Genetic Polymorphisms

Introduction

Genetic polymorphisms are different forms of a DNA sequence. "Poly" means many, and "morph" means form. Polymorphisms are a type of genetic diversity within a population's gene pool. They can be used to map (locate) genes such as those causing a disease, and they can help match two samples of DNA to determine if they come from the same source.

The DNA sequence of exactly the same region on a chromosome is remarkably similar among chromosomes carried by many different individuals from around the world

In randomly chosen segment of human DNA about 1000 base pairs in length, contains (on average), only one base pair that varies between the two homologous chromosomes inherited from the parents (assuming the parents are unrelated)!!!

When a variant is so common that it constitutes more than 1% of chromosomes in the general population the variant constitutes what is known as genetic polymorphism

Polymorphism ca be caused by: Deletion, duplication, triplication and so on of hundreds of millions of base pairs of DNA Or it can also be changes in one or a few bases in the DNA located between genes or within introns Sequence changes may also be located in the coding sequence of genes themselves and result in different protein variants that may lead in turn to different phenotypes Others are in regulatory regions and may also be important in determining phenotype by affecting transcription or mRNA stability

Single nucleotide polymorphism - SNPs (simplest and most common of all polymorphisms) SNPs usually have only 2 alleles corresponding to the two different bases occupying a particular location in the genome 99.9% of the DNA sequence is identical between any two individuals. Of the remaining 0.1% difference, 80% is represented by SNPs For example, two sequenced DNA fragments from different individuals, AAGCCTA to AAGCTTA, contain a difference in a single nucleotide. In this case we say that there are two alleles : C and T Within a population, SNPs can be assigned a minor allele frequency — the lowest allele frequency at a locus that is observed in a particular population. (As an extension of the Human Genome Project, scientists focus (besides other topics) to identification of millions of SNPs, that are gathered to public databases. It is technically feasible now to type 500 000 SNPs in a single DNA sample, which should accelerate the identification of the alleles responsible for a wide range of common diseases) The fact that SNPs are common does not mean that they must be neutral and without effect on health or prolonged existence what it does mean is that any effect of common SNPs must be a delicate altering of disease susceptibility rather than a direct cause of serious illness

RFLP = restriction fragment length polymorphism Restriction endonucleases are bacterial enzymes that cleave DNA at a specific sequence. They protect bacteria against bacteriophage infection (or perhaps have no function) TAGCCATCGGTA-CGTACTCAATGATCA ATCGGTAGCCATGC-ATGAGTTACTAGT TAGCCATCGGTAAGTACTCAATGATCA ATCGGTAGCCATTCATGAGTTACTAGT

Insertion-deletion (indels) polymorphism - Insertion – deletion of between 2 and 100 nucleotides - Half of all indels are referred to as simple because they have only two alleles presence or absence of inserted or deleted segment… - The other half are multiallelic due to a variable numbers of a segment of DNA that is repeated in tandem at a particular location (Multiallelic indels are further subdivided into microsatellite and minisatellite)

MICROSATELLITES

Microsatellites (or STR = short tandem repeats, SSR = simple sequence repeats) TAGCCATCGGTACACACACACACACACAGTGCTTCAGTAGC TAGCCATCGGTACACACACACACAGTGCTTCAGTAGCGT The different alleles in the microsatellite polymorphism are a result of differing numbers of repeated nucleotide units contained within one microsatellite A microsatellite locus often has many alleles present in the population and can be readily genotyped by determining the length of the PCR fragment generated by primers that border the microsatellite repeat If a detectable polymorphism is situated closely enough to a locus, which harbours a fundamental mutation for the studied disease, the polymorphism will be linked to the mutated allele. In most cases, the polymorphism will be passed together with the mutated allele from the parents to the offspring („co segregation“). Thus, the polymorphism can be used as a „marker“for the disease even without exact knowledge of its molecular basis.

MINISATELLITES Variable number tandem repeats

Sources: Lecture's note. Emery's Elements of Medical Genetics. Color Atlas of Genetics.