Human genetics

Article by:

Thompson, Margaret W. Departments of Medical Genetics and Pediatrics, University of Toronto and the Hospital for Sick Children, Toronto, Ontario, Canada.

Denniston, Carter L. Department of Medical Genetics, University of Wisconsin, Madison, Wisconsin.

Vogel, Friedrich Institute for Anthropology and Human Genetics, Im Neuenheimer Feld, Germany.

Relethford, John H. Department of Anthropology, State University of New York, Oneonta, New York.

Nyhan, William L. Department of Pediatrics, University of California, San Diego, California.

Povey, Sue Galton Laboratory, University College, London, United Kingdom.

Publication year: 2014

DOI: http://dx.doi.org/10.1036/1097-8542.324600 (http://dx.doi.org/10.1036/1097-8542.324600)

Content

A discipline concerned with genetically determined resemblances and differences among human beings. The idea that certain physical and mental characteristics, normal or abnormal, can “run in families” goes back to ancient times, though the mechanism by which heredity operates remained mostly unknown until the twentieth century. Formerly, genetics was thought to be concerned only with the familial transmission of rare and insignificant characteristics, but its fundamental biological role is now apparent. Genes, the units of heredity, have two unique properties: they are self-replicating, and they carry in their biochemical structure the codes for protein synthesis. Consequently, genes play the double role of transmitting genetic information from generation to generation and of governing all the activities of living cells. *See also:* **Gene (/content/gene/284400)**

Expansion of human genetic knowledge has come from several different directions and has had major consequences for human health and for the understanding of the place of humans in nature. Technological advances in the visualization of human chromosomes have shown that abnormalities of chromosome number or structure are surprisingly common and of many different kinds, and that they account for birth defects or mental impairment in many individuals as well as for numerous early spontaneous abortions. Progress in molecular biology has clarified the molecular structure of chromosomes and their constituent genes and the mechanisms of deoxyribonucleic acid (DNA) synthesis and protein synthesis, as well as the ways in which change in the molecular structure of a gene can lead to a disease. Concern about possible genetic damage through environmental agents, particularly ionizing radiation, and the possible harmful effects of hazardous substances in the environment on prenatal development has also stimulated research in human genetics. The medical aspects of human genetics have become prominent as nonhereditary causes of ill health or early death, such as infectious disease or nutritional

deficiency, have declined, at least in developed countries. It has been estimated, for example, that half of all admissions to pediatric hospitals involve disorders that are completely or partly genetic. Detailed knowledge of the distribution of many genetic traits in individuals, families, and populations has expanded, providing fresh insights into human origins and the process of evolution.

Chromosome and gene structure

In normal humans, the nucleus of each normal cell contains 46 chromosomes, which comprise 23 different pairs. A karyotype (**Fig. 1**) may be prepared from any type of cell that will divide in cell culture, such as white blood cells, skin, bone marrow, or cells from amniotic fluid, for example. Of each chromosome pair, one is paternal and the other maternal in origin. In turn, only one member of each pair is handed on through the reproductive cell (egg or sperm) to each child. Thus, each egg or sperm has only 23 chromosomes, the haploid number; fusion of egg and sperm at fertilization will restore the double, or diploid, chromosome number of 46. *See also:* **Chromosome (/content/chromosome/134900)**; **Fertilization (animal) (/content/fertilization-animal/254900)**

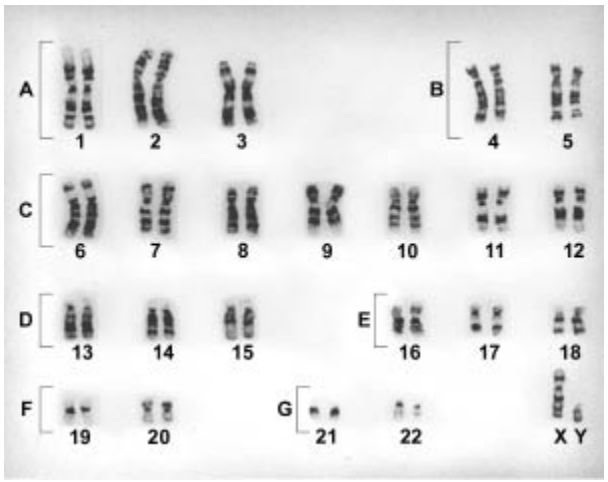

Fig. 1 Normal male karyotype. The chromosomes are classified in seven groups labeled A–G. The chromosome pairs are individually labeled. (*Photomicrograph courtesy of R. G. Worton*; *from J. S. Thompson and M. W. Thompson, Genetics in Medicine, 3d ed., W. B. Saunders, Philadelphia, 1980*)

The segregation of chromosome pairs during meiosis allows for a large amount of "shuffling" of genetic material as it is passed down the generation. With 23 pairs of chromosomes, the total possible number of different chromosome combinations in the gametes is 2^{23}, or about 8 million. Two parents can provide $2^{23} \times 2^{23}$ different chromosome combinations. This enormous source of variation is multiplied still further by the mechanism of crossing over, in which homologous chromosomes exchange segments during meiosis. *See also:* **Crossing-over (genetics) (/content/crossing-over-genetics/168850)**; **Meiosis (/content/meiosis/413500)**

Twenty-two of the 23 chromosome pairs, the autosomes, are alike in both sexes; the other pair comprises the sex chromosomes. A female has a pair of X chromosomes; a male (Fig. 1) has a single X, paired with a Y chromosome that he has inherited from his father and will transmit to each of his sons. Sex is determined at fertilization, and depends on whether the egg (which has a single X chromosome) is fertilized by an X-bearing or a Y-bearing sperm. *See also:* **Sex determination (/content/sex-determination/617400)**

All chromosomes are composed of DNA. It is estimated that the total length of DNA in the haploid set of human chromosomes is 1.7 m (5.7 ft), which is 10,000 times the length of the chromosomes at metaphase. Genes are segments of DNA. There may be as many as 20,000–25,000 genes on the 46 human chromosomes, of which about 3000 have been

identified and about 1500 have been mapped to specific chromosomal locations. *See also:* **Deoxyribonucleic acid (DNA) (/content/deoxyribonucleic-acid-dna/186500)**; **Genetic mapping (/content/genetic-mapping/285200)**

The genetic information is coded in DNA in the form of triplets of four bases: two purines, adenine (A) and guanine (G), and two pyrimidines, thymine (T) and cytosine (C). Each triplet combination (codon) codes for a specific amino acid. The sequence of bases in a specific gene dictates the sequence of amino acids in the specific protein coded by that gene. *See also:* **Genetic code (/content/genetic-code/284900)**

A gene initiates synthesis of a protein by transcription of the DNA into messenger ribonucleic acid (mRNA), which is a single-stranded molecule complementary to the DNA. After processing within the cell nucleus, the mRNA moves into the cytoplasm, where it binds to ribosomes. The translation of RNA into protein takes place on the ribosomes, which are small particles in the cytoplasm composed of protein and a special type of RNA, ribosomal RNA. Another type of RNA, transfer RNA, of which there is one type for each amino acid, moves amino acids from the cytoplasm to the mRNA molecule, where they are aligned in the sequence dictated by the genetic code carried in the mRNA. The sequence of amino acids forms a polypeptide, which is released from the ribosome when complete. An average protein contains about 500 amino acids; thus an average gene contains about 1500 base pairs. Since the haploid chromosome set contains 3 billion base pairs, the chromosomes must contain a large excess of DNA in addition to the DNA in the constituent genes. *See also:* **Protein (/content/protein/550200)**; **Ribonucleic acid (RNA) (/content/ribonucleic-acid-rna/589000)**; **Ribosomes (/content/ribosomes/589200)**

A typical gene is not a simple uninterrupted length of DNA. Most genes are made up of coding sequences (exons) separated by noncoding regions (intervening sequences, or introns). Transcription of the gene into RNA begins at an initiation site in advance of the first coding sequence and terminates beyond the end of the last sequence. After the entire DNA segment has been transcribed into mRNA, the intervening sequences are removed and the coding sequences are spliced together before the mRNA is translated into a polypeptide (**Fig. 2**). *See also:* **Exon (/content/exon/248350)**; **Intron (/content/intron/350450)**

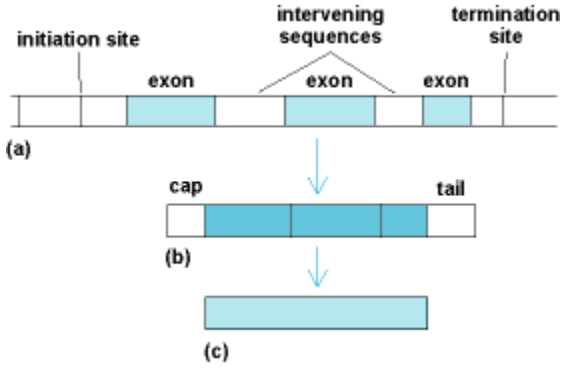

Fig. 2 Simplified diagram showing (*a*) the structure of a gene, (*b*) the mature messenger RNA after the two intervening sequences have been spliced out and the "cap" and "tail" added, and (*c*) the polypeptide translated from the messenger RNA.

Any gene occupies a specific chromosomal position, or locus. The alternative genes at a particular locus are said to be alleles. If a pair of alleles are identical, the individual is homozygous; if they are different, the individual is heterozygous. *See also:* **Allele (/content/allele/024000)**

Mutation

Genetic variation has its origin in mutation. Although in broad terms any change in DNA is a mutation, whether it is a microscopically detectable change in the structure of a chromosome or a single base change in the genetic code, the term is usually applied to stable changes in DNA that alter the genetic code and thus lead to synthesis of an altered protein.

Mutation can occur in reproductive cells or somatic cells, but the genetically significant ones are those that occur in reproductive cells and can therefore be transmitted to future generations. Natural selection acts upon the genetic diversity generated by mutation to preserve beneficial mutations and eliminate deleterious ones.

