Science Development
2025, Vol. 6, No. 1, pp. 17-24
https://doi.org/10.11648/j.scidev.20250601.12

Review Article

Exploring the Intricacies of Crop Yield Performance Through Genomics

Melkam Anteneh*

Debre Zeit Agricultural Research Center, Ethiopian Institute of Agricultural Research, Debre Zeit, Ethiopia

Abstract

Genomic analysis is central to our effort to decode and enhance crop yield performance, an endeavor that, while promising, comes with its challenges. Yield traits are notoriously intricate; they weave a complex tapestry of genetic and environmental interactions that require extensive datasets for meaningful analysis. This research paper explores the complicated nature of these yield traits, emphasizing the urgency of developing large-scale databases and understanding the nuanced interplay between genotype and environment. By identifying genes linked to markers associated with yields, we can streamline breeding programs, making them faster and more precise. However, we must be candid about the limitations inherent in genomic analysis. It is undeniably powerful for boosting crop productivity, but recognizing its boundaries is equally essential. It improves our data analysis techniques and fosters a comprehensive understanding of yield genetics. SNP markers on high-density arrays may indicate genetic associations with phenotypic variation, but GBS-based genotyping methods may be better suited for identifying causal genetic variants in complex crop species, influenced by rare alleles not adequately represented on SNP arrays. Only then can we promote advancements in plant breeding, ultimately ushering in an era of increased crop productivity. Techniques like genome-wide association studies (GWAS), QTL mapping, and marker-assisted selection help accelerate breeding programs by directly targeting specific genes or loci responsible for these traits. High-throughput genotyping allows for a detailed assessment of genetic variation within and between crop populations. We study genome size, heterozygosity, and identify regions of the genome associated with traits of interest such as yield, stress tolerance, or disease resistance. We detect genomic regions where natural or artificial selection has favored specific alleles, leading to reduced genetic diversity and altered patterns within and between populations. We introduce and utilize genetic variation within specific regions of the genome. Understanding the genetic mechanisms that result in increased vigor and performance in hybrids compared to their inbred parents is critical. Generally, genomic-assisted breeding (GAB) revolutionizes crop improvement by using modern molecular tools to enhance accuracy and efficiency in plant breeding. GAB leverages techniques like marker-assisted selection, association mapping, and genomic selection to identify desirable traits, genes, and genomic regions associated with specific traits, ultimately accelerating the development of new crop varieties.

Keywords

Agronomic Traits, Crop Diversity, DNA Sequence, and Genetic Architecture

---

*Corresponding author: jarcmelkam@gmail.com (Melkam Anteneh)

Received: 19 May 2025; **Accepted:** 7 June 2025; **Published:** 30 June 2025

1. Introduction

The genome, often described as the blueprint of life, holds the intricate code that defines every living organism on Earth. Composed of DNA (deoxyribonucleic acid), this molecular marvel serves as a comprehensive instruction manual that dictates the development, function, and uniqueness of each organism. The study of genomes has revolutionized biology, medicine, and our understanding of evolution, offering profound insights into both the unity and diversity of life forms. At its core, the genome consists of a sequence of nucleotide bases—adenine (A), cytosine (C), guanine (G), and thymine (T)—arranged in a double helix structure. This structure, famously elucidated by James Watson and Francis Crick in 1953, not only clarifies the physical basis of heredity but also underscores the astonishing complexity encoded within its relatively simple form. The Human Genome Project (HGP), a landmark international endeavor completed in 2003, marked a watershed moment in genomic research. By mapping and sequencing the entire human genome, scientists unlocked a treasure trove of information. Genomes serve as the blueprint of life, containing the complete set of genetic instructions necessary for organisms' development, function, and regulation. The journey to unravel these intricate blueprints began with the pioneering work of scientists such as James Watson, Francis Crick, and Rosalind Franklin, who elucidated the double helix structure of DNA in the 1950s. This discovery laid the foundation for modern genomics, providing a molecular framework for understanding how genetic information is stored and transmitted. Genomics plays a crucial role in agriculture, specifically for crop improvement, livestock breeding, aquaculture, and conservation efforts. It identifies genes associated with desirable traits, enhances productivity, and provides insights into genetic diversity among species. By studying genomes, scientists can pinpoint populations at risk of extinction, assess the impact of habitat loss and climate change on genetic diversity, and develop strategies for conserving endangered species and restoring degraded ecosystems [1].

High-resolution genome analysis technologies and next-generation DNA sequencing are revolutionizing crop genetics. These technologies offer insights into crop genome structural diversity, making sequence variations more accessible and affordable. High-density genotyping arrays are crucial for genetic mapping, association studies, and genomic selection strategies. Next-generation genomics offers high-resolution analyses of crop diversity, identifying novel genetic diversities for improved yields and heterosis. This article discusses the key advancements in genetic engineering that find direct ways to crop yield improvement. It considers the potential of different Genomic technologies and their applications in staple crops.

