Glossary

A

A site The ribosomal site to which new aminoacyl-tRNAs bind, with the exception of the initiator tRNA.

AAUAAA An element of the eukaryotic polyadenylation signal that dictates cleavage and polyadenylation about 20 nucleotides downstream.

abasic site A deoxyribose in a DNA strand that has lost its base through DNA damage.

Ac *Activator*, a maize transposon that encodes a transposase required for movement and can activate transposition of an inactive transposon like *Ds*.

acceptor stem The region of a tRNA molecule formed by base-pairing between the 5′ and 3′ ends of the molecule that becomes charged with an amino acid.

acentric chromosome A chromosome with no centromere.

acetylation The post-translational modification of a protein by the addition of an acetyl group(s) (e.g. histone acetylation).

acid A substance that releases an H^+ ion or proton in solution.

acid blob A transactivation domain in a transcription factor that is rich in acidic amino acids.

acrocentric chromosome A chromosome with the centromere near one end.

activator A protein that binds to an enhancer or other DNA regulatory element and activates transcription from a nearby promoter.

active site The part of an enzyme at which substrate molecules bind and are converted into their reaction products.

acylated tRNA A tRNA molecule to which an amino acid is linked.

adenine (A) A nitrogenous purine base found in DNA and RNA. Adenine always pairs with thymine in DNA and uracil in RNA.

adeno-associated virus A parvovirus that is being tested for use as a gene therapy vector.

adenosine A nucleoside containing the base adenine.

adenosine deaminase acting on RNA (ADAR) An RNA editing enzyme that deaminates certain adenosines in RNAs, converting them to inosines.

adenosine deaminase deficiency (ADA) A severe immunodeficiency disease that results from a lack of the enzyme adenosine deaminase. It usually leads to death within the first few months of life.

adenovirus A group of DNA-containing viruses which cause respiratory disease, including one form of the common cold. Genetically modified adenovirus vectors have been used in gene therapy trials to treat cystic fibrosis, cancer, and other diseases.

A-DNA An alternate form of the right-handed DNA double helix found at low relative humidity, with 11 base pairs per helical turn. It is unlikely that A-DNA is present in any length sections in cells. This is the form adopted in solution by an RNA–DNA hybrid.

ADP-ribosylation The enzymatic transfer of an ADP-ribose residue from NAD^+ to a specific amino acid of the acceptor protein by an ADP-ribosyltransferase.

affinity chromatography Chromatography that relies on affinity between the substance of interest and another substance immobilized on a resin.

AFLP Amplified fragment length polymorphism. The PCR amplification of restriction fragments to which adaptor oligomer sequences have been attached. Pronounced "a-flip."

agarose A polysaccharide extracted from seaweed; used as a gelling agent in electrophoresis of nucleic acids; its value is that few molecules bind to it, so it does not interfere with electrophoretic movement.

allele One of a number of alternative forms of a given gene (or DNA sequence) that can occupy a given genetic locus on a chromosome. Different alleles produce variation in inherited characteristics such as hair color or blood type. In an individual, one form of the allele (the dominant one) may be expressed more than another form (the recessive one).

allelic exclusion See monoallelic gene expression.
allolactose A rearranged version of lactose with a β-1,6-galactosidic bond; the real inducer of the *lac* operon.
allosteric protein Any protein whose activity is altered by a change of shape induced by binding a small molecule.
allosteric regulation Ligand-induced conformational (shape) changes that alter the activity of a protein.
alpha (α)-helix A fundamental unit of protein folding in which successive amino acids form a right-handed helical structure held together by hydrogen bonding between the amino and carboxyl components of the peptide bonds in successive loops of the helix.
alternative splicing Splicing the same pre-mRNA in two or more ways to yield two or more different mRNAs that produce two or more different protein products.
***Alu* element** A human nonautonomous retrotransposon that contains the AGCT sequence recognized by the restriction enzyme *AluI*. Also known as a short interspersed nuclear element (SINE).
amino acids A group of 20 common and two rare kinds of small molecules that link together in long chains to form proteins. Often referred to as the "building blocks" of proteins.
amino terminus The end of a polypeptide with a free amino group; i.e. the end at which protein synthesis began.
aminoacyl-tRNA An activated amino acid linked through a high-energy anhydride bond to the phosphate group of AMP. Created by aminoacyl-tRNA synthetase in the first step in tRNA charging.
aminoacyl-tRNA synthetase The enzyme that links a tRNA to its cognate amino acid.
A-minor motif A folding motif in RNA in which single-stranded adenosines make tertiary contacts with the minor grooves of RNA double helices by hydrogen bonding and van der Waals contacts.
ampicillin A derivative of penicillin that blocks synthesis of the peptidoglycan layer that lies between the inner and outer membranes of *E. coli*, and kills dividing cells.
amyloid fibrils Insoluble toxic deposits of aggregated proteins that accumulate in cells in certain types of protein misfolding diseases, such as Alzheimer's, Parkinson's, Huntington's, type II diabetes, and Creutzfeldt–Jakob disease.
anaphase A short stage of cell division in which the chromosomes move to the poles.
Angelman syndrome A neurodevelopmental disorder caused by defects in genomic imprinting in a region of chromosome 15, or by classic mutations in the UBE3A gene which is involved in the ubiquitin-mediated protein degradation pathway.
animal model See model organism.
annealing of DNA The process of bringing back together the two separate strands of denatured DNA to re-form a double helix. This is also referred to as renaturation.
annotated genes Genes or gene-like sequences from a genomic sequencing project that are at least partially characterized.
anode Positive (red) electrode.
antibody A blood protein that is produced in response to and counteracts an antigen with great specificity. Antibodies are produced in response to disease and help the body fight against the particular disease and develop immunity.
anticodon A 3-base sequence in tRNA that base pairs with a specific codon in mRNA.
anticodon loop The loop drawn by convention at the bottom of the tRNA molecule which contains the anticodon.
antigen A substance recognized and bound by an antibody.
antiparallel The relative polarities of the two strands in a DNA double helix (or in any double-stranded nucleic acid) ; if one strand runs 5′ → 3′, top to bottom, the other runs 3′ → 5′.
antisense DNA strand The noncoding strand in double-stranded DNA. The antisense strand serves as the template for mRNA synthesis.
antisense oligonucleotides Short oligonucleotides designed to selectively bind to a specific mRNA and induce antisense-mediated inhibition of gene expression. Either the mRNA strand in the hybrid duplex is cleaved by RNase H or translation is blocked.

antisense RNA An RNA complementary to an mRNA.
antiserum Blood serum containing an antibody or antibodies directed against a particular antigen.
AP endonuclease An enzyme that cuts a strand of DNA on the 5′ side of an apurinic or apyrimidinic site (AP site).
AP1 A transcription activator that is a heterodimer composed of one molecule each of Fos and Jun, or of a Jun-Jun homodimer.
APE1 (AP endonuclease 1) A mammalian enzyme that uses 3′ → 5′ endonuclease activity to edit the errors made by DNA polymerase β during base excision repair.
apoenzyme Inactive protein part of an enzyme remaining after removal of the prosthetic group or other necessary accessory protein subunit (see holoenzyme).
apolipoproteins Proteins involved in the metabolism of cholesterol.
apoptosis Programmed cell death, an organism's normal method of disposing of damaged, unwanted, or unneeded cells.
aptamer The domain of a riboswitch that selectively binds to the target metabolite.
apurinic site (AP site) A deoxyribose sugar in a DNA strand that has lost its purine base.
apyrimidinic site (AP site) A deoxyribose sugar in a DNA strand that has lost its pyrimidine base.
AraC The negative regulator of the arabinose operon.
Archaea A third domain of life. Although prokaryotic, they are as phylogenetically distinct from bacteria as they are from eukaryotes. The Archaea typically live in extreme environments, e.g. hot springs and salt lakes. Some are methane producers.
Argonaute One of the components of the RISC complex that has "Slicer" activity and cleaves the target mRNA during RNAi.
Artemis An endonuclease involved in end-processing of double-strand break sites in DNA during repair by nonhomologous end-joining.
A-T rich DNA sequences that have many adenine-thymine bases that promote DNA melting; e.g. at origins of replication and promoter regions.
ataxia Loss of muscle control of voluntary movements.
ataxia-telangiectasia A rare fatal disease involving a damaged immune system, unsteady walk, premature aging, and a strong predisposition to some kinds of cancer. People who possess only one copy of the gene, called ATM, do not have the disease, but may be predisposed to cancer and unusually sensitive to radiation.
ataxia-telangiectasia mutated (ATM) A serine-threonine kinase in the nucleus that is a key signal transducer for recruitment of factors involved in double-strand break repair by homologous recombination; e.g. ATM kinase activity is induced by exposure of cells to ionizing radiation.
atomic force microscopy (AFM) A technique in which the deflection of a cantilever tip as it scans the surface of a molecule is used to produce a topographic image of the molecule's surface.
ATP synthase An enzyme that generates ATP using the free energy of an electrochemical gradient of protons across a membrane.
ATPase An enzyme that hydrolyzes ATP, releasing energy for other cellular activities.
***att* sites** Sites on bacteriophage and host DNA where recombination occurs, allowing integration of the phage DNA into the host genome.
autonomous replicating sequence (ARS) A yeast origin of replication.
autoradiography A technique in which a radioactive sample is allowed to expose a photographic emulsion; e.g. an X-ray film.
autosomal dominant A pattern of Mendelian inheritance whereby an affected individual possesses one copy of a mutant allele and one normal allele. Individuals with autosomal dominant diseases have a 50–50 chance of passing the mutant allele and hence the disorder onto their children.
autosome Any chromosome other than a sex chromosome. Humans have 22 pairs of autosomes.
auxotroph An organism that requires specific supplementary compounds in order to grow (e.g. vitamins, amino acids, nucleotides, etc.) because it is unable to synthesize them.

B

backbone (DNA) The repeating sugar-phosphate groups that form the two strands of the DNA double helix.

backbone damage Types of DNA damage that include the formation of abasic sites and double-strand breaks.

bacteria A single-celled prokaryotic organism. Bacteria are found throughout nature and can be beneficial or pathogenic.

bacterial artificial chromosome (BAC) A vector based on the *E. coli* F plasmid capable of holding foreign DNA inserts up to 300,000 base pairs. Once the foreign DNA has been cloned and transformed into the host bacteria, many copies of it can be made.

bacteriophage A virus that infects bacteria (literally "bacterium eater"); often abbreviated as phage for short.

basal transcription Low levels of gene transcription determined by the frequency with which RNA polymerase binds a promoter and initiates transcription in the absence of transcription activators. The basal level is some 20–40-fold lower than activated levels of transcription.

base In general, a substance that accepts a H^+ ion or protein in solution. In molecular biology, the term typically refers to a cyclic, nitrogen-containing compound linked to deoxyribose in DNA and to ribose in RNA. See also nitrogenous base.

base analog A compound that substitutes for normal bases; e.g. 5-bromouracil.

base excision repair (BER) A repair pathway that removes a damaged base by a DNA glycosylase, then cleaves the 5′ side of the resulting abasic site by an AP endonuclease. The abasic sugar-phosphate and downstream bases are removed and the gap is filled by DNA polymerase and DNA ligase.

base pair (bp) A single pair of complementary nucleotides from opposite strands of the DNA double helix (A-T or G-C). The number of base pairs is used as a measure of the length of a double-stranded DNA molecule.

base pairing Weak bonding between purine and pyrimidine bases within nucleic acids. Normally adenine (A) pairs with thymine (T) by two hydrogen bonds in DNA, or with uracil (U) in RNA. Cytosine (C) pairs with guanine (G) by three hydrogen bonds in both DNA and RNA. Such base pairing is also referred to as "Watson–Crick" or "complementary" base pairing.

base stacking The tendency of the hydrophobic faces of the paired nitrogenous bases in DNA to stack on top of one another in such a way in solution as to minimize contact with the water molecules. A double-stranded DNA molecule thus has a hydrophobic core composed of stacked bases that exclude the maximum amount of water from the interior of the double helix.

basic helix-loop-helix (bHLH) motif An HLH motif coupled to a basic motif. When two bHLH proteins dimerize through their HLH motifs by forming a coiled coil, the basic motifs form a DNA-binding domain and are in position to interact with a specific region of DNA.

basic leucine zipper (bZIP) motif A leucine zipper motif coupled to a basic motif. When two bZIP proteins dimerize through their leucine zippers the basic motifs form a DNA-binding domain and are in the correct orientation to interact with a specific region of DNA.

B-DNA The standard Watson–Crick model of DNA. The predominant form of the DNA double helix within cells, favored at high relative humidity and in solution. A right-handed helix in which the bases are stacked almost exactly perpendicular to the main axis with 10.5 bases per turn.

beta (β)-galactosidase An enzyme that breaks the bond between the two constituent sugars of lactose.

beta (β)-lactamase An enzyme that breaks down ampicillin and renders a bacterium resistant to the antibiotic.

beta (β)-pleated sheet (or β-sheet) A fundamental unit of protein folding that involves extended amino acid chains in a protein that interact by hydrogen bonding. The chains are packed side by side to create an accordian-like appearance.

beta (β)-strand One of the polypeptide strands in a β-pleated sheet.

bidirectional DNA replication Replication that occurs in both directions at the same time from a common starting point, or origin of replication.

bioinformatics A discipline combining information technology with biotechnology; i.e. an area of computer science devoted to collecting, organizing, and analyzing DNA and protein sequences, and the data being generated by genomics and proteomics laboratories.

biolistics A method in which tiny metal pellets are coated with DNA and shot into cells.

BLAST A program that searches a database for DNA or protein sequence and displays how the query sequence lines up with the database sequences.

blue-white screening A method for distinguishing between bacteria transformed with recombinant and nonrecombinant vector DNA when the vector contains a multiple-cloning site that interrupts the *lacZ* gene. When grown on selective medium containing a colorless chromogenic compound X-gal, recombinant colonies are white. Nonrecombinant colonies are blue because β-galactosidase is produced and hydrolyzes X-gal to produced a blue-colored product.

bovine spongiform encephalopathy "Mad cow" disease; a prion infection in cattle.

branch migration Lateral motion of the branch of a Holliday junction during recombination.

BRCA1/BRCA2 The first breast cancer susceptibility genes to be identified. Mutated forms of these genes are believed to be responsible for about half the cases of inherited breast cancer, especially those that occur in younger women. Both are tumor suppressor genes involved in repair of DNA double-strand breaks by homologous recombination.

BRE TFIIB recognition element; a core promoter element in some RNA polymerase II gene promoters.

bromodomain A protein domain that binds specifically to acetylated lysine residues on other proteins, such as histones.

C

cancer Diseases in which abnormal cells divide and grow unchecked. Cancer can spread from its original site to other parts of the body and can also be fatal if not treated adequately.

candidate gene A gene, located in a chromosome region suspected of being involved in a disease, whose protein product suggests that it could be the disease gene in question.

canonical sequence See consensus sequence.

5′ cap A methylated guanosine bound through a 5′ → 5′ triphosphate linkage to the 5′ end of a eukaryotic mRNA, an hnRNA, or an snRNA.

CAP (catabolite activator protein) A protein which, together with cAMP, activates operons that are subject to catabolite repression. Also known as CRP.

cap-binding protein (CBP) A protein that associates with the 5′ cap on eukaryotic mRNA and allows the mRNA to bind to a ribosome. Also known as eIF4F.

carboxyl-terminal domain (CTD) The carboxyl-terminal region of the largest subunit of RNA polymerase II. Consists of dozens of repeats of a heptamer rich in serines and threonines and acts as a "landing pad" for RNA processing factors.

carboxyl terminus The end of a polypeptide with a free carboxyl group.

carcinogen Chemical compounds that are linked to sporadic human cancers, such as benzo(*a*)pyrene in cigarette smoke.

carcinoma Any of the various types of cancerous tumors that form in the epithelial tissue, the tissue forming the outer layer of the body surface and lining the digestive tract. Examples of this kind of cancer include breast, lung, and prostate cancer.

carrier An individual who possesses one copy of a mutant allele that causes disease only when two copies are present. Although carriers are not affected by the disease, two carriers can produce a child who has the disease.

cartilage hair hypoplasia A rare autosomal recessive form of dwarfism caused by mutations in the RNA component of RNase MRP; characterized by short limbs, short stature, fine sparse hair, impaired cellular immunity, anemia, and predisposition to several cancers.

catalytic center The active site of an enzyme, where catalysis takes place.

cathode Negative (black) electrode.

CCAAT-binding transcription factor (CTF) A transcription factor that binds to the CCAAT box.

CCAAT box An upstream DNA sequence motif, having the sequence CCAAT, found in many eukaryotic proximal promoters recognized by RNA polymerase II.

Cdc6 Cell division cycle 6. A component of the pre-replication complex that is essential for the initiation of DNA replication in eukaryotes.

cDNA (complementary DNA) A DNA copy of an RNA, made by reverse transcription.

cDNA library A collection of DNA sequences generated from mRNA expressed in a particular cell or tissue type at a given time. This type of library contains only protein-coding DNA sequences and does not include any noncoding DNA (e.g. regulatory elements or introns).

cell The basic structural unit of any living organism. It is a small, watery, compartment filled with chemicals and a complete copy of the organism's genome.

cell cycle The changes that take place in a cell in the period between its formation as one of the products of cell division and its own division and which in all cells includes DNA replication. In eukaryotic cells the cell cycle is divided into phases termed G1, S, G2, and M. G1 is the period immediately after mitosis and cell division when the newly formed cell is in the diploid state. S is the phase of DNA synthesis. This phase is followed by G2 when the cell is in a tetraploid state. Mitosis (M) follows to restore the diploid state, accompanied by cell division.

cell division splitting of a cell into two complete new cells, by binary fission in bacteria and other prokaryotes, and by division of both nucleus and cytoplasm in eukaryotic cells. The rapid cell divisions that occur during early embryogenesis are referred to as cleavages.

centimorgan A measure of genetic distance that tells how far apart two genes are and that yields a 1% recombination frequency between two markers. Generally 1 centimorgan equals about 1 million base pairs.