A very large amount of genetic variation exists in the human population. Everyone carries many mutations, some newly acquired but others inherited through innumerable generations. Though the exact number is unknown, it is likely that everyone is heterozygous at numerous loci, perhaps as many as 20%. *See also:* **Mutation (/content/mutation/441200)**

Single-gene inheritance

The patterns of inheritance of characteristics determined by single genes or gene pairs depend on two conditions: (1) whether the gene concerned is on an autosome (autosomal) or on the X chromosome (X-linked); (2) whether the gene is dominant, that is, expressed in heterozygotes (when it is present on only one member of a chromosomal pair and has a normal allele) or is recessive (expressed only in homozygotes, when it is present on both chromosomes). *See also:* **Dominance (/content/dominance/203000)**

Margaret W. Thompson

Quantitative inheritance

A quantitative trait is one that is under the control of many factors, both genetic and environmental, each of which contributes only a small amount to the total variability of the trait. The phenotype may show continuous variation (for example, height and skin color), or quasicontinuous variation (taking only integer values, such as the number of ridges in a fingerprint), or it may be discontinuous (a presence/absence trait, such as diabetes or mental retardation). With discontinuous traits, it is assumed that there exists an underlying continuous variable and that individuals having a value of this variable above (or below) a threshold possess the trait.

A trait that “runs in families” is said to be familial. However, not all familial traits are hereditary because relatives tend to share common environments as well as common genes. It is the major task of quantitative genetics to disentangle the effects of environment and heredity, but this task is not easy in humans where environment and heredity are often confounded.

The variability of almost any trait is partly genetic and partly environmental. A rough measure of the relative importance of heredity and environment is an index called heritability. R. A. Fisher showed that the total variance (a statistical measure of variability) of a trait can be partitioned into a genetic variance and an environmental variance, at least in simple cases. The quantitative model can be expressed as:

$$\text{Variance (phenotypic)} = \text{variance (genetic)} + \text{variance (environmental)} \quad (1)$$

Heritability is then defined as:

$$\text{Heritability} = \frac{\text{variance (genetic)}}{\text{variance (phenotypic)}} \quad (2)$$

In humans, the heritability of height is about 0.75. That is, about 75% of the total variance in height is due to variability in genes that affect height and 25% is due to exposure to different environments. Expressed in another way, the difference in height between two individuals is, on the average, 75% genetic and 25% environmental.

There are actually two kinds of heritability. The one described above, usually called broad heritability, measures the total

effect of heredity. The other, narrow heritability, is the proportion of total phenotypic variance resulting from the additive effects of genes. Narrow heritability is a more subtle concept but is actually more useful. For example, the correlation between parent and child is equal to one-half the narrow heritability plus their environmental correlation. Equations of this sort relate heritability to observable phenotypic correlations, thus providing a means of estimating the relative importance of heredity and environment.

Carter Denniston

Hereditary Diseases

Medicine is an important field for the practical application of human genetics; medical genetics has become an integral part of preventive medicine (that is, genetic counseling, including prenatal diagnostics). It has contributed increasingly to systematics of disease, diagnostics, and even therapy. Many external causes of disease, such as infections, have been brought under control in the twentieth century; therefore, doctors can devote much of their skill to treating diseases from internal sources, that is, hereditary diseases, or those brought about by interaction of genetic predispositions with certain stress factors in the environment. Widespread use of genetic knowledge in medical practice has important consequences for basic science: problems posed by the numerous and often unexpected observations in medical genetics help in developing basic theory, and suggest new approaches in research.

Hereditary diseases may be subdivided into three classes: chromosomal diseases; hereditary diseases with simple, mendelian modes of inheritance; and multifactorial diseases.

Chromosomal diseases

One out of 200 newborns suffers from an abnormality that is caused by a microscopically visible deviation in the number or structure of chromosomes. Such chromosomal aberrations are much more common (≈40%) among spontaneous miscarriages. About 10–20% of all recognized pregnancies terminate in spontaneous miscarriage and many more embryos die during the first weeks of pregnancy, when fetal loss goes unnoticed. Hence, a large fraction of all human zygotes are abnormal chromosomally and die early. *See also:* **Chromosome aberration (/content/chromosome-aberration/135000)**

The most important clinical abnormality among the survivors is Down syndrome—a condition due to trisomy of chromosome 21, one of the smallest human chromosomes. The term trisomy means that this chromosome is present not twice but three times; the entire chromosome complement therefore comprises 47, not 46, chromosomes. Down syndrome occurs one to two times in every 1000 births; its pattern of abnormalities derives from an imbalance of gene action during embryonic development. Down syndrome is a good example of a characteristic pattern of abnormalities that is produced by a single genetic defect. Such patterns, recognizable to the experienced observer as syndromes, are found not only in chromosomal diseases but in hereditary diseases with simple (mendelian) modes of inheritance as well. *See also:* **Down syndrome (/content/down-syndrome/204400)**

Other autosomal aberrations observed in living newborns that lead to characteristic syndromes include trisomies 13 and 18 (both very rare), and a variety of structural aberrations such as translocations (exchanges of chromosomal segments between different chromosomes) and deletions (losses of chromosome segments). Translocations normally have no influence on the health status of the individual if there is no gain or loss of chromosomal material (these are called balanced translocations). However, carriers of balanced translocations usually run a high risk of having “unbalanced” offspring—children in whom the same translocation causes gain or loss of genetic material, and who suffer from a characteristic malformation syndrome.

Clinical syndromes caused by specific aberrations vary, but certain clinical signs are common: low birth weights (small for

date); a peculiar face; delayed general, and especially mental, development, often leading to severe mental deficiency; and multiple malformations, including abnormal development of limbs, heart, and kidneys. Single malformations in children who otherwise develop normally are not typical for a chromosomal aberration. *See also:* **Congenital anomalies (/content/congenital-anomalies/156500)**

Less severe signs than those caused by autosomal aberrations are found in individuals with abnormalities in number (and, sometimes, structure) of sex chromosomes. This is because, in individuals having more than one X chromosome, the additional X chromosomes are inactivated early in pregnancy. For example, in women, one of the two X chromosomes is always inactivated. Inactivation occurs at random so that every normal woman is a mosaic of cells in which either one or the other X chromosome is active. Additional X chromosomes that an individual may have received will also be inactivated; in trisomies, genetic imbalance is thus avoided to a certain degree. However, inactivation is not complete; therefore, individuals with trisomies—for example, XXY (Klinefelter syndrome), XXX (triple-X-syndrome), or XYY—or monosomies (XO; Turner syndrome) often show abnormal sexual development, intelligence, or behavior.

In some individuals, chromosomal aberrations are found only in some cells. Clinical signs in these cases are often milder. Sometimes, a new mutation giving rise to a structural chromosomal aberration may occur in a somatic tissue, leading, for example, to a translocation only in one cell and its descendants. Sometimes, especially when a chromosome break involved in this translocation has affected an oncogene, these cells may have a selective advantage, and develop into a malignant tumor. An example is the translocation between chromosomes 22 and 9 found in chronic myelonic leukemia. *See also:* **Mosaicism (/content/mosaicism/435500)**; **Oncogenes (/content/oncogenes/468950)**

Diagnosis of chromosomal aberrations, especially in newborns with multiple malformations, individuals with disturbances of sexual development, and parents suffering from multiple miscarriages, are of practical importance, since in many cases monitoring of future pregnancies by prenatal diagnosis is possible.

Diseases with mendelian inheritance

In contrast to chromosomal aberrations, the genetic defects in hereditary diseases with simple, mendelian modes of inheritance cannot be recognized by microscopic examination; as a rule, they must be inferred more indirectly from the phenotype and the pattern of inheritance in pedigrees. The defects are found in the molecular structure of the DNA. Often, one base pair only is altered, although sometimes more complex molecular changes, such as deletions of some bases or abnormal recombination, are involved. Methods of molecular biology have permitted in some cases direct analyses of such defects at the gene level.

Approximately 1% of all newborns have, or will develop during their lives, a hereditary disease showing a simple mendelian mode of inheritance.

Mendel called an allele “dominant” when the homozygote and the heterozygote were indistinguishable phenotypically; in experimental genetics, the same convention is still being used. In medical genetics, the terms dominant and recessive are not used so strictly. A condition is called dominant if the heterozygotes deviate in a clearly recognizable way from the normal homozygotes, in most cases by showing an abnormality. Since such dominant mutations are usually rare, almost no homozygotes are observed and their clinical condition is, in most cases, unknown. In exceptional instances, homozygotes for dominant conditions have been described; they have usually shown more severe clinical signs than the heterozygotes.

In **Fig. 3**, there are four mendelian crosses for one pair of alleles A and A′. Cross no. 2, the backcross between the normal ' homozygote and the heterozygote, which leads to 1:1 segregation, is found most commonly in autosomal dominant ' inheritance. Crosses 3 and 4 are extremely rare, since homozygotes for an abnormal allele are usually rare. In autosomal-

recessive inheritance, backcrosses between normal homozygotes and heterozygotes (cross no. 2) are also common; however, since heterozygotes are normal phenotypically, such crosses will go unnoticed in most instances. The most common cross leading to homozygous, clinically affected offspring is the intercross between two heterozygotes, AA′, which results in 1:2:1 segregation.

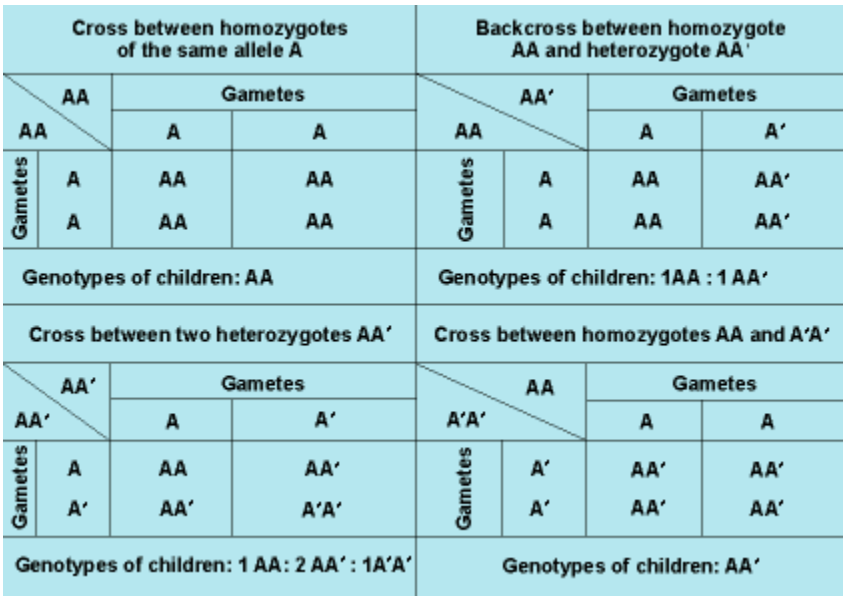

Fig. 3 Examples of mendelian crosses using two alleles, A and A′.

