A Plant Genetics Primer

Basic Terminology

Steven E Smith and Kandres Halbrook |

ABSTRACT

We present a basic introduction to genetic terminology that has implications for growers and users of native plants. An understanding of underlying genetic concepts is essential for ensuring that plant material of an appropriate source is planted on restoration sites.

KEY WORDS
terminology, concepts, glossary

NOMENCLATURE
USDA NRCS (2004)

Genetics is the field of study that deals with variation and its transmission from parents to their progeny, that is, **inheritance**. Understanding the principles of genetics may be important in many activities involving native plants and restoration tactics. For example, familiarity with genetics may be helpful when devising such things as seed collection and increase strategies, describing a suitable range of environments for planting a particular collection, or determining whether deliberate selection of plants is appropriate, and if so, how best to conduct it. The goal of this primer is to explain some of the basic concepts of genetics applicable to problems that might be encountered by restorationists, land managers, seed and plant producers, plant breeders, propagators, and collectors working with native plants.

GENES, CHROMOSOMES, AND PLOIDY

Genes are the fundamental physical and functional units of inheritance. They are essentially segments of the molecule DNA, the chemical sequence of which is ultimately responsible for producing all growth and development processes of an organism. Particular genes usually occur at a specific location, referred to as a **locus** (plural **loci**), on **chromosomes.** Chromosomes are the threadlike DNA and protein-based structures in cells whose function is the orderly duplication and distribution of genes during cell division.

Each species generally has a characteristic number of chromosomes within each of its body (somatic, nonreproductive) cells. For example, bluebunch wheatgrass (*Pseudoroegneria spicata* [Pursh] A. Löve [Poaceae]), a common grass of the Intermountain and Rocky Mountain regions, has 14 chromosomes in somatic cells. These are in 7 pairs with 1 chromosome of each pair derived from each parent. These pairs of chromosomes are known as **homologs** or homologous chromosomes. Thus, 7 chromosomes from the sperm and 7 from the egg unite to form a plant. This species is considered **diploid** because there are 2 copies of each chromosome. Collectively, the total chromosomal DNA of an organism is referred to as its **genome** with each set of chromosomes making up a single genome. A diploid organism, therefore, has 2 genomes, one from each parent. It is important to note that "genome" has also come to refer to the total chromosomal DNA of a species. Reproductive cells, such as the sperm and the egg and the other cells in pollen and embryo sacs are **haploid,** meaning they contain half the number of chromosomes of a diploid cell. As mentioned above, particular genes usually occur at a specific location on a chromosome and the order of the genes on the chromosome is normally consistent for all members of a species.

Many plant species are **polyploid** because their body cells have more than 2 copies of each chromosome. In fact, many plant genera may contain some polyploid members. In certain types of plants, such as the irises, saltbushes, sagebrushes, and sunflowers, various **ploidy levels** may occur within different species and this can prevent the production of viable offspring following matings (**crosses** or **hybridization**) between individuals with different chromosome numbers.

Photo by Steven E Smith

Figure 1. A portion of an ear of corn (*Zea mays* L.) exhibiting numerous grains with a variety of endosperm color patterns (phenotypes). These phenotypes represent examples of easily recognizable morphological markers that represent different genetic constitutions (genotypes) present among these grains.

ALLELES, GENOTYPE, PHENOTYPE, AND GENETIC DIVERSITY

The specific composition of DNA at each gene is known as an **allele.** Alleles may differ between chromosomes in a diploid species or among chromosomes in polyploid species. Multiple alleles of a gene may be generated by mutations, which are structural or chemical changes in DNA. The different forms of individual alleles are typically symbolized using single letters (for example, A, a, B^1, B^2). **Genetic variation** (**diversity**) is the result of differences among individuals in a group; it is ultimately a function of allelic differences resulting from mutations, and the random assortment of alleles during **meiosis.** Meiosis is a special type of cell division that ultimately leads to the production of sperm and egg.

An individual's **genotype** is the set of alleles it possesses at a certain locus or over particular loci. Genotype can be expanded to reflect allelic constitution over all of the loci in the individual. In contrast, the appearance or performance of an individual is known as its **phenotype.** If different alleles exist at a locus in an individual, the locus is considered **heterozygous** and the individual is a **heterozygote** for that locus. If alleles at a locus are the same, the locus is **homozygous** and the organism is a **homozygote**.

