117–1120.
216. Rowley, J. D., 1980, Chromosome abnormalities in human leukemia, *Ann. Rev. Genet.* **14**:17–39.
217. Ruddle, F. H., 1972, Linkage analysis using somatic cell hybrids, *Adv. Hum. Genet.* **3**:173–235.

218. Ruddle, F. H., and Creagan, R. P., 1975, Parasexual approaches to the genetics of man, *Ann. Rev. Genet.* **9**:407–486.
219. Ruddle, F. H., and McBride, O. W., 1977, New approaches to cell genetics cotransfer of linked genetic markers by chromosome mediated gene transfer, in: *The Molecular Biology of the Mammalian Genetic Apparatus* (P. Tso, ed.), pp. 163–169, Elsevier, North-Holland.
220. Sakaguchi, A. Y., and Shows, T. B., 1982, Coronavirus 229E susceptibility in man–mouse hybrids is located on human chromosome 15. *Somat. Cell Genet.* **8**:83–94.
221. Sakaguchi, A. Y., Naylor, S. L., Quinto, C., Rutter, W. J., and Shows, T. B., 1982, The chymotrypsinogen B gene (*CTRB*) is on human chromosome 16, *Cytogenet. Cell Genet.*, in press.
222. Sakaguchi, A. Y., Naylor, S. L., Schmickel, R. D., and Shows, T. B., 1982, Assignment of an arbitrary restriction fragment, ARF-1, to human chromosome 6, *Cytogenet. Cell Genet.*, in press.
223. Sakano, H., Rogers, J. H., Huppi, K., Brack, C., Traunecker, A., Maki, R., Wall, R., and Tonegawa, S., 1979, Domains and the hinge region of an immunoglobulin heavy chain are encoded in separate DNA segments, *Nature* **277**:627–633.
224. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H., and Roe, B. A., 1980, Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing, *J. Mol. Biol.* **143**:161–178.
225. Schafer, M., and White, R., 1982, Three random loci in the human genome with base pair change polymorphisms, *Cytogenet. Cell Genet.*, in press.
226. Schmeckpeper, B. J., Willard, H. F., and Smith, K. D., 1981, Isolation and characterization of cloned human DNA fragments carrying reiterated sequences common to both autosomes and the X chromosome, *Nucleic Acids Res.* **9**:1853–1872.
227. Schmid, C. W., and Deininger, P. L., 1975, Sequence organization of the human genome, *Cell* **6**:345–358.
228. Scott, A. F., Phillips, J. A., and Migeon, B. R., 1979, DNA restriction endonuclease analysis for localization of human β- and δ-globin genes on chromosome 11, *Proc. Nat. Acad. Sci. USA* **76**:4563–4565.
229. Shows, T. B., 1972, Genetics of human–mouse somatic cell hybrids: Linkage of human genes for lactate dehydrogenase-*A* and esterase-A_4, *Proc. Nat. Acad. Sci. USA* **69**:348–352.
230. Shows, T. B., 1972, Genetics of human–mouse somatic cell hybrids: Linkage of human genes for malate dehydrogenase and isocitrate dehydrogenase, *Biochem. Genet.* **7**:193–204.
231. Shows, T. B., 1974, Somatic cell genetics of enzyme markers associated with three human linkage groups, in: *Somatic Cell Hybridization* (R. Davidson and F. de la Cruz, eds.), pp. 15–21, Raven Press, New York.
232. Shows, T. B., 1974, Synteny of human genes for glutamic oxaloacetic transaminase and hexokinase in somatic cell hybrids, *Cytogenet. Cell Genet.* **13**:143–145.
233. Shows, T. B., 1975, Genetics, expression, and characterization of isozymes in somatic cell hybrids, in: *Isozymes, III. Developmental Biology* (C. L. Markert, ed.), pp. 619–636, Academic Press, New York.
234. Shows, T. B., 1977, Genetic and structural dissection of human enzymes and enzyme defects using somatic cell hybrids, in: *Isozymes: Current Topics in Biological and Medical Research, Vol. 2* (M. C. Rattazzi, J. G. Scandalios, and G. S. Whitt, eds.), pp. 107–158, Alan R. Liss, New York.
235. Shows, T. B., 1978, Mapping the human genome and metabolic diseases, in: *Birth*

Defects (J. W. Littlefield and J. DeGrouchy, eds.), pp. 66–84, Excerpta Medica, Amsterdam.

236. Shows, T. B., 1979, The X chromosome gene map, in: *Genetic Mechanisms of Sexual Development* (H. L. Vallet and I. H. Porter, eds.), pp. 253–269, Academic Press, New York.
237. Shows, T. B., and Brown, J. A., 1975, Human X-linked genes regionally mapped utilizing X-autosome translocations and somatic cell hybrids, *Proc. Nat. Acad. Sci. USA* **72**:2125–2129.
238. Shows, T. B., and McAlpine, P. J., 1979, The 1979 catalog of human genes and chromosome assignments, *Cytogenet. Cell Genet.* **25**:117–127.

