Using molecular marker technology in studies on plant genetic diversity

Introduction

Contents

- Genetic diversity
- Plant genetic resources
- Measuring genetic variation
- Genetic markers:
  - Description
  - Types
  - Desirable properties
- Comparing major techniques:
  - Technologies
  - Costs

Genetic diversity

- Genetic diversity refers to variation in the:
  - DNA sequence,
  - Amount of DNA per cell or
  - Number and structure of chromosomes
- Genetic diversity is the result of selection, mutation, migration, genetic drift and/or recombination. All these phenomena cause changes in gene and allele frequencies, leading to the evolution of populations

Genetic diversity refers to the variation of genes within species, that is, the heritable variation within and between populations of organisms. In the end, all variation resides in the sequence of the four base pairs that compose the DNA molecule and, as such, constitute the genetic code. Other kinds of genetic diversity can also be identified at all levels of organisation in the nucleus, including the amount of DNA per cell, chromosome number and DNA structure.

The generation of new genetic variation occurs continuously in individuals through chromosomal and gene mutations, which, in organisms with sexual reproduction, are propagated by recombination. Genetic variation is also influenced by selection. The consequences of these phenomena are changes in gene and allele frequencies that account for the evolution of populations. Similar situations can occur through artificial selection such as breeding.

Plant genetic resources

- Plant genetic resources comprise the present genetic variation that is potentially useful for the future of humankind
- Plant genetic resources include:
  - Wild relatives of crops
  - Wild species
  - Traditional varieties and/or landraces
  - Commercial cultivars, hybrids or breeding lines
- We must conserve plant genetic resources for their eventual use

Plant genetic resources comprise the present genetic variation that is potentially useful for the future of humankind. These resources include traditional varieties, landraces, commercial cultivars, hybrids, and other plant materials developed through breeding; wild relatives of crop species; and others that could be used in the future for either agriculture or environmental benefits. Hence, plant genetic resources should be conserved, with the ultimate reason being to eventually use them as a source of potentially useful genetic variation.

Measuring genetic variation

- Efficient conservation and use of plant genetic resources require thorough assessment of the genetic variation they comprise

- Genetic variation can be measured at two levels:
  - Phenotype—the combination of individual traits resulting from a genotype and its interaction with the environment
  - Genotype—the particular genetic make-up of an organism

To conserve and use genetic variation, it should first be assessed, that is, the extent and its distribution need to be determined. Variation can be evaluated on the phenotypic and genotypic levels. Assessment of phenotypic variation focuses on morphological traits—those characteristics that define the shape and appearance of a set of individuals. Some of these traits can be considered as 'genetic' if their presence in related individuals is heritable and not dependent on the environment, meaning that they are associated with a particular DNA sequence.

Assessment of genotypic variation is at the level of the DNA molecule responsible for transmitting genetic information. The DNA molecule is composed of nucleotides, which are organised in a double-helix configuration in increasing levels of complexity up to the chromosomal units.

Genetic markers: description

- Genetic markers identify characteristics of the phenotype and/or genotype of an individual
- Their inheritance can be followed through generations

A genetic marker is a measurable character that can detect variation in either a protein or DNA sequence. A difference, whether phenotypic or genotypic, may act as a genetic marker if it identifies characteristics of an individual's genotype and/or phenotype, and if its inheritance can be followed through different generations.

A genetic trait may not have necessarily observable consequences on an individual's performance. Sometimes, however, this trait may be linked to, or correlated with, other traits that are more difficult to measure and do affect the individual's performance. In such cases, these unobservable genetic traits may be used as genetic markers for the linked traits because they indirectly indicate the presence of the characteristics of interest. The two measures can be correlated, using an analysis of inheritance and studying the distribution of the characteristics in both parents and offspring.