2. Review of Related Research

The following summary's literature section covers earlier studies and their conclusions about the discovery of crop diversity, unravelling the genetic bases of heterosis, and future applications for breeding and crop improvement. A review has been done to boost production and create demand for improved technologies because of the crucial role it play in food and nutritional security [2].

2.1. Genotyping-by-Sequencing in Crop Plants for Discovery of DNA Sequence Diversity

Novel genomic technologies are essential for modern crop improvement, consistently enhancing efficiency and affordability [3]. The complete genome sequence of Arabidopsis thaliana, a model plant, was revealed fifteen years ago, encompassing over 125 mega-base pairs [4]. Rice sequencing, which has a larger size of 430 mega-base pairs, followed closely behind [5]. Analyzing various genome-wide markers can accurately predict a plant's performance [6]. Genomic selection is quickly gaining popularity as a strategy for breeding vital crops such as maize, wheat, rice, and barley [7-9].

Fast forward to our current era, where the meteoric rise of next-generation sequencing (NGS) platforms has unlocked access to reference sequences for several vital crop species, each with sprawling and complex genomes. We have maize weighing in at a staggering 2500 megabases [10], sorghum at 730 megabases [11], soybean at 1115 megabases [12], potato at 850 megabases [13], barley reaching 5100 megabases [14], and rapeseed even hitting 1200 megabases [15]. Moreover, if you can imagine the colossal hexaploid genome of bread wheat, which boasts an astounding 17000 megabases, it’s fascinating to note that innovative techniques like flow cytometry and synthetic mapping, along with NGS, have significantly paved the way for creating a chromosome-based draft genome sequence [16].

The field is buzzing with promise and possibility, as each revelation propels us further into the future of agriculture. The landscape of plant breeding is evolving rapidly, thanks to high-resolution genomic data that are becoming increasingly central to the field. This wealth of information helps breeders dive deep into the genetic makeup of germplasm, allowing them to pinpoint the genes associated with crucial agronomic traits and assess the breeding potential of individuals within various programs. This, in turn, accelerates the selection process for superior varieties [17]. The emergence of reduced-representation genotyping-by-sequencing (GBS) methods, whether or not a complete reference genome is in hand, has paved a promising pathway for high-throughput genomic re-sequencing. This marvel of modern technology makes it feasible to analyze vast plant populations at ever-declining costs [18].

Take maize, for example; [19] harnessed the GBS technique initially outlined by [20] to uncover nearly 700,000 genome-wide SNPs within a diverse panel of 2,815 inbred lines sourced from breeding programs worldwide. Such extensive sequencing datasets provide researchers with an extraordinary level of insight into genetic diversity and population structures. They reveal the intricate history of recombination and the variety of alleles present in different breeding pools [21, 22]. The versatility of GBS methods for genetic analysis has been demonstrated in numerous vital crops, such as rice [23], barley [20], and potatoes [24], illustrating their transformative impact on agricultural genetics.

2.2. Targeting of Genetic Variants Associated with Agronomic Traits

Targeted discovery of molecular variations in protein-coding genomic regions is enabled by reduced-representation sequencing techniques like transcriptome sequencing or exome capture [25]. Hundreds of thousands of molecular variations in gene-coding areas can be rapidly identified through reference-based assembly of target-captured or transcriptome sequence data in a test panel. The high-density genic polymorphism data produced by these methods can support high-resolution genetic resource search, marker-assisted breeding, genomic selection, and quantitative trait dissection [26].

Custom-designed capture probes or tailed sequencing primers offer flexibility in targeting specific chromosome regions that encompass quantitative trait loci, known pathways from related model plants, or candidate genes associated with traits of interest across a particular species. For example, [27] show how to use bead-based capture technology to re-sequence a panel of more than 30 genes involved in regulating the flowering time pathway, while [28] and [29] describe how tailed PCR can target specific genes for next-generation re-sequencing in large numbers of individuals.

Using a microarray-based capture device, [30] provides an example of re-sequencing genetic diversity that covers significant meta-QTL in a key crop. Mascher et al., [31] explain how exome capture sequencing can help in the cost-effective identification of coding sequence variants, even in very large genomes like that of barley. Additionally, [32] introduces the concept of associative transcriptomics, in which polymorphic SNP data from transcriptome sequencing in a diversity panel are associated with phenotypic variation for QTL identification.

Overall, these techniques showed strong potential for mapping and identifying the genes that underlie quantitative traits, as well as for creating closely related, sequence-based breeding markers. Furthermore, they demonstrate the wide variety of technical platforms for sequence capture that are available, allowing for highly flexible scaling of re-sequencing operations to manage anywhere from a few to many genes in large test populations at low cost.

