The application of representational difference analysis and plant differentiation

By

BJ VORSTER

Dissertation submitted in partial fulfillment of the requirements for the degree

MAGISTER SCIENTIA

Forestry and Agricultural Biotechnology Institute (FABI)
Department of Botany

in the

Faculty of Natural and Agricultural Sciences
University of Pretoria

SUPERVISOR: PROF. K.J. KUNERT
CO-SUPERVISOR: PROF. A-M. OBERHOLSTER

November 2003

DECLARATION

I, the undersigned, hereby declare that the thesis submitted herewith for the degree Magister Scientiae to the University of Pretoria, contains my own independent work and has not been submitted for any other degree at any other university.

BJ Vorster

November 2003

TABLE OF CONTENTS

ABSTRACT

The methods used in analyzing plant genomes have changed dramatically in the last century, from extensive breeding experiments using crosses and backcrosses to techniques that identify molecular differences in the DNA sequence of different genomes and studying these differences much more intensely. Representational Difference Analysis (RDA) is one such molecular technique. Originally developed in cancer research, it has recently also been applied to study plant genomes. RDA uses subtractive hybridization to isolate genomic regions that differ between two genomes, allowing the isolation of unique genetic sequences from one genome. Since RDA has only been used on a very limited scale in plant genome analysis, the aim of this project was to test the potential of this technique on the level of plant identification and the application of the technique on large and small plant genomes. Eleven subtraction products were isolated from date palm using the restriction enzyme *Bam*HI for genome differentiation and production of genome representations. These subtraction products were shown to be dispersed repetitive sequences. It was show that out of these subtraction products, Dp41, was much more abundant in the genome of one date palm cultivar compared to the other cultivar used in this study. Analysis of this element in a number of date palm plants indicates that Dp41 possibly represents a hot-spot for stress-induced mutations. In addition three additional subtraction products were isolated from date palm using the methylation sensitive restriction enzyme *Hpa*II for the creation of genome representations. Two of these showed homology to rRNA genes. In a third application subtraction products were isolated from the very large genome of *Pinus strobus*. Due to the size of the pine genome the complexity and number of the obtained subtraction products could not be used for analysis. A difference in genome methylation between the different lines could however be demonstrated.

OPSOMMING

Die analise van plantgenome het drasties gedurende die afgelope eeu verander. Uitgebreide telingseksperimente, wat gebruik maak van kruising en terugkruisings word aangevul met tegnieke wat molekulêre verskille in die nukleotiedvolgorde van verskeie genome identifiseer en in diepte ondersoek. Verteenwoordigende verskilanaliese (RDA) is een van hierdie tegnieke. Alhoewel oorspronklik ontwikkel vir kankernavorsing, word dit suksesvol in die bestudeering van plantgenome aangewend. RDA isoleer unieke geenvolgordes van een genoom deur die verskille tussen twee genome te identifiseer deur middel van substraksie-hibidisasie. Die doel van hierdie projek is om die potensiaal van RDA in die identifiseering van plante en die toepassing daarvan op groot en klein plantgenome te toets, aangesien die gebruik van RDA op die huidige oomblik baie beperk is in plantnavorsing. Elf substraksie-produkte is uit datelpalms geïsoleer deur die snydingsensiem *Bam*HI te gebruik vir genoom-differensiasie. Hierdie substraksie-produkte is verspreide herhalende nukleotiedvolgordes waarvan een, Dp41, baie meer algemeen voorgekom het in die genoom van een van die dadel-kultivars wat bestudeer is, in vergelyking met die ander kultivar. Verdere analise van hierdie element in 'n aantal dadelpalms het getoon dat Dp41 'n moontelike teikengebied vir stress-geïnduseerde mutasies mag wees. Die gebruik van die snydingsensiem *Hpa*II, wat sensetief vir DNS-metilering is, het nog drie substraksie-produkte opgelewer, waarvan twee ooreenkomste met rRNS gene getoon het. Ten derde is daar ook geslaag om substraksie-produkte uit die baie groot genoom van *Pinus strobus* te isoleer. As gevolg van die grootte van die genoom was die hoeveelheid en kompleksiteit van die subtraksie-produkte te groot om geanaliseer te word. 'n Verskil in genoom-metileering tussen die versikilinde lyne kon egter gedemonstreer word.