Genetic markers: types

- Morphological traits
- Protein (biochemical) markers
- DNA (molecular) markers

Morphological traits

- Advantages:
  - Readily available
  - Usually require only simple equipment
  - Form the most direct measure of phenotype
- Disadvantages:
  - Require expertise on crop and/or species
  - Subject to environmental influences
  - Limited in number

Traditionally, diversity within and between populations was determined by assessing differences in morphology. These measures have the advantage of being readily available, do not require sophisticated equipment and are the most direct measure of phenotype, thus they are available for immediate use, an important attribute. However, morphological determinations need to be taken by an expert in the species, they are subject to changes due to environmental factors and may vary at different developmental stages and their number is limited.

Protein (biochemical) markers

- Based on the migrational properties of proteins, which allow separation by electrophoresis
- Detected by specific histochemical assays
- Advantages:
  - Require relatively simple equipment
  - A robust complement to the morphological assessment of variation
- Disadvantages:
  - Subject to environmental influences
  - Limited in number

To overcome the limitations of morphological traits, other markers have been developed at both the protein level (phenotype) and the DNA level (genotype). Protein markers are usually named 'biochemical markers' but, more and more, they are mistakenly considered as a common class under the so-called 'molecular markers'.

Protein markers (seed storage proteins and isozymes) are generated through electrophoresis, taking advantage of the migrational properties of proteins and enzymes, and revealed by histochemical stains specific to the enzymes being assayed.

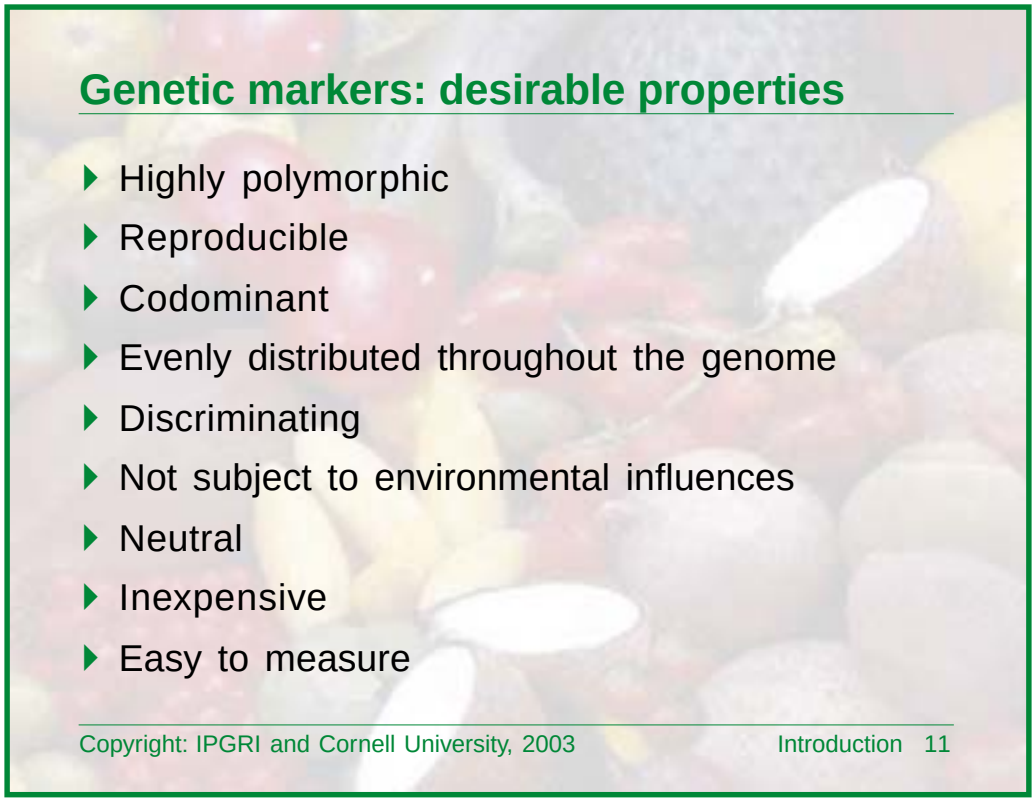
Detecting polymorphisms—detectable differences at a given marker occurring among individuals—in protein markers is a technique that shares some of the advantages of using morphological ones. However, protein markers are also limited by being influenced by the environment and changes in different developmental stages. Even so, isozymes are a robust complement to the simple morphometric analysis of variation.