Autosomal-dominant inheritance

This type of inheritance pattern is often determined by studying the history of a trait among a group of relatives. Such a history, called a pedigree, can be represented in a standard chart (**Fig. 4**). A pedigree in which a rare autosomal-dominant condition is transmitted through four generations is shown in **Fig. 5**. The mutant allele is located on an autosome; the affected individuals are heterozygous. Since one of the two alleles at this gene locus is altered by mutations and each of the two alleles has a 50% probability of being transmitted to a child (Fig. 3, cross no. 2), each child has a 50% risk of being affected. Since this mutation is autosomal, it is transmitted independently of the sex chromosomes and the risk of being affected is not influenced by the sex of parents or child.

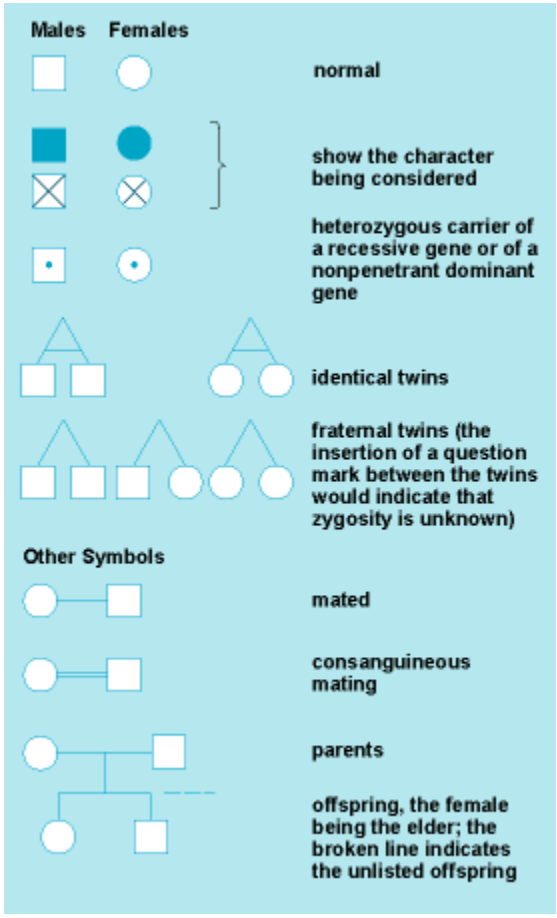

Fig. 4 Some symbols in common use in the United States for presenting human pedigrees.

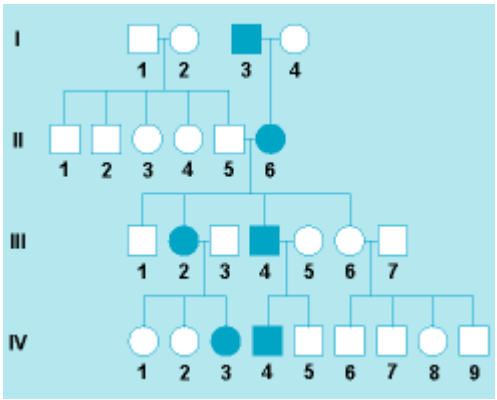

Fig. 5 Typical pedigree pattern of a rare autosomaldominant trait: classical achondroplastic dwarfism. Those with this anomaly, caused by a defect in growth of the long bones, have extremely short arms and legs, but are otherwise normal. I–IV are generations; shaded individuals manifest the trait.

The pedigree in Fig. 5 demonstrates the characteristic features of autosomal-dominant inheritance. For conditions such as achondroplastic dwarfism, however, such large pedigrees are the exception rather than the rule. In most cases, known pedigrees extend over two generations only. It is also common that one person is the only affected individual in a family. These individuals owe their abnormal allele to a fresh mutation in the germ cell of one of their parents, but the risk for their children to become affected is, nevertheless, 50% (**Fig. 6**). The fraction of new mutants among all carriers of a certain dominant disease must be high if the disease impairs average reproduction of its carriers: the harmful alleles are eventually eliminated from the population since many carriers have no children. When the carriers are so severely affected that they have no children at all, each new mutant will be eliminated in the first generation, and all cases in a population are new mutants. On the other hand, large pedigrees with many affected individuals are usually observed when the anomaly is

relatively harmless or manifests itself later in life, at a time when the carriers already have had their children. An example is Huntington's disease, a severe degenerative disease of the brain: in most gene carriers, the first clinical signs become visible only between 40 and 50 years of age; they die after many years of progressive deterioration of brain function.

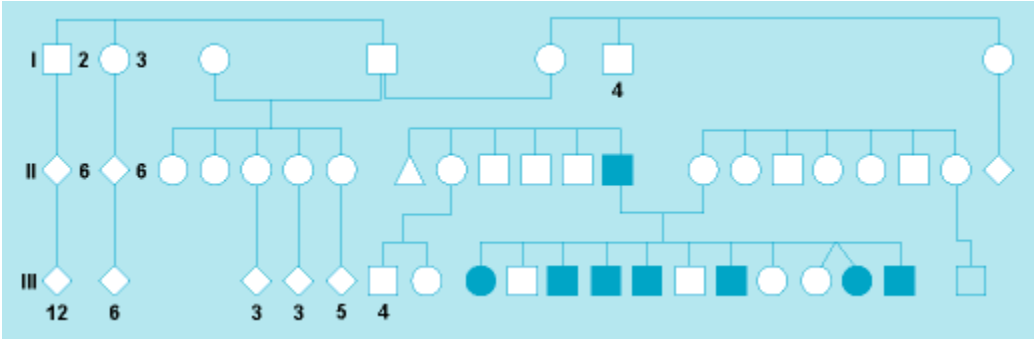

Fig. 6 Pedigree with new mutation to autosomal-dominant aniridia (defect of irises). The mutation must have occurred in the germ cell of the father or mother of the patient in generation II. Diamonds represent the indicated number of unaffected offspring. (*After F. Vogel and A. G. Motulsky, Human Genetics: Patterns and Approaches, Springer-Verlag, Berlin/New York, 1979*)

In some dominant conditions, the harmful phenotype may not be expressed in a gene carrier (this is called incomplete penetrance), or clinical signs may vary in severeness between carriers (called variable expressivity). Penetrance and expressivity may be influenced by other genetic factors—sometimes, for example, by the sex of the affected person; in other instances, the constitution of the “normal” allele has been implicated. Environmental conditions may occasionally be important. In most cases, however, the reasons are unknown.

Autosomal-recessive inheritance

A pedigree for albinism, that is, with an autosomal-recessive trait, is shown in **Fig. 7**. The affected individuals (IV, 2 and 6) have two mutant alleles, one from each parent; they are homozygous for the albinism gene. Hence, their unaffected parents, III, 4 and 5, must both be heterozygous for this allele. The probability of two individuals having the same allele increases when some of their genes come from a common ancestor (that is, when they are relatives). Indeed, the mothers of the parents in Fig. 4 are sisters; hence, the parents themselves are first cousins. An increase of matings between relatives, for example, first cousins, in comparison with the population average is characteristic for rare autosomal-recessive diseases. There is a segregation of 1:2:1 between homozygotes AA, heterozygotes Aa, and homozygotes aa. This means that every child from a mating of two heterozygotes (Fig. 3, cross no. 3) has a 25% risk of being homozygous for the mutant allele, and thus affected.

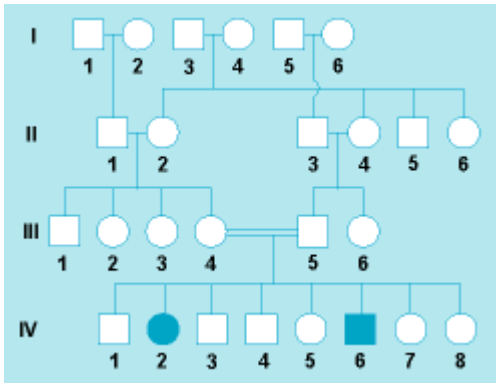

Fig. 7 Typical pedigree pattern of a rare autosomal-recessive trait, one form of albinism. Note, however, that sibships with so many sibs as well as marriages between first cousins have become rare in populations of industrialized countries.

In the first decades of the twentieth century, when the first autonomal-recessive conditions such as albinism were discovered, both criteria—an increase of matings between relatives, and appearance of the disease in 25% of children—were often observed, and were useful for genetic analysis. In the 1980s, both indicators have become much rarer: with the average number of one to two children per marriage that is observed in many industrialized countries, there will be one affected sib only in the great majority of all sibships. Since the fraction of first-cousin marriages is only two to three per 1000 marriages, even a tenfold increase does not lead to an appreciable fraction among parents of such patients. Hence, recognizing an autosomal-recessive mode of inheritance has become more difficult.

As a rule of thumb, autosomal-dominant mutations often lead to structural anomalies such as malformation syndromes, ' whereas autosomal-recessive mutations can often be traced back to a biochemical abnormality, for example, an enzyme ' defect. '

X-linked inheritance

X-linked modes of inheritance occur when the mutant allele is located on the X chromosome. The family patterns result from the mechanism of phenotypic sex determination: children receiving a Y chromosome from their fathers become males, and those receiving the paternal X become females. The mothers contribute an X chromosome to children irrespective of their sex. Therefore, all daughters and no sons receive an X-linked mutant allele from their fathers, whereas sons as well as daughters may receive such an allele from the mothers, the chance always being 50%, if the mother is heterozygous for the trait.

