

## Appendix G

# Glossary

---

**Alleles:** Alternative forms of a genetic locus; alleles are inherited separately from each parent (e.g., at a locus for eye color there might be alleles resulting in blue or brown eyes).

**Amino acid:** Any of a group of 20 molecules that combine to form proteins in living things. The sequence of amino acids in a protein is determined by the genetic code.

**Autoradiography:** A technique that uses X-ray film to visualize radioactively labeled molecules or fragments of molecules; used in analyzing length and number of DNA fragments after they are separated by gel electrophoresis.

**Autosome:** A chromosome not involved in sex determination. The diploid human genome consists of 46 chromosomes, 22 pairs of autosomes and 1 pair of sex chromosomes.

**Base pair:** Two nucleotides (adenosine and thymidine or guanosine and cytidine) held together by weak bonds. Two strands of DNA are held together in the shape of a double helix by the bonds between base pairs.

**Centimorgan:** A unit of measure of recombination frequency. One centimorgan is equal to a 1 percent chance that a genetic locus will be separated from a marker due to recombination in a single generation. In human beings, 1 centimorgan is equivalent, on average, to 1 million base pairs.

**Cloning:** The process of asexually producing a group of cells (clones), all genetically identical to the original ancestor. In recombinant DNA technology, the process of using a variety of DNA manipulation procedures to produce multiple copies of a single gene or segment of DNA.

**Complementary DNA, cDNA:** DNA that is synthesized from a messenger RNA template; the single-strand form is often used as a probe in physical mapping.

**Contigs:** Groups of clones representing overlapping, or contiguous, regions of a genome.

**Crossing over:** The breaking during meiosis of one maternal and one paternal chromosome, the exchanging of corresponding sections of DNA, and the rejoining of the chromosomes.

**C-value paradox:** The lack of correlation between the amount of DNA in a haploid genome and the biological complexity of the organism. (C-value refers to haploid genome size.)

**Determinism:** The theory that for every action taken there are causal mechanisms such that no other action was possible.

**Diploid:** A full set of genetic material (two paired sets of chromosomes), one from each parental set. All cells except sperm and egg cells have a diploid set of chromosomes. The diploid human genome has 46 chromosomes. Compare *haploid*.

**DNA, deoxyribonucleic acid:** The molecule that encodes genetic information. DNA is a double-stranded molecule held together by weak bonds between base pairs of nucleotides. There are four nucleotides in DNA: adenosine (A), guanosine (G), cytidine (C), and thymidine (T). In nature, base pairs form only between A and T and between G and C, thus the sequence of each single strand can be deduced from that of its partner.

**DNA probes:** Segments of single-strand DNA that are labeled with a radioactive or other chemical marker and used to identify complementary sequences of DNA by hybridizing with them. See *hybridization*.

**DNA sequence:** The relative order of base pairs, whether in a stretch of DNA, a gene, a chromosome, or an entire genome.

**Domain:** A discrete portion of a protein with its own function. The combination of domains in a single protein determines its unique overall function.

**Double helix:** The shape in which two linear strands of DNA are bonded together.

**Electrophoresis:** A method of separating large molecules (such as DNA fragments or proteins) from a mixture of similar molecules. An electric current is passed through a medium containing the mixture, and each kind of molecule travels through the medium at a different rate, depending on its electrical charge and size. Separation is based on these differences.

**Enzyme:** A protein that acts as a catalyst, speeding the rate at which a biochemical reaction proceeds but not altering its direction or nature.

**Eukaryote:** Cell or organism with membrane-bound, structurally discrete nucleus and other well-developed subcellular compartments. Eukaryotes include all organisms except viruses, bacteria, and blue-green algae. Compare *prokaryote*.

**Eugenics:** Attempts to improve hereditary qualities through selective breeding. See *positive eugenics*, *negative eugenics*, *eugenics of normalcy*.

**Eugenics of normalcy:** Policies and programs intended to ensure that each individual has at least a minimum number of normal genes.

**Exons:** The protein-coding DNA sequences of a gene. Compare *introns*.

**Gamete** Mature male or female reproductive cell with a haploid set of chromosomes (23); that is, a sperm or ovum.

**Gene** The fundamental physical and functional unit of heredity. A gene is an ordered sequence of nucleotides located in a particular position on a particular chromosome. See *gene expression*.

**Gene expression** The process by which a gene's blueprint is converted into the structures present and operating in the cell. Expressed genes include those that are transcribed into mRNA and then translated into protein and those that are transcribed into RNA but not translated into protein (e.g., transfer and ribosomal RNAs).