Central Dogma, the The hypothesis stated by Francis Crick in 1957 that "once information has passed into protein it cannot get out again." Describes the flow of information involving the genetic material.

centromere The constricted region near the center of a human chromosome. This is the region of the chromosome where the two sister chromatids are joined to one another. The centromere attaches chromosomes to the spindle fibers during mitosis and ensures that sister chromatids segregate correctly to daughter cells.

chaperones Proteins that bind to unfolded proteins or RNA molecules and help them fold properly.

Chargaff's rules The three regular relationships among the molar concentrations of the different bases in DNA: (1) The number of A residues in all DNA samples is equal to the number of T residues; i.e. [A] = [T], where the molar concentration of the base is denoted by the symbol for the base enclosed in square brackets. (2) The number of G residues equals that of C; i.e. [G] = [C]. (3) The amount of purine bases equals that of the pyrimidine bases: [A] + [G] = [T] + [C].

charging Coupling a tRNA with its cognate amino acid.

chemiluminescent The property of the product of an enzyme-catalyzed reaction that emits light and is therefore easily assayed.

chloramphenicol acetyl transferase (CAT) An enzyme encoded by a bacterial gene that adds acetyl groups to the antibiotic chloramphenicol. The CAT gene is frequently used as a reporter gene in eukaryotic transcription and translation experiments.

chloroplast DNA (cpDNA) The genetic material of the chloroplast, the organelle involved in photosynthesis in higher plants, some protozoans, and algae.

chromatids Copies of a chromosome produced during cell division.

chromatin Chromosomal DNA with its associated proteins.

chromatin assembly factor 1 (CAF-1) A protein that brings histones to the DNA replication fork via direction interaction with PCNA.

chromatin immunoprecipitation (ChIP) assay A method for purifying chromatin containing a protein of interest by immmunoprecipitating the chromatin with an antibody directed against that protein or against an epitope tag attached to the protein.

chromatin modification complexes Multiprotein complexes that modify histones post-translationally in ways that allow greater access or restrict access of other proteins to DNA.

chromatin remodeling ATP-dependent alterations of the structure of the nucleosomes that either move the nucleosome, remove them entirely, or modify their composition.

chromatin remodeling complexes Multiprotein complexes of the yeast SWI/SNF family (or their mammalian homologs BRG1 and BRM) and related families that contain ATP-dependent DNA unwinding activities.

chromatography A group of techniques for separating molecules based on their relative affinities for a mobile and a stationary phase; e.g. in ion-exchange chromatography, the charged resin is the stationary phase, and the buffer of increasing ionic strength is the mobile phase.

chromodomain A conserved region found in proteins involved in heterochromatin formation that binds to methylated histones.

chromogenic substrate A substrate that produces a colored product when acted on by an enzyme.

chromosome One of the threadlike "packages" of genes and other DNA in the nucleus of a cell. Different kinds of organisms have different numbers of chromosomes. Humans have 23 pairs of chromosomes, 46 in all. Each parent contributes one chromosome to each pair, so children get half of their chromosomes from their mothers and half from their fathers.

chronic wasting disease A type of transmissible spongiform encephalopathy that infects elk and deer.

***cis*-acting** A term that describes a DNA sequence element, such as an enhancer, a promoter, or an operator, that must be on the same chromosome in order to influence a gene's activity.

***cis*-splicing** Ordinary splicing in which the exons are on the same pre-mRNA.

cleavage factors I and II (CFI and CFII) RNA-binding proteins involved in cleavage of pre-mRNA at the polyadenylation site.

cleavage poly(A) specificity factor (CPSF) A protein that recognizes the AAUAAA part of the polyadenylation signal in a pre-mRNA and stimulates cleavage.

cleavage stimulation factor (CstF) A protein that recognizes the GU-rich part of the polyadenylation signal in a pre-mRNA and stimulates cleavage.

cloning The process of making copies of a specific piece of DNA, usually a gene. When molecular biologists speak of cloning, they generally do not mean the process of making genetically identical copies of an entire organism. Sometimes called molecular cloning to distinguish it from cloning by nuclear transfer.

cloning by nuclear transfer The production of genetically identical animals by nuclear transfer from adult somatic cells to unfertilized eggs.

coactivator A protein that increases transcriptional activity through protein–protein interactions without binding DNA directly. Operationally defined as a component required for activator-directed (activated) transcription, but dispensible for activator-independent (basal) transcription; includes chromatin remodeling and modification complexes.

Cockayne syndrome A disease characterized by photosensitivity, cataracts, deafness, and severe mental retardation; linked to defects in transcription-coupled repair.

codon A 3-base sequence in mRNA which specifies a single amino acid to be added into a polypeptide chain or causes termination of translation.

codon bias The frequency with which different codons are used varies significantly between different organisms and between proteins expressed at high or low levels within the same organism.

cofactor Any nonprotein substance required by a protein (or ribozyme) for biological activity such as prosthetic groups and, especially in enzyme-catalyzed reactions, other compounds which are not consumed in the process and are found unchanged at the end of the reaction.

cognate An antigen recognized by its corresponding antibody, a tRNA recognized by a particular aminoacyl-tRNA synthetase, etc.

coiled coil A protein motif in which two α-helices (coils) wind around each other. Coiled coils may form within one protein or between two separate proteins.

coimmunoprecipitation assay A technique used to analyze protein–protein interactions *in vivo*. Proteins from cell extracts are reacted with a specific antibody or antiserum against one of the proteins, then immunoprecipitated by centrifugation. The precipitated proteins are usually detected by Western blot analysis.

colony hybridization A procedure for selecting a baterial clone (colony) containing a gene of interest. DNAs from a large number of clones are tested with a labeled probe that hybridizes to the gene of interest.

competent cells Host bacterial cells chemically treated to make their membranes leaky; used for transformation with foreign DNA.

complementary base pairing See base pairing.

complementation group Various mutations which do not form a wild-type (normal) phenotype after crossing.

concatemers DNAs of multiple genome length.

congenital Any trait or condition that exists from birth.

consensus sequence The "ideal" form of a DNA sequence found in slightly different forms in different organisms, but which is believed to have the same function. The consensus sequence gives for each position the nucleotide most often found. Sometimes called a canonical sequence.

constant region (C) The region of an antibody (immunoglobulin) that is basically the same from one antibody to the next.

constitutive gene expression A gene is always turned on.

contig A chromosome map showing the locations of those regions of a chromosome where contiguous DNA segments overlap. Contig maps are important because they allow study of a complete, and often large, segment of the genome, by examining a series of overlapping clones which then provide an unbroken series of information about that region.

core histone One of four highly conserved histones (H2A, H2B, H3, H4) that are present in the nucleosome as an octamer composed of a dimer of histones H2A and H2B at each end and a tetramer of histones H3 and H4 in the center, around which 146 bp of genomic DNA are wound.

core histone octamer (or core octamer) A particle composed of a dimer of histones H2A and H2B at each end and a tetramer of histones H3 and H4 in the center, around which 146 bp of genomic DNA are wound. The core octamer lacks the linker histone and linker DNA that would be present in the complete nucleosome.

core promoter An approximately 60 bp DNA sequence overlapping the transcription start site (+1) that serves as the recognition site for RNA polymerase II and general transcription factors.

corepressors Proteins that decrease transcriptional activity through protein–protein interactions without binding DNA directly; include chromatin remodeling and modification complexes.

cos The cohesive ends of the linear lambda (λ) phage DNA.

cosmid A vector designed for cloning large DNA inserts. It contains the *cos* sites of lambda (λ) phage DNA, so it can be packaged into λ heads, and a plasmid origin of replication so it can replicate in bacteria as a plasmid.

cotranscriptonal cleavage (CoTC) The cleavage of a growing transcript downstream of the polyadenylation site that is part of the transcription termination process.

counts per minute (cpm) The average number of scintillations detected per minute by a liquid scintillation counter. Generally, this is disintegrations per minute (dpm) times the efficiency of the counter.

coupling The triggering of a large conformational change in both an enzyme and its substrate upon binding. Also known as induced fit.

CpG island A region of DNA containing many unmethylated CpG sequences that is usually associated with active genes.

Creutzfeldt–Jakob disease (CJD) A sporadic form of transmissible spongiform encephalopathy. In this disease PrP^C misfolds spontaneously and then by "autoinfection" generates more prions.

cross-over Physical exchange between DNAs that occurs during recombination.
cruciform A cross-shaped paired stem-loop formation that can form in regions of inverted repeats in DNA.
cryoelectron microscopy A protein or other macromolecule sample is snap-frozen at extremely low temperature in liquid helium so that it is fully hydrated without water crystals; it is then examined in a cryoelectron microscope.
cyclic AMP (cAMP) An adenine nucleotide with a cyclic phosphodiester linkage between the 3′ and 5′ carbons.
cyclin-dependent kinases (CDKs) A family of kinases whose activity is regulated by both phosphorylation and interaction with regulatory protein called cyclins. The CDKs are important in regulating progression through the cell cycle.
cyclins A family of proteins that accumulate gradually during interphase and are abruptly destroyed during mitosis; regulatory subunits of the cyclin-dependent kinases.
cyclobutane-pyrimidine dimers See pyrimidine dimers.
cystic fibrosis A hereditary disease whose symptoms usually appear shortly after birth, caused by mutations in the cystic fibrosis transmembrane regulator (CFTR) protein. They include faulty digestion, breathing difficulties and respiratory infections due to mucus accumulation, and excessive loss of salt in sweat. In the past, cystic fibrosis was almost always fatal in childhood, but treatment is now so improved that patients commonly live to their twenties and beyond.
cytidine A nucleoside containing the base cytosine.
cytogenetic map The visual appearance of a chromosome when stained and examined under a microscope. Particularly important are visually distinct regions, called light and dark bands, which give each of the chromosomes a unique appearance and allows characterization of chromosomal alterations. This feature allows a person's chromosomes to be studied in a clinical test known as a karyotype.
cytoplasmic polyadenylation element (CPE) A sequence in the 3′ UTR of an mRNA (consensus, UUUUUAU) that is involved in cytoplasmic polyadenylation.
cytosine One of the four bases in DNA that make up the letters ATGC, cytosine is the "C." The others are adenine, guanine, and thymine. Cytosine always pairs with guanine.
cytosine deamination The removal of an amino group (NH_2) from a cytosine in DNA, in which the amino group is replaced by a carbonyl group (C–O). This converts cytosine to uracil.

D

D loop (mtDNA) Displacement loop. A region of 500–600 nucleotides where replication begins in mitochondrial DNA.
Dalton The term used to describe the molecular weight of proteins. 1 Dalton is equivalent to 1 atomic mass unit.
DEAD box protein A member of a family of proteins containing the sequence Asp-Glu-Ala-Asp and having RNA helicase activity.
deadenylation The removal of AMP residues from poly(A) in the cytoplasm.
deamination The conversion of one base to another by replacement of the amino group with, for example, an oxygen; e.g. the conversion of cytosine in DNA to uracil, a base that should only be present in RNA, and the conversion of 5-methylcytosine to thymine.
decoding Interactions between codons and anticodons on the ribosome that lead to binding of the correct aminoacyl-tRNA.
degenerate code In the genetic code more than one codon can stand for a single amino acid.
deletion A mutation involving a loss of one or more base pairs of DNA from a chromosome. Deletion of a gene or part of a gene can lead to a disease or abnormality.
denaturation of DNA The unwinding and separation of the two strands of double-stranded DNA; also referred to as melting.
denaturation of protein Disruption of the three-dimensional structure of the protein without breaking any covalent bonds.

density gradient centrifugation A solution of cesium chloride (CsCl) containing a DNA sample is spun in an ultracentrifuge at high speed for several hours. An equilibrium between centrifugal force and diffusion occurs, such that a gradient forms with a high concentration of CsCl at the bottom of the tube and a low concentration at the top. DNA forms a band in the tube at the point where the density is the same as that of CsCl.

deoxynucleoside triphosphates (dNTPs) The building blocks of DNA: dATP, dCTP, dGTP, dTTP.

deoxyribonucleic acid (DNA) The chemical inside the nucleus of a cell that carries the genetic instructions for making living organisms.

deoxyribose The pentose (five-carbon) sugar present in a nucleotide subunit of DNA.

diabetes mellitus Two types of a highly variable disorder in which abnormalities in the ability to make and/or use the hormone insulin interfere with the process of turning dietary carbohydrates into glucose, the body's fuel. Type I is known as insulin-dependent diabetes mellitus, and type II is known as non-insulin-dependent diabetes mellitus.

Dicer The member of the RNAse III family that cleaves double-stranded RNA into pieces about 21 bp long (called siRNAs) during the RNAi process.

dideoxynucleotide A nucleotide, lacking the oxygens at both the 2′ and 3′ positions (hence "dideoxy"), used to terminate DNA chain elongation in DNA sequencing.

digoxygenin A plant steroid isolated from foxglove (*Digitalis*) used in nonradioactive labeling techniques.

dimer (protein) A complex of two polypeptides. These can be the same (homodimer) or different (heterodimer).

dimerization domain The region of a protein that interacts with another protein to form a dimer.

diploid The number of chromosomes in most cells except the gametes. In humans, the diploid number is 46.

directional cloning Insertion of a foreign DNA into two different restriction sites of a vector, such that the orientation of the insert can be predetermined.

disintegrations per minute (dpm) The average number of radioactive emissions produced each minute by a sample.

DNA-binding domain The part of a DNA-binding protein that makes specific contacts with a target site on the DNA.

DNA glycosylase An enzyme that breaks the glycosidic bond between a damaged base and its sugar to form an abasic site.

DNA ligase An enzyme that joins two double-stranded DNAs end to end.

DNA looping The process by which DNA-binding proteins can interact simultaneously with one another and with remote sites on DNA, by causing the DNA in between the sites to form a loop.

DNA melting The unwinding and separation of double-stranded DNA; also referred to as denaturation.

DNA methyltransferase 1 (DNMT1) An enzyme that adds methyl groups to hemimethylated DNA substrates during DNA replication.

DNA methyltransferase 3 (DNMT3) Together with its cofactor DNMT3L, this enzyme is required to establish genomic imprinting *de novo* in sperm and oocytes.

DNA microarray A way of studying how large numbers of genes interact with each other and how a cell's regulatory networks control vast batteries of genes simultaneously. The method uses a robot to precisely apply tiny droplets containing functional DNA to glass slides. Researchers then synthesize fluorescently labeled cDNA from mRNA isolated from the cells they are studying. The labeled probes are allowed to bind to cDNA strands on the slides. The slides are put into a scanning microscope that can measure the brightness of each fluorescent dot; brightness reveals how much of a specific DNA fragment is present, an indicator of how active the gene was in the original cell type.

DNA photolyase The enzyme that catalyzes photoreactivation by breaking pyrimidine dimers.

DNA polymerase An enzyme that synthesizes DNA by linking together deoxynucleoside monophosphates (dNMPs) in the order dictated by the complementary sequence of nucleotides in a template DNA strand. There are two major classes of DNA polymerases: those involved in genome replication (replicative polymerases) and those involved in DNA repair.

DNA polymerase I An *E. coli* DNA polymerase that plays a role in primer removal and gap filling between Okazaki fragments and in the nucleotide excision repair pathway.

DNA polymerase II An *E. coli* DNA polymerase that is involved in DNA repair mechanisms.

DNA polymerase III The *E. coli* replicative polymerase that catalyzes genome replication.

DNA polymerase IV The *E. coli* DNA polymerase that mediates translesion DNA synthesis. Also known as DinB (encoded by the *dinB* gene).

DNA polymerase V The *E. coli* DNA polymerase that mediates translesion DNA synthesis. Also known as the $UmuD'_2C$ complex (encoded by the *umuDC* operon).

DNA polymerase α The eukaryotic DNA polymerase involved in priming DNA synthesis during replication and repair.

DNA polymerase β The eukaryotic DNA polymerase involved in high fidelity base excision and double-strand break repair.

DNA polymerase σ The eukaryotic DNA polymerase involved in high fidelity replication of the leading strand during replication and repair.

DNA polymerase ε The eukaryotic DNA polymerase involved in high fidelity replication of the lagging strand during replication and repair.

DNA polymerase γ The eukaryotic DNA polymerase involved in high fidelity replication and repair of mitochondrial DNA.

DNA polymerase η The eukaryotic DNA polymerase involved in high fidelity translesion DNA synthesis (relatively accurate replication past thymine-thymine dimers).

DNA polymerase ι A eukaryotic DNA polymerase involved in error-prone translesion DNA synthesis (required during meiosis).

DNA polymerase κ A eukaryotic DNA polymerase involved in error-prone translesion DNA synthesis and in double-strand break repair by nonhomologous end-joining.

DNA polymerase λ A eukaryotic DNA polymerase involved in error-prone translesion DNA synthesis.

