Gene speak

Allele Alternative version of a particular gene. Humans carry two sets of most genes, one inherited from each parent, so a single individual may carry either two of the same, or two different alleles at a given locus, each one inherited separately from one parent.

Amino acid Any of a class of 20 molecules that are combined to form proteins.

Autosome A chromosome not involved in sex determination. The diploid human genome contains 22 pairs of autosomes, and 1 pair of sex chromosomes. Compare with **sex chromosome**.

Base pair (bp) Two nitrogenous bases (adenine and thymine or guanine and cytosine) held together by weak bonds. See **DNA**, **nucleotide**.

Bioinformatics The study of genetic and other biological information using computer and statistical techniques. In genome projects, bioinformatics includes the development of methods to search databases, analyse DNA sequence information, and predict protein sequence and structure from DNA sequence data.

Biotechnology The industry that sprang up in the 1970s from the discovery of bacterial enzymes that cut DNA at specific sites in the genome and rejoin them, enabling human genes, for example, to be 'cloned' in bacteria, and the protein product isolated and purified in large quantities.

Centromere The compact region at the centre of a chromosome.

Chromosome A rod-shaped structure inside the nucleus of a cell which contains a densely packed continuous strand of DNA. Different organisms have different numbers of chromosomes. The diploid human genome consists of 23 pairs of chromosomes, 46 in all: 22 pairs of autosomes and 2 sex chromosomes. See **autosome**, **sex chromosome**.

Clone An exact copy made of biological material, such as a DNA segment (a gene or other region).

Cloning The process of generating multiple, exact copies of a particular piece of DNA to allow it to be sequenced or studied in some other way.

Conserved sequence A sequence of DNA (or an amino acid sequence in a protein) that has remained essentially unchanged throughout evolution, usually because of functional constraints.

DNA (deoxyribonucleic acid) The molecule that encodes genetic information. The four nucleotides in DNA contain the bases: adenine (A), guanine (G), cytosine (C) and thymine (T). Two strands of DNA are held together in the shape of a double helix by bonds between base pairs of nucleotides, where A pairs with T and G with C. See **nucleotide**, **base pair**.

Diploid A full set of genetic material, consisting of paired chromosomes, one from each parental set. Most animal cells except the gametes have a diploid set of chromosomes. Compare with **haploid**.

Draft sequence DNA sequence in which the order of bases is sequenced at least four to five times (an accuracy of 99.9%), which enables the reassembling of DNA fragments in their original order. Some segments can be missing or in the wrong order or orientation. Compare with **finished sequence**.

Enzyme A protein that specifically catalyses reactions between biological molecules.

Euchromatin The gene-rich regions of a genome. Compare with **heterochromatin**.

Eukaryote An organism whose cells have a complex internal structure, including a nucleus. Animals, plants and fungi are all eukaryotes. Compare with **prokaryotes**.

Excision The process by which enzymes cut out and remove a portion of DNA, for example one that is recognized as containing a mutation.

Exon The protein-coding DNA sequence of a gene. Compare with **intron**.

Finished sequence DNA sequence in which bases are identified to an accuracy of 99.99% and are placed in the right order and orientation along a chromosome with almost no gaps.

FISH (fluorescence *in situ* hybridization) A process that vividly paints chromosomes or portions of chromosomes with fluorescent molecules. This technique is useful for identifying chromosomal abnormalities and gene mapping.

Gamete Mature male or female reproductive cell (sperm or egg) with a **haploid** set of chromosomes.

Gene The fundamental physical and functional unit of heredity. A gene is an ordered sequence of nucleotides located at a particular position on a given chromosome that encodes a specific functional product.

Genetic code The sequence of nucleotides, coded in triplets (codons) along the messenger RNA, that determines the sequence of amino acids in protein synthesis. The DNA sequence of a gene can be used to predict the messenger RNA sequence, and the genetic code can in turn be used to predict the amino acid sequence.

Genome The complete genetic material of an organism; the entire DNA sequence.

Genomics The study of genomes and their sets of genes.

Genotype The set of genes that an individual carries; usually refers to the particular pair of **alleles** that a person has at a given region of the genome. Genotype refers to what is inherited (e.g. an allele for brown eyes), whereas **phenotype** refers to what is expressed (brown eyes in this case).

Haploid A single set of chromosomes (half the full set of genetic material) present in the egg and sperm cells of animals and in the egg and pollen cells of plants. Compare with **diploid**.

Haplotype A particular combination of **alleles** or sequence variations that are closely linked – that is, are likely to be inherited together – on the same chromosome.