DNA (molecular) markers

- Polymorphisms detected in the DNA sequence of the nucleus and organelles
- Advantages:
  - Not subject to environmental influences
  - Potentially unlimited in number
  - Objective measure of variation
- Major disadvantage is the need for technically more complex equipment

DNA polymorphisms can be detected in nuclear and organellar DNA, which is found in mitochondria and chloroplasts. Molecular markers concern the DNA molecule itself and, as such, are considered to be objective measures of variation. They are not subject to environmental influences; tests can be carried out at any time during plant development; and, best of all, they have the potential of existing in unlimited numbers, covering the entire genome.

Many different types of molecular markers with different properties exist, as we show below.



A good marker is:

- *Polymorphic*, that is, it is variable among individuals. The degree of polymorphism detected depends on the technology used to measure it.
- *Reproducible in any laboratory experiment*, whether within experimental events in the same laboratory or between different laboratories performing identical experiments.
- *Codominant*. Depending on the type of application, the selected technology must be able to detect the marker's different forms, distinguishing between homozygotes and heterozygotes (codominant inheritance). A heterozygous individual shows simultaneously the combined genotype of the two homozygous parents.
- *Evenly distributed throughout the genome*. The more distributed and dense genome coverage is, the better the assessment of polymorphism.
- *Discriminating*, that is, able to detect differences between closely related individuals.
- *Not subject to environmental influences*. The inference of a marker's genotype should be independent of the environment in which the individual lives or its developmental stage.
- *Neutral*. The allele present at the marker locus is independent of, and has no effect on, the selection pressure exerted on the individual. This is usually an assumption, because no data are usually available to confirm or deny this property.
- *Inexpensive*. Easy, fast and cheap in detecting across numerous individuals. If possible, the equipment should be of multipurpose use in the experiment.

Comparing major techniques

Markers	Number	Codominant	Polymorphism	Locus Specificity	Technicity	Cost
Isozymes	< 90	Yes	Low	Yes	Low	Low
RFLP	Unlimited	Yes	Medium	Yes	High	Medium
RAPD	Unlimited	No	Medium	No	Low	Low
DAF	Unlimited	No	Very high	No	Low	Low
AP-PCR	Unlimited	No	Very high	No	Low	Low
Microsatellites	Unlimited	Yes	Very high	Yes	Low [a]	Low [a]
SCAR	Unlimited [b]	Yes/no	Low/medium	Yes	Medium	Low
CAPS	Unlimited [b]	Yes	Low/medium	Yes	Medium	Low
ISSR	Unlimited	No	High	Yes	Low/medium	Low/medium
AFLP	Unlimited	No	High	No	Medium	Medium
Sequencing	Unlimited	Yes	High	Yes/no	High	High
EST	Unlimited	Yes	Low/medium	Yes	Medium	Medium
SNP	Unlimited	Yes	Very high	Yes	High	High

[a] When microsatellites have already been identified and primers designed.
[b] Depending on other markers already available.

Here we compare major techniques that use biochemical and molecular markers to identify genetic diversity. Criteria used to assign levels within each column (yes/no, low/medium/high, etc.) are based on experience and results described in the literature. We cannot provide a number, or even a range, for each item and its technology because results are highly dependent on the species under study. However, the table gives an objective notion of how techniques can be compared among themselves for a given species and the items in the columns.