DNA polymerase μ A eukaryotic DNA polymerase involved in error-prone double-stranded DNA break repair by nonhomologous end-joining.

DNA polymerase υ A eukaryotic DNA polymerase thought to be involved in error-prone DNA cross-link repair.

DNA polymerase θ A eukaryotic DNA polymerase involved in error-prone repair of DNA interstrand cross-links.

DNA polymerase ξ A eukaryotic DNA polymerase involved in error-prone translesion DNA synthesis (thymine dimer bypass).

DNA replication The process by which the DNA double helix unwinds and makes an exact copy of itself.

DNA sequencing Determining the exact order of the base pairs in a segment of DNA.

DNA transposons Transposable elements with a DNA intermediate during transposition.

DNA typing The use of highly variable regions of DNA to identify particular individuals (also called DNA fingerprinting).

DNA-PK (DNA protein kinase) The key enzyme in eukaryotic double-strand break repair by nonhomologous end-joining.

$DNA\text{-}PK_{CS}$ The catalytic subunit of DNA-PK.

DNase Deoxyribonuclease, an enzyme that degrades DNA.

DNase footprinting A method of detecting the binding site for a protein on DNA by observing the DNA region this protein protects from cleavage by DNase.

DNase hypersensitive sites Regions of chromatin that are about a hundred times more susceptible to cleavage by DNase than the rest of the chromatin. Active genes tend to be DNase-sensitive.

dominant An allele or trait that expresses its phenotype when heterozygous with a recessive allele; for example, a disease, even though the patient's genome possesses only one copy. With a dominant gene, the chance of passing on the gene (and therefore the disease) to children is 50–50 in each pregnancy.

dominant negative mutation A mutation that yields a protein that is not only inactive but disrupts the activity of wild-type protein made in the same cell, by forming a heterodimer, for example.

double helix The structural arrangement of DNA, which looks something like a very long ladder twisted into a helix, or coil. The sides of the "ladder" are formed by a backbone of sugar and phosphate molecules, and the "rungs" consist of nucleotide bases joined weakly in the middle by hydrogen bonds and stabilized by base stacking interactions.

double-strand break Damage to the DNA backbone in which nicks are made in both strands of the DNA double helix; induced by agents such as reactive oxygen species, ionizing radiation, and chemicals that generate reactive oxygen species (free radicals).

double-strand break repair Repair of double-strand breaks in DNA either by homologous recombination or nonhomologous end-joining.

downstream DNA sequence after the start of transcription (+1) in the direction of the 3′ end; numbered with positive numbers (e.g. +35).

DPE Downstream promoter element; an RNA polymerase II core promoter element in some gene promoters.

Drosha The endonuclease the cleaves a primary miRNA transcript in the nucleus to form a shorter hairpin pre-miRNA.

Ds *Disassociation*, a defective transposable element found in maize, which relies on an *Ac* element for transposition.

dsDNA Double-stranded DNA.

dsRNA Double-stranded RNA.

duplex DNA Double-stranded DNA.

duplication A particular kind of mutation: production of one or more copies of any piece of DNA, including a gene or even an entire chromosome.

duplicative transposition Transposition in which a DNA sequence replicates, so one copy remains in the original location as another copy moves to the new site. Also called replicative transposition.

dyskeratosis congenita A rare inherited, premature aging disease linked to loss of telomerase activity either through mutations in the dyskerin gene or the telomerase RNA gene.

dyskerin A pseudouridine synthase that binds to many snoRNAs and is proposed to play a role in ribosomal RNA base modification.

E

E site The exit site to which deacylated tRNAs bind on their way out of the ribosome.

EF-Tu The prokaryotic translation elongation factor that, along with GTP, carries aminoacyl-tRNAs (except fMet-tRNA$_f^{Met}$ to the ribosome A site).

eIFs The set of eukaryotic translation initiation factors.

electrophoretic mobility shift assay (EMSA) An assay for DNA-protein binding. A short labeled DNA is mixed with a protein and electrophoresed. If the DNA binds the protein, its electrophoretic mobility is greatly decreased.

electroporation The use of a strong electric current to introduce DNA into cells.

Elongator A protein complex that facilitates transcript elongation by RNA polymerase II.

empty vector A term sometimes used to refer to a cloning vector without an insert.

encode To contain the information for making an RNA or polypeptide. A gene can encode a functional RNA or a polypeptide.

end-filling Filling in the recessed 3′-end of a double-stranded DNA using deoxynucleoside triphosphates (dNTPs) and a DNA polymerase; often used to label the 3′-end of a DNA strand.

endonuclease An enzyme that cleaves the phosphodiester bond joining adjacent nucleotides at an internal site within a polynucleotide.

endoplasmic reticulum A network of membranes in the cell on which proteins destined for secretion from the cell are synthesized.

enhanceosome The complex formed by enhancers coupled to their activators, involving protein–protein interactions and DNA looping.

enhancer A DNA regulatory element that increases gene promoter activity. Enhancers are usually 700–1000 bp or more away from the start of transcription and can be downstream, upstream, or within an intron, and can function in either orientation relative to the promoter.

enzyme A molecule that catalyzes or increases the rate, or velocity, of a biochemical reaction without itself being changed in the overall process. Most enzymes are globular proteins, but some RNA molecules have the properties of enzymes (see ribozyme).

enzyme-linked immunosorbent assay (ELISA) An immunoassay used to quantify antigen–antibody reactions by combining the specificity of antibodies with the sensitivity of simple enzyme assays.

epigenetics The study of mitotically and/or meiotically heritable changes in gene expression without changes in gene sequence; e.g. through changes in the pattern of DNA methylation and histone modifications.

epitope A region of an antigen to which an antibody can bind.

epitope tagging Using genetic means to attach a small group of amino acids (an epitope tag) to a protein. This enables the protein to be purified readily by immunoprecipitation with the antibody that recognizes the epitope tag.

eRF1 The eukaryotic release factor that recognizes all three stop codons and releases the completed polypeptide from the ribosome.

eRF3 The eukaryotic release factor with ribosome-dependent GTPase activity that interacts with eRF1 in releasing polypeptides from the ribosome.

***Escherichia coli* (*E. coli*)** An intestinal bacterium; the conventional host bacterium used in molecular biology research.

ethidium bromide A fluorescent dye that intercalates between the bases of nucleic acids.

euchromatin Chromatin that is open and accessible to RNA polymerase, and at least potentially active. These regions stain lightly and are thought to contain most of the genes.

eukaryote An organism whose cells have nuclei.

exons Sequences that are ligated together after excision during RNA splicing; typically the expressed sequences of an mRNA. The term is used to refer both to the corresponding sequence in the DNA and the RNA.

exonuclease An enzyme that removes dNMPs or NMPs from the end of a polynucleotide chain inward by breaking the terminal phosphodiester bond.

exosome A protein complex that degrades RNAs. Different exosomes are found in the nucleus and the cytoplasm.

exportin A soluble receptor that mediates nuclear export of a NES-bearing cargo.

expressed sequence tag (EST) Partial cDNA sequences generated by amplifying cellular mRNA by RT-PCR.

expression platform The domain of a riboswitch that converts metabolite binding into changes in gene expression via changes in RNA folding. The expression platform has the potential to form alternative antiterminator and terminator hairpins (see transcriptional attenuation).

expression site A locus on a chromosome where a gene can be moved to be expressed efficiently. For example, the expression site for VSG genes in trypanosomes is at the end (telomere) of a chromosome.

expression vector A cloning vector that allows expression of a cloned gene in bacterial or eukaryotic cells.

extrachromosomal A molecule of DNA, such as a plasmid, that remains separate from the host cell chromosome.

F

F plasmid An *E. coli* plasmid that allows conjugation between bacterial cells.

F_1 The progeny (i.e. the first filial generation) of a cross between two parental types that differ at one of more genes.

F_2 The progeny (i.e. the second filial generation) of a cross between two F_1 individuals or the progeny of a self-fertilized F_1.

FACT A protein that facilitates transcription elongation by eukaryotic RNA polymerases through nucleosomes arrays.

familial dysautonomia An autosomal recessive disorder of the sensory and autonomic nervous system caused by a splice site mutation in a subunit of the Elongator complex.

fatal familial insomnia An inherited autosomal dominant form of transmissible spongiform encephalopathy which involves a mutated PrP^C gene with a greater tendency to spontaneously misfold to the prion form.

fibroblasts A type of cell found just underneath the surface of the skin. Fibroblasts are part of the support structure for tissues and organs.

fibrous protein A protein with a long filamentous or "rod-like" structure; major structural components of cells.

five (5)-prime (5′) The ends of a DNA or RNA chain are chemically distinct and are designated by the symbols 3′ and 5′. The symbol 5′ refers to the carbon in the ribose or deoxyribose sugar ring to which a phosphate (PO_4) functional group is attached.

FlapEndoNuclease I (FEN-1) An nuclease that acts in association with PCNA to degrade the RNA primer during eukaryotic DNA replication.

fluorescence *in situ* hybridization (FISH) A process which hybridizes a fluorescent probe to whole chromosomes to determine the location of a gene or other DNA sequence within a chromosome. This technique is useful for identifying chromosomal abnormalities and gene mapping. Also used to determine the precise localization of RNA transcripts within a cell.

fluorescence resonance energy transfer (FRET) A method for detecting protein–protein interactions *in situ* by a nonradiative process whereby energy from an excited donor fluorophore is transferred to an acceptor fluorophore that is within approximately 10 nm of the excited fluorophore.

fluorophore A group of atoms in a molecule responsible for absorbing light energy and producing the color of the compound; i.e. the region of a molecule exhibiting fluorescence.

fragile X syndrome After Down syndrome, the second most frequent genetic cause of mental retardation. The disorder is one of a group of diseases that results from trinucleotide repeat expansion. In fragile X, the repeating triplet is CGG in a gene called FMR1 on the X chromosome. There are ordinarily fewer than 55 copies of the repeat. When the number of repeats exceeds 200, the expansion is called a "full mutation" and expression of the FMR1 gene is inhibited.

frameshift mutation An insertion or deletion of one or two bases in the coding reigon of a gene, which changes the reading frame of the corresponding mRNA.

free radicals Very reactive chemical substances with an unpaired electron that can cause DNA damage (e.g. reactive oxygen species such as O_2^- and H_2O_2).

Friedreich's ataxia A rare inherited neurological disease characterized by the progressive loss of voluntary muscular coordination and heart enlargement. The disorder is caused by a GAA trinucleotide repeat expansion in the first intron of the frataxin gene.

fusion protein A protein resulting from the expression of a recombinant DNA containing two open reading frames (ORFs) fused together. One or both of the ORFs can be incomplete.

G

G + C content The base composition of DNA, defined as the "percent G + C," differs among species but is constant in all cells of an organism within a species. The G + C content can vary from 22 to 73%, depending on the organism.

G protein A protein that is activated by binding to GTP and inactivated by hydrolysis of the bound GTP to GDP by an inherent GTPase activity.

gamete A haploid sex cell.

gametogenesis The process by which gametes are produced.

gel electrophoresis The process in which molecules (such as proteins, DNA, or RNA fragments) can be separated according to size and electrical charge by applying an electric current to them. The current

forces the molecules through pores in a thin layer agarose or polyacrylamide. The gel can be made so that its pores are just the right dimensions for separating molecules within a specific range of sizes and shapes. Smaller fragments usually travel further than large ones.

gel filtration A column chromatography method for separating molecules according to their sizes. Small molecules enter the beads of the gel and so take longer to move through the column than larger molecules, which cannot enter the beads.

gene The basic unit of heredity. A complete chromosomal segment responsible for making a functional product. Three essential features of a gene are: the expression of a product, the requirement that it be functional, and the inclusion of both coding and regulatory regions.

gene amplification An increase in the number of copies of any particular piece of DNA. A tumor cell often amplifies or copies DNA segments as a result of cell signals and sometimes environmental events.

gene cloning Generating many copies of a gene by inserting into an organism, such as a bacterium, where it can replicate along with the host.

gene conversion The conversion of one gene's sequence to that of another by homologous recombination.

gene expression The process by which gene products (RNA transcripts or protein) are made from the instructions encoded in DNA.

gene mapping Determining the relative positions of genes on a chromosome and the distance between them.

gene targeting The replacement or mutation of a particular gene that provides a means for creating strains of "knockout" organisms with mutations in virtually any genes.

gene therapy An evolving technique used to treat inherited diseases. The medical procedure involves either replacing, manipulating, or supplementing nonfunctional genes with healthy genes. Also referred to as somatic cell gene therapy.

gene transfer Insertion of unrelated DNA into the cells of an organism. Most techniques involve the use of a vector, such as a specially modified virus that can take the gene along when it enters the cell.

general transcription factors A set of five RNA polymerase II transcription factors, denoted TFIIB, TFIID, TFIIE, and TFIIH that are responsible for promoter recognition and for unwinding the promoter DNA. RNA polymerase II is absolutely dependent on these auxiliary transcription factors for the initiation of transcription.

general transcription machinery General, but diverse, components of large multiprotein RNA polymerase machines required for promoter recognition and the catalysis of RNA synthesis.

genetic code (ATGC) The set of 64 codons and the amino acids (or termination) for which they stand. Each gene's code combines the four bases in various ways to spell out three-letter "words" that specify which amino acid is needed at every step in making a protein.

genetic counseling A short-term educational counseling process for individuals and families who have a genetic disease or who are at risk for such a disease. Genetic counseling provides patients with information about their condition and helps them make informed decisions.

genetic linkage The physical association of genes on the same chromosome.

genetic map Also known as a linkage map; a chromosome map of a species that shows the position of its known genes and/or markers relative to each other, rather than as specific physical points on each chromosome.

genetic marker A segment of DNA with an identifiable physical location on a chromosome and whose inheritance can be followed. A marker can be a gene, or it can be some section of DNA with no known function. Because DNA segments that lie near each other on a chromosome tend to be inherited together, markers are often used as indirect ways of tracking the inheritance pattern of a gene that has not yet been identified, but whose approximate location is known.

genetic screening Testing a population group to identify a subset of individuals at high risk for having or transmitting a specific genetic disorder.

genome One complete set of genetic information contained within an organism or a cell; i.e. the single, circular chromosome of a bacterium is its genome. The term is often used interchangebly with the terms genomic DNA, chromosomal DNA, or nuclear DNA (to distinguish it from organelle or plasmid DNA).

genomic imprinting The nonrandom expression of a gene from only one of the two parental chromosomes, regulated by differential methylation.

genomic library A set of clones containing DNA fragments derived directly from a genome, rather than from mRNA.

genomics The comprehensive study of whole sets of genes and their interactions rather than single genes or proteins.

genotype The genetic identity or allelic constitution of an individual that does not show as outward characteristics. The genotypes at locus B in a diploid individual may be BB, Bb, or bb.

germline Inherited material that comes from the eggs or sperm and is passed on to offspring.

Gerstmann–Sträussler–Scheinker syndrome An inherited autosomal dominant form of transmissible spongiform encephalopathy which involves a mutated PrP^C gene with a greater tendency to spontaneously misfold to the prion form.

global genome repair Nucleotide excision repair pathway responsible for recognizing DNA damage in the whole genome.

globular protein A protein which folds into a roughly spherical shape. Most enzymes are globular proteins.

glucocorticoid receptor A nuclear receptor that mediates a highly abbreviated signal transduction pathway by activating target genes in response to ligand, leading to many diverse cellular responses, ranging from increases in blood sugar to anti-inflammatory action.

glucose A simple, six-carbon sugar used by many forms of life as an energy source.

glycoprotein A protein with a carbohydrate group attached post-translationally.

Golgi apparatus A membranous organelle that packages newly synthesized proteins for export from the cell.

green fluorescent protein A naturally fluorescent protein from the jelly fish *Aequorea victoria* commonly used as a reporter gene.

Greig cephalopolysyndactyly syndrome A very rare autosomal dominant disorder that is characterized by physical abnormalities affecting the fingers, toes, head, and face; caused by mutations in the GLI3 gene which is part of the Sonic hedgehog signal transduction pathway.

group I introns Self-splicing introns in which splicing is initiated by a free guanosine or guanosine nucleotide.

group II introns Self-splicing introns in which splicing is initiated by formation of a lariat-shaped intermediate.

guanine (G) One of the four bases in DNA. Guanine always pairs with cytosine.

guanosine A nucleoside containing the base guanine.

guide RNAs (editing) Small RNAs that bind to regions of an mRNA precursor in trypanosomes and serve as templates for editing a region upstream.

gyrase A topoisomerase that introduces negative superhelical turns into DNA and relaxes the positive superhelical strain created by unwinding the *E. coli* DNA during replication.

H

H strand Heavy strand of mitochondrial DNA, based on the relative buoyant density by density gradient centrifigution.

