5Gs for crop genetic improvement

Rajeev K Varshney[1], Pallavi Sinha[1], Vikas K Singh[2],
Arvind Kumar[3], Qifa Zhang[4] and Jeffrey L Bennetzen[5]

Check for updates

Here we propose a 5G breeding approach for bringing much-needed disruptive changes to crop improvement. These 5Gs are Genome assembly, Germplasm characterization, Gene function identification, Genomic breeding (GB), and Gene editing (GE). In our view, it is important to have genome assemblies available for each crop and a deep collection of germplasm characterized at sequencing and agronomic levels for identification of marker-trait associations and superior haplotypes. Systems biology and sequencing-based mapping approaches can be used to identify genes involved in pathways leading to the expression of a trait, thereby providing diagnostic markers for target traits. These genes, markers, haplotypes, and genome-wide sequencing data may be utilized in GB and GE methodologies in combination with a rapid cycle breeding strategy.

Addresses
[1] Center of Excellence in Genomics & Systems Biology, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, 502324, India
[2] International Rice Research Institute, South Asia Hub, ICRISAT, Hyderabad, 502324, India
[3] IRRI South Asia Regional Center, NSRTC Campus, G.T. Road, Collectry Farm, P.O. Industrial Estate, Varanasi, 221006, India
[4] National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, 430070, China
[5] Department of Genetics, University of Georgia, Athens, GA 30602, USA

Corresponding author: Varshney, Rajeev K (r.k.varshney@cgiar.org)

Introduction

Dramatic and rapid climate change will cause extreme weather, including droughts, floods and other disasters. Food production will suffer greatly from these changes. The nearly 80 percent of the world's population that are poor and live in rural areas typically rely on local agriculture for their survival [1]. It has been predicted that, on average, global yields of major crops will be reduced 6.0% in wheat, 3.2% in rice, 7.4% in maize, and 3.1% in soybean for every degree Celsius increase in global mean temperature [2]. In this regard, the CGIAR system (https://www.cgiar.org/) initiated a 'Two Degree Initiative for Food and Agriculture'. This initiative is targeted on assisting ∼200 million small scale food producers across the globe to adapt at the speed and scale needed for the current pace of climate change. Improving access to climate-smart technologies and practices, including this development of climate-resilient high yielding varieties and their rapid availability to farmers' fields, will provide an opportunity to achieve climate smart solutions [3].

Crop improvement for food and nutritional security, especially in the context of continuous population growth and such challenges as climate change and water scarcity, have become important global concerns [4••]. Facing these threats, current crop breeding strategies will not yield a sufficient rate of crop improvement to meet demands in the short-term or long-term future. Hence, we propose a 5G breeding strategy to dramatically accelerate crop genetic improvement. The 1st G is *G*enome assembly for each crop species, the 2nd G is *G*ermplasm characterized at genomic and agronomic levels, the 3rd G is *G*ene function identification, the 4th G is *G*enomic breeding methodologies, and the 5th G is *G*ene editing technologies (Figure 1).

In the following sections, we describe the 5Gs for enhancing crop improvement. We conclude with a discussion of the current challenges and opportunities for integrating the 5Gs into crop improvement.

1st G: Genome assembly

Advances in next-generation sequencing (NGS) technologies coupled with improved genome assembly algorithms have facilitated the *de novo* assemblies of >264 plant genomes, including such crops as rice, maize, wheat, barley, soybean, cotton, sorghum, tomato, pigeonpea, chickpea, and groundnut. The quality of these genome assemblies varies tremendously, from nearly finished genomes to draft genomes with hundreds of unoriented sequence scaffolds. A few meet the platinum genome standard, including assemblies with full-chromosome scaffolds and haplotypes resolved across the entire genome, preferably including strong links to the genetic map. However, most plant genome assemblies are draft genomes. Recent advancements in sequencing technologies, particularly long read generation and physical map linkages, can now often generate chromosome-scale, fully

Figure 1
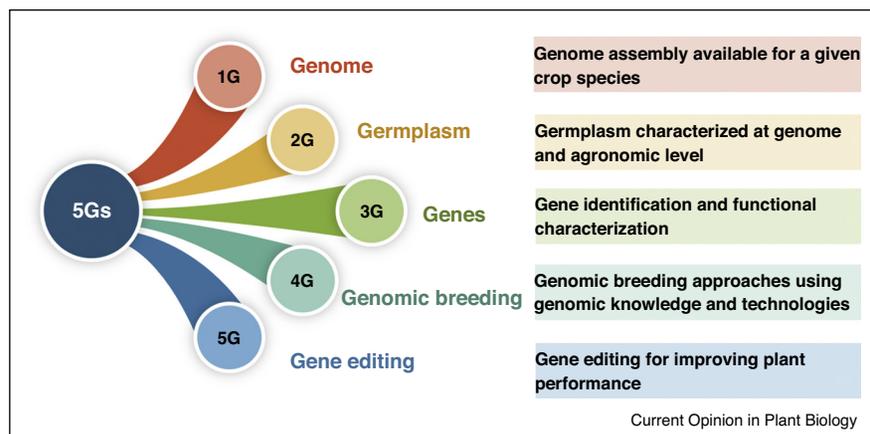
5Gs for crop genetic improvement.

phased diploid genome assemblies for any species at the platinum genome level [5].

