GLOSSARY

Basic glossary on genetic epidemiology

N Malats, F Calafell

....

This is the second of a series of three glossaries on genetic concepts used in epidemiological research that the journal is publishing with the objective of helping the reader "walk" around the journal.

....

The first glossary, on basic molecular genetic terms,[1] provided the basis to understand the concepts presented here.

In this section, we refer to the most basic and commonly used genetic terms that epidemiologists interested in this kind of research need to know. The concepts defined here are the pillars on which genetic epidemiology builds up its methodology.

Again, the list is not exhaustive and an attempt has been made to provide you with concise definitions. Hence, we encourage the interested reader to further "explore" the classic bibliography on genetics and genetic epidemiology[2–5] to complete their knowledge on this topic.

CANDIDATE GENE
A known *gene* suspected to be associated with the disease of interest on the basis of the biological function of its protein.

COMPLEX TRAIT (POLYGENIC TRAIT, MULTIFACTORIAL TRAIT)
Any *phenotype* that results from the effect of multiple *genes* at two or more loci, with possible environmental influences too. Examples are: obesity, hypertension, hypercholesterolaemia, skin pigmentation, cancer, etc.

Discontinuous trait
Trait that is either present or absent, such as birth defects and common behavioural disorders. The threshold model is used to explain discontinuous traits: a protein level has a continuous distribution but the *phenotype* does not appear until a certain threshold is reached.

Continuous trait
Measurable trait that is always present and that follows a normal distribution in the population. For example: height, weight, and blood pressure.

CO-SEGREGATION
The tendency of two traits to be jointly inherited.

DNA REPAIR
Major cell defence system against DNA damage produced by environmental and endogenous compounds. There are several different repair pathways and several enzymes (some of them *polymorphic*) involved in each way. Abnormalities in these processes have been implicated in cancer and aging.

EPISTASIS
Gene interaction and, particularly, interaction between different alleles at different genes. Epistasis can occur at the same step or at different stages of the same biochemical pathway.

FAMILIAL AGGREGATION
A tendency of a disease to cluster in families, which is generally taken as evidence for the existence of a genetic aetiological mechanism, or environmental factors common to family members, or a combination of both. Ascertainment bias should be seriously considered.

FAMILIAL ANTICIPATION
The younger age of appearance of a late onset trait in successive generations. A typical effect in *repeat expansions*, in which severity of the disease is proportional to repeat length, which tends to grow in each transmission.

FOUNDER EFFECT
A change in the population *allele* frequency that occurs when a subpopulation is established by a small number of individuals. The change occurs only by chance because the members of the new population are a random subsample that may deviate from the overall allele frequencies. Such changes are stronger in smaller founder populations, given the higher sampling variance.

GENETIC HETEROGENEITY
Process by which a phenotype can be caused by different loci. A complex example is epilepsy, which may be attributable to different causes in different individuals: single *gene* disorders, *multifactorial inheritance*, chromosomal disorders, or even brain injuries. The last case is a *phenocopy*.

GENOTYPE
The genetic constitution of an organism, which is modulated by the environment before being expressed as a phenotype.

HAPLOTYPE
Set of allelic states found at neighbouring loci in a chromosome, as inherited from a parent. Haplotypes can be broken down by *recombination*. A haplotype shared among unrelated individuals affected with a genetic disease may indicate that a gene causing the disease maps to that genomic region.

See end of article for authors' affiliations

....

Correspondence to:
Dr N Malats, Institut Municipal d'Investigació Mèdica, Carrer del Dr Aiguader 80, E-08003, Barcelona, Spain;
nuria@imim.es

....

HARDY-WEINBERG EQUILIBRIUM

State in which the *allele* and *genotype* frequencies do not change from one generation to the next in a population. It requires random mating and the absence of selection, mutation, migration, and genetic drift. In Hardy-Weinberg equilibrium, allele and genotype frequencies are related through the **Hardy-Weinberg law**: for a locus with two alleles P, Q at frequencies p and q respectively, homozygotes for P are found at frequency p^2, homozygotes for Q have a frequency q^2, and heterozygotes are found at a frequency $2pq$. Although conditions for Hardy-Weinberg equilibrium are seldom strictly met, genotype frequencies are usually consistent with the Hardy-Weinberg law. Some useful software packages to test whether a set of genotypic frequencies conforms to Hardy-Weinberg are Arlequin (http://anthropologie.unige.ch/arlequin/) and Genepop (http://wbiomed.curtin.edu.au/genepop/), among others.