ASSESSING GENOTYPE WITH MARKERS

It is often useful to describe the genotype of a particular individual or of individuals in a population and a number of methods now exist to do so. The simplest method uses morphological markers. A **morphological marker** is an easily observable genetically determined trait that identifies, or *marks*, different genotypes. Flower and fruit or seed color are commonly used morphological markers in plants (Figure 1). Although morphological markers are generally easy to use, their usefulness is limited by their availability. That is, there are typically insufficient morphological markers in a

particular species to be able to uniquely identify individual genotypes. Most well-characterized morphological markers are only available in cultivated species that have been intensively studied genetically.

Molecular markers identify differences in the genetic code (DNA) among individuals. The process of using molecular markers to identify individuals is known commonly as DNA fingerprinting. There are many techniques used to fingerprint, however, the basic concept is the same for most approaches. Plant DNA is segmented by certain enzymes called restriction endonucleases. Different endonucleases recognize different segments of the DNA code. When a particular endonuclease finds its DNA code segment in the plant DNA, it makes a cut. After the DNA has been cut, the DNA segments, now called fragments, are separated by gel electrophoresis. Electrophoresis separates the fragments by size with the largest fragment appearing at the top of the gel and successively smaller fragments below (Figure 2). A stain is added to the gel that reacts with the DNA fragments causing the DNA to become visible. A gel is typically divided into columns with DNA from one plant in each column. If all the plants examined are genetically identical then the location of the fragments in each column in the gel will be identical. Conversely, if there are genetic differences among the plants, one or more columns will have variable numbers or sizes of fragments. Hence, DNA fingerprinting is ultimately based on variation in the length of DNA fragments between two or more individuals of a species.

Another molecular marker system directly uses gene *products* as markers. Proteins are the products of genes and thus differences among particular proteins reflect differences in the underlying DNA that produced them. Isoenyzmes, commonly known as isozymes, are multiple forms of a single enzyme that catalyze the same reaction but they differ in their genetic sequence. That is, isozymes are the product of different alleles of the gene for the enzyme. To use isozymes for genotyping, live plant tissue is ground up and run through gel electrophoresis. Enzymes catalyze reactions; therefore, if the substrate necessary for a specific enzyme to work is added to the gel, the enzyme will react. If a stain or dye is also added, the location of the enzyme in the gel can be detected. As with DNA markers, the location of the enzyme in the gel may denote differences among individuals.

An important consideration when describing differences in markers among members of a population is that the genetic differences detected may or may not be adaptively significant. That is, DNA markers may be able to exquisitely distinguish between 2 individual genotypes, but this does not indicate that the difference between them corresponds to any quantifiable difference in their ability to function in, or even to survive in, their environment. Such variation is called **neutral variation** because it does not appear to be related to the performance of the organism.

MATING SYSTEMS AND GENETIC VARIATION

Plant species exhibit great variability in sexual reproduction and this has significant effects on the genetic constitution of individuals, especially the average heterozygosity or homozygosity that

