Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in
Plant Biology

Breeding-assisted genomics

Jesse Poland[1,2]

The revolution of inexpensive sequencing has ushered in an unprecedented age of genomics. The promise of using this technology to accelerate plant breeding is being realized with a vision of genomics-assisted breeding that will lead to rapid genetic gain for expensive and difficult traits. The reality is now that robust phenotypic data is an increasing limiting resource to complement the current wealth of genomic information. While genomics has been hailed as the discipline to fundamentally change the scope of plant breeding, a more symbiotic relationship is likely to emerge. In the context of developing and evaluating large populations needed for functional genomics, none excel in this area more than plant breeders. While genetic studies have long relied on dedicated, well-structured populations, the resources dedicated to these populations in the context of readily available, inexpensive genotyping is making this philosophy less tractable relative to directly focusing functional genomics on material in breeding programs. Through shifting effort for basic genomic studies from dedicated structured populations, to capturing the entire scope of genetic determinants in breeding lines, we can move towards not only furthering our understanding of functional genomics in plants, but also rapidly improving crops for increased food security, availability and nutrition.

Addresses
[1] Wheat Genetics Resource Center, Department of Plant Pathology, Kansas State University, Manhattan, KS 66506, United States
[2] Department of Agronomy, Kansas State University, Manhattan, KS 66506, United States

Corresponding author: Poland, Jesse (jpoland@ksu.edu)

Current Opinion in Plant Biology 2015, **24**:119–124

This review comes from a themed issue on **Genome studies and molecular genetics**

Edited by **Insuk Lee** and **Todd C Mockler**

http://dx.doi.org/10.1016/j.pbi.2015.02.009

Introduction

In the land of inexpensive sequencing, breeders will be royalty. As readily available genotyping permeates biology, functional genomics has developed into a whole-genome association paradigm [1–3] with current studies generating complete genomic sequence of the populations under investigation [4,5]. Likewise, genomic profiling is becoming commonplace in breeding with many programs completely genotyping the full complement of new material coming through the program [6]. In this reality, where sequencing data is easily obtained on everything, those who know how to develop, maintain and evaluate large populations hold the power to drive our understanding of the genome through laying the foundation to evaluate phenotypes on the scale needed to complement full genomic sequence.

Early quantitative genetics studies focused on the idea of valuable genotyping information with relatively tractable phenotypes. The idea of 'selective genotyping' [7] shortly transitioned to the concept of 'selective phenotyping' [8,9] as it became apparent that genotyping costs were quickly decreasing while phenotypic evaluation would remain costly. This trend has only accelerated; even with current efforts to develop high-throughput phenotyping, the throughput and decreasing cost of genotyping is outpacing developments on the phenomics side.

In this context, there has been increasing focus on developing high-throughput phenotyping to complement the advancements in high-throughput genotyping. As highlighted below, emerging developments in this field show great promise and it is not unreasonable to think that phenotyping capacity will soon keep pace with genotyping. At this point, genomics studies will only become limited by the number and size of populations that can be developed, maintained and planted. It is in this area that breeding programs excel and this resource will be key to the future of plant genomics.

Advances in genomics-assisted breeding

The realization of inexpensive genetic markers through array-based genotyping [10•,11] and genotyping-by-sequencing (GBS) [12–15] has paved the way for applying genomics-assisted breeding (i.e. genomic selection) in plant breeding programs. First proposed by Meuwissen *et al.* in 2001 for application in animal breeding [16], the approach of using genome-wide molecular markers to predict overall performance of individuals in breeding populations has now been demonstrated across many plant species [17–21]. These studies have collectively shown that prediction accuracy for complex traits is sufficient to merit many applications of genomic selection across breeding programs. Reports in multiple crops, including maize [22], soybean [19] and wheat [23], have also demonstrated the use of the very inexpensive GBS platform as a suitable tool for genotyping in a genomic selection context.

