We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter 2

Integration of Next-generation Sequencing Technologies with Comparative Genomics in Cereals

Thandeka N. Sikhakhane, Sandiswa Figlan, Learnmore Mwadzingeni, Rodomiro Ortiz and Toi J. Tsilo

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/61763

Abstract

Cereals are the major sources of calories worldwide. Their production should be high to achieve food security, despite the projected increase in global population. Genomics research may enhance cereal productivity. Genomics immensely benefits from robust next-generation sequencing (NGS) techniques, which produce vast amounts of sequence data in a time and cost-efficient way. Research has demonstrated that gene sequences among closely related species that share common ancestry have remained well conserved over millions of years of evolution. Comparative genomics allows for comparison of genome sequences across different species, with the implication that genomes with large sizes can be investigated using closely related species with smaller genomes. This offers prospects of studying genes in a single species and, in turn, gaining information on their functions in other related species. Comparative genomics is expected to provide invaluable information on the control of gene function in complex cereal genomes, and also in designing molecular markers across related species. This chapter discusses advances in sequencing technologies, their application in cereal genomics and their potential contribution to the understanding of the relationships between the different cereal genomes and their phenotypes.

Keywords: Bioinformatics, Cereals, Comparative genomics, Next-generation sequencing, Synteny

1. Introduction

Significant limitations to cereal crop production and productivity pose a threat to global food security since these crops are the main sources of calories that support the ever-growing human population. Despite the significant progress that has been made in the improvement of edible

INTECH
open science | open minds

© 2016 The Author(s). Licensee InTech. This chapter is distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

yield through classical breeding techniques, the current rates of increase in grain yield in several major cereal crops are still too slow to catch up with the increasing demand of the growing population [1, 2]. This is likely to get worse according to the projected climate change scenarios [3], as it also affects biotic stresses such as pests, diseases and weeds, and abiotic stresses including drought, extreme temperatures, salinity and nutrient deficiencies [4-6]. Although there are various strategies to cope with these constraints, Kole [7] suggested the use of genomics-assisted breeding as an effective and economic strategy.

Despite the sustainability of breeding resilient crops, there are still several genomic constraints to genome-based selection and stress resistance improvement, particularly for multigenic traits. A poor understanding of the genetic basis and the regulatory mechanisms of various stresses is among the major challenges for successful genetic manipulation through gene introgression, gene pyramiding, gene stacking or gene silencing. Additionally, more diagnostic genetic markers are necessary to improve the current limited success in marker application in both foreground and background selection. These challenges are related to the fact that genomes of some cereal crops are not yet fully sequenced and annotated, either because the crops have been under-researched or the genomes are huge and structurally complex. For instance, the hexaploid wheat (*Triticum aestivum*) genome is the largest (about 17 billion nucleotides) among cultivated cereals, and is multifaceted by repetitive DNA sequences [8]. Furthermore, dissection of the genetic and regulatory mechanisms of host plant resistance is complicated because most traits of interest are multigenic and thus influenced by several genes with additive and nonadditive gene effects. Hence, tools that detect the genetic variation at the genome sequence level allow all genes controlling particular traits to be investigated for various genetic applications to realize phenotypic gains from genetic manipulation.

Enhanced application of next-generation sequencing (NGS) techniques in cereal crops is revolutionizing and speeding up plant breeding. The advances that have been made so far in the use of NGS, particularly with the human genome in the field of medicine, and on various model crops through plant biotechnology, envisions the following in cereals and other crops: first, complete sequencing of small and less complex plant genomes is increasingly becoming possible as costs have dropped significantly and more sequences are being generated in a shorter time than before. Secondly, the genetic mechanisms of particular traits in huge and complex plant genomes can now be investigated using small and less complex genomes of related plants sharing conserved regions through comparative genomics. This will potentially identify genes or quantitative trait loci (QTL) and putative single nucleotide polymorphism (SNP) markers for genome-wide association mapping and annotation of genomes. This chapter discusses the advances made in improving sequencing technologies and how these advances can assist in generating complete sequences for the improvement of genome-aided selection. This will also assist in identifying the unique sequences responsible for the major differences existing among cereals.

2. The need for high-throughput genome and transcriptome sequencing

Since the discovery of the DNA molecule by Friedrich Miescher in 1869 [9], and the subsequent exposition of its double-helical structure by Watson and Crick in 1953, significant knowledge

http://dx.doi.org/10.5772/61763

has been gained on the flow of genetic information. Understanding how this genetic information influences the phenotype (trait) of interest has, however, remained a challenge. This is mainly because the overall instruction contributing to the phenotype is not restricted to the coding region but is also influenced by some posttranscriptional modifications controlled by noncoding DNA [10-12]. Also, multigenic traits are influenced by complex interactions of alleles at different loci, having major or minor influence [13]. These, together with differential genotype-by-environment interactions, add to the structural and functional complexity of most cereal genomes that are multifaceted by repetitive DNA sequences, transposable elements and polyploid genomes, as in the case of wheat and finger millet (*Eleusine coracana*) [8, 14]. Whole genome and transcriptome sequencing therefore become a necessity so that all the genomic and transcriptomic variation can be detected. NGS and various 'omic' technologies, including genomics, transcriptomics, proteomics, metabolomics and phenomics, offer prospects towards whole-genome annotation; particularly in cereals that have small and less complex genomes. This will simplify comparative genomics and evolutionary genetic research, which will enhance the manipulation and exploitation of important genes for cereal improvement.