The availability of a genome assembly provides an opportunity to develop genomics tools and technologies for such applications as trait discovery and molecular breeding. All genetic variation can be described, including SNPs, insertions, deletions, transversions, copy number variations and epigenetic changes [6]. These variants are useful in the development of customized SNP arrays [7], that can be utilized for development of saturated genetic maps and QTL identification. Sequence variant information defines haplotypes [8], which can then be employed for overcoming or taking advantage of linkage disequilibrium in a breeding program [9]. Genome assembly information is also vital for developing a gene expression atlas, proteome maps, metabolome maps, and epigenome maps.

With the ongoing and deep reductions in sequencing costs, large-scale re-sequencing projects have been initiated in several crops. For instance, 3010, 994 and 429 germplasm accessions have been re-sequenced in rice [10••], pearl millet [11••], and chickpea [12], respectively. Such projects generate 'big data' that pose storage and computational challenges. These challenges include compilation, curation, complex data analyses, visualization, retrieval and sharing [13]. To accelerate use of genome sequence information in next-generation breeding, customized informatics platforms are needed. In this context, some initiatives/platforms such as SNPSeek (for rice) (https://snp-seek.irri.org/_snp.zul), Genomic Open-source Breeding Informatics Initiative (GOBII) (http://cbsugobii05.tc.cornell.edu/wordpress/) and Excellence in Breeding Platform (EiB) (https://excellenceinbreeding.org/) have become available. These platforms will be vital to breeders for mining superior alleles/haplotypes, thus identifying the most-suitable parental lines for breeding populations.

2nd G: Germplasm characterization

During the course of crop domestication and breeding, cultivar genetic diversity is narrowed for all traits [14], but national and international 'genebanks' (germplasm repositories/germplasm banks) provide a rich source of diverse alleles that may be vital for future crop improvement. The ~1750 plant germplasm banks worldwide hold ~7.4 million accessions (www.fao.org), but <2% of these materials have been used as plant genetic resources (PGRs), although these few uses have led to major crop improvements [15••]. One of the reasons behind this limited use of PGRs is the overwhelming number of accessions that have no trait or other genetic information. Therefore, we propose characterization of as many accessions as possible at both genomic and agronomic levels. If the phenotyping is performed at specific nursery locations, and with community-established criteria, the provided information will allow deep genome wide association studies (GWAS) and identification of GXE effects. This provides the information to determine the potential agronomic value of particular alleles and accessions that will allow informed decision-making in breeding programs.

While NGS-based approaches have allowed comprehensive sequencing of large germplasm collections in several crops, field phenotyping lags dramatically. For instance, whole-genome re-sequencing (WGRS) has investigated 3010 rice accessions [10••], while genotyping-by-sequencing (GBS) has been utilized to characterize 44 624 wheat breeding lines [16•] and 20 000 wild and domesticated barley accessions [17]. These studies are initial examples of how genomics and informatics technologies can characterize large crop germplasm collections [18]. These studies are providing genome-wide

variant information and insights on population structure, crop domestication, and so on. However, for mining useful genetic information, it is imperative to phenotype the collections. NGS technologies together with some phenotyping have been utilized in a few crops for identification of marker-trait associations, including rice [19*], foxtail millet [20], pigeonpea [21*], pearl millet [11**], cotton [22], rapeseed [23], chickpea [12] and grape [24]. These studies have provided information on the genetic architecture of agriculturally important traits and the identification of valuable alleles for morphological, agronomic, developmental and quality-related traits. In the future, sequencing of entire germplasm collections present in genebanks and association with phenotypes should be a primary component for all crop-breeding programs.

Large-scale germplasm characterization also provides information on the presence of haplotypes at a particular locus for a given trait that can be used in haplotype-based breeding strategies ([25**], see later) or the genomic selection approach. Similarly, deleterious effect mutations (genetic load) can also be identified [26], and then can be purged by marker-assisted selection or gene editing, as suggested by Johnsson *et al.* [27*]. Eventually, superior parental lines will be identified with the best alleles at each locus, including minimum genetic load, and introduced into breeding programs with a plan to optimize the best allelic combinations. As an early step towards this optimal goal, current haplotype information can be used to select parents for nested association mapping (NAM) and multi-parent advanced generation inter-cross (MAGIC) populations for high-resolution gene:trait discovery.