Research Objectives

There are many molecular techniques available that can be used to do genome analysis, each with their own advantages and disadvantages. In this MSc study one such technique was evaluated for its suitability in plant genome analysis. Representational difference analysis (RDA), which was developed in human cancer research to isolate differences between cancerous and non-cancerous cells, have recently been applied to study plant genomes, but very little is still known about the potential of RDA in plant genome analysis. The research objectives were therefore (1) to evaluate its potential in plant identification on the variety level by using methylation insensitive and sensitive restriction enzymes for the generation of genome representations, and (2) the evaluation of the RDA technique when applied to large genomes such as the pine genome. The experiments carried out in this MSc project focused in particular on (1) the isolation and characterization of possible genomic variations between closely related date palm varieties (2) the evaluation of the usefulness of these isolated subtraction products to be used as markers for cultivar identification and differentiation, (3) the isolation and characterization of differences due to methylation changes in the date palm genome, and (4) the ability of the RDA technique to subtract large genomes with a high complexity.

Thesis Composition

The following theses composes of seven chapters and discuss the application of representational difference analysis (RDA) to genome analysis in plants. Since (RDA) is a technique developed to isolate differences between cancerous and non-cancerous cells in humans, the hypothesis was that it could be used to isolate differences between closely related cultivars or cell-lines to isolate differences that could be useful in the creation of molecular markers. In this study RDA was used to isolate subtraction products from the genomes of two date palm cultivars as well as from the genome of white pine. It is also used to isolate subtraction products using methylation sensitive and non-sensitive restriction enzymes. A breakdown of the various chapters is provided below.

Chapter 1 of this thesis presents an introduction into molecular biology regarding the theories and philosophy forming the basis of this field of study and how it developed. It also gives an overview as to how plant genomes are organized and the processes involved in creating genome variation and evolution. It also deals with some of the most popular techniques used today in plant genome analysis, random amplified polymorphic DNAs (RAPD), and discusses the advantages and disadvantages of each technique focusing to a larger extent on representational difference analysis. **Chapter 2** presents the results obtained using RAPD to differentiate between two date palm cultivars, *'Medjool'* and *'Barhee'*. This includes two different methods for genomic DNA isolation, RAPD analysis, the cloning of a polymorphic DNA fragment, its sequence and the design of Sequence characterized amplified region (SCAR) primers for this fragment. **Chapter 3** focuses on the application of RDA on date palm. This chapter outlines the results obtained from application of this subtractive technique executed on genomic DNA digested with restriction enzymes including the isolation and cloning of subtraction products. **Chapter 4** focusus on the characterization of the RDA subtraction products using bioinformatic tools to conduct sequence homology searches and to do sequence alignments. The results obtained from the detailed characterization of one subtraction product DP41 is also presented in this chapter. **Chapter 5** deals with genome methylation. It consists of two parts, firstly detecting genome methylation in date palm using RDA, and secondly using RDA to detect genome methylation in the much larger and complex genome of *Pinus strobus*. It also includes a comparison of genome

methylation in *P. strobus* embryogenic cultures grown on tissue culture medium containing either a low or high concentration of 2,4-D, a synthetic plant growth regulator known to influence genome methylation during tissue culture. In **Chapter 6** the results of the various experiments are discussed and **Chapter 7** the **Achievements and conclusion** outlines the scientific achievements made by this thesis and also the failures are mentioned and discussed, and an overview about possible future research activities is provided. **References** lists all the literature used and sited in this thesis and finally, in the **Annex** the methods, composition of the tissue culture media as well as plasmid maps, and general buffers and solutions used in the various protocols, and lastly an article that was published from this results.