The most important X-linked mode of inheritance is the recessive one (**Fig. 8**). Here, the males (referred to as hemizygotes since they have only one allele) are affected, since they have no normal allele. The female heterozygotes, on the other hand, will be unaffected, since the one normal allele is sufficient for maintaining function. A classical example is hemophilia A, in which one of the serum factors necessary for normal blood clotting is inactive or lacking. (The disease can now be controlled by repeated substitution of the deficient blood factor—a good example for phenotypic therapy of a hereditary disease by substitution of a deficient gene product.) As shown in Fig. 8*a*, male family members are affected, whereas their sisters and daughters, while being unaffected themselves, transmit the mutant gene to half their sons. Only in very rare instances, when a hemophilic patient marries a heterozygous carrier, are homozygous females observed (Fig. 8*b*). As for autosomal-dominant diseases, the classic pedigrees are rare for severe X-linked conditions. Again, many of them are new mutants, and most actual pedigrees are small.

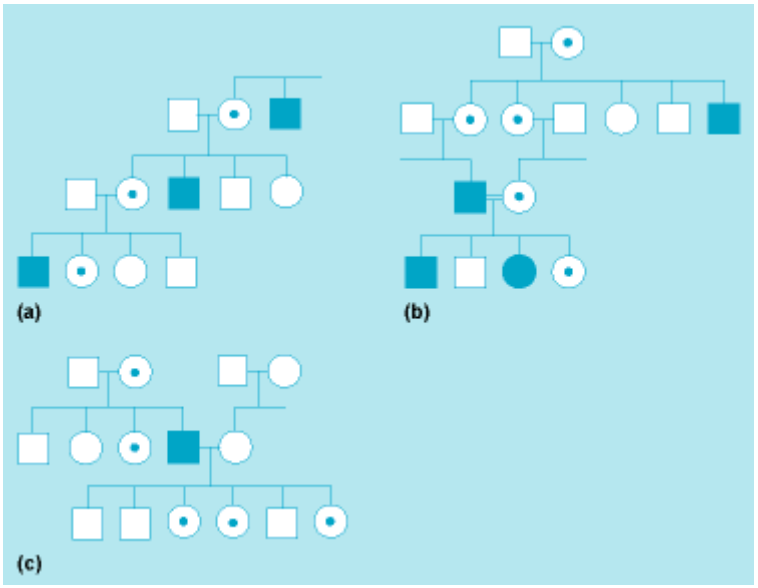

Fig. 8 Typical pedigree patterns of a sex-linked recessive trait, hemophilia A. (*a*) Only males are affected and usually come from unaffected mothers who often have affected fathers or brothers. (*b*) An affected female can result from an affected father and a carrier woman. (*c*) All daughters of affected males are carriers.

Some X-linked conditions are dominant, such that female heterozygotes express the abnormal trait. As a rule, however, their clinical signs are milder than those found in male hemizygotes. In quite a few such conditions, male hemizygotes are so severely damaged that they die even before birth, and are aborted. Hence, (almost) exclusively female patients are observed; there is a 1:1 ratio of normal and affected daughters of affected mothers, sex ratio among liveborn children is shifted in favor of girls, and miscarriages are common. Mutant genes located on the Y chromosome would be expected to be transmitted from the father to sons only, but to all sons. No such example has been confirmed. *See also:* **Sex-linked inheritance (/content/sex-linked-inheritance/617600)**

Multifactorial diseases

There are thousands of hereditary diseases with simple mendelian modes of inheritance, but most common anomalies and ' diseases are influenced by genetic variability at more than one gene locus. Most congenital malformations, such as ' congenital heart disease, cleft lip and palate, neural tube defects, and many others, fall into this category, as do the ' constitutional diseases, such as diabetes mellitus, coronary heart disease, anomalies of the immune response, and many ' mental diseases, including schizophrenia and affective disorders. '

All of these conditions are common and often increase in frequency with advanced age. In industrialized societies, almost everybody dies sooner or later of such a multifactorial disease, in which neither a chromosome aberration nor a simple mode of inheritance can be detected. The influence of genetic constitution can be indicated in a variety of ways, including a higher concordance of monozygotic as compared with dizygotic twins; a higher incidence of the same disease among relatives of affected probands than in the general population; and similarity of adopted children with their biological, not their adoptive, parents. However, there is no simple 1:1 relationship between genotype and phenotype in these cases; a variety of genetic factors may be involved. Moreover, the environment usually contributes significantly to the disease risk, which can be concluded from the observation that monozygotic twins, while being affected concordantly more often than dizygotics, are discordant in many cases.

In the absence of a more penetrating analysis, such a situation is described preliminarily by the genetic model of multifactorial inheritance in populations with a threshold effect (**Fig. 9**). The conceptual background of the multifactorial model may be described as follows: many undefined genes cooperate in creating a disease liability, which is distributed normally among individuals of a population. When this liability exceeds a threshold, the individual will be clinically affected. Instead of a sharp threshold, as shown in Fig. 9, a threshold area may be assumed in which manifestation depends on additional environmental factors. From such a genetic model, and the theory of quantitative genetics, some conclusions may be derived that are often found to apply more or less to empirical data; for example, that close relatives are affected more often than remote relatives, or that relatives of more severely affected probands run a higher risk than those of less severely affected ones. However, the multifactorial model does not attempt to identify, much less to characterize, the individual genes involved in a disease liability. The genotype is treated as “black box.”

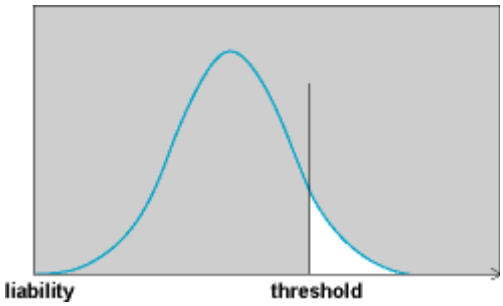

Fig. 9 Multifactorial inheritance in combination with a threshold effect. The figure shows the distribution of disease liabilities in a population (ordinate: number of individuals; abscissa: degree of disease liability). This distribution is assumed to be normal: many individuals have an average liability; in a few, liability is low; and in some, it is high. Individuals having liabilities on the right-hand side of the threshold (white area) are affected. (*After F. Vogel and A. G. Motulsky, Human Genetics: Problems and Approaches, Springer-Verlag, Berlin/New York, 1979*)

The multifactorial model poses questions rather than answering them. A meaningful answer can come only from analysis of the contribution of single genes (and specific environmental factors) to such disease liabilities. Some such analyses have been successful. The ABO blood groups, for example, influence liability to many common diseases, such as stomach cancer and cancers of some other organs; variants of the major histocompatibility gene (HLA) contribute to the liability for some rheumatic and autoimmune diseases; and $alpha_1$-antitrypsin variants are involved in chronic obstructive pulmonary disease. Here, the risk increases when the carriers of such alleles suffer from repeated bronchial infections—for example, heavy cigarette smokers or laborers working in a dusty environment. This is a good example of a specific interaction between a genetic liability and an environmental stress factor. Such examples are studied in a special branch of medical genetics that is known as ecogenetics.

Heterogeneity for autosomal-recessive mutations, while not leading normally to disease, may contribute to disease liabilities under certain external conditions. Often, specific disease units with clear-cut mendelian modes of inheritance have been identified by a combination of clinical, biochemical, and genetic methods in a minority of individuals and families within a bulk of conditions hitherto described as multifactorial. There are good chances that future research will help to open the “black boxes” of multifactorial inheritance step by step by analysis of single genes, the mechanisms of their action, and the ways in which they interact with each other and with specific stress factors in the environment.

Friedrich Vogel

Population Genetics

Population genetics is the mathematical basis of evolutionary theory. It is concerned with the frequency of genes and genotypes in a population, their relationship, and how they change over time. A major concern is the frequency of alleles, which are different forms of a gene or DNA sequence. Many principles of population genetics are easily understood using a simple model of a locus with two alleles, *A* and *a*. There are three possible genotypes: *AA, Aa*, and *aa*. A starting point in population genetics is to determine the relative frequency of each allele. It is conventional to use the symbol *p* to refer to the relative frequency of the *A* allele, and the symbol *q* to refer to the relative frequency of the *a* allele. If each genotype can be uniquely identified, the allele frequencies can be determined by simple counting. The number of individuals with each genotype are first counted and expressed as a relative frequency. If, for example, there are 18 individuals with genotype *AA*, 24 with genotype *Aa*, and 8 with genotype *aa*, there will be a total of 50 individuals, and the relative frequencies will be *AA* = 18/50 = 0.36, *Aa* = 24/50 = 0.48, and *aa* = 8/50 = 0.16. The frequency of the *A* allele is determined by taking the frequency of genotype *AA* plus half the frequency of genotype *Aa*, giving $p = 0.36 + 0.48/2 = 0.6$. Likewise, the frequency of the *a* allele is determined by taking the frequency of genotype *aa* plus half the frequency of genotype *Aa*, giving $q = 0.16 + 0.48/2 = 0.4$. Note that $p + q = 1$. *See also:* **Population genetics (/content/population-genetics/538200)**

Genotype frequencies

Given allele frequencies *p* and *q*, it is possible to predict the genotype frequencies in the next generation under different assumptions regarding mating.

Panmixis

Panmixis refers to random mating with respect to genotype. Under this model, any male of reproductive age is equally likely to mate with any female of reproductive age in terms of the particular gene or trait being analyzed. In other words, a male with genotype *AA* is just as likely to mate with a female of genotype *AA* as with a female of genotype *Aa* or *aa*. Under these conditions, the distribution of genotype frequencies in the next generation is a function of probability, known as Hardy-Weinberg equilibrium.