238a. Shows, T. B., and McAlpine, P. J., 1982, The 1981 catalog of assigned human genetic markers, *Cytogenet. Cell Genet.*, in press.

239. Shows, T. B., and Ruddle, F. H., 1968, Function of the lactate dehydrogenase *B* gene in erythrocytes: Evidence for control by a regulatory gene, *Proc. Nat. Acad. Sci. USA* **61**:574–581.
240. Shows, T. B., and Sakaguchi, A. Y., 1980, Gene transfer and gene mapping in mammalian cells in culture, *In Vitro* **16**:55–76.
241. Shows, T. B., Tashian, R. E., Brewer, G. J., and Dern, R. J., 1964, Erythrocyte glucose-6-phosphate dehydrogenase in Caucasians: New inherited variant, *Science* **145**:1056.
242. Shows, T. B., Ruddle, F. H., and Roderick, T. H., 1969, Phosphoglucomutase electrophoretic variants in the mouse, *Biochem. Genet.* **3**:25–35.
243. Shows, T. B., Chapman, V. M., and Ruddle, F. H., 1970, Mitochondrial malate dehydrogenase and malic enzyme: Mendelian inherited electrophoretic variants in the mouse, *Biochem. Genet.* **4**:707–718.
244. Shows, T. B., Champion, M. J., and Lalley, P. A. 1977, Expression of enzyme deficiencies in somatic cell hybrids: A strategy for the genetic and structural dissection of fatal lysosomal enzyme diseases, *Am. J. Hum. Genet.* **29**:99A.
245. Shows, T. B., Brown, J. A., Eddy, R. L., Byers, M. G., Haley, L. L., Cooper, E. S., and Goggin, A. P., 1978, Assignment of peptidase S (*PEPS*) to chromosome 4 in man using somatic cell hybrids, *Hum. Genet.* **43**:119–125.
246. Shows, T. B., Alper, C. A., Bootsma, D., Dorf, M., Douglas, T., Huisman, T., Kit, S., Klinger, H. P., Kozak, C., Lalley, P. A., Lindsley, D., McAlpine, P. J., McDougall, J. K., Meera Khan, P., Meisler, M., Morton, N. E., Opitz, J. M., Partridge, C. W., Payne, R., Roderick, T. H., Rubinstein, P., Ruddle, F. H., Shaw, M., Spranger, J. W., and Weiss, K., 1979, International system for human gene nomenclature, *Cytogenet. Cell Genet.* **25**:96–116.
247. Shows, T. B., Scrafford-Wolff, L., Brown, J. A., and Meisler, M. H., 1979, G_{M1}-Gangliosidosis: Chromosome 3 assignment of the β-galactosidase-A gene (βGAL_A), *Somat. Cell Genet.* **5**:147–158.
248. Shows, T. B., Mueller, O. T., Honey, N. K., Wright, C. E., and Miller, A. L., 1982, Genetic heterogeneity within I-cell disease is demonstrated by complementation of lysosomal enzyme processing mutants, *Am. J. Med. Genet.*, in press.
249. Sinet, P. M., Allard, D., Lejeune, S., and Jerome, H., 1974, Augmentation d'activite de la superoxyde dismutase erythrocytaire dans la trisomie pour le chromosome 21, *C. R. Acad. Sci.* **280**:3267–3270.
250. Slightom, J. L., Blechl, A. E., and Smithies, O., 1980, Human fetal $^{G}\gamma$- and $^{A}\gamma$-globin genes: Complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes, *Cell* **21**:627–638.