H5N1 virus An avian influenza virus strain that poses the threat of an influenza pandemic if it gains the ability to transmit readily from human to human.

hairpin A structure that resembles a bobby pin (hairpin), formed by intramolecular base pairing in an inverted repeat of a single-stranded DNA or RNA.

hairpin ribozyme A small ribozyme found in some virusoids that has a hairpin secondary structure.

half-life The time it takes for half of a population of molecules to disappear (turn-over).

hammerhead ribozyme A small ribozyme so called for its secondary structure of three helices in a T-shape; the most frequently found catalytic motif in plant pathogenic RNAs such as viroids.

haploid The number of chromosomes in a sperm or egg cell (gamete), half the diploid number.
haploinsufficiency A situation in which the protein produced by a single copy of an otherwise normal gene is not sufficient to assure normal function.
haplotype A cluster of alleles on a single chromosome.
Hayflick Limit The point at which cultured human and animal cells normally stop dividing because they have a limited capacity for replication.
heat shock proteins Proteins whose expression is significantly increased in response to environmental stress, including heat; promote protein folding and aid in the destruction of misfolded proteins.
helicase An enzyme that unwinds a polynucleotide double helix (either DNA or RNA).
helix-turn-helix (HTH) A structural motif in certain DNA-binding proteins, especially those from prokaryotes, that fits into the DNA major groove and gives the protein its DNA-binding capacity and specificity. A specialized type of HTH in eukaryotes is the homeodomain, present in DNA-binding proteins important for development.
helper virus A virus that supplies the functions lacking in a defective virus, allowing the latter to replicate.
hematopoietic stem cell An unspecialized precursor cell that will develop into a mature blood cell.
hemizygous When a new transgenic locus is present in only one member of a particular chromosome pair.
hemoglobin The oxygen-carrying protein in red blood cells.
hemophilia A sex-linked inherited bleeding disorder that generally only affects males. The disorder is characterized by a tendency to bleed spontaneously or at the slightest injury because of the lack of certain clotting factors in the blood.
hepatitis delta virus (HDV) RNA ribozyme A small ribozyme that is a viroid-like satellite virus of the human hepatitis B virus, which when present causes an exceptionally strong type of hepatitis in infected patients.
hereditary nonpolyposis colon cancer (HNPCC) A common form of human hereditary colon cancer, caused by defects in mismatch repair genes.
heredity The transmission of characteristics from parent to offspring by means of genetics.
heterochromatin Chromatin that is condensed and inactive.
heteroduplex A double-stranded polynucleotide whose two strands are not completely complementary.
heterogeneous nuclear RNA (hnRNA) A class of large, heterogeneous-size RNAs found in the nucleus, including unspliced mRNA precursors.
heterologous probe A probe that is similar to, but not exactly the same as, the nucleic acid sequence of interest.
heteroplasmy A condition in which both mutant and normal mitochondrial DNA coexist within the same cell.
heterozygous Possessing two different forms of a particular gene, one inherited from each parent.
high-throughput analysis Methods for whole genome and proteome analysis on a large scale.
highly conserved sequence A DNA sequence that is very similar in several different kinds of organisms. These cross-species similarities are regarded as evidence that a specific gene performs some basic function essential to many forms of life and that evolution has therefore conserved its structure by permitting few mutations to accumulate in it.
Hispanic thalessemia A rare disorder characterized by partial or complete deletions of the locus control region of the β-globin gene cluster.
histone Any one of a set of small, positively charged, basic proteins (H1, H2A, H2B, H3, H4), rich in arginine and lysine, bound to DNA in eukaryote chromosomes to form nucleosomes (see also linker histone and core histone).
histone acetyltransferase (HAT) An enzyme that transfers acetyl groups from acetyl CoA to core histones.
histone code A hypothesis proposing that covalent post-translational modifications of histones are read by the cell and lead to a complex, combinatorial transcriptional output.
histone deacetylase (HDAC) An enzyme that removes acetyl groups from core histones.

histone demethylase An enzyme that removes methyl groups from core histones (e.g. lysine specific demethylase 1, LSD1).

histone methyltransferase (HMT) A chromodomain-containing enzyme that transfers methyl groups to core histones.

hit and run model A model proposing that transcriptional activation reflects the probability that all components required for activation will meet at a certain chromatin site (the "hit"), i.e. a transcription complex is assembled in a stochastic fashion from freely diffusible proteins, and that the binding of these proteins is transient (the "run").

HMG protein An architectural nuclear protein with a high electrophoretic mobility (high-mobility group). Some of the HMG proteins are involved in transcriptional regulation.

Holliday junction The branched DNA structure formed by the first strand exchange during recombination.

holoenzyme A complete, fully functional enzyme molecule, consisting of a protein portion (apoenzyme), a nonprotein prosthetic group(s), or any other regulatory or accessory protein subunit if appropriate.

homeobox A sequence of about 180 bp that encodes a homeodomain; found in homeotic genes and other development-controlling genes in eukaryotes.

homeodomain A 60-amino acid domain of a DNA-binding protein with a type of helix-turn-helix domain that allows the protein to bind tightly to a specific DNA region.

homeotic gene A gene in which a mutation causes the transformation of one body part to another.

homologous chromosomes Chromosomes that are identical in size, shape, banding pattern and, except for allelic differences, genetic composition.

homologous probe A probe that is exactly complementary to the nucleic acid sequence of interest.

homologous recombination (DNA repair) Repair of double-strand breaks by retrieving genetic information from an undamaged homologous chromosome. In cases where the two chromosomes are not exactly homologous, gene conversion may take place.

homologous recombination (meiosis) The exchange of pieces of DNA during the formation of eggs and sperm. Recombination allows the chromosomes to shuffle their genetic material, increasing the potential of genetic diversity. Homologous recombination is also known as crossing-over.

homologs Genes that have evolved from a common ancestral gene. Includes orthologs and paralogs.

homoplasmy The normal condition in which all the mitochondrial DNA (mtDNA) within the cells of an individual are identical.

homozygous Possessing two identical forms of a particular gene, one inherited from each parent.

Hoogsteen base pairs A-T and G-C base pairs that have altered patterns of hydrogen bonding compared with Watson–Crick base pairs.

hopping (protein) A protein moves between binding sites on DNA through three-dimensional spaces, by dissociating from its initial site before reassociating elsewhere in the same DNA chain; the main mode of translocation over long distances.

hormone response elements Enhancers that respond to nuclear receptors bound to their ligands; e.g. the glucocorticoid response element (GRE).

hotspots Sites on chromosomes where mutations arise at a higher frequency than other regions of the DNA.

housekeeping genes Genes that code for proteins needed for basic processes in all kinds of cells.

human artificial chromosome (HAC) A vector used to transfer or express large fragments of human DNA. HACs behave and are constructed like human chromosomes.

human endogenous retrovirus (HERV) Transposition-defective LTR-containing retrotransposons in human cells.

Human Genome Project An international research project to map each human gene and to completely sequence human DNA.

human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS) AIDS was first reported in 1981 in the USA and has since become a major epidemic, killing nearly 12 million

people and infecting more than 30 million others worldwide. The disease is caused by HIV, a retrovirus that destroys the body's ability to fight infections and certain cancers.

Huntington's disease A degenerative brain disorder that usually appears in mid-life, caused by a trinucleotide repeat expansion. Its symptoms, which include involuntary movement of the face and limbs, mood swings, and forgetfulness, get worse as the disease progresses. It is generally fatal within 20 years.

Hutchinson–Gilford progeria syndrome A premature aging syndrome caused by a splicing mutation in the lamin A gene.

hybrid dysgenesis A phenomenon observed in *Drosophila* in which the hybrid offspring of two certain parental strains suffer so much chromosomal damage that they are sterile, or dysgenic.

hybridization (or DNA or RNA) Complementary base pairing of two single strands of DNA or RNA from two different sources.

hydrogen bonds Very weak bonds that involve the sharing of a hydrogen between two electronegative atoms, such as an oxygen and nitrogen.

hydrophilic Water-attracting or attracted to water, as polar groups on compounds such as lipids and proteins.

hydrophobic Water-repelling or repelled by water (literally "water-hating"), as nonpolar groups on lipids, proteins, etc. which tend to aggregate, excluding water from between them.

hydroxyl radicals OH units with an unpaired election. They are highly reactive and can cause DNA damage.

hyperchromicity The phenomenon in which the absorption of UV light increases as double-stranded DNA denatures to become single-stranded. Native double-stranded DNA absorbs less light at 260 nm by about 40% than does the equivalent amount of single-stranded DNA.

I

ionizing radiation Radiation that can attack (ionize) the deoxyribose sugar in the DNA backbone directly or indirectly by generating reactive oxygen species; e.g. X-rays, radioactive materials.

immunocytochemistry *In situ* analysis of intracellular protein expression and localization using enzyme-conjugated secondary antibodies to the primary antibody against the target protein.

immunofluorescence assay *In situ* analysis of protein expression and localization using fluorescently labeled antibodies. When fluorescently tagged primary antibodies are used for detection, the technique is called direct immunofluorescence assay. When fluorescently tagged secondary antibodies are used for detection, the technique is called indirect immunofluorescence assay.

immunoglobulin (antibody) A protein that binds very specifically to an invading substance and alerts the body's immune defenses to destroy the invader.

immunohistochemistry *In situ* analysis of protein expression and localization in organs using enzyme-conjugated secondary antibodies to the primary antibody against the target protein.

immunoprecipitation A technique in which labeled proteins are reacted with a specific antibody or antiserum bound to resin beads, then precipitated by centrifugation. The precipitated proteins are usually detected by electrophoresis and autodiography.

importin A soluble receptor that mediates nuclear import of nuclear localization sequence (NLS)-bearing cargo.

imprinting control regions (ICRs) Specific intergenic regions responsible for establishing the differential imprint and for maintenance of genomic imprinting during development.

***in situ* hybridization** The base pairing of a labeled probe to metaphase chromosomes on a microscope slide, or to RNA to determine the precise localization within a cell.

in vitro Studies performed in cells or tissues grown in culture, or in cell extracts or synthetic mixtures of cell components.

in vivo Studies performed within a living organism.

incision Nicking a DNA strand with an endonuclease.

independent assortment A principle discovered by Mendel, which states that genes on different chromosomes are inherited independently.

induced mutation A mutation that occurs as a result of interaction of DNA with an outside agent or mutagen that causes DNA damage.
inducer A substance that releases negative control of on operon.
induction The synthesis of enzymes in response to the appearance of a specific substrate.
inherited Transmitted through genes from parents to offspring.
initiation factor A protein that helps catalyze the initiation of translation.
initiator (Inr) A core promoter element surrounding the transcription start site that is important in the efficiency of transcription from some RNA polymerase II gene promoters, especially those lacking TATA boxes.
inosine (I) A nucleoside containing the base hypoxanthine, which base pairs with cytosine; a common modified base found in RNA.
insert A foreign DNA molecule ligated into a vector.
insertion A type of chromosomal abnormality in which a DNA sequence is inserted into a gene, disrupting the normal structure and function of that gene.
insertion sequence (IS) A type of transposon found in bacteria, containing only inverted terminal repeats and the genes needed for transposition.
insulator A DNA regulatory element that acts as a chromatin boundary marker between regions of heterochromatin and euchromatin and can block enhancer or silencer activity of neighboring genes.
integrase An enzyme that integrates one nucleic acid into another; e.g. the provirus of a retrovirus into the host genome.
intellectual property rights Patents, copyrights, and trademarks.
intercalate To insert between two base pairs in DNA.
interferon A double-stranded RNA-activated antiviral protein with various effects on the cell.
intermediate A substrate–product in a biochemical pathway.
internal ribosome entry sequence (IRES) A sequence to which a ribosome can bind and begin translating in the middle of a transcript, without having to scan from the 5′-end.
interphase The stage of the cell cycle during which DNA is synthesized but the chromosomes are not visible.
introns The sequences that remain physically separated after excision during RNA splicing; may encode snoRNAs or miRNAs. The term is used to refer both to the corresponding sequence in the DNA and the RNA.
inverted repeat A symmetrical sequence of DNA, reading the same forward on one strand and backward on the opposite strand. Sometimes referred to as a palindrome.
IPTG Isopropylthiogalactoside. A sulfur-containing lactose analog that is not metabolized by β-galactosidase; used as an inducer of the *lac* operon in the laboratory.
isoelectric focusing Electrophoresing a mixture of proteins through a pH gradient until each protein stops at the pH that matches its isoelectric point. The proteins can no longer move toward the anode or cathode because they have no net charge at the isoelectric point.
isoelectric point The pH at which a protein has no net charge.
isoschizomers Two or more restriction endonucleases that recognize and cut the same restriction site.
ISWI A family of coactivators (imitation SWI/SNF) that help remodel chromatin by moving nucleosomes.

J

J A modified base in trypanosomes that replaces thymine.
joining region (J) The segment of an immunoglobulin gene encoding the last 13 amino acids of the variable region. One of the several joining regions is joined by a chromosomal rearrangement to the rest of the variable region, introducing extra variability into the immunoglobulin gene.

K

karyopherins Family of soluble receptors that mediate nuclear import and export.
karyotype The chromosomal complement of an individual, including the number of chromosomes and any abnormalities. The term is also used to refer to a photograph of an individual's chromosomes.

Kearns–Sayre syndrome A disease linked to mitochondrial DNA (mtDNA) mutations that is characterized by paralysis of eye muscles, progressive muscle degeneration, heart disease, hearing loss, diabetes, and kidney failure.
kilobase pair (kb or kbp) A unit of length use for DNA corresponding to 1000 base pairs.
kilodalton (kD) 1000 Daltons.
kinase An enzyme that catalyzes the addition of phosphate groups to a substrate.
kinetic experiment An experiment that measures the speed (kinetics) of a reaction.
kinetoplast DNA The mitochondrial DNA of trypanosomes; consists of minicircles and maxicircles.
kink A sharp bend in a double-stranded DNA made possible by unstacking of bases.
kink-turn motif An asymmetrical internal loop embedded in a RNA double helix that has a sharp bend in the phosphodiester backbone of the three-nucleotide bulge associated with this motif.
Klenow fragment A fragment of *E. coli* DNA polymerase I, created by cleaving with a protease, that lacks the 5′ → 3′ exonuclease activity of the parent enzyme.
knockdown Repression of gene expression by RNAi.
knockout Inactivation of specific genes by removal of the sequence. Knockouts are often created in laboratory organisms such as yeast or mice by gene targeting so that scientists can study the knockout organism as a model for a particular disease.
known genes Genes from a genomic sequencing project whose sequences are identical to previously characterized genes.
Kozak consensus sequence The sequence context of a eukaryotic translation initiation signal.
Ku The ATPase regulatory subunit of DNA-PK. Binds to double-stranded DNA ends created by chromosome breaks and protects them until repair by nonhomologous end-joining can occur.
Kuru The first form of transmissible spongiform encephalopathy described in humans that was rampant at one time in New Guinea as a result of ritual cannabalism.

L

L strand Light strand of mitochondrial DNA; relative buoyant density determined by density gradient centrifigution.
L1 An abundant human LINE (long interspersed nuclear element) that comprises about 15% of the human genome.
***lac* operon** The operon that encodes enzymes that permit a bacterial cell to metabolize the milk sugar lactose.
Lac repressor A protein, the product of the *E. coli lacI* gene, that forms a tetramer that binds the *lac* operator and thereby represses the *lac* operon.
lacA The *E. coli* gene that encodes galactoside transacetylase.
lacI The *E. coli* gene that encodes the *lac* repressor.
lactose A dissaccharide composed of two simple sugars, glucose and galactose.
lacY The *E. coli* gene that encodes galactoside permease.
lacZ The *E. coli* gene that encodes β-galactosidase.
lagging strand The strand that is made discontinuously in semidiscontiuous DNA replication.
large T antigen The major product of the SV40 viral early region. A DNA helicase that binds to the viral origin of replication and unwinds DNA during replication. Causes malignant transformation of mammalian cells.
lariat The name given the lasso-shaped intermediate in certain splicing reactions.
leader A sequence of untranslated bases at the 5′ end of an mRNA, the 5′ untranslated region (5′-UTR).
leading strand The strand that is made continuously in semidiscontinuous DNA replication.
Leber's hereditary optic neuropathy (LHON) A form of young-adult blindness linked to a small inherited mutation in a mitochondrial gene.
leucine zipper A domain in a DNA-binding protein that includes several leucines spaced at regular intervals. Involved in dimerization with another leucine zipper protein to form a dimer that can then bind DNA.

leukemia Cancer of the developing blood cells in the bone marrow. Leukemia leads to rampant overproduction of white blood cells (leukocytes); symptoms usually include anemia, fever, enlarged liver, spleen, and/or lymph nodes.
library A collection of cloned DNA, usually from a specific organism.
licensing protein complex Mcm2-7 bound to an origin of replication; only licensed origins containing Mcm2-7 can initiate a pair of replication forks.
ligation The joining of linear DNA fragments together with covalent (phosphodiester) bonds.
linkage The association of genes and/or markers that lie near each other on a chromosome, and the likelihood of having one gene and/or marker transmitted with another through meiosis. Linked genes and markers tend to be inherited together.
linker histone One of the more variable histones (H1, H5, H1°) that occurs between core octamers, where the DNA enters and exits the nucleosome.
linker scanning mutagenesis Creation of clustered mutations by replacing small segments of DNA with synthetic oligonucleotides (linkers).
linking number (L) The number of times each strand (chain) crosses the other in a double-stranded DNA circle.
lipoprotein A protein with a lipid group attached post-translationally.
liposome A phospholipid-bound vesicle; often used to introduce DNA into cells.
liquid scintillation counting A techique for measuring the degree of radioactivity in a substance by surrounding it with scintillation fluid, a liquid containing a fluor that emits photons when excited by radioactive emissions.
liver cirrhosis Heavy scarring of the liver; common in chronic alcoholics.
locus (loci, pl.) The place on a chromosome where a specific gene is located; used synonymously with the term gene in many instances.
locus control region (LCR) A chromatin region, such as that associated with the globin genes, that ensures activity of the associated genes, regardless of chromatin location.
LOD score A statistical estimate of whether two loci are likely to lie near each other on a chromosome and are therefore likely to be inherited together. A LOD score of 3 or more is generally taken to indicate that the two loci are close.
long interspersed nuclear elements (LINEs) The most abundant non-LTR retrotransposons in mammals.
long-range regulatory elements Regulatory DNA sequences in multicellular eukaryotes that can work over distances of 100 kb or more from the gene promoter.
long terminal repeats (LTRs) Regions of several hundred base pairs of DNA found at both ends of the provirus of a retrovirus.
LTR-containing retrotransposons A retrotransposon with LTRs at both ends. Replicates its DNA like the provirus of a retrovirus except no transmissible viral particle is involved.
luciferase An enzyme that converts luciferin to a bioluminescent product that emits light and is therefore easily assayed. The firefly luciferase gene is often used as a reporter gene in eukaryotic transcription and translation experiments.
lymphocyte A small white blood cell that plays a major role in defending the body against disease. There are two main types of lymphocytes: B cells, which make antibodies that attack bacteria and toxins, and T cells, which attack body cells themselves when they have been taken over by viruses or become cancerous.
lysis Rupturing the membrane of a cell, as by a virulent bacteriophage.
lysogen A bacterium harboring a prophage.

