Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Plant Biology

New genomic approaches for enhancing maize genetic improvement

Ning Yang and Jianbing Yan

Maize (*Zea mays*) is one of the most widely grown crops in the world, with an annual global production of over 1147 million tons. Genomics approaches are thought to be the best solution for accelerating yield improvement to meet the challenges of a growing population and global climate change. Here, we review current approaches to the exploration of novel genetic variation in genomes, DNA modifications, and transcription levels of cultivated maize, landraces, and wild relatives. We discuss applications of genetic engineering to maize yield improvement and highlight future directions for maize genomics studies.

Address
National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, China

Corresponding authors:
Yang, Ning (ningy@mail.hzau.edu.cn), Yan, Jianbing (yjianbing@mail.hzau.edu.cn)

Current Opinion in Plant Biology 2020, **60**:101977

This review comes from a themed issue on **Plant biotechnology**

Edited by **Yiping Qi** and **Jing-Ke Weng**

https://doi.org/10.1016/j.pbi.2020.11.002

Introduction

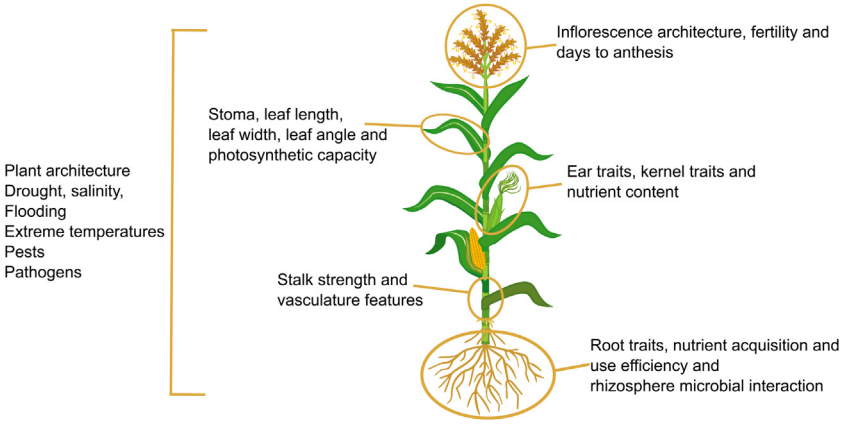
Improving grain yield has always been a primary target of maize breeding and has become more challenging against the backdrop of climate stress and increasing global populations. In recent decades, improvements in maize yield have relied on high rates of fertilizer and pesticide application and on breeding for yield and fitness in particular climatic regions. The former approach had many adverse effects on sustainable development [1], and the latter produced maize varieties that were unable to meet the challenge of a highly variable climate. High maize yields are determined by numerous traits, including the capacity to tolerate biotic and abiotic stress (Figure 1). Balancing the trade-offs among these traits is key to obtaining high and stable maize yields and requires a clear mechanistic understanding of genes, pathways, natural variation, and local adaptations. Genomics-based breeding is thought to be the best solution for overcoming these challenges and ensuring sustainable increases in maize yield [2].

Genetic variations that can be acted upon by human or natural selection hold the key to maize yield improvement, and the assembly of a reference genome can facilitate mining genetic variations in maize. The first maize reference genome (B73 RefGen_v1) based on BAC library and Sanger sequencing technology was published in 2009 [3]. This initial reference genome promoted substantial progress in maize genomics by accelerating genome-wide SNPs (single nucleotide polymorphisms) surveys and subsequent marker–trait association studies (QTL/GWAS) to link genetic and phenotypic variations. As a result, our understanding of the genetic architecture of yield-related traits [4–6], abiotic stress-related traits [7], stalk strength [8•], plant architecture [9,10], disease resistance [11,12], and flowering time [13] was greatly extended. Although this initial reference genome enabled rapid progress in maize genomics, it contained many gaps and more than 100 000 small contigs due to the large amount of repetitive sequences [3], many of which were incorrectly ordered and oriented, complicating the detailed analysis of specific loci crucial to our understanding of phenotypic variations [14]. However, the development of single-molecule sequencing technologies now enables the production of high-quality reference genomes, multiple reference genomes, and pangenome graphs. In addition to identifying genetic variations, understanding the relationships among gene expression, epigenetic modifications, chromatin interactions, and metabolic, proteome, and phenotype variations can further enrich our knowledge. With the development of high-quality genome assemblies, the accurate characterization of genomic diversity at multiple levels, and the precise association of genetic variants with yield-related traits, we can now accelerate maize yield improvement using genome editing guided by mechanistic understanding.

High-throughput short-read sequencing permits efficient mining of genetic variation

As the cost of short-read sequencing decreased, maize HapMaps were developed with larger sample sizes and higher densities of SNPs and indels. HapMap1 [15] consisted of 3.3 million SNPs and indels from 27 lines, HapMap2 [16] consisted of 55 million SNPs from 103 lines, and HapMap3 [17] consisted of 83 million SNPs from 1218 lines. These datasets have greatly enhanced our understanding of maize domestication

Figure 1



Overview of the key traits that determine maize yield.

and improvement. For example, HapMap2 identified 484 domestication and 695 improvement sweeps which helped Sosso *et al.* to understand the selection feature of a domestication gene, *ZmSWEET4c*, which affected carbohydrate import into seeds and directly determined seed size [18]. During the last decade, there has also been significant progress in genome-wide association studies based on DNA sequencing of various large-scale populations. We will not summarize these studies in detail, as a number of related reviews are already available [19–23]. Although high-throughput short-read sequencing has accelerated the discovery of natural genetic variants among diverse maize germplasms, it also has several unavoidable limitations: most accurately identified variants consist of SNPs and small indels but without structure variations (SVs); genome assemblies based on high-throughput short-read sequencing contain many gaps and hundreds of thousands of small contigs due to the complexity of the maize genome [24,25]; and most of the variant calls are based on a single reference genome which could not represent all the genetic diversity of maize [26,27••].