Recent advances in genomics have led to the development of various sequencing-based rapid trait mapping approaches such as BSR-Seq [28], MutMap [29], QTLseq [30] and Indel-seq [31]. NGS technologies have enabled modification and improvement of traditionally tricky, time-consuming bulked segregant analysis (BSA, [32]) into rapid and whole-genome sequence-based high-resolution trait mapping [33*]. Due to the availability of genome assemblies, inexpensive high-throughput WGRS pipelines have become available, so that the use of sequence-based trait mapping approaches has become possible in several crop species. Following this approach, sequencing-based trait mapping can be broadly grouped into two classes: i) trait mapping through pooled sequencing of populations, and ii) trait mapping through complete sequencing of populations. Several examples of NGS-based trait mapping have been reported in crops [34]. This kind of trait mapping has several advantages over traditional marker-based mapping. For instance, in addition to taking much less time, these approaches identify genes or even quantitative trait nucleotides (QTN) for a given trait. In several cases, such QTNs have been converted into diagnostic markers. We believe

that genes and markers identified by using these approaches will have a uniquely high prediction/diagnostic power for breeding applications.

3rd G: Gene function identification

Using a range of functional genomics and trait mapping approaches, a large number of candidate genes with associated molecular markers for traits of interest have been identified in many crops. For instance, various -omics platforms were established in the past that have allowed the functional characterization of about 2296 genes controlling major traits in rice [35,36]. However, in most crops, the great majority of candidate genes, identified through transcriptomic approaches and/or mapping, are far from confirmation. Moreover, the molecular mechanisms of their potential agronomic values need to be understood in detail.

Systems Biology is an emerging holistic approach that proposes full understanding of biological systems by combining -omics approaches such as genomics, transcriptomics, epigenomics, proteomics, and metabolomics, together with modeling and high-performance computational analysis [37*]. In brief, systems biology is the study of an organism and/or trait, viewed as an integrated and interacting network of genes, proteins, and biochemical reactions, including the inputs from various internal and external environments. One goal of systems biology is to discover emergent properties derived from molecular interactions that will further our understanding of the entirety of processes that occur in a biological system. In furtherance of this goal, gene expression atlases [38–42], epigenome maps [43–45], proteome maps [46–48] and metabolome maps [49–51] have been developed in some crop species. Availability of these resources will accelerate the use of systems biology approaches to understand the molecular mechanism of complex traits such as drought tolerance [52] or heterosis [53]. Once traits are associated with particular pathways, and superior alleles identified, then breeders can employ a deeper understanding of plant biology to predict parental and allelic combinations that will uncover improved agronomic traits.

4th G: Genomic breeding (GB)

Genomic breeding involves approaches that use multi-omics data, knowledge resources, genes and technologies generated by genomics research for breeding the genomes to enhance crop breeding programmes. [35]. Although some methods of GB such as marker-assisted selection (MAS), marker-assisted backcrossing (MABC) and marker-assisted recurrent selection (MARS) have been used for breeding in several crops, it is important to have GB methodologies well-integrated into most or all crop breeding programs. In addition to above-mentioned GB methodologies, some new approaches such as forward breeding (FB), haplotype-based breeding (HBB) and genomic selection (GS), coupled with speed breeding

(SB), have also been suggested for enhancing the precision, efficiency and rate of acquired genetic gain in crop breeding [34]. While diagnostic markers associated with genes and major effect QTL are required for MAS, MABC and FB, superior haplotypes at a given locus for a target trait need to be identified for HBB. The GS approach, in contrast, does not need markers specifically associated with a trait because breeding lines are selected for crossing and advancing generations based on genomic-estimated breeding values calculated from genome-wide marker data.

Considering the breeding objectives, any of above-mentioned GB approaches can be chosen for crop improvement. For example, if breeders need to select parental lines or introgress some major effect QTL for a target trait, MAS and MABC approaches can be used. MABC is useful to introgress a few loci (<10) for improving elite varieties. This approach has been extensively used to develop a large number of breeding lines for commercial release in public and private sectors. The FB approach will be the best option when early generations of segregating populations (e.g. F_2 generations) are used to advance plants carrying the target QTL/gene. The MARS approach is useful to introgress from 10 to 40 loci through intercrossing elite × elite parents to develop superior lines with an optimum combination of superior alleles [34].

