

## Genetics Glossary

Susan A. Ehrlich<sup>1</sup>

**3'**. The end of a nucleoside defined by the number 3 carbon atom of the ribose or deoxyribose sugar as opposed to the **5'** end defined by the number 5 carbon atom. The nucleoside - phosphate - nucleoside - phosphate - nucleoside - phosphate ... chain of DNA or RNA thus has polarity; one end is the 3' end and one end is the 5' end. DNA replication is from 5' to 3' on the template.

**5' cap**. A chemically modified guanine nucleotide added to the 5' end of a growing mRNA molecule.

**Å**. Angstrom unit. Named after nineteenth-century Swedish physicist Anders Ångstrom, 1Å equals 1 ten-billionth meter or there are 254 million Ås per inch.

**Adenine (A)**. A purine base; one of the four molecules containing nitrogen present in DNA and RNA. Designated by the letter A, it binds to thymine (T) or uracil (U).

**Adenosine Triphosphate (ATP)**. A nucleotide triphosphate that upon hydrolysis results in energy available for such processes as muscle contraction and synthesis of macromolecules, including proteins and carbohydrates.

**Aerobic**. Requiring oxygen for growth.

**Affinity**. The binding power of an antibody with an antigen.

**Agarose**. The neutral gelling fraction of agar (a polysaccharide extracted from certain seaweed) commonly used in gel electrophoresis.

**Algorithm**. A step-by-step process for solving a problem.

**Allele**. From “allelomorph,” one of a series of alternative forms of a gene (or VNTR) at a specific locus in a genome.

**Allele frequency**. The proportion of a particular allele among the chromosomes carried by individuals in a population.

---

<sup>1</sup>Copyright © 2005 by author, who wishes to thank Dr. Raymond F. Gesteland for his editorial assistance.

**Allele-specific oligonucleotide.** A short DNA sequence, usually 18-20 nucleotides, that can hybridize with either disease-causing or normal DNA sequences but not both.

**Allogenic.** Of the same species but with a different genotype.

**Alu.** A major group of dispersed repetitive DNA sequences; a family of repeat DNA sequences, cleaved by the restriction enzyme Alu I, dispersed throughout the genomes of many animal species. A human being has approximately 500,000 copies at 300 base pairs each.

**Amino acid.** The major building blocks of polypeptides; any of a class of 20 molecules combined to form proteins in living things; the building blocks of proteins coded by triplets of bases in the DNA blueprint: alanine (Ala), arginine (Arg), asparagine (Asn), aspartic acid (Asp), cysteine (Cys), glutamine (Gln), glutamic acid (Gla), glycine (Gly), histidine (His), isoleucine (Ile), leucine (Leu), lysine (Lys), methionine (Met), phenylalanine (Phe), proline (Pro), serine (Ser), threonine (Thr), tryptophan (Trp), tyrosine (Tyr), valine (Val). For example, the mRNA transcript code AUG is for the amino acid methionine, whereas CGU, CGC, CGA, CGG, AGA and AGG code for arginine, UUG codes for leucine and GGG codes for glycine. In certain situations, UGA codes for selenocysteine or the 21<sup>st</sup> amino acid.

The sequence of amino acids in a protein and hence protein function are determined by the genetic code.

**Amplification.** An increase in the number of copies of a specific DNA sequence, usually by PCR.

**Amplified fragment length polymorphism (AMP-FLP).** PCR-amplified restriction fragment lengths consisting of a variable number of tandem repeats.

**Anaerobic.** Growing in the absence of oxygen.

**Antibody.** A protein produced for body defense in response to an antigen.

**Anticodon.** A three-nucleotide sequence in a tRNA molecule that undergoes complementary base pairing with an mRNA codon.

**Antigen.** From “antibody generator,” a molecule, usually a protein, capable of stimulating an antibody response for body defense.

**Antisense.** The strand of DNA that would produce a mirror image (antisense) messenger RNA that is opposite in sequence to one directing protein synthesis. Antisense technology is used to selectively turn off production of certain proteins.

**Antiserum.** Blood serum containing specific antibodies against an antigen used to confer immunity to a disease.

**Autorad.** From “autoradiogram,” the resultant x-ray film after having been exposed to a radioactive source. A DNA probe tagged with a radioactive isotope such as radioactive phosphorus ( $^{32}\text{P}$ ) will expose an x-ray film where the probe hybridizes to complementary sequences on the blot in contact with the film.

**Autosome.** A chromosome not involved in sex determination. The diploid human genome consists of 46 chromosomes, 22 pair of autosomes and one pair of sex chromosomes (X and Y).

**Bacteriophage** or “phage”. A virus that reproduces in bacteria.

**Bacterium.** Any of a large group of microscopic organisms with a simple cell structure, some of which manufacture their own food, some of which live as parasites on other organisms and some of which live on decaying matter.

**Bands.** Visibly darkened areas on autorads that represent the location of DNA fragments on a gel; alternating dark and light areas visible on chromosomes after certain types of stains are used.

**Band shifting.** The phenomenon in which DNA fragments in one lane of an electrophoresis gel migrate more rapidly than fragments in a second lane.

**Base.** One of four DNA nucleotides, adenine (A), cytosine (C), guanine (G) or thymine (T). In RNA, it is uracil (U) instead of thymine.

**Base analog.** A molecule that can mimic the chemical behavior of one of the DNA bases. Base analogs are a type of mutagen, some of which are used in chemotherapy.

**Base pair (bp).** Two complementary nucleotides joined by hydrogen bonds binding DNA or RNA complementary strands; base-pairing occurs between A and T (DNA) or U (RNA) and between C and G.

**Base-pair substitution.** The replacement of one base pair by another, a type of mutation.

**Base sequence.** The order of bases in a DNA or RNA molecule.

**Bin.** In VNTR profiling, a range of base pairs (DNA fragment lengths). When a database is divided into **fixed bins**, the proportion of bands within each bin is determined and the relevant proportions are used in estimating the profile frequency. In a **floating bin** method of estimating a profile frequency, the bin is centered on the base-pair length of the allele in question and the width of the bin can be defined by the laboratory's matching rule, e.g.,  $\pm 5\%$  of band size. **Binning** is grouping VNTR alleles into sets of similar sizes because the alleles' lengths are too similar to differentiate.

**Biochemical marker** or **biomarker.** A substance whose detection indicates a biochemical (chemical, molecular or physical) activity or change in the body, its tissues or cells; it may be monitored as a quantifiable indicator in the assessment of a disorder or condition.

**Blastocyst.** The 4-5-day-old ball of undifferentiated cells from which a prospective embryo develops.

**Capillary electrophoresis.** A method for separating DNA fragments according to their lengths. A long, narrow tube is filled with an entangled polymer or comparable sieving medium and an electric field is applied to pull DNA fragments placed at one end of the tube through the medium. The procedure is faster and uses smaller samples than slab gel electrophoresis, and it can be automated.

**Catalyst.** An agent (such as an enzyme) that facilitates a reaction but is not changed during the reaction.

**Ceiling principle.** A procedure for setting a minimum DNA profile frequency proposed in 1992 by a National Academy of Science committee. One hundred persons from each of 15-20 genetically homogeneous populations spanning the range of racial and ethnic groups in the United States are sampled. For each allele, the higher frequency among the groups sampled (or 5%, whichever is larger) is used in calculating the profile frequency.

**Cell differentiation.** The process by which descendants of a common parental cell achieve specialized structure and function.

**Cell division.** The process whereby a mother cell gives rise to two identical daughter

cells (mitosis) or four gametes (meiosis).