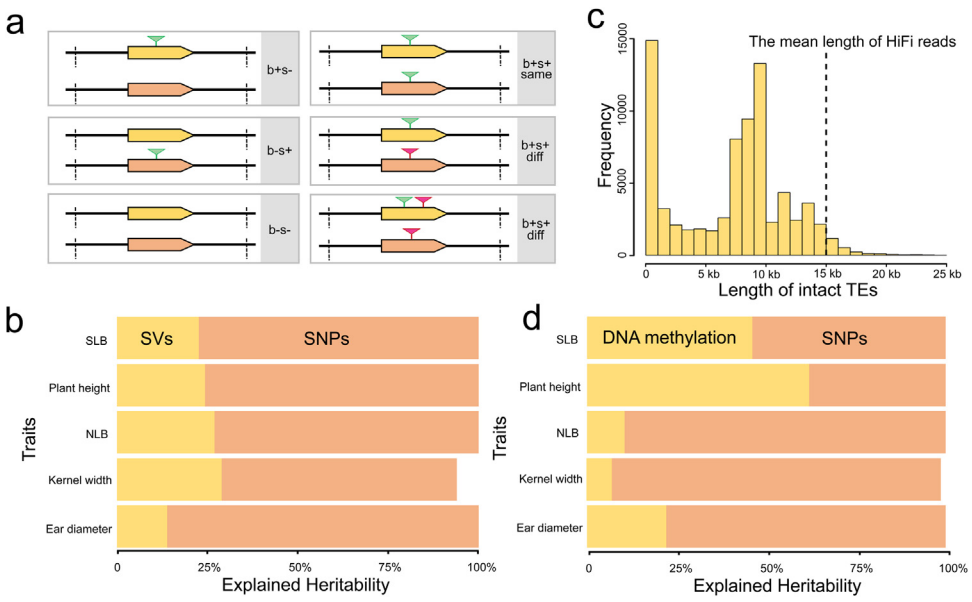
Multiple high-quality reference genomes are needed for maize genome research

The quality of the B73 reference genome has been vastly improved [14] by the development of long-read sequencing, which is dominated by Pacific Biosciences' (PacBio) single molecule real-time (SMRT) sequencing and Oxford Nanopore Technologies' (ONT) nanopore sequencing. As a result, researchers have identified more than 130 000 intact transposable elements (TEs) [14]. However, a single reference genome is insufficient for mining and utilizing the full spectrum of genetic diversity in maize. 92.5% of orthologous gene pairs have at least one TE insertion variation between SK and B73 (Figure 2a), and a comparison of the B73 and Mo17 genomes revealed that more than 10% of genes do not have homologs in the alternate genome [26]. With advances in technology and reduced costs of long-read sequencing, more high-quality genomes (contig N50 > 1 Mb) have been released since 2017. These include the B73 Ref_V4 [14], Mo17 [26], SK [27••], K0326Y [28], and NC358 [29] genomes. As more and more genomes are assembled, pangenome graphs are a promising format in which to present the diversity of multiple reference genomes [30•,31,32]. A good example is the first maize Practical Haplotype Graph (PHG) database built by Franco *et al.* [33••], which consists of a pangenome haplotype database from 27 high-quality maize assemblies (the founders of the NAM population). The PHG has shown its ability to generate accurate genotype calls using either GBS or NGS reads, although it is a very early version that will benefit from further improvements, especially in the non-coding regions.

Structural variations play important roles in trait differences

In case studies, most of which began with map-based cloning, growing evidence has shown that SVs affect important yield-related traits such as tillering, flowering time, kernel row number, disease resistance, drought resistance, and protein quality [34–39]. Multiple reference genomes open the door to the investigation of SVs and their phenotypic effects at the genome scale, which may partly explain the phenomenon of missing heritability [27••]. Comparison of the SK genome (tropical germplasm with small kernels) with the B73 and Mo17 genomes (temperate germplasm) identified 386 014 SVs, of which 80 614 polymorphic SVs were genotyped in 521 inbred lines using high-depth DNA resequencing data [27••]. 21.9% of polymorphic SVs could not be represented by neighboring SNPs via linkage disequilibrium, indicating that the contribution of SVs to genetic

Figure 2



Complexity of the maize genome and the contributions of different variations to phenotypic differences. **(a)** A diagram showing how we classified TE insertions in gene bodies. We considered only the families of the TE insertions but not their lengths. For example, b + s− means that the TE insertion is present in the B73 gene but not in its SK ortholog. The green and red triangles represent TEs from different families. TE insertions upstream and downstream (±2 kb) were calculated. **(b)** The proportion of phenotypic heritability explained by SNPs and SVs. SLB, southern corn leaf blight. NLB, northern corn leaf blight. Phenotype data were downloaded from http://maizego.org/Resources.html. One million randomly selected SNPs and all SVs were obtained from a previous study [27••]. Kinship matrices of SNPs and SVs [27••] were calculated separately using Tassle5 [40] and then input to LDAK [41] to calculate the heritability of SNPs and SVs. **(c)** The length distribution of annotated intact TEs in the B73_RefV4 genome. **(d)** The proportion of phenotypic heritability explained by SNPs and DNA methylation variations. DNA methylation variation data were obtained from a previous study [55•]. The kinship matrix of DNA methylation variation was calculated using OSCA [57].

diversity cannot be ignored. Furthermore, we found that up to 30.8% heritability of kernel width and up to 27.0% heritability of northern corn leaf blight (NLB) were explained by SVs (Figure 2b.). This finding is surprising and provides direct insight into the effects of SVs on traits. In addition, the SK genome permitted researchers to efficiently identify the causal variant (an 8.9-kb insertion) of a maize kernel-size and weight QTL, *qHKW1* [27••]. Using only one reference genome and a traditional map-based cloning strategy, *qHKW1* would be difficult to clone and its functional variation would be difficult to detect. These results emphasize that multiple reference genomes can help us to accurately identify SVs and rapidly understand their phenotypic effects.

Long-read sequencing enables high quality genome assembly and variation mining

Currently, long-read sequencing (SMRT and ONT) has overcome early limitations in price, accuracy, and throughput and now permits longer read lengths. For example, the PacBio HiFi sequencing method yields a highly accurate long-read sequencing maize dataset with read lengths averaging 15.6 kb, accuracies greater than 99.5%, and a throughput >24 Gb HiFi reads per cell (SMRT Cell 8 M) [42] at a cost of ~$1500 USD. These changes will revolutionize genomic studies in multiple ways, especially for complex genomes such as maize. First, SVs detected by comparing limited numbers of genomes are unlikely to fully represent the substantial genetic diversity of maize. Long-read resequencing now enables a broad survey of SVs at the population scale and will give us a more comprehensive understanding of SVs, such as that recently achieved for tomato [43]. Second, although the majority of TEs are thought to be silent in maize, more and more studies have found that a subset of TEs are transcriptionally active with tissue-specific expression [44] and can be induced by stress conditions [45•]. Because of the highly repetitive nature of most TEs, many things about TEs remain unknown in maize, including their evolutionary relationship with SVs and their potential influence on gene expression or regulation [46]. We annotated B73_RefV4 using a joint TE annotation pipeline, EDTA [47]. 96.9% of intact TEs were shorter than 15 kb (Figure 2c), which means that HiFi reads can span repetitive regions where short reads cannot easily be uniquely mapped.