NGS technologies are one of the available tools that can produce complete sequences for diverse research at the DNA and RNA level within and across species. Firstly, this will make it easy to obtain the entire DNA, coding and noncoding regions. Secondly, this will simplify studies on the whole transcriptome, including RNAs involved in protein synthesis such as the messenger, ribosomal, signal recognition particle, transfer and transfer-messenger RNAs and other RNAs involved in posttranscriptional modifications, such as small RNAs [15]. Quantification of such transcripts through NGS under various stress conditions will precisely determine the levels of gene expression within and across different species.

3. Advances in sequencing technologies

Since the pioneering of genome sequencing through technologies such as Sanger sequencing [16], significant advances have been made to resolve the limitations of the early technologies. This has seen the development of more sophisticated sequencing technologies that allow *de novo* genome sequencing, generating vast amounts of data in a short period at low costs. Table 1 summarizes the advances made in sequencing technology development, from the advent of the chain termination sequencing [16], to prominent NGS technologies including Roche/454 sequencing [17], Illumina (Solexa) sequencing [18], sequencing by oligonucleotide ligation and detection (SOLiD) [19], the single molecule sequence pioneered by Helicos Biosciences [20] and Ion Torrent sequencing [21]. These technological advances are instrumental in whole-genome research and are expected to simplify comparative genomics within species and across distantly related cereals and grasses. Several modifications are available for each of these technologies and fine-tuned protocols are constantly being developed to address some of the current limitations.

Although NGS technologies have enormous prospective benefits, they come with their own limitations that need to be addressed to realize their full potential. Key among these drawbacks are the bioinformatic and computational challenges related to storage, image analysis, base

calling and integration of the large amounts of data that are generated in several terabytes per day. Apparently, the large amount of sequence data that is being generated on a daily basis in cereal genomics cannot be transformed into information that is useful for the detection of important genomic variants within and among species or in identifying genes that are differentially expressed under particular stress conditions. Hence, investment in computational and high-throughput bioinformatic equipment and human resources and combining the various NGS technologies will allow the data generated using different NGS techniques by various laboratories to be related and used to build onto each other. Unlike traditional marker technologies, NGS is currently dissociated from phenomics, yet it should be complementary to high-throughput phenotyping in order to relate sequence variations to traits of interest for progressive discoveries through genome-wide association mapping, particularly for multigenic traits like adaptation to drought in complex cereal genomes [22]. Additionally, NGS technologies are still associated with high error rates [23] and short read lengths that limit data analysis accuracy. This further confuses detection and distinction of sequence variations including large amounts of duplications, deletions, inversions and chromosomal rearrangements that characterize cereal genomes.

Technologies (Developer)	**Year**	**Sequencing chemistry**	**Throughput**	**Read length**	**References**
Sanger sequencing (Frederick Sanger and team)	1977	Involves DNA polymerase based selective amplicon-termination of in vitro DNA replication by radioactively or fluorescently labeled di-deoxynucleotide triphosphates, followed by electrophoresis and UV or X-ray spectra detection of DNA sequences. Major limitations of the Sanger technique Since the technique relies on cloning vectors, there is potential for a mix up of the target sequences with some DNA portions from the clonal vector. Additionally, it requires a lot of labor and space since multiplexing is not possible.	Up to 84 Kb per about 3 hr run	Up to 1000 bp	[16]
Roche/454 (Life Sciences)	2004	First NGS technique This is a sequencing by synthesis (SBS) technique where DNA fragments attached to adapters annealed to beads are PCR amplified using adapter specific primers. Addition of each dNTP is associated with the release of a pyrophosphate, which is converted to ATP energy used to produce an optical signal (light). The light allows reading of the beads to which the dNTP is added, hence deducing the sequence (Pyrosequencing).	700 Mb of sequence data per 23 hr run	Up to 1,000 bp	[17]

Technologies (Developer)	Year	Sequencing chemistry	Throughput	Read length	References
Illumina/ Solexa (Illumina -Inc)	2006	This sequencing by synthesis technology uses dNTPs with reversible dye-terminators ensuring that DNA polymerase enzyme adds only one base to a growing DNA strand. The terminators are removed after the images of the four dNTPs added to the growing sequence are recorded, and the cycle is repeated.	20 to 130 Gb in 15 to 30 hr run time	Up to 3000 bp	[18]
ABI SOLiD (Life Technologies)	2006	Clonal bead populations prepared from a library of same species of DNA fragments each with a universal P1 adapters are attached on the surface of magnetic beads. On the universal adapters, primers are hybridized, on which a set of four dye-labeled di-base probes compete for ligation. A series of ligation cycles is followed by cleavage of the extension product then the template is reset for the next cycle of ligation by annealing a primer complimentary to the next adapter.	At least 20 Gb in about 3.5 days run time	Up to 50 bp	[19]
Helicos (Helicos Biosciences)	2009	This is a single molecule fluorescent sequencing technique which achieves direct DNA or RNA sequencing without amplification through imaging light emitting single molecules corresponding to each nucleotide base. DNA sequencing is achieved through an imaging system. The technology identifies the exact sequence of a piece of DNA and does not require PCR amplification, thus have reduced amplification bias.	21 to 35 Gb per 8 days run	35 bp average	[20]
Ion Torrent (Life Technologies)	2010	This sequencing by synthesis technique involves detection of a pH change caused by hydrogen ions released when a dNTP complementary to the leading unpaired template nucleotide is added to the growing strand. The electrical pulses transmitted to a computer in this process are subsequently translated into a DNA sequence	Up to 2 Gb per 2.3 to 7.3 hr run	35 to 400 bp (average: 200 bp)	[21]

Table 1. Evolution of next-generation sequencing technologies.