**Cell line.** Cells that are made to grow continuously outside of the living organism from which they were taken.

**Centimorgan (cM).** A unit of measure of recombination frequency. One cM equals a 1% chance that a marker at one locus will be separated from a marker at a second locus due to crossing over in a single generation. In a human being, on average, one cM equates to one million base pairs.

**Central dogma.** The theory of DNA → RNA → protein flow of genetic information.

**Centromere.** The region of a chromosome that separates the two arms; centromeres are the sites of attachment of spindle fibers during cell division.

**Chimera.** An individual composed of two different cell lines originally derived from the union of different zygotes or by genetically distinct cells in the embryo.

**Chromatid.** One of the two attached daughter strands of a mother chromosome observable during mitosis or meiosis.

**Chromatin.** The combination of proteins, e.g., histones, and nucleic acids of which chromosomes are made.

**Chromatin loop.** A unit of DNA coiling consisting of a group of solenoids. Each loop is approximately 100 kb in size.

**Chromosome.** A threadlike structure consisting of chromatin by which hereditary information is physically transmitted from one generation to the next; a nuclear structure in eukaryotes that carries a portion of the genome. A human being has 46 chromosomes per nucleus, 22 homologous pairs of autosomes and two sex chromosomes. Prokaryotes have only one circular chromosome. Genes are arranged along the DNA of chromosomes.

**Chromosome abnormalities.** A major group of genetic diseases consisting of microscopically observable alterations of chromosome number or structure.

**Chromosome banding.** The process of applying specific stains to chromosomes in order to produce characteristic patterns of bands, e.g., G-banding.

**Clastogen.** A substance that can induce chromosome breakage.

**Cline.** Any gradual change in a particular characteristic or feature (phenotype).

**Clone.** A series of identical DNA fragments created by recombinant DNA techniques; also refers to identical cells or organisms that are descended from a single common ancestor.

**Coding DNA.** A small fraction (5%) of the human genome containing the blueprint for encoding proteins and a few special RNA molecules. The remainder of the DNA is “non-coding.”

**CODIS (Combined DNA Index System).** A collection of databases of STR and other loci of felons maintained by the Federal Bureau of Investigation.

**Codon.** A nucleotide triplet; a group of three mRNA bases, each of which specifies an amino acid to be inserted at a specified position in a protein.

**Complementary base pairing.** A fundamental process in which, ordinarily, A pairs only with T and C pairs only with G; sometimes known as “Watson-Crick” pairing.

**Complementary DNA (cDNA).** DNA formed by the enzyme reverse transcriptase copying the sequence in an RNA molecule.

**Complementary sequence.** A nucleic acid base sequence forming a double-stranded structure by matching base pairs, e.g., the complementary antiparallel sequence to GTAC is CATG.

**Compound heterozygote.** An individual who is heterozygous for two different disease-causing mutations at a locus. Compound heterozygotes for recessive disease mutations usually are affected with the disorder.

**Concordant.** When two individuals have the same trait.

**Conjugation.** The sexual reproduction of bacterial cells in which there is a one-way exchange of genetic material between the cells in contact.

**Conserved sequence.** A base sequence in a DNA molecule that has remained primarily unchanged in the evolutionary process.

**Contig.** From the word “contiguous,” a group of cloned DNA fragments representing overlapping regions of a given chromosome. A **contig map** represents the relative order of a complete chromosomal segment.

**Covalent.** A strong chemical bond formed between atoms, e.g., oxygen, O=O; water, H-O-H.

**Crossing over.** Exchange of homologous chromosome segments during meiosis, produces recombination.

**Cross-hybridization.** The binding of a probe to a DNA sequence other than the intended target sequence. This occurs because of homology between the probe and the sequence and because low-stringency hybridization wash conditions are followed.

**Cytoplasm.** That portion of the cell within the cell membrane.

**Cytosine (C).** A pyrimidine base; one of the four molecules containing nitrogen present in DNA and RNA. Designated by the letter C, it binds to guanine (G).

**D-loop.** A portion of the mitochondrial genome known as the “control region” or “displacement loop” instrumental in the regulation and initiation of mtDNA gene products.

**Daughter cells.** Cells resulting from the division of a parent cell.

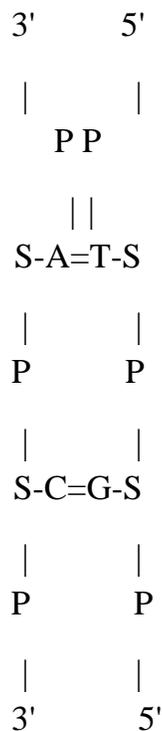
**Degenerate genetic code.** More than one RNA nucleotide triplet of bases coding for the same amino acid.

**Degradation.** The breaking down of macromolecules by chemical, enzymatic or physical means.

**Deletion.** The loss of one or more nucleotides from a DNA strand; may result in a gene mutation. A **deletion map** is a description of a specific chromosome using defined mutations as markers.

**Denaturation.** The process of separating the complementary double strands of DNA, as by heating, to form single strands in preparation for hybridization with biological probes.

**Deoxyribonucleic acid (DNA).** The molecule of heredity, DNA is composed of deoxyribonucleotide building blocks, each containing a base (adenine (A), thymine (T), cytosine (C) or guanine (G)), a deoxyribose sugar (S), and a phosphate group (P). The DNA molecule forms a double helix with the nucleotides of each strand held together through phosphate molecules at the 3' and 5' carbons of the sugar and the antiparallel complementary strands held by hydrogen (H) bonds between the base pairs, A=T and C=G. The two chains twine coaxially in a right-handed screw, one up and the other down; the diameter is twenty Ås in diameter, and a complete turn of the screw is 34Ås, each base pair flat in the middle with 3.4 Ås, a tenth of a revolution, separating each pair. During replication, the hydrogen bonds break; the complementary strands separate, and each acts as a template for the production of its complement.



**DNAase (deoxyribonuclease).** An enzyme capable of cleaving DNA into small fragments.

**DNA band.** A DNA fragment or allele on a Southern-blot autorad. With reference to an identity profile, a band is a tandem repeat DNA sequence (allele) produced by cleaving a genome into fragments with a restriction enzyme having recognition sites flanking the allele and usually at millions of other genome locations.

**DNA blot.** A membrane (usually nylon) with covalently bound single-stranded DNA.

**\*DNA fingerprinting.** The use of restriction enzymes to measure the genetic variations among individuals.

**DNA haplotype.** A pattern of DNA polymorphisms.

**DNA hybridization.** The formation of a double-stranded nucleic acid molecule from two separate strands; also applies to a molecular technique that uses one nucleic acid strand to locate another.

**DNA identification analysis.** The characterization of one or more features of an individual's genome by developing a DNA fragment band (allele) pattern. If a sufficient number of different-size bands are analyzed, the resultant bar-code profile will be unique for each individual except identical twins.

**DNA looping.** The formation of looped structures in DNA; sometimes permits the interaction of various regulatory elements.

**DNA mismatch repair.** A type of DNA repair in which nucleotide mismatches, i.e., violations of the A-T, C-G complementary base-pairing rule, are corrected by specialized repair enzymes.

**DNA polymerase.** An enzyme that catalyzes the linking of deoxyribonucleotide triphosphates using DNA as a template.

**DNA probe.** A piece of nucleic acid labeled with a radioactive isotope, dye or enzyme used to locate a particular nucleotide sequence or gene on a DNA molecule.

**DNA profile.** The alleles at each locus revealing the population variation in genome sequences; a series of DNA polymorphisms, usually VNTRs or microsatellites, typed in an individual. Because these polymorphisms are highly variable, the combined genotypes are useful in identifying individuals for forensic purposes.

