Genetic aspects of Multiple Sclerosis
Boon, Maartje

IMPORTANT NOTE: You are advised to consult the publisher’s version (publisher’s PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2006

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Glossary of genetic and statistical terms

**Affected relative pair**: two genetically related individuals, each of whom is affected with the trait in question. The most common types of affected relative pairs are affected sibling pairs, affected cousins and affected avuncular pairs.

**Affected sib(ling) pair**: see affected relative pair

**Allele**: variant of a gene or marker due to changes at the DNA level

**Ascertainment**: The selection of individuals for inclusion in a genetic study

**Association study**: study that tests whether a genetic marker or markers occur(s) more frequently in cases than in controls.

**Base pair**: measurement of length for DNA. Includes a nucleotide base with its complementary base, i.e. adenine bound to thymine or guanine bound to cytosine

**Candidate gene**: A gene that has been implicated in susceptibility to or influencing the course of a given disease

**cM (CentiMorgan)**: Measure of genetic distance, equivalent to 1% recombination

**Chromosome**: Macromolecular complex of DNA and proteins.

**Concordant**: A pair of relatives, most often twins, of which both members exhibit the same phenotype i.e. they both do or they both do not have the disease or phenotype in question.

**Crossing-over**: Reciprocal breaking and rejoining of homologous chromosomes during meiosis that results in exchange of chromosomal segments.

**Discordant**: A pair of relatives, most often twins, of which both members exhibit different phenotypes; opposite of concordant.

**DNA**: Deoxyribonucleic acid, double-stranded helix consisting of complementary strands of nucleotides
**Epistasis**: process of two or more genes interacting with each other

**Exon**: part of a gene that is expressed by transcription into mRNA

**Gene**: sequence of DNA nucleotide bases, coding for a polypeptide. Individual unit of heredity

**Genetic drift**: random process leading to increase in frequency of some alleles and decrease of others over time

**Genome**: sum of all the genetic information of an organism.

**Genotype**: The observed alleles at a genetic locus for an individual.

**Haplotype**: the ordered arrangement of alleles on a chromosome.

**Hardy-Weinberg equilibrium (HWE)**: indicates how many homozygotes and heterozygotes at a given locus are expected in a population on the basis of allele frequencies. Testing for HWE is often used as a test for the quality of the genotype-data offered for statistical analysis. Deviations from HWE are often caused by systematic errors in genotyping.

**Heterozygote**: organism with two different alleles at a given locus.

**Heterozygosity**: also informativeness: the heterozygosity of a marker locus indicates which percentage of individuals is heterozygous at the locus.

**Homozygote**: organism with two identical alleles at a given locus.

**Identity-by-descent**: Two alleles are identical and it can be established that they are inherited from a common ancestor

**Identity-by-state**: Two alleles are identical.

**Imprinting**: differential effects of maternally and paternally derived DNA

**Intron**: part of the gene that is not transcribed into mRNA.

**kb (kilobase)**: one thousand basepairs in a sequence of DNA
**Linkage**: Co-inheritance of two or more loci because of close proximity on the chromosome. In this case, the recombination fraction (theta) between the loci is smaller than $\frac{1}{2}$.

**Linkage analysis**: statistical method aiming to estimate the recombination fraction and to test whether it is significantly smaller than $\frac{1}{2}$.

**Linkage disequilibrium (LD)**: a measure of whether alleles at two loci co-exist in a population in a non-random fashion. Alleles that are in LD are found together on the same haplotype more often than would be expected by chance.

**Lod score**: base 10 logarithm of the odds for linkage

**Lod score method**: also known as maximum likelihood (ML) method, statistical method testing whether a set of data indicates that two loci are linked or unlinked. For an assumed value of theta, the likelihood (probability of occurrence) of the given family data is calculated. Theta is varied and that value of theta where the likelihood of the observed data is the highest is called the maximum likelihood estimate (MLE).

**Mapping**: process of determining the position of a locus on the chromosome relative to other loci

**Marker**: see DNA marker

**Mb (megabase)**: one million basepairs in a sequence of DNA

**Meiosis**: type of cell division creating germ cells

**Mendelian inheritance**: transmission of traits according to the laws of inheritance first formulated by Gregor Mendel in 1865. The Law of Segregation describes the separation of the two alleles at one genetic locus into different germ cells. The Law of Independent Assortment describes that most genetic loci behave as though they were independent of one another. The resulting patterns of inheritance are autosomal dominant, recessive or X-linked (which may be dominant or recessive).
**Microsatellites**: repeat DNA segments which comprise short head-to-tail tandem repeats giving the variable number of tandem repeat (VTNR) type polymorphism of which the basic core repeat unit involves a two to four nucleotide base pair repeat motif (e.g. CACACA, also written (CA)n).

**Multifactorial disorders**: diseases that result from an interaction of environmental factors with multiple genes at different loci.

**Mutation**: alteration of the DNA sequence caused by substitution, deletion, insertion or duplication of base pairs.

**Nucleotide**: the monomeric component of DNA/RNA comprising a base, a pentose sugar and a phosphate group. The bases are adenine (A), thymine (T), uracil (U), guanine (G) and cytosine (C), the pentose sugars deoxyribose and ribose. The sequence of base pairs determines the genetic code.

**Oligogenic**: caused by the influence of a few genes.

**PCR**: polymerase chain reaction: technique for amplifying stretches of DNA using mainly template DNA, primers, an amplifying enzyme and nucleotides. The process consists of repeated cycles of separation of the two strands of the double helix of DNA, annealing of primers, amplification and again separation.

**Penetrance**: The frequency with which individuals carrying a given gene variant will show manifestations associated with that variant. If the penetrance of a given disease allele is 100%, all individuals carrying that allele will express the associated disorder.

**Phenocopies**: individuals showing the same phenotype as a result of different underlying genetic and or environmental causes.

**Population stratification**: The presence of multiple subgroups with different allele frequencies within a population. These different underlying allele frequencies in sampled subgroups might be independent of the disease and lead to erroneous conclusions of linkage disequilibrium or relevance to the disease.
**Power** is defined as the probability of rejecting the null hypothesis when it is false or of concluding the alternative hypothesis when it is true. In general terms: power is the capability of a study to detect a given difference of a given size if the difference really exists. Power is calculated as $1-\beta$ (see: Type II error).

**Proband**: Individual through whom a family is ascertained for a genetic study, most often an affected individual.

**Recombination**: Event during meiosis: homologous chromosomes exchange parts when they are paired to be segregated to different germ cells. Thus, new combinations of genes are formed. Also called crossing-over.

**Recombination fraction**: (also theta, $\theta$) the proportion of recombinant offspring. When two loci are very close, theta is small (close to zero). When two loci are far apart on the same chromosome or on different chromosomes, theta is equal to $\frac{1}{2}$, meaning that the two loci do not co-inherit more often than expected by chance.

**Relative Risk (RR)**: the ratio of the disease rate in relatives of affected individuals to the disease rate in relatives of controls.

**Type I error**: error of hypothesis testing: rejecting the null hypothesis when it is really true. $\alpha$ is the probability of a type I error.

**Type II error**: error of hypothesis testing: failing to reject the null hypothesis when it is false (or not rejecting the null hypothesis when the alternative hypothesis is true). $\beta$ is the probability of a type II error.

**RNA**: Ribonucleic acid, transcript of DNA executing various functions (among which messenger for translation into protein, transport of amino acids, ribosomal activities).

**Segregation analysis**: method of genetic analysis, testing whether an observed pattern of phenotypes in families is compatible with a specific model of inheritance.

**Sequencing**: Process of determining the identity and order of nucleotides in DNA or RNA.