M

Mad cow disease See bovine spongiform encephalopathy.
Mad-Max A mammalian transcriptional repressor; a heterodimer of Mad and Max proteins.
MALDI-TOF Matrix Assisted Laser Desorption/Ionization – Time Of Flight mass spectrometry. A technique for measuring the mass of peptides. The time of flight of ionized peptides down a tube

toward a detector is inversely proportional to their mass and directly proportional to the charge on the protein.

mammalian artificial chromosome (MAC) A vector that contains a multiple cloning site for very large foreign DNA inserts, centromeric sequences, sequences that can initiate DNA replication and telomeric sequences. MACs segregate with the host cell's chromosomes during cell division.

major groove In B-form DNA, the larger of two continuous indentations running along the outside of the double helix.

maple syrup urine disease A metabolic disorder inherited as an autosomal recessive which affects the metabolism of the amino acids leucine, isoleucine, and valine leading to an accumulation of keto acids that gives the urine a sweet odor resembling maple syrup and interferes with brain function; caused by a defect in a component of the multienzyme branched-chain α-keto acid dehydrogenase complex.

mapping The process of deducing schematic representations of DNA. Three types of DNA maps can be constructed: physical maps, genetic maps, and cytogenetic maps, with the key distinguishing feature among these three types being the landmarks on which they are based.

mariner A defective, inactive human transposon that once transposed by direct DNA replication.

marker Also known as a genetic marker, a segment of DNA with an identifiable physical location on a chromosome whose inheritance can be followed. A marker can be a gene, or it can be some section of DNA with no known function. Because DNA segments that lie near each other on a chromosome tend to be inherited together, markers are often used as indirect ways of tracking the inheritance pattern of genes that have not yet been identified, but whose approximate locations are known.

maternal genes Genes that are expressed during oogenesis in the mother.

matrix attachment regions (MARs) AT-rich regions of DNA typically located near enhancers that organize the genome into chromatin loops with an average loop size of 70 kb; involved in regulating tissue specificity and developmental control of gene expression by recruiting transcription factors and providing a landing platform for chromatin-remodeling enzymes.

Mcm2-7 Minichromosome maintenance proteins. A hexameric (six-subunit) component of the pre-replication complex that is essential for the initiation of DNA replication in eukaryotes and has helicase activity.

Mediator A 20-subunit complex which transduces regulatory information from activator and repressor proteins to RNA polymerase II.

medium Nutritive material in which microorganisms, cells, and tissues are grown in the laboratory, plural "media."

megabase pair (Mb or Mbp) The unit length used for DNA corresponding to 1,000,000 base pairs.

meiosis Cell division that produces gametes (or spores) having half the number of chromosomes of the parental cell.

melanoma Cancer of the cells in the skin that produce melanin, a brown pigment. Melanoma often begins in a mole.

melting temperature (T_m) The temperature at which half the bases in a double-stranded DNA sample have denatured.

membrane protein A protein that folds into characteristic transmembrane helical structures and is embedded in the cell membrane.

Mendelian inheritance Manner in which genes and traits are passed from parents to children. Examples of Mendelian inheritance include autosomal dominant, autosomal recessive, and sex-linked genes.

messenger RNA (mRNA) The template for protein synthesis that binds to ribosomes. Each set of three bases, called codons, specifies a certain amino acid in the sequence of amino acids that comprise the protein. The sequence of a strand of mRNA is based on the sequence of a complementary strand of DNA.

metacentric chromosome A chromosome with the centromere in the middle.

metalloenzyme An enzyme in which binding of divalent cations (e.g. Mg^{2+}) in the active site is critical for its folding into an active conformation.

metaphase The phase of mitosis, or cell division, when the chromosomes align along the center of the cell. Because metaphase chromosomes are highly condensed, these chromosomes are used for gene mapping and identifying chromosomal aberrations.

metastasis The process by which cancer cells travel from the tissue of origin to other parts of the body where they begin to grow and replace normal tissue.

metazoan A multicellular animal.

7-methyl guanosine The capping nucleoside at the beginning of a eukaryotic mRNA.

micrococcal nuclease A nuclease that degrades the DNA between nucleosomes, leaving the nucleosomal DNA intact.

microRNA (miRNA) A short RNA molecule encoded by a cellular gene that folds into a hairpin to create a double-stranded RNA that then triggers the RNAi machinery. miRNAs play a central role in post-transcriptional gene regulation.

microsatellite See short tandem repeats.

microsatellite instability An increase in the accumulation of mutations in the microsatellite regions of DNA, leading to variability in the number or repeats; common in tumors.

mid-blastula transition The point in embryonic development at which the zygote's own genes become active.

minimal medium Culture medium containing a basic set of nutrients only, on which normal wild-type organisms can grow, but which cannot suport the growth of metabolic mutants.

minisatellite A short sequence of (usually) 12 bp or more repeated over and over in tandem; also known as variable number tandem repeats.

minor groove In B-form DNA, the smaller of the two continuous indentations running along the outside of the double helix.

minus 10 box (–10 box) An *E. coli* promoter element centered about 10 bp upstream of the start of transcription.

minus 35 box (–35 box) An *E. coli* promoter element centered about 35 bp upstream of the start of transcription.

mismatch repair The correction of a mismatched base incorporated by mistake during DNA replication; mediated by MutSα or MutSβ which recognize the mismatch, MutLα, exonuclease EXO1 which excises a large region of DNA around the machinery, and the DNA replication machinery which fills in the gap.

missense mutation A change in a codon that results in an amino acid change in the corresponding protein.

mitochondrial DNA (mtDNA) The genetic material of the mitochondria, the organelles that generate energy for the cell.

mitogen A substance, such as a growth factor that stimulates cell division.

mitogen-activated protein kinase (MAPK) A protein kinase that is activated by phosphorylation as a result of a signal transduction pathway initiated by a mitogen such as a growth factor.

mitosis Cell division that produces two daughter cells having nuclei identical to the parental cell.

model organism An organism chosen for study of another organism because of any or all of the following factors: small genome size, short generation time, and ease of manipulation in genetic experiments. A model organism is useful for medical research because it has specific characteristics that resemble a human disease or disorder. Scientists can create animal models, usually laboratory mice, by transferring new genes into them.

molecular biology The study of biological phenomena at the molecular level, in particular the study of the molecular structure of DNA and the information it encodes, and the biochemical basis of gene expression and its regulation.

molecular chaperones See chaperones.

molecular cloning See cloning.

molecular machines The large dynamic macromolecular assemblages that regulate the expression of genetic information in eukaryotes.

monoallelic gene expression Preferential transcription from a single allele in a cell; also known as allelic exclusion.

monoclonal antibodies Identical antibodies to a specific epitope of a protein produced by a clone originating from one cell.

monosomy Possessing only one copy of a particular chromosome instead of the normal two copies.

morpholino oligonucleotides Modified DNA analogs with an altered backbone linkage that lacks a negative charge; used in antisense-mediated inhibition of gene expression.

mouse model A laboratory mouse useful for medical research because it has specific characteristics that resemble a human disease or disorder. Mouse models can be created by transferring new genes into mice or by inactivating certain existing genes in them.

MRN complex A complex comprising exonuclease Mre11/Rad50/NBS1 that is recruited to double-strand break sites in DNA and initiates repair.

MTE Motif ten element; an RNA polymerase II core promoter element found in a few gene promoters.

multiple cloning site (MCS) A region in a cloning vector that contains several unique restriction sites in tandem. Any of these can be used for inserting foreign DNA.

multiplex PCR Simultaneous amplification of many targets of interest in one reaction by using more than one pair of primers.

mutagen Any chemical agent that causes an increase in the rate of mutation above the spontaneous background.

mutant An organism (or genetic system) that has suffered at least one mutation compared to the wild-type (normal).

mutation A permanent structural alteration in DNA. In most cases, DNA changes either have no effect or cause harm, but occasionally a mutation can improve an organism's chance of surviving and passing the beneficial change on to its descendants.

N

negative control A control system in which gene expression is turned off unless a controlling element (e.g. repressor) is removed.

neurofibromatosis An inherited progressive disorder in which tumors form on peripheral nerves. The tumors can result in loss of hearing and vision, cancer, epilepsy, bone deformities, and learning disabilities.

N-formyl methionine The initiating amino acid in prokaryotic translation.

nick A single-stranded break of a phosphodiester bond in DNA.

nitrogenous base Nitrogen-containing molecules having the chemical properties of a base that are components of nucleotides. Two of the bases, adenine (A) and guanine (G), have a double carbon-nitrogen ring structure. The other three bases, thymine (T), cytosine (C), and uracil (U), have a single-ring structure. Thymine is found in DNA only, while uracil is specific for RNA.

nonautonomous retrotransposon A non-LTR retrotranposon that encodes no proteins, so it depends on other retrotransposons for transposition activity.

noncanonical (non-Watson–Crick) base pairs Unconventional base pairs that form in RNA double helices such as the GU wobble, the sheared GA pair, the Reverse Hoogsteen pair, and the GA imino pair.

nonhomologous end-joining (NHEJ) A eukaryotic mechanism for repairing double-strand breaks in DNA. Double-strand breaks are rejoined via direct ligation of the DNA ends without any requirement for sequence homology.

non-LTR retrotransposons A retrotransposon that lacks LTRs and replicates by a mechanism different from that used by the LTR-containing retrotransposons.

nonpermissive conditions Those conditions under which a conditional mutant gene product cannot function (e.g. the inhibitory temperature for a temperature-sensitive mutant).

nonsense codons UAG, UAA, UGA These stop codons terminate protein synthesis by the ribosome.

nonsense-mediated mRNA decay (NMD) A eukaryotic system for degrading mRNAs with premature termination (stop) codons.

nonsense mutation A single DNA base substitution resulting in a stop codon.
nonspecific binding (protein) The first contact of a DNA-binding protein with DNA involves interactions with the the sugar-phosphate backbone and not the bases.
nontemplate DNA strand The strand complementary to the template strand. Also called the coding strand or sense strand.
nontranscribed spacer (NTS) A DNA region lying between two ribosomal RNA precursor genes in a tandemly repeated cluster of such genes.
Northern blot Transfer of RNA fragments to a support medium (see Southern blot). A technique used to identify and locate mRNA sequences that are complementary to a piece of DNA (or antisense RNA) called a probe.
novel variant of CJD (vCJD) A human form of transmissible spongiform encephalopathy linked to eating beef products from a cow with bovine spongiform encephalopathy (mad cow disease).
nuclear export sequence (NES) A specific amino acid sequence that interacts with an exportin and mediates nuclear export of the protein via the nuclear pore complex; typically hydrophobic leucine-rich sequences.
nuclear factor kappa B (NF-κB) A transcription factor that activates expression of the many genes involved in the immune system and stress response in cells.
nuclear lamina A protein meshwork underlying the nuclear membrane that is primarily composed of the intermediate filament proteins lamins A, B, and C.
nuclear localization sequence A specific amino acid sequence that interacts with an importin and targets the protein to the nucleus via the nuclear pore complex; typically rich in lysine and arginine.
nuclear magnetic resonance (NMR) spectroscopy A technique in which a concentrated protein solution is placed in a magnetic field and the effects of different radiofrequencies on the resonances of different atoms are measured.
nuclear matrix A structural organizer within the cell nucleus. Operationally defined as a branched meshwork of insoluble filamentous proteins within the nucleus that remains after digestion with high salt, nucleases, and detergent.
nuclear pore complexes Large multiprotein complexes embedded in the nuclear membrane that serve as a selective gateway for the exchange of material between the nucleus and cytoplasm.
nuclear receptor A transcription factor that interacts with ligand, such as glucocorticoids, thyroid hormone, vitamin D, or retinoic acid and binds to an enhancer or silencer to stimulate or repress transcription, respectively.
nucleic acids A long chain-like polymer (DNA or RNA) of repeating subunits called nucleotides.
nucleoid In prokaryotes, the condensed ovoid region of the cell containing the chromosomal DNA.
nucleolus A nonmembrane-bound cell organelle found in the nucleus that contains the ribosomal RNA genes and is the site of ribosome assembly.
nucleoside A nitrogenous base chemically linked to one molecule of a five-carbon sugar at the 1′ carbon of the sugar – either ribose (RNA) or deoxyribose (DNA).
nucleoside triphosphates (NTPs) The building blocks of RNA: ATP, CTP, GTP, UTP. Also referred to as ribonucleoside triphosphates.
nucleosome A repeating structural element in eukaryotic chromosomes, composed of the core octamer of histones (two each of histones H2A, H2B, H3, and H4) plus one molecule of the linker histone (H1) with approximately 180 bp DNA wrapped around.
nucleotide (nt) One of the structural components, or building blocks, of DNA and RNA. A nucleotide is composed of three parts: a five-carbon sugar, at least one phosphate group, and a nitrogenous base.
nucleotide excision repair A repair pathway in which DNA damage such as a pyrimidine dimer is recognized by the cooperative binding of RPA, XPA, XPC, and the TFIIH complex (including the helicases XPB and XPD). XPG endonuclease then cuts the DNA strand on either side of the damaged base, removing an oligonucleotide that contains the damage. The gap is filled in with DNA polymerase and DNA ligase.

nucleotide substitution A type of mutation in which a nucleotide pair in a DNA duplex is replaced with a different nucleotide pair. Mutations that alter a single nucleotide pair are called point mutations.
nucleus The central cell structure that houses the chromosomes.

O

O_6-methylguanine methyl transferase A suicide enzyme that accepts methyl or ethyl groups from alkylated DNA bases and thereby reverses the DNA damage.
Okazaki fragments Small DNA fragments, 1000–2000 bases long, created by discontinuous synthesis of the lagging strand.
oligo(dT) cellulose affinity chromatography A method for purifying poly (A+) RNA by binding it to oligo(dT) cellulose in buffer at relatively high ionic strength, and eluting it at low ionic strength.
oligonucleotide (oligo) A short sequence of single-stranded DNA or RNA. Oligos are often used as probes for detecting complementary DNA or RNA because they bind readily to their complements.
oncogene A gene whose product is capable of causing the transformation of normal cells into cancer cells through gain of function.
open reading frame (ORF) A reading frame that is uninterrupted by translation stop codons.
operator A DNA element found in prokaryotes that binds tightly to a specific repressor and thereby regulates the expression of adjacent genes.
operon A group of genes coordinately controlled by an operator.
oriC The *E. coli* origin of replication.
origin firing Initiation of a bidirectional pair of replication forks.
origin of replication The unique site on chromosomal DNA (replicon) where replication begins, and a bidirectional pair of replication forks initiate.
origin recognition complex A eukaryotic ATP-regulated DNA binding complex composed of six polypeptide subunits that binds to origins of replication and then recruits cdc6, Cdt1, and Mcm2-7 proteins, other components of the pre-replication complex that are essential for the initiation of DNA replication. The SV40 large T antigen functions as a viral ORC comparable to the cellular ORC.
orthologs Homologous genes in different species that have evolved from a common ancestral gene.
overexpression Production of a large quantity of recombinant proteins in a bacterial or eukaryotic host cell.
8-oxoguanine (oxo8) A damaged guanine base containing an extra oxygen atom; it can form a Hoogsteen base pair with adenine, thereby giving rise to a GC to TA transversion.