In breeding programs, continual selection of favorable individuals leads to discarding a large majority of the population during early generations. While molecular markers remained costly, the investment of resources in genotyping lines and individuals that would not be retained was not justified. However, with the rapid decrease in genotyping costs, the paradigm is rapidly changing. The marginal investment to genotype the entirety of new breeding lines has largely dropped below the cost to develop those lines. Therefore, rapidly growing numbers of breeding lines are being genotyped and are soon to approach the entirety of the breeding program.

Power

The power to understand quantitative traits at a functional level is largely facilitated by large population sizes. This has been highlighted and demonstrated many times over in the context of plant genomics. Through simulation studies, Yu *et al.* [24] clearly demonstrated the power of population size for finding small effect QTL in the maize nested association mapping (NAM) population. This power was fully realized when studies from NAM demonstrated the ability to identify and map small effect QTL across many different traits [5,25–27]. Many of these studies showed extreme power to identify minor QTL with estimated small effect sizes even less than the observation unit (e.g. <1 day for flowering time [25], <5% diseased leaf area for Northern Leaf Blight [27]). Comprised of 5000 inbred lines, the maize NAM is an order of magnitude larger than previous mapping populations and clearly represents a considerable investment of resources in development. Similar studies investigating and demonstrating the power of large populations have also emerged.

The power of large populations such as NAM for dissection of complex traits is now well-recognized [24]. While the maize NAM was many times larger than previous studies, this must be put in context of breeding programs which routinely operate on a scale that is orders of magnitude larger than typical quantitative genetics studies (Table 1). As such, breeding programs represent population sizes of the magnitude needed to have the superior power for detection of minor QTL and complex genetic architectures, while at the same time being directly relevant to agriculture and food security.

The potential of breeding programs

The dynamics of a breeding program offer opportunities for a range of different studies in plant genomics, just a sample of which is highlighted here (Figure 1). Current efforts in understanding genetic diversity and climatic adaptation [28•], selection theory, recombination studies [29], and *de novo* mutation [30] all fit within the context of the mechanics and germplasm found in breeding programs.

Many studies are emerging, highlighting the power of artificial selection to understand evolutionary processes while identifying underlying traits. Recent studies have implicated suites of genes for domestication and crop improvement ranging from plant growth and reproductive timing, to morphology, architecture and physiology [31]. In the context of wild and improved breeding germplasm, this understanding of selection has been demonstrated through mapping of selective sweeps with identification of flowering time and phenology loci in wheat [10•]. Within the context of breeding programs, these types of analyses could be conducted directly, implicating not only important regions of major change for domestication when paired with genomic information from wild relative, but also for targets of selection to local environments. Overall, it is well recognized that the selection process in breeding is well suited to understand the process of phenotypic evolution.

The need for phenotypes

Breeders are notoriously good at phenotyping. A released variety is very unlikely to be too tall or too short as this is

Table 1

Size of selected public sector wheat breeding programs relative to genetic populations. The numbers for breeding programs represent a *per year* estimate

	Populations	Inbred lines	Observation plot	Full yield plots
Busemeyer *et al.* [37]	4	647		1920
Maize NAM[a]	25	4700	48,000[b]	
KSU wheat[c] breeding	1000	5500[d]		15,000
CSU wheat[e] breeding	900–1400	4800[f]	30,000	17,000
CIMMYT bread wheat breeding[g]	3600–4000	46,000	147,000	47,000

[a] Buckler *et al.* (2009).

[b] Largest growout of the original maize NAM in a single year (Edward Buckler, personal communication).

[c] Kansas State University, Hard Winter Wheat Breeding program (Allan Fritz, personal communication).

[d] Inbred lines in the KSU program include 2500 F_5 inbred lines and 3000 double haploid lines derived from approximately 300 populations.

[e] Colorado State University, Hard Winter Wheat Breeding program (Scott Haley, personal communication).

[f] Inbred lines in the CSCU program include ~1800 F_5 inbred lines and 3000 double haploid lines.

[g] International Maize and Wheat Improvement Center, Spring Bread Wheat Program (Ravi Singh, personal communication).