2.3. Accessing Crop Diversity with High-Density Genotyping Arrays

In addition to their immediate uses in genetic mapping, QTL dissection, and diversity characterization, NGS technologies have laid the groundwork for the creation of high-density SNP genotyping platforms, which are high-throughput tools for genetic analysis of sizable experimental and breeding populations [33]. High-capacity SNP arrays are now available for a wide variety of plant species and are being used extensively in breeding major crops such as rapeseed (60k SNPs) [34], soya bean (50k SNPs) [35], rice (51.5k SNPs) [36, 37], maize (50–600k SNPs) [38, 39], or sorghum (3k and 90k SNPs) [40, 41].

At present, the Infinium platform from Illumina Inc. (San Diego, CA) and the Axiom technology from Affymetrix Inc. (Santa Clara, CA) are the most widely used platforms for large-scale SNP genotyping in crop plants [42]. Fixed genotyping chips are often preferred to GBS technologies for scenarios aiming to generate structured data sets of common sequence variants at low cost, with minimal bio-informatic input, for example, within an ongoing breeding program [17]. On the other hand, to be effective, a fixed SNP genotyping platform must apply to a wide range of different genotypes; hence, the alleles of the chosen SNPs must be representative even for diverse germplasm.

SNP markers on a fixed, high-density array are still likely to exhibit genetic associations with phenotypic variation through LD to the causal genes [43]. However, due to the relatively high extent of linkage disequilibrium (LD) throughout the genomes of most crops, GBS-based genotyping methods may be more appropriate for identifying true, causal genetic variants for phenotypes with a complex genetic architecture, as these are typically influenced in crop species by rare alleles that may not be adequately represented on an SNP array [44].

2.4. Genome-Scale Characterization of Crop Diversity

High-throughput genotyping techniques are a key enabling technology for complicated trait dissection by genome-wide association studies (GWAS). Detailed molecular characterization of breeding material, providing thorough information on population genetic characteristics and their links to natural and artificial selection for essential traits, is a crucial requirement for the creation of new, superior cultivars. Due to extensive human-mediated selection during plant breeding, especially in chromosome regions that include agronomically significant loci, modern crop varieties typically have biased allele frequency spectra and lower levels of genetic variation than their wild counterparts [45]. This leads to elevated levels of LD overall, chromosome-wise, and/or region-specific in the associated genomic locations. Furthermore, agricultural selection has produced larger haplotypes with extended homozygosity, which can sometimes cover large chromosome segments [45, 21, 22]. Due to directional selection's enrichment of specific allele variations in gene pools and the resulting loss of genetic variety, crop domestication resulted in genetic bottlenecks that led to noticeable selective sweeps in all major crops [46]. Linkage drag—the inadvertent co-selection of undesired gene variants that are strongly connected to specific loci of interest makes this especially problematic [47].

Because broad variation provides breeders with the foundation to choose varieties that continuously improve yield performance, and genetic diversity is the fundamental key to breeding success, these directional selection footprints pose a serious challenge to crop improvement by deteriorating genetically determined phenotypic variation. High-throughput population genomic studies can help address this dilemma by offering a comprehensive molecular basis for identifying and genomic introgression of new variation into chromosome regions surrounding directionally selected loci. They can identify and designate genome regions that require the most rejuvenation with novel diversity using high-resolution technologies, as well as employing genomic selection strategies [48] for enriching depleted gene pools.

2.5. Identifying and Overcoming Signatures of Selection

Jiao et al. [49] analyzed genome sequences from 278 maize inbred lines sourced from China and the US to illustrate the structural development of the maize genome throughout domestication and human utilization, highlighting significant genetic changes resulting from modern breeding. They identified approximately 400 loci exhibiting selective sweeps based on around 4.8 million SNPs present in at least 50% of the population, marking the progress of domestication and breeding. Their findings indicated that modern breeding has led to an increase in rare alleles among elite cultivars, suggesting that the proportion of rare alleles could serve as a selection index in future maize breeding efforts. Mace et al. [45] conducted a high-resolution study of genome-scale diversity and directional selection in sorghum using genome sequences from 44 highly diverse accessions that reflect the diversity of the primary gene pool. They discovered strong selection signatures surrounding key genes associated with domestication and agricultural traits such as plant height, seed color, and maturity across several gene pools. Their research uncovered untapped genetic variation within allopatric gene pools but noted reduced diversity in landraces and improved germplasm.

Huang et al. [50] have resequenced 446 distinct accessions and 1083 cultivated (O. indica and O. japonica) species of the wild relative of rice (Oryza rufipogon). In this investigation, 55 domestication-derived selective sweeps were identified. When comparing wild rice populations to farmed rice, nearly 8 million SNP markers showed exceptionally high levels of allelic variation. These findings suggest genomic target regions for future rice-breeding efforts focused on reintroducing genetic diversity. The efficiency of introgression techniques employing exotic plant resources in many crops has been hampered by the co-selection of undesirable loci caused by linkage drag, and the resolution with which introgressed DNA segments could be tracked on a molecular level was extremely low when concepts for marker-assisted back-crossing were first introduced to breeding.