Beyond diversity at the DNA level

Growing evidence indicates that genetic variants influence complex traits by modulating gene expression;

typical examples in maize include *tb1*, *ZmCCT9*, *ZmCCT10*, and others [34–36]. In addition to these case studies, several studies have delineated the gene expression regulatory landscape of complex traits. For instance, the complex regulatory network of the developing maize kernel has been revealed by RNA sequencing of diverse maize inbred lines [48]. Similarly, Liu *et al.* characterized a drought-responsive regulatory network by RNA sequencing of 224 maize inbred lines under three water regimes [49]. Kremling *et al.* demonstrated that the dysregulation of gene expression caused by rare deleterious mutations, obtained from RNA sequencing of 255 lines from seven tissues, reduced seed-weight fitness [50]. A transcriptome-wide association study (TWAS) was also recently developed to identify expression–trait associations [51]. A recent study demonstrated that TWAS combined with GWAS increased the power to detect known genes and aided in prioritizing likely causal genes. Significantly associated genes identified by the combination of GWAS and TWAS explained more heritable variation for a majority of traits than did TWAS or GWAS alone [52].

Variations in DNA and chromatin modifications can also affect gene expression and play an important role in complex traits. DNA methylation, the best-studied example of chromatin variation in maize, has been shown to be heritable. The inheritance of DNA methylation levels across generations is quite stable, even in recombinant inbred line (RIL)/near-isogenic line (NIL) populations [53,54]. A recent study showed that epigenetic diversity may provide additional sources of variation that can be captured for maize improvement [55•]. Greater than sixty percent of the differentially methylated regions (DMRs) were not tagged by genetic variations (SNPs and SVs), consistent with previous studies in smaller populations [54,56], and association analysis of 986 metabolic traits using DMRs suggested that DNA methylation was associated with phenotypic variation in 156 traits [55•]. Using the same dataset, we found that DMRs explained 7–61% of the heritability of yield-related traits (Figure 2d). Epigenetic variation can create heritable variation in functional traits, underscoring the potential for applying epigenetics to crop improvement by developing new combinations of epialleles and DNA alleles that provide desired traits.

Three-dimensional configurations such as genome-wide chromatin interactions [58,59] and open chromatin [60] are crucial for gene regulation, but the variation in chromatin states within a population is largely unknown.

Genetic resources in teosintes and landraces should not be ignored

Abiotic and biotic stresses act as driving forces for maize evolution by affecting diversity at all levels [49,61,62] and producing different kinds of local adaptations. Studies of local adaptation that aim to understand how maize responds to the environment can explore genetic resources that can meet the challenges of climate change [63] and extreme weather events [64]. However, domestication and improvement bottlenecks have reduced the genetic diversity of maize by 20% and <5%, respectively [65]. The narrow genetic base of cultivated maize is becoming a major impediment to maize improvement efforts. The wild relatives of maize, collectively called teosintes, are native to Mexico and Central America and have adapted to a diverse range of environments, as reviewed previously [66]; they exhibit many unique biotic and abiotic tolerance traits that are absent in cultivated maize [67]. QTL mapping analysis using TeoNAM, derived from crossing five teosinte inbreds (*Zea mays* ssp. *parviglumis* and *Z. mays* ssp. *mexicana*) to the maize inbred line W22, identified three QTL for kernel row number and ear length in which the teosinte allele contributes to a more maize-like phenotype [68]. *Zea nicaraguensis* is a flood-tolerant species that grows on the Pacific coast of Nicaragua [69]. It shows the special trait of constitutive aerenchyma formation (CAF), although aerenchyma are not formed under well-drained soil conditions. Gong *et al.* produced an introgression line with four QTL for CAF from *Z. nicaraguensis* in a maize background and demonstrated that CAF could reduce the oxygen deficiency stress caused by waterlogging [70]. Tian *et al.* cloned one large-effect alleles (*UPA2*) for leaf angle in teosinte that were absent in maize. Introgressing the teosinte allele into modern hybrids increased yield by permitting growth at high density [71••]. There are many unused favorable alleles in maize wild relatives. Further research in this area will not only help us improve maize grain yield, but also enable us to better cope with the challenges of future climate change.

Other important genetic resources exist in landraces. During the worldwide spread of maize, humans have guided its adaptation to a vast range of climatic and ecological conditions through the selection of local landraces [72]. Introgression of alleles from landraces to an elite maize variety improved drought tolerance and produced higher yields [73]. However, few studies have identified the genes that underlie local adaptations [74,75]. Local adaptation studies face a number of difficulties: the collection of stress-tolerant landraces, the need for laborious field trials in different environments, the capture of causal genes by large SVs such as inversions [76], and the lack of corresponding reference genomes.

The main staple crops were domesticated approximately 10 000 years ago to meet the specific needs of humans [77,78] and were improved in different global regions [79•]. Although domesticated species and origins vary, people have had similar goals and have tended to select traits that confer high yields, nutrient density, and ease of cultivation. Large populations of cultivated crops and

their wild relatives have now been sequenced [80–83]. Understanding the mechanisms of convergent domestication or improvement in maize [79•,84] and other crops can permit knowledge-driven genetic improvement.

