

Using molecular marker technology in studies on plant genetic diversity

Glossary

Accession: A sample of a crop variety collected at a specific location and time; may be of any size.

AFLP: Amplified fragment length polymorphism. A highly sensitive method for detecting polymorphisms in DNA. DNA first undergoes restriction enzyme digestion, then a subset of DNA fragments is selected for PCR amplification and visualisation.

Allele: One of the alternative forms of a gene that can exist at a single locus.

Allele frequency: A measure of the commonness with which an allele is found in a population, or its proportional share of all alleles of a gene.

Allozyme: An isozyme whose synthesis is controlled by codominant alleles of one gene.

Amino acids: Bi-functional organic compounds that contain a basic amino group (-NH₂) and an acidic carboxyl group (-COOH).

Annealing: Spontaneous alignment of two single DNA strands to form a double helix.

AP-PCR: Arbitrarily primed polymerase chain reaction. A technique for amplifying anonymous stretches of DNA, using PCR. Related to RAPD.

Apomixis: The production of an embryo in the absence of meiosis. Apomictic higher plants produce asexual seeds, derived only from maternal tissue.

Arbitrary primer: A short oligonucleotide primer used in certain PCR methods to initiate DNA synthesis at random locations on the target DNA.

Autogamy: Transfer of pollen from the anther of a flower to the stigma of the same flower or, sometimes, to that of a genetically identical flower (as of the same plant or clone). Ability of many plant species to naturally and successfully fertilise within one individual. Also called self-pollination.

Autoradiography: A technique where radioactively labelled molecules are visualised through exposure to X-ray film.

Bacteriophage: A virus that infects bacteria. Genetically altered forms are used as vectors in cloning DNA.

Base: The chemical unit that characterises a nucleotide. In DNA, the bases found are adenine (A), guanine (G), thymine (T) and cytosine (C). In RNA, the bases are adenine, guanine, uracil (U) and cytosine.

Base pair: Two nucleotide bases on different strands of a nucleic acid molecule that are held together by hydrogen bonds. Bases can pair in only one way: adenine with thymine (DNA) or uracil (RNA), and guanine with cytosine (DNA). *See also* Complementary sequence

Base sequence: The order of nucleotide bases in DNA or RNA.

Bottleneck: A brief reduction in population size that usually leads to random genetic drift.

CAPS: Cleaved amplified polymorphic sequence, also known as PCR-RFLP, a technique for detecting polymorphisms at a particular locus. A locus undergoes PCR amplification and polymorphisms are detected by differences in restriction fragment sizes between individuals.

cDNA or complementary DNA: DNA transcribed from an RNA molecule by an enzyme called reverse transcriptase.

Centromere: The constricted region of a chromosome to which spindle fibres attach during cell division.

Character: Or trait, an attribute of individuals within a species for which various heritable differences can be defined.

Chromatin: Complex of DNA and protein of the interphase cell.

Chromosome: A linearly continuous arrangement of genes and other DNA, and associated proteins and RNA.

Cloning: In molecular biology: the process of using DNA manipulation procedures to produce multiple copies of a single gene or DNA segment.

Codominance: The situation in which a heterozygous individual exhibits the phenotypes of both alleles of a particular gene.

Complementary sequence: The sequence of a DNA or RNA strand to which a given nucleotide sequence can bond to form a double-stranded structure, e.g. TAGGAT is the complementary sequence to ATCCCTA where A = adenine, C = cytosine, G = guanine and T = thymine. *See also* Base pair.

cpDNA: Chloroplast DNA.

Cytoplasmic inheritance: Inheritance of the genes found in cytoplasmic organelles (viz. chloroplasts or mitochondria).

DAF: DNA amplification fingerprinting. A technique for amplifying anonymous stretches of DNA, using PCR. Related to RAPD.

DArT: Diversity array technology.

Deletion: A particular kind of mutation involving the loss of some DNA from a chromosome.

Dendrogram: Any branching diagram that shows, by means of lines shaped like U's a hierarchy of categories or objects based on the degree of similarity or number of shared characters. Often, the length of each U represents the distance between the two objects being connected.

Denaturation: The separation of the two strands of the DNA double helix, or the severe disruption of a complex molecule without breaking the major bonds of its chains.

DGGE: Denaturing gradient gel electrophoresis. A method for separating DNA fragments according to their mobility under increasingly denaturing conditions (usually increasing formamide/urea concentrations). *See also* heteroduplex analysis, SSCP and TGGE.