**DNA repair.** A process in which mistakes in the DNA sequence are altered to recreate the original sequence.

**DNA replication.** The use of existing DNA as a template for the synthesis of new DNA strands.

**DNA sequence.** The order of DNA bases along a DNA molecule.

**Depurination.** The process of partial DNA hydrolysis by acid at purine (adenine and guanine) sites, resulting in the cleavage of large DNA fragments into smaller pieces. This process improves Southern transfer.

**Derivative chromosome.** A chromosome that is the result of a translocation of one part of a chromosome to another.

**Dialysis.** The process of separating different-size molecules in solution by means of their differential transfer across a porous membrane; commonly used to remove salt from solutions of macromolecules.

**Dideoxy method.** A technique for sequencing DNA in which dideoxynucleotides that terminate replication are incorporated into the replicated DNA strands.

**Differentiation.** The process of biochemical and structural changes by which cells become specialized in form and function.

**Digested DNA.** DNA cleaved by the action of restriction enzymes or other DNAses.

**Diploid.** Having two copies of each chromosome. See haploid.

**Discordant.** When two individuals do not share the same trait.

**Disomy.** The presence in a cell of two chromosomes derived from a single parent and none from the other parent.

**Dispersed repetitive DNA.** A class of repeated DNA sequences in which single repeats are scattered throughout the genome.

**Dizygotic.** Twins produced from two separate zygotes (fraternal twins).

**Domain.** A discrete portion of a protein with its own function.

**Dominant.** Allele the effect of which is the same in single copy (heterozygotes) as in double copy (homozygotes).

**Dominant negative.** A type of mutation in which the altered protein product in a heterozygote forms a complex with the normal protein product produced by the

homologous normal gene, thus disabling it.

**Dot blot.** A DNA analysis system in which sample DNA is directly pipetted onto a membrane, as opposed to the Southern-blot procedure of enzymatic digestion, electrophoresis and Southern transfer.

**Double helix.** A term often used to describe the “twisted ladder” shape that two linear strands of DNA assume when bonded together.

**Double-stranded DNA breaks.** A type of DNA breakage in which both strands are broken at a specific location.

**Duplication.** The presence of an extra copy of chromosome material.

**Electrophoresis.** The process of separating charged molecules, e.g., negatively charged DNA fragments in a porous medium such as agarose, by the application of an electric field. The DNA migrates through the medium at different rates according to length.

**Electroporation.** The use of an electric field to create reversible small holes in a cell wall or membrane through which foreign DNA can pass; this DNA can then integrate into the cell’s genome.

**Embryonic stem cells.** Cells that can give rise to any type of differentiated cell.

**Endogamy.** Reproduction by the fusion of gametes of similar ancestry.

**Endonuclease.** An enzyme that cleaves the phosphodiester bond within a nucleotide chain.

**Enhancer.** Regulatory DNA sequence that interacts with specific transcription factors to increase the transcription of genes.

**Enzyme.** A protein that catalyzes (speeds up) a specific chemical reaction without being changed or consumed in the process.

**Ethidium bromide.** A molecule that can intercalate into DNA double helices used to identify the presence of DNA in a sample by its fluorescence under ultraviolet light.

**Eukaryote.** A multi-cellular organism having true membrane-bound nuclei containing chromosomes that undergo mitosis.

**Euploid.** Cells whose chromosome number is a multiple of (in human beings 23).

**Exogenous DNA.** DNA originating outside an organism.

**Exon.** Portion of a gene that encodes amino-acid sequences and is retained after the primary mRNA transcript is spliced. Exons consist of the code signals for (1) the initiation of RNA transcription and ribosomal accommodation, (2) termination of translation and addition of poly-A tail.

**Exon trapping.** Method for isolating exons in a fragment of genomic DNA by using an in vitro cell system to artificially splice out the introns.

**Exonuclease.** An enzyme that digests DNA strands starting at their termini.

**Expanded repeat.** A type of mutation in which a tandem dinucleotide, trinucleotide, tetranucleotide ... repeat increases in number, e.g., Huntington disease.

**Expression.** The manifestation of a gene.

**Extra-chromosomal inheritance.** Cytoplasmic inheritance of DNA via organelles such as mitochondria or plasmids in female gametes. The human egg, for example, transmits approximately 10kbp mitochondria.

**Extra-nuclear DNA.** DNA located in organelles such as mitochondria and plasmids; also referred to as cytoplasmic DNA and its inheritance as maternal or cytoplasmic since the organelles are transmitted only from the female via gamete cytoplasm.

**False match.** Two samples of DNA that have different profiles could be declared to match if, instead of measuring the distinct DNA in each sample, there is an error in handling or preparing samples such that the DNA from a single sample is analyzed twice or the alleles are too similar to be distinguished.

**Founder Effect.** Arises when a new and isolated environment is invaded by only a few members of a species, which then multiply rapidly, the result of which is that there is a sharp loss of genetic variation compared with the parent population. The new population then may be distinctively different, genetically and phenotypically. In extreme cases, founder effect may lead to the speciation and subsequent evolution of new species.

**Frameshift mutation.** An alteration of DNA in which a duplication or deletion occurs in a coding region that is not a multiple of three base pairs, thus altering the frame of readout.

**Fusion.** The joinder of the membrane of two cells, thereby creating a daughter cell that contains the components and hence some of the same properties from each parent.

**Fusion gene.** A gene that results from a combination of two genes or parts of two genes.

**Gamete.** The haploid germ cell or reproductive cell, ova or sperm.

**Gel.** Semisolid porous matrix, usually agarose or acrylamide, used in electrophoresis to separate molecules based on its sieving properties.

**Gel electrophoresis.** The process of sorting DNA fragments by size by applying an electric current to a gel. The different-sized fragments move at different rates through the gel.

**Gene.** The fundamental unit of heredity; a set of nucleotide base pairs or sequence of DNA nucleotides on a chromosome containing the sequence encoding a specific function.

**Gene amplification.** The increase within a cell of the number of copies of a given gene.

**Gene expression.** The process through which a gene's encoded information is converted into cell-operating structures. Expressed genes include both those that are transcribed into mRNA and then protein and those that are not translated into protein, e.g., transfer and ribosomal RNA.

**Gene family.** A group of genes that are similar in DNA sequence and have evolved from a single common ancestral gene; may or may not be located in the same chromosome region.

**Gene flow.** The exchange of genes between different organisms usually through generational inheritance.

**Gene frequency.** The relative frequency of an occurrence of a particular allele in a population.

**Gene mapping.** Determination of the relative positions of genes on a DNA molecule (chromosome or plasmid) and of the distance in linkage units or physical units between them.

**Gene product.** The biochemical material resulting from the expression of a gene used to measure how active a gene is.

**Gene sequencing.** The determination of the sequence of nucleotide bases in a strand of DNA.

**Gene therapy.** The replacement of a defective gene in an organism suffering from a genetic disease.

**Genetic code.** The combination of 64 mRNA codons that specify the 20 amino acids.

**Genetic drift.** An evolutionary process in which gene frequencies change as a result of random fluctuations in the transmission of genes from one generation to the next. Drift is greater in smaller populations.

**Genetic mapping.** The ordering of genes on chromosomes according to recombination frequency.

**Genetics.** The study of the patterns, processes and mechanisms of inheritance of biological characteristics.

**Genome.** All of the genetic materials in the chromosomes of a particular organism. Its haploid size is generally given as the total number of base pairs.