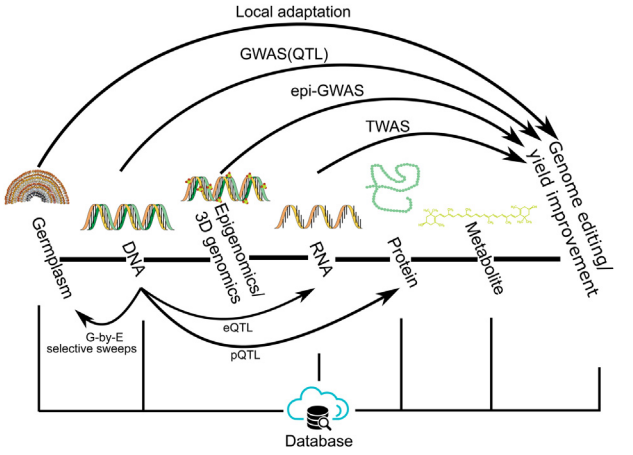
Mining omics in an integrated database

Comprehensive databases that store, maintain, analyze, and visualize the multi-omics data mentioned above can greatly accelerate maize genomic research. The most famous maize database is MaizeGDB [85], and its functions have been comprehensively reviewed [86]. The integration of multi-omics data generated from the same shareable individual/panel makes the best use of multi-omics information for maize improvement. A newly built database, ZEAMAP [87•], not only incorporates multiple reference genomes, annotations, comparative genomics, and transcriptomes, but also integrates chromatin interactions, high quality SNPs/SVs, DNA methylation, phenotypes, metabolomics, genetic maps, and genetic mapping loci from the same maize panel [88]. Hundreds, even thousands, of diverse and representative maize lines have been deeply resequenced by different labs across the world [27••,89–91]. However, each lab may use a different reference genome version and a different variant calling pipeline due to ongoing software development or personal preference. This creates a barrier to cross referencing, and data utilization is not maximized. In the future, it would be helpful to construct a comprehensive maize variation map that integrates all sequencing data worldwide. More and more high-quality genome assemblies will be generated, including assemblies for teosinte, landraces, diverse maize inbred lines, and population scale resequencing based on long reads. A true pangenome graph, a standardized variant identification pipeline, and an integrated database are in great demand.

Genetic engineering and mutant libraries

Genetic engineering can create substantial positive changes in complex traits such as grain yield [92••] and disease resistance [93]. These applications of transformative engineering were driven by mechanistic understanding. However, it is difficult to identify the exact causal gene for a desired trait using a forward-genetics approach, as the results always give a list of candidate genes in an interval that may include tens or hundreds of genes. Genome-wide mutant libraries are essential for identifying the functional genes. Traditional maize mutant libraries are based on random mutations induced by transposons [94,95] and/or ethyl-methanesulfonate mutagenesis [96] about which had been well reviewed previously [97]. However, lines derived from these libraries may harbor mutations at multiple loci, and many generations are required to stabilize loss-of-function mutations and determine the causal mutation that underlies a phenotype. Fortunately, homozygous mutants can be obtained in a single generation by CRISPR/Cas9-based editing. Moreover, by locating the sgRNA, researchers can easily link phenotype with genotype [98••]. Such genome-wide targeted mutagenesis has been performed in rice [99,100]. A recent study developed a CRISPR/Cas9-based editing platform adapted to high-throughput gene targeting in maize [101]. Although only 743 genes were targeted, the study was a good start towards developing genome-wide targeted mutagenesis in maize. However, low efficiency of maize genetic transformation [86] and high acquisition costs are still significant constraints that require the joint efforts of not only the maize community but also commercial companies and funding sponsors. The construction of a whole-genome CRISPR mutation library will revolutionize maize functional genomics research.

Figure 3



A schematic diagram showing that every biological level can be exploited to identify the causal variant underlying a phenotypic consequence and that these levels also complement each other.

Conclusions

Genomic approaches are thought to be one of the most powerful solutions for accelerating crop yield improvement to meet increasing global demands. With the rapid development of genomic technologies, all kinds of diversity in different environments can be comprehensively explored (Figure 3). Multi-omics data will be analyzed in a more systematic and integrated way to target genes that determine maize yield (Figure 3). Efficient [102] and genotype-independent [103] maize genome editing guided by mechanistic understanding and enabled by genomics approaches will make possible precise maize breeding to meet the growing and diversified demands of the future (Figure 3).

Conflict of interest statement

Nothing declared.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (31961133002, 31900494) and Young Elite Scientists Sponsorship

Program by CAST (2019QNRC001). We would like to thank Dr Songtao Gui from Huazhong Agricutural University for drawing picture of maize individual in Figure 1 and Jing Xu from Huazhong Agricultural University for sharing the DNA methylation variation data.

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. United States Environmental Protection Agency (EPA): *Report on the Environment-chemicals Used on Land*. 2020. Washington, DC.

2. Abberton M, Batley J, Bentley A, Bryant J, Cai H, Cockram J, de Oliveira AC, Cseke LJ, Dempewolf H, De Pace C *et al.*: **Global agricultural intensification during climate change: a role for genomics**. *Plant Biotechnol J* 2016, **14**:1095-1098.

3. Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA *et al.*: **The B73 maize genome: complexity, diversity, and dynamics**. *Science* 2009, **326**:1112-1115.

4. Yang N, Lu Y, Yang X, Huang J, Zhou Y, Ali F, Wen W, Liu J, Li J, Yan J: **Genome wide association studies using a new nonparametric model reveal the genetic architecture of 17 agronomic traits in an enlarged maize association panel**. *PLoS Genet* 2014, **10**:e1004573.

5. Xiao Y, Tong H, Yang X, Xu S, Pan Q, Qiao F, Raihan MS, Luo Y, Liu H, Zhang X *et al.*: **Genome-wide dissection of the maize ear genetic architecture using multiple populations**. *New Phytol* 2016, **210**:1095-1106.

6. Liu J, Huang J, Guo H, Lan L, Wang H, Xu Y, Yang X, Li W, Tong H, Xiao Y *et al.*: **The conserved and unique genetic architecture of kernel size and weight in maize and rice**. *Plant Physiol* 2017, **175**:774-785.

7. Inghelandt DV, Frey FP, Ries D, Stich B: **QTL mapping and genome-wide prediction of heat tolerance in multiple connected populations of temperate maize**. *Sci Rep* 2019, **9**:14418.

8. • Zhang Z, Zhang X, Lin Z, Wang J, Liu H, Zhou L, Zhong S, Li Y, Zhu C, Lai J *et al.*: **A large transposon insertion in the *stiff1*promoter increases stalk strength in maize**. *Plant Cell* 2020, **32**:152-165

This study cloned the *stiff1* gene, and one TE insertion in *stiff1* increased stalk strength and determined stalk lodging. Stalk lodging resistance can increase maize yield under high-density planting.

9. Pan Q, Xu Y, Li K, Peng Y, Zhan W, Li W, Li L, Yan J: **The genetic basis of plant architecture in 10 maize recombinant inbred line populations**. *Plant Physiol* 2017, **175**:858-873.

10. Zhang X, Huang C, Wu D, Qiao F, Li W, Duan L, Wang K, Xiao Y, Chen G, Liu Q *et al.*: **High-throughput phenotyping and QTL mapping reveals the genetic architecture of maize plant growth**. *Plant Physiol* 2017, **173**:1554-1564.