Recent re-sequencing of germplasm collections in a few crops has facilitated identification of a small number of strong marker-trait associations and haplotypes for target traits [54,55]. 'Haplotype assembly' was proposed as one new approach for developing improved crops through assembling superior haplotypes of the targeted traits [25••]. 'Superior haplotypes', in which the phenotypic performance of the group of individuals sharing a haplotype ('specific haplotype group'), can be identified. The identified superior haplotypes then can be utilized in the breeding program through haplotype-assisted breeding.

GS is an approach using genome-wide selection with a large number of markers [56]. GS works upon defined 'genomic estimated breeding values' (GEBVs) that are calculated from the genotypic and phenotypic dataset of a 'training population.' This approach has a higher accuracy of prediction of elite genetic materials in the initial generations and permits shorter breeding cycles. GS, reviewed by Crossa *et al.* [57•], has been extensively used in several crops. Very recently, Watson *et al.* [58••] introduced the concept of 'speed breeding' by giving plants light for 22 hours and dark for only 2 hours. Speed breeding shortens generation times, and thus has been proposed or is now being used for many crops [59]. In fact, speed breeding has also been suggested to be coupled with GS in a process called SpeedGS, for rapid development of new breeding lines [1]. GS combined with

superior haplotypes (Haplo-GS) is another new and promising approach for the rapid development of new breeding lines.

5th G: Gene editing (GE)

GE has emerged as a powerful approach for improving plant performance and the development of various abiotic and biotic stress tolerance lines. With the recent discovery of *Cas9* guide RNA and availability of functional genomics data coupled with advances in bioinformatics pipelines, targets are being identified and subjected to editing. A large number of genes with significant phenotypic effects have been cloned and functionally characterized. As a result, GE has been used to generate useful traits in such crops as rice, maize, wheat, sugarcane, soybean, potato, sorghum, orange, cucumber, tomato, flax, and cassava, for traits like herbicide resistance, drought tolerance, thermo-sensitive genic male sterility, disease resistance and altered product quality, including some in the process of commercial release [60]. For instance, Oliva *et al.* [61] edited promoters of *SWEET11*, *SWEET13* and *SWEET14* at effector-binding elements recognized by the pathogen *Xanthomonas oryza* pv *oryzae*, a causal agent for rice bacterial blight. These experiments generated rice plants that are broadly resistant to the pathogen. To enhance the durability and management of resistance, Eom *et al.* [62] developed a kit to trace the disease, its virulence and resistance alleles. However, the stewardship of gene-edited lines in combination with an appropriate deployment strategy is essential to meet environmental health and safety standards. There remains a lack of clarity as to the GMO or non-GMO status of such germplasm in many countries [63]. It is anticipated that legislation and a better-educated public will soon allow the benefits of this research to reach the farming community [64].

It is also important to mention that the GE approach is not only useful to create novel alleles, it can also be used for the promotion of superior alleles [65] and removal of deleterious effect alleles [27•] identified through large-scale sequencing efforts. Furthermore, it has been suggested that a reverse domestication approach could be pursued for new crops or current crops by editing genes related to domestication traits in wild species. This could provide crop diversification and make available superior lines with enhanced stress resistances. As this approach may require several cycles of editing and line fixation, 'ExpressEdit' approaches that combine speed breeding with GE have been suggested [1].

Conclusions and prospects

Although components of the described 5Gs are being used in public and private crop improvement programs in several developed countries, comprehensive 5G integration is lacking, especially in developing countries. However, we are hopeful that recent advances in sequencing,

phenotyping and data science will accelerate utilization of the 5G strategy in coordinated crop improvement programs worldwide. In this context, capacity building of young scientists in developing countries is required in 5G breeding to handle, analyze and interpret the enormous data sets from sequencing, genotyping, phenotyping, -omics and systems biology studies pursued across large-scale germplasm collections. In particular, training on breeder-friendly pipelines, analytical and decision support tools and databases related to identification of variants and haplotype, diversity analysis, sequencing-based trait mapping, identification of GE targets and implementation of GB methodologies will be very helpful. In summary, a comprehensively applied 5G breeding can enhance the precision, efficiency and effectiveness of breeding programs to develop climate-resilient, high-yielding and nutritious varieties while delivering a high rate of genetic gain in any breeding program, including in developing countries where these gains are most needed.

Conflict of interest statement

Nothing declared.

Acknowledgements

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- •• of outstanding interest

1. FAO: *Agriculture and Climate Change – Challenges and Opportunities at the Global and Local Level – Collaboration on Climate-Smart Agriculture*. Food and Agriculture Organization (FAO) of the United Nations; 2019. Rome 52 pp.

2. Zhao C, Liu B, Piao S, Wang X, Lobell DB, Huang Y, Huang M, Yao Y, Bassu S, Ciais P et al.: **Temperature increase reduces global yields of major crops in four independent estimates**. *Proc Natl Acad Sci U S A* 2017, **114**:9326-9331.

