

## Genomics tools for accelerating plant breeding

Alan H. Schulman<sup>1,2)</sup>

*<sup>1)</sup>Plant Genomics, Biotechnology and Food Research, MTT, ET-talo, 31600 Jokioinen*

*<sup>2)</sup>MTT/BI Plant Genomics Laboratory, Institute of Biotechnology, University of Helsinki, P.O. Box 65, 00014 Helsinki*

Plant breeding is facing simultaneous challenges from a changing climate, increasing prices for non-renewable inputs, and evolving consumer demands. Biotic and abiotic stresses are increasing with climate change; sustainable and stable production will require higher yields with lower inputs; consumers expect health-promoting, safe and traceable food. To meet these challenges, breeding requires more efficient tools with which to unlock and apply existing genetic diversity, to understand the relationship between genotype and phenotype, and to apply the approaches of biotechnology where appropriate. The growing genomics toolbox, based on genome projects for crop plants, offers much promise for acceleration of plant breeding. Here, these approaches will be explored with an emphasis on barley and wheat, which are the key cereal crops of Europe.

plant breeding, genomics, abiotic stress, disease, molecular markers

## Challenges faced by plant breeding

We are currently facing a conjunction of demographic, climatic, natural resource, and economic trends that pose unprecedented challenges for agriculture and the research base that supports it. The world population will double by 2067 at current rates; we will need to double food production in less than 60 years and simultaneously alleviate the poverty that makes the increasingly expensive food unavailable to so many; over 900 M are undernourished and 5 M children under the age of 5 die annually of hunger.

Food production could be increased by planting more land with food crops. However, there is little suitable new land to claim; total arable land has not increased in 50 years, more land will be required for dwellings, and land degradation combined with sea level rises will lead to net loss of arable land. Since the 1950s, 2000 M ha of land have been degraded. In addition to sea level rises, global warming is threatening agricultural production with droughts, heat waves, plant diseases shifting to new areas, more storms and climatic variability, and a general shift in growth conditions. It is estimated that a 2°C increase in temperature in Asia, Africa, and Latin America will lead to a 20 - 40 % decrease in crop yields. Crop production in Finland and its Nordic neighbors likewise face major challenges from climate change, affecting diseases, plant growth rhythms, and water availability.

Another solution might be to increase productivity of the key grains through better or more fertilizers or irrigation. According to the FAO, Denmark, UK, and France are currently producing at theoretically maximum attainable yields for the available wheat varieties. Some major wheat producers, such as the Ukraine or Poland could produce more with better management practices. Increasing irrigation is difficult due to expense, competing demands for water, and the drought brought on by global warming. Most of Africa suffers from a lack of fertility and lack of fertilizer; in Bangladesh, production is limited by boron deficiencies in a soil, something cheaply remedied. However, the prices of most agricultural inputs are rising steeply, or are very limited. High quality phosphate rock, together with nitrate and potassium a key plant nutrient, is more limited than oil; poor quality phosphate is loaded with toxic cadmium. Nitrates are produced by an energy-intensive process that causes fertilizer prices to rise as does the price of oil. The same increasingly expensive oil is used for the tractors needed to till the land. The problem of non-renewable agricultural inputs, besides threatening productivity, drives up production and food prices. Organic agriculture is not productive enough with today's plant varieties and depends too heavily on manure (itself from non-organic sources), to offer a world-wide solution for solving the food crisis.

Adding to the problem is the good intention of increasing the use of renewable energy sources. The International Energy Agency predicts that by 2050, 27% of all fuels for transport will be biofuels. Biofuel production whether based on rapeseed or palm oil, sugar cane, cassava, or maize is currently claiming crops and land that could feed people, driving up food prices. In 2010, 98 % of cassava chips exported from Thailand, the world's largest cassava exporter, went to China to make biofuel.

## Why can genomics help breeding?

Crop production is not only a question of environment, but also of the genetic background of the plants. Crop yield, insect and disease resistance, drought and flooding tolerance, efficiency of converting fertilizer to improved yield, and nutritional quality are all dependent on the particular combination of genes in the variety a farmer plants. Breeders attempt to produce the best combination of genes for the use and growing conditions they envisage. Genetic crossing and selection of offspring based on their performance is, however, a painstaking process; keeping pace with the rapid changes in environment and demand is a major challenge.

What is needed is a way to help plant breeding be more efficient. Especially for traits such as disease and drought resistance and ability to grow on unfertilized soils, the genetic diversity of the wild relatives of crops needs to be used, present a great challenge for conventional plant breeding. Over the past few years, however, a quiet revolution has been taking place. The genome projects, in which all of the genes and their order is discovered, has finally taken on the major food crops.

