

Toward a Cactus Genome Project: Utilizing Information from Well-Studied Model Plants to Map the *Opuntia* Chromosomes

Andrew H. Paterson
Department of Soil and Crop Science
Texas A&M University
College Station, TX, 77843-2474

Christine Richardson
Professor of Agriculture and Director, Plant Genome Mapping Laboratory
Department of Soil and Crop Science
Texas A&M University
College Station, TX, 77843-2474

e-mail: ahp2343@bioch.tamu.edu
FAX 1409-845-0456

Introduction

Because most forms of life on Earth share a common origin, the genes which encode key functions in different organisms are often very similar. During the past decade, the arrangement of genes along the chromosomes of many important organisms have been described. Such "genetic maps" are similar to roadmaps, showing the locations of individual genes and distances along the DNA between different genes, as a roadmap shows the location of cities along the highway. My laboratory is engaged in the development of genetic maps for many important crops and model systems. I will discuss the potential utility of developing a genetic map for cactus, and suggest some approaches by which cactus improvement might be able to benefit from a large body of genetic information already available for other plants. Because of the novel morphology and physiology of cactus, it has an important contribution to make in improving our overall understanding of plant biology, and we are actively seeking collaborators and sponsors to further genetic understanding of cactus. Because of its established importance as a food and feed crop, I will focus on the genus *Opuntia* (incl. *Nopalea*).

What is a genetic map, and why is it useful?

To the agricultural scientist, a genetic map is a very important resource, both as an encyclopedia for basic research into the evolution and development of a plant or animal and as a tool for engineering new genotypes (cultivars) to meet specific needs of producers, processors, or consumers. Once the location of an important gene on the genetic map is determined, an enormous amount of future time, effort, and cost can be saved. The DNA of an organism is essentially identical at all stages of development (from seed to mature plant). Therefore, the DNA can be assayed for the presence or absence of the gene, and the scientist can select target genotypes based on "DNA fingerprinting" even before the characteristic that is being improved can be observed. This has enormous advantages, for example:

- One can select seedlings which are likely to have superior fruit quality and propagate only these seedlings

- One can select for tolerance of rare events such as freezing tolerance, without having to conduct repeated, expensive field trials.
- One can document the genetic identity (paternity and maternity) of particular genotypes, so that the products of investments in research and development can be identified and protected.
- One can use genes which have been previously identified in other organisms to evaluate their roles in cactus biology, genetics, development, and productivity.

“Mapping” the Genes of Cactus

The information content of DNA is much like that of a language. The DNA language is comprised of four “letters” (nucleotides, abbreviated A, C, T, and G) arranged into long “words” or genes. Small variations in the “spelling” of the words (genes) cause organisms to be different from one another.

Different plant species have different amounts of DNA. The DNA of most plants includes from 100 million to 10 billion letters, organized into 10,000 to 100,000 genes. Cactus is probably approximately in the center between these extremes, in the range of 1 to 10 billion letters. Although it may be possible in the future to consider complete characterization of the cactus genome, at this time it would be as impractical as trying to memorize a complete roadmap of the USA, including every dirt road and driveway. However, a useful summary of the genes that comprise cactus can be obtained efficiently by making a “genetic map”, with mileposts (“DNA markers”) to orient the traveller (scientist). Using the genetic map, one can identify particular regions of the genome that carry important genes, and develop detailed “local” maps of these regions. If necessary, the “street address” of an important gene can eventually be obtained, and the gene can be isolated. Once isolated, the precise “spelling” of the gene can be determined, and alternative “spellings” can be re-introduced into the plant to seek improvements.

Prerequisites for Establishing a Genetic Map of Cactus

The first step in establishing a genetic map of cactus is to identify two diverse genotypes, which can be mated to produce viable, fertile, sexual offspring. Preferably, these genotypes would have the same chromosome number and, ideally, would be diploid (with only two sets of each chromosome, such as is true of humans and many animals. However, many plants often have multiple sets of chromosomes.). Further, it would be most sensible if these individuals also differed in many important agricultural traits—for example, if one parent had superior fruit quality, and the other had superior cold tolerance.

For this purpose, clones from a hybrid plant, or apomictic seeds (i.e., seeds derived from maternal tissue) will not be useful. However, apomixis, in particular, is of much interest in many plant taxa, and cactus may have an important role to play in better understanding apomixis.