**Gene families:** Groups of closely related genes that make similar products.

**Gene product:** The biochemical material, either RNA or protein, made by a gene. The amount of gene product is used to measure how active a gene is; abnormal amounts can be correlated with disease-causing genes.

**Genetic code:** The sequence of nucleotides, coded in triplets along the mRNA, that determines the sequence of amino acids in protein synthesis. The DNA sequence of a gene can be used to predict the mRNA sequence, and the genetic code can in turn be used to predict the amino acid sequence.

**Genetic engineering technologies:** See *recombinant DNA technologies*.

**Genetic linkage map:** A map of the relative positions of genetic loci on a chromosome, determined on the basis of how often the loci are inherited together. Distance is measured in centimorgans.

**Genetics:** The study of the patterns of inheritance of specific traits.

**Genomtx** All the genetic material in the chromosomes of a particular organism; its size is generally given as its total number of base pairs.

**Genome projects:** Research and technology development efforts aimed at mapping and sequencing some or all of the genome of human beings and other organisms.

**Genomic library:** A collection of clones made from a set of overlapping DNA fragments representing the entire genome of an organism. Compare *library*.

**Haploid:** A single set of chromosomes (half the full set of genetic material), present in the egg and sperm cells of animals and in the pollen cells of plants. Human beings have 23 chromosomes in their reproductive cells. Compare *diploid*.

**Homeo box:** A short stretch of nucleotides whose sequence is virtually identical in all the genes that contain it. It has been found in many organisms, from fruit flies to human beings. It appears to determine

when particular groups of genes are expressed in the development of the fruit fly.

**Human gene therapy:** Insertion of normal DNA directly into cells to correct a genetic defect.

**Human Genome Initiative:** Collective name for several projects begun in 1986 by DOE to 1) create an ordered set of DNA segments from known chromosomal locations, 2) develop new computational methods for analyzing genetic map and DNA sequence data, and 3) develop new techniques and instruments for detecting and analyzing DNA.

**Hybridization** The process of joining two complementary strands of DNA, or of DNA and RNA, together to form a double-stranded molecule.

**Informatic:** The study of the application of computer and statistical techniques to the management of information. In genome projects, informatics includes the development of methods to search databases quickly, to analyze DNA sequence information, and to predict protein sequence and structure from DNA sequence data.

**International technology transfer:** Movement of inventions and technical know-how across national borders.

**Introns:** The DNA sequences interrupting the protein-coding sequences of a gene that are transcribed into mRNA but are cut out of the message before it is translated into protein. Compare *exons*.

**Karyotypcx** A photomicrograph of an individual's chromosomes arranged in a standard format showing the number, size, and shape of each chromosome; used in low-resolution physical mapping to correlate gross chromosomal abnormalities with the characteristics of specific diseases.

**Library:** A collection of clones in no obvious order whose relationship can be established by physical mapping. Compare *genomic library*.

**Linkage** The proximity of two or more markers (e.g., genes, RFLP markers) on a chromosome; the closer together the markers are, the lower the probability that they will be separated during meiosis and hence the greater the probability that they will be inherited together.

**Locus:** The position on a chromosome of a gene or other chromosome marker; also, the DNA at that position. Some restrict use of locus to regions of DNA that are expressed. See *gene expression*.

**Marker** An identifiable physical location on a chromosome (e.g., restriction enzyme cutting site, gene, RFLP marker) whose inheritance can be monitored. Markers can be expressed regions of DNA (genes) or some segment of DNA with no known coding function but whose pattern of inheritance can be determined.

**Meiosis:** The process of two consecutive cell divisions in the diploid progenitors of sex cells. Meiosis results in four rather than two daughter cells, each with a haploid set of chromosomes.

**Messenger RNA, mRNA:** A class of RNA produced by transcribing the DNA sequence of a gene. The mRNA molecule carries messages specific to each of the 20 amino acids. Its role in protein synthesis is to transmit instructions from DNA sequences (in the nucleus of the cell) to the ribosomes (in the cytoplasm of the cell).

**Multifactorial or multigenic disorders** See *polygenic disorders*.

**Mutation:** Any change in DNA sequence that results in a new characteristic that can be inherited. Compare *polymorphism*.

**Negative eugenics:** Policies and programs intended to reduce the occurrence of genetically determined disease.

**Nucleotide:** A subunit of DNA or RNA consisting of a nitrogenous base (adenine, guanine, thymine, or cytosine in DNA; adenine, guanine, uracil, or cytosine in RNA), a phosphate molecule, and a sugar molecule (deoxyribose in DNA and ribose in RNA). Thousands of nucleotides are linked to form the DNA or RNA molecule. See *DNA, base pair, RNA*.