Heterochromatin Compact, gene-poor regions of a genome, with abundant simple sequence repeats. Compare with **euchromatin**.

Intron The DNA sequence interrupting the protein-coding sequence of a gene; this sequence is transcribed into RNA but is cut out before the RNA is transcribed. Compare with **exon**.

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

Kilobase (kb) A unit of length of DNA fragments equal to 1000 nucleotides.

Library A unordered collection of clones whose relationship to each other can be established by physical mapping.

Linkage The proximity of two or more markers (e.g. genes) on a chromosome.

Marker An identifiable physical location or landmark on a chromosome (e.g. a restriction enzyme cleavage site) whose inheritance through generations can be monitored.

Megabase (Mb) Equal to 1 million nucleotides.

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

mRNA (messenger RNA) The form of RNA that serves as a template for protein synthesis.

Mitosis The process of nuclear division in cells that produces two daughter cells genetically identical to each other and to the parent cell.

Mutation An alteration in the sequence of DNA, for example the substitution of one nucleotide base for another.

Nucleotide Subunit of DNA or RNA comprising a nitrogenous base (adenine, guanine, thymine or cytosine in DNA; adenine, guanine, uracil or cytosine in RNA), a phosphate molecule and a sugar molecule (deoxyribose in DNA and ribose in RNA). Nucleotides are linked to form the strands of a DNA or RNA molecule. Short stretches of nucleotides are called oligonucleotides. See **base pair**.

PCR The polymerase chain reaction: a technique that produces millions of copies of a short stretch of DNA in a matter of hours, providing enough material for the detection, for example, of a single mutation in a biological specimen such as a single cell taken from a human embryo.

Peptide A sequence of amino acids that is shorter than a protein.

Phenotype The observable properties and physical characteristics of an organism.

Polymorphism A difference in DNA sequence between individuals. To be called a polymor-

phism, a variant should be present in a significant number of people in the population.
Prokaryote A cell or organism lacking a membrane-bound, structurally discrete nucleus and other subcellular compartments. Bacteria are prokaryotes. Compare with **eukaryote**.
Protein A large molecule composed of one or more chains of amino acids in a specific order; the order being determined by the base sequence of nucleotides in the gene coding for the protein.
Recombinant DNA A combination of DNA molecules of different origin that are joined using recombinant DNA technologies.
Recombination The process by which DNA is exchanged between pairs of equivalent chromosomes during egg and sperm formation. Recombination has the effect of making the chromosomes of the offspring distinct from those of the parents.
Replication The process of copying DNA prior to each cell division.
RNA (ribonucleic acid) A molecule with a similar structure to DNA that plays an important role in protein synthesis and other chemical activities of the cell. There are several types of RNA molecules, including **mRNA**, which acts as an intermediary molecule between DNA and protein.
Sex chromosome The X or Y chromosome in humans. Determines the sex of an individual. Females have two X chromosomes in diploid cells; males have an X and a Y chromosome. The sex chromosomes comprise the 23rd chromosome pair of a human genome. Compare with **autosome**.
SNP (single nucleotide polymorphism) A polymorphism caused by the change of a single nucleotide. Most genetic variation between individuals is due to SNPs.
Space group The symmetry arrangement of the **unit cell** (the basic repeating unit) of a crystal. For DNA, the space group is known as C2, or 'face-centred monoclinic', meaning that the unit cell ('monoclinic' in shape) contains two components (the two sugar-phosphate backbones) that look the same after rotation through 180 degrees about an axis which is at right angles to the direction of the DNA molecule.
Splicing The process that removes **introns** (non protein-coding portions) from transcribed **mRNA**. **Exons** (protein-coding portions) can also be removed. Depending on which exons are removed, different proteins can be made from the same initial RNA or gene. Different proteins created in this way are 'splice variants' or 'alternatively spliced'.
Transcription The process of copying a gene into **mRNA**. This is the first step by which a gene dictates the production of a protein, although not all transcripts generate proteins. Compare **translation**.
Translation The process of using an **mRNA** sequence to build a protein. The **mRNA** serves as a template on which transfer RNA molecules, carrying amino acids, are lined up. The amino acids are then linked together to form a protein chain.
Unit cell The smallest possible volume of a crystal that can be repeated infinitely in three dimensions. The atoms within the unit cell have a fixed geometry relative to each other, for example representing the single turn of the double helix of DNA. The repetition of unit cells intensifies the **X-ray diffraction** pattern to a detectable level. For DNA the unit cell is described as 'monoclinic', meaning that the arrangement of atoms has a squashed box-like shape in which the sides have different lengths in each dimension, and the corner angle of one dimension is greater than a 90 degrees right angle.
X-ray diffraction The scattering of X-rays by the electron clouds of atoms. The resulting pattern of spots, detected by an X-ray sensitive surface, can be recorded as a two-dimensional electron density map. By rotating the fibre or crystal specimen, the two-dimensional maps created from them can be combined to produce a set of three-dimensional spatial coordinates, revealing the positions of atoms within the **unit cell** of the specimen.

