



BIOTECH Basics

Agricultural Genomics: *Improving crops through genomic research*

What you need to know:

- Plants, like other living organisms, rely on the instructions contained in DNA
- Changes in the DNA recipe influence many agriculturally important traits
- Genomics helps scientists identify these DNA changes
- Plant breeders use this information to develop higher-yielding, more flavorful or disease-resistant crops
- These genetically informed varieties strengthen our planet's food security

References:

- G. P. Morris, P. Ramu, S. P. Deshpande, et al. Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proceedings of the National Academy of Sciences*, 2012; DOI: 10.1073/pnas.1215985110
- L. G. Maron, C. T. Guimaraes, M. Kirst, et al. Aluminum tolerance in maize is associated with higher MATE1 gene copy number. *Proceedings of the National Academy of Sciences*, 2013; DOI: 10.1073/pnas.1220766110
- C. Sautenac, W. Zhang, A. Salcedo, et al. Identification of Wheat Gene Sr35 That Confers Resistance to Ug99 Stem Rust Race Group. *Science*, 2013 DOI: 10.1126/science.1239022
- H. Huo, P. Dahal, K. Kunusoth, et al. Expression of 9-cis-Epoxycarotenoid Dioxygenase4 Is Essential for Thermoinhibition of Lettuce Seed Germination but Not for Seed Development or Stress Tolerance. *The Plant Cell*, 2013; DOI:10.1105/tpc.112.108902
- J. C. Motamayor, K. Mockaitis, J. Schmutz, et al. The genome sequence of the most widely cultivated cacao type and its use to identify candidate genes regulating pod color. *Genome Biology*, 2013; 14 (6): R53 DOI: 10.1186/gb-2013-14-6-r53

Most of the topics traditionally covered in this series discuss genetics and genomics from a human perspective – understanding how changes in the DNA recipe impact traits such as eye color, personality and health. However, DNA is a universal code and studying the genome (the entire set of genetic instructions found in a cell) of other organisms provides important insight into all life on Earth. This edition of *Biotech Basics* focuses on genomic research in plants, with an emphasis on crops critical to our food supply.

Eighty percent of human caloric intake is derived from fewer than a dozen of the approximately 50,000 edible species of plants. Rice, corn and wheat account for two-thirds of consumption. The plants commonly used for agricultural purposes are genetically “exclusive” strains – obtained from centuries of cross breeding for larger yields, easier cultivation and efficient harvesting. Essentially, farmers cross-pollinate two strains of a plant in hopes of producing an offspring with the best qualities from both. Imagine working with Canadian farmers to developing a new strain of sweet corn that sprouts while the ground is still cold. This could involve crossing an existing variety of late growing supersweet corn with one that germinates in cool soil but tastes relatively starchy. The goal is to bring together the early sprouting DNA change from the germination gene with the high sugar version of the sweetness gene. This is a laborious process - multiple crosses are usually required and the correct combination can only be identified when the corn has grown and the ears are ripe.

As modern strains generally cross only a handful of genetic changes, the vast majority of genetic variation at the molecular level remains unexplored. Thanks to technological advances during the last decade, scientists and plant breeders are unearthing this genetic information and identifying new DNA changes associated with agriculturally important traits. These studies have given rise to the field of *agricultural genomics*.

Plant genomes

Like most organisms, plants rely on DNA – composed of four chemical building blocks (abbreviated A, T, G and C), assembled into long strands with the familiar double-helix structure. These strands are tightly wrapped around various proteins and condensed into structures known as chromosomes. Specific stretches of the DNA contain instructions (genes). The majority of plant genomic sequence does not code for genes, but contains highly repetitive sequences of unknown function. Genome sizes vary widely in plants and total chromosome number ranges from 2 to nearly 600. For comparison, consider the human genome with 46 chromosomes. Genome characteristics for several common crops are displayed in table 1.

Modern DNA-based techniques are paving the way for greater understanding of gene function, the impact of genetic variation and their potential applications for crop improvement. The genomes of several key plants have been deciphered in the past few years. Much of this work has taken place in the HudsonAlpha Genome Sequencing Center (www.hagsc.org).

Mining the diversity of seed banks

Having a reference genome is a critical first step to better understanding each gene's function. However, as mentioned earlier, genetic variation exists among different strains of the same plant. Studying this variation links DNA change to distinguishing features of the plant. Working together, farmers and scientists have gathered seeds from domesticated crops, heirloom strains no longer commercially grown and wild relatives of crops. These are stored in seed banks throughout the globe. Such collections are valuable resources for classifying genetic variation across key crops. This includes:

Sorghum:

A particularly drought-tolerant grain first domesticated in Africa some 10,000 years ago, sorghum is a critical component of sub-Saharan diets. Scientists have cataloged the genetic variation present in 971 strains gathered

from global seed banks. Of particular interest were DNA changes associated with the panicle, the part of the plant where the grains form and grow. Dense rows of grain maximize crop yield, but looser spacing allows the grains to dry more efficiently, reducing crop loss from moisture-associated diseases.

