

# Glossary of terms

provided by J. E. SWINBURNE and S. BLOTT to accompany

## Genes and respiratory disease: a first step on a long journey

S. L. EWART and N. E. ROBINSON\*

*Department of Large Animal Clinical Sciences, College of Veterinary Medicine, Michigan State University, East Lansing, Michigan 4882-1314, USA.*

**Allele:** At any one gene an individual inherits one paternal copy and one maternal copy. These 2 copies are referred to as alleles of the gene. Different alleles may be present in the population; individuals carrying 2 copies that are the same are referred to as 'homozygous' while individuals carrying 2 different copies are 'heterozygous'.

**Amplicon:** The amplified DNA product of PCR.

**Association analysis:** An analysis that looks for a significant statistical relationship between a marker allele in the population and a particular trait of interest. Markers that show significant association indicate the position of the causal gene.

**BAC 'Bacterial Artificial Chromosome' clone:** A vector used to clone DNA fragments (100–300 kb insert size; average, 150 kb) in *Escherichia coli* cells. Based on naturally occurring F-factor plasmid found in the bacterium *E. coli*.

**Candidate gene:** A gene known to be located in the region of interest whose product has biochemical properties, or is implicated in a related disease aetiology, suggesting that it may prove to be the disease gene being sought.

**CentiMorgan (cM):** A unit of measure of genetic recombination frequency. 1 cM corresponds to a 1% frequency of recombinant progeny. In the human, and other mammals of a similar genome size, 1 cM ~ 1 Mb (10<sup>6</sup> bp).

**Chromosome-wise significance threshold:** The level of significance of a mapping test statistic is determined according to whether the whole genome or a single chromosome is being scanned. The corresponding significance thresholds are designated 'genome-wise' or 'chromosome-wise'. The P value associated with the observed test statistic is determined using *permutation*, which is a computer-based method that re-samples from the data, randomly matching genotypes with phenotypes, to generate the null distribution of the test statistic.

**FISH 'Fluorescence In Situ Hybridisation':** A molecular microscopy technique used in chromosome studies. FISH employs fluorescent dyes that glow under ultraviolet light to locate hybridised molecular probes on specific chromosomes or chromosome regions. FISH thus vividly paints chromosomes or

portions of chromosomes with fluorescent molecules. This technique is useful for identifying chromosomal abnormalities and in gene mapping. For example, a FISH probe to chromosome 21 permits one to 'fish' for cells with trisomy 21, an extra chromosome 21, which is the cause of Down's syndrome.

**G-banding:** Technique for producing banding patterns on eukaryotic chromosomes. Staining with Giemsa stain, after pretreating chromosomes with trypsin, produces bands. Each homologous chromosome pair has a unique pattern of g-bands, enabling the recognition of particular chromosomes using microscopy.

**Genotype:** Used as a noun, the combination of alleles carried by an individual or as a verb, the process of identifying the alleles carried by the individual.

**Half-sib family:** A family of individuals who share one common parent (usually the sire). Large paternal half-sib families are commonly found in animal populations and are advantageous for genetic mapping.

**Haplotype:** The combination of alleles at adjacent loci on either the paternal or maternal chromosome inherited by an individual.

**HOARSI (Horse Owner Assessed Respiratory Signs Index):** An owner assessment of clinical symptoms of RAO, reported via a questionnaire, which is used to assign an index of 1, 2, 3 or 4 representing healthy, mild, moderate and severe symptoms.

**Linkage analysis:** An analysis based on families that tracks the inheritance of markers within those families and identifies co-inheritance of markers with the trait of interest. Markers that are linked to the trait indicate the chromosome and position of the causal gene.

**Locus (sing.)/Loci (pl.):** The position of a gene or segment of DNA on a chromosome. The term may be used to describe a gene, a part of a gene, or a DNA segment located between genes.

**Locus heterogeneity:** The gene or locus underlying a trait is different between individuals.

**Marker:** A segment of polymorphic DNA whose position in the

genome is known. Panels of markers are commonly used to identify the location of a gene causing an inherited trait.

*Microsatellite markers:* Areas of tandemly repeated DNA within the genome where the repeating unit is very small, usually 1–6 nucleotides. These are numerous and usually polymorphic within a population and can therefore be used for forensics, identity testing, paternity testing, genetic mapping etc.

*Nick translation:* The use of enzymes to break DNA and repolymerise small sections of the molecule, usually for the purpose of labelling the DNA with a radioactive or biochemically tagged nucleotide.

*PCR 'Polymerase Chain Reaction':* A technique for amplifying DNA, making it easier to isolate, clone and sequence.

*Phenotype:* The expressed characteristics or appearance of an individual.

*Polygenic disorder or Complex trait:* A characteristic that is influenced by multiple genes and the environment - the genes underlying these traits are sometimes referred to as *quantitative trait loci (QTL)*.

*Polymorphism:* Genetic variation at a gene; the existence of more than one allele (or genetic variant) in the population.

*Primer:* A short synthetic DNA sequence from which DNA replication can initiate.

*Recombination events:* During meiosis the maternal and paternal chromosomes of an individual undergo crossing-over. Exchanges are made between the 2 chromosomes so that the single copy of the chromosome inherited by the offspring is a mix of the paternal and maternal grandparents' genes. Each episode of crossing-over

where genes are exchanged along the chromosome is referred to as a recombination.

*Regression Interval Mapping:* A method of linkage analysis for complex traits or QTL that uses half-sib families and marker information to test each interval (between flanking markers) in the genome for presence of a QTL. Heterozygous markers in the sire are tested to see whether there is a significant difference in mean trait values between progeny inheriting one allele vs. progeny inheriting the other allele. The test statistic generated is an *F-statistic*.

*Single nucleotide polymorphism:* a type of genetic marker characterised by different nucleotides at a single base pair.

*Short-arm/long-arm of chromosome:* Each chromosome has a p and q arm; p (petit) is the short arm and q is the long arm. A region known as the centromere, which is a pinched region of the chromosome, separates the arms.

*Somatic cell hybrid panel:* A panel of cell lines derived usually from a rodent, which carry fragments of chromosomes, generated by radiation, from a second species. Genetic markers which are present on the same chromosome segment will amplify, using PCR, in the same subset of cell lines. Using these data new genetic markers can be assigned to specific chromosomes.

*Syntenic:* Located on the same chromosome segment.

*Syntenic mapping:* Assigning markers to a chromosome by means of for example, a somatic cell hybrid panel.

*Taq polymerase:* The thermostable DNA polymerase from *Thermus aquaticus (Taq)* has been the most extensively used enzyme in PCR. *T. aquaticus* was first isolated from a hot spring in Yellowstone National Park.