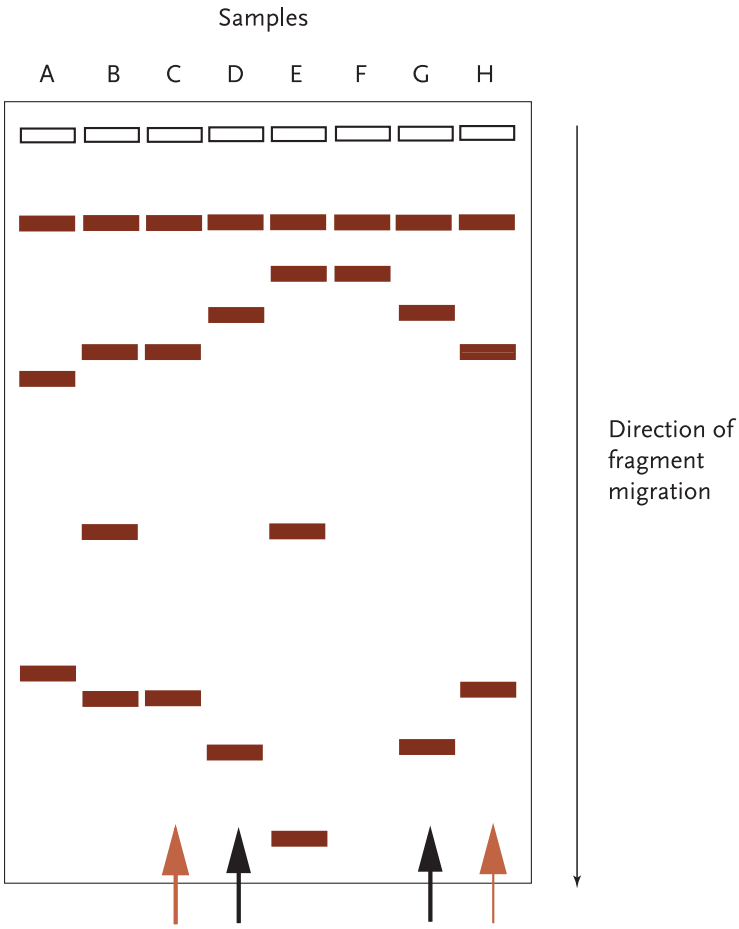

Figure 2. Depiction of electrophoresis gel containing DNA from 8 plants. The DNA has been cut by a restriction endonuclease, the fragments separated based on their size under electrical current in the gel, and then stained. The number and size of fragments can be used to characterize genetic differences among the plants. In this example, 6 genotypes can be identified. Samples A, B, E, and F represent unique genotypes while samples C and H (brown arrows), and D and G (black arrows) share genotypes.

they possess. Reproduction may range from complete **self-pollination**, where seeds are derived from pollen that has come only from the same plant, to complete **cross-pollination**, where pollen that leads to seed is received only from different plants. Individuals in self-pollinated species tend to be highly homozygous. When heterozygosity does occur, either as a result of mutation or cross-pollination between genetically different individuals, repeated self-pollination tends to rapidly reduce the amount of heterozygosity present within offspring of a given individual (Figure 3). Alternatively, in species that tend to exhibit more cross-pollination, heterozygosity is typically much more common and is generally preserved over generations of sexual reproduction.

More than 400 plant species, including many grasses and members of the sunflower and rose families exhibit **apomixis**, which is an asexual reproductive process where seeds are not derived from the union of sperm and egg (**fertilization**) but result from the development of a body cell into an embryo. Plants derived from seeds produced by means of apomixis are genetically equal to the parental plant that produced them. In apomictic species, a single genotype may occupy very large areas and therefore populations may contain relatively little genetic variation.

ALLELES IN GAMETES
(half of each type expected)

	A^1	A^2
A^1	A^1A^1	A^1A^2
A^2	A^1A^2	A^2A^2

Figure 3. Allelic constitutions (genotypes) expected at a single locus in offspring produced by self-pollination of an A^1A^2 heterozygote individual. Among these offspring (shaded area), half are homozygous (black) and half are heterozygous (white). If each of the resulting offspring undergo another round of self-pollination, homozygous individuals—half of the original total—will produce only homozygous offspring while heterozygous individuals will again produce half homozygous and half heterozygous progenies. Thus, each generation of self-pollination will reduce overall heterozygosity at this locus across all plants by half.

While individual species are often characterized as either self- or cross-pollinated, it is more precise to think of reproduction as a continuum ranging from apomixis to self- to cross-pollinated. The mode of reproduction may fluctuate within a particular group of plants because of factors in the environment and the genetic makeup of the plants themselves. For example, many apomictic species exhibit **facultative apomixis**, wherein some seeds result from apomixis while the remainder are produced from typical fertilization. Unfortunately, our understanding of mating systems is rudimentary in most non-cultivated plant species and this may greatly complicate seed collection and increase procedures or such things as genetic interpretation of plant performance in these species.

The ways in which genetic variation exists within and among plant **populations** (members of a species that are reproductively isolated and genetically distinct) is also highly dependent on mating system characteristics. The actual numerical size of a population is largely related to the dispersal distances of pollen and seeds, the density of reproductive plants, and the rate of cross-pollination. Populations of self-pollinated species tend to be made up of many small subpopulations where neighboring plants are closely related to each other genetically due to self-pollination. This contrasts with cross-pollinated species where populations are relatively more homogeneous and most often contain more genetic variation than similarly sized populations of self-pollinated species.