3. Atlin GN, Cairns JE, Das B: **Rapid breeding and varietal replacement are critical to adaptation of cropping systems in the developing world to climate change**. *Glob Food Secur* 2017, **12**:31-37.

4. Hickey LT, Hafeez Amber N, Robinson H, Jackson SA, Leal-
•• Bertioli SCM, Tester M, Gao C, Godwin ID, Hayes BJ, Wulff BBH: **Breeding crops to feed 10 billion**. *Nat Biotechnol* 2019, **37**:744-754.
The paper discusses how various tools and technologies, including speed breeding and express gene editing, can be utilized to dramatically enhance crop improvement.

5. Zhang X, Zhang S, Zhao Q, Ming R, Tang H: **Assembly of allele-aware, chromosomal-scale autopolyploid genomes based on Hi-C data**. *Nat Plants* 2019, **5**:833-845.

6. Johannes F, Schmitz RJ: **Spontaneous epimutations in plants**. *New Phytol* 2019, **221**:1253-1259.

7. Rasheed A, Hao Y, Xia X, Khan A, Xu Y, Varshney RK, He Z: **Crop breeding chips and genotyping platforms: progress, challenges, and perspectives**. *Mol Plant* 2017, **10**:1047-1064.

8. Kim SA, Cho CS, Kim SR, Bull SB, Yoo YJ: **A new haplotype block detection method for dense genome sequencing data based on interval graph modeling of clusters of highly correlated SNPs**. *Bioinformatics* 2018, **34**:388-397.

9. Voss-Fels KP, Stahl A, Wittkop B, Lichthardt C, Nagler S, Rose T, Chen TW, Zetzsche H, Seddig S, Baig MM et al.: **Breeding improves wheat productivity under contrasting agrochemical input levels**. *Nat Plants* 2019, **17**:1.

10. Wang W, Mauleon R, Hu Z, Chebotarov D, Tai S, Wu Z, Li M,
•• Zheng T, Fuentes RR, Zhang F et al.: **Genomic variation in 3,010 diverse accessions of Asian cultivated rice**. *Nature* 2018, **557**:43-01049.
An important initiative in rice whole-genome re-sequencing, with 3010 accessions originating from 86 countries. This study opens the path for big data genomics in crop plants.

11. Varshney RK, Shi C, Thudi M, Mariac C, Wallace J, Qi P, Zhang H,
•• Zhao Y, Wang X, Rathore A et al.: **Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments**. *Nat Biotechnol* 2017, **35**:969-976.
This paper presents the vision of using sequencing-based technologies for germplasm identification, trait mapping and breeding. This is the first genome sequence paper for pearl millet and includes re-sequencing of 994 pearl millet accessions, enabling insights into population structure, genetic diversity, and domestication.

12. Varshney RK, Thudi M, Roorkiwal M, He W, Upadhyaya HD, Yang W, Bajaj P, Cubry P, Rathore A, Jian J: **Resequencing of 429 chickpea accessions from 45 countries provides insights into genome diversity, domestication and agronomic traits**. *Nat Genet* 2019, **51**:857-864.

13. Stephens ZD, Lee SY, Faghri F, Campbell RH, Zhai C, Efron MJ et al.: **Big data: astronomical or genomical?** *PLoS Biol* 2015, **13**: e1002195 http://dx.doi.org/10.1371/journal.pbio.1002195.

14. Smýkal P, Nelson MN, Berger JD, von Wettberg EJB: **The impact of genetic changes during crop domestication**. *Agronomy* 2018, **8**:119 http://dx.doi.org/10.3390/agronomy8070119.

15. Janzen GM, Wang L, Hufford MB: **The extent of adaptive wild
•• introgression in crops**. *New Phytol* 2019, **221**:1279-1288.
This is a comprehensive review paper about the use of wild relatives in crop breeding.

16. Juliana P, Poland J, Huerta-Espino J, Shrestha S, Crossa J,
• Crespo-Herrera L, Toledo FH, Govindan V, Mondal S, Kumar U et al.: **Improving grain yield, stress resilience and quality of bread wheat using large-scale genomics**. *Nat Genet* 2019, **51**:1530-1539.
This presents the vision of how large-scale sequence-based genotyping (>40 000 individuals) could be useful for trait mapping.

17. Milner SG, Jost M, Taketa S, Mazón ER, Himmelbach A, Oppermann M, Weise S, Knüpffer H, Basterrechea M et al.: **Genebank genomics highlights the diversity of a global barley collection**. *Nat Genet* 2019, **51**:319-326.