**Genome scan.** A gene-mapping approach in which many markers from the human genome are tested for linkage with a disease phenotype.

**Genomic sequence.** The order of bases that constitute a particular fragment of DNA in a genome.

**Genomics.** The study of genes and their functions.

**Genotype.** The genetic makeup of an organism; the particular forms (alleles) of a set of genes possessed by an organism; an individual's allelic constitution at a locus.

**Genotype frequency.** The proportion of individuals in a population that carries a specific genotype.

**Genotype, single-locus.** The alleles an organism possesses at a particular site in its genome.

**Genotype, multi-locus.** The alleles an organism possesses at several sites in its genome.

**Germ cell.** Sex cell, the haploid ova or sperm.

**Germ-line.** Cells responsible for the production of gametes.

**Guanine (G).** A purine base; one of the four molecules containing nitrogen present in DNA and RNA. Designated by the letter G, it binds to C.

**Hae III.** A particular restriction enzyme, derived from *Haemophilus influenza*.

**Haploid.** A single set of chromosomes present in the sperm and egg cells, 23 being the haploid number in a human being. When a sperm cell fertilizes an egg cell, the number of chromosomes double (the diploid number).

**Haplotype.** From haploid genotype, the allelic constitution of multiple loci on a single chromosome.

**Hardy-Weinberg Principle.** Specifies equilibrium relationship between gene frequencies and genotype frequencies in a population, e.g., in a large random intrabreeding population not subjected to excessive selection or mutation the gene and genotype frequencies will remain constant over time; a condition in which the allele frequencies within a large random intrabreeding population are unrelated to patterns of mating. In this condition, the occurrence of alleles from each parent will be independent and have a joint frequency estimated by the product rule.

**Helper cells.** See packaging cells.

**Hemizygous.** A gene present in only a single copy, most commonly referring to genes on the single male X chromosome but can refer to other genes in the haploid state.

**Heredity.** The transmission of characteristics from one generation to the next.

**Heritability.** The proportion of population variance in a trait that can be ascribed to genetic factors.

**Heterogeneity, allelic.** Describes conditions in which different alleles at a locus can

produce variable expression of a disease. Depending on phenotype definition, allelic heterogeneity may cause two distinct diseases as in Duchenne and Becker muscular dystrophy.

**Heterogeneity, locus.** Describes diseases in which mutations at distinct loci can produce the same disease phenotype, e.g., retinitis pigmentosa, osteogenesis imperfecta.

**Heterologous.** Refers to segments of DNA derived from different sources.

**Heteroplasty.** The condition in which some copies of mitochondrial DNA in the same individual have different base pairs at certain points.

**Heterozygosity.** The presence of different alleles at a locus or loci on homologous chromosomes.

**Heterozygote.** An individual who has two different alleles at a locus.

**High-resolution banding.** Chromosome banding using prophase or prometaphase chromosomes which are more extended than metaphase chromosomes and thus yield more bands and greater resolution.

**Histone.** The protein core around which DNA is wound in a chromosome.

**Homologies.** Similarities in DNA or protein sequences between molecules.

**Homologous.** DNA or amino-acid sequences that are recognizably similar to one another; the chromosome pairs found in diploid organisms. The human being has 22 homologous pairs of autosomes plus two sex chromosomes per nucleus. The members of each pair have an identical sequence of genes; however, the alleles at corresponding loci may be identical (homozygous) or different (heterozygous).

**Homologs.** Chromosomes that are homologous.

**Homozygote.** An individual in whom the two alleles at a locus are the same.

**Hormone.** A chemical or protein that acts as a messenger or stimulatory signal, relaying instructions to modulate certain physiological activities.

**Housekeeping genes.** Genes whose protein products are required for general cellular maintenance or metabolism and therefore expressed in all cells.

**Human artificial chromosome.** A synthetic chromosome consisting of a centromere and telomeres and an insert of human DNA that can be 5-10 Mb in size.

**Hybridization.** The pairing of complementary strands of DNA, or DNA and RNA, by matching at base-pair sites. For example, a primer with the sequence AGGTCT would bond with the complementary sequence TCCAGA on a DNA fragment.

**Hydrogen bond.** A relatively weak bond between a hydrogen (H) atom, covalently bound to a nitrogen (N) or oxygen (O) atom and another atom. These bonds can be broken by increasing temperature.

**Hypervariable region.** A segment of a chromosome characterized by considerable variation in the number of alleles at a locus or loci.

**Imprinting, genomic.** Process in which genetic material is expressed differently when inherited from the mother than when inherited from the father, e.g., Angelman's syndrome (mother) versus Prader-Willi syndrome (father).

**In situ hybridization.** Molecular gene-mapping technique in which labeled probes are hybridized to stained metaphase chromosomes and then imaged to reveal the position of the probe.

**In vitro.** Means "in glass" and refers to a biological process performed outside a living organism, e.g., in the laboratory.

**In vivo.** Refers to a biological process within a living organism.

**Independent assortment.** One of Mendel's fundamental principles: Alleles at different loci are transmitted independently of one another, although this has proven to be true only if they are sufficiently distant from one another.

**Inducer.** A molecule or substance that increases the rate of expression of a specific gene.

**Insertion.** Addition of one or more nucleotides into a DNA strand. This may result in a gene mutation.

**Intergenic.** Nucleotide sequences located between genes.

**Interphase.** Portion of the cell cycle that alternates with meiosis or mitosis. DNA is replicated and repaired during this phase.

**Intron.** A sequence of DNA between two exons within a gene that is transcribed into mature mRNA but excised and degraded prior to translation and does not code for a protein. A number of introns of variable length may be separating the exons of a gene.

**Inversion.** A structural rearrangement of a chromosome in which two breaks occur followed by the reinsertion of the chromosome segment but in reversed order. **Paracentric** does not include the centromere. **Pericentric** includes the centromere.

**Isochromosome.** A structural chromosome rearrangement caused by the division of a chromosome along an axis perpendicular to the usual axis of division resulting in chromosomes with either two short arms or two long arms.

**Karyotype.** A display of chromosomes ordered according to length and banding pattern.

**Kilobase (kb).** Unit of length for DNA fragments equal to 1000 nucleotides or base pairs.

**Lamarckism.** A (discredited) theory that adaptations to the environment will cause heritable changes.

**Ligase.** An enzyme used to join DNA or RNA segments.

**Linkage.** The proximity of two or more markers on a chromosome; the closer the markers, the lower the probability that they will be separated during DNA replication or repair and the greater the probability that they will be inherited together.

**Linkage disequilibrium.** A specific allele of one locus being associated or linked to a specific allele or marker of another locus on the same chromosome with a greater frequency than expected by chance.

**Linkage equilibrium.** A condition in which the occurrence of alleles at different loci is independent.

**Linkage map.** A map of the relative positions of loci on a chromosome determined on the basis of how often the loci are inherited together and measured in cM.

**Localize.** Determination of the original locus of a gene or other chromosome marker on a chromosome.

**Locus** (plural: **loci**). The specific physical location of a gene or other chromosome marker on a chromosome; the DNA at that position.

**Long interspersed elements (LINEs).** A class of dispersed repetitive DNA in which each repeat is relatively long, up to 7 kb.

**Long-range restriction map.** Depiction of the positions on a chromosome of restriction enzyme cutting sites used as markers of specific areas along the chromosome; details the positions on the DNA molecule that are cut by particular restriction enzymes.

**Macrorestriction map.** Map depicting the order of and distance between the sites at which restriction enzymes cleave chromosomes.