DHPLC: Denaturing high-performance liquid chromatography. This method can detect sequence variations of a single base pair.

DNA: Deoxyribonucleic acid, a double chain of linked nucleotides (having deoxyribose as the sugar component), which is the fundamental molecule of which genes are composed.

DNA fingerprint: A unique pattern of DNA fragments as revealed by Southern hybridisation or by PCR.

DNA polymerase: Any enzyme with the ability to synthesise new DNA strands, using a DNA template.

DNA sequence: The order of nucleotide bases in the DNA molecule.

Dominant allele: An allele that expresses its phenotypic effect even when heterozygous with a recessive allele (*qv.*). That is, if A is dominant over a, then AA and Aa have the same phenotype.

Double helix: The structure of DNA first proposed by Watson and Crick, with two linked helices joined by hydrogen bonds between paired bases.

Ecotype: A population or strain of an organism that is adapted to a particular habitat.

Electrophoresis: A technique for separating the components of a mixture of molecules (proteins, DNA or RNA) by size as a result of an electric field within a support gel.

Enzyme: A protein that functions as a catalyst of biochemical reactions.

EST: Expressed sequence tag. A small part of the active part of a gene, made from cDNA. It can be used as a marker, to search the rest of the gene or to locate it in a larger segment of DNA.

Eukaryote: A cell or organism with a distinct, membrane-bound nucleus and other differentiated subcellular components.

***ex situ* conservation:** (1) A conservation method that entails the removal of germplasm resources (seed, pollen, sperm, individual organisms) from their original habitat or natural environment. (2) Keeping components of biodiversity alive outside their original habitat or natural environment. *Cf. in situ* conservation.

F₁, F₂, F₃, ... : A shorthand notation used to denote the different generations involved in breeding experiments. F₁ is the first filial generation, that is, the progeny of the parental cross; F₂ is the second filial generation, that is, the progeny of self-fertilising or intercrossing F₁ individuals, and so on.

Flanking regions: The DNA sequences extending on either side of a specific gene or locus.

Gene: The basic physical and functional unit of heredity, which passes information from one generation to the next. It is a segment of DNA that includes a transcribed section and a regulatory element that allows its transcription.

Genebank: A facility established for the *ex situ* conservation of individuals (seeds), tissues, or reproductive cells of plants or animals.

Gene flow: The exchange of genes between different but (usually) related populations.

Gene mapping: The determination of the relative positions of genes on a chromosome or plasmid and the distance between them.

Genepool: The sum total of genes, with all their variations, possessed by a particular species at a particular time.

Genetic drift: Change in allele frequency from one generation to another within a population, due to the sampling of finite numbers of genes that is inevitable in all finite-sized populations. The smaller the population, the greater is the genetic drift, with the result that some alleles are lost, and genetic diversity is reduced.

Genetic linkage: The proximity of two or more genes on a chromosome so that they tend to be inherited together.

Genetic marker: An allele, a band in a gel or trait that serves experimentally as a probe to identify an individual or one of its characteristics.

Genetic resources: The genes found in plants and animals that are of actual or potential value to people.

Genome: The entire complement of genetic material in an organism.

Genotype: The specific allele composition of either of the entire cell or, more commonly, of a certain gene or set of genes.

Germplasm: The total genetic variability available to a population of organisms as represented by germ cells, seeds, etc.

Haplotype: A specific allelic constitution at a number of loci within a defined linkage block.

Heredity: The process by which genetic traits are passed from parents to offspring.

Heteroduplex: A double-stranded DNA molecule or a DNA/RNA hybrid where each strand is from a different source.

Heteroduplex analysis: The study of the mobility of heteroduplex DNA under polyacrylamide gel electrophoresis. The reduced mobility of heteroduplex DNA compared with homoduplex DNA is proportional to the degree of divergence of the sequences. See *also* DGGE, SSCP and TGGE.

Heterozygous gene pair: A gene pair having two different alleles in the two chromosome sets of a diploid individual, for example, A a or A¹ A².

Homoduplex DNA: A double-stranded DNA molecule where both strands are from the same source.

Homologous: Corresponding or alike in structure, position or origin.

Homozygous gene pair: A gene pair having identical alleles in both copies of the chromosome set, for example, AA or aa.

Hybrid: Either (1) a heterozygous individual, or (2) a progeny individual from a cross between parents with different genotypes.