INHERITANCE

Pattern followed by the transmission from generation to generation of a given *phenotype*, usually a disease.

Complex inheritance (non-Mendelian inheritance)

Variability in *phenotype* expression that is attributed both to the inheritance of combinations of alleles at multiple loci and to environment exposures.

Multifactorial inheritance

Complex inheritance in which multiple genes are involved jointly with environmental influences.

Polygenic inheritance

Complex inheritance in which multiple genes but no environmental factors are involved.

Mendelian inheritance

Simple pattern of inheritance that follows the rules set out by Mendel. Mendelian traits are determined by just one genetic locus, with complete *penetrance* and no *phenocopies*. Mendelian inheritance can be *dominant*, *recessive*, or *sex linked*.

Dominant inheritance

Type of inheritance in which one copy of an abnormal gene is sufficient to cause disease (for example, Huntington's disease). If *penetrance* is complete, the abnormal gene is inherited from a parent who also has the disorder and every generation in the family has members with the disorder.

Recessive inheritance

Type of inheritance in which two abnormal copies of the gene must be present for the individual to be affected (for example, cystic fibrosis). Each parent contributes one abnormal copy of the gene to the child who has the disorder. *Heterozygous* individuals (such as the parents of the affected) are called carriers of the disorder because they have one normal and one abnormal copy of the gene, but they do not show symptoms of the disorder.

Sex linked inheritance

Type of inheritance followed by the traits caused by genes located on the X or (rarely) on the Y chromosomes. X linked disorders can also be recessive or (very rarely) dominant. When the abnormal gene that is responsible for a recessive disorder is located on the X chromosome (for example, haemophilia) usually only males are affected because they do not possess a second, normal, copy of the gene. Such males are called hemizygous. X linked dominant inheritance (for example, Rett syndrome) follows a pattern similar to autosomal dominant inheritance except that more females are affected than males.

LINKAGE

The phenomenon whereby *phenotypes* and *alleles* at one or more marker *alleles* tend to be inherited together more often than expected. Linkage usually means that a gene contributing partially or completely to the *phenotype* (a genetic disease, for instance) maps in the vicinity of the *markers*.

MARKER

Any neutral *polymorphism* used in *linkage* or *association analysis*.

MEIOSIS

Process whereby four haploid germ cells (gametes) are produced from a diploid parent cell for sexual reproduction. During meiosis crossovers occur between *homologous chromosomes* so that each chromosome found in the gamete consists of a patchwork of material from both members of the pair.

METABOLISM (OF EXOGENOUS AND ENDOGENOUS CHEMICAL COMPOUNDS)

Cellular system of enzymes (most of them polymorphic) that activates and deactivates chemical compounds through chemical radicals. Metabolic enzymes are classified in two groups according their most important function, activating or deactivating, and in several families.

MITOSIS

Asexual reproduction of a somatic cell in which the two daughter cells each have a genetic makeup that is identical to that of the parent cell.

PENETRANCE

The likelihood, or probability, that a particular genotype will be expressed in the *phenotype*. A penetrance of 100% means that the associated phenotype always occurs when the corresponding genotype is present. Similarly, if only 30% of those carrying a particular allele (such as a disease-causing mutation) exhibit a phenotype (the disease), the penetrance is 30%.

PHENOCOPY

An environmentally caused *phenotype* that mimics a genetic trait. For example, epilepsy can be caused by mutations in single genes (with *genetic heterogeneity*), and, among other causes, by brain injury, which produces a phenocopy of genetic epilepsy.

PHENOTYPE

Expressed traits or characteristics of an organism, regardless of whether or to what extent the traits are the result of genotype or environment, or of the interaction of both. For example, hair colour, weight, or the presence or absence of a disease.