Herzog et al. [51] demonstrated the intrinsic suitability of high-throughput genotyping arrays for the targeted introgression of donor chromosomal regions into recipient genotypes through a simulation study using maize data. Our capability to introduce novel loci into elite cultivars with minimal linkage drag is enhanced by employing exotic germplasm for the focused restoration of genetic diversity in crop breeding populations, which is facilitated by the detection of introgressed DNA fragments through comprehensive molecular marker information.

The comparative impact of breeding and artificial selection for key agronomic qualities on LD and haplotype structure can be revealed through large-scale molecular data from genome-wide SNP markers, even in complex polyploid genomes. For instance, divergent bread wheat gene pools exhibit stark variations in local LD around loci related to critical adaptation and grain quality parameters [22]. Qian et al. [21] demonstrated considerable sub-genomic bias for selection signals during breeding for essential seed quality variables in the extensively duplicated A and C subgenomes of rapeseed.

Significant genetic variability degradation in C-subgenome QTL was shown to reflect a much decreased diversity and a lower recombination rate, which hinders breeding progress and heterotic potential. According to [52], this data provides breeders with essential information to develop strategies for precisely restoring diversity in these regions, such as employing de novo polyploidization to increase recombination. Detailed comparison of local linkage disequilibrium (LD) decay on a 13-cM segment of Triticum aestivum chromosome 1B in a population of 460 international wheat accessions, comparing the local genetic diversity within five sub-populations representing distinct breeding pools. This example demonstrates how strong directional selection in distinct breeding pools can lead to highly distinct patterns of LD. Densely spaced SNP markers can assist back-crossing programs to enrich diversity-poor regions.

2.6. Enriching Sub-Genomic Diversity for Improvement of Heterotic Potential

A foundation for enhancing low-diversity chromosome regions is established by converting high-resolution genome data into structured breeding populations from sequenced founder lines, helping to address linkage drag associated with strong selection signatures [22]. Increasing diversity within hybrid breeding pools into chromosomal segments that exhibit high adaptation signatures can improve heterosis. Heterotic haplotype capture (HHC), as described by [52], utilizes whole-genome profiling to identify and enhance genomic areas with low diversity, integrating them into hybrid breeding programs to specifically bolster heterosis. The HHC method allows for the introduction and characterization of new diversity at a high-resolution, sub-chromosomal level, aiding in the restoration of eroded diversity in highly selected genomic regions. By generating large, structured breeding populations like nested association mapping panels, HHC employs varied, fully sequenced founder lines [53, 54]. Genotyping an entire population with a high-density SNP array facilitates the detection of crossover breakpoints in individuals at an unprecedented resolution and allows for the imputation of sequenced haplotypes across large populations using parental sequence data.

3. Unravelling the Genetic Basis of Heterosis

Alongside advanced molecular characterization and the use of crop breeding germplasm, innovative genotyping technologies are enhancing our understanding of the genetic basis of heterosis. Hybrid crops that leverage heterosis for improved yield and stability have significantly boosted agricultural production in recent decades. However, the molecular and genetic mechanisms behind heterosis remain not fully understood, despite its vital role in food security and the increasing use of hybrid vigor in inbred crops like bread wheat. Current genome profiling in extensive hybrid populations enables detailed analysis of the loci and genes related to heterotic traits. The dominance hypothesis suggests that the combination of dominant alleles from both parents results in heterosis [55]. According to this theory, an inbred line that aggregates beneficial loci from both parental lines should perform similarly to a hybrid; yet, creating such inbred lines is a challenging and lengthy process [56]. Recently, [57] conducted a study assessing the genetic factors influencing hybrid vigor across 38 agronomic traits, employing extensive genome-wide sequencing on a population of 1495 elite hybrid rice varieties and their inbred parental lines.

The genomic impacts on phenotypic traits from combining multiple loci lead to the accumulation of rare superior alleles with beneficial effects. In hybrid rice production, the heterosis effect mainly results from dominant complementation. It is anticipated that using forward and background selection alongside high-throughput genome screening technologies [58, 51] will significantly enhance breeding gains through effective hybrid vigor exploitation. The concept of genomic hybrid breeding was introduced in rice by [13], utilizing a genome-based prediction strategy informed by genome sequence data to estimate the F1 progeny's performance. This approach, based on over 250,000 SNP markers obtained from re-sequencing 210 parental inbred lines in a training set of 278 hybrids, demonstrated the capability of marker-directed estimation of F1 hybrid yields. The top 100 predicted hybrids out of 21,945 combinations were estimated to surpass the overall yield average by 16%, indicating considerable improvements over conventional breeding methods and accelerating hybrid rice production [13].