Bibliography

Crick, Francis Harry Compton, *What Mad Pursuit: A Personal View of Scientific Discovery* (New York: Basic Books division of Harper Collins, 1988)
– 'The Double Helix: A Personal View', *Nature*, 248: 766–9, 1974
Delhanty, Joy D.A., and Wells, Dagan, 'Preimplantation Genetic Diagnosis: Applications for Molecular Medicine', *Trends in Molecular Medicine*, 7: 23–30, 2001
Friedman, Jeffrey M., 'Obesity in the New Millennium', *Nature*, 404: 632–4, 2000
Thangaraj, Kumarasamy, *et al.*, 'Genetic Footprints of the Andaman Islanders', *Current Biology*, 10 December 2002
Hendrickson, Wayne, 'Synchrotron Crystallography', *Trends in Biochemical Sciences*, 25: December 2000
Judson, Horace Freeland, *The Eighth Day of Creation: Makers of the Revolution in Biology* (London: Jonathan Cape, 1979, Penguin, 1995; expanded edition: New York: Cold Spring Harbor Laboratory Press, 1996)
Klug, Aaron, 'Rosalind Franklin and the Discovery of the Structure of DNA', *Nature*, 219: 808–10, 1968
– 'Rosalind Franklin and the Double Helix', *Nature*, 248: 787–8, 1974
Maddox, Brenda, *Rosalind Franklin: The Dark Lady of DNA*, HarperCollinsPublishers, 2002
Nowlan, William, 'A Rational View of Insurance and Genetic Discrimination', *Science*, 297: 195–6, 2002
Pauling, Linus, 'Molecular Basis of Biological Specificity', *Nature*, 248: 769–71, 1974
Perutz, Max F., correspondence about the MRC report of December 1952, *Science*, 164: 1537–8, 1969
Rosenberg, N., *et al.*, 'Genetic Structure of Human Populations', *Science*, 298: 2381–5 2002
Rothenberg, Karen, H., and Terry, Sharon F., 'Before It's Too Late – Addressing Fear of Genetic Information', *Science*, 297: 196–7, 2002
Sulston, John, and Ferry, Georgina, *The Common Thread: A Story of Science, Politics, Ethics and the Human Genome* (London, Sydney, Auckland, Parktown: Bantam Press, 2002)
The Mouse Genome special supplement, *Nature*, 420: 509–80, 2002
The Human Genome special supplement, *Nature*, 409: 813–958, 2001
Thomas, Mark, *et al.*, 'Y Chromosomes Travelling South: The Cohen Modal Haplotype and the Origins of the Lemba – The "Black Jews of Southern Africa"', *American Journal of Human Genetics*, 66: 674–86, 2000
Venter, J. Craig, *et al.*, *Science*, 291: 1304–51, 2001
Watson, James D., *The Double Helix: A Personal Account of the Discovery of the Structure of DNA* (New York: Athenaeum, 1968; London: reissued with an introduction by Steve Jones, London: Weidenfeld and Nicolson, 1997, Penguin Books, 1999)
– *Genes, Girls and Gamow* (Oxford: Oxford University Press, 2001)
– *A Passion for DNA: Genes, Genomes and Society* (New York: Cold Spring Harbor Laboratory Press, 2000; Oxford: Oxford University Press, 2000)
Weale, Michael E., *et al.*, 'Y Chromosome Evidence for Anglo-Saxon Mass Migration', *Molecular Biology and Evolution*, 19 (7): 1008–21, 2002

Useful websites

Nobel Prizes: http://www.nobel.se
http://www.mendel-museum.org
DNA from the Beginning: http://www.DNAftb.org
DNA Interactive: http://www.dnai.org
BBC Gene Stories: http://www.bbc.co.uk/print/health/genes
The X-ray Century: http://www.emory.edu/X-RAYS/century_06.html
Linus Pauling Institute: http://lpi.oregonstate.edu/
Blazing a Genetic Trail: a report from the Howard Hughes Medical Institute http://www.hhmi.org/genetictrail/h100.html

Index