Hardy-Weinberg equilibrium

Given allele frequencies *p* and *q*, corresponding to alleles *A* and *a*, the expected genotype frequencies under panmixis in the next generation are $AA = p^2$, $Aa = 2pq$, and $aa = q^2$. Under Hardy-Weinberg equilibrium, the genotype and allele frequencies will remain constant generation after generation. Hardy-Weinberg equilibrium provides a baseline model from which to make predictions about nonrandom mating and evolutionary change. *See also:* **Hardy-Weinberg formula (/content/hardy-weinberg-formula/308200)**

Nonrandom mating

If individuals are not mating at random with respect to genotype, then there is no panmixis. There are two basic forms of nonrandom mating: inbreeding and assortative mating. Inbreeding is the mating of individuals that are closely related. In humans, this usually means a couple that are more closely related than third cousins are. Inbreeding changes genotype frequencies by increasing the proportion of homozygotes (*AA, aa*) and decreasing the proportion of heterozygotes (*Aa*). Assortative mating refers to nonrandom mating based on phenotype. Positive assortative mating refers to mating between individuals that are phenotypically similar, such as between two tall people or two people with blond hair. Positive assortative mating also increases the frequency of homozygotes. Negative assortative mating occurs when mating is between individuals that are phenotypically different, and will result in a decrease in homozygotes. Overall, nonrandom mating affects the genotype frequencies but does not change the allele frequencies.

Evolutionary forces

Allele frequencies will remain the same, generation after generation, under both panmixis and nonrandom mating. Evolution, defined here as a change in allele frequency over time, can be caused by four evolutionary forces: mutation, natural selection, gene flow, and genetic drift. Nonrandom mating does not lead to allele frequency changes, although it can affect the rate at which allele frequencies change.

Mutation

Mutation is a random change in the genetic code. It can consist of a small point mutation, where a single nucleotide is changed in the DNA, or larger units of DNA may be changed. Mutation rates are often very difficult to estimate, but most seem to be rather low, with probabilities ranging from about 10^{-4} to 10^{-8} per locus per generation. Mutation introduces new genetic variants into a population and is thus required for any evolutionary change. By itself, however, mutation does not cause rapid evolutionary change; the other evolutionary forces act upon this new variation to increase or decrease the frequency of a mutant allele.

Natural selection

Each genotype can be characterized by its relative probability of survival and reproduction, known as fitness (*w*), and the probability of not surviving or reproducing, known as the selection coefficient (*s*), such that $w + s = 1$. For example, if the probability of an individual with genotype *aa* surviving and reproducing is one-fourth that of individuals with genotypes *AA* or *aa*, then the fitness of *aa* = 0.25 and the selection coefficient of *aa* = 1 − 0.25 = 0.75. Fitness values depend on characteristics of the specific locus and the local environment. Hardy-Weinberg equilibrium assumes that the fitness values of all genotypes are the same. When this is not the case, then natural selection can cause changes in allele frequency through the process of differential survival. There are several different types of selection, each with different outcomes.

Selection against recessive homozygotes. This occurs when the allele *a* is recessive and the recessive homozygous genotype (*aa*) has a lower fitness relative to the other genotypes (*AA, Aa*), which have equal fitness. Under this condition, the frequency of the *a* allele will decline over time to approach a value of zero (although balanced to some extent by continuing mutation from *A* to *a*). The change in allele frequency per generation is a function of the selection coefficient (*s*) for *aa* and the initial frequency (*q*) of the *a* allele. The frequency of *a* will decline each generation by the amount $-spq^2/(1 - sq^2)$. When the allele *a* is lethal, as is the case with many genetic diseases, then $s = 1$ (all individuals with *aa* are selected against), and the decrease in *a* each generation is equal to $-q^2/(1 + q)$. The lethal allele will not be eliminated in a single generation because, even though all individuals with *aa* are selected against, individuals with genotype *Aa* continue to contribute *a* alleles into the population.

Selection for recessive homozygotes. In this case, the fitness of the recessive homozygote is higher than that of the other two genotypes, and the frequency of the dominant allele will decline over time. If the dominant allele is completely lethal, then it will be eliminated in a single generation. If it is not completely lethal, then it will take time for the dominant allele to be reduced to near zero.

Selection against the heterozygote. In this case, the fitness of the heterozygote (*Aa*) is lower than the fitness of the homozygous genotypes (*AA, aa*). Both alleles will be selected against, and the frequency will change in the direction of the initially most common allele. For example, if selection against the heterozygote starts from a condition where $p = 0.7$ and $q = 0.3$, then selection will drive the population to the state of $p = 1.0$ and $q = 0.0$. If, however, *p* is initially less than *q*, then the reverse will occur. For example, if selection against the heterozygote starts with $p = 0.2$ and $q = 0.8$, then selection will ultimately result in $p = 0.0$ and $q = 1.0$. In the unlikely event that the initial allele frequencies are exactly equal ($p = q = 0.5$), the allele frequencies will stay the same until another evolutionary force changes them initially.

Balanced polymorphisms and selection for the heterozygote. All of the above examples involve allele frequencies moving toward 1 or 0. In some cases, the ultimate fate of natural selection is a balance between selective forces producing a set of intermediate allele frequencies that remain in equilibrium. Such balanced polymorphisms result when the heterozygote is selected for because the fitness of the heterozygote is higher than that of the two homozygous genotypes. Selection against *AA* results in the elimination of some *A* alleles, while selection against *aa* results in the elimination of some *a* alleles. At the same time, the selection for the heterozygote results in maintaining some *A* alleles and some *a* alleles. Unlike the previous examples of selection, selection for the heterozygote simultaneously selects for and against both alleles. The result is a balance in allele frequencies determined by the fitness values of the homozygotes. When the heterozygote is the most fit, it has, by definition, a relative fitness of 1. Given a fitness of $1 - s$ for genotype *AA* and a fitness of $1 - t$ for genotype *aa*, selection for the heterozygote will reach an equilibrium where $p = t/(s + t)$ and $q = s/(s + t)$. These values are the allele frequencies where overall survival is maximized. The classic example of a balanced polymorphism is human hemoglobin, a locus with a normal hemoglobin allele (*A*) and the sickle-cell allele (*S*). Individuals with the genotype *SS* have the genetic disease sickle-cell anemia, which is frequently fatal early in life. In many populations, this disease leads to selection against the *SS* genotype and the removal of the *S* allele. In environments with epidemic malaria, heterozygous people (*AS*) are the

most fit because they have greater resistance to malaria due to the presence of one *S* allele, but do not suffer from sickle-cell anemia. As a result, the *S* allele is maintained at a relatively high frequency (usually 5–20%) in malarial environments. *See also:* **Polymorphism (genetics) (/content/polymorphism-genetics/535500)**

Gene flow

Gene flow occurs when populations share genes through the process of mating with someone in another population. An allele absent in one population can be introduced from migrants from another population where the allele is present. Over time, gene flow acts to make populations increasingly genetically similar. Gene flow helps keep populations within a single species; for new species to form, gene flow must be eliminated or severely reduced. Further genetic change must occur for a new species to form, since reduction of gene flow is a necessary, though not sufficient, cause of speciation.

Genetic drift

Genetic drift refers to random fluctuations in allele frequency from one generation to the next because of chance. As an example, consider that if you flip a coin 10 times you expect to get, on average, 5 heads and 5 tails. Because of chance, you might actually get 3 heads and 7 tails, 8 heads and 2 tails, or any other possible combination adding up to 10. Reproduction in a population works in the same manner, so alleles may not be passed along to the next generation in the exact same frequencies as in the parental generation. If p and q are the allele frequencies in the parental generation, the average expectation in the next generation is also p and q, but the variance around these values is equal to $pq/2N$, where N is the number of reproductive adults. When population size is very large, the variance will be small, so there is little chance for genetic drift in a single generation. When N is relatively small (generally less than 200), the variance is large and there is a greater chance for genetic drift. Genetic drift is not directional; an allele frequency can increase, decrease, or remain the same. The probability of change is related to population size. Over time, all populations drift, and will continue to do so until, by chance, an allele becomes extinct (frequency = 0) or fixed (frequency = 1).

Interaction of evolutionary forces

The four evolutionary forces have been discussed one at a time. In reality, all four forces can operate simultaneously and can interact with each other in different ways. For example, a new mutation can increase or decrease in frequency because of natural selection and/or genetic drift, and it can be transmitted to another population via gene flow. In some cases, a harmful allele can actually increase in frequency because of genetic drift. New neutral mutations are frequently eliminated by genetic drift, but may occasionally increase because of random chance and become fixed in a population. Some evolutionary forces, such as drift and selection, will act to make populations genetically different, whereas gene flow can counter such differences. The interaction of evolutionary forces, and their net effect on allele frequency change, is the focus of many methods and studies of population genetics.

John H. Relethford

Biochemical Genetics

Biochemical genetics began with the study of inborn errors of metabolism. These are diseases of the body chemistry in which a small molecule such as a sugar or amino acid accumulates in body fluids because an enzyme responsible for its metabolic breakdown is deficient. This molecular defect is the result of mutation in the gene coding for the enzyme protein. The accumulated molecule, dependent on its nature, is responsible for the causation of a highly specific pattern of disease. There was explosive growth of knowledge in this field with the discovery of many inborn errors of amino acid metabolism following the recognition that amino acids could be separated chromatographically and that they could be identified and quantified because they became purple when reacted with ninhydrin. More recently, the development of gas chromatography–mass

spectrometry led to logarithmic growth in the elucidation of organic acidemias (abnormal acidity of the blood) and disorders of fatty acid metabolism.

The field of biochemical genetics expanded during this period with the recognition that similar heritable defective enzymes ' interfere with the breakdown of very large molecules, such as mucopolysaccharides and the complex lipids that are such ' prominent components of brain substance. The resultant storage disorders present with extreme alterations in morphology ' and bony structure and with neurodegenerative disease. '

Advances in the methodology of molecular biology have permitted the study and detection of disease at the level of the mutation in the DNA. This capability continues to broaden the scope of biochemical genetics. It has been particularly rewarding in the analysis of mutations in the mitochondrial genome. This technology is elucidating a previously unrecognized spectrum of mitochondrial disorders, in which abnormal energy metabolism is often characterized clinically by lactic acidemia.