4. Further Applications for Breeding and Crop Improvement

High-throughput, high-density genome profiling tools enable rapid and cost-effective characterization of crop genomes with high precision. DNA sequence variants offer diverse practical applications, allowing breeders to perform targeted germplasm exchanges among gene pools, thereby enhancing diversity and heterosis. Crop wild relatives act as valuable genetic resources to restore variation lost during domestication and breeding [44, 59]. However, breeders often hesitate to utilize novel diversity from distant gene pools due to adaptation issues and potential performance drawbacks. High-density genome data can alleviate this by facilitating more efficient genomic selection strategies that accurately predict performance using genome-wide marker combinations [6]. Improving crop performance and breeding processes is closely tied to enhancing recombination in regions lacking diversity. Recent studies have demonstrated that high-resolution genotype data can assist in this enhancement, as shown by high-coverage sequencing of rice offspring from a biparental cross.

5. Conclusions and Outlook

The rising global population is expected to significantly increase the demand for key crops, while plant breeders face challenges in developing high-yield varieties due to limited agricultural land and the pressures of climate change. Genetic diversity is essential for effective breeding, and advanced genomics tools present new opportunities for accelerating crop improvement and enhancing genetic gain, especially for traits that are traditionally difficult to assess phenotypically. Techniques like GBS and array-based SNP genotyping enable efficient genome profiling of large plant populations at decreasing costs, providing breeders with detailed molecular insights. This supports the mapping of useful genes and QTL, particularly as structured populations like NAM or HHC panels are utilized. High-resolution analysis of breeding germplasm aids breeders in understanding genetic diversity at a sub-genomic level, which is crucial for enriching gene pools diminished by selective breeding practices. Chromosome segments associated with key traits often exhibit strong linkage disequilibrium (LD), potentially leading to yield reductions due to linkage drag. An important aspect of improving heterotic performance in major crops involves the strategic introgression of alleles from exotic germplasm into high-yield varieties.

Abbreviations

DNA	Deoxyribonucleic Acid
GWAS	Genome-wide Association Studies
GBS	Genotype by Sequencing
NAM	Nested Association Mapping
NGS	Next-Generation Sequencing
QTL	Quantitative Trait Locus
HHC	High-Throughput Hybrid Cross
SNP	Single-nucleotide Polymorphisms
LD	Linkage Disequilibrium

Author Contributions

Melkam Anteneh is the sole author. The author read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Wu, Bi, Wei Hu, and Yong Zhong Xing (2018). "The history and prospect of rice genetic breeding in China." Hered 40: 841-857.

[2] Abreham Reda (2015). Lentil (Lens culinaris Medikus) Current status and prospects of production in Ethiopia, Department of Wollo University, Ethiopia.

[3] Poland, J. A. and Rife, T. W. (2012). Genotyping-by-Sequencing for plant breeding and genetics. *Plant Genome*, 5, 92.

[4] The Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, 408, 796– 815.

[5] Goff S. A., Ricke D., Lan T. H., Presting G., Wang R., Dunn M., Glazebrook J., Sessions A., Oeller P., Varma H., (2002). A draft sequence of the rice genome (Oryza sativa L. ssp. japonica) Science; 296: 92–100. https://doi.org/10.1126/science.1068275

[6] Heffner, E. L., Sorrells, M. E. and Jannink, J. L. (2009). Genomic selection for crop improvement. *Crop Sci.* 49, 1–12.

[7] Lin, Z., Hayes, B. J., and Daetwyler, H. D. (2014). Genomic selection in crops, trees, and forages: a review. *Crop Pasture Sci.* 65, 1177.

[8] Poland, J., Endelman, J., Dawson, J., Rutkoski, J., Wu, S., Manes, Y., Dreisigacker, S., Crossa, J., Sánchez-Villeda, H., Sorrells, M. and Jannink, J. L. (2012). Genomic selection in wheat breeding using genotyping-by-sequencing. *Plant Genome*, 5, 103–113.

[9] Riedelsheimer, C., Czedik-Eysenberg, A., Grieder, C., Lisec, J., Technow, F., Sulpice, R., Altmann, T., Stitt, M., Willmitzer, L. and Melchinger, A. E. (2012). Genomic and metabolic prediction of complex heterotic traits in hybrid maize. *Nat. Genet.* 44, 217– 220.