Inbreeding results from matings between individuals that are related to each other and therefore possess alleles derived from a common ancestor(s). Self-pollination is the most extreme example of such a mating (Figure 3), but cross-pollination between other kinds of relatives will also lead to inbreeding. Inbreeding can occur even with random cross-pollination provided the matings take place between relatives. In many cross-pollinated species, inbreeding may lead to an overall decline in plant performance that is known as **inbreeding depression** (see also **outbreeding depression** below). Species that are predominately self-pollinated typically do not display significant inbreeding depression. **Heterosis**, or **hybrid vigor**, can be thought of as the opposite of inbreeding depression. Heterosis is observed when the performance of offspring from a mating far exceeds that of the parents. Typically, heterosis is observed in offspring produced from matings between 2 genetically different parents that will tend to produce highly heterozygous offspring. In many cases, heterosis is greatest when the parents involved are somewhat inbred but differ substantially in their allelic constitutions.

A major concern in many seed collection and increase projects with predominately cross-pollinated species is to collect seeds from a sufficiently large number of plants so that matings between relatives, and therefore inbreeding, are unlikely during seed production. Collecting seeds from at least 50 individuals per population is recommended to reduce the chances for inbreeding depression in these species.

PHENOTYPE, GENOTYPE, AND THE ENVIRONMENT

It is important to realize that not all variation observed within a group of plants is caused exclusively by genetic

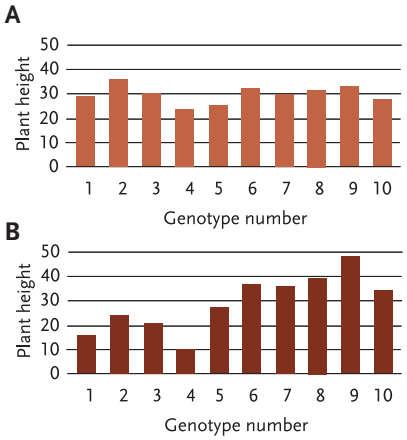

Figure 4. Effect of environmental variability on plant height in vegetative propagules of 10 genotypes. **A.** Genotypes grown at a site with a relatively consistent environment. Here most of the variation among the genotypes would be due to genetic differences, and broad-sense heritability of plant height (the proportion of the total amount of variation in height that is due to genetic factors relative to that due to non-genetic effects) would be relatively high. **B.** Genotypes grown at a site where the environment is more variable with plant growth conditions generally improving as one moves from genotype 1 to 10, for example, mean soil moisture increases throughout the season. Variation among genotypes is due to both genetic (see Figure 5) and relatively large environmental effects, and broad-sense heritability of plant height is lower than in **A**.

differences among individual plants. The environment in which each plant grows will have an effect on the phenotype that it expresses and this effect can be substantial. It is possible in some cases to experimentally partition variation for a given trait into its genetic and environmental sources. Following such an exercise, it is possible to calculate **broad-sense heritability,** which is the proportion of the total amount of variation for a particular trait that is due to genetic factors relative to that due to non-genetic effects (Figure 4). One of the major challenges in plant breeding, and in the assessment of adaptation more generally, is to understand the extent to which environmental variability may be masking true genetic differences in performance among individuals or populations (Figure 5). Moreover, in many cases, individual plants are able to alter their growth or development in response to some environmental signal. For example, many annual plants are able to rapidly shift from vegetative growth to production of flowers during an extended drought. A change in phenotype in response to environmental factors such as this represents an example of **phenotypic plasticity** and this also complicates evaluation of true genetic merit.

EVOLUTION OF ADAPTATION

Genetic variation that exists among plants within and among populations is ultimately due to actions of **evolution**, a collection of processes that results in a change in allele frequencies over time. Members of a given population may exhibit particularly good performance at a location and would be considered **locally adapted**. Such a locally adapted population may be referred to as an **ecotype** (**provenance** or **ecological race**), often using the location as the ecotype's name. The Lassen ecotype of antelope bitterbrush (*Purshia tridentata* [Pursh] DC.), collected in Lassen County, California, represents an example of a commercially available ecotype. Evolution leading to local adaptation is primarily the result of **natural selection**. Natural selection is based on **fitness**, which is measured by the tendency for an individual and its offspring to survive. Fitness may be associated with a plant's longevity, fertility, and its ability to produce highly viable offspring. If variation for a trait associated with fitness is genetically based or "**heritable**" (heritability > 0), then individuals with higher fitness will leave more offspring on average than those with lower fitness. This will result in the genetic constitution of the population becoming more similar to that of the more fit individuals. In many ways, **artificial selection** as practiced in plant breeding programs is directly analogous to evolution through natural selection. Plant breeders **select** (retain) plants with desired phenotypes or genotypes and these individuals contribute alleles to the next generation. Over generations, the frequencies of alleles associated with the desired phenotypes therefore increase within the population.