**Major gene.** A single locus responsible for a trait.

**Manifesting heterozygote.** An individual who is heterozygous for a recessive trait but displays the trait. Most commonly used to describe females heterozygous for an X-linked trait who display the trait due to X inactivation.

**Marker.** A DNA sequence or gene of known location on a chromosome and phenotype that is used as a point of reference in mapping other loci.

**Mass spectroscopy.** The separation of molecules according to their molecular mass. In the version for analyzing DNA, small quantities of PCR-amplified fragments are irradiated with a laser to form gaseous ions that in an electric field traverse a fixed distance. Heavier ions have longer times of flight, and the process is known as “matrix-assisted laser desorption-ionization time-of-flight mass spectroscopy” or MALDI-TOF-MS.

**Match.** The presence of the same allele or alleles in two samples. Two DNA profiles are declared to match when they are indistinguishable in genetic type. For loci with discrete alleles, two samples match when they display the same set of alleles. For RFLP testing of VNTRs, two samples match when the pattern of the bands is similar and the positions of the corresponding bands at each locus fall within a preset distance (the **match window**).

**Megabase (Mb).** Unit of DNA equal to one million nucleotides and approximately equal to one cM.

**Meiosis.** Cell-division process in which haploid gametes are formed from diploid germ cells.

**Meiotic failure.** Aberrant meiosis in which a diploid gamete is produced rather than the normal haploid gamete.

**Melt.** The process of disrupting the hydrogen bonds linking complementary DNA strands producing the two single strands.

**Mendelian.** Referring to Gregor Mendel, describes a trait that is attributable to a single gene, assorting according to **Mendel's laws**: Segregation: during meiosis only one member of each homologous chromosome pair is transferred to a specific gamete; independent assortment: during meiosis the members of the different homologous chromosome pairs sort independently when transferred to a specific gamete, for example, AA' and BB' homologous chromosome pairs could give rise to AB, AB', A'B or A'B' possible gametes in a ratio of 1:1:1:1.

**Messenger RNA (mRNA).** RNA molecule formed from the transcription of DNA. Prior to intron splicing, mRNA is termed a primary transcript; after splicing, the mature transcript proceeds to the cytoplasm where it is translated into an amino acid sequence.

**Metabolism.** All of the biochemical activities that are carried out by an organism to maintain life.

**Metacentric.** A chromosome in which the centromere is located approximately in the middle of the chromosome arm.

**Metaphase.** A stage of mitosis and meiosis in which homologous chromosomes are arranged along the equatorial plane or metaphase plate of the cell, the stage at which chromosomes are maximally condensed and most easily visualized.

**Methylation (<sup>me</sup>).** One form of methylation, the most common in mammals, involves the conversion of cytosine to 5-methyl cytosine. Methylation can prevent cleavage of DNA at a restriction enzyme restriction site, e.g., Hpa cleaves at CCGG but not at C<sup>me</sup>CGG. Methylation of DNA alters its transcriptional potential.

**Microdeletion.** A chromosome deletion too small to be visible under a microscope, e.g.,

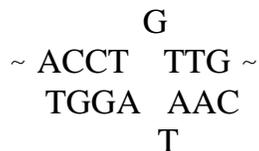
Prader-Willi syndrome, DiGeorge syndrome.

**Microsatellite.** An STR; a type of satellite DNA that consists of small repeat units (usually 2-5 bp) occurring in tandem.

**Microsatellite repeat polymorphism.** A type of genetic variation in populations consisting of differing numbers of microsatellite repeat units at a locus.

**Minisatellites.** A VNTR; a type of satellite DNA that consists of tandem repeat units that are each 20-70 bp in length. A variation in the number of minisatellite repeats is the basis of VNTR polymorphisms.

**Mismatch.** Also known as **mispairing**. Bases that do not match in “complementary” DNA strands. Depending on the blot wash stringency conditions, some mismatches can be tolerated between hybridized sample and probe DNA complementary regions. For example:



**Missense mutation.** A codon change resulting in a different amino acid in a protein.

**Mitochondrion** (plural: **mitochondria**). A DNA-containing cytoplasmic structure (organelle) of nucleated (eukaryotic) cells that is the site of the energy-producing reactions within the cell, the sites of ATP production. Mitochondria contain their own DNA (**mtDNA**) inherited only from the mother.

**Mitosis.** The process whereby a somatic cell nucleus after chromosomal replication divides to form two identical nuclei.

**Mobile elements.** DNA sequences capable of inserting themselves into other locations in the genome; **transposons**.

**Modifier gene.** A gene that alters the expression or activity of a gene at another locus.

**Monoclonal.** A group of cells consisting of a single clone, i.e., all cells are derived from the same single ancestral cell.

**Monogenic.** Describing a single-gene or Mendelian trait.

**Monomorphic.** A gene or DNA characteristic almost always found in only one form in a population.

**Monomorphic bands.** Each different-size monomorphic fragment is detected by cleaving genomic DNA with a specific restriction enzyme and hybridizing with a specific monomorphic probe. These fragments provide markers for use in quality control, especially as related to band shifts.

**Monosomy.** An aneuploid condition in which a specific chromosome is present in only a single copy, giving a human being a total of 45 chromosomes.

**Monozygotic.** Twins produced from a single zygote which later splits; the genomes are therefore identical.

**Morphogenesis.** The developmental process or formation of a cell, organ or organism.

**Mosaic.** Two or more genetically different cell lines in a single individual.

**Multi-allele.** Refers to a number of different possible alleles at a specific locus.

**Multi-factorial trait or disease.** A trait or disease resulting from the interaction of multiple genetic and environmental factors.

**Multi-locus.** Refers to a number of different loci or positions in the genome.

**Multi-locus probe.** A (rarely used) probe that marks multiple sites (loci). RFLP analysis using a multi-locus probe will yield an autorad showing a pattern of 30 or more bands.

**Multi-plexing.** Typing several loci simultaneously, thereby increasing sequencing speed.

**Multi-point mapping.** A type of genetic mapping in which the recombination frequencies among three or more loci are estimated simultaneously.

**Mutation.** Any heritable change in DNA sequence.

**Natural selection.** An evolutionary process in which individuals with favorable genotypes produce relatively greater numbers of surviving offspring that compete well.

**Neutral mutation.** Any change in the sequence of genomic DNA that does not affect to an extent that is detectable in an individual organism.

**Noncoding DNA.** DNA that does not encode a product or protein.

**Nonsense mutation.** A type of mutation in which an mRNA stop codon is produced resulting in the premature termination of translation.

**Northern blotting.** A gene expression assay in which mRNA on a blot is hybridized with a labeled probe.

**Nuclease.** An enzyme that breaks nucleic acids into their constituent nucleotides by cleaving the chemical bonds.

**Nucleic acid.** A nucleotide polymer of which DNA and RNA are the major types. Nucleic acid has three constituents: a sugar, ribose, built from a pentagonal ring of carbon atoms; a phosphate or phosphorus atom surrounded by four oxygen atoms; and one of the five bases – guanine, adenine, cytosine or thymine/uracil. In **DNA**, as opposed to **RNA**, the ribose lacks a fringe oxygen atom – thus de-oxy-ribose nucleic acid. The phosphates link and space the sugars, the third carbon of one sugar ring to the fifth carbon of the ring beyond, etcetera, and the base is linked to the sugar.

**Nucleoside.** A unit of nucleic acid composed of ribose or deoxyribose and a purine or pyrimidine base.

**Nucleosome.** A structural unit of chromatin in which 140-50 bp of DNA are wrapped around a core unit of eight histone molecules.