Corn:

In the tropics and subtropics where corn is a primary food crop, plants are often grown in acidic soils. These soil types dissolve aluminum from the clays within the ground. This is toxic to growing roots, stunting the growth and development of the plant. Aluminum toxicity is a major threat to food security in food-producing tropical regions. A few corn varieties have been identified that tolerate aluminum-rich soils. A genomic comparison between aluminum-tolerant and susceptible corn strains uncovered key differences at the MATE1 gene. When aluminum is present, this gene is activated, expressing a protein that exports citric acid from root tip cells out into the soil. The citric acid binds to aluminum, limiting its toxicity. Intriguingly, aluminum-tolerant strains have three copies of the MATE1 gene, compared to a single copy for the susceptible lines. The extra copies encode more protein, which exports additional citric acid into the soil to provide greater protection to the roots. A study of 126 corn strains identified the triplicated gene in only three lines, all of which originated in the acidic soils of South America.

Exploring family trees

The wild relatives of modern crops, while agriculturally less productive than their domesticated cousins, have survived thousands of years of extreme environmental conditions. Often this success is due to subtle changes in genetic code. Genomic approaches locate and characterize these changes. Conventional plant breeding techniques incorporate these genetic traits into modern crops.

Wheat:

Over a third of the global population depends on bread wheat for survival. Bread

wheat is part of a larger family of plants tracing ancestry back to an ancient strain of wild wheat. Over thousands of years, this strain was domesticated, crossed with other grain-like plants and cultivated to express specific characteristics and traits leading to the many types of wheat present today.

Genomics is making it easier to untangle this complex family tree. The genomes of related wheat strains and ancestral plants have recently been sequenced. This information is proving especially useful for identifying genes that confer disease resistance. For example, wheat stem rust is a devastating disease that can quickly turn a healthy field of wheat into a blackened mass of twisted stems and shriveled grain. Caused by a fungus, the stem rust is carried on the wind, allowing it to spread rapidly from field to field. Fifty years ago, the introduction of rust-resistant wheat strains helped reduce the global incidence of famine. Unfortunately, the fungal DNA frequently mutates and some of these random changes allow it to infect what were previously resistant strains of wheat. One mutated version, known as Ug99, currently threatens the wheat growing regions of Africa and Central Asia.

As a first step to developing new strains of bread wheat, scientists analyzed the genome of an ancestral wheat species resistant to Ug99. A gene called Sr35 appears to confer this immunity. When the genomes of modern and ancestral wheat were compared, the Sr35 gene was found to be missing in bread wheat and other modern varieties. Initial studies suggest that transferring this gene into bread wheat will confer effective resistance to Ug99, preventing this devastating form of wheat rust.

Lettuce:

A similar comparative approach identified a genetic control for lettuce seed germination. Commercial varieties of lettuce, predominantly grown in California and Arizona, do not sprout in hot weather. During the summer months, lettuce growers must pre-chill either the seeds or the soil to improve yields. A wild ancestor of lettuce germinates across a wide range of temperatures. Genomic studies

comparing DNA sequences detected a genetic regulator of a plant hormone that inhibits hot weather germination. The gene is silent in the wild ancestor, but active in modern varieties. This information can be used to develop warm-germinating varieties of lettuce.

Improving selection efficiency

Recognizing a specific trait-based DNA change allows the inheritance of that trait to be tracked. Hundreds of seedlings can be rapidly screened for the presence of genetic markers that confer drought or disease resistance, increase the size of fruit or improve the flavor profile. This identifies the desirable plants long before they reach maturity, offering a considerable cost and time savings over traditional methods that require direct analysis of the plant's fruit or grain.

Cacao:

The beans of the cacao tree are harvested and processed to produce chocolate. Although not a key agricultural crop from a nutritional perspective, millions around the world would rank chocolate near the top of their "must have" list. There are several varieties of the Cacao tree, ranging in yield and flavor profile. A popular variety of the Cacao tree was recently sequenced by HudsonAlpha's Genome Sequencing Center as part of a collaborative project with Mars Incorporated. Identifying the combination of genetic changes associated with high yield and pleasing flavor speeds the selection and breeding process, resulting in productive trees that produce high quality cocoa.

The challenges ahead

The demand for crop production is rising due to increased human population, greater worldwide meat and dairy consumption and the expanding role of biofuels. Studies suggest that agricultural production must double between 2005 and 2050 to meet this growing need.

Developing new, high-yield seeds adapted for present and future environmental conditions is a cornerstone of increased food production. This begins with the ability to locate and characterize agriculturally important versions of specific genes. These discoveries can then be shared with farmers and commercial plant breeders who are developing new varieties of crops. Such a collaborative approach blends the emerging field of genomics with the ancient practice of agriculture, increasing yields and strengthening global food security.

- Neil Lamb, Ph.D.

director of educational outreach
HudsonAlpha Institute for Biotechnology

Table 1: Comparing the genomes of various crops

Crop	Genome size (in millions of bases)	Number of chromosomes	Number of genes
Bread Wheat	17,000	42 (6 sets of 7)	~ 95,000
Corn	2,300	20 (2 sets of 10)	~ 32,000
Rice	430	24 (2 sets of 12)	~ 41,000
Soybean	1,115	20 (2 sets of 10)	~ 46,500
Sorghum	697	20 (2 sets of 10)	~ 35,000