4. Application of next-generation sequencing in cereal biotechnology

Among the major cereals, the relatively small rice (*Oryza sativa*) genome (~389 Mb) has long been fully sequenced by the International Rice Genome Project [24]. Kawahara [25] recently demonstrated, however, the robustness of NGS technologies by revising the rice genome using the Illumina and Roche 454 pyrosequencing platforms. Their study noted some errors in the initial assembly. This research provides sufficient evidence that high quality and validated reference genomes can be produced among most cereals through resequencing using NGS technologies. Also, a recent whole genome-wide study of the hexaploid wheat genome (~17 Gb) using the Roche/454 pyrosequencing technology reviewed the capacity of NGS technologies to resequence huge and complex genomes and to identify SNPs for dissection of quantitative traits [26]. Similarly, Illumina sequencing was recently used to quantify the transposable element (TE) content in the complex maize *(Zea mays)* genome (~2.3 Gb) [27] and to estimate their potential contribution to the genome size differences between the cultivated species and its close relative, *Zea luxurians* [28]. The latter also reported high proportions of conserved TE families between the two species, revealing the potential of NGS technologies to enhance evolutionary and comparative genomic studies. Other major cereals whose genomes have been sequenced and are expected to further benefit from NGS technologies include barley *(Hordeum vulgare)* (~5.1 Gb) [29] and sorghum *(Sorghum bicolor)* (~730 Mb) [30].

Minor and under-researched cereals such as the allotetraploid finger millet *(Eleusine coracana*) —which has a genome size of about 1.76 Gb [31]—and the diploids, pearl millet *(Pennisetum glaucum*)—with a genome size of about 4.6 Gb [32]—and tef *(Eragrostis tef)*—with a 714 to 733 Mb genome [33]—have not received much benefit from NGS technologies. However, these crops are expected to benefit from the African Orphan Crops Consortium that has the mandate to use the latest scientific equipment and techniques to sequence, assemble and annotate genomes of under-researched crops [34]. These minor crops are renowned for their adaptation to various biotic and abiotic stresses, particularly drought. Thus, sequencing or resequencing their genomes will potentially expose huge amounts of relevant genetic information for cereal improvement. NGS technologies will have great application in comparing genomic features of cereal crops through comparative genomic research.

5. Comparative genomics in cereal crops

Core questions unanswered with traditional cereal biotechnology approaches include: (1) What are the genetic foundations that underlie the similarities between different grass species or individuals within a species? (2) What are the genetic variations responsible for the detected phenotypic differences? Comparative genomics is the branch of biology in which DNA sequence information from genomes of different life forms are compared in an effort to directly answer these questions. It was founded mainly on various ideas. Firstly, comprehensive analysis and comparison of whole genomes can uncover the essentially conserved and the important variable components of any set of genomes [35]. Secondly, differences in genome sequence (genotype) contribute to differences in genome function and therefore explain differences between phenotypic traits [36]. The application of comparative genomic informa-

tion on various plants including cereals has, however, been a challenge previously because of the large genome sizes of most species, which are complicated by high rates of structural rearrangements mainly due to transposable elements, duplications and inversions [35], as listed in Table 2.

Species	Clade (Subfamily, Tribe)	Ploidy level	Genome size	Repetitive DNA and retrotransposon content	References
Rice (*Oryza sativa)*	Ehrhartoideae, Oryzeae	*2n=2x*=24	420 to 460 Mb	~35%	[41]
Sorghum (*Sorghum bicolor)*	Panicoideae, Andropogoneae	*2n=2x*=20	~730 Mb	~61%	[30]
Pearl millet (*Pennisetum glaucum)*	Panicoideae, Paniceae	*2n=2x*=14	~4.6 Gb		[32]
Finger millet (*Eleusine coracana)*	Chloridoideae, Eragrostideae	*2n=4x*=40	3.3 to 3.8 Gb		[31]
Maize *(Zea mays)*	Panicoideae, Andropogoneae	*2n=2x*=20	~2.3 Gb	~78%	[27]
Barley (*Hordeum vulgare)*	Pooideae, Triticeae	*2n=2x*=14	~5.1 Gb	~76%	[29]
Rye *(Secale cereale)*	Pooideae, Triticeae	*2n=2x*=14	~7.9 Mb	~92%	No reference available
Bread wheat (*Triticum aestivum)*	Pooideae, Triticeae	*2n=6x*=42	~17 Gb	~80%	[8]
Oat (*Avena sativa)*	Pooideae, Triticeae	*2n=6x*=42	~11.3 Gb		No reference available
Tef (*Eragrostis tef)*	Chloridoideae,	*2n=4x*=40	~672 Mb		[33]

Table 2. Genome size, structure and genomic resources of major cereal species.