P

P element A transposable element of *Drosophila*, responsible for hybrid dysgenesis; used as a tool by molecular biologists for mutagenesis studies (transposon tagging).
P site The ribosomal site to which a peptidyl tRNA is bound at the time a new aminoacyl-tRNA enters the ribosome.
p53 A tumor suppressor gene which normally regulates the cell cycle and protects the cell from damage to its genome. Mutations in this gene cause cells to develop cancerous abnormalities.
palindrome See inverted repeat.
pandemic A worldwide epidemic of an infectious disease.
panediting Extensive editing of a pre-mRNA; occurs in trypanosome kinetoplast transcripts.
paralogs Homologous genes that have evolved by gene duplication within a species.
Parkinson's disease Common progressive neurological disorder that results from degeneration of nerve cells in a region of the brain that controls movement. The first symptom of the disease is usually tremor of a limb, especially when the body is at rest.
patent When applied to molecular biology, the government regulations or requirements conferring the right or title to an individual or organization to genes if there has been substantial human intervention.
pathogenic Causing disease; e.g. a parasite (especially a microorganism) in relation to a particular host.

pathway (biochemical) A series of biochemical reactions in which the product of one reaction (an intermediate) becomes the substrate for the next reaction.
pause sites DNA sites where an RNA polymerase pauses before continuing elongation.
pBR322 One of the original plasmid vectors for gene cloning.
pedigree A simplified diagram of a family's genealogy that shows family members' relationships to each other and how a particular trait or disease has been inherited.
peptide Two or more amino acids joined by a peptide bond.
peptide bond Covalent bond joining the α-amino group of one amino acid to the carboxyl group of another with the loss of a water molecule, and which is the bond linking amino acids together in a polypeptide chain.
peptidyl transferase An enzymatic activity that is an integral part of the large ribosomal RNA in the ribosome that catalyzes the formation of peptide bonds during protein synthesis.
permissive conditions Those conditions under which a conditional mutant gene product can function (i.e. the permissive temperature for a temperature-sensitive mutant).
Pfu polymerase A thermostable DNA polymerase from *Pyrococcus furiosus* used in PCR that has greater fidelity than Taq polymerase.
phage See bacteriophage.
phage mu A phage of *E. coli* that replicates by transposition.
phenotype The biochemical, behavioral, morphological, or other properties of an organism.
phosphatase An enzyme that catalyzes the removal of a phosphate group from a substrate.
phosphodiester bond The covalent sugar-phosphate bond (–O–P–) that forms the linkage between adjacent nucleotides in a nucleic acid (DNA or RNA). The hydroxyl group on the 3′ carbon of a sugar of one nucleotide forms an ester bond to the phosphate group on the 5′ carbon of another nucleotide, eliminating a molecule of water. The bond is also referred to as a 5′ → 3′ phosphodiester bond, indicating the polarity of the strand.
phosphorimaging A technique for measuring the degree of radioactivity of a substance (e.g. on a blot) electronically without using X-ray film, using an instrument called a phosphorimager.
phosphorylation (protein) The post-translational modification of a protein by addition of a phosphate group at serine, threonine, or tyrosine residues.
photoreactivation Direct repair of a pyrimidine dimer by DNA photolyase.
physical map A chromosome map of a species that shows the specific physical locations of its genes and/or markers on each chromosome; e.g. based on physical characteristics such as restriction sites. Physical maps are particularly important when searching for disease genes by positional cloning strategies and for DNA sequencing.
plaque A hole that a virus makes on a layer of host cells by infecting and either killing the cells or slowing their growth.
plasmid A small, double-stranded circular (or linear) DNA molecule carried by bacteria, some fungi, and some higher plants. They are extrachromosomal, independent, and self-replicating.
pleiotropy Multiple phenotypic manifestations.
point mutation An alteration (substitution) of one, or a very small number, of contiguous bases.
polar molecule A molecule with an asymmetrical distribution of charge across the molecule.
poly(A)-binding protein (PABP) A protein that binds the poly(A) tails at the end of the pre-mRNA. There are nuclear (PABPN) and cytoplasmic (PABPC) forms.
poly(A)+ RNA RNA that contains a poly(A) tail at its 3′ end.
poly(A)− RNA RNA that does not contain a poly(A) tail at its 3′ end.
poly(A) tail Polyadenylic acid tail. The string of about 200 As added to the end of a typical eukaryotic mRNA.
polyacrylamide A cross-linked polymer of acrylamide; used in electrophoretic separation of nucleic acids and proteins.
polyadenylation Addition of poly(A) to the 3′ end of an RNA.

polyadenylation signal The set of RNA sequences that mediate the cleavage and polyadenylation of a transcript. An AAUAAA sequence followed 20–30 nucleotides downstream by a GU-rich region, then a U-rich region is the canonical cleavage signal. After cleavage, the AAUAAA sequence is the polyadenylation signal.

polycistronic message An mRNA containing information from more than one gene.

polyclonal antibodies A mixture of antibodies with different specificities to different epitopes.

Polycomb group proteins A family of proteins that mediate silencing of homeobox genes by altering the higher order structure of chromatin.

polydactyly An abnormality in which a person is born with more than the normal number of fingers or toes.

polymerase chain reaction (PCR) Amplification of a specific region of DNA using primers that flank that region and repeated cycles of DNA polymerase activity. A fast, inexpensive technique for making an unlimited number of copies of any piece of DNA.

polymerase switching The hand-off ("trading places") of the DNA template from one DNA polymerase to another between primer synthesis and elongation, and during translesion DNA synthesis.

polymorphism A common variation in the sequence of DNA among individuals.

polynucleotide A polymer composed of nucleotide subunits; either DNA or RNA.

polypeptide A polymer composed of amino acid subunits; a single protein chain.

polysome A messenger RNA attached to several ribosomes.

position effect A phenomenon in which expression of a transgene is unpredictable; it varies with the chromosomal site of integration.

positional cloning A process which, through gene mapping techniques, is able to locate a gene responsible for a disease or other genetic traits when little or no information is known about the biochemical basis of the disease or trait.

positive control A control system in which gene expression depends on the presence of a positive effector such as CAP (and cAMP).

post-transcriptional control Control of gene expression that occurs when transcripts are processed by capping, splicing, editing, nuclear export, and degradation pathways.

post-transcriptional gene silencing (PTGS) See RNA interference.

post-translational modification The set of changes that occur in a protein after it is synthesized (e.g. modification by phosphorylation).

Prader–Willi syndrome A neurodevelopmental disorder caused by defects in genomic imprinting in a region of chromosome 15.

preinitiation complex (PIC) The combination of eukaryotic RNA polymerase and general transcription factors assembled at a promoter just before transcription begins.

pre-replication complex A complex of ORC, cdc6, Cdt1, and Mcm2-7 assembled at a eukaryotic origin of replication that is required for the initiation of a replication fork.

primary antibody The first set of antibodies made for a specific epitope.

primary structure The sequence of amino acids in a polypeptide, or of nucleotides in DNA or RNA.

primary transcript The initial, unprocessed RNA product of a gene.

primase The enzyme that catalyzes the formation of an RNA primer during DNA replication.

prime See five (5)-prime (5′) and three (3)-prime (3′).

primer (DNA replication) A small piece of RNA that provides the free 3′-OH end needed for DNA replication to begin.

primer (techniques) A short oligonucleotide sequence (DNA) used in laboratory techniques that involve DNA polymerase, e.g. the polymerase chain reaction, DNA sequencing, random primed DNA labeling, etc.

prion Proteinaceous infectious particle. The causative agent of transmissible spongiform encephalopathies. A prion is a misfolded form of a normal cellular protein that propagates through promoting misfolding of the normal protein into the infectious prion form. Prions can be sporadic, inherited, or infectious.

probe (nucleic acid) A piece of DNA or RNA, labeled with a tracer (typically radioactive), that allows a molecular biologist to track the hybridization of the probe to an unknown DNA or RNA. For example, a radioactive probe can be used to identify an unknown DNA band on a Southern blot.

processivity The tendency of an enzyme to remain bound to one or more of its substrates during repetitions of the catalytic process. For example, the longer an RNA or DNA polymerase continues making its product without dissociating from its template, the more processive it is.

programmed gene rearrangements Rearrangements of DNA that regulate the expression of some genes.

prokaryotes Microorganisms that lack nuclei; comprising eubacteria, cyanobacteria, and Archaea.

proliferating cell nuclear antigen (PCNA) A eukaryotic protein that acts as a "sliding clamp" to increase processivity of DNA polymerase during DNA replication, along with many other functions in the cell.

promoter The collection of DNA sequence elements, including the core promoter and promoter proximal elements, that are required for initiation of transcription or that increase the frequency of initiation only when positioned near the transcriptional start site.

promoter clearance The process by which an RNA polymerase moves away from a promoter after initiation of transcription.

promoter strength The relative frequency of transcription initiation; related to the affinity of RNA polymerase for the promoter region.

pronucleus The nucleus of a sperm or an egg prior to fertilization. Sperm and egg cells carry half the number of chromosomes of other nonreproductive cells. When the pronucleus of a sperm fuses with the pronucleus of an egg, their chromosomes combine and become part of a single nucleus in the resulting embryo, containing a full set of chromosomes.

proofreading (aminoacyl-tRNA synthetase) The process by which aminoacyl adenylates and, less commonly, aminoacyl-tRNAs are hydrolyzed if their amino acids are too small for the synthetase.

proofreading (DNA) The process used by DNA polymerase to check the accuracy of DNA replication as it occurs and to replace a mispaired nucleotide with the correct one. DNA polymerase has $3' \rightarrow 5'$ exonuclease activity which mediates this process.

proofreading (protein synthesis) The process by which aminoacyl-tRNAs are cross-checked on the ribosome for correctness before the amino acids are incorporated into the growing polypeptide chain.

proofreading (RNA) The backtracking of RNA polymerase along the DNA template followed by cleavage of the most recently added base(s).

prophage A phage genome integrated into the host's genome.

prophase Early stage of cell division in which chromosomes condense and become visible.

prosthetic group Non-protein chemical group (e.g. metal ion or heme) bound to a protein, as in many enzymes, usually forming part of the active site, and essential for biological activity (see also cofactor).

protease An enzyme that cleaves other proteins.

proteasome A collection of proteins (sedimentation coefficient 26S) that proteolytically degrades a ubiquitinylated protein.

Protection of Telomeres (POT1) A protein involved in regulation of telomerase activity; it binds the 3′ single-stranded DNA tail of telomeres and participates in forming a folded chromatin structure that prevents access of telomerase to telomeres.

protein A large complex polymer or polypeptide made up of one or more chains of amino acid subunits. Proteins perform a wide variety of activities in the cell. Sometimes the term protein denotes a functional collection of more than one polypeptide (e.g. the hemoglobin protein as a quaternary structure that consists of four polypeptide chains).

protein arrays A high-throughput technique that allows rapid analysis of protein activity on a proteomic scale.

protein kinase A A serine-threonine-specific protein kinase whose activity is stimulated by cAMP.

protein sequencing Determining the sequence of amino acids in a protein.

proteome The complete set of proteins encoded by the genome.

proteomics The comprehensive study of the full set of proteins encoded by a genome.

proto-oncogene A cellular gene with the potential to give rise to an oncogene through inappropriate activation.

provirus A double-stranded DNA copy of a retroviral RNA genome, which inserts into the host cell genome.

proximal promoter element Regulatory element located just upstream of the core promoter and usually within 70–200 bp of the start of transcription (e.g. the CCAAT box and GC box); increases the frequency of initiation of transcription but only when positioned near the transcriptional start site.

PrP^C A normal cellular protein that when misfolded can become an infectious prion.

PrP^{Sc} Scrapie prion protein, the misfolded infectious form of the PrP^C protein.

pseudogene A sequence of DNA that is very similar to a normal gene but that has been altered slightly so it is not expressed. Such genes were probably once functional but over time acquired one or more mutations that rendered them incapable of producing a protein product.

pseudoknot A motif that forms in RNA when a single-stranded loop base pairs with a complementary sequence outside this loop and folds into a three-dimensional structure by coaxial stacking.

pseudouridine A modified nucleoside found in RNA, in which the ribose is joined to the 5-carbon instead of the 1-nitrogen of the uracil base.

pUC vectors Plasmid vectors based on pBR322, containing an ampicillin resistance gene, and a multiple cloning site that interrupts the *lacZ* gene, which enables blue/white screening for inserts.

pull-down assay A procedure for analysis of protein–protein interactions *in vitro*. For example, a glutathione S-transferase (GST) pull-down assay tests interaction between a GST-tagged protein (the bait) and another protein (the prey).

pulse-chase The process of giving a short period (the pulse) of radioactive precursor so that a substance such as protein becomes radioactive, then adding an excess of unlabeled precursor to "chase" the radioactivity out of the substance as protein turnover occurs.

pulse-field gel electrophoresis (PFGE) An electrophoresis technique in which the electric field is repeatedly reversed. Allows separation of very large pieces of DNA, up to several Mb in size.

pulse labeling Providing a radioactive precursor for only a short time. For example, DNA can be pulse labeled by incubating cells for a short time in radioactive thymidine.

purine The nitrogenous bases adenine (A) and thymine (T) which have a double carbon-nitrogen ring structure.

pyrimidine The nitrogenous bases thymine (T), cytosine (C), and uracil (U) which have a single carbon-nitrogen ring structure.

pyrimidine dimers Two adjacent pyrimidines in one DNA strand linked covalently, interrupting their base pairing with purines in the opposite strand. The main DNA damage caused by UV light is the formation of T-T dimers.

Q

quaternary structure The way two or more polypeptides interact to form a functional complex protein.

quenching Quickly chilling heat-denatured DNA to keep it denatured and prevent the single strands from annealing.

R

R-looping A classic technique for visualizing hybrids between DNA and RNA by electron microscopy. An R loop is formed when an RNA hybridizes to one strand of DNA and displaces the other strand as a loop.

Rad50, Rad52, Rad55, Rad57 Proteins involved in repair of DNA double-stranded breaks by homologous recombination.

Ran A member of the superfamily of GTP-binding proteins that act as molecular switches cycling between GDP- and GTP-bound states.

random primed labeling A method of incorporating radioactive nucleotides along the length of a fragment of DNA.

random walk (protein) The equal probabilities for forward and reverse steps as a protein slides by linear diffusion along DNA over short distances from a nonspecific binding site.

RAPD Randomly amplified polymorphic DNA analysis. A PCR reaction where the primers consist of random sequences. Pronounced "rapid."

reactive oxygen species See free radicals.

reading frame One of three possible ways the triplet codons in an mRNA can be translated, depending on where translation begins. A natural mRNA generally has only one correct reading frame.

recessive An allele or trait that does not express its phenotype when heterozygous with a dominant allele. Also, a genetic disorder that appears only in patients who have received two copies of a mutant gene, one from each parent.

recognition helix The α-helix in a DNA-binding motif of a DNA-binding protein that fits into the major groove of its DNA target; the sequence specific contacts define the specificity of the protein.

recombinant DNA The product of recombination between two (or more) fragments of DNA. Can occur naturally in a cell through genetic processes such as crossing-over, but the term is generally reserved for DNA molecules produced by joining segments derived from different biological sources (i.e. constructed by molecular biologists *in vitro*).

recombinant DNA technology A variety of techniques that molecular biologists use to manipulate DNA molecules to study the expression of genes.

recombination Reassortment of genes or alleles in new combinations. Occurs by crossing-over between or within DNAs.

recombination signal sequence (RSS) A specific sequence at a recombination junction, recognized by the recombination apparatus during immunoglobulin and T-cell receptor gene maturation.

recruitment Promoting the binding of a substance to a complex. Often refers to enhancing the binding of RNA polymerase or transcription factors to a promoter, or the binding of replication factors to an origin of replication.

relaxed DNA A DNA molecule that is not supercoiled.

release factor A protein that causes termination of translation at stop codons.

renaturation of DNA The process by which when heated solutions of denatured DNA are slowly cooled, single strands often meet complementary strands and form a new double helix. This is also referred to as annealing.

repetitive DNA DNA sequences that are repeated may times in a haploid genome.

replica plating Transferring colonies of bacteria or other cells from one culture plate to another, usually using a plating tool coated with a soft material that can pick up cells from one plate and place them in the same relative positions on the second plate (or membrane).

replication factor C (RFC) The ATP-dependent "clamp loader" that loads PCNA onto the DNA during replication.

replication factories Clusters of 40 to many hundreds of active replication forks in discrete subnuclear compartments or foci.

replication fork The point where the two parental DNA strands separate to allow replication.

replication initiation point (RIP) mapping A procedure that allows the detection of start sites for DNA synthesis at the nucleotide level.

replication licensing A system in eukaryotes that ensures that DNA replicates only once per cell cycle. Mediated through tight regulation of the formation and activation of pre-replication complexes by the levels of cyclin-dependent kinases.

replication protein A (RPA) A mammalian single-stranded DNA-binding protein that binds to single-stranded DNA and keeps it from base pairing with a complementary strand, used during DNA replication.

replicative transposition See duplicative transposition.