Acknowledgements

I would like to thank the University of Pretoria and the National Research Foundation for financial assistance that allowed me to conduct and complete my research. Special thanks also to Prof. Karl Kunert for years of patience and support in teaching and sharing his endless amount of knowledge with me. My sincere gratitude and appreciation also goes to Prof Anna- Maria Oberholster for her help and advice as well as to Prof Chris Cullis who allowed me to work in his laboratory, sharing his knowledge and spending time teaching and guiding me. My appreciation also goes to Dr Krystyna Klimaszewska and the Canadian Forestry Services for their assistance and support and allowing me to use their facilities.

This thesis would not have been completed if not for the help and support from friends, family and lab colleagues. Thanks for all your advice and support throughout my project.

ABBREVIATIONS, DEFINITIONS AND SYMBOLS

°C	Degrees Celsius
%	Percentage
µg	microgram
µl	microlitre
A	adenine
ABI	Applied Biosystems (Perkin Elmer)
AFLP	Amplified fragment length polymorphism
Amp	Ampicillin
App.	Appendix
BA	Benzyladenine
bp	base pair
C	cytosine
CH_3COONH_4	Ammonium acetate
CpG	Cytosine and guanine dinucleotide
CTAB	Cetyltrimethylammonium bromide
2,4-D	2,4-dichloropenxyacetic acid
dH_2O	Distilled water
DNA	Deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetra acetic acid
EPPS	*N*-(2-hydroxyethy)piperazine-*N'*-(3-propanesulfonic acid)
g	Grams
G	guanine
H	Hour
H_2O	Water

IPTG	Isopropyl-[beta]-D-thiogalactopyranoside
L	Litre
LB	Luria Bertani
LTR	Long terminal repeat
M	Molar
mer	Oligomer
mg	Milligrams
$MgCl_2$	Magnesium cloride
ml	Millilitres
mM	Millimolar
mRNA	messenger Ribonucleic Acid
MS – RDA	Methylation sensitive Representational Difference Analysis
NaCl	Sodium chloride
ng	nano gram
nmol	nano mole
NaOAc	Sodium acetate
NaOH	Sodium hydroxide
oligo	Oligomer
PCR	Polymerase Chain Reaction
pH	Log hydrogen ion concentration
QTL	Quantitative Trait Locus
RAPD	Random Amplified Polymorphic DNA
RDA	Representational Difference Analysis
RDA-WEEC	RDA-With Elimination of Excessive Clones
rDNA	Ribosomal DNA
RE	Restriction Endonuclease
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonucleic Acid
RNaseA	Ribonuclease A
rpm	revolutions per minute
SDS	Sodium Dodecyl Sulphate
sec	Second(s)
ss	single stranded
SCAR	Sequence Characterized Amplified Regions

SSC	Saline sodium citrate
T	Thymine
T_a	Annealing Temperature
TAE	Tris-acetate EDTA
Taq	*Thermus aquaticus* DNA polymerase
TE	Tris EDTA
T_e	Extension Temperature
Tet	Tetracyclin
T_m	Melting Temperature
Tris	2-amino-2-(hydromethyl) propane-1, 3 diol
tRNA	transfer RNA
U	Unit
UV	Ultraviolet
X-Gal	5-Bromo-4-chloro-indoyl-[beta]-D galactoside

Term	Definition
Adenine (A)	purine base that pairs with thymine (T) in DNA.
Agarose	the neutral gelling fraction of agar commonly used in gel electrophoresis.
Algorithm	a step by step process for solving a problem.
Alignment	the juxtaposition of amino acids or nucleotides in homologous molecules that are assumed to contain residues that are all derived from a single common ancestral residue.
Aliquot	a subsample of a reagent; to divide into several subsamples.
Allele	one of a series of possible alternative forms of a given gene differing in DNA sequence and affecting the structure and/or function of a single product (RNA and/or protein).
Alu family	a short (300 bp), interspersed DNA sequence repeated about 500,000 times in the human genome, and characterized by containing a distinctive *Alu*I restriction site.