18. Langridge P, Waugh R: **Harnessing the potential of germplasm collections**. *Nat Genet* 2019, **51**:200-201.

19. Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y, Li C, Zhu C, Lu T,
• Zhang Z et al.: **Genome-wide association studies of 14 agronomic traits in rice landraces**. *Nat Genet* 2010, **42**:961-967.
This study demonstrates the first use of whole-genome resequencing for constructing the comprehensive haplotype map for GWAS of complex traits in rice. It is also the first application of GWAS in crops.

20. Jia G, Huang X, Zhi H, Zhao Y, Zhao Q, Li W, Chai Y, Yang L, Liu K, Lu H, Zhu C: **A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*)**. *Nat Genet* 2013, **45**:957.

21. Varshney RK, Saxena RK, Upadhyaya HD, Khan AW, Yu Y, Kim C,
• Rathore A, Kim D, Kim J, An S et al.: **Whole-genome resequencing of 292 pigeonpea accessions identifies genomic regions associated with domestication and agronomic traits**. *Nat Genet* 2017, **49**:1082-1088.

This is an early paper demonstrating the use of NGS-based technologies for assembling genome in legume crops and identified domestication-related genes and marker-trait associations.

22. Ma Z, He S, Wang X, Sun J, Zhang Y, Zhang G, Wu L, Li Z, Liu Z, Sun G *et al.*: **Resequencing a core collection of upland cotton identifies genomic variation and loci influencing fiber quality and yield**. *Nat Genet* 2018, **50**:803-813.

23. Lu K, Wei L, Li X, Wang Y, Wu J, Liu M, Zhang C, Chen Z, Xiao Z, Jian H *et al.*: **Whole-genome resequencing reveals *Brassica napus* origin and genetic loci involved in its improvement**. *Nat Commun* 2019, **10**:1154 http://dx.doi.org/10.1038/s41467-019-09134-9.

24. Liang Z, Duan S, Sheng J, Zhu S, Ni X, Shao J, Liu C, Nick P, Du F, Fan P *et al.*: **Whole-genome resequencing of 472 *Vitis* accessions for grapevine diversity and demographic history analyses**. *Nat Commun* 2019, **10**:1190 http://dx.doi.org/10.1038/s41467-019-09135-8.

25. Bevan MW, Uauy C, Wulff BBH, Zhou J, Krasileva K, Clark MD:
•• **Genomic innovation for crop improvement**. *Nature* 2017, **543**:346-354.
This pioneering paper discussed how genomics can provide new foundations for crop-breeding systems and proposed the haplotype-assembly concept.

26. Ramu P, Esuma W, Kawuki R, Rabbi IY, Egesi C, Bredeson JV, Bart RS, Verma J, Buckler ES, Lu F: **Cassava haplotype map highlights fixation of deleterious mutations during clonal propagation**. *Nat Genet* 2017, **49**:959-963.

27. Johnsson M, Gaynor RC, Jenko J, Gorjanc G, de Koning DJ,
• Hickey JM: **Removal of alleles by genome editing (RAGE) against deleterious load**. *Genet Sel Evol* 2019, **51**:14 http://dx.doi.org/10.1186/s12711-019-0456-8.
This paper proposes that deleterious allele loads can be eliminated by genome editing.

28. Liu S, Yeh CT, Tang HM, Nettleton D, Schnable PS: **Gene mapping via bulked segregant RNA-Seq (BSR-Seq)**. *PLoS One* 2012, **7**:e36406.

29. Abe A, Kosugi S, Yoshida K, Natsume S, Takagi H, Kanzaki H, Matsumura H, Yoshida K, Mitsuoka C, Tamiru M *et al.*: **Genome sequencing reveals agronomically important loci in rice using MutMap**. *Nat Biotechnol* 2012, **30**:174-178.

30. Takagi H, Abe A, Yoshida K, Kosugi S, Natsume S, Mitsuoka C, Uemura A, Utsushi H, Tamiru M, Takuno S *et al.*: **QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations**. *Plant J* 2013, **74**:174-183.

31. Singh VK, Khan AW, Saxena RK, Sinha P, Kale SM, Parupalli S, Kumar V, Chitikineni A, Vechalapu S, Sameer Kumar CV *et al.*: **Indel-seq: a fast-forward genetics approach for identification of trait-associated putative candidate genomic regions and its application in pigeonpea (*Cajanus cajan*)**. *Plant Biotechnol J* 2017, **15**:906-914.

32. Michelmore RW, Paran I, Kesseli RV: **Identification of markers linked to disease resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations**. *Proc Natl Acad Sci U S A* 1991, **88**:9828-9832.

33. Schlotterer C, Tobler R, Kofler R, Nolte V: **Sequencing pools of**
• **individuals - mining genome-wide polymorphism data without big funding**. *Nat Rev Genet* 2014, **15**:749-763.
Pioneering paper describes how a pool-based sequencing approach could be useful in rapid marker identification with less input cost.

