Chapter 1
Cereal Genomics: Excitements, Challenges and Opportunities

Pushpendra K. Gupta and Rajeev K. Varshney

1.1 Introduction

Cereals constitute the most important food crops of the world, occupying ~680 million hectares of land and producing ~2,295 million tonnes of food grain globally (October 4, 2012; http://www.fao.org/worldfoodsituation/wfs-home/csdb/en/), even though this production is lower than that for the year 2011 (2,340 million tonnes). Cereals are also an excellent source of fat-soluble vitamin E (an essential antioxidant) and contain 20–30 % of our daily mineral requirement (including selenium, calcium, zinc, and copper). Among all crops, cereals also provide 60 % of calories and proteins for the growing human population, which is estimated to reach 9.2 billion level in the year 2050. The projected need for annual cereal production in 2050 is ~3,000 million tonnes, so that at least a 30 % increase in the annual cereal grain production would be needed to meet this demand; this translates into an annual growth rate of ~0.70 %, which should not be difficult to achieve if the present growth rate of ~1.0 % is maintained. It has also been noticed that although during the last 4–5 decades the annual production of cereals (sum of wheat, milled rice and coarse grains) has been steadily increasing, the rate of growth in this production has shown a fatigue, with the growth rate falling from 3.7 % p.a. in the 1960s, to 2.5 % in 1970s, 1.4 % in 1980s and 1.1 % in 1990s, this growth rate sometimes also being negative (years 2006–2007, 2010–2011; see Fig. 1.1). The rate of growth in

---

P. K. Gupta (✉)
Department of Genetics and Plant Breeding, CCS University, Meerut 250004, India
e-mail: pkgupta36@gmail.com

R. K. Varshney (✉)
Center of Excellence in Genomics (CEG), International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502324, India
e-mail: r.k.varshney@cgiar.org

R. K. Varshney
CGIAR Generation Challenge Programme (GCP), c/o CIMMYT, 06600 Mexico, DF, Mexico

P. K. Gupta and R. K. Varshney (eds.), *Cereal Genomics II*,
DOI: 10.1007/978-94-007-6401-9_1,

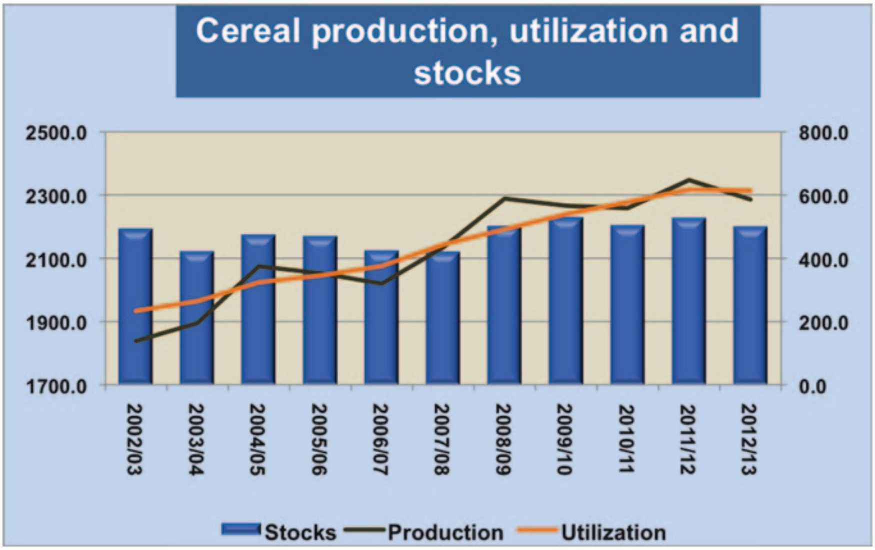

Fig. 1.1 Production, utilization and stocks of cereal grains globally (*source* http://www.fao.org/worldfoodsituation/wfs-home/csdb/en/)

yield (production per unit area) has also slowed down in recent years, so that there is a need to improve further the yield per unit area also. While total annual production of leading cereal crops such as rice, maize, and wheat has increased, the annual production of other cereal crops like barley, oats and rye has declined (Table 1.1). Since the productivity and production of cereals are not stable and may even decrease, depending on weather condition, and because the demand will increase in future due to population pressure, cereal workers can not be complacent and will need to keep on working for enhancing both, yield and production. It is anticipated that this may not be possible with the conventional plant breeding, and genomics-based technologies will have to supplement to meet this challenge. With the shrinking land area, water shortage and the projected climate change, the task is certainly not going to be easy.