11. Chen G, Wang X, Long S, Jaqueth J, Li B, Yan J, Ding J: **Mapping of QTL conferring resistance to northern corn leaf blight using high-density SNPs in maize**. *Mol Breed* 2016, **36**:4.

12. Chen J, Shrestha R, Ding J, Zheng H, Mu C, Wu J, Mahuku G: **Genome-wide association study and QTL mapping reveal genomic loci associated with *Fusarium* ear rot resistance in tropical maize germplasm**. *G3 (Bethesda)* 2016, **6**:3803-3815.

13. Swarts K, Bauer E, Glaubitz JC, Ho T, Johnson L, Li Y, Li Y, Miller Z, Romay C, Schöen CC *et al.*: **A large scale joint analysis of flowering time reveals independent temperate adaptations in maize**. *bioRxiv* 2016 http://dx.doi.org/10.1101/086082. 086082.

14. Jiao Y, Peluso P, Shi J, Liang T, Stitzer MC, Wang B, Campbell MS, Stein JC, Wei X, Chin CS *et al.*: **Improved maize reference genome with single-molecule technologies**. *Nature* 2017, **546**:524-527.

15. Gore MA, Chia JM, Elshire RJ, Sun Q, Ersoz ES, Hurwitz BL, Peiffer JA, McMullen MD, Grills GS, Ross-Ibarra J *et al.*: **A first-generation haplotype map of maize**. *Science* 2009, **326**:1115-1117.

16. Chia JM, Song C, Bradbury PJ, Costich D, de Leon N, Doebley J, Elshire RJ, Gaut B, Geller L, Glaubitz JC *et al.*: **Maize HapMap2 identifies extant variation from a genome in flux**. *Nat Genet* 2012, **44**:803-807.

17. Bukowski R, Guo X, Lu Y, Zou C, He B, Rong Z, Wang B, Xu D, Yang B, Xie C *et al.*: **Construction of the third-generation *Zea mays* haplotype map**. *Gigascience* 2018, **7**:1-12.

18. Sosso D, Luo D, Li QB, Sasse J, Yang J, Gendrot G, Suzuki M, Koch KE, McCarty DR, Chourey PS *et al.*: **Seed filling in domesticated maize and rice depends on SWEET-mediated hexose transport**. *Nat Genet* 2015, **47**:1489-1493.

19. Liu HJ, Wang X, Xiao Y, Luo J, Qiao F, Yang W, Zhang R, Meng Y, Sun J, Yan S *et al.*: **CUBIC: an atlas of genetic architecture promises directed maize improvement**. *Genome Biol* 2020, **21**:20.

20. Li H, Peng Z, Yang X, Wang W, Fu J, Wang J, Han Y, Chai Y, Guo T, Yang N *et al.*: **Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels**. *Nat Genet* 2013, **45**:43-50.

21. Xiao Y, Liu H, Wu L, Warburton M, Yan J: **Genome-wide association studies in maize: praise and stargaze**. *Mol Plant* 2017, **10**:359-374.

22. Gage JL, Monier B, Giri A, Buckler ES: **Ten years of the maize nested association mapping population: impact, limitations, and future directions**. *Plant Cell* 2020, **32**:2083-2093.

23. Dell'Acqua M, Gatti DM, Pea G, Cattonaro F, Coppens F, Magris G, Hlaing AL, Aung HH, Nelissen H, Baute J *et al.*: **Genetic properties of the MAGIC maize population: a new platform for high definition QTL mapping in *Zea mays***. *Genome Biol* 2015, **16**:167.

24. Haberer G, Kamal N, Bauer E, Gundlach H, Fischer I, Seidel MA, Spannagl M, Marcon C, Ruban A, Urbany C *et al.*: **European maize genomes highlight intraspecies variation in repeat and gene content**. *Nat Genet* 2020, **52**:950-957.

25. Li C, Song W, Luo Y, Gao S, Zhang R, Shi Z, Wang X, Wang R, Wang F, Wang J *et al.*: **The HuangZaoSi maize genome provides insights into genomic variation and improvement history of maize**. *Mol Plant* 2019, **12**:402-409.

26. Sun S, Zhou Y, Chen J, Shi J, Zhao H, Zhao H, Song W, Zhang M, Cui Y, Dong X *et al.*: **Extensive intraspecific gene order and gene structural variations between Mo17 and other maize genomes**. *Nat Genet* 2018, **50**:1289-1295.

27. •• Yang N, Liu J, Gao Q, Gui S, Chen L, Yang L, Huang J, Deng T, Luo J, He L *et al.*: **Genome assembly of a tropical maize inbred line provides insights into structural variation and crop improvement**. *Nat Genet* 2019, **51**:1052-1059

This study constructed the first SV map at the population scale in maize and showed that ~21% of SVs could not be represented by SNPs.

28. Li C, Xiang X, Huang Y, Zhou Y, An D, Dong J, Zhao C, Liu H, Li Y, Wang Q *et al.*: **Long-read sequencing reveals genomic structural variations that underlie creation of quality protein maize**. *Nat Commun* 2020, **11**:17.

29. Ou S, Liu J, Chougule KM, Fungtammasan A, Seetharam AS, Stein JC, Llaca V, Manchanda N, Gilbert AM, Wei S *et al.*: **Effect of sequence depth and length in long-read assembly of the maize inbred NC358**. *Nat Commun* 2020, **11**:2288.

30. • Rakocevic G, Semenyuk V, Lee WP, Spencer J, Browning J, Johnson IJ, Arsenijevic V, Nadj J, Ghose K, Suciu MC *et al.*: **Fast and accurate genomic analyses using genome graphs**. *Nat Genet* 2019, **51**:354-362

In this study, the first human reference genome graph was built and shown to improve variant calling accuracy.

31. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL: **Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype**. *Nat Biotechnol* 2019, **37**:907-915.

32. Eizenga JM, Novak AM, Sibbesen JA, Heumos S, Ghaffaari A, Hickey G, Chang X, Seaman JD, Rounthwaite R, Ebler J *et al.*:

Pangenome graphs. *Annu Rev Genomics Hum Genet* 2020, **21**:139-162.

33. •• Franco JA, Gage JL, Bradbury PJ, Johnson LC, Miller Z, Buckler ES, Romay MC: **A maize practical haplotype graph leverages diverse NAM assemblies**. *bioRxiv* 2020 http://dx.doi.org/10.1101/2020.08.31.268425. 268425
This was the first attempt to build a reference genome graph in maize based on 27 maize genomes.