**Oncogene** A gene, one or more forms of which is associated with cancer. Many oncogenes are involved, directly or indirectly, in controlling the rate of cell growth.

**Physical map:** A map of the locations of identifiable landmarks on DNA (e.g., restriction enzyme cutting sites, genes, RFLP markers), regardless of inheritance. Distance is measured in base pairs. For the human genome, the lowest-resolution physical map is the banding patterns of the 24 different chromosomes; the highest-resolution map would be the complete nucleotide sequence of the chromosomes.

**Polygenic disorder** Genetic disorders resulting from the combined action of alleles of more than one gene (e.g., heart disease, diabetes, and some cancers). Although such disorders are inherited, they depend on the simultaneous presence of several alleles, thus the hereditary patterns are usually more complex than those of single-gene disorders. Compare *single-gene disorders*.

**polymorphism:** Difference in DNA sequence among individuals. Genetic variations occurring in more than 1 percent of a population would be considered useful polymorphisms for genetic linkage analysis. Compare *mutation*.

**Positive eugenics:** The achievement of systematic or planned genetic changes to improve individuals or their offspring.

**Prokaryote:** Cell or organism lacking membrane-bound, structurally discrete nucleus and subcellular compartments. Bacteria are examples. Compare *eukaryote*.

**Protein:** A large molecule composed of chains of smaller molecules (amino acids) in a specific sequence; the sequence is determined by the sequence of nucleotides in the gene coding for the protein. Proteins are required for the structure, function, and regulation of the body's cells, tissues, and organs, and each protein has a unique function. Examples are hormones, enzymes, and antibodies.

**Recombinant DNA technologies:** Procedures used to join together DNA segments in a cell-free system (an environment outside of a cell or organism). A recombinant DNA molecule can enter a cell and replicate there, either autonomously or after it has become integrated into a cellular chromosome.

**Replication:** The synthesis of new DNA strands from existing DNA. In human beings and other eukaryotes, replication occurs in the nucleus of the cell.

**Resolution:** Degree of molecular detail on a physical map of DNA, ranging from low to high.

**Restriction enzyme, endonuclease:** A protein that recognizes specific, short nucleotide sequences and cuts DNA at those sites. There are over 400 such enzymes in bacteria that recognize over 100 different DNA sequences. See *restriction enzyme cutting site*.

**Restriction enzyme cutting site:** A specific nucleotide sequence of DNA at which a restriction enzyme cuts the DNA. Some sites occur frequently in DNA (e.g., every several hundred base pairs), others much less frequently (e.g., every 10,000 base pairs).

**RFLP, restriction fragment length polymorphism:** Variation in DNA fragment sizes cut by restriction enzymes; polymorphic sequences that are responsible for RFLPs are used as markers on genetic linkage maps.

**Ribosomal RNA, rRNA:** A class of RNA found in the ribosomes of cells.

**RNA, ribonucleic acid:** A chemical found in the nucleus and cytoplasm of cells; it plays an important role in protein synthesis and other chemical activities of the cell. The structure of RNA is similar to that of DNA. There are several classes of RNA molecules, including messenger RNA, transfer RNA, ribosomal RNA, and other small RNAs, each serving a different purpose.

**Sex chromosomes:** The X and Y chromosomes in human beings that determine the sex of an individual. Females have two X chromosomes in diploid cells; males have an X and a Y chromosome.

**Single-gene disorders:** Hereditary disorders caused

by a single gene (e.g., Duchenne muscular dystrophy, retinoblastoma, sickle cell disease). Compare *polygenic disorders*.

**Somatic cells:** Any cells in the body except reproductive cells and their precursors.

**Technology transfer** The process of converting scientific knowledge into useful products.

**Transcription:** The synthesis of mRNA from a sequence of DNA (a gene); the first step in gene expression. Compare *translation*.

**Transfer RNA tRNA:** A class of RNA having structures with triplet nucleotide sequences that are complementary to the triplet nucleotide coding se-

quences of mRNA. The role of tRNAs in protein synthesis is to bond with amino acids and transfer them to the ribosomes, where proteins are synthesized according to the instructions carried by mRNA.

**Translation:** The process in which the genetic code carried by mRNA directs the synthesis of proteins from amino acids. Compare *transcription*.

**Vector:** DNA molecule originating from a virus, a bacterium, or the cell of a higher organism used to carry additional DNA base pairs; vectors introduce foreign DNA into host cells, where it can be reproduced in large quantities. Examples are plasmids, cosmids, and yeast artificial chromosomes.