[10] Schnable P. S., Ware D., Fulton R. S., Stein J. C., Wei F., Pasternak S., Liang C., Zhang J., Fulton L., Graves T. A., (2009). The B73 maize genome: complexity, diversity, and dynamics. Science.; 326: 1112–1115. https://doi.org/10.1126/science.1178534

[11] Paterson A. H., Bowers J. E., Bruggmann R., Dubchak I., Grimwood J., Gundlach H., Haberer G., Hellsten U., Mitros T., Poliakov A., (2009). The Sorghum bicolor genome and the diversification of grasses. Nature; 457: 551–556. https://doi.org/10.1038/nature07723

[12] Schmutz, J., Cannon, S., Schlueter, J., (2010). Genome sequence of the paleopolyploid soybean. *Nature* 463, 178–183. https://doi.org/10.1038/nature08670

[13] Xu, S., Zhu, D. and Zhang, Q. (2014). Predicting hybrid performance in rice using genomic best linear unbiased prediction. *Proc. Nat. Acad.. Sci. USA*, 111, 12456–12461.

[14] Mayer, K. F. X., Waugh, R., Brown, J. W. S., Schulman, A., Langridge, P., Platzer, M., Fincher, G. B., Muehlbauer, G. J., Sato, K., Close, T. J., Wise, R. P. and Stein, N. (2012). A physical, genetic, and functional sequence assembly of the barley genome. *Nature*, 491, 711– 716.

[15] Chalhoub B, Denoeud F, Liu S, (2014). Early allopolyploid evolution in the post-Neolithic Brassica napus oilseed genome. Science. 2014 Aug 22; 345(6199): 950-3. https://doi.org/10.1126/science.1253435

[16] International Wheat Genome Sequencing Consortium (2014). A chromosome-based draft sequence of the hexaploidy bread wheat (*Triticum aestivum)* genome. *Science*, 345, 1251788.

[17] Varshney, R. K., Terauchi, R. and McCouch, S. R. (2014). Harvesting the promising fruits of genomics: applying genome sequencing technologies to crop breeding. *PLoS Biol.* 12, e1001883.

[18] Deschamps, S., Llaca, V. and May, G. D. (2012). Genotyping-by-sequencing in plants. *Biology*, 1, 460– 483.

[19] Romay, M. C., Millard, M. J., Glaubitz, J. C., Peiffer, J. A., Swarts, K. L., Casstevens, T. M., Elshire, R. J., Acharya, C. B., Mitchell, S. E., Flint-Garcia, S. A., McMullen, M. D., Holland, J. B., Buckler, E. S. and Gardner, C. A. (2013). Comprehensive genotyping of the USA national maize inbred seed bank. *Genome Biol.* 14, R55.

[20] Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S. and Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high-diversity species. *PLoS ONE*, 6, e19379.

[21] Qian, L., Qian, W. and Snowdon, R. J. (2014). Sub-genomic selection patterns as a signature of breeding in the allopolyploid *Brassica napus* genome. *BMC Genom.* 15, 1170.

[22] Voss-Fels, K., Frisch, M., Qian, L., Kontowski, S., Friedt, W., Gottwald, S. and Snowdon, R. J. (2015). Subgenomic diversity patterns caused by directional selection in bread wheat gene pools. *Plant Genome*, 8, https://doi.org/10.3835/plantgenome2015.03.0013

[23] Spindel, J., Wright, M., Chen, C., Cobb, J., Gage, J., Harrington, S., Lorieux, M., Ahmadi, N., and McCouch, S. (2013). Bridging the genotyping gap: using genotyping by sequencing (GBS) to add high-density SNP markers and new value to traditional biparental mapping and breeding populations. *Theor. Appl. Genet.* 126, 2699– 2716.

[24] Uitdewilligen, J. G. A. M. L., Wolters, A. A., D'hoop, B. B., Borm, T. J. A. and Visser, R. G. F. (2013). A next-generation sequencing method for genotyping-by-sequencing of highly heterozygous autotetraploid potato. *PLoS ONE*, 8, e62355.

[25] Ku, C. -S., Wu, M., Cooper, D. N., Naidoo, N., Pawitan, Y., Pang, B., Iacopetta, B. and Soong, R. (2012). Exome versus transcriptome sequencing in identifying coding region variants. *Expert Rev. Mol. Diagn.* 12, 241– 251.

[26] Bolger, M. E., Weisshaar, B., Scholz, U., Stein, N., Usadel, B. and Mayer, Klaus. F. X. (2014). Plant genome sequencing - applications for crop improvement. *Curr. Opin. Biotech.* 26, 31– 37.

[27] Schiessl, S., Samans, B., Hüttel, B., Reinhard, R. and Snowdon, R. J. (2014). Capturing sequence variation among flowering-time regulatory gene homologues in the allopolyploid crop species *Brassica napus. Front. Plant Sci.* 5, 404.

[28] Gholami, M., Bekele, W. A., Schondelmaier, J., and Snowdon, R. J. (2012). A tailed PCR procedure for cost-effective, two-order multiplex sequencing of candidate genes in polyploid plants. *Plant Biotech. J.* 10, 635– 645.

[29] Rife, T. W., Wu, S., Bowden, R., and Poland, J. A. (2015). Spiked GBS: a unified, open platform for single-marker genotyping and whole-genome profiling. *BMC Genom.* 16, 169.