Genomics, meaning the analysis of the role and properties of all genes in an organism simultaneously, is beginning to make possible the accurate and rapid use of germplasm resources and the acceleration of breeding.

### **A genomics toolkit for plant breeding**

The requisite scientific capabilities and resources are now converging for key agricultural species, including the cereals, oilseed crops, fiber crops, and tuber crops, with the horticultural crops following closely behind. These methods and resources together have led to the development of a genomics toolkit needed for achieving the key breeding goals of the 21<sup>st</sup> century: improved quality and functionality, adaptation to novel end-uses, sustainability and abiotic (climatic) stress and disease resistance.

#### ***Tools for assessing genetic diversity and using it in breeding***

The key tools for assessing and applying genetic diversity in breeding are molecular markers. Molecular or DNA markers, which in essence deliver a rapid and informative summary of the genotype, a genetic “fingerprint” (Fig. 1) of the germplasm or crossing parents and offspring, are increasingly used by the breeding community in Marker Assisted Selection. Development and deployment of more rapid and robust, less expensive, and denser molecular markers is a continuing area of development and attracts great interest both for basic research, pre-competitive but applied work, and in practical breeding and germplasm management. There will be a continuing need both to sample the genetic diversity of the genome and have the ability to recognize specific allelic forms. The most commonly used methods currently are microsatellites,

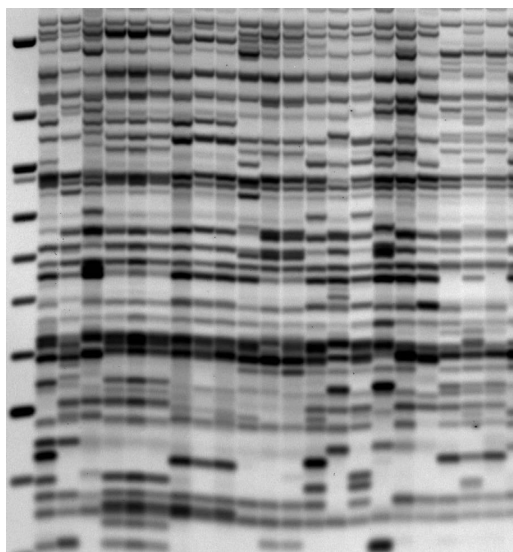


Figure 1. Retrotransposon-based fingerprint of barley varieties.

AFLP, retrotransposons and single nucleotide polymorphism (SNP). The area undergoing perhaps the most rapid development is SNPs. For example, a recent effort in barley has in 2011 produced a platform for determining the allelic state at 9000 SNP markers simultaneously, producing the densest genetic map of barley ever made. This map has allowed the identification of many genes responsible for key agronomic, sustainability, and quality traits.

#### ***Tools for accessing the genome for sustainability and quality***

A logical extension of the development of molecular markers, which sample the genome at intervals of one every 500 000 nucleotides or so, is the determination of the entire genome – “sequencing” – of crop plants. Rice was the first crop plant to be sequenced; this has been followed by many others. For Finland, the key crops of barley, wheat, potato, and turnip rape are in the process of being sequenced or being released to the scientific community. For breeding, having a genome sequence allows the content and order of all genes to be known, allowing estimates to be made, based on predicted gene function, of the likelihood of getting desired traits in combination through crosses. Moreover, the genome sequence allows gene discovery, which is the determination and validation of the genes behind desired traits. Gene discovery is needed for the efficient analysis and application of genetic diversity. For example, new resistance genes for plant diseases constantly need to be discovered and brought into plant breeding programs to overcome the ever-evolving pathogens. Knowing the identity of a resistance gene on the DNA level enables rapid screening of wild germplasm from regions where both the ancestor of a crop species and its disease evolved and then accelerated introgression of the resistance gene into elite cultivars through the application of molecular markers. Because genes, particularly in the cereals are largely in the same order from species to species, we are now able to apply the knowledge gained in already-sequenced genomes such as rice,

the model *Brachypodium*, sorghum, and maize to predict where genes will be in Finnish crops such as barley and wheat. An example of this approach to find a gene of interest through “positional cloning” is shown in Figure 2. In this case, we developed molecular markers based on sequenced genomes in order to genetically map a disease resistance gene. We then were able to transfer this information onto a physical piece of DNA, called a BAC clone, in order to find the gene itself. Ultimately, having access to the genome and knowing the breeding value of each gene will allow effective Genome Wide Selection to be applied in plant breeding.