The application of comparative genomics for crop improvement has evolved over time. In the grass family, significant research provided remarkable and comprehensive datasets demonstrating high degree of collinearity or synteny among genomes at chromosome (macro) and gene (micro) levels [37, 38]. *Synteny*, from the Greek *syn* (together with) and *taenia* (ribbon), refers to loci contained within the same chromosome. Collinearity, on the other hand, refers to some degree of conservation of gene order between chromosomes of different species or between nonhomologous chromosomes of a single species [39]. A large number of sequences within the grass family has remained considerably conserved at the genome level over millions of years of evolution, irrespective of the differences in ploidy level, chromosome number and haploid DNA content [37]. This conservation of gene content and order at the megabase level makes it easy to use species with small genome sizes such as *Arabidopsis* and rice as model species for studying similar gene contents in other related species. Their applications include allele discovery, positional cloning, and comparative studies in related species [40]. There is, however, limited synteny and gene homology between *Arabidopsis* and rice, but an extensive collinearity between the latter and other grasses, thereby suggesting that rice is an appropriate

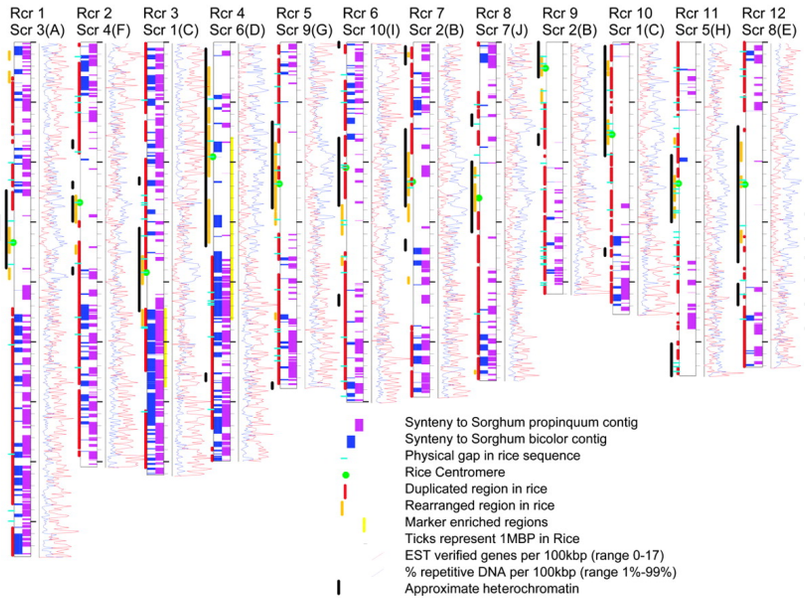
grass model species for cereal comparative genomics [41]. In this case, rice and purple false brome (*Brachypodium distachyon)* (genome size ~355 Mb), both of which are from the grass family, serve as functional model species for cereal comparative genomics owing to their small and fully sequenced genomes. Moreover, *Brachypodium* showed conservation of gene content and family structure with rice and sorghum [42]. A phylogenetic study carried out on seven grass species also revealed a close evolutionary relationship of *Brachypodium* with maize, barley and wheat based on 335 commonly shared sequences [43].

Microcollinearity has numerous interesting applications in cereal genome analysis including the transfer of genetic markers between species and the identification of candidate genes across species borders [44]. It is possible, due to such advances, to intensively study, decipher and understand the genetic makeup of the cereal genomes including those of rice, maize, wheat, barley and sorghum [30, 45-47]. Comparing the gene sequences of these cereal crops is the initial step towards understanding their morphological and functional similarities and differences. Comparative analysis research has been extended to the DNA sequence (micro) level, to allow the investigation of conservation of coding and noncoding regions as well as characterization of molecular mechanisms of genome evolution [38].

6. Several examples of macro- and microcollinearity in cereal crops

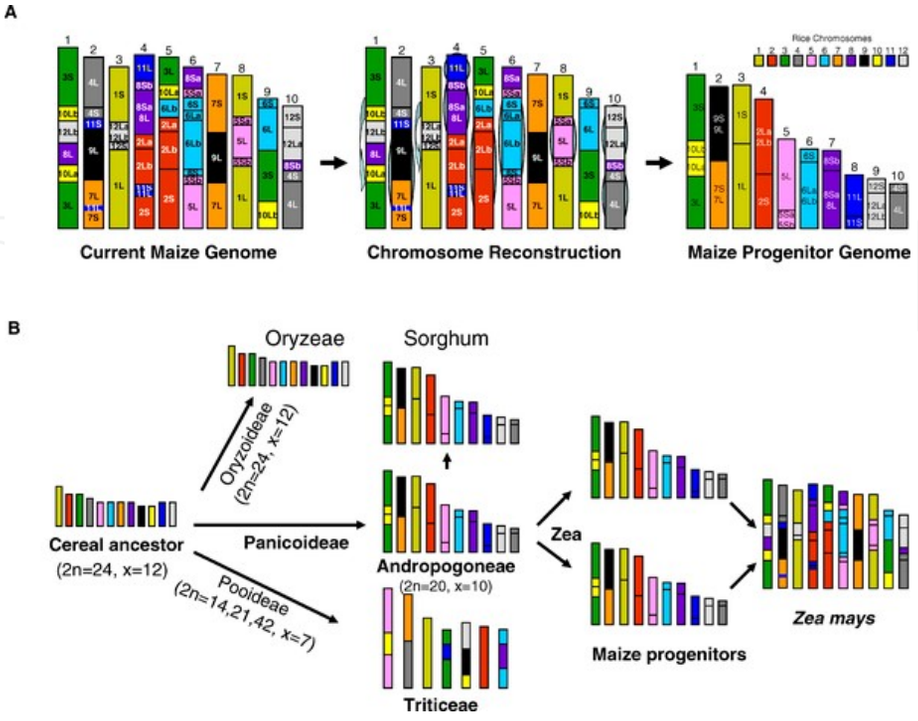
The advent of molecular markers and molecular mapping allowed researchers to conduct comparative mapping research, comparing gene orders and content of genes and markers along chromosomes of related species. The first research of large-scale restriction fragment length polymorphism (RFLP) mapping in several economically important crop genomes included the genomes of wheat, rice, maize, oat and barley. They are benchmarks for the discovery of collinearity in the grass family [44]. Hence, in the past, exploiting RFLPs to compare genomes was a valuable method as the markers made it possible to map, for the first time, a huge number of randomly distributed polymorphic loci in a single population and provided the foundation for efficient, whole-genome studies at the molecular level [48]. The application of RFLP technology in comparative genome analysis studies revealed that an extensive commonality in gene content and arrangement was a basic chromosomal property, thus prompting the idea that the genetic map could be used to tie all grasses into a single model system. This led to the construction of a consensus grass map based on 25 rice linkage blocks [37, 38]. The resolution of the genetic maps, however, proved to be very low with an average of one marker in every 5 to 10 centimorgans (cM), allowing the detection of only large rearrangements. The RFLP markers used to construct the maps were also low-copy, therefore limiting the detection of small deletions, inversions and whole or partial genome duplication events [49]. The use of RFLP markers for comparative mapping also had difficulty to assess orthologous (derived from a common ancestor by speciation) and paralogous (derived by duplication within one genome) relationships in gene families. Having these challenges associated with traditional genotyping, the NGS techniques discussed above are expected to advance comparative genomics because they provide actual DNA sequences that allow interspecies or intergeneric comparisons.