[30] Clarke, W. E., Federico, M. L., Gajardo, H. A., Gerhardt, D. J., Higgins, E., Sharpe, A. G., Snowdon, R. J., Parkin, I. A. P., and Iniguez-Luy, F.L. (2013). Genomic DNA enrichment using sequence capture microarrays: a novel approach to discover sequence nucleotide polymorphisms (SNP) in *Brassica napus* L. *PLoS ONE*, 8, e81992.

[31] Mascher, M., Richmond, T. A., Gerhardt, D. J., Himmelbach, A., Clissold, L., Sampath, D., Ayling, S., Steuernagel, B., Pfeifer, M., D'Ascenzo, M., Akhunov, E. D., Hedley, P. E., Gonzales, A. M., Morrell, P. L., Kilian, B., Blattner, F. R., Scholz, U., Mayer, K. F., Flavell, A. J., Muehlbauer, G. J., Waugh, R., Jeddeloh, J. A. und Stein, N. (2013). Barley whole exome capture: ein Werkzeug für genetics Forschung im Genus Hordeum und darüber hinaus. Plant J. 76, 494–505.

[32] Harper, A. L., Trick, M., Higgins, J., Fraser, F., Clissold, L., Wells, R., Hattori, C., Werner, P., and Bancroft, I. (2012). Associative transcriptomics of traits in the polyploid crop species *Brassica napus. Nat. Biotech.* 30, 798–802.

[33] Ganal, M. W., Polley, A., Graner, E. -M., Plieske, J., Wieseke, R., Luerssen, H. and Durstewitz, G. (2012). Large SNP arrays for genotyping in crop plants. *J. Biosci.* 37, 821– 828.

[34] Edwards, D., Batley, J. and Snowdon, R. J. (2013). Accessing complex crop genomes with next-generation sequencing. *Theor Appl Genet*. 126, 1– 11.

[35] Song, Q., Hyten, D. L., Jia, G., Quigley, C. V., Fickus, E. W., Nelson, R. L. and Cregan, P. B. (2013). Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. *PLoS ONE*, 8, e54985.

[36] Chen, H., Xie, W., He, H., Yu, H., Chen, W., Li, J., Yu, R., Yao, Y., Zhang, W., He, Y., Tang, X., Zhou, F., Deng, X. W. and Zhang, Q. (2014). A high-density SNP genotyping array for rice biology and molecular breeding. *Mol. Plant*, 7, 541– 553.

[37] Zhao, K., Tung, C. -W., Eizenga, G. C., Wright, M. H., Ali, M. L., Price, A. H., Norton, G. J., Islam, M. R., Reynolds, A., Mezey, J., McClung, A. M., Bustamante, C. D. and McCouch, S. R. (2011). Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa. Nat. Commun.* 2, 467.

[38] Ganal, M. W., Durstewitz, G., Polley, A., Bérard, A., Buckler, E. S., Charcosset, A., Clarke, J. D., Graner, E. -M., Hansen, M., Joets, J., Le Paslier, M. -C., McMullen, M. D., Montalent, P., Rose, M., Schön, C. -C., Sun, Q., Walter, H., Martin, O. C. and Falque, M. (2011). A large maize (*Zea mays* L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. *PLoS ONE*, 6, e28334.

[39] Unterseer, S., Bauer, E., Haberer, G., Seidel, M., Knaak, C., Ouzunova, M., Meitinger, T., Strom, T. M., Fries, R., Pausch, H., Bertani, C., Davassi, A., Mayer, K. F. and Schön, C. -C. (2014). A powerful tool for genome analysis in maize: development and evaluation of the high-density 600 k SNP genotyping array. *BMC Genom.* 15, 823.

[40] Bekele, W. A., Wieckhorst, S., Friedt, W. and Snowdon, R. J. (2013). High-throughput genomics in sorghum: from whole-genome resequencing to a SNP screening array. *Plant Biotech. J.* 11, 1112– 1125.

[41] Wieckhorst, S., Bekele, W., Kloiber-Maitz, M., Schulz-Streeck, T., Knaak, C., Ouzunova, M., Davassi, A. and Snowdon, R. (2015). A high-density SNP genotyping array for genome-based breeding of energy sorghum for central Europe. Plant and Animal Genome XXIII Conference, San Diego, CA, USA, January 10-14, 2015. https://pag.confex.com/pag/xxiii/webprogram/Paper15935.html

[42] Thomson, M. J. (2014). High-throughput SNP genotyping to accelerate crop improvement. *Plant Breed. Biotech.* 2, 195– 212.

[43] Wray, N. R., Yang, J., Hayes, B. J., Price, A. L., Goddard, M. E., and Visscher, P. M. (2013). Pitfalls of predicting complex traits from SNPs. *Nat. Rev. Genet.* 14, 507–515.