The fidelity of seed or other propagule increase activities may be affected by natural selection. For example, economical levels of seed production in many plant species can only be achieved in environments that are warmer and drier than those where the seeds will eventually be sown. If such seed production environments differ significantly from the environments of use and differential mortality or fertility of plants occurs during seed production, then natural selection for adaptation to the seed production environment may take place and lead to genetic changes in the population being increased. Such changes associated with natural selection during increase are known as **genetic shifts**. In some crop species, special rules have been established to limit the number of generations or the length of time that a population may be grown for seed production purposes outside its area of adaptation to reduce the potential for genetic shift.

Local adaptation may be the result of the isolated action of individual fitness-associated alleles at particular loci. It may also be due, however, to the evolution of complexes of alleles at many different loci that interact in a coordinated fashion to produce high fitness. When such **co-adapted gene complexes** are broken up, as could occur following hybridization of individuals from ecotypes from very different environments, the fitness of the resulting progeny may be much lower than that of either parental ecotype. This condition characterized by reduced fitness is known as **outbreeding depression**. Outbreeding depression is generally a consideration when a non-local ecotype is introduced into an area where local ecotypes exist. Hybridization between the ecotypes could result in outbreeding depression due to simple replacement of the local alleles with those from elsewhere leading

Photo by Steven E Smith

Figure 5. Field evaluation of different genotypes (individual plants) of whiplash pappusgrass (*Pappophorum vaginatum* Buckl. [Poaceae]) from a single population in an experiment at the USDA Natural Resources Conservation Service, Tucson Plant Materials Center.

to reductions in fitness. Occurrences such as this represent the most common type of **genetic pollution** that may be associated with revegetation projects with native plants. Still, simple introduction of a non-adapted ecotype in the absence of significant hybridization with local ecotypes could also be thought of as an example of genetic pollution.

Acceptance of the importance of local adaptation, concerns about outbreeding depression, and a desire to increase the likelihood of success in revegetation represent the primary bases for the development of **seed transfer rules** or **seed (transfer) zones**. These describe the particular life or climatic zone, or such things as the elevation or distance from the site of collection that propagules of particular ecotypes may be planted to increase the probability of successful performance. Seed transfer rules are relatively well developed for commercial forestry species. Developing realistic seed transfer rules is difficult for many other wildland species because too little is known regarding the genetic basis for local adaptation, especially in environments that have been altered by human activities.

Processes other than natural selection can result in genetic changes in plant populations. One of the most common of these encountered in many plant collection and propagation activities is **genetic (random) drift**. This occurs when particular alleles are very rare or when population size becomes very small, as may occur with a near-extinction event or during seed collection. Under these circumstances, small random samples will tend to differ substantially in allele frequencies from the larger base population simply because of their restricted size. As populations become very small these changes can be considerable. For example, in a population of 1000 homozygous individuals, assume there are 10 different alleles at a particular locus with each present at roughly the same frequency (about 0.1). If a small sample, say of only 8 plants, is randomly taken from this population, then chances are quite high (about 43%) that at least one of the 10 alleles will be completely excluded from the sample and that the frequencies of the other alleles will differ markedly from 0.1 simply by chance alone. As the size of the sample increases, we would expect the sample and the original population to be increasingly similar genetically and therefore less genetic change due to genetic drift to occur. A desire to minimize genetic drift is one reason why most plant and seed collectors try to sample at least 50 individuals per population. **Founder effects** are genetic changes due to genetic drift that are associated with the establishment of new populations from a small number of individuals.

REFERENCE

USDA NRCS. 2004. The PLANTS database, version 3.5. URL: http://www.plants.usda.gov (accessed 22 Aug 2004). Baton Rouge (LA): National Plant Data Center.

AUTHOR INFORMATION

Steven E Smith
Associate Professor
azalfalf@ag.arizona.edu

Kandres Halbrook
Graduate Research Associate
halbrook@u.arizona.edu

School of Natural Resources
University of Arizona
Tucson, AZ 85721

Allele
Specific composition of *DNA* at each *gene*.