Bacterial colony	a clone of bacterial cells.
Base pair	a pair of hydrogen-bonded nucleotides that join the two strands of a DNA double helix. In a double-stranded DNA molecule, adenine (A) forms a base pair with thymine (T), and guanine (G) pairs with cytosine (C).
Base sequence	order of bases in a DNA molecule.
C-value	a measure of haploid DNA content per cell.
Cathode	the negative electrode in an electrolytic cell (such as an electrophoresis chamber) toward which cations migrate.
cDNA	complementary DNA, often refering to a cDNA library made with mRNA and the enzyme reverse transcriptase.
cDNA clone	double-stranded DNA sequence that is complementary to a specific RNA and inserted in a cloning vector such as a plasmid.
Chromatid	the eukaryotic chromosome prior to replication, or one of the two longitudinal subunits of a chromosome after replication, joined by a centromere.
Chromosome	structure containing DNA and proteins in the cell nucleus.
Clone	1.*verb.* to insert a piece of DNA into a vector for subsequent amplification and isolation of that specific piece; 2.*noun.* a piece of DNA composed of a vector and its insert.
Cloning vector	biological carriers such as plasmids, bacteriophage, or cosmids used to amplify an inserted DNA sequence.
Coding strand	DNA strand that is transcribed into mRNA.
Colony - hybridizatrion	method for detecting bacteria that carry a vector with a desired inserted sequence.
Complementary sequence	a sequence of nucleotides related by the base-pairing rules. For example, in DNA a sequence A-G-T in one strand is complementary to T-C-A in the other strand. A given sequence defines the complementary sequence.

DNA reannealing	double-stranded DNA separates into single strands when heated which reanneal back into double strands when temperature is lowered.
DNase	deoxyribonuclease; enzyme that cleaves phosphodiester bonds in DNA, to break the molecule into pieces.
Dot (slot)-blot	a DNA analysis system where sample DNA is directly pipetted onto a membrane, as opposed to the Southern blot procedure of enzymatic digestion, electrophoresis, and Southern transfer.
Double-stranded helix	three-dimensional shape exhibited by two complementary base -paired DNA strands.
Electrophoresis	the separation of macromolecules in the presence of an electric current. Electrophoresis is routinely used to separate both proteins and DNA fragments; allozymes are separated based on differences in net charge, whereas DNA fragments are separated based on differences in size.
Epigenetic	all processes relating to the expression and interaction of genes.
Fingerprinting	separation of the DNA of an individual into defined fragments the lengths of which are determined by the spacing of given restriction of enzyme sites. Numbers and lengths of fragments form a unique 'DNA fingerprint' for an individual.
Gene	a sequence of DNA that functions as a unit (e.g., coding for a specific protein).
Gene family	set of very similar genes derived by duplication of an ancestral gene and subsequent minor alteration in each gene in the family.
Genome	the sum total of all the DNA on a haploid set of chromosomes in the nucleus of an individual, including both coding and non-coding sequences.
Genotype	genetic consititution of an individual organism.
Guanine (G)	purine base that pairs with cytosine in DNA.

Haploid	chromosome number in the gametes of a species, symbolized by "n".
Heterochromatin	chromosomal segments or whole chromosomes that generally exhibit a condensed state throughout interphase and late replication.
Homology	common ancestry of two or more genes or gene products (or portions thereof).
Hotspot	region in DNA where mutations occur at exceptionally high frequency.
Hybridization	formation of a double-stranded nucleic acid molecule from complementary single-stranded molecules.
Hybridization stringencies	the fidelity with which single strands of DNA reanneal depends on the stringency of hybridization determined by temperature and ionic conditions.
Hybridize	to induce the pairing of complementary DNA strands, often from different individuals or species, to form a DNA-DNA hybrid molecule.
Hypervariability	extreme genetic variations between individuals in certain genomic sequences.
Hyper variable region	a segment of a chromosome characterized by considerable variation in the number of tandem repeats at one or more loci.
In vitro	means "in glass" and refers to a biological process carried out in the laboratory separate from an organism.
Insert	the DNA of interest that has been cloned, or inserted, into a vector.
Insertion	placemment of additional nucleotide pairs in a specific site in DNA.
Intergenic	nucleotide sequences located between genes.