[44] Huang, X. and Han, B. (2014). Natural variations and genome-wide association studies in crop plants. *Annu. Rev.. Plant Biol.* 65, 531– 551.

[45] Mace, E. S., Tai, S., Gilding, E. K., Li, Y., Prentis, P. J., Bian, L., Campbell, B. C., Hu, W., Innes, D. J., Han, X., Cruickshank, A., Dai, C., Frère, C., Zhang, H., Hunt, C. H., Wang, X., Shatte, T., Wang, M., Su, Z., Li, J., Lin, X., Godwin, I. D., Jordan, D. R. and Wang, J. (2013). Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum. *Nat. Commun.* 4, 2320.

[46] Shi, J. and Lai, J. (2015). Patterns of genomic changes with crop domestication and breeding. *Curr. Opin. Plant Biol.* 24, 47– 53.

[47] Langridge, P. and Fleury, D. (2011). Making the most of 'omics' for crop breeding. *Trends Biotechnol.* 29, 33– 40.

[48] Jannink, J. L., Lorenz, A. J. and I wata, H. (2010). Genomic selection in plant breeding: from theory to practice. *Brief. Funct. Genomics*, 9, 166– 177.

[49] Jiao, Y., Zhao, H., Ren, L., Song, W., Zeng, B., Guo, J., Wang, B., Liu, Z., Chen, J., Li, W., Zhang, M., Xie, S. and Lai, J. (2012). Genome-wide genetic changes during modern breeding of maize. *Nat. Genet.* 44, 812– 815.

[50] Huang, X., Kurata, N., Wei, X., Wang, Z.-X., Wang, A., Zhao, Q., Zhao, Y., Liu, K., Lu, H., Li, W., Guo, Y., Lu, Y., Zhou, C., Fan, D., Weng, Q., Zhu, C., Huang, T., Zhang, L., Wang, Y., Feng, L., Furuumi, H., Kubo, T., Miyabayashi, T., Yuan, X., Xu, Q., Dong, G., Zhan, Q., Li, C., Fujiyama, A., Toyoda, A., Lu, T., Feng, Q., Qian, Q., Li, J. and Han, B. (2012). A map of rice genome variation reveals the origin of cultivated rice. *Nature*, 490, 497– 501.

[51] Herzog, E., Falke, K. C., Presterl, T., Scheuermann, D., Ouzunova, M. and Frisch, M. (2014). Selection strategies for the development of maize introgression populations. *PLoS ONE*, 9, e92429.

[52] Snowdon, R. J., Abbadi, A., Kox, T., Schmutzer, T. and Leckband, G. (2015). Heterotic haplotype capture: precision breeding for hybrid performance. *Trends Plant Sci.* 20, 410– 413.

[53] Gore, M. Q., Chia, J. M., Elshire, R. J., Sun, Q., Ersoz, E. S., Hurwitz, B. L., Peiffer, J. L., McMullen, M. D., Grills, G. S., Ross-Ibarra, J., Ware, D. H. and Buckler, E. S. (2009). A first-generation haplotype map of maize. *Science*, 326, 1115–1117.

[54] Yu, J., Holland, J. B., McMullen, M. D. and Buckler, E. S. (2008). Genetic design and statistical power of nested association mapping in maize. *Genetics*, 178, 539–551.

[55] Jones DF. Dominance of linked factors as a means of accounting for heterosis. Proc Natl Acad Sci U S A. 1917: 3(4): 310–312. https://doi.org/10.1073/pnas.3.4.310

[56] Schnable PS, Springer NM. Progress toward understanding heterosis in crop plants. Annu Rev Plant Biol. 2013: 64(1): 71–88.
https://doi.org/10.1146/annurev-arplant-042110-103827

[57] Huang, X., Yang, S., Gong, J., Zhao, Y., Feng, Q., Gong, H., Li, W., Zhan, Q., Cheng, B., Xia, J., Chen, N., Hao, Z., Liu, K., Zhu, C., Huang, T., Zhao, Q., Zhang, L., Fan, D., Zhou, C., Lu, Y., Weng, Q., Wang, Z.-X., Li, J. and Han, B. (2015). Genomic analysis of hybrid rice varieties reveals numerous superior alleles that contribute to heterosis. *Nat. Comm.* 6, 6258.

[58] Herzog, E. and Frisch, M. (2011). Selection strategies for marker-assisted back crossing with high-throughput marker systems. *Theor. Appl. Genet.* 123, 251– 260.

[59] Lopes, M. S., El-Basyoni, I., Baenziger, P. S., Singh, S., Royo, C., Ozbek, K., Aktas, H., Ozer, E., Ozdemir, F., Manickavelu, A., Ban, T. and Vikram, P. (2015). Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. *J. Exp. Bot.* 66, 3477–3486.
https://doi.org/10.1093/jxb/erv122