34. Studer A, Zhao Q, Ross-Ibarra J, Doebley J: **Identification of a functional transposon insertion in the maize domestication gene *tb1***. *Nat Genet* 2011, **43**:1160-1163.

35. Huang C, Sun H, Xu D, Chen Q, Liang Y, Wang X, Xu G, Tian J, Wang C, Li D *et al.*: ***ZmCCT9* enhances maize adaptation to higher latitudes**. *Proc Natl Acad Sci U S A* 2018, **115**:E334-E341.

36. Yang Q, Li Z, Li W, Ku L, Wang C, Ye J, Li K, Yang N, Li Y, Zhong T *et al.*: **CACTA-like transposable element in ZmCCT attenuated photoperiod sensitivity and accelerated the postdomestication spread of maize**. *Proc Natl Acad Sci U S A* 2013, **110**:16969-16974.

37. Zuo W, Chao Q, Zhang N, Ye J, Tan G, Li B, Xing Y, Zhang B, Liu H, Fengler KA *et al.*: **A maize wall-associated kinase confers quantitative resistance to head smut**. *Nat Genet* 2015, **47**:151-157.

38. Liu L, Du Y, Shen X, Li M, Sun W, Huang J, Liu Z, Tao Y, Zheng Y, Yan J *et al.*: **KRN4 controls quantitative variation in maize kernel row number**. *PLoS Genet* 2015, **11**:e1005670.

39. Mao H, Wang H, Liu S, Li Z, Yang X, Yan J, Li J, Tran LS, Qin F: **A transposable element in a NAC gene is associated with drought tolerance in maize seedlings**. *Nat Commun* 2015, **6**:8326.

40. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES: **TASSEL: software for association mapping of complex traits in diverse samples**. *Bioinformatics* 2007, **23**:2633-2635.

41. Speed D, Cai N, UCLEB Consortium, Johnson MR, Nejentsev S, Balding DJ: **Reevaluation of SNP heritability in complex human traits**. *Nat Genet* 2017, **49**:986-992.

42. Hon T, Mars K, Young G, Tsai YC, Karalius JW, Landolin JM, Maurer N, Kudrna D, Hardigan A, Steiner CC *et al.*: **Highly accurate long-read HiFi sequencing data for five complex genomes**. *bioRxiv* 2020 http://dx.doi.org/10.1101/2020.05.04.077180. 077180.

43. Alonge M, Wang X, Benoit M, Soyk S, Pereira L, Zhang L, Suresh H, Ramakrishnan S, Maumus F, Ciren D *et al.*: **Major impacts of widespread structural variation on gene expression and crop improvement in tomato**. *Cell* 2020, **182**:145-161.

44. Anderson SN, Stitzer MC, Zhou P, Ross-Ibarra J, Hirsch CD, Springer NM *et al.*: **Dynamic patterns of transcript abundance of transposable element families in maize**. *G3 (Bethesda)* 2019, **9**:3673-3682.

45. • Liang Z, Anderson SN, Noshay JM, Crisp PA, Enders TA, Springer NM: **Genetic and epigenetic contributions to variation in transposable element expression responses to abiotic stress in maize**. *bioRxiv* 2020 http://dx.doi.org/10.1101/2020.08.26.268102. 268102
This study showed that TE families lacking DNA methylation exhibited stress-responsive expression.

46. Chuong EB, Elde NC, Feschotte C: **Regulatory activities of transposable elements: from conflicts to benefits**. *Nat Rev Genet* 2017, **18**:71-86.

47. Ou S, Su W, Liao Y, Chougule K, Agda JRA, Hellinga AJ, Lugo CSB, Elliott TA, Ware D, Peterson T *et al.*: **Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline**. *Genome Biol* 2019, **20**:275.

48. Fu J, Cheng Y, Linghu J, Yang X, Kang L, Zhang Z, Zhang J, He C, Du X, Peng Z *et al.*: **RNA sequencing reveals the complex regulatory network in the maize kernel**. *Nat Commun* 2013, **4**:2832.

49. Liu S, Li C, Wang H, Wang S, Yang S, Liu X, Yan J, Li B, Beatty M, Zastrow-Hayes G *et al.*: **Mapping regulatory variants controlling gene expression in drought response and tolerance in maize**. *Genome Biol* 2020, **21**:163.

50. Kremling KAG, Chen SY, Su MH, Lepak NK, Romay MC, Swarts KL, Lu F, Lorant A, Bradbury PJ, Buckler ES: **Dysregulation of expression correlates with rare-allele burden and fitness loss in maize**. *Nature* 2018, **555**:520-523.

51. Wainberg M, Sinnott-Armstrong N, Mancuso N, Barbeira AN, Knowles DA, Golan D, Ermel R, Ruusalepp A, Quertermous T, Hao K *et al.*: **Opportunities and challenges for transcriptome-wide association studies**. *Nat Genet* 2019, **51**:592-599.

52. Kremling KAG, Diepenbrock CH, Gore MA, Buckler ES, Bandillo NB: **Transcriptome-wide association supplements genome-wide association in *Zea mays***. *G3 (Bethesda)* 2019, **9**:3023-3033.

53. Li Q, Eichten SR, Hermanson PJ, Springer NM: **Inheritance patterns and stability of DNA methylation variation in maize near-isogenic lines**. *Genetics* 2014, **196**:667-676.

54. Eichten SR, Briskine R, Song J, Li Q, Swanson-Wagner R, Hermanson PJ, Waters AJ, Starr E, West PT, Tiffin P *et al.*: **Epigenetic and genetic influences on DNA methylation variation in maize populations**. *Plant Cell* 2013, **25**:2783-2797.

55. • Xu J, Chen G, Hermanson PJ, Xu Q, Sun C, Chen W, Kan Q, Li M, Crisp PA, Yan J *et al.*: **Population-level analysis reveals the widespread occurrence and phenotypic consequence of DNA methylation variation not tagged by genetic variation in maize**. *Genome Biol* 2019, **20**:243
Differentially methylated regions (DMRs) could not be totally represented by SNPs/SVs. DNA methylation variations affected maize phenotypes such as metabolites and could be used as new bio-markers for genome-wide association studies.

56. Eichten SR, Swanson-Wagner RA, Schnable JC, Waters AJ, Hermanson PJ, Liu S, Yeh CT, Jia Y, Gendler K, Freeling M *et al.*: **Heritable epigenetic variation among maize inbreds**. *PLoS Genet* 2011, **7**:e1002372.