(GENETIC) POLYMORPHISM

Genome segment (locus), within or outside a *gene*, in which alternate forms (*alleles*) are present. In population genetics, variation is polymorphic if all alleles are found at frequencies >1%. In clinical genetics, a polymorphism refers to any genetic variation not known to be a direct cause of disease, in contrast with a *mutation*. However, the distinction between mutation and polymorphism in the latter sense may be rather fuzzy, as the path from genetic variation to disease can be sometimes very complex. In molecular epidemiology, *metabolic* and *DNA repair* gene polymorphisms are some of the markers (indicators) used to explore genetic susceptibility to develop a

disease. They are considered under the hypothesis that they can affect the development of the disease only in the presence of an environmental risk factor.

RECOMBINATION

Presence in the offspring of allelic combinations in a chromosome (that is, *haplotypes*) not present in the parents as a result of **crossing over** (*see meiosis*). The average probability of recombination is 1% per million base pairs, although this figure varies greatly across the genome.

Contributors

Both authors have contributed equally to the manuscript.

....

Authors' affiliations
N Malats, Institut Municipal d'Investigació Mèdica, Barcelona, Spain
F Calafell, Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, Barcelona, Spain

REFERENCES
1 **Calafell F**, Malats N. Glossary on basic molecular genetics. *J Epidemiol Community Health* 2003;**57**:398–400.
2 **Various authors**. Genetic epidemiology. *Epidemiol Rev* 1997;**19**. (Monographic issue on genetic epidemiology).
3 **Hartwell LL**, Hood L, Goldberg ML, *et al. Genetics, from genes to genomes*. Boston: McGraw Hill, 2000. (Includes a glossary).
4 **King RC**, Stansfield WD. *A dictionary of genetics*. Oxford: Oxford University Press, 1997. (Extensive but needs updating).
5 **Strachan T**, Read AP. *Human molecular genetics*. Oxford: Bios Publishers, 2001. (Includes an extensive glossary).

SPEAKER'S CORNER

The urge for evidence based knowledge

Clearly, the context of health sciences is increasingly been influenced by a general commitment related to interventions intending to improve the health conditions of individuals and populations brought by the so called evidence based knowledge. You can easily notice a proliferation of evidence based issues and debates in scientific health (including community health) journals.

Our view about the challenges to an evidence based knowledge does not refer to the impending difficulties of, sometimes, not having confirmed evidences on specific matters, or, even when they are available, there is uncertainty on which of them would be the best choice. The emphasis here is placed in the debatable conviction of the supremacy of some kind of knowledge over others that does not seem to require a plain belief in the truth itself to support it.

This standpoint needs some kind of a reasonably grounded concept of truth (even if it is a provisional one). Especially when anyone says that they possess the techniques to establish what is the actual best approximation to the truth. None the less, truth as a hierarchical quality of knowledge is, eventually, a social relationship whereby hegemony and domination are acting decisively. Rational truth will, in the end, guarantee the certainty that rational order is the only ticket that should enable mankind to arrive—sooner or later—to the ultimate and universal truth, whatever it may be.

You can notice in language that there are verbal forms for the opposite nouns of truth/veracity: falsehood—to falsify; lie—to lie. But there is not an equivalent verb for "truth". We can only verify the truthfulness of something. Perhaps because it seems that there is always the supposition that its "concrete" existence as a fact can, sooner or later, be established. The truth extracted by scientific means seems to have an ontological dimension in itself. Only, it would be necessary in a due course of time, to discover it, reveal it...

Well, a theory of truth tries to sustain a constant and secure superiority of some set of beliefs over others based upon an idea that these beliefs were achieved through reliable procedures and/or were put forward by special kind of people that can be trusted sufficiently to be followed. Therefore, a rhetoric of power granted by a rationalistic mode of reasoning and its undeniable successes is always at stake.

L D Castiel
National School of Public Health, Oswaldo Cruz Foundation, Rua Leopoldo Bulhões 1480/802, Rio de Janeiro, Brazil 21041-210; castiel@ensp.fiocruz.br