replicon All the DNA replicated from one origin of replication.
reporter gene A gene attached to a promoter or translation start site and used to measure the activity of the resulting transcription or translation. The reporter gene serves as an easily assayed counterpart for the gene it replaces.
repressor A protein that regulates a gene by turning it off.
resolution The final step in recombination, in which the second pair of strands is broken.
resolvase An endonuclease that nicks two DNA strands to resolve a Holliday junction after branch migration.
restriction endonuclease An enzyme that recognizes specific base sequences of double-stranded DNA and cuts the DNA at or near thoses sites. Restriction endonucleases are often called restriction enzymes.
restriction fragment A piece of DNA cut from a larger DNA by a restriction endonuclease.
restriction fragment length polymorphism (RFLP) A variation from one individual to the next in the number of cuttting sites for a given restriction endonuclease in a given genetic locus. Such variations affect the size of the resulting fragments. These sequences can be used as markers on physical maps and linkage maps. RFLP is pronounced "rif-lip."
restriction map A map that shows the locations of restriction sites in a region of DNA.
restriction-modification system The combination of a restriction endonuclease and the DNA methylase that recognizes the same DNA site in bacteria.
restriction site A sequence of nucleotides recognized and cut by a restriction endonuclease.
retinoblastoma (Rb) protein A tumor suppressor protein that acts as a cell cycle master switch.
retrohoming The process by which a mobile group I or group II intron in one gene can transpose to an intronless version of the same gene elsewhere in the genome.
retrotransposon A transposable element such as Ty in yeast that transposes via a retrovirus-like mechanism.
retrovirus A type of virus that contains RNA as its genetic material and whose replication depends on formation of a DNA provirus by reverse transcription. Retroviruses can cause many diseases, including some cancers and AIDS.
Rett syndrome A neurodevelopmental disorder caused by mutations in the gene coding for methyl-CpG-binding protein 2.
Rev1 A eukaryotic DNA polymerase involved in error-prone repair of abasic sites; its deoxycytidyl transferase activity inserts a C across from a nucleotide lacking a base.
reverse transcriptase An RNA-dependent DNA polymerase which is the enzyme commonly found in retroviruses that catalyzes reverse transcription.
reverse transcriptase PCR (RT-PCR) A PCR method that begins with the synthesis of cDNA from an mRNA template, using reverse transcriptase. The cDNA then serves as the template for conventional PCR.
reverse transcription Synthesis of a DNA using an RNA template.
rho (ρ) A protein that is needed for transcription termination at certain terminators in *E. coli* and its bacteriophages.
rho-dependent terminator A terminator in *E. coli* and its bacteriophages that requires rho for its activity.
rho-independent terminator A terminator in *E. coli* and its bacteriophages that does not require rho for its function. Also called an intrinsic terminator.
ribonuclease (RNase) An enzyme that degrades RNA.
ribonucleic acid (RNA) A polymer composed of ribonucleotides linked together by phosphodiester bonds. In RNA, the letter U, which stands for uracil, is substituted for T in the genetic code.
ribonucleoprotein particle (RNP) An RNA–protein complex.
ribonucleoside triphosphates (NTPs) See nucleoside triphosphates.
riboprobe A labeled RNA probe, commonly used in RNase protection assays and *in situ* hybridization.
ribose The pentose (5-carbon) sugar present in a nucleotide subunit of RNA.
ribose zipper motif An RNA motif that involves hydrogen bonding between the 2′-OH of a ribose in one helix and the 2-oxygen of a pyrimidine base (or the 3-nitrogen of a purine base) of the other helix between their respective minor groove surfaces.
ribosomal RNA (rRNA) The RNA molecules contained in ribosomes.

ribosomal subunit (30s) The small bacterial ribosomal subunit, involved in mRNA decoding.
ribosomal subunit (40s) The small eukaryotic ribosomal subunit, involved in mRNA decoding.
ribosomal subunit (50s) The large bacterial ribosomal subunit, involved in peptide bond formation.
ribosomal subunit (60s) The large eukaryotic ribosomal subunit, involved in peptide bond formation.
ribosome An RNA–protein complex that translates mRNAs to produce proteins (70S in bacteria and 80S in eukaryotes).
riboswitch A switchable "on–off" RNA element that selectively binds metabolites or functions as a sensor for signals as diverse as temperature or salt concentration and controls gene expression.
riboswitch ribozyme A riboswitch that is a metabolite-responsive ribozyme.
ribozyme A catalytic RNA molecule (RNA enzyme). Ribozymes can catalyze a number of chemical reactions that take place in living cells, ranging from cleavage of phosphodiester bonds to peptide bond formation.
RNA-directed RNA polymerase (RdRP) The enzyme that elongates the siRNA primers, using target mRNA as a template, thus providing more substrate for Dicer and amplying the siRNA.
RNA helicase An enzyme that can unwind a double-stranded RNA, or a double-stranded region within RNA.
RNA-induced silencing complex (RISC) The RNase complex that degrades target mRNA during RNAi; contains endonuclease "Slicer" activity.
RNA interference (RNAi) Control of gene expression by specific mRNA degradation or translational repression caused by insertion of a double-stranded RNA into a cell.
RNA ligase An enzyme that can join two pieces of RNA, such as the two pieces of pre-tRNA created by splicing out an intron.
RNA polymerase In general, the enzyme that directs transcription, or synthesis of RNA. Specifically, the bacterial RNA polymerase.
RNA polymerase I The eukaryotic RNA polymerase that directs transcription of ribosomal RNA.
RNA polymerase II The eukaryotic RNA polymerase that directs transcription of mRNA, snoRNAs, some snRNAs, and miRNAs.
RNA polymerase III The eukaryotic RNA polymerase that directs transcription of tRNA, 5S rRNA, and some snRNAs.
RNA polymerase IV The eukaryotic RNA polymerase (described in plants only) that directs transcription of siRNAs involved in heterochromatin formation.
RNA polymerase, chloroplast The multiple eukaryotic RNA polymerases (both chloroplast and nuclear-encoded) that direct transcription of chloroplast genes.
RNA polymerase, mitochondrial The eukaryotic RNA polymerase that directs transcription of mitochondrial genes.
RNA processing Modifying an initial transcript to its mature form by cleavage, splicing, capping, polyadenylation, editing, etc.
RNA splicing The process of removing introns from a primary transcript and attaching the exons to one another.
RNA world A hypothetical stage in the evolution of life some four billion years ago when RNA both carried the genetic information and catalyzed its own replication.
RNase MRP RNase mitochondrial RNA processing. The enzyme that cleaves the RNA primer during mtDNA replication and plays a role in processing 5.8S rRNA in the nucleolus.
RNase P The enzyme that cleaves the extra nucleotides from the 5′ end of a tRNA precursor. Most forms of bacterial RNase P have a catalytic RNA subunit.
RNase protection assay (RPA) A method for detecting and quantifying specific mRNA transcripts in a complex mixture of RNA and for mapping internal and external boundaries in mRNA, such as the start of transcription. Hybridization of a labeled RNA probe to a transcript protects a part of the probe from digestion with RNase. The amount of probe protected by the transcript is proportional to the concentration of the transcript.

rolling circle replication A mechanism of replication in which one strand of a double-stranded circular DNA remains intact and serves as the template for elongation of the other strand at a nick.

rRNA (5S) The smaller of the vertebrate rRNAs found in the large (60S) ribosomal subunit; transcribed from a separate gene. Sometimes referred to as 5S RNA.

rRNA (5.8S) The smallest of the eukaryotic rRNAs derived from the 45S or 40S precursor. Found in the large (60S) ribosomal subunit; base paired to the 28S rRNA.

rRNA (18S) The vertebrate rRNA found in the small (40S) ribosomal subunit.

rRNA (28S) The largest vertebrate rRNA found in the large (60S) ribosomal subunit, base paired to the 5.8S rRNA.

rRNA precursor (45S or 40S) The large rRNA precursor in vertebrates which contains the 28S, 18S, and 5.8S rRNA sequences.

Rubenstein–Taybi syndrome A rare autosomal dominant disease characterized by facial abnormalities, broad digits, stunted growth, and mental retardation, caused by mutations in the gene coding for CBP which is a coactivator with HAT activity.

run-off transcription A method for quantifying the extent of transcription of a particular gene *in vitro*. The RNA polymerase "runs off" the end of a truncated linear gene template to give a short RNA product of predictable length. The abundance of the run-off product is a measure of the extent of transcription of the gene *in vitro*.

S

S1 nuclease A nuclease specific for single-stranded RNA and DNA.

satellite RNA A subviral pathogen found in plants and animals that replicates only in the presence of a helper virus. Some of the larger satellite RNAs may encode a protein.

scanning A model of translation initiation in eukaryotes that invokes a 40S ribosomal subunit binding to the 5′ end of the mRNA and scanning, or sliding along, the mRNA until it finds the first start codon in a good context for initiation.

scrapie The prototype prion disease first described in sheep and goats.

screen A genetic sorting procedure that allows one to distinguish desired organisms from unwanted ones, but does not automatically remove the undesired ones.

SDS-PAGE SDS-polyacrylamide gel electrophoresis; the electrophoretic separation of proteins in the presence of a strong, negatively charged detergent (sodium dodecyl sulfate).

secondary antibody Antibodies directed against all primary antibodies of a given species (e.g. anti-rabbit); typically conjugated (covalently bonded) to detectable tags.

secondary structure The local folding of a polypeptide or nucleic acid.

sedimentation coefficient (S) A measure of the rate at which a molecule or particle travels toward the bottom of a centrifuge tube under the influence of a centrifugal force. Also known as a Svedberg unit.

selection A genetic sorting procedure that eliminates unwanted organisms of cell types, usually by preventing their growth or killing them.

semiconservative replication DNA replication in which the two strands of a parental duplex separate completely and pair with new progeny strands. One parental strand is therefore conserved in each progeny duplex.

semidiscontinuous replication A mechanism of DNA replication in which one strand is made continuously and the other is made discontinuously.

senescence An irreversible state of cellular aging, characterized by continued cell viability without further cell division.

sequencing Determining the amino acid sequence of a protein, or the base sequence of a DNA or RNA.

severe combined immunodeficiency (SCID) A disease affecting the immune system. SCID is fatal if affected individuals do not receive bone marrow transplants.

sex chromosome One of the two chromosomes that specify an organism's genetic sex. Humans have two kinds of sex chromosomes, one called X and the other Y. Normal females possess two X chromosomes and normal males one X and one Y.

sex-linked Located on the X chromosome. Sex-linked (or X-linked) diseases are generally seen only in males.

SH2 domain A phosphotyrosine-binding domain found in many signal transduction proteins.

SH3 domain A proline-rich-helix binding domain that mediates protein–protein interactions.

Shine–Dalgarno sequence A G-rich sequence (consensus = AGGAGGU) that is complementary to the 3′ end of *E. coli* 16S rRNA. Base pairing between these two sequences helps the ribosome bind an mRNA.

short interspersed nuclear elements (SINEs) Nonautonomous retrotransposons that do not encode the enzymatic activities necessary for their retrotransposition. The only active SINEs in the human genome are members of the abundant *Alu* element family.

short tandem repeats (STRs) Repetitive stretches of short sequences of DNA (usually 2–4 bp) used as genetic markers. A given STR is found in varying lengths scattered around a eukaryotic genome. STR analysis is currently the most widely used DNA typing procedure in forensic genetics.

shotgun sequencing An approach used to decode an organism's genome by cutting it into smaller fragments of DNA which can be sequenced individually. The sequences of these fragments are then ordered, based on overlaps in the genetic code, and finally reassembled into the complete sequence. The 'whole genome shotgun' method is applied to the entire genome all at once.

sickle cell disease A genetic disease in which abnormal β-globin is produced, seen most commonly in people of African ancestry. The disorder is caused by a single amino acid change. This mutation causes the protein to aggregate under low-oxygen conditions and the red blood cells take on a sickle shape, rather than their characteristic donut shape. Individuals who suffer from sickle cell disease are chronically anemic and experience significant damage to their heart, lungs, and kidneys.

sigma (σ) The prokaryotic RNA polymerase subunit that confers specificity of transcription; i.e. the ability to recognize specific gene promoters.

σ^{70} The principle sigma factor of *E. coli*.

signal peptide A stretch of about 20 amino acids, usually at the amino terminus of a polypeptide, that helps to anchor the nascent polypeptide and its ribosome in the endoplasmic reticulum. Polypeptides with a signal peptide are destined for packaging in the Golgi apparatus and are usually secreted from the cell.

signal recognition particle (SRP) Mediates protein targeting to the endoplasmic reticulum.

signal transduction pathway A biochemical pathway that connects a signal, such as a growth factor binding to the cell surface, with an intracellular effect, usually gene activation or repression.

silencer A DNA regulatory element that decreases gene promoter activity. Silencers are usually 700–1000 bp or more away from the start of transcription and can be downstream, upstream, or within an intron, and can function in either orientation relative to the promoter.

silent mutation Mutations that cause no detectable change in an organism, even in a haploid organism or in a homozygote.

single-copy DNA DNA sequences that are present once, or only a few times, in a haploid genome.

single nucleotide polymorphisms (SNPs) A single-nucleotide difference between two or more individuals at a particular genetic locus. Common, but minute, variations that occur in human DNA at a frequency of one every 1000 bases. SNP is pronounced "snip."

single-stranded DNA-binding protein (SSB) Binds to single-stranded DNA and keeps it from base pairing with a complementary strand, used during DNA replication.

SIR2, SIR3, SIR4 Proteins associated with, and required for, the formation of yeast heterochromatin, including at the telomeres.

siRNAs (short interfering RNAs) The short pieces (21–28 nt) of double-stranded RNA created by Dicer during the RNAi process.

site-directed mutagenesis A method for introducing specific, predetermined alterations into a cloned gene.

site-specific recombination Recombination that always occurs in the same place and depends on limited sequence similarity between the recombining DNAs.

7SL RNA A small RNA involved in recognizing the signal peptides on proteins destined for secretion.

sliding (protein) Helical movement of a DNA-binding protein along the DNA due to tracking along a groove of the DNA over short distances.

slipped structures An unusual secondary structure in DNA that can occur at tandem repeats in which misalignment of repeats leads to single-stranded loops of DNA.

small nuclear RNAs (snRNAs) A set of small uracil (U)-rich RNAs found in the nucleus, associated with proteins to form small nuclear ribonucleoproteins (snRNPs), which participate in splicing of pre-mRNAs.

small nucleolar RNAs (snoRNAs) A set of hundreds of small RNAs found in the nucleolus. A small subset of the snoRNAs, associated with proteins in small nucleolar ribonucleoproteins (snoRNPs), participate in processing and base modification of the large ribosomal RNA precursor. Pronounced "snow" RNA.

snoRNP An snoRNA with its associated proteins. Pronounced "snorp."

snRNP An snRNA with its associated proteins. Pronounced "snurp."

somatic cells All body cells, except the reproductive (sex) cells.

somatic mutation A mutation that affects only somatic cells, so it cannot be passed on to progeny.

SOS response The activation of a group of genes that helps *E. coli* cells respond to environmental stress such as chemical mutagens or radiation that can cause DNA damage.

Southern blot Transferring DNA fragments separated by gel electrophoresis to a support medium such as nitrocellulose or nylon membrane, in preparation for hybridization to a labeled probed. Used to identify and locate DNA sequences which are complementary to the probe.

spacer DNA DNA sequences found between, or sometimes within, repeated genes such as the ribosomal RNA genes.

spliced leader (SL) The independently synthesized 35 nt leader that is trans-spliced to surface antigen mRNA coding regions in trypanosomes.

spliceosome The large RNA–protein complex on which splicing of nuclear pre-mRNA precursors occurs; composed of five snRNPs and ~200 proteins.

splicing The process of linking together two RNA exons while removing the intron that lies between them.

splicing factors Proteins in addition to snRNPs that are essential for splicing nuclear pre-mRNAs.

spontaneous mutation A mutation that occurs as a result of natural processes in cells; e.g. DNA replication errors.

spore A specialized haploid cell formed sexually in plants or fungi, or asexually by fungi. The latter can either serve as a gamete, or germinate to produce a new haploid cell. In bacteria, the term refers to a specialized cell formed asexually in response to stressful conditions. Such a spore is resistant to environmental stress.

sporulation Formation of spores.

squelching Inhibition of one activator by increasing the concentration of a second one. This may occur through competition for a common factor that is at a limiting concentration.

SR proteins A group of RNA-binding proteins having an abundance of serine (S) and arginine (R).

stem cells Undifferentiated cells that can undergo unlimited division and can give rise to one or several different cell types.

sticky ends Single-stranded ends of double-stranded DNAs that are complementary and can therefore base pair and "stick" together by hydrogen bonding.

stop codon One of three codons (UAG, UAA, UGA) that code for termination of translation.

strand-coupled model An alternate model for mammalian mitochondrial DNA replication in which replication initiates at multiple sites and occurs by a semidiscontinuous, bidirectional mode of DNA replication.

strand-displacent model An alternate model for mammalian mitochondrial DNA replication in which replication is unidirectional around the circular DNA molecule and there is one replication fork for each strand. Also known as the strand-asynchronous model.

streptavidin A protein made by *Streptomyces* bacteria that binds to biotin.

stringency (of hybridization) The combination of factors (temperature, salt, and organic solvent) that influence the ability of two nucleic acid strands (DNA or RNA) to hybridize. At high stringency,

only perfectly complementary strands will hybridize. At reduced stringency, some mismatches can be tolerated.

SWR1 complex A chromatin remodeling complex family named after its ATPase subunit Swr1 (for Swi2/Snf2 related) that transfers H2A.Z-H2B dimers to nucleosomes in exchange for H2A-H2B dimers.

substitution Replacement of one nucleotide in a DNA sequence by another nucleotide or replacement of one amino acid in a protein by another amino acid.

suicide gene A strategy for making cancer cells more vulnerable to chemotherapy. One approach has been to link parts of genes expressed in cancer cells to other genes for enzymes not found in mammals that can convert a harmless substance into one that is toxic to the tumor.

SUMO (small ubiquitin-related modifier) A small polypeptide that can be attached to other proteins and modify their function.

sumoylation The attachment of SUMO to other proteins.

supercoiled DNA A circular DNA molecule or loop doman within linear DNA that is under torsional stress, and the double helix coils around itself like a twisted rubber band. The DNA has become either overwound (positive supercoiling) or underwound (negative supercoiling) with respect to the number of complete turns of the DNA double helix.

supershift The extra electrophoretic gel mobility shift observed when a new protein (such as an antibody) binds a protein–DNA complex.

suppression Compensation by one mutation for the effects of another.

suppressor mutation A mutation that reverses the effects of mutation in the same or another gene.

SV40 Simian virus 40. A DNA tumor virus (papovavirus) with a small circular genome, capable of causing tumors in certain rodents. Used extensively as a model organism for eukaryotic molecular biology.