34. Varshney RK, Pandey MK, Bohra A, Singh VK, Thudi M, Saxena RK: **Toward the sequence-based breeding in legumes in the post-genome sequencing era**. *Theor Appl Genet* 2019, **132**:797-816.

35. Wing RA, Purugganan MD, Zhang Q: **The rice genome revolution: from an ancient grain to green super rice**. *Nat Rev Genet* 2018, **19**:505 517.

36. Li Y, Xiao J, Chen L, Huang X, Cheng Z, Han B, Zhang Q, Wu C: **Rice functional genomics research: past decade and future**. *Mol Plant* 2018, **11**:359-380.

37. Lavarenne J, Guyomarc'h S, Sallaud C, Gantet P, Lucas M: **The**
• **spring of systems biology-driven breeding**. *Trends Plant Sci* 2018, **23**:706-720.
This opinion paper connects systems biology research with breeding programs.

38. Nobuta K, Venu RC, Lu C, Beló A, Vemaraju K, Kulkarni K, Wang W, Pillay M, Green PJ, Wang GL, Meyers BC: **An expression atlas of rice mRNAs and small RNAs**. *Nat Biotechnol* 2007, **25**:473-477.

39. Pazhamala LT, Purohit S, Saxena RK, Garg V, Krishnamurthy L, Verdier J, Varshney RK: **Gene expression atlas of pigeonpea and its application to gain insights into genes associated with pollen fertility implicated in seed formation**. *J Exp Bot* 2017, **68**:2037-2054.

40. Kudapa H, Garg V, Chitikineni A, Varshney RK: **The RNA-Seq-based high resolution gene expression atlas of chickpea (*Cicer arietinum* L.) reveals dynamic spatio-temporal changes associated with growth and development**. *Plant Cell Environ* 2018, **41**:2209-2225.

41. Shinozaki Y, Nicolas P, Fernandez-Pozo N, Ma Q, Evanich DJ, Shi Y, Xu Y, Zheng Y, Snyder SI, Martin Laetitia BB *et al.*: **High-resolution spatiotemporal transcriptome mapping of tomato fruit development and ripening**. *Nat Commun* 2018, **9**:364.

42. Hoopes GM, Hamilton JP, Wood JC, Esteban E, Pasha A, Vaillancourt B, Provart NJ, Buell CR: **An updated gene atlas for maize reveals organ-specific and stress-induced genes**. *Plant J* 2019, **97**:1154-1167.

43. Junaid A, Kumar H, Rao AR, Patil AN, Singh NK, Gaikwad K: **Unravelling the epigenomic interactions between parental inbreds resulting in an altered hybrid methylome in pigeonpea**. *DNA Res* 2018, **25**:361-373.

44. Li Z, Wang M, Lin K, Xie Y, Guo J, Ye L, Zhuang Y, Teng W, Ran X, Tong Y *et al.*: **The bread wheat epigenomic map reveals distinct chromatin architectural and evolutionary features of functional genetic elements**. *Genome Biol* 2019, **20**:139.

45. Peng Y, Xiong D, Zhao L, Ouyang W, Wang S, Sun J, Zhang Q, Guan P, Xie L, Li W, Li G, Yan J, Li X: **Chromatin interaction maps reveal genetic regulation for quantitative traits in maize**. *Nat Commun* 2019, **10**:2632.

46. Duncan O, Trösch J, Fenske R, Taylor NL, Millar AH: **Resource: mapping the *Triticum aestivum* proteome**. *Plant J* 2017, **89**:601-616.

47. Barua P, Lande NV, Subba P, Gayen D, Pinto S, Keshava Prasad TS, Chakraborty S, Chakraborty N: **Dehydration-responsive nuclear proteome landscape of chickpea (*Cicer arietinum* L.) reveals phosphorylation-mediated regulation of stress response**. *Plant Cell Environ* 2019, **42**:230-244.

48. Jiang LG, Li B, Liu SX, Wang HW, Li CP, Song SH, Beatty M, Zastrow-Hayes G, Yang XH, Qin F *et al.*: **Characterization of proteome variation during modern maize breeding**. *Mol Cell Proteomics* 2019, **18**:263-276.

49. Okazaki Y, Saito K: **Integrated metabolomics and phytochemical genomics approaches for studies on rice**. *Gigascience* 2016, **5**:11.

50. Chen J, Wang J, Chen W, Sun W, Peng M, Yuan Z, Shen S, Xie K, Jin C, Sun Y *et al.*: **Metabolome analysis of multi-connected biparental chromosome segment substitution line populations**. *Plant Physiol* 2018, **178**:612-625.