Traditional genome analyses have provided sufficient evidence that cereal genomes share conserved regions at either macro or micro levels. For example, a comparative genomics study on rice and maize indicated high levels of collinearity between the two genomes with some chromosomes or their arms—accounting for at least 67% of the two genomes—having almost similar gene order and sequences [46]. Similarly, large proportions of conserved regions between rice and wheat chromosomes were identified with major differences arising from chromosomal rearrangements [40, 45]. Conservation of about 24% of grass-specific gene orders have been reported in sorghum [30], including high collinearity with rice [50]. Thus, sorghum can also serve as a model species for cereal genomic studies due to its relatively small genome size and wide adaptability. High levels of microcollinearity have been demonstrated between chromosome 6 of rice and the telomeric regions of barley chromosome 1P, which further confirm the usefulness of mapping the small rice genome for map-based cloning of important genes in complex genomes [47]. Figures 1 and 2 illustrate the conservation of synteny and collinearity among different cereals by revealing the syntenic relationships between chromosomes of cereal crops. Furthermore, Figure 2B reveals that the 10 maize progenitor chromosomes and the 10 linkage groups of sorghum appear to be similar, thus exposing their evolutionary divergence from rice that could be their common ancestor before speciation [51]. The study of such evolutionary relationships and changes that occurred after cereals diverged from their progenitors will further be enhanced through comparative genomics integrated with NGS and next-next or third-generation sequencing techniques, which can generate more resolute physical maps. Availability of updated genome sequences will expose the multiple breaks in collinearity occurring in the genome compositions due to structural rearrangements caused by transposable elements, inversions, deletions and duplications. The macro- and microcollinearities described in this section are exposed by the observed phenotypic similarities that exist among different cereal species.



Source: Bowers, et al. [50].

Figure 1. Microsynteny conservation between sorghum and rice.



Source: Wei [51].

Figure 2. Conservation and changes in rice, maize, sorghum and wheat chromosomes during cereal speciation.

7. Phenotypic commonality in cereals

The conservation of synteny and collinearity of genes among cereals is highly attributed to the common phenotypic features or characteristics that are evidence that they share common ancestry, while their differences mainly stem from chromosomal rearrangements and polyploidization as shown in Figure 2. Their morphological similarity (Figure 3) also shows evidence that they share common ancestry. Based on phenotype alone, most also share similar rooting system, leaf venation, flowering habits, tillering, inflorescences, physiological behavior such as vernalization requirements, and adaptation to biotic and abiotic stresses. For example, some cereals are hosts of common diseases, as in the case of maize streak virus (MSV), wheat streak mosaic virus (WSMV) and rusts [52, 53], while others are nonhosts, as in the case of rice to rusts. The differences in phenotype and genome structure among all these species could be due to mutations, breaks in collinearity and loss of synteny that occurred in their genomes over millions of years. Such differences can be traced through comparative genomic analysis, particularly with the aid of high-throughput sequencing techniques. Likewise, the similarity in phenotype and genome structure could be due to sharing a common ancestry (Figures 2 and 3). This finding therefore reveals some phenotypes along with gene orders and sequences that have been conserved over millions of years.