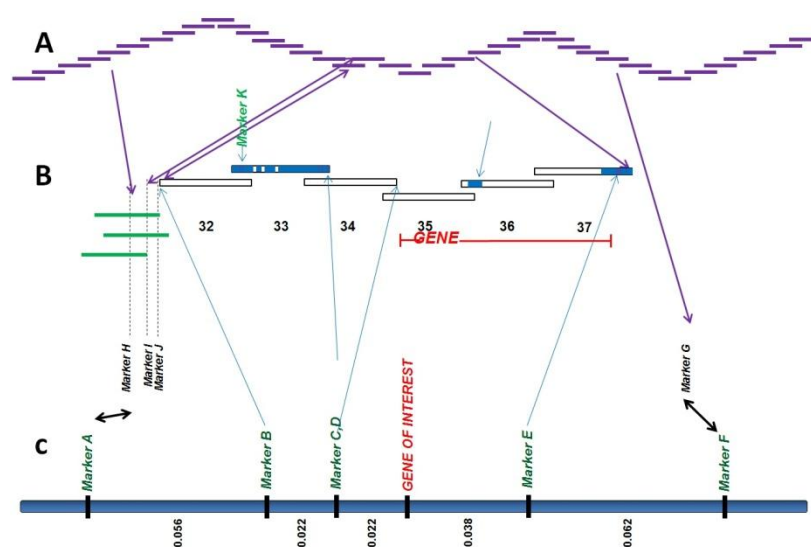


Figure 2. Scheme for map-based cloning of a gene of interest. (A) BAC library of large DNA inserts ordered according to position on chromosome. (B). BAC clones chosen for sequencing based on presence of genetic markers on them. Physical position of gene is shown (C). Genetic markers and recombination distances on fine map, flanking gene of interest.

### Genomics Approaches for Understanding Crop Traits

Many or most traits of interest that control the sustainability and value of the crop, ranging from biotic and abiotic stress resistance to grain or tuber quality and plant architecture, are derived from the action of multiple, interacting genes. Understanding and improving these traits necessarily requires genomics tools. Technologies to identify differentially expressed genes and to analyse gene expression levels on a genomic scale are being applied and developed for the major crops, in order to determine which are the genes that need to be focused on in diversity mining and in breeding programs. For example abiotic stresses, particularly drought, low temperature, and salinity, place major limits on cereal productivity. Crop species belonging to the tribe Triticeae represent, world-wide, the main foodstuff sources for men and animals; they are cultivated from the Arctic to the Sahara margins. Such a great geographic range already suggests that the Triticeae genomes should contain genes for wide environmental adaptability and good stress resistance. The identification of the genetic components of stress tolerance is, therefore, a requirement to ensure further breeding progress since the traditional selection process has met only limited success due to genotype  $\times$  environment interactions. Complex traits such as grain quality or plant architecture is comprised of many factors, among them starch structure, protein content and the presence of epitopes for allergic response, phytate and mineral content and availability, fibre content, and carotenoid content. Genomics offers tools for identifying candidate genes controlling these traits, positionally cloning them, verifying their function, and developing closely linked markers for the deployment of the best alleles through MAS. This strategy represents a major goal for the next decade. These traits may be conferred by multiple genes whose products cause major changes in cell physiology.

### Tools for Testing and Validation of Candidate Genes

A necessary step, following the identification of the gene behind a key trait needed for agriculture, is the validation of that gene's function. This is needed because even in the “model” crops and species such as rice, *Brachypodium*, and *Arabidopsis*, where the genome has been sequenced, no one has rigorously established the function of every gene through physiological and biochemical studies. The function is often inferred through structural comparisons to other known genes, but this frequently leads to circular reasoning or errors due to misleading similarities. To overcome these

problems, various “reverse genetics” tools are being developed and applied. These include mutagenesis and gene silencing methods that allow the specific knock-out of individual genes in plants carrying the trait of interest in order to see the effect on plant performance. In addition, other methods, particularly transgenic introduction of the gene to a plant not having the trait of interest allow evidence to be gathered for a gene’s function.

## **Conclusions**

Genomics will not, by itself, solve world hunger, tame the population explosion, or bring world peace. However, it will help give breeders a fighting chance to provide the varieties that farmers and consumers need. The Green Revolution, begun with the work of Nobel Peace Prize Laureate Norman Borlaug in Mexico in the 1940s, staved off famine in India and elsewhere, and raised the productivity of rice and wheat by up to ten-fold. It gave the world some extra time as the population continued to increase, although this time has not in the end been used effectively to solve the world food supply. The Green Revolution was, surprisingly, based on only a few genes, particularly those controlling plant height, which enabled grain to use fertilizer without lodging. What is needed now is to move from the Green Revolution to the Gene Revolution, based on knowledge and breeding of the genome as a whole. For this reason, progress in plant breeding will increasingly rely on the sort of techniques previously found in the domain of basic research, as the line between basic and applied is increasingly blurred.