RFLP	Restriction fragment length polymorphism
RAPD	Random amplified polymorphic DNA
DAF	DNA amplification fingerprinting
AP-PCR	Arbitrarily primed polymerase chain reaction
SCAR	Sequence-characterised amplified region
CAPS	Cleaved amplified polymorphic sequence
ISSR	Inter-simple sequence repeat
AFLP	Amplified fragment length polymorphism
EST	Expressed sequence tag
SNP	Single nucleotide polymorphism

Costs: how major technologies differ

Procedure	Item	Est. costs (USD, 2002), 96 samples	RFLP	SSR	RAPD	AFLP	Comments
DNA extractions			5.40-101.40	5.40	5.40	5.40	96 micropreps vs. 96 large preps
	Centrifuge and/or microcentrifuge tubes	2.40					
	Or, 96-well plate	3.50					
	Or, 50-ml tubes (large preps)	24.00					
	Extraction buffer	2.00-50.00					96 micropreps vs. 96 large preps
	Misc. supplies (alcohol, lysis buffer,tips, etc.)	1.00-25.00					
Agarose electrophoresis			23.00	[23.00]	23.00	[23.00]	[ ] = optional, quality checking only
	Agarose	10.00 (2 gels)					
	Running buffer	13.00 (2 gels)					
Acrylamide gel electrophoresis				1.50	1.50		
	Acrylamide	0.75 (1 gel)					
	Urea	0.75 (1 gel)					
	Running buffer	-[a]					
Sequencing gel electrophoresis				4.06		4.06	
	Fluorescent primers	2.56					
	Gel (*see* Acrylamide gel electrophoresis)	1.50					
Ethidium bromide method of visualisation			0.20-2.00	0.20-2.00	0.20-2.00		
	Ethidium bromide						
	Photograph of gel	0.20-2.00 (2 gels)					High-density paper vs. polaroid

[a] - = negligible cost.

Continued on slide 14

Costs: how major technologies differ (continued)

Procedure	Item	Est. costs (USD, 2002), 96 samples	RFLP[a]	SSR	RAPD	AFLP	Comments
PCR				74.30	74.30	74.30	
	Taq polymerase	50					1.5 U/rxn. $$ will drop sharply when patent ends
	dNTPs	24					
	primers	0.30-0.40					
	PCR buffer with $MgCl_2$						
Restriction enzyme digests			0.30-30.00				
	Enzymes	0.30-30.00					~10 units/sample, highly variable prices
Southern blotting			30.50				
	Buffers (HCl, NaOH)	4 (2 gels)					
	Nylon membrane	24 (2 gels)					Cost is reduced with every reuse of the filter
	Whatman paper	2.50 (2 gels)					
Hybridisation			3.35				Hybridising one probe to 96 samples
	Buffer	-[b]					
	ST DNA (or other blocker)	0.30					
	LS labelling mix	0.30					
	Radioisotope (^{32}P)	2.50					
	Miscellaneous reagents, tubes	0.25					
TOTALS			62.75-190.25	85.16-110.26	104.40-106.20	83.76-106.76	
(per 96 samples)			(MAS vs. mapping)				

[a] = PCR not needed for the RFLP procedure. Probes can be PCR'd to save time, minimal cost
[b] - = negligible cost

Costs: estimates of general costs

Procedure	Item	Cost (USD, 2002, est.)	Approx. cost per sample (USD, 2002, est.)	Comments
DNA extractions				
	Centrifuge and/or microcentrifuge tubes	$25/1000	0.075	
	Or, 96-well plate	$3.50	0.04	
	Extraction buffer	($2/litre)	0.02-0.50	microprep vs.large prep
	Tris	$270/5 Kg		
	Sorbitol	$70/5 Kg		
	Misc. (EDTA, etc.)			
	Misc. supplies (alcohol, lysis buffer, tips,etc.)		0.01-0.25	
	Drill and pestle (optional)	$100		
	Genogrinder (optional)	$8000		
	Leaf crusher (optional)	$500		
Agarose electrophoresis				
	Agarose	$365/500g	0.11	6 g/gel, ~40 samples/gel
	Running buffer (Tris, EDTA, NaAc)	$6.50/litre	0.05	
	Horizontal gel system	$400		
	Power supply	$400 (2 outlets)		
Acrylamide gel electrophoresis				~96 samples/gel
	Acrylamide	$30/100 ml		
	Urea	$42/500 g		
	Vertical gel system	$1000		
	Power supply	$400 (2 outlets)		
	Gel dryer	$1500		
Sequencing gel electrophoresis				
	Fluorescent primers	$40/1500 samples	0.03	
	Gel	$3		
	Size standards	$12/gel		
	Software for gel analysis	< $100,000		
	Sequencing gel equipment	$100,000		

Continued on slide 16

The costs presented here assume that a basic laboratory is available and set up with glassware, plasticware, heater/stirrers, pH meter, balance, refrigerator, freezer, distilled water, pipettes and tips, centrifuge.