Interspersed repetitive DNA	short segments of DNA that have been found in hundreds of thousands of copies interspersed throughout the genome, rather than being serially repeated like satellite DNAs.
Inversion	a reversed chromosome segment.
Inverted repeats	symmetrical nucleotide sequence of DNA that is repeated in opposite orientations on same molecule.
Isochizomer	restriction endonuclease with the same recogniton sequence as another restriction endonuclease.
Label	to chemically "tag" a piece of DNA so that it can be visualized in some manner. Most commonly, DNA is labeled with a radioactive element, enabling detection of its presence using x-ray film. Labeling may also be achieved by colorimetric methods or by chemiluminesence. Visualization of larger amounts of DNA requires less sensitivity and is often possible using stains such as ethidium bromide.
Ligase	an enzyme that covalently joins two single stranded DNA molecules when annealed end to end on a DNA template.
Ligation	enzymatically catalyzed formation of a phosphodiester bond that links two DNA molecules.
Locus	a specific position on a chromosome.
Long-terminal repeats (LTR)	nucleotide sequence that is repeated at the end of a DNA molecule.
Melt	the process of disrupting the hydrogen bonds linking complementary DNA strands.
Melting temperature (*Tm*)	midpoint of the heat denaturation curve for double-stranded DNA.
messenger RNA (mRNA)	single-stranded template RNA that contains information for amino acid sequence of the protein.
Methylation	the chemical process of adding a methyl group to a molecule.

Methylation (me)	one form of methylation, the most common in mammals, invloves the conversion of cytosine to 5-methyl cytosine. Methylation can prevent cleavage of DNA at a restriction enzyme recognition site, for example, Hpa II cleaves at C^CG G
Microsatellite	a class of repetitive DNA. Microsatellites are simple sequence repeats two to eight nucleotides in length. For example the repeat unit can be simply "CA", and might exist in a tandem array (CACACACACA...) be highly polymorphic.
Minisatellite	tandem array of from 10 to 50 copies of a non-coding length (typically 10 to 100 nucleotides) of DNA. Arrays on different chromosomes are usually with different numbers of repeated copies giving rise to unique individual DNA fingerprints.
Mismatch	bases that do not match in "complementary" DNA strands. Depending in the blot wash stringency conditions, some mismatch can be tolerated between hybridized sample and probe DNA complementary
Mobile elements	lengths of DNA that can move from one position to another in the genome.
Multigene family	a collection of identical or near identical genes in the genome. The numbers of gene copies and their distribution amongst chromosomes varies widely between species depending on the gene family in question.
Mutagen	an agent that causes changes in the nucleotide sequence of DNA.
Mutant	organism that carries a modified inherited gene.
Mutation	change in the nucleotide sequence of DNA that is inherited.
Nuclear genome	the portion of the genome contained in the nucleus of eukaryotes, i.e., the chromosomes.
Nuclease	enzyme that cleaves phophordiester bonds in nucleic acids.