57. Zhang F, Chen W, Zhu Z, Zhang Q, Nabais MF, Qi T, Deary IJ, Wray NR, Visscher PM, McRae AF, Yang J: **OSCA: a tool for omic-data-based complex trait analysis**. *Genome Biol* 2019, **20**:107.

58. Li E, Liu H, Huang L, Zhang X, Dong X, Song W, Zhao H, Lai J: **Long-range interactions between proximal and distal regulatory regions in maize**. *Nat Commun* 2019, **10**:2633.

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

60. Rodgers-Melnick E, Vera DL, Bass HW, Buckler ES: **Open chromatin reveals the functional maize genome**. *Proc Natl Acad Sci U S A* 2016, **113**:E3177-E3184.

61. Makarevitch I, Waters AJ, West PT, Stitzer M, Hirsch CN, Ross-Ibarra J, Springer NM: **Transposable elements contribute to activation of maize genes in response to abiotic stress**. *PLoS Genet* 2015, **11**:e1004915.

62. Forestan C, Farinati S, Zambelli F, Pavesi G, Rossi V, Varotto S: **Epigenetic signatures of stress adaptation and flowering regulation in response to extended drought and recovery in *Zea mays***. *Plant Cell Environ* 2020, **43**:55-75.

63. Bailey-Serres J, Parker JE, Ainsworth EA, Oldroyd GED, Schroeder JI: **Genetic strategies for improving crop yields**. *Nature* 2019, **575**:109-118.

64. Lesk C, Rowhani P, Ramankutty N: **Influence of extreme weather disasters on global crop production**. *Nature* 2016, **529**:84-87.

65. Hufford MB, Xu X, van Heerwaarden J, Pyhäjärvi T, Chia JM, Cartwright RA, Elshire RJ, Glaubitz JC, Guill KE, Kaeppler SM *et al.*: **Comparative population genomics of maize domestication and improvement**. *Nat Genet* 2012, **44**:808-811.

66. Hufford MB, Bilinski P, Pyhäjärvi T, Ross-Ibarra J: **Teosinte as a model system for population and ecological genomics**. *Trends Genet* 2012, **28**:606-615.

67. Mammadov J, Buyyarapu R, Guttikonda SK, Parliament K, Abdurakhmonov IY, Kumpatla SP: **Wild relatives of maize, rice, cotton, and soybean: treasure troves for tolerance to biotic and abiotic stresses**. *Front Plant Sci* 2018, **9**:886.

68. Chen Q, Yang CJ, York AM, Xue W, Daskalska LL, DeValk CA, Krueger KW, Lawton SB, Spiegelberg BG, Schnell JM *et al.*: **TeoNAM: a nested association mapping population for domestication and agronomic trait analysis in maize**. *Genetics* 2019, **213**:1065-1078.

69. Iltis HH, Benz FB: ***Zea nicaraguensis* (Poaceae), a New Teosinte from Pacific Coastal Nicaragua**. *Novon* 2000, **10**:382-390.

70. Gong F, Takahashi H, Omori F, Wang W, Mano Y, Nakazono M: **QTLs for constitutive aerenchyma from *Zea nicaraguensis* improve tolerance of maize to root-zone oxygen deficiency**. *J Exp Bot* 2019, **70**:6475-6487.

71. •• Tian J, Wang C, Xia J, Wu L, Xu G, Wu W, Li D, Qin W, Han X, Chen Q *et al.*: **Teosinte ligule allele narrows plant architecture and enhances high-density maize yields**. *Science* 2019, **365**:658-664

This study first showed the great potential of teosinte genetic resources and demonstrated that a teosinte allele that decreased leaf angle could increase maize yield by ~20% by increasing planting density.

72. Camacho Villa TC, Maxted N, Scholten M, Ford-Lloyd B: **Defining and identifying crop landraces**. *Plant Genet Res* 2005, **3**:373-384.

73. Meseka S, Fakorede M, Ajala S, Badu-Apraku B, Menkir A: **Introgression of alleles from maize landraces to improve drought tolerance in an unadapted germplasm**. *J Crop Improv* 2013, **27**:96-112.

74. Romero Navarro JA, Willcox M, Burgueño J, Romay C, Swarts K, Trachsel S, Preciado E, Terron A, Delgado HV, Vidal V *et al.*: **A study of allelic diversity underlying flowering-time adaptation in maize landraces**. *Nat Genet* 2017, **49**:476-480.

75. Crow T, Ta J, Nojoomi S, Aguilar-Rangel MR, Rodrǵuez JVT, Gates D, Rellan-Alvarez R, Sawers R, Runcie D: **Gene regulatory effects of a large chromosomal inversion in highland maize**. *bioRxiv* 2020 http://dx.doi.org/10.1101/861583. 861583.

76. Pyhäjärvi T, Hufford MB, Mezmouk S, Ross-Ibarra J: **Complex patterns of local adaptation in teosinte**. *Genome Biol Evol* 2013, **5**:1594-1609.

77. Doebley JF, Gaut BS, Smith BD: **The molecular genetics of crop domestication**. *Cell* 2006, **127**:1309-1321.

78. Linares OF: **African rice (*Oryza glaberrima*): history and future potential**. *Proc Natl Acad Sci U S A* 2002, **99**:16360-16365.

79. • Wang B, Lin Z, Li X, Zhao Y, Zhao B, Wu G, Ma X, Wang H, Xie Y, Li Q *et al.*: **Genome-wide selection and genetic improvement during modern maize breeding**. *Nat Genet* 2020, **52**:565-571

This study sequenced 350 elite inbred lines representing multiple eras of germplasm from both China and America. The authors found several convergent phenotypic changes in both countries and used evolutionary genomics approaches to identify 160 loci underlying the convergent phenotypes.

80. Huang X, Kurata N, Wei X, Wang ZX, Wang A, Zhao Q, Zhao Y, Liu K, Lu H, Li W *et al.*: **A map of rice genome variation reveals the origin of cultivated rice**. *Nature* 2012, **490**:497-501.

81. Wang M, Li W, Fang C, Xu F, Liu Y, Wang Z, Yang R, Zhang M, Liu S, Lu S *et al.*: **Parallel selection on a dormancy gene during domestication of crops from multiple families**. *Nat Genet* 2018, **50**:1435-1441.