These costs vary between countries, and have been calculated according to U.S. prices. Because costs in many countries will be higher, the ones shown here should be taken as indicative and used to compare different technologies and alternative equipment.

Costs: estimates of general costs (continued)

Procedure	Item	Cost (USD, 2002, est.)	Approx. cost per sample (USD, 2002, est.)	Comments
Ethidium bromide method of visualisation				
	Ethidium bromide	$100/10 g	-[a]	Can be reused, to last a long time
	Transilluminator	$1000-$2000		
	Photography equipment	$3000-12,000		Highly variable system types
	Photograph of gel	$0.10-$1.00	0.0025-0.025	High-density paper vs. polaroid; ~40 samples/gel
PCR				
	Thermocycler	$7000-$23,000		96-well - tetrad (four 96-well plates)
	Taq polymerase	$170/500 units	0.51	1.5 U/rxn. $$ will drop sharply when patent ends
	dNTPs	$250/set of 4	0.25	
	Primers	$15-$25	0.0025-0.0042	
	PCR buffer with $MgCl_2$			
Restriction enzyme digests				
	Enzymes	$0.003-$0.30/unit	0.03-3.00	~10 units/sample, highly variable prices
	Incubator	$500-$1000		
Southern blotting				
	Buffers (HCl, NaOH)	$1.00/litre (avg.)		
	Nylon membrane	$180/roll	6.00/filter (20x10cm)	
	Whatman paper	$124/pack 100 large sheets		
	Blotting system (sponges, tray)	$15		
Hybridisation				
	Buffer			
	ST DNA (or other blocker)	$306/10 g	0.30	0.15 ml. 1 ml used/50 ml hybrid'n buffer
	LS labelling mix	$125/50 unit	0.30	
	Shaking incubator	$2500		
	Radioisotope (^{32}P)	$120/100 µl	2.40	
	Misc. reagents			
	Environmental expenses			
	Waste removal	$500/\month		
	Shielding	$300/\station		

[a] - = negligible cost

In summary

- Defining strategies for good conservation and use requires assessing the variation found in genetic resources
- Plant genetic diversity can be measured through genetic markers—morphological, biochemical and molecular
- No single marker meets all the desired properties
- Choice of technique depends on the nature of the biological question being addressed

By now you should know

- What genetic variation is and how it can be measured
- The main advantages and disadvantages of the different types of genetic markers
- What constitutes the desirable properties of genetic markers

Basic references

Ayad, W.G., T. Hodgkin, A. Jaradat and V. Ramanatha Rao. 1997. Molecular genetic techniques for plant genetic resources. Report of an IPGRI workshop, 9-11 October 1995, Rome. IPGRI, Rome.

Brown, S.M. and S. Kresovich. 1996. Molecular characterization for plant genetic resources conservation. Pp. 85-93 in Genome Mapping in Plants (H. Paterson, ed.). R.G. Landes Company, Georgetown, TX.

Karp, A., K.J. Edwards, M. Bruford, S. Funk, B. Vosman, M. Morgante, O. Seberg, A. Kremer, P. Boursot, P. Arctander, D. Tautz and G.M. Hewitt. 1997. Molecular technologies for biodiversity evaluation: opportunities and challenges. Nature Biotechnol. 15:625-628.

Karp, A., S. Kresovich, K.V. Bhat, W.G. Ayad and T. Hodgkin. 1997. Molecular tools in plant genetics resources conservation: a guide to the technologies. Technical Bulletin No. 2. IPGRI, Rome.

Next

Protein-based technologies
Protein basics

- Protein-based technologies
  - Isozymes
- DNA-based technologies
- Complementary technologies
- Final considerations
- Glossary