251. Smith, M., Krinsky, A. M., Arredondo-Vega, F., Wang, A.-L., and Hirschhorn, K., 1982, Confirmation of the assignment of genes for human immunoglobulin heavy chains to chromosome 14, *Cytogenet. Cell Genet.*, in press.
252. Solomon, E., Goodfellow, P., Chambers, S., Sperr, N., Hobart, M. J., Rabbits, T. H., and Povey, S., 1982, Confirmation of the assignment of immunoglobulin heavy chain genes to chromosome 14, using cloned DNA as molecular probes, *Cytogenet. Cell Genet.*, in press.
253. Sood, A. K., Pereira, D., and Weissman, S. M., 1981, Isolation and partial nucleotide sequence of a cDNA clone for human histocompatibility antigen HLA-B by use of an oligodeoxynucleotide primer, *Proc. Nat. Acad. Sci. USA* **78**:616–620.
254. Southern, E. M., 1975, Detection of specific sequences among DNA fragments separated by gel electrophoresis, *J. Mol. Biol.* **98**:503–517.
255. Spritz, R. A., Jagadeeswaran, P., Choudary, P. V., Biro, P. A., Elder, J. T., DeRiel, J. K., Manley, J. L., Gefter, M. L., Forget, B. G., and Weissman, S. M., 1981, Base substitution in an intervening sequence of a β^+-thalassemic human globin gene, *Proc. Nat. Acad. Sci. USA* **78**:2455–2459.
256. Steffensen, D. M., 1975, Human histone genes mapped to chromosome 7, *Cytogenet. Cell Genet.* **25**:211.
257. Steffensen, D. M., Prensky, W., Mutton, D., and Hamerton, J. L., 1975, Mapping the human 5S RNA gene on chromosome 1 using translocations, *Cytogenet. Cell Genet.* **14**:264–268.
258. Steinmetz, M., Moore, K. W., Frelinger, J. G., Sher, B. T., Shen, F.-W., Boyse, E. A., and Hood, L., 1981, A pseudogene homologous to mouse transplantation antigens: Transplantation antigens are encoded by eight exons that correlate with protein domains, *Cell* **25**:683–692.
259. Surrey, S., Ohene-Frempong, K., Rappaport, E., Atwater, J., and Schwatz, E., 1980, Linkage of $\alpha^{\text{G-Philadelphia}}$ to α-thalassemia in African-Americans, *Proc. Nat. Acad. Sci. USA* **77**:4885–4889.
260. Szybalski, W. S., Szybalska, E. H., and Ragni, G., 1962, Genetic studies with human cell lines, *Nat. Cancer Inst. Monogr.* **7**:75–88.
261. Tashima, M., Calabretta, B., Torelli, G., Scofield, M., Maizel, A., and Saunders, G. F., 1981, Presence of a highly repetitive and widely dispersed DNA sequence in the human genome, *Proc. Nat. Acad. Sci. USA* **78**:1508–1512.
262. Tiemeier, D. C., Tilghman, S. M., Polsky, F. I., Seidman, J. G., Leder, A., Edgell, M. H., and Leder, P., 1978, A comparison of two cloned mouse β-globin genes and their surrounding and intervening sequences, *Cell* **14**:237–245.
263. Timmis, K. N., Cohen, S. N., and Cabello, F. C., 1978, DNA cloning and the analysis of plasmid structure and function, *Prog. Mol. Subcell. Biol.* **6**:1–58.
264. Ulbrich, A., Dull, T. J., Gray, A., Brosius, J., and Sures, I., 1980, Genetic variation in the human insulin gene, *Science* **209**:612–615.
265. Vanin, E. F., Goldberg, G. I., Tucker, P. W., and Smithies, O., 1980, A mouse α-globin-related pseudogene lacking intervening sequences, *Nature* **286**:222–226.
266. Van Someren, H., Beyersbergen, von Henegouvien, H., Westerveld, A., and Bootsma, D., 1974, Synteny of the human loci for fumerate hydratase and UDPG pyrophosphorylase with chromosome 1 markers in somatic cell hybrids, *Cytogenet. Cell Genet.* **13**:551–557.
267. Verma, I. M., Temple, G. F., Fan, H., and Baltimore, D., 1972, *In vitro* synthesis of DNA complementary to rabbit reticulocyte 10S RNA, *Nature New Biol.* **235**:163–167.
268. Wallace, R. B., Johnson, M. J., Hirose, T., Miyake, T., Kawashima, E. H., and

Itakura, K., 1981, The use of synthetic oligonucleotides as hybridization probes. II. Hybridization of oligonucleotides of mixed sequence to rabbit β-globin DNA, *Nucleic Acids Res.* **9**:879–894.

269. Weiss, M. C., and Green, H., 1967, Human–mouse hybrid cell lines containing partial complements of human chromosomes and functioning human genes, *Proc. Nat. Acad. Sci. USA* **58**:1104–1111.
270. Wellauer, P. K., and Dawid, I. B., 1979, Isolation and sequence organization of human ribosomal DNA, *J. Mol. Biol.* **128**:289–303.
271. Wigler, M., Pellicer, A., Silverstein, S., Axel, R., Urlaub, G., and Chasin, L., 1979, DNA-mediated transfer of the adenine phosphoribosyltransferase locus into mammalian cells, *Proc. Nat. Acad. Sci. USA* **76**:1373–1376.
272. Willecke, K., Lange, R., Kruger, A., and Reber, T., 1976, Cotransfer of two linked human genes into cultured mouse cells, *Proc. Nat. Acad. Sci. USA* **73**:1274–1278.
273. Wolf, S. F., Mareni, C. E., and Migeon, B. R., 1980, Isolation and characterization of cloned DNA sequences that hybridize to the human X chromosome, *Cell* **21**:95–102.
274. Wullems, G. J., van der Horst, J., and Bootsma, D., 1977, Transfer of the human gene coding for thymidine kinase and galactokinase to Chinese hamster cells and human–Chinese hamster cell hybrids, *Som. Cell Genet.* **3**:281–293.
275. Wyman, A. R., and White, R., 1980, A highly polymorphic locus in human DNA, *Proc. Nat. Acad. Sci. USA* **77**:6754–6758.
276. Young, B. D., Krumlauf, R., Wiedemann, L., and Jeanpierre, M., 1982, Cloning of a DNA library from human chromosome no. 22, in: Abstract book, 6th International Congress of Human Genetics, Israel.
277. Yunis, J. J., Sawyer, J. R., and Ball, D. W., 1978, Characterization of banding patterns of metaphase-prophase G-banded chromosomes and their use in gene mapping, *Cytogenet. Cell Genet.* **22**:679–683.