**Nucleotide.** A unit of nucleic acid composed of phosphate, ribose or deoxyribose, and a purine or pyrimidine base.

**Nucleotide excision repair.** A type of DNA repair in which altered groups of nucleotides are removed and replaced with properly pairing nucleotides.

**Nucleus.** The cellular organelle in eukaryotes that contains the genetic material.

**Obligate carrier.** An individual who necessarily possesses a disease-causing gene but may or may not be affected with the disease phenotype.

**Oligonucleotide.** A DNA sequence of fewer than 100 nucleotides often used as a primer or a probe in PCR.

**Oncogene.** A gene that can transform cells into a highly proliferative state causing cancer.

**Oogenesis.** The process in which ova are produced.

**Operator.** The region of a chromosome, adjacent to the operon, where a repressor or inducer protein binds to modulate transcription of the operon.

**Operon.** A sequence of genes responsible for synthesizing a set of enzymes needed for a pathway of biosynthesis of a molecule. An operon is controlled by an operator.

**pH.** A measure of the acidity or alkalinity of a solution.

**Packaging cells and helper cells.** Packaging cells are those in which replication-deficient viruses are placed so that the replication machinery of the cells can encapsulate foreign genes. Helper cells provide missing functions of defective viruses so that the cells can make viral copies.

**Palindrome.** A DNA sequence the complementary sequence of which is the same if read backwards, e.g., 5' AATGCGCATT 3'.  
TTACGCGTAA

**Panmixia.** Describes a population in which individuals mate at random with respect to a specific genotype.

**Pathogen.** A disease-causing agent such as a virus or bacterium.

**Pattern formation.** The spatial arrangement of differentiated cells to form tissues and organs during embryonic development.

**Penetrance.** In a population, the proportion of individuals possessing a disease-causing genotype who express the disease phenotype. When this proportion is less than 100%, the disease genotype is said to have reduced or incomplete penetrance.

**Peptide.** Two or more amino acids joined by a peptide bond.

**Pharmacogenetics** is the study of the hereditary basis for differences in a population's

response to a drug. **Pharmacogenomics** describes the science of using genomic technology to identify drug targets and study genetic variations that have an impact on the safety and efficacy of drugs.

**Phenocopy.** A phenotype that resembles the phenotype produced by a specific gene but is due instead to a different, typically non-genetic, factor.

**Phenotype.** The physical make-up of an individual as defined by genetic and non-genetic factors.

**Physical mapping.** The determination of physical distances between genes using cytogenetic and molecular techniques; the mapping of the locations as separated by base pairs of identifiable DNA landmarks, e.g., genes, restriction enzyme cutting sites, regardless of inheritance. For the human genome, the lowest-resolution physical map is the banding patterns on the different chromosomes.

**Plasmid.** An extra-chromosomal circular double-stranded DNA element native to certain bacteria, capable of replication. Plasmids are used as vehicles to replicate cloned (recombinant) DNA sequences.

**Pleiotropy.** Describes genes that have multiple phenotypic effects, e.g., Marfan syndrome, cystic fibrosis.

**Pluripotent cells.** Having the capacity to become any kind of cell or tissue in the body. Embryonic stem cells and cells of the inner cell mass are pluripotent.

**Polarity.** Direction, e.g., definition of anterior versus posterior in axis specification or 3' versus 5' of nucleic acids.

**Poly-A tail.** The adenine (A) nucleotide polymer often attached to the 3' end of primary-mRNA.

**Polyclonal.** Derived from different types of cells.

**Polygenic.** Describes a trait caused by the combined additive effects of alleles of multiple genes, e.g., certain types of heart disease, some cancers, diabetes. Although a polygenic disorder is inherited, it depends on the simultaneous presence of several alleles and therefore the hereditary patterns usually are more complex than those of single-gene disorders.

**Polymerase.** The general term for enzymes that carry out the synthesis of nucleic acids.

**Polymerase Chain Reaction (PCR).** A technique for amplifying a large number of copies of a specific DNA sequence flanked by two oligonucleotide primers. The DNA is alternately heated and cooled in the presence of DNA polymerase and free nucleotides so that the specified DNA segment is denatured, hybridized with primers and extended by DNA polymerase.

**Polymerase, DNA or RNA.** Enzymes that catalyze the synthesis of nucleic acids on preexisting nucleic acid templates, assembling RNA from ribonucleotides or DNA from deoxyribonucleotides.

**Polymorphism.** A difference in DNA sequence among individuals. A locus in which two or more alleles have gene frequencies greater than 0.01 in a population is considered a useful polymorphism for genetic linkage analysis. When this criterion is not fulfilled, the locus is considered to be **monomorphic**.

**Polypeptide.** A series of amino acids linked together by peptide bonds.

**Polyploidy.** Chromosome abnormality in which the number of chromosomes in a human cell is a multiple of 23 greater than two, the diploid number, e.g., **triploidy** in which the individual has three copies of each chromosome.

**Positional cloning.** A technique to identify genes based on their location on a chromosome.

**Post-translational modification.** Various types of additions and alterations of a polypeptide that take place after mRNA is translated into a polypeptide.

**Prehybridization.** The process of incubating a DNA blot with a hybridization solution containing complex DNA to in part block cross-hybridizing sites. This precedes the addition of the labeled probe.

**Primary transcript.** The pre-mRNA molecule directly after its transcription from DNA. A **mature mRNA transcript** is formed from the primary transcript when the introns are spliced out and poly-A is added.

**Primer.** An oligonucleotide that flanks either side of a DNA fragment and provides a point for complementary nucleotides to attach and replicate the DNA strand in the PCR process.

**Primer extension.** Part of the PCR process in which DNA polymerase extends the DNA sequence beginning at an oligonucleotide primer.

**Probe.** Single-stranded DNA or RNA of a specific base sequence, labeled either radioactively (usually used for RFLP analysis) or biochemically (usually used for PCR-based analysis) that is used to detect the complementary base sequence by hybridization.

**Prokaryote.** A unicellular organism lacking a membrane-bound nucleus, e.g., bacteria.

**Promoter.** A DNA site to which RNA polymerase will bind and initiate transcription.

**Protein.** A large molecule composed of one or more chains of amino acids in a specific order, the order determined by the base sequence of nucleotides in the gene coding for the protein. Proteins are required for the structure, function and regulation of the body cells, tissues and organs, and each protein has unique functions, e.g., hormones, enzymes and antibodies.

**Protein electrophoresis.** A technique in which amino acid variations are identified on the basis of charge differences that cause differential mobility of polypeptides in an electrical field.

**Proteomics.** Each cell produces thousands of proteins, each of which has a specific function. The collection of proteins in a cell is the proteome, and, unlike the genome, which is constant irrespective of cell type, the proteome varies from one cell type to the next. The science of proteomics seeks to identify the protein profile of each cell type, assess protein differences and uncover not only each protein's specific function but how each protein interacts with other proteins.

**Pseudogene.** A gene that is highly similar in sequence to another gene or genes but has been rendered transcriptionally or translationally inactive by mutations.

**Pulsed-field gel electrophoresis.** A type of electrophoresis suitable for large DNA fragments. The fragment is moved through a gel by alternating pulses of electricity across fields that are 90 degrees in orientation from one another.

**Punnett Square.** A grid developed by English geneticist Reginald Crundall Punnett (1875-1967) specifying the genotypes that can arise from the gametes contributed by a mating pair of individuals, i.e., a table that shows the possible genetic outcomes of a mating.

**Purine.** The two DNA and RNA bases, adenine and guanine, that consist of two carbon-nitrogen rings.

**Pyrimidine.** The bases cytosine and thymine in DNA and cytosine and uracil in RNA that consist of single carbon-nitrogen rings.