Hybridisation: In molecular biology: the binding of complementary DNA and/or RNA sequences to form a double-stranded structure.

Insertion: A type of chromosomal abnormality in which a DNA sequence is inserted into a gene, disrupting the normal structure and function of that gene.

***in situ* conservation:** A conservation method that attempts to preserve the genetic integrity of gene resources by conserving them within the evolutionary dynamic ecosystems of the original habitat or natural environment. *Cf. ex situ* conservation

Isozyme: Multiple forms of an enzyme whose synthesis is controlled by more than one gene.

ISSR: Inter-simple sequence repeat. ISSR primers are anchored at their 3' ends to direct the amplification of the genomic segments between the ISSRs.

Landrace: A crop cultivar or animal breed that has evolved with and has been genetically improved by traditional farmers without influence from modern breeding practices.

Library: A collection of DNA clones obtained from one DNA donor.

Ligase: A type of enzyme that can rejoin a broken phosphodiester bond in a nucleic acid.

Ligation: The process of joining two or more DNA fragments together.

Locus (*pl. loci*): The specific place on a chromosome where a gene or particular piece of DNA is located.

MAAP: Multiple arbitrary amplicon profiling. A collective term for PCR techniques using arbitrary primers.

Marker: An identifiable physical location on a chromosome whose inheritance can be monitored (e.g. gene, restriction enzyme site or RFLP marker).

Melting temperature (T_m): Midpoint of the temperature range over which DNA is denatured.

Microarray: Small spots of DNA fixed to glass slides or nylon membranes. This technology is based on the hybridisation between short oligonucleotide probes and complementary DNA sequences.

Microsatellite DNA: A type of repetitive DNA based on very short repeats such as dinucleotides, trinucleotides or tetranucleotides. *See also* Repetitive DNA.

Minisatellite DNA: A type of repetitive DNA sequence based on short repeats with a unique common core. *See also* Repetitive DNA.

mtDNA: Mitochondrial DNA.

Multiplexing: Simultaneously performing several different reactions in the same reaction tube to increase efficiency.

Mutation: A permanent structural alteration in DNA. In most cases, DNA changes either have no effect or cause harm, but occasionally a mutation can improve an organism's chance of surviving and passing on the beneficial change to its descendants.

Nitrogen bases: Molecules that are important components of nucleic acids, composed of nitrogen-containing ring structures. Hydrogen bonds between bases link the two strands of the DNA double helix.

Nuclease: An enzyme that cleaves phosphodiester bonds, which link adjacent nucleotides in DNA and/or RNA. An exonuclease progressively cleaves from the end of the substrate molecule; an endonuclease cleaves at internal sites within the substrate molecule.

Nucleotide: A molecule composed of a nitrogen base, a sugar and a phosphate group. Nucleotides are the building blocks of nucleic acids.

Oligonucleotide: A short segment of DNA that is synthesised artificially.

Pedigree: A simplified diagram of a family's genealogy that shows family members' relationships to each other and how a particular trait or disease has been inherited.

PCR: Polymerase chain reaction. A method for amplifying a DNA sequence in large amounts, using a heat-stable polymerase and suitable primers to direct the amplification of the desired region of DNA.

PCR-RFLP: Alternative name for the technique known as 'cleaved amplified polymorphic sequence' or CAPS *qv*.

Peptide: Two or more amino acids joined by a peptide bond.

Phenotype: Either (1) the form taken by a trait (or group of traits) in a particular individual; or (2) the detectable external appearance of a specific genotype.

Plasmid: An extra chromosome molecule of DNA that is able to replicate autonomously.

Point mutation: A change in a single base pair of DNA.

Polymer: A molecule having repeated subunits.

Polymerase: General term for enzymes that carry out the synthesis of nucleic acid, using a pre-existing nucleic acid template and appropriate nucleotides (*viz.* ribonucleotides for RNA and deoxyribonucleotides for DNA).

Polymorphism: The appearance of different forms associated with various alleles of one gene or homologous of one chromosome.

Polypeptide: A protein, which is a chain of linked amino acids.

Primer: A short DNA or RNA fragment annealed to a single-stranded DNA and to which further nucleotides can be added by DNA polymerase.

Probe: A finite nucleic acid piece that can be used to identify specific DNA segments bearing its complementary sequence.

Protein: A polymer of amino acids joined by peptide bonds and which may comprise two or more polypeptide chains.