Nucleolar organizer region	a region on a chromosome that contains the ribosomal RNA genes and associated spacers.
Nucleotide	one of the monomeric units from which DNA molecules are constructed, consisting of a purine or pyrimidine base, a pentose sugar, and a phosphoric acid group. The nucleotides of DNA are deoxyadenylic acid, thymidylic acid, deoxygualinic acid, and deoxycytidylic acid. The nucleotides are often referred to interchangeably with their corresponding nitrogenous base, i.e., the nucleotide deoxyadenylic acid is often referred to as adenine (represented in a sequence by "A").
ORF	open reading frame; start and stop codons are present around a DNA sequence.
Pellet	the button of particulate material formed after a suspension has been centrifuged.
Phenotype	the physical make-up of an individual as define by genetic and non-genetic factors.
Plasmid	a self-replicating extrachromosomal genetic element found in a variety of bacterial species that generally confers some advantage to the host cell (i.e., resistance to antibiotics, etc.). Plasmids are double-stranded, circular DNA molecules ranging in size from 1 to 200 thousand base pairs commonly used as vectors for cloning.
point mutation	a mutation involving a single nucleotide substitution.
Polymerase	an enzyme that assembles the subunits of macromolecules. DNA polymerases have the ability to synthesize the complementary strand of a single stranded DNA template. Synthesis only extends from existing double-stranded sequence across single-stranded template; thus synthesis can be controlled in reactions such as PCR or sequencing using oligonucleotide primers that will anneal to form short double stranded sections contiguous to specific regions of interest.

Polymerase Chain Reaction (PCR)	a series of thermal cycles of denaturation, annealing of primers, and primer extension catalyzed by a thermostable DNA polymerase, in which a target DNA fragment is amplified exponentially; primers that have nucleotide sequences complementary to the DNA that flanks the target region are added to sample DNA along with a heat-stable DNA polymerase. The DNA is heated to separate the complementary strands and then cooled to let the primers bind to the flanking sequences. The polymerase initiates synthesis of complementary DNA. The reaction is allowed to proceed for a series of replication cycles. Twenty cycles will yield a millionfold amplification; thirty cycles will yield an amplification factor of one billion.
Polymorphism	intraspecific variation. On the DNA level, this refers to differences in base pair sequence between two individuals.
Primers	short pieces of single stranded DNA (10-30 bp) annealed to the 5' end of a DNA template used to initiate synthesis of the complementary strand of the template piece of DNA. Primers can be designed so that they will bind only to a very specific region of the DNA, and will thus initiate synthesis of a targeted sequence (as in PCR or DNA sequencing).
Probe	in molecular biology, any biochemical molecule that can be used to distinguish a specific molecule of interest apart from others of its kind. A DNA probe is a segment of DNA with a sequence homologous to the DNA of interest. If the probe is labeled, the sequence can be visualized independently from the rest of the DNA in the sample.
Purine	a heterocyclic ring molecule with various side chains. (Adenine and guanine are purines found in DNA and RNA.)
Pyrimidine	a heterocyclic ring molecule with six carbons and various side chains. (Cytosine and uracil are pyrimidines found in RNA, while cytosine and thymine are found in DNA.)

RAPD (random amplified polymorphic DNA) An analysis of PCR products produced by short non-specific primers. High levels of polymorphism are often observed using this technique.

RDNA ribosomal DNA; the genes for several classes of ribosomal RNA molecules that go into the construction of ribosomes, usually in long tandem arrays in the chromosomes.

Repetitive DNA nucleotide sequences occurring repeatedly in chromosomal DNA. Repetitive DNA can belong to the highly repetitive or middle repetitive categories. The highly repetitive fraction contains sequences of several nucleotides repeated millions of times. Middle repetitive DNA consists of segments 1 to 500 base pairs in length repeated 100 to 10,000 times each.

Restriction enzyme (endonuclease) an enzyme that cleaves double-stranded DNA. Type I restriction endonucleases are not sequence-specific; type II restriction endonucleases cleave DNA at particular recognition sequences (typically 4-6 bp palindromes).The enzymes are named by an acronym that indicates the bacterial species from which they were isolated, followed by a Roman numeral that gives the chronological order of discovery when more than one enzyme came from the same source. DNA fragments produced by certain enzymes, such as *Eco*RI, can anneal with any other fragment produced by that enzyme. This property allows splicing of foreign genes into *E. coli* plasmids or bacteriophage vectors.