**Random match.** A match in the DNA profiles of two samples of DNA when one is drawn at random from the population.

**Random-match probability.** The chance of a random match. As often used, this refers to the probability of a true match when DNA being compared to evidentiary DNA comes from a person drawn at random from the population. A random-true-match probability reveals the probability of a true match when samples of DNA come from unrelated persons.

**Random mating.** Like panmixia, the members of a population are said to mate randomly with respect to particular genes when the choice of mates is independent of the alleles.

**Receptor.** A cell-surface component that binds to extra-cellular molecules.

**Recessive.** A trait phenotypically expressed only when an allele is present in the homozygous state. The recessive allele is masked by a dominant allele when the two occur together in a heterozygote.

**Recombinant DNA.** DNA formed by the union of two heterologous DNA molecules, e.g., the ligation of a human growth hormone gene into a plasmid.

**Recombinant DNA technology.** Procedure used to join DNA sequences in an in vitro cell-free system. Under appropriate conditions, a recombinant DNA molecule can enter a cell and replicate there either autonomously or after it has become integrated into a cellular chromosome.

**Recombination.** Combinations of genes in offspring different from those in the parents due to independent assortment and during crossing-over.

**Recombination frequency.** The proportion of meioses in which recombinants between two loci are observed used to estimate genetic distances between loci.

**Reduction division.** The first stage of meiosis (meiosis I) in which the chromosome

number is reduced from diploid to haploid.

**Redundancy, genetic.** The existence of alternate genetic mechanisms or pathways that can compensate when another mechanism or pathway is disabled.

**Regulatory gene.** A gene that acts to control the expression of other genes.

**Regulatory sequence.** A DNA base sequence that controls gene expression.

**Repetitive DNA.** DNA sequences that are found in multiple copies in the genome, dispersed or repeated in tandem.

**Repetitive sequence.** A repeated series of bases in a DNA molecule.

**Replication.** The process in which the double-stranded DNA molecule is duplicated.

**Replication bubble.** Replication structures that occur in multiple locations on a chromosome allowing replication to proceed more rapidly.

**Replication origin.** The point at which replication begins on a DNA strand. In eukaryotes, each chromosome has numerous replication origins.

**Replicon.** A segment of DNA that can replicate independently.

**Repressor.** A protein that binds to an operator adjacent to a structural gene, inhibiting transcription of that gene.

**Restriction digest.** Process in which DNA is exposed to a restriction enzyme causing it to be cleaved into restriction fragments.

**Restriction enzyme.** An enzyme that breaks DNA in highly sequence-specific locations.

**Restriction enzyme cutting site.** A specific nucleotide sequence of DNA at which a particular restriction enzyme cuts the DNA.

**Restriction enzyme, endonuclease.** A protein that recognizes specific, short nucleotide sequences and cuts DNA uniquely at those sites.

**Restriction fragment.** A fragment of DNA resulting from digestion with a restriction enzyme.

**Restriction Fragment Length Polymorphism (RFLP).** Variations in DNA sequence in populations detected by digesting DNA with a restriction endonuclease, electrophoresing the resulting restriction fragments, transferring the fragments to a solid medium (blot) and hybridizing the DNA on the blot with a labeled probe. The polymorphism arises from the loss or creation of a restriction enzyme site.

**Restriction site.** A DNA sequence marking the location at which a specific restriction endonuclease cuts DNA into fragments also known as a **recognition site**.

**Restriction site polymorphism (RSP).** A variation in DNA sequence that is due to the presence or absence of a restriction site. This type of polymorphism is the basis for most traditional RFLPs.

**Retrovirus.** A type of RNA virus that can reverse transcribe its RNA into DNA for insertion into the genome of a host cell, useful as a vector for gene therapy.

**Reverse banding (R-banding).** A chromosome banding technique in which chromosomes are heated in a phosphate buffer; produces dark and light bands in patterns that are the reverse of those produced by G-banding.

**Reverse dot blot.** A detection method used to identify SNPs in which DNA probes are affixed to a membrane and amplified DNA is passed over the probes to see if it contains the complementary sequence.

**Reverse transcriptase (RNA-dependent DNA polymerase).** An enzyme that transcribes RNA into DNA (hence reverse).

**Ribonucleic acid (RNA).** Single-stranded nucleic acid molecules composed of a sugar (ribose), phosphate group and a series of bases: adenine, guanine, cytosine and uracil. The types of RNA include mRNA, rRNA and tRNA.

**Ribosomal RNA (rRNA).** RNA molecules that together with specific proteins form the subunits of ribosomes.

**Ribosome.** The site of translation of mature mRNA into amino-acid sequences.

**RNAase.** An enzyme capable of degrading RNA.

**RNA interference.** A process that uses RNA sequence to block gene expression.

**RNA polymerase.** Enzyme that binds to a promoter site and synthesizes mRNA from a DNA template.

**Ribosome.** A cellular component containing protein and RNA that is involved in protein synthesis.

**Ribozyme.** An mRNA molecule that has catalytic activity.

**Ring chromosome.** A structural chromosome abnormality formed when both ends of a chromosome are lost and the new ends fuse together.

**Satellite DNA.** Repetitive DNA located in telomeres and centromeres.

**Sensitivity.** The proportion of affected individuals who are correctly identified by a test; true positives.

**Sequence-specific oligonucleotide (SSO) probe** also known as an allele-specific oligonucleotide probe. Oligonucleotide probes used in a PCR-associated detection technique to identify the presence or absence of certain base-pair sequences identifying different alleles. The probes are visualized by an array of dots rather than the electrophoretograms associated with RFLP analysis.

**Sequence tagged site (STS).** DNA sequences of several hundred bp in size flanked by PCR primers. Because the chromosome location has been established, the site is useful as an indicator of a physical position on the genome.

**Sequencing.** Determination of the order of base pairs in a DNA or RNA molecule or the order of amino acids in a protein.

**Sex chromosomes (X and Y chromosomes).** Chromosomes that are different in the two sexes and involved in sex determination. A female has two X chromosomes in diploid cells, and a male has an X chromosome and a Y chromosome. The sex chromosomes comprise the 23<sup>rd</sup> chromosome pair in a karyotype.

**Sex-influenced.** A trait the expression of which is modified by the gender of the

individual possessing the trait.

**Sex-linked.** A trait that is expressed in only one gender.

**Short interspersed elements (SINEs).** A class of dispersed repetitive DNA in which each repeat is relatively short.

**Short tandem repeats (STR).** Multiple copies of an identical DNA sequence arranged in direct succession in a particular region of a chromosome.

**Shotgun method.** A method involving randomly sequencing tiny cloned pieces of the genome with no prior knowledge of from where on a chromosome the piece originated. In **directed strategies**, the sequenced pieces of DNA come from adjacent stretches of a chromosome.

**Signal transduction.** Process in which biochemical messages are transmitted from the cell surface to the nucleus.

**Silencer.** A DNA sequence that binds to specific transcription factors to decrease or repress the activity of certain genes.

**Silent substitution.** DNA sequence change that does not change the amino-acid sequence because of the degeneracy of the genetic code.

**Single-copy DNA.** DNA sequences that occur only once in the haploid genome.

**Single-gene disorder or trait.** A feature or disease that is caused by a mutant allele of a single gene, e.g., Duchenne muscular dystrophy, retinoblastoma, sickle-cell disease.

**Single-locus probe.** A probe that only marks a specific site (locus). RFLP analysis using a single-locus probe will yield an autorad showing one band if the individual is homozygous, two bands if heterozygous.