RAPD: Random amplified polymorphic DNA. A technique for amplifying anonymous stretches of DNA, using PCR with arbitrary primers.

Recessive allele: An allele whose phenotypic effect is not expressed in the heterozygous state, but is masked by the dominant allele (*qv.*).

Recombination: Also known as crossing over. The production of a DNA molecule with segments derived from more than one parental DNA molecule. In eukaryotes, this is achieved by the reciprocal exchange of DNA between non-sister chromatids within a homologous pair of chromosomes during prophase of the first meiotic division. Recombination allows the chromosomes to rearrange their genetic material, thereby increasing the potential of genetic diversity.

Repetitive DNA: A stretch of DNA consisting of multiple repeats of a motif. *See also* Microsatellite DNA, Minisatellite DNA and Satellite DNA.

Restriction enzyme: An endonuclease that will recognise a specific target sequence and cut the DNA chain at that point.

Restriction fragment: A DNA fragment that has been cut by a restriction enzyme.

Restriction site: The specific nucleotide sequence of DNA at which a particular restriction enzyme cuts the DNA.

Reverse transcriptase: An enzyme, requiring a DNA primer, that catalyses the synthesis of a complementary DNA strand from an RNA template.

RFLP: Restriction fragment length polymorphism. Variation between individuals as detected by differences in DNA fragment sizes after restriction digestion.

RNA: An organic acid containing repeating nucleotide units of adenine (A), guanine (G), cytosine (C) and uracil (U), whose ribose components are linked with phosphodiester bonds.

Satellite DNA: Very high repetition (1000 to more than 100,000 copies) of a basic motif or repeat unit (commonly 100 to 300 base pairs) that occurs at a few loci on the genome. *See also* Repetitive DNA.

SCAR: Sequence characterized amplified region. A SCAR is a locus representing a single RAPD fragment that has been sequenced. Primers specific to the locus can be designed and used in PCR amplification.

Segregation: Genetically, it refers to the production of two separate phenotypes corresponding to the two alleles of a gene.

Sequencing: The determination of the order of nucleotides in a DNA or RNA molecule, or of the order of amino acids in a protein.

SNP: Single nucleotide polymorphism. Polymorphisms resulting from single-base substitutions between homologous sequences.

Somaclonal variation: Variation found in vegetative cells dividing mitotically in culture.

Southern blotting: A procedure in which DNA fragments, separated by electrophoresis, are transferred to membrane filters for detecting specific base sequences by radiolabelled complementary probes. Also known as Southern hybridisation.

SSCP: Single-stranded conformational polymorphism. A method for distinguishing between similar sized DNA fragments according to the mobility of the single-stranded DNA under polyacrylamide gel electrophoresis. *See also* DGGE, heteroduplex analysis, and TGGE.

SSR: Simple-sequence repeats. *See* Microsatellite DNA.

SSRP: Simple-sequence repeat polymorphism. *See* Microsatellite DNA.

STMS: Sequence-tagged microsatellite sites. Primers constructed from the flanking regions of microsatellite DNA, and which can be used in PCR reactions to amplify the repeat region.

Structural gene: Any gene that codes for a protein.

STS: Sequence-tagged site. A general term given to a marker that is defined by its primer's exact location and order of bases.

Synteny: Said to occur where all loci are positioned on the same chromosome. Loci may not appear to be linked by conventional genetic tests for linkage but can still be syntenic.

Tandem repeat: Multiple copies of the same base sequence on a chromosome.

Telomeres: The distal ends of each chromosomal arm, involved in the replication and stability of linear DNA molecules.

Template: A molecule that serves as the pattern for synthesising another molecule, e.g. a single-stranded DNA molecule can be used as a template to synthesise the complementary nucleotide strand.

TGGE: Thermal gel gradient electrophoresis. A method for separating DNA fragments according to their mobility under increasingly hot denaturing conditions. *See also* DGGE, heteroduplex analysis and SSCP.

Transcription: The synthesis of complementary RNA, using a DNA template.

Transposable element: A genetic element that has the ability to move from one site on a chromosome to another. Some resemble, and many originate from, retroviruses.

Vector: The agent (e.g. plasmid or virus) used to carry a cloned DNA segment.

VNTR: Variable number of tandem repeat. A class of polymorphism characterised by the highly variable copy number of identical or closely related sequences.

Wild type: The type or form of an organism or gene that occurs most frequently in nature. Often refers to how organisms or genes are found naturally, that is, in the wild, before researchers induced mutations.