82. Hufford MB, Xu X, van Heerwaarden J, Pyhäjärvi T, Chia JM, Cartwright RA, Elshire RJ, Glaubitz JC, Guill KE, Kaeppler SM *et al.*: **Comparative population genomics of maize domestication and improvement**. *Nat Genet* 2012, **44**:808-811.

83. Zhou Y, Zhao X, Li Y, Xu J, Bi A, Kang L, Chen H, Wang Y, Wang YG, Liu S *et al.*: **Convergence within divergence: insights of wheat adaptation from Triticum population sequencing**. *bioRxiv* 2020 http://dx.doi.org/10.1101/2020.03.21.001362. 001362.

84. Wang L, Josephs EB, Lee KM, Roberts LM, Álvarez RR, Ross-Ibarra J, Hufford MB: **Molecular parallelism underlies convergent highland adaptation of maize landraces**. *bioRxiv* 2020 http://dx.doi.org/10.1101/2020.07.31.227629. 227629.

85. Portwood JL 2nd, Woodhouse MR, Cannon EK, Gardiner JM, Harper LC, Schaeffer ML, Walsh JR, Sen TZ, Cho KT, Schott DA *et al.*: **MaizeGDB 2018: the maize multi-genome genetics and genomics database**. *Nucleic Acids Res* 2019, **47**:D1146-D1154.

86. Liu J, Fernie AR, Yan J: **The past, present, and future of maize improvement: domestication, genomics, and functional genomic routes toward crop enhancement**. *Plant Commun* 2020, **1**:13.

87. • Gui S, Yang L, Li J, Luo J, Xu X, Yuan J, Chen L, Li W, Yang X, Wu S *et al.*: **ZEAMAP, a comprehensive database adapted to the maize multi-omics era**. *iScience* 2020, **23**:101241

This study built an integrated database of genomic variations, transcriptome, epigenomics and so on from a same maize association mapping panel and developed visual tools for comparison of multi-omics in the same window.

88. Yang X, Gao S, Xu S, Zhang Z, Prasanna BM, Li L, Li J, Yan J: **Characterization of a global germplasm collection and its potential utilization for analysis of complex quantitative traits in maize**. *Mol Breed* 2011, **28**:511-526.

89. Brandenburg JT, Mary-Huard T, Rigaill G, Hearne SJ, Corti H, Joets J, Vitte C, Charcosset A, Nicolas SD, Tenaillon MI: **Independent introductions and admixtures have contributed to adaptation of European maize and its American counterparts**. *PLoS Genet* 2017, **13**:e1006666.

90. Kistler L, Maezumi SY, Gregorio de Souza J, Przelomska NAS, Malaquias Costa F, Smith O, Loiselle H, Ramos-Madrigal J, Wales N, Ribeiro ER *et al.*: **Multiproxy evidence highlights a complex evolutionary legacy of maize in South America**. *Science* 2018, **362**:1309-1313.

91. Wang L, Beissinger TM, Lorant A, Ross-Ibarra C, Ross-Ibarra J, Hufford MB: **The interplay of demography and selection during maize domestication and expansion**. *Genome Biol* 2017, **18**:215.

92. •• Wu J, Lawit SJ, Weers B, Sun J, Mongar N, Van Hemert J, Melo R, Meng X, Rupe M, Clapp J *et al.*: **Overexpression of *zmm28* increases maize grain yield in the field**. *Proc Natl Acad Sci U S A* 2019, **116**:23850-23858

This is an excellent case study demonstrating how genome editing guided by mechanistic understanding can be used to improve maize yield. Increasing and extending the expression of a maize MADS-box transcription factor gene, zmm28, increased maize plant growth, photosynthetic capacity, and nitrogen utilization.

93. USDA/APHIS Letter (2017) https://www.aphis.usda.gov/biotechnology/downloads/reg_loi/17-076-01_air_inquiry_cbidel.pdf.

94. McCarty DR, Settles AM, Suzuki M, Tan BC, Latshaw S, Porch T, Robin K, Baier J, Avigne W, Lai J *et al.*: **Steady-state transposon mutagenesis in inbred maize**. *Plant J* 2005, **44**:52-61.

95. Cowperthwaite M, Park W, Xu Z, Yan X, Maurais SC, Dooner HK: **Use of the transposon Ac as a gene-searching engine in the maize genome**. *Plant Cell* 2002, **14**:713-726.

96. Lu X, Liu J, Ren W, Yang Q, Chai Z, Chen R, Wang L, Zhao J, Lang Z, Wang H *et al.*: **Gene-indexed mutations in maize**. *Mol Plant* 2018, **11**:496-504.

97. Nannas NJ, Dawe RK: **Genetic and genomic toolbox of *Zea mays***. *Genetics* 2015, **199**:655-669.

98. •• Chen K, Wang Y, Zhang R, Zhang H, Gao C: **CRISPR/Cas genome editing and precision plant breeding in agriculture**. *Annu Rev Plant Biol* 2019, **70**:667-697

This is a comprehensive review on the use of CRISPR/Cas to accelerate functional genomics research in plants and improve crop yield.

99. Lu Y, Ye X, Guo R, Huang J, Wang W, Tang J, Tan L, Zhu JK, Chu C, Qian Y: **Genome-wide targeted mutagenesis in rice using the CRISPR/Cas9 system**. *Mol Plant* 2017, **10**:1242-1245.

100. Meng X, Yu H, Zhang Y, Zhuang F, Song X, Gao S, Gao C, Li J: **Construction of a genome-wide mutant library in rice using CRISPR/Cas9**. *Mol Plant* 2017, **10**:1238-1241.

101. Liu H, Jian L, Xu J, Zhang Q, Zhang M, Jin M, Peng Y, Yan J, Han B, Liu J *et al.*: **High-throughput CRISPR/Cas9 mutagenesis streamlines trait gene identification in maize**. *Plant Cell* 2020, **32**:1397-1413.

102. Lowe K, Wu E, Wang N, Hoerster G, Hastings C, Cho MJ, Scelonge C, Lenderts B, Chamberlin M, Cushatt J *et al.*: **Morphogenic regulators *Baby boom* and *Wuschel*improve monocot transformation**. *Plant Cell* 2016, **28**:1998-2015.

103. Lowe K, La Rota M, Hoerster G, Hastings C, Wang N, Chamberlin M, Wu E, Jones T, Gordon-Kamm W: **Rapid genotype "independent" *Zea mays* L. (maize) transformation via direct somatic embryogenesis**. *In Vitro Cell Dev Biol Plant* 2018, **54**:240-252.