From these two individuals, we would like to establish about 50 progeny for making the genetic map, and another 250 progeny of the exact same parents for application of the genetic map. First, a genetic map can be developed quickly and efficiently based on a minimum of 50 progeny. Subsequently, a subset of DNA markers that are representative of the genetic map can be applied to the 250 progeny to test for association to important agricultural traits, such as fruit quality or cold tolerance. Cactus is a particularly facile experimental system in this regard because plants can be clonally replicated, and measurements can, in principle, be accumulated from many environments differing in altitude, latitude, temperature, rainfall, soil type, and conditions.

Clonal replication will also enable different researchers at different locations to accumulate large amounts of information about the same genetic population.

Because of the tremendous amount of work which is likely to be invested in the genetic mapping population, the choice of parents for this population should be given much attention. Ideally, the parents should be of common ploidy (chromosome number), but highly divergent genetically. We can be confident of mapping genes which control traits for which the parents differ markedly, so choice of divergent parents will maximize opportunities to identify agriculturally important traits. Further, divergence at the phenotypic level is often paralleled by DNA-level divergence, meaning that lab analysis will also be easier. A preliminary screen for divergent parents might involve choice of a "core set" of genotypes that contain particular traits of agronomic interest and represent geographically distinct populations. Once such a core set is identified, cytological analysis may help to identify genotypes of common ploidy, and molecular analysis can be employed to screen DNA variation.

Comparative Mapping: Alignment of the Cactus Chromosomes with Those of Well-Studied Plants to Quickly Gain a Better Understanding of the Cactus Genome

Perhaps the most important development in the field of plant genome analysis in the last five years has been the finding that many higher plants share common gene order along sizable portions of their chromosomes. Recently, we (Paterson et al., 1996) have shown that even plants as distantly related as monocots and dicots, which diverged from a common ancestor about 200 million years ago, retain common gene order along small chromosomal segments. As few as 200 chromosomal rearrangements may distinguish the compact genome of *Arabidopsis*, from the genome of sorghum, rice, and other monocots.

The genome of *Arabidopsis* may provide both molecular tools and genetic information that we can use to accelerate our progress with cactus. Because cactus is a dicot, and much more closely related to *Arabidopsis* than other species we have studied, such as sorghum, the chromosomes of cactus should be much more similar to those of *Arabidopsis*. Moreover, more than 10,000 genes have already been cloned and sequenced from *Arabidopsis*, and many of these may prove useful for genetic analysis of cactus.

Our general strategy will be to use DNA probes from *Arabidopsis* for initial genetic analysis of cactus. Because we have already identified a large number of *Arabidopsis* probes that are suitable for genetic analysis in sorghum, these will provide an excellent starting point.

Alignment of the cactus chromosomes with those of *Arabidopsis* and other plants, will greatly accelerate our progress. Recent results have shown that the genes that control traits of major agricultural importance often occur in corresponding locations in plant taxa diverged by 65 million years or more (Paterson et al., 1995). By aligning the genomes of cactus and *Arabidopsis*, we can investigate the possibility that genes that have already been identified and cloned, may be responsible for important traits such as thornlessness, photoperiodic regulation of flowering, plant stature (height), and others. Use of the enormous body of prior genetic information that exists for other plants should greatly accelerate "rediscovery" of the cactus "versions" of these genes.

Summary and Prospects

Like a roadmap, development of a genetic map for cactus will help us to quickly and efficiently proceed to our future destinations. Based on prior experience with other plants, a cactus genome project will quickly provide tools important both to basic investigations of physiology and

development of this remarkable plant, and to breeding genotypes suitable for commercial agriculture. A large body of technology, information, and tools stand poised to contribute to understanding cactus biology. We look forward to developing collaborations with others in the community to identify and develop appropriate genetic materials, measure important phenotypes on these materials, and investigate cactus at the DNA level.

Literature Cited

Paterson, A. H., Y. R. Lin, Z. Li, K. F. Schertz, J. F. Doebley, S. R. M. Pinson, S. C. Liu, J. W. Stansel, J. E. Irvine 1995. Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. *Science* 269:1714-1718.

Paterson, A. H., T. H. Lan, K. P. Reischmann, J. L. Chang, Y. R. Lin, S. C. Liu, M. D. Burow, S. P. Kowalski, C. S. Katsar, T. A. DelMonte, K. A. Feldmann, K. F. Schertz, J. F. Wendell 1996. Toward a unified map of higher plant chromosomes, transcending the monocot-dicot divergence. *Nature Genetics*. (in press).