51. Zhou S, Kremling KA, Bandillo N, Richter A, Zhang YK, Ahern KR, Artyukhin AB, Hui JX, Younkin GC, Schroeder FC *et al.*: **Metabolome-scale genome-wide association studies reveal chemical diversity and genetic control of maize specialized metabolites**. *Plant Cell* 2019, **31**:937-955.

52. Miao Z, Han Z, Zhang T, Chen S, Ma C: **A systems approach to a spatio-temporal understanding of the drought stress response in maize**. *Sci Rep* 2017, **7**:6590.

53. Li D, Huang Z, Song S, Xin Y, Mao D, Lv Q, Zhou M, Tian D, Tang M, Wu Q *et al.*: **Integrated analysis of phenome, genome, and transcriptome of hybrid rice uncovered multiple heterosis-related loci for yield increase**. *Proc Natl Acad Sci U S A* 2016, **113**:E6026-E6035.

54. Abbai R, Singh VK, Nachimuthu VV, Sinha P, Selvaraj R, Vipparla AK, Singh AK, Singh UM, Varshney RK, Kumar A: **Haplotype analysis of key genes governing grain yield and quality traits across 3K RG panel reveals scope for the development of tailor-made rice with enhanced genetic gains**. *Plant Biotechnol J* 2019, **17**:1612-1622.

55. Kuroha T, Nagai K, Gamuyao R, Wang DR, Furuta T, Nakamori M, Kitaoka T, Adachi K, Minami A *et al.*: **Ethylene-gibberellin signaling underlies adaptation of rice to periodic flooding**. *Science* 2018, **361**:181-186.

56. Meuwissen THE, Hayes BJ, Goddard ME: **Prediction of total genetic value using genome-wide dense marker maps**. *Genetics* 2001, **157**:1819-1829.

57. Crossa J, Pérez-Rodríguez P, Cuevas J, Montesinos-López O,
• Jarquín D, de Los Campos G, Burgueño J, GonzálezCamacho JM, Pérez-Elizalde S, Beyene Y *et al.*: **Genomic selection in plant breeding: methods, models, and perspectives**. *Trends Plant Sci* 2017, **22**:961-975.
This is an outstanding paper that discusses how genomic selection can be utilized in plant breeding.

58. Watson A, Ghosh S, Williams MJ, Cuddy WS, Simmonds J,
•• Rey MD, Hatta MA, Hinchliffe A, Steed A, Reynolds D *et al.*: **Speed breeding is a powerful tool to accelerate crop research and breeding**. *Nat Plants* 2018, **4**:23-29.
This is the game-changer manuscript for how rapid cyclic breeding could be useful for rapid delivery of improved varieties.

59. Ghosh S, Watson A, Gonzalez-Navarro OE, Ramirez-Gonzalez RH, Yanes L, Mendoza-Suárez M, Simmonds J, Wells R, Rayner T,
Green P: **Speed breeding in growth chambers and glasshouses for crop breeding and model plant research**. *Nat Protoc* 2018, **12**:2944-2963.

60. Zhang Y, Massel K, Godwin ID, Gao C: **Applications and potential of genome editing in crop improvement**. *Genome Biol* 2018, **1**:13 http://dx.doi.org/10.1186/s13059-019-1622-6.

61. Oliva R, Ji C, Atienza-Grande G, Huguet JC, Perez-Quintero A, Li T, Eom JS, Li C, Nguyen H, Liu B *et al.*: **Broad-spectrum resistance to bacterial blight in rice using genome-editing**. *Nat Biotechnol* 2019, **37**:1344-1350 http://dx.doi.org/10.1038/s41587-019-0267-z.

62. Eom JS, Luo D, Atienza-Grande G, Yang J, Ji C, Luu VT, Huguet-Tapia JC, Char SN, Liu B, Nguyen H *et al.*: **Diagnostic kit for rice blight resistance**. *Nat Biotechnol* 2019, **37**:1372-1379 http://dx.doi.org/10.1038/s41587-019-0268-y.

63. Schulman AH, Oksman-Caldentey KM, Teeri TH: **European court of justice delivers no justice to Europe on genome-edited crops**. *Plant Biotechnol J* 2020, **18**:8-10 http://dx.doi.org/10.1111/pbi.13200.

64. Varshney RK, Godwin I, Mohapatra T, Jones JDG, McCouch SR: **A SWEET solution to rice blight**. *Nat Biotechnol* 2019, **37**:1280-1282 http://dx.doi.org/10.1038/s41587-019-0302-0.

65. Jenko J, Gorjanc G, Cleveland MA, Varshney RK, Whitelaw CB, Woolliams JA, Hickey JM: **Potential of promotion of alleles by genome editing to improve quantitative traits in livestock breeding programs**. *Genet Sel Evol* 2015, **2**:47-55.