Restriction fragment length polymorphism (RFLP) a polymorphism in an individual, population, or species defined by restriction fragments of a distinctive length. Usually caused by gain or loss of a restriction site, but may result from an insertion or deletion of a fragment of DNA between two conserved restriction sites.

restriction sites short motifs of DNA capable of being recognized by a restriction enzyme leading to the cutting of the DNA molecule into separate fragments. Each restriction enzyme has a unique cutting site.

RNA (ribonucleic acid)	polynucleotide that contains ribose sugar.
RNAse	an enzyme capable of degrading RNA.
RRNA	ribosomal RNA, the nucleic acid component of ribosomes, which functions in translation of proteins from mRNA.
Single copy DNA	usually refers to sequences that appear only once in the entire genome. Specific genes or sequences that are single copy are much more difficult to isolate because they represent such a small percentage of the total DNA of an organism.
Single-copy gene	genes for which only two alleles exist (one from each parent) in a diploid cell.
Size marker	DNA fragments of known molecular weight and base pair length, such as l-phage digested with the restriction enzyme *Hind*III, run on electrophoresis gels for the determination of DNA sample fragment sizes.
Slippage	a mechanism of DNA turnover by which gains-and losses occur of short motifs (usually less than 10 nucleotides) in DNA helix leading to pure and cryptic DNA simplicity.
Southern blot	a membrane onto which DNA has been transferred directly from an electrophoretic gel.
Stringency	in DNA-DNA hybridization or DNA-RNA hybridization, the conditions of the hybridization (such as temperature and concentration of chemical additives) that determine the degree of similarity that will result in formation of hybrid molecules.
Tandem array	multiple copies of a sequence of DNA that are arranged one after another in series. Repeat units can be short nucleotide sequences or entire sets of genes.
***Taq* polymerase**	a thermostable DNA polymerase from *Thermus aquaticus*, thermophilic bacterium. Used for amplification via the polymerase chain reaction.

Thymine	pyrimide base found in DNA.
thymine dimer	adjacent thymine residues in DNA that have been chemically linked, usually by the action of ultraviolet irradiation.
Transform	to cause bacterial cells to take up a plasmid host.
Translocation	change in the position of a chromosome segment within a genome.
Transposable element	a genomic element that can move from site to site in the genome of an organism, either through direct DNA copying (at least in prokaryotes) or reverse transcription from an RNA intermediate (probably the usual mechanism in eukaryotes).
Transposition	see mobile elements. Duplicative transposition occurs when a given DNA region replicates and the extra copy moves to another position in the genome. Non-duplicative transposition occurs when the DNA region moves from one position to another: no extra copies are involved.
Transposon	a segment of DNA flanked by transposable elements that is capable of moving its location in the genome.
Uracil	pyrimidine base in RNA that appears in place of the thymine found in DNA.
Vector	a self-replicating DNA molecule that exists with, but is separate from the genome of the host cell. Many different vectors have been identified and genetically engineered for use in molecular biology. DNA inserted into a vector will be replicated along with the vector. In this manner, DNA of interest can be obtained in large quantities i.e., cloned. For example, the human insulin gene can be cloned into the plasmid vector pBR 322 which in turn will replicate in *E. coli* cultures.

References

Freifelder D. and Malacinski GM. 1993. Essentials of Molecular Biology 2nd Ed. Jones and Bartlett, Boston, Massachusetts.

Hillis DM. and Morits C. 1990. Molecular Systematics. Sinauer, Sunderland, Massachusetts.

Hoelzel AR. and Dover GA. 1991. Molecular Genetic Ecology. Oxord University Press, New York. D. Rickwood, ed.

Kirby LT. 1990. DNA fingerprinting: An introduction. Stockton Press, New York.

INDEX OF FIGURES

INDEX OF TABLES