Apomixis
Seeds are not derived from *fertilization* but result from the development of a *somatic cell*.

Artificial selection
People retain individuals with desired *phenotypes* or *genotypes* and these individuals contribute *alleles* to the next generation.

Chromosome
Threadlike *DNA* and protein-based structures in cells whose function is the duplication and distribution of *genes* during cell division.

Co-adapted gene complex
Complexes of *alleles* at many different *loci* that interact to produce high *fitness*.

Cross-pollination
Pollination where pollen that leads to seed is received only from different plants. See: *self-pollination*.

Diploid
Two copies of each *chromosome* within each cell. See: *haploid*, *polyploid*.

DNA
Molecule that encodes genetic information.

Ecotype
Locally adapted population from a given environment or locale. Also known as a provenance, or ecological race.

Evolution
Processes resulting in change in *allele* frequencies over time.

Facultative apomixis
Condition where some seeds result from *apomixis* while the remainder are produced from typical *fertilization*.

Fertilization
Union of sperm and egg.

Fitness
Measure of the tendency for an individual and its offspring to survive and reproduce.

Gene
Fundamental physical and functional units of *inheritance*.

Genetic drift
Genetic change that occurs as a result of formulating new *populations* from a relatively small number of individuals.

Genetic pollution
Introduction of *alleles* from a foreign *population*.

Genetic shift
Genetic changes associated with *natural selection* during the production of planting seeds.

Genetic variation (diversity)
Differences among organisms due to factors that can be transmitted from parents to offspring.

Genetics
Field of study dealing with variation among organisms and its *inheritance*.

Genome
Total *chromosomal DNA* of an organism.

Genotype
Alleles an organism possesses at a certain *locus* or over particular loci. See: *phenotype*.

Haploid
One copy of each *chromosome* within a cell. See: *diploid*, *polyploid*.

Heritability
Proportion of the total amount of *variation* for a trait that is due to genetic factors relative to that due to non-genetic effects.

Heritable
Condition where *heritability* is greater than zero.

Heterosis
Overall performance of offspring from a *mating* exceeds that of the parents. Also known as hybrid vigor.

Heterozygous
Different *alleles* exist at a *locus*. See: *homozygous*.

Homologous chromosomes
Chromosomes (usually pairs) derived from maternal and paternal parents.

Homozygous
Same *alleles* exist at a *locus*. See: *heterozygous*.

Hybridization
The process of joining *genomes* from two different individuals.

Inbreeding
Condition resulting from *matings* between individuals that possess *alleles* derived from a common ancestor(s).

Inbreeding depression
Overall decline in plant *fitness* that is associated with *inbreeding*.

Inheritance
Transmission of variation from parents to offspring.

Local adaptation
Ability to successfully establish and persist over extended periods of time in a given location.

Locus
Particular location of a *gene* on a *chromosome*.

Mating
Pollination of flower(s) that may lead to fertilization.

Meiosis
Type of cell division that results in reduction in *chromosome* number and that ultimately leads to the production of sperm and egg.

Molecular marker
Differences in *DNA* that can be used to identify different *genotypes*. See: *morphological marker*.

Morphological marker
Easily observable genetically determined trait that identifies different *genotypes*. See: *molecular marker*.

Natural selection
Process resulting in *alleles* from individuals with higher *fitness* becoming more common.

Neutral variation
Genetic variation that is unrelated to an organism's *phenotype*.

Outbreeding depression
Overall *fitness* of offspring from a *mating* is much less than that of the parents.

Phenotype
Appearance or performance of an organism. See: *genotype*.

Phenotypic plasticity
Extent to which a *phenotype* changes in response to changes in the environment.

Ploidy
Number of sets of *chromosomes* within an organism's cells. See: *diploid*, *haploid*, *polyploid*.

Polyploid
More than two copies of each *chromosome* within a cell. See: *haploid*, *diploid*.

Population
Members of a species that are reproductively or ecologically isolated and genetically distinct.

Seed transfer rules/zones
Outline the particular environment where propagules of particular *ecotypes* may be planted to increase the probability of successful performance.

Self-pollination
Pollination where the resulting seeds are derived from pollen that has come only from the same plant. See: *cross-pollination*.

Somatic cell
Body or nonreproductive cell.