Inheritance mechanism

The majority of hereditary disorders of metabolism are inherited in an autosomal-recessive fashion. In these families, each parent carries a single mutant gene on one chromosome and a normal gene on the other. Most of these mutations are rare. In populations where consanguinity is common, rare recessive diseases are seen with relative frequency, and affected individuals are homozygous for the same mutation. In populations with more genetic diversity, most affected individuals carry two different mutations in the same gene. Some metabolic diseases are coded for by genes on the X chromosome. Most of these disorders are fully recessive, and so affected individuals are all males, while females carrying the gene are clinically normal. Some genes, such as that for ornithine transcarbamylase (an enzyme of the urea cycle), are expressed as X-linked dominants in which most females are detectable as metabolically abnormal, and some are affected as severely as males. The disorders that result from mutations in the mitochondrial genome are inherited in nonmendelian fashion because mitochondrial DNA is inherited only from the mother. Those that carry a mutation are heteroplasmic; that is, each carries a mixed population of mitochondria, some with the mutation and some without. Expression of clinical disease and its degree of severity are functions of the numbers of abnormal genomes, and they may differ greatly among siblings, because the ovum results from a funneling that may lead to concentration or dilution of the prevalence of the mutation in maternal cells.

Inborn errors of amino acid metabolism

Phenylketonuria (PKU) is a prototypic biochemical genetic disorder. It is an autosomally recessive disorder in which mutations demonstrated in a sizable number of families lead, when present in the genes on both chromosomes, to defective activity of the enzyme that catalyzes the first step in the metabolism of phenylalanine. This results in accumulation of phenylalanine and a recognizable clinical disease whose most prominent feature is severe retardation of mental development. This was the disease with which programs of neonatal screening were developed throughout the world. Determination of the concentration of phenylalanine in a spot of blood obtained from the heel of the baby permits a diagnosis of phenylketonuria early enough to initiate a diet sufficiently restricted in its content of phenylalanine that mental retardation is prevented. This was a wonderful development in public health and preventive medicine. Programs of neonatal screening are expanding throughout the developed world as new methodology permits the detection and treatment of many more inborn errors of metabolism.

The mutations that cause phenylketonuria result in deficient activity of the enzyme phenylalanine hydroxylase, which normally catalyzes the conversion of phenylalanine to tyrosine. Thus, levels of phenylalanine soar and those of tyrosine diminish. When phenylalanine concentrations are high, the amino acid is converted to a number of products including phenylpyruvic acid, a phenylketone. A green color results when ferric chloride is added to solutions containing phenylpyruvic acid, and this once led to the detection of the compound in the urine of two mentally retarded siblings, the index cases of phenylketonuria. Patients with this disease are typically blond and blue-eyed, because the abnormal chemistry interferes with the development

of pigment, in addition to affecting brain development. *See also:* **Phenylketonuria (/content/phenylketonuria/506900)**

X-linked metabolic disorders

The most commonly encountered of the inborn errors of purine metabolism is the Lesch-Nyhan disease. The gene is on the end of the long arm of the X chromosome. Very many mutations have been discovered, a distinct one for virtually every affected family, and most mutations lead to a complete absence of enzyme activity. The enzyme is hypoxanthine-guanine phosphoribosyl transferase (HPRT), which catalyzes the conversion of hypoxanthine and guanine to their respective nucleotides, inosinic acid and guanylic acid. Inosinic acid is also converted to guanylic acid and to adenylic acid, and adenylic acid and guanylic acid are converted to adenosine triphosphate and guanosine triphosphate and their deoxynucleotides—the building blocks of the nucleic acids, RNA and DNA. HPRT is called a salvage enzyme because it is used in the reclaiming of purines resulting from the breakdown of cellular DNA and RNA. The salvage pathways contrast with the de novo pathway of purine nucleotide synthesis in which inosinic acid is made in stepwise fashion from small molecules such as carbon dioxide, ammonia, and glycine. When HPRT is deficient, the phosphoribosylpyrophosphate that serves as the ribose phosphate to form inosinic acid and guanylic acid accumulates and drives the de novo pathway into overactivity; the excess purine synthesized ends up as uric acid.

In this way, patients with HPRT deficiency have high concentrations of uric acid in blood and urine, as do patients with gout. Like patients with gout, these patients may develop acute arthritis, kidney stones, and renal failure. They also have a distinct neurologic disease in which retardation of motor development is associated with spasticity like that of patients with cerebral palsy, and involuntary movements and posturing called chroeoathetosis and dystonia. An even more striking feature of this disease is that the patients develop compulsive-aggressive behavior, the most prominent aspect of which is self-injury through biting. Most patients have loss of tissue about the lips, and most bite their fingers, sometimes with partial amputation. The disease promises to be rewarding in providing chemical explanations of behavior, but so far the linkage between the enzyme defect and the behavioral phenotype has been elusive.

Mitochondrial disease

The diseases that result from mutation in mitochondrial DNA have been recognized as such only since the 1990s. They result from point mutations, deletions, and other rearrangements. It seems likely that there are many more entities yet to be discovered. A majority of these disorders express themselves chemically in elevated concentrations of lactic acid in the blood or cerebrospinal fluid. The importance of the processes of energy metabolism to the central nervous system is underscored by the fact that levels in the cerebrospinal fluid are often higher than those in the blood and may be elevated in cerebrospinal fluid when the blood is normal. Many of the disorders are known as mitochondrial myopathies (diseases of muscles) because skeletal myopathy or cardiomyopathy are characteristic features. The histologic hallmark of these mitochondrial myopathies is the presence of ragged red fibers when the tissue is processed with the Gomori stain. These fibers result from the aggregation of mitochondria, which increase in quantity as compensation for diminished function, and which appear abnormal in size and shape when viewed with electron microscopy. Many of these disorders are known by acronyms, such as MERRF, which stands for mitochondrial encephalomyelopathy with ragged red fibers. An example of disease that results from mutation in mitochondrial DNA is mitochondrial encephalomyelopathy, lactic acidemia, and strokelike episodes (MELAS).

MELAS presents classically with strokelike episodes. Myopathy may have been present earlier, as elucidated by a careful history, or the patient may have been unaware of it until it is discovered upon physical examination. The episode is clinically a typical stroke with hemiplegia (paralysis on one side of the body) or other neurologic manifestation dependent on the area of brain infarcted. They are called strokelike episodes because there is no demonstrated occlusion of a vessel. Consistent with this, the symptoms are sometimes transient, but they also may leave permanent neurologic evidence. Patients are characteristically short in stature. Levels of lactic acid, while elevated, are often not very high; levels in the blood may be

normal, but those in the cerebrospinal fluid are seldom normal. Some patients have convulsions, and ultimately there is evidence of encephalopathy and neurodegeneration. Dementia is a late consequence. Additional features are migraine and noninsulin-dependent diabetes mellitus. The disease is generally recognized by the occurrence of the classic picture in a family member. Once the mutation is identified, it is then usually found in other members of the family, some of whom are asymptomatic, while others may have diabetes or migraine and no other signs of disease. The disease is caused by a point mutation in the mitochondrial gene for the transfer RNA for the amino acid leucine. The mutation generally interferes with protein synthesis in the mitochondria, disrupting the activity of a number of enzymes of the electron transport chain. *See also:* **Mitochondria (/content/mitochondria/428200)**

Significance

Biochemical genetics has been a rich source of insight into fundamental principles in biology and medicine. The concept that biochemical and clinical abnormality could result from a defect in a single enzymatic step in metabolism was first enunciated by A. E. Garrod around the turn of the century. This was the first statement of the one gene–one enzyme hypothesis in which genes function to determine the structure of proteins, whose structure determines function. These concepts, developed long before the elucidation of the DNA as the genetic material, have stood the test of time. The functions of mitochondrial DNA and the variety of disease its mutation causes represent a recent chapter in the growth of knowledge.

William L. Nyhan

Chromosome Mapping

Human chromosome mapping, the localization of human genes to specific chromosomes or regions of chromosomes, has undergone explosive growth since the mid-1970s. Nearly 1500 genes have been assigned to their respective autosomes, well over a hundred genes to the X chromosome, and several functional genes to the Y chromosome. Assuming that the total number of human genes is about 20,000–25,000, this represents about 6–7.5% of the total number. There seems no reason in principle why a complete human gene map should not eventually be achieved. The current state of the map is summarized in **Fig. 10**. For each chromosome, only a few genes, usually those of some clinical interest, have been named. A finding of general interest which has emerged is that the genetic maps of the great apes are almost identical to those of humans, and many groups of genes which are syntenic (on the same chromosome) in humans are syntenic even in the mouse. *See also:* **Proteins, evolution of (/content/proteins-evolution-of/550400)**