http://dx.doi.org/10.5772/61763

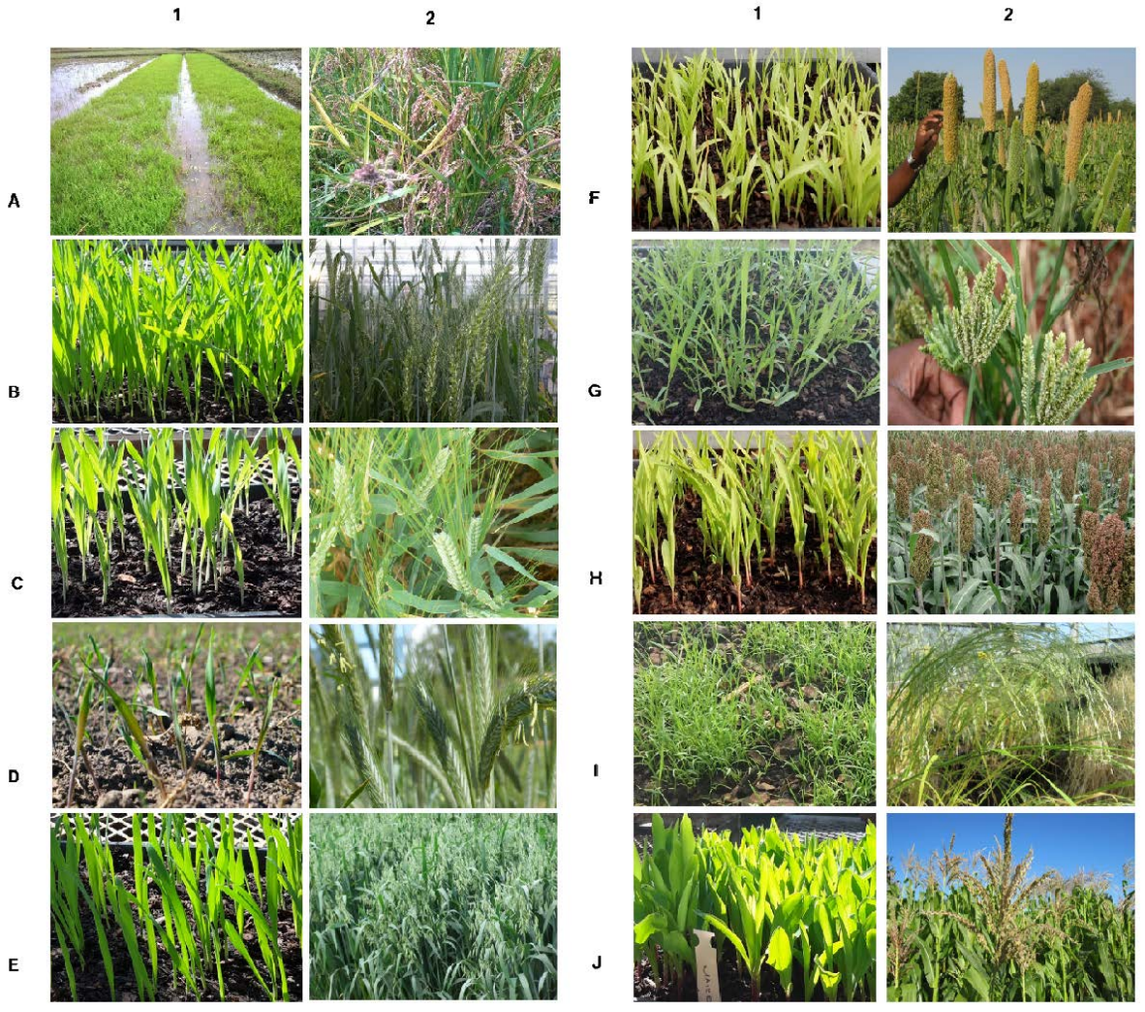


Figure 3. Phenotypic commonality in cereals. Similar seedlings: (A1) rice seedlings, (B1) wheat seedlings, (C1) barley seedlings, (D1) rye seedlings, (E1) oat seedlings, (F1) pearl millet seedlings, (G1) finger millet seedlings, (H1) sorghum seedlings, (I1) tef seedlings, (J1) maize seedlings. Similar heads: (A2) rice heads, (B2) wheat heads, (C2) barley heads, (D2) rye heads, (E2) oat heads, (F2) pearl millet heads, (G2) finger millet heads, (H2) sorghum heads, (I2) tef heads, (J2) maize heads.

8. Outlook

Plant species have highly conserved regions at DNA sequence level, whereas the bulk of the large genomes consist of repetitive DNA sequences, most of which are species-specific. Comparative genomics have opened new avenues for map-based positional cloning of genes encoding important traits on large and intricate genomes through investigating small and less complex genomes. In grasses, rice and *Brachypodium* have been identified as model species for such research since they have small and stable genomes. This, however, requires the integration of NGS techniques so that all the conserved and nonconserved regions can be fully sequenced and annotated with the aid of other "omic" technologies. Hence, the future of

comparative genomics studies in cereals will largely rely on cost-effective sequencing technologies along with computational systems that handle large numbers of sequences, thus allowing effective sequence comparisons across species of interest. The substantial evidence regarding a common ancestry of cereals—based on genome and morphological structures—led to the successful use of the genome sequence of one species to share a light on the function of that sequence in other related species. A wide adoption of this approach across different cereals will speed up gains and generate useful databases and datasets for effective cereal breeding. Furthermore, researchers will be able to use other widely adapted cereals like sorghum and some of the under-researched cereals as models for sequencing genes and alleles responsible for unique traits such as wide adaptation to stress-prone environments due to increased sequencing throughput. There is, however, a need to invest in advanced computational and bioinformatics tools to handle and analyze huge datasets that will be generated through these technology advances.

Acknowledgements

The authors would like to thank Damien Shumbusha (Rwanda Agriculture Board), Andre Malan (Agricultural Research Council - Small Grain Institute), Hussein Shimelis (African Centre for Crop Improvement, University of KwaZulu-Natal), Caleb Souta (Seed Co. Limited, Zimbabwe), Lydia Ndinelao Horn (Ministry of Agriculture, Water and Forestry, Namibia), Cousin Musvosvi (Department of Crop Science, University of Zimbabwe) and Kingstone Mashingaidze (Agricultural research Council –Grain Crops Institute) for providing seeds or images used to illustrate phenotypic commonalities in cereals.