SWI/SNF A family of coactivators that help to remodel chromatin by disrupting nucleosome cores.

syndactyly Webbing and/or fusion of the fingers and toes.

syndrome The group or recognizable pattern of symptoms or abnormalities that indicate a particular trait or disease.

synteny Preservation of gene order in different organisms.

systemic lupus erythematosus (SLE) An autoimmune disorder in which individuals form antibodies against their own proteins. Symptoms range from mild to severe and include swollen glands, joint pain, fatigue, skin rashes, damage to internal organs, etc.

T

T antigen See large T antigen.

T cell receptor Antigen-binding proteins on the surface of T cells. Composed of two heavy (β) and two light (α) chains.

t loop A loop formed in the telomere at the end of a eukaryotic chromosome.

T loop The loop in a tRNA molecule that contains the nearly invariant sequence TΨC, where Ψ is pseudouridine.

tandem mass spectrometry (MS/MS) A process in which peptide ions are bombarded with rare gas atoms which provide energy required to cleave amide bonds of the peptide backbone. The process produces a collection of peptide ion fragments that differ in mass by a single amino acid. Measurement of the mass-to-charge ratios of the fragments allows the amino acid sequence to be read.

tandem repeat Two or more adjacent, approximate copies of a pattern of nucleotides, arranged in a head to tail fashion (sometimes called a direct repeat).

Taq polymerase A heat-resistant DNA polymerase obtained from the thermophilic bacterium *Thermus aquaticus*.

target mRNA The mRNA that is targeted and degraded during RNAi.

TATA box An element with the consensus sequence TATAAA that begins about 30 bp upstream of the start of transcription in many eukaryotic promoters recognized by RNA polymerase II.

TATA-box binding protein (TBP) A subunit of the TFIID complex in RNA polymerase II transcription. Binds to the TATA box.

TBP-associated factor (TAF) A protein associated with TBP in the TFIID complex.

T-DNA The tumor-inducing part of the Ti plasmid.

telocentric chromosome A chromosome with the centromere at one end.

telomerase An enzyme that can extend the ends of telomeres after DNA replication. The protein subunit (TERT) has reverse transcriptase activity, and the RNA component (TERC) serves as the template for synthesis of new telomere repeats.

telomerase reverse transcriptase (TERT) The protein component of the enzyme telomerase.

telomerase RNA component (TERC) The RNA component of the enzyme telomerase.

telomere Specialized structures that cap the end of chromosomes and prevent them from being joined to each other. Telomeres are composed of tandem repeats of a simple guanine (G)-rich sequence (e.g. in humans, several thousand repeats of the sequence TTAGGG).

telomere shortening The progressive shortening of the ends of chromosomes in cells without telomerase activity.

telophase The last stage of cell division, in which the nuclear membrane forms and encloses the chromosomes in the daughter cell.

temperature-sensitive mutation A mutation that causes a product to be made that is defective at high temperature (the nonpermissive temperature) but functional at low temperature (the permissive temperature).

template A polynucleotide (RNA or DNA) that serves as a guide for making a complementary polynucleotide.

template DNA strand The DNA strand of a gene that is complementary to the RNA product of the gene; that is, the strand that serves as the template for making the RNA. Also called the anticoding strand or antisense strand.

teratogen A substance that causes abnormal development of an organism.

terminal transferase An enzyme that adds deoxynucleoside triphosphophates, one at a time to the 3′ end of DNA.

terminal uridylyl transferase (TUTase) An enzyme that adds UMP residues to pre-mRNAs during RNA editing in trypanosomes.

tertiary structure The overall three-dimensional shape of a polypeptide or RNA.

tetraloop An RNA stem-loop structure that is stabilized due to special base-stacking interactions in the loop.

tetramer (protein) A complex of four polypeptides.

thermal cycler A machine that performs PCR reactions automatically by repeatedly cycling among the three temperatures required by DNA denaturation, primer annealing, and primer extension.

theta (θ) structure A replication intermediate formed during replication of circular double-stranded DNA in bacteria.

three (3)-prime (3′) The ends of a DNA or RNA chain are chemically distinct and are designated by the symbols 3′ and 5′. The symbol 3′ refers to the carbon in the ribose or deoxyribose sugar ring to which a hydroxyl (OH) functional group is attached.

thymidine A nucleoside containing the base thymine.

thymine (T) One of the four bases in DNA. Thymine always pairs with adenine.

thymine dimer Two adjacent thymines in one DNA strand linked covalently, whose base pairing with adenines in the opposite strand is interrupted; the most common form of DNA damage induced by UV irradiation.

thyroid hormone receptor A nuclear receptor that in the presence of thyroid hormone either activates or respresses the expression of thyroid hormone-responsive genes.

Ti plasmid The tumor-inducing plasmid from *Agrobacterium tumefaciens*. Used as a vector to carry foreign genes into plant cells.

toe print assay A primer extension assay that locates the edge of a protein (e.g. the ribosome) bound to DNA or RNA.

topoisomer A form of DNA that has the same sequence as another DNA molecule yet differs in its linkage number. Also known as a topological isomer.

topoisomerase An enzyme that converts (isomerizes) one topoisomer of DNA to another.

***trans*-acting** A term that describes a genetic element, such as a repressor gene or transcription factor gene, that can be on a separate chromosome and still influence another gene. These *trans*-acting genes function by producing a diffusible substance that can act at a distance.

transactivation domain The part of a transcription factor that stimulates transcription by protein–protein interactions.

transcribed spacer A region encoding a part of an rRNA precursor that is removed during processing to produce the mature rRNAs.

transcript An RNA copy of a gene.

transcription The process by which an RNA copy of a gene is made.

transcription bubble The region of locally melted DNA that follows the RNA polymerase as it synthesizes the RNA transcript.

transcription-coupled repair Nucleotide excision repair pathway responsible for recognizing DNA damage in the transcribed strand of active genes.

transcription factors Sequence-specific DNA binding proteins that bind to gene promoters and long-range regulatory elements and mediate gene-specific transcriptional activation or repression.

transcription terminator A specific DNA sequence that signals transcription to terminate.

transcription unit A region of DNA bounded by a promoter and a terminator that is transcribed as a single unit; it may be polycistronic and contain multiple coding regions.

transcriptional attenuation A regulatory process in some bacterial biosynthetic operons in which translation of a leader sequence in mRNA terminates transcription at a terminator hairpin instead of at the competing antiterminator hairpin.

transcriptome The sum of all the different transcripts an organism can make in its lifetime.

transcriptomics The global study of an organism's transcripts.

transesterification A reaction that simultaneously breaks one ester bond and creates another; for example, the reactions that take place during RNA splicing.

transfection Introduction of foreign DNA into eukaryotic cells, either for a short duration (transient) or for long-term analysis (stable integration).

transfer RNA (tRNA) A relatively small RNA molecule that is "charged" with an amino acid and delivers to the ribosome the appropriate amino acid via interaction of the tRNA anticodon with the mRNA codon.

transferrin An iron-carrier protein that imports iron into cells via the transferrin receptor.

transformation The introduction of foreign DNA into bacterial cells.

transgene A foreign gene transferred into an organism, making the recipient a transgenic organism.

transgenic organism An experimentally produced organism that carries transferred genetic material (the transgene) that has been inserted into its genome at a random site, usually by injecting the foreign DNA into the nucleus of a fertilized embryo.

transition A mutation in which a pyrimidine replaces a pyrimidine, or a purine replaces a purine.

translation The process by which ribosomes decode an RNA message (mRNA) to synthesize a protein.

translesion synthesis (TLS) A mechanism for bypassing DNA damage by replicating through it; mediated by specialized low-fidelity, "error-prone" DNA polymerases.

translocation (chromosome) Breakage and removal of a large segment of DNA from one chromosome, followed by the segment's attachment to a different chromosome.

translocation (protein synthesis) The translation elongation step, following peptide bond formation, that involves moving an mRNA one codon length through the ribosome and bringing a new codon into the ribosome A site.

transmissible spongiform encephalopathies A rare brain-wasting disease in mammals caused by infectious prions and characterized by sponge-like holes in the brain, dementia, and loss of muscle control of voluntary movements (ataxia).

transmission genetics The study of the transmission of genes from one generation to the next.

transposable element A DNA element that can move from one genomic location to another.

transposase The name for a collection of enzymes, encoded by a transposon, that catalyze transposition.

transposition The movement of a DNA element from one DNA location to another.

transposon See transposable element and DNA transposons.

***trans*-splicing** Splicing together two RNA fragments transcribed from separate transcription units.

transversion A mutation in which a pyrimidine replaces a purine, or vice versa.

trichothiodystrophy An autosomal recessive disorder in which individuals have an exaggerated sensitivity to light, brittle hair and nails, and dry scaly skin; linked to defects in nucleotide excision repair genes.

trinucleotide repeats Repetitive regions of three bases of DNA with unusual gene instability. Dynamic expansion of trinucleotide repeats (e.g. GAAGAAGAA) leads to certain genetic neuorological disorders such as fragile X syndrome, Huntington's disease, Friedreich's ataxia, etc.

triple helix DNA A secondary structure that can form at purine-pyrimidine stretches in DNA with mirror repeat symmetry in which a third strand of DNA joins the first two to form a triplex DNA.

trisomy Possessing three copies of a particular chromosome instead of the normal two copies.

tritium A radioactive isotope of hydrogen (^{3}H).

tRNA charging The process of coupling tRNA with its cognate amino acid, catalyzed by aminoacyl-tRNA synthetase.

tRNA endonuclease The enzyme that cuts an intron out of a tRNA precursor.

$tRNA_f^{Met}$ The tRNA responsible for initiating protein synthesis in prokaryotes.

$tRNA_i^{Met}$ The tRNA responsible for initiating protein synthesis in eukaryotes.

$tRNA_m^{Met}$ The tRNA that inserts methionines that are coded for within a protein sequence.

***trp* operon** The operon that encodes the enzymes needed to make the amino acid tryptophan.

trypanosomes Protozoa that parasitize both mammals and tsete flies; the latter spread the disease by biting mammals; infection leads to the disease called sleeping sickness.

TTAGG Repeat Binding Factor 1 and 2 (TRF1, TRF2) Proteins involved in regulation of telomerase activity; they count the number of G-rich repeats and when telomeres become overly long, inhibit further telomerase activity by transferring POT1 to the single-stranded overhang at the telomere tip.

tumor suppressor gene A protective gene that normally limits the growth of tumors. When a tumor suppressor is mutated and undergoes a loss of function, it may fail to keep a cancer from growing. BRCA1, p53, and the retinoblastoma protein are well-known tumor suppressor genes.

turns (protein) Relatively short loops of amino acids that do not exhibit a defined secondary structure themselves, but are essential for the overall folding of a protein.

two-dimensional gel electrophoresis A high-resolution method for separating proteins. The proteins are separated in the first dimension by isoelectric focusing. Then they are separated in the second dimension by SDS-PAGE.

U

U1 snRNP The first snRNP that recognizes the 5′-splice site in a nuclear pre-mRNA.

U2-associated factor (U2AF) A splicing factor that helps recognize the correct AG at the 3′-splice site by binding to both the polypyrimidine tract in the 3′-splice signal and the AG; composed of two subunits, U2AF35 and U2AF65.

U2 snRNP The snRNP that recognizes the branchpoint in a nuclear pre-mRNA.

U4 snRNP The first snRNP whose RNA base pairs with the RNA in the U6 snRNP until U6 snRNP is needed to splice a nuclear pre-mRNA.

U4atac A minor snRNA that participates in splicing variant introns and plays the same role as U4 snRNA.
U5 snRNP The snRNP that associates with both the 5′ and 3′ exon–intron junctions, thus helping to bring the two exons together during nuclear pre-mRNA splicing.
U6 snRNP The snRNP whose RNA base pairs with both the 5′ splice site and with the RNA in U2 snRNP in the spliceosome during nuclear pre-mRNA splicing.
U6atac A minor snRNA that participates in splicing variant introns and plays the same role as U6 snRNA.
U11 snRNA A minor snRNA that participates in splicing variant introns and plays the same role as U1 snRNA.
U12 snRNA A minor snRNA that participates in splicing variant introns and plays the same role as U2 snRNA.
ubiquitin A small polypeptide that can be attached to proteins and modifies their function. Polyubiquitinylation marks a protein for destruction by the proteasome.
ubiquitinylation The addition of ubiquitin to a protein. Sometimes called ubiquitination.
ultraviolet (UV) radiation Radiation found in sunlight. Can induce pyrimidine dimers in DNA.
undermethylated region A region of a gene or its flanking region that is relatively poor in, or devoid of, methyl groups. Also known as hypomethylation.
unidirectional DNA replication Replication that occurs in one direction, with only one active replication fork.
untranslated region (UTR) A region at the 5′ or 3′ end of an mRNA that lies outside the coding region, so it is not translated (i.e. before a start codon or after a stop codon).
upstream DNA sequence before the start of transcription (+1) in the direction of the 5′ end; numbered with negative numbers (e.g. –35).
upstream activating sequence (UAS) An enhancer for yeast genes.
uracil (U) One of the four bases in RNA. Uracil replaces thymine, which is the fourth base in DNA. Like thymine, uracil always pairs with adenine.
uridine A nucleoside containing the base uracil.

V

variable number tandem repeat (VNTR) A type of restriction fragment length polymorphism (RFLP) that includes tandem repeats of a minisatellite between the restriction sites.
variable region (V) The region of an antibody that binds specifically to a foreign substance or antigen. It varies considerably from one specific antibody to another.
variable surface glycoprotein (VSG) The antigen that coats a trypanosome. The parasite can vary the nature of this coat by switching which VSG gene is expressed in a telomere expression site.
V(D)J joining The assembly of active immunoglobulin or T-cell receptor genes by recombination involving separate V and J or V, D, and J segments in the embryonic genes.
vector DNA An agent, such as a virus or a plasmid, phage, or artificial chromosome, that carries a modified or foreign gene. Used as a carrier in gene cloning experiments. When used in gene therapy, a vector delivers the desired gene to a target cell.
vegetative cell A cell that is reproducing by division, rather than sporulating or reproducing asexually.
viroid A subviral RNA pathogen that causes infectious disease in higher plants.
virus A minute intracellular obligate parasite. A virus particle consists of a core of nucleic acid, which may be DNA or RNA, surrounded by a protein coat, and in some viruses a further lipid/glycoprotein envelope. It is unable to multiply or express its genes outside a host cell as it requires host cell enzymes to aid DNA replication, transcription, and translation.
virusoid A subviral pathogen that causes infectious disease in higher plants. It consists of an RNA molecule that does not encode any proteins and depends on a helper virus for replication.
VS ribozyme Varkud Satellite RNA. A small ribozyme transcribed from plasmid found in the mitochondria of some strains of *Neurospora crassa*, a filamentous fungus.

W

Watson–Crick base pairing See base pairing.

Western blot Electrophoresing proteins, then blotting them to a membrane and reacting them with a specific antibody or antiserum. The antibody is detected with a labeled secondary antibody. This technique is used to identify and locate proteins based on their ability to bind to specific antibodies.

wobble The ability of the third base of a codon to shift slightly to form a non-Watson–Crick base pair with the first base of an anticodon, thus allowing a tRNA to translate more than one codon.

wobble base pair A base pair formed by wobble (e.g. a G-U or A-I base pair).

wobble position The third base of a codon, where wobble base pairing is permitted.

wyosine A highly modified guanine nucleoside found in tRNA.

X

X chromosome inactivation The process by which one of the two X chromosomes in females is randomly inactivated in each cell.

X inactivation specific transcript (XIST) An RNA transcript that is expressed from and coats the inactive X chromosome.

xeroderma pigmentosum (XP) A disease characterized by extreme sensitivity to sunlight. Caused by a defect in nucleotide excision repair genes (XPA-XPG) or translesion synthesis (XP-V).

X-ray crystallography A method for determining the three-dimensional structure of molecules by measuring the diffraction of X-rays by crystals of a molecule or molecules.

X-ray diffraction See X-ray crystallography.

X-rays High-energy radiation. X-rays can ionize cellular components and can cause DNA double-strand breaks.

Y

yeast artificial chromosome (YAC) Extremely large segments of DNA from another species spliced into artificial chromosomes of yeast. YACs are used to clone up to one million bases of foreign DNA into a host cell, where the DNA is propagated along with the host cell chromosomes.

yeast mating type switching A simple form of cellular differentiation in which two mating types are defined by the expression of one of two gene cassettes; switching occurs through DNA rearrangement by homologous recombination.

yeast two-hybrid assay An assay for interaction between two proteins. One protein (the bait) is produced from a fusion protein with a DNA-binding domain from another protein. The other protein (the prey) is produced from a fusion protein with a transactivation domain. If the two fusion proteins interact within a yeast cell, they form a transcriptional activator that can activate one or more reporter genes.

Z

Z-DNA An alternate left-handed form of the DNA double helix with 12 base pairs per turn whose backbone has a zig-zag appearance. This form is stabilized by stretches of alternating purines and pyrimidines. Short sections of Z-DNA may function as regulatory elements within cells.

zinc finger A DNA-binding motif that contains a zinc ion complexed to four amino acid chains, usually the side chains of two cysteines and two histidines, or four cysteines. The motif is roughly finger-shaped and inserts into the DNA major groove, where it makes specific protein–DNA contacts.

zygote cell formed from the union of two gametes or reproduction cells.