**Single nucleotide polymorphism (SNP).** A substitution, insertion or deletion of a single base pair at a given point in the genome.

**Single-strand conformation polymorphism (SSCP).** A technique for detecting variation in DNA sequence by running single-stranded DNA fragments through a non-denaturing gel. Fragments with differing secondary structure (conformation) caused by sequence variation will migrate at different rates.

**Sister chromatids.** The two identical strands of a duplicated chromosome joined by a single centromere.

**Solenoid.** A structure of coiled DNA consisting of approximately six nucleosomes.

**Somatic cells.** All cells of eukaryotes excluding gametes and their precursors or, in other words cells other than sex or germ cells.

**Somatic cell gene therapy.** Therapy involving the insertion of genes into somatic cells for therapeutic purposes.

**Somatic cell nuclear transfer.** The transfer of a nucleus from a fully differentiated cell into an egg that has had its nucleus removed.

**Southern transfer or blot.** Transfer by absorption of DNA fragments separated in electrophoretic gels to a solid membrane such as nitrocellulose for detection of specific base sequences by radiolabeled complementary probes.

**Specificity.** The proportion of unaffected individuals who are correctly identified by a test; true negatives.

**Spliced mRNA.** mRNA after removal of the intron regions from the primary messenger and linking of the exon (coding) portions.

**Splice site mutation.** DNA sequence alterations in donor or acceptor sites or in the consensus sites near them that produces altered exon splicing such that portions of exons are deleted or portions of introns are included in the mature mRNA transcript.

**Splicing.** The removal of introns and joining of exons to form a continuous coding sequence in messenger RNA.

**Stem cells.** Embryonic cells that still are undifferentiated.

**Stop codon.** mRNA base triplets UGA, UAG or UAA signalling the end of mRNA translation into protein.

**Stringency.** The buffer salt concentration and temperature used in the DNA blot wash post-hybridization process. As these parameters are changed, the degree of the binding of the probe to target DNA changes.

**Strip.** Removal by melting of hydrogen bonds of hybridized probe from DNA blots.

**Structural gene.** Genes that encode protein products.

**Submetacentric.** A chromosome in which the centromere is located closer to one end of the chromosome arm than the other.

**Substrate.** Material acted on by an enzyme.

**Syndrome.** A pattern of multiple primary malformations or defects due to a single underlying cause, e.g., Down syndrome.

**Tandem repeat sequences.** Multiple copies of the same base sequence located directly next to each other on a chromosome, used as a marker in physical mapping.

**Taq polymerase.** A DNA polymerase isolated from the bacterium *Thermus aquaticus* that lives in hot springs. This enzyme is capable of withstanding high temperatures and therefore very useful in the PCR process.

**Target DNA.** The DNA sequence to be hybridized to a specific probe.

**Targeted disruption.** The disabling of a specific gene so that it is not expressed.

**Telomerase.** An enzyme that replaces the DNA sequences in telomeres during cell division.

**Telomere.** The tip of a chromosome.

**Template.** The single-stranded DNA blueprint for complementary strand assembly; the DNA strand from which mRNA is transcribed.

**Termination sequence.** The DNA sequence that signals the cessation of transcription.

**Terminator.** Sequence of DNA bases that tells the RNA polymerase to stop synthesizing RNA.

**Thymine (T).** A pyrimidine base, one of the four molecules containing nitrogen present in DNA. Designated by the letter T, it always binds with A.

**Totipotent.** The state of a cell that can give rise to any and all adult cell types.

**Transcription.** The process in which an mRNA sequence is synthesized from a DNA template.

**Transcription factor.** Protein that binds to DNA to influence and regulate transcription. General: class required for transcription of all structural genes. Specific: class that activates only specific genes at specific times.

**Transdifferentiation.** The process whereby a specialized cell de-differentiates and re-differentiates into a different cell type.

**Transduction.** The transfer of genetic material from one cell to another by means of a virus or phage vector.

**Transfection.** The transfer of a DNA sequence into a cell.

**Transfer RNA (tRNA).** RNA molecules that brings amino acids to ribosomes for protein production during mRNA translation. The anticodon portion of the tRNA binds to a complementary mRNA codon, and the 3' end of the tRNA molecule attaches to a specific amino acid.

**Transformation.** The process by which the genetic material carried by an individual cell is altered by the incorporation of exogenous DNA into its genome.

**Translation.** The process by which the information on a messenger RNA molecule is used to direct the synthesis of a protein.

**Transgenic.** Refers to an organism into which a gene has been introduced from an organism of another species.

**Translation.** The process in which an amino-acid sequence is assembled according to the sequence specified by the mature mRNA transcript.

**Translocation.** The exchange of genetic material between non-homologous chromosomes.

**Translocation, reciprocal.** A translocation resulting from breaks on two different chromosomes and a subsequent exchange of material. Carriers of reciprocal translocations maintain the normal number of chromosomes and normal amount of chromosome material.

**Translocation, Robertsonian.** A translocation in which the long arms of two acrocentric chromosomes are fused at the centromere; the short arms of each chromosome are lost. The carrier has 45 chromosomes instead of 46 but is phenotypically normal because the short arms contain no essential genetic material.

**Transposon.** See **mobile element**.

**Triplet code.** A code in which a series of three successive DNA or RNA nucleotide bases specify a particular amino acid.

**Trisomy.** An aneuploid condition in which the individual has an extra copy of one chromosome for a total of 47 chromosomes in each cell. **Partial trisomy.** Chromosomal abnormality in which a portion of a chromosome is present in three copies, may be produced by reciprocal translocation or unequal crossover.

**True match.** Two samples of DNA that have the same profile should match when tested. If there is no error in the labeling, handling and analysis of the samples and in the report of the results, a match is a true match.

**Unequal crossover.** Crossing over between improperly aligned DNA sequences producing deletions or duplications of genetic material.

**Uniparental disomy.** Condition in which two copies of one chromosome are derived from a single parent, and no copies are derived from the other parent.

**Uracil (U).** The pyrimidine base in RNA that appears in place of thymine in DNA.

**Variable expression.** A trait in which the same genotype may produce phenotypes of varying severity or expression, e.g., neurofibromatosis type I.

**Variable number tandem repeats (VNTRs).** A type of polymorphism created by variations in the number of minisatellite repeats in a defined region. The number of repeats varies from individual to individual, thus providing a basis for individual recognition.

**Vector.** An engineered DNA molecule into which any other DNA molecule can be inserted that then can be returned to an organism and replicated in order to amplify the DNA. The vector needs an origin for replication and usually has a gene, such as for antibiotic resistance, that permits selection for organisms carrying the vector and DNA of

interest. This is the key technology for cloning DNA molecules.

**Virus.** A non-cellular biological entity that can reproduce only within a host cell, it consists of nucleic acid covered by protein.

**Wash.** The process of removing non-bound or loosely bound probe from blots after hybridization, used to reduce background interference.

**X chromosome.** A chromosome responsible for sex determination. Two copies are present in the genome of the homogametic sex and one copy in the heterogametic sex. The human female has two X chromosomes and the male has one X chromosome and a Y chromosome.

**X inactivation.** Process in which genes from one X chromosome in each cell of the female embryo are rendered transcriptionally inactive.

**X-linked.** Genes on the X chromosome.

**Y chromosome.** A chromosome responsible for sex determination in the heterogametic sex.

**Zygoty.** Twin development from one or two zygotes. If one, the twins are identical (monozygotic); if two, they are fraternal (dizygotic).

**Zygote.** The diploid cell resulting from the union of a haploid egg and sperm, i.e., the diploid fertilized ovum.