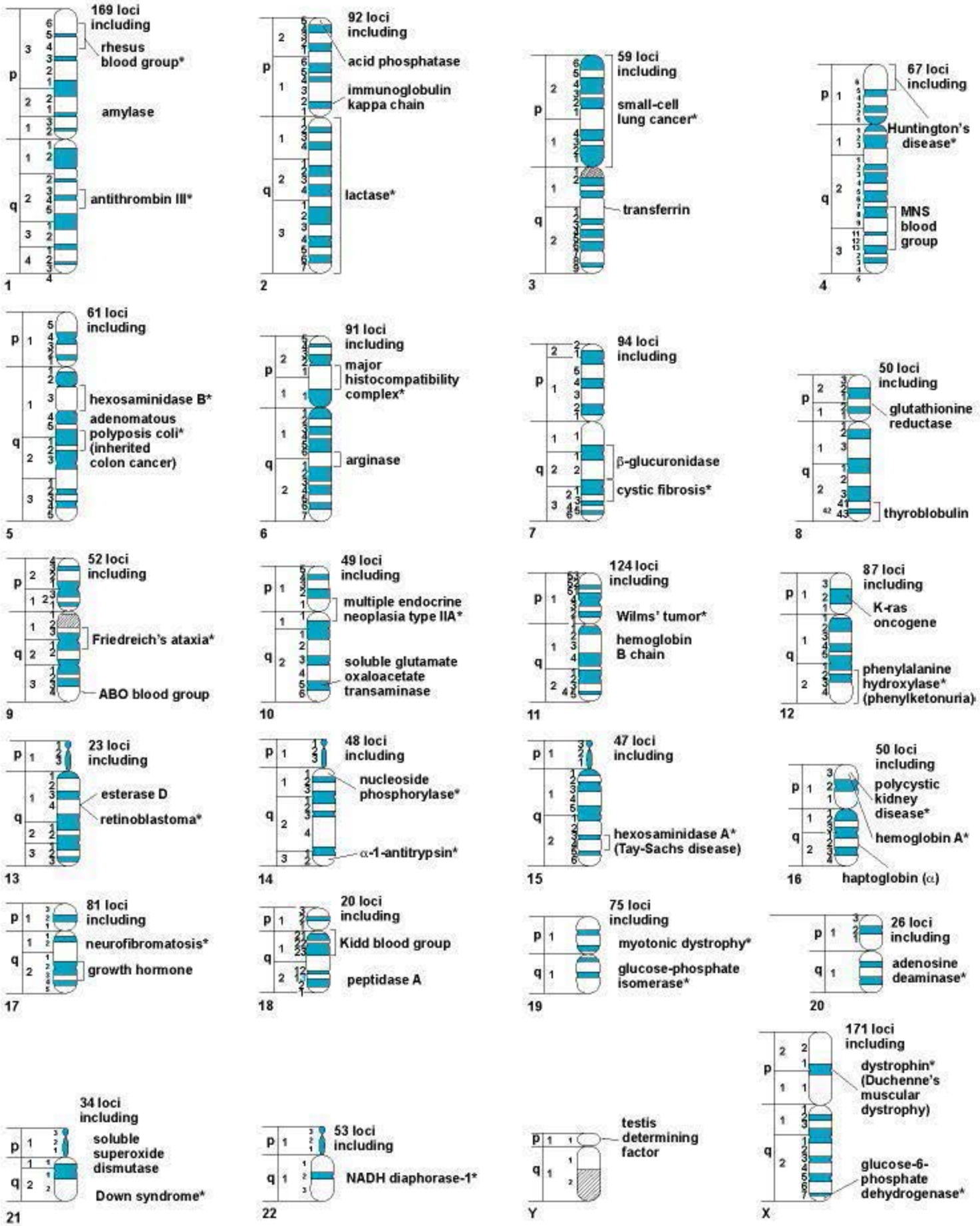

Fig. 10 Diagrammatic representation of the human chromosomes as observed by banding techniques. The number of genes already mapped to each chromosome is indicated beside it, and some examples of particular gene loci are shown. The exact position of these loci is marked where this is known. An asterisk indicates that the locus so marked is the site of one or more mutations associated with inherited disease or is of major clinical importance, such as ABO blood groups.

Most of the information about localization of human genes has been obtained either by somatic cell hybrids or by family studies. Advances in DNA technology have greatly increased the potential of both methods, and have also allowed the development of in situ hybridization, which now accounts for at least one-third of new gene assignments.

Somatic cell hybrids

If human and rodent somatic cells are grown together in tissue culture and treated with certain viruses or chemicals, it is possible to select hybrid uninucleate cells which have resulted from the fusion of human with rodent cells. The most commonly used combinations are human and mouse, or human and Chinese hamster. In both these types of hybrids, on extended subculture, some of the human chromosomes are usually lost preferentially. It is therefore possible to obtain a set of clonal (derived from a single cell) hybrid cell lines, each containing a full set of rodent chromosomes together with a unique subset of human chromosomes.

Any gene whose product is distinguishable in human and rodent and which is expressed in cultured cells can be assigned to a particular chromosome by testing a relatively small number of hybrids for the presence or absence of a given human gene product. This can then be correlated with the presence or absence of some other human gene product or chromosome. The mouse and human gene products, most of which are proteins, can often be separated by electrophoresis. This relies on the fact that the homologous proteins in the two species are not identical but carry slightly different charges owing to differences in their amino acid sequences. **Figure 11** shows an example of electrophoresis of adenylate kinase in human and hamster controls and in six independent human-hamster hybrids. This experiment shows that the two major forms of adenylate kinase on the human cells are coded by different genes on different chromosomes. By chromosome analysis of a number of independently obtained hybrids, one can correlate the synthesis by the hybrid cells of either or both forms of the enzyme with the presence of two particular human chromosomes: actually, one gene (AK_1) is on chromosome 9, the other (AK_2) on chromosome 1. Many gene products distinguished in other ways, such as cell surface antigens, have also been mapped by using such hybrids.

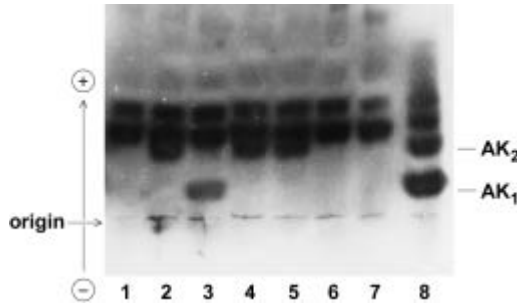

Fig. 11 Photograph of adenylate kinase isozymes separated by starch gel electrophoresis in human control (channel 8), hamster control (channel 7), and six independent hybrid clones, showing independent segregation of human AK_1 (positive in the hybrid in channel 3) and AK_2 (positive in the hybrids in channels 2, 4, and 5).

Originally the only genes that could be mapped in hybrids were those whose products could be detected in these cells. These would not, for example, include genes such as those coding for hemoglobin or insulin, which are only produced in specialized cell types, although presumably the genes themselves are present in all cells. However, it is now possible to isolate many of the genes themselves. The DNA sequence can then be labeled, usually with radioactivity, and used as a probe to search for similar sequences in the hybrids. The presence of the human sequence is then correlated with that of a particular chromosome in exactly the same way as described above for gene products.

This approach is not confined to DNA sequences of known function, and more than 3300 random DNA fragments have also been assigned to particular chromosomes. Although these are of less intrinsic interest, they are proving invaluable as genetic

markers in family studies. If random fragments from some particular chromosome are needed, these can be generated either from a hybrid containing the appropriate single human chromosome, or by direct sorting of the human chromosomes using a fluorescent activated cell sorter. In both cases, it is necessary to check with a hybrid panel that the fragments obtained are indeed from the right chromosome.

The use of human parental cells containing rearranged chromosomes such as translocations or deletions, or the observation of spontaneous or induced chromosome rearrangements arising in hybrid cells, also allows the assignment of genes to particular regions of chromosomes. *See also:* **Somatic cell genetics (/content/somatic-cell-genetics/636300)**

In situ hybridization

The chromosomes of normal human lymphocytes can be readily visualized during the metaphase stage of cell division. In the case of genes involving localized repeated sequences, a DNA probe radioactively labeled to high activity can be used directly for hybridization to specific chromosomes in metaphase spreads, which are then autoradiographed. It has also become possible to label probes with sufficient specific activity to detect sequences present only as single copies (such as most of those coding for proteins), and thus, in one experiment, to localize the gene to a particular region of a particular chromosome. Developments in nonradioactive labeling of DNA with fluorescent dyes and the use of a confocal laser scanning microscope have greatly improved the precision of gene localization by this technique.

Family studies

Unlike somatic cell hybrids, gene mapping by family studies depends on finding genetic differences between people and observing the way these differences are inherited. A special case is the X chromosome because most genes on the X chromosome are inherited in a distinctive way—they never pass from father to son. It is therefore relatively easy to establish that a particular disease or distinctive trait is X-linked; for a long time, many such diseases have been recognized. A well-known example is the hemophilia gene that was carried by Queen Victoria and caused the disease in many of the royal families of Europe. The assignment of genes to particular autosomes by family studies presents a more formidable problem and it was only in 1969 that the first confident assignment, that of the Duffy blood group to chromosome 1, was made. *See also:* **Hemophilia (/content/hemophilia/313900)**

For many years, workers have analyzed the patterns of inheritance of two or more gene loci in the same family to establish whether the genes are linked, that is, carried on the same chromosome pair close enough together to show nonindependent segregation. If the genes are linked, a measure of the intergene distance can be derived by the frequency with which they are separated by crossing over. A linkage group found in this way was that of the ABO blood groups with one of the forms of adenylate kinase. The assignment of the adenylate kinase gene to chromosome 9 by somatic cell hybrids thus allowed the indirect assignment of the gene determining the ABO blood group to the same chromosome.

Direct assignment of genes to chromosomes by using family studies can be accomplished by studying concurrently the inheritance of genetically determined traits and of chromosome polymorphisms or rearrangements. Individuals carrying unbalanced chromosome abnormalities involving deletions or duplications of specific regions of chromosomes can also be useful for mapping by gene dosage. However, quantitative differences indicating one, two, or three copies of a gene are difficult to distinguish from secondary effects, and from unusual normal alleles which code for enzymes of different activity. This method has been most successful in regional localization when the gene is already known to lie on a particular abnormal chromosome.

The efficiency of family studies in assigning genes to chromosomes depends very much on the proportion of the human genome which is within measurable genetic distance of “good” genetic markers. In this context, a good genetic marker might

be defined as one determined by a single gene locus, easily typed and showing a large amount of normal variation, so that a high proportion of individuals are heterozygous. Before the advent of DNA technology, the genetic markers available have been limited and in practice have been confined to those expressed in readily obtained tissues such as blood. Thus, the chance of assigning any new gene to its chromosome even from study of a large number of families was small, one estimate being about 8%. Even in this early work, however, family studies had an essential role in establishing the fine details of gene arrangement, and in the construction of a true genetic map where the known amount of recombination between two genes gives a definite value or map distance. In most chromosome regions, it has been found that recombination, and hence map distance, between genes is almost twice as great in females as in males. *See also:* **Recombination (genetics) (/content/recombination-genetics/575500)**