Author details

Thandeka N. Sikhakhane[1,2], Sandiswa Figlan[1], Learnmore Mwadzingeni[1], Rodomiro Ortiz[3] and Toi J. Tsilo[1,2,4*]

*Address all correspondence to: tsilot@arc.agric.za

1 Agricultural Research Council, Small Grain Institute, Bethlehem, South Africa

2 Department of Life and Consumer Sciences, University of South Africa, Pretoria, South Africa

3 Department of Plant Breeding, Swedish University of Agricultural Sciences, Alnarp, Sweden

4 Department of Plant Production, University of Venda, Thohoyandou, South Africa

References

[1] Ray DK, et al. Yield trends are insufficient to double global crop production by 2050. 2013. DOI: 10.1371/journal.pone.0066428

[2] Lobell DB and CB Field. Global scale climate–crop yield relationships and the impacts of recent warming. Environmental research letters, 2007. 2(1): p. 014002. DOI: 10.1088/1748-9326/2/1/014002

[3] Rosenzweig C, et al. Climate change and extreme weather events; implications for food production, plant diseases, and pests. Global change & human health, 2001. 2(2): p. 90-104.

[4] Waddington SR, et al. Getting the focus right: production constraints for six major food crops in Asian and African farming systems. Food security, 2010. 2(1): p. 27-48. DOI: 10.1007/s12571-010-0053-8

[5] Chakraborty S and AC Newton. Climate change, plant diseases and food security: an overview. Plant Pathology, 2011. 60(1): p. 2-14. DOI: 10.1111/j.1365-3059.2010.02411.x

[6] Duveiller E, RP Singh, and JM Nicol. The challenges of maintaining wheat productivity: pests, diseases, and potential epidemics. Euphytica, 2007. 157(3): p. 417-430. DOI: 10.1007/s10681-007-9380-z

[7] Kole C, et al. Application of genomics-assisted breeding for generation of climate resilient crops: progress and prospects. Frontiers in Plant Science, 2015. 6. DOI: 10.3389/fpls.2015.00563

[8] Paux E, et al. Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B. The Plant Journal, 2006. 48(3): p. 463-474. DOI: 10.1111/j.1365-313X.2006.02891.x

[9] Dahm R. Friedrich Miescher and the discovery of DNA. Developmental biology, 2005. 278(2): p. 274-288. DOI: 10.1016/j.ydbio.2004.11.028

[10] Cogoni C and G Macino. Post-transcriptional gene silencing across kingdoms. Current opinion in genetics & development, 2000. 10(6): p. 638-643. DOI: 10.1016/S0959-437X(00)00134-9

[11] Depicker A and M Van Montagu. Post-transcriptional gene silencing in plants. Current opinion in cell biology, 1997. 9(3): p. 373-382. DOI: 0955067400900373

[12] Saze H, et al. DNA methylation in plants: relationship to small RNAs and histone modifications, and functions in transposon inactivation. Plant and Cell Physiology, 2012. 53(5): p. 766-784. DOI: 10.1093/pcp/pcs008

[13] Alonso-Blanco C and B Méndez-Vigo. Genetic architecture of naturally occurring quantitative traits in plants: an updated synthesis. Current opinion in plant biology, 2014. 18: p. 37-43. DOI: 10.1016/j.pbi.2014.01.002

[14] Edwards D, J Batley, and RJ Snowdon. Accessing complex crop genomes with next-generation sequencing. Theoretical and Applied Genetics, 2013. 126(1): p. 1-11. DOI: 10.1007/s00122-012-1964-x

[15] Taft RJ, et al. Non-coding RNAs: regulators of disease. The Journal of pathology, 2010. 220(2): p. 126-139. DOI: 10.1002/path.2638

[16] Sanger F, S Nicklen, and AR Coulson. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences, 1977. 74(12): p. 5463-5467.

[17] Roche. GS FLX+ System. Sanger-like read lengths - the power of next-gen throughput. Roche Diagnostics GmbH; Available from: http://454.com/downloads/GSFLXApplicationFlyer_FINALv2.pdf [Accessed: 2015-09-05].

[18] Illumina. Next-Generation Sequencing 2015; Available from: http://www.illumina.com/ [Accessed: 2015-09-05].

[19] TFS. SOLiD® Next-Generation Sequencing 2015; Available from: https://www.lifetechnologies.com/za/en/home/life-science/sequencing/next-generation-sequencing/solid-next-generation-sequencing.html [Accessed: 2015-09-05].

[20] SeqLL. Helicos Genetic Analysis Aool 2014; Available from: http://seqll.com/ [Accessed: 210-09-05].

[21] TFS. Ion Personal Genome Machine® (PGM™) System 2015; Available from: http://www.lifetechnologies.com/order/catalog/product/4462921 [Accessed: 2015-09-05].

[22] Mwadzingeni L, et al. Breeding wheat for drought tolerance: Progress and technologies. Journal of Integrative Agriculture, 2015. DOI: 10.1016/s2095-3119(15)61102-9

[23] Harismendy O, et al. Evaluation of next generation sequencing platforms for population targeted sequencing studies. Genome biol, 2009. 10(3): p. R32.