In the 1980s, there was a dramatic increase in the number of genetic markers available, allowing the construction of at least partial genetic maps of all the human chromosomes. As a result, most of the human genome is now within reach of known genetic markers. This has come about through the use of DNA probes in conjunction with bacterial enzymes known as restriction endonucleases. These enzymes recognize certain specified short sequences of DNA, usually 4 or 5 bases long, and then cut the DNA at all the places where this sequence occurs. Many different enzymes can be extracted from different bacteria, each recognizing a different sequence. Human DNA from different individuals is treated with an appropriate enzyme, which breaks it into many fragments. These are then separated by size on an agarose gel and the fragments containing sequences that are complementary to the DNA of the radioactively labeled probe can be identified (**Fig. 12**). The number and size of the fragments depend on the distribution of sites for that particular restriction enzyme in the region of the genome to which the probe hybridizes, and the sizes commonly detected are between 500 bases and 15,000 bases (15 kilobases, or kb). In this way, variation in the base sequences can be detected not only in the region of the gene which codes for the protein but also in the large regions of noncoding DNA which form intervening and flanking sequences. Variants that are detected in this way (restriction fragment length polymorphism) behave as single codominant alleles, and it appears that for almost any DNA probe (whether a random fragment or a known gene) such variation can be found if a sufficient number of restriction enzymes are tried. *See also:* **Human genome (/content/human-genome/757575)**; **Restriction enzyme (/content/restriction-enzyme/584150)**

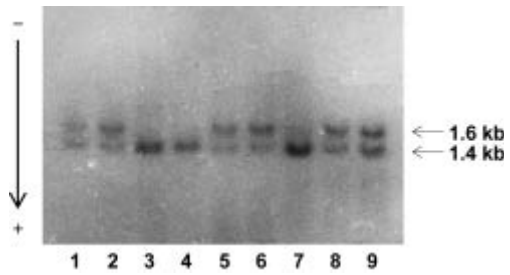

Fig. 12 Autoradiograph showing polymorphism defined by a random DNA fragment from human chromosome 7 and the restriction enzyme *HinfI*. Three individuals (tracks 3, 4, 7) are homozygous for the common allele producing a fragment 1400 bases long (1.4 kb). The remaining samples are from heterozygotes in which one chromosome produces the common pattern and the other a longer fragment (1.6 kb). (*Photograph courtesy of Dr. Ben Carritt*)

Another very useful type of genetic marker usually found in noncoding regions of DNA is the VNTR (variable number of tandem repeats). In this case, sequences recognized by one probe occur at the same place but in variable numbers on the chromosomes of different individuals. If the DNA is digested with any enzyme that does not cut within the repeat, the size of the fragment recognized by probing after agarose gel electrophoresis will vary between individuals and between the two homologous chromosomes of the same individual chromosomes. At some of these hypervariable loci, 99% of individuals are heterozygous, so every family is informative for linkage analysis with that marker.

Family studies for mapping disease genes

The frequency of the different polymorphisms described above means that virtually the whole genome is within reach of good genetic markers. Family studies on diseases with clear mendelian inheritance, but in which the basic biochemical defect is unknown, have become a very practical undertaking. Such studies have allowed the localization of the genes that cause many autosomal-dominant disorders such as Huntington's disease and adult-onset polycystic kidney disease. It is still much more complex to localize a gene that causes a disease inherited in an autosomal-recessive manner if the heterozygotes cannot be identified, but in the case of cystic fibrosis, a common disease in populations of European origin, this has been accomplished (Fig. 10).

There are several ways in which localizing such genes may help in understanding a disease, in some cases by demonstrating that more than one gene is involved and in others by providing a starting point for reverse genetics. Once the approximate position of a disease gene has been found, it is possible to generate more genetic markers in the same region of the chromosome and to “walk” or “jump” along the chromosome to find the actual disease gene and the mutations that cause the disease, allowing more certain diagnosis and better genetic counseling. A study of the protein coded for by the gene in question should then offer a deeper understanding of the disease and how it may be treated. Understanding a disease by reverse genetics was first successful for Duchenne muscular dystrophy and chronic granulomatous disease, both of which are inherited on the X chromosome. The most dramatic success of this approach has been the identification of the gene that is defective in cystic fibrosis.

A few diseases are inherited only through the mother, but are equally severe in males and females. (This is in contrast to X-linked inheritance, in which genes can pass from father to daughter but not from father to son.) Some of these diseases, including Leber's optic atrophy, are due to mutations of the mitochondria. The mitochondrion is a small circular piece of DNA, only about 16 kb long, that is found in many copies in each cell. The whole DNA sequence of human mitochondrial DNA is known and all the genes have been identified, and so in some ways it can be regarded as a very small chromosome.

Ongoing research

Several techniques should greatly speed up progress in human gene mapping. One is the polymerase chain reaction, which allows the selective amplification of a small region of DNA of interest from a single drop of blood—or even a single cell—without complex purification. One or more particular genes can then be examined quickly and easily in very large numbers of people or in many individual sperms from a single informative individual. The latter technique is the genetic equivalent of sampling hundreds of children from the same father. Another approach, which can be used in conjunction with polymerase chain reaction, is the direct preparation of DNA from a small piece, such as a single band, of a chromosome dissected from a metaphase spread. Therefore, the chromosomal origin of clones that are derived from such pieces is precisely known. *See also:* **Polymerase chain reaction (PCR) (/content/polymerase-chain-reaction-pcr/900192)**

Another exciting field of study is that of oncogenes. These are sequences similar to those of ceres, and in some circumstances these DNA sequences can cause changes in mouse cells similar to those seen in cancer cells. Although their relationship to human cancer is still not clear, the coincidence of the map positions of some of these oncogenes and the breakpoints of characteristic chromosome rearrangements seen in some tumors is under investigation. *See also:* **Cancer (medicine) (/content/cancer-medicine/105800)**

It is important to realize the different scale of mapping that can be achieved by the different approaches. Although the banding patterns shown in Fig. 1 are those usually obtained during the metaphase stage of cell division, higher resolution giving at least 800 bands for the whole genome can be obtained by examining the chromosomes earlier, during prophase, when they are much longer. This has led to a more precise definition of breakpoints and of map positions, although the smallest visible band still represents approximately 1 million base pairs, and two genes that appear to be quite close together

in family studies may be several million base pairs apart. Therefore, there is a considerable difference in scale between molecular mapping (for example, of the various defects in beta-hemoglobin) and the physical and genetic chromosome mapping described here. This gap is bridged by a technique known as pulsed-field gel electrophoresis in which so-called rare cutting restriction enzymes are used to generate very large fragments of DNA (up to 1 million base pairs long), which are separated in gels by applying alternating cycles of electric fields in different directions. Some of these fragments are found to carry more than one gene, and so it gives an exact measure of the distance between them. Eventually, overlapping fragments of DNA spanning all the human chromosomes should be identified and most of it should be sequenced. The human genetic map will then be essentially complete and should lead to better diagnosis, genetic counseling, and, eventually, treatment of inherited diseases. The human genetic map should also allow a greater understanding of the genetic changes in cells that underlie other diseases, especially cancers. *See also:* **Genetics (/content/genetics/285300)**

Sue Povey

Bibliography

K. E. Davies (ed.), *Genome Analysis: A Practical Approach*, 1988

D. S. Falconer and T. F. C. Mackay, *Introduction to Quantitative Genetics*, 4th ed., 1996

F. C. Fraser and J. J. Nora, *Genetics of Man*, 2d ed., 1986

H. Harris, *The Principles of Human Biochemical Genetics*, 3d ed., 1981

D. L. Hartl and A. G. Clark, *Principles of Population Genetics*, 4th ed., 2007

B. S. Kerem et al., Identification of the cystic fibrosis gene: Genetic analysis, *Science*, 245:1073–1080, 1989 DOI: 10.1126/science.2570460 (http://dx.doi.org/10.1126/science.2570460)

M. Levitan, *Textbook of Human Genetics*, 3d ed., 1988

V. A. McKusick, *Mendelian Inheritance in Man*, 12th ed., 1998

J. J. Nora, *Medical Genetics: Principles and Practices*, 4th ed., 1993

W. L. Nyhan and P. T. Ozand, *Atlas of Metabolic Diseases*, 1998

R. R. Race and R. Sanger, *Blood Groups in Man*, 6th ed., 1975

J. H. Relethford, *The Human Species: An Introduction to Biological Anthropology*, 7th ed., 2008

J. R. Riordan et al., Identification of the cystic fibrosis gene: Cloning and characterization of complementary DNA, *Science*, 245:1066–1073, 1989 DOI: 10.1126/science.2475911 (http://dx.doi.org/10.1126/science.2475911)

C. R. Scriver et al., *The Metabolic and Molecular Basis of Inherited Disease*, 8th ed., 2000

T. B. Shows et al., Mapping the human genome, cloned genes, DNA polymorphisms, and inherited disease, *Adv. Hum. Genet.*, 12:341–452, 1982 DOI: 10.1007/978-1-4615-8315-8_5 (http://dx.doi.org/10.1007/978-1-4615-8315-8_5)

M. Speicher, S. E. Antonarakis, and A. G. Motulsky (eds.), *Vogel and Motulsky's Human Genetics: Problems and Approaches*, 4th ed., 2010

H. E. Sutton, *An Introduction to Human Genetics*, 4th ed., 1988

F. Vogel, Clinical consequences of heterozygosity for autosomal-recessive diseases, *Clin. Genet.*, 25:381–415, 1984 DOI: 10.1111/j.1399-0004.1984.tb02009.x (http://dx.doi.org/10.1111/j.1399-0004.1984.tb02009.x)

Additional Readings

R. Hodge, *Human Genetics: Race, Population, and Disease*, Facts on File, New York, 2010

J. F. Hughes et al., Chimpanzee and human Y chromosomes are remarkably divergent in structure and gene content, *Nature*, 463(728):536–539, 2010 DOI: 10.1038/nature08700 (http://dx.doi.org/10.1038/nature08700)

B. R. Korf and M. B. Irons, *Human Genetics and Genomics*, 4th ed., John Wiley & Sons, Chichester, West Sussex, UK, 2013

J. C. Roach et al., Analysis of genetic inheritance in a family quartet by whole-genome sequencing, *Science*, ' 328(5978):636–639, 2010 DOI: 10.1126/science.1186802 (http://dx.doi.org/10.1126/science.1186802) '

P. N. Robinson and S. Mundlos, The human phenotype ontology, *Clin. Genet.*, 77(6):525–534, 2010 DOI: ' 10.1111/j.1399-0004.2010.01436.x (http://dx.doi.org/10.1111/j.1399-0004.2010.01436.x) '

L. E. Rosenberg and D. D. Rosenberg, *Human Genes and Genomes: Science, Health, Society*, Academic Press, Waltham, MA, 2012