[24] IRGSP. The map-based sequence of the rice genome. Nature, 2005. 436(7052): p. 793-800. DOI: 10.1038/nature03895

[25] Kawahara Y, et al. Improvement of the Oryza sativa Nipponbare reference genome using next generation sequence and optical map data. Rice, 2013. 6(1): p. 4.

[26] Brenchley R, et al. Analysis of the bread wheat genome using whole-genome shotgun sequencing. Nature, 2012. 491(7426): p. 705-710. DOI: 10.1038/nature11650

[27] Schnable PS, et al. The B73 maize genome: complexity, diversity, and dynamics. science, 2009. 326(5956): p. 1112-1115. DOI: 10.1126/science.1178534

[28] Tenaillon MI, et al. Genome size and transposable element content as determined by high-throughput sequencing in maize and Zea luxurians. Genome biology and evolution, 2011. 3: p. 219-229. DOI: 10.1093/gbe/evr008

[29] IBGSC. A physical, genetic and functional sequence assembly of the barley genome. Nature, 2012. 491(7426): p. 711-716. DOI: 10.1038/nature11543

[30] Paterson AH, et al. The Sorghum bicolor genome and the diversification of grasses. Nature, 2009. 457(7229): p. 551-556. DOI: doi:10.1038/nature07723

[31] Mysore KS and V Baird. Nuclear DNA content in species of Eleusine (Gramineae): A critical re-evaluation using laser flow cytometry. Plant Systematics and Evolution, 1997. 207(1-2): p. 1-11. DOI: 10.1007/BF00985206

[32] Martel E, et al. Genome size variation and basic chromosome number in pearl millet and fourteen related Pennisetum species. Journal of Heredity, 1997. 88(2): p. 139-143.

[33] Ayele M, et al. Flow cytometric analysis of nuclear genome of the Ethiopian cereal tef [Eragrostis tef (Zucc.) Trotter]. Genetica, 1996. 98(2): p. 211-215.

[34] AOCC. Improving nutrition through African Orphan Crops. 2013; Available from: http://www.mars.com/global/african-orphan-crops.aspx [Accessed: 2015-09-05].

[35] Bennetzen JL and M Chen. Grass genomic synteny illuminates plant genome function and evolution. Rice, 2008. 1(2): p. 109-118.

[36] Schuit F. Comparative genomics: beyond the horizon of the next research grant. Diabetologia, 2015: p. 1-5. DOI: 10.1007/s00125-015-3620-x

[37] Devos KM and MD Gale. Genome relationships: the grass model in current research. The Plant Cell, 2000. 12(5): p. 637-646.

[38] Feuillet C and B Keller. Comparative genomics in the grass family: molecular characterization of grass genome structure and evolution. Annals of botany, 2002. 89(1): p. 3-10. DOI: 10.1093/aob.2002.mcafv0a0il8a

[39] Eckardt NA. Everything in Its Place: Conservation of Gene Order among Distantly Related Plant Species. The Plant Cell, 2001. 13(4): p. 723-725. DOI: 10.2307/3871334

[40] Sorrells ME, et al. Comparative DNA sequence analysis of wheat and rice genomes. Genome research, 2003. 13(8): p. 1818-1827.

[41] Goff SA, et al. A draft sequence of the rice genome (Oryza sativa L. ssp. japonica). Science, 2002. 296(5565): p. 92-100.

[42] Vogel JP, et al. Genome sequencing and analysis of the model grass Brachypodium distachyon. Nature, 2010. 463(7282): p. 763-768. DOI: 10.1038/nature08747

[43] Huo N, et al. The nuclear genome of Brachypodium distachyon: analysis of BAC end sequences. Functional & integrative genomics, 2008. 8(2): p. 135-147. DOI: 10.1007/s10142-007-0062-7

[44] Smilde WD, et al. New evidence for the synteny of rice chromosome 1 and barley chromosome 3H from rice expressed sequence tags. Genome, 2001. 44(3): p. 361-367. DOI: 10.1139/gen-44-3-361

[45] Ahn S, et al. Homoeologous relationships of rice, wheat and maize chromosomes. Molecular and General Genetics MGG, 1993. 241(5-6): p. 483-490.

[46] Ahn S and SD Tanksley. Comparative linkage maps of the rice and maize genomes. Proceedings of the National Academy of Sciences, 1993. 90(17): p. 7980-7984. http://www.pnas.org/content/90/17/7980.abstract

[47] Kilian A, et al. Rice-barley synteny and its application to saturation mapping of the barley Rpg1 region. Nucleic acids research, 1995. 23(14): p. 2729-2733. DOI: 10.1093/nar/23.14.2729

[48] McCouch SR. Genomics and synteny. Plant Physiology, 2001. 125(1): p. 152-155.

[49] Paterson AH, et al. Comparative genomics of plant chromosomes. The Plant Cell, 2000. 12(9): p. 1523-1539.

[50] Bowers JE, et al. Comparative physical mapping links conservation of microsynteny to chromosome structure and recombination in grasses. Proceedings of the National Academy of Sciences of the United States of America, 2005. 102(37): p. 13206-13211. DOI: 10.1073pnas.0502365102

[51] Wei F, et al. Physical and genetic structure of the maize genome reflects its complex evolutionary history. PLoS Genet, 2007. 3(7): p. e123. DOI: 10.1371/journal.pgen.0030123

[52] Cummins GB, The rust fungi of cereals, grasses and bamboos. 1971: Berlin, Springer-Verlag.

[53] Brakke M. Wheat streak mosaic virus. CMI/AAB descriptions of plant viruses, 1971. 48: p. 1-4.

IntechOpen