The volume '*Cereal Genomics*' that was edited by the authors and published in 2004 contained useful articles written by eminent scientists in different areas of cereal genomics research. It served a useful purpose of making available all information on cereal genomics at one place (Gupta and Varshney 2004), and was well received by cereal workers globally. However, this volume has become out-of-date, since during the last eight years, cereal genomics research progressed at unprecedented pace (Table 1.1). Sufficient additional information has become available making it necessary to have another fresh look on the present status and future possibilities of cereal genomics research. For instance, whole genome sequences became available not only for rice (Goff et al. 2002; Yu et al. 2002), maize (Schnable et al. 2009) and sorghum (Paterson et al. 2009), the three major cereal crops, but also for *Brachypodium distichum* (TIBI 2010), a newly identified model grass species. Significant progress has also been made in sequencing

Table 1.1 A comparative overview of the crops discussed in cereal genomics volume I and II

Cereal species	Biological name	Chromsome number	Genome size (Mbp)[a]	Status year[b]	Yield (tonnes/hectare)[c]	Production (million tonnes)[c]	Number of ESTs available in public domain[d]	Availability of resources
Barley	*Hordeum vulgare*	$2n = 2x = 14$	5,000	2012	2.6	124	5,01,838	Extensive genetic and physical maps; first glimpse of genome sequence is now available (the International Barley Genome Sequencing Consortium (IBGSC) 2012)
				2004	2.5	132	3,56,848	Both genetic and physical maps (based on translocation breakpoints) were available
Maize	*Zea mays*	$2n = 2x = 20$	2,500	2012	5.2	840	20,19,137	Genome sequencing and re-sequencing for several hundred accessions have become available, hapmap available; GWAS analysis undertaken; GS in progress
				2004	4.3	603	3,93,719	Extensive genetic (including transcript map) and BAC-based physical maps were available
Oats	*Avena sativa*	$2n = 6x = 42$	11,400	2012	2.2	20	25,344	Saturated genetic maps have become available
				2004	1.9	27	574	Genetic maps (but not saturated) were available
Rice	*Oryza sativa*	$2n = 2x = 24$	430	2012	4.4	696	12,52,989	Several drafts for whole genome sequence available, resequencing of several thousand accessions has been completed/in progress, GWAS undertaken; GS in progress
				2004	3.9	580	2,83,935	Extensive genetic (including transcript) and YAC/BAC-based physical maps, as well as 4 drafts of complete genome sequences available
Rye	*Secale cereale*	$2n = 2x = 14$	8,400	2012	2.3	12	9,298	Extensive genetic and physical maps available
				2004	2.2	21	9,194	Genetic maps (but not saturated) available

(continued)

Table 1.1 (continued)

Cereal species	Biological name	Chromsome number	Genome size (Mbp)[a]	Status year[b]	Yield (tonnes/ hectare)[c]	Production (million tonnes)[c]	Number of ESTs available in public domain[d]	Availability of resources
Sorghum	*Sorghum bicolor*	$2n = 2x = 20$	750	2012	1.4	56	2,09,835	Genome sequence has become available and re-sequencing of several accessions initiated
				2004	1.3	55	1,61,813	Integrated cytogenetic, genetic and physical maps were available
Wheat	*Triticum aestivum*	$2n = 6x = 42$	16,000	2012	3.0	654	12,86,173	Extensive genetic and physical map available, Gene space of individual chromosomes captured and genome sequencing in process; GWAS and GS conducted
				2004	2.7	568	5,49,926	Extensive genetic as well as deletion lines-based physical maps were available

[a] As per Bennett and Leitch (1995)

[b] Comparative status of productivity, production and genomic resources between 2004 and 2012 has been shown

[c] As per FAO wesbsite http://apps.fao.org; accessed in 2004 and 2012

[d] As per dbEST release 030504 and 181012- http://www.ncbi.nlm.nih.gov/dbEST/dbEST summary.html

genome or gene space in wheat (Paux et al. 2008; Berkman et al. 2011) and barley (Mayer et al. 2011; IBGSC 2012).

During recent years, another major technology development is the availability of next generation sequencing (NGS), which included the second and third generation high throughput and cost-effective sequencing systems (Thudi et al. 2012). These NGS platforms revolutionized genomics research not only in cereals, but in all living systems including humans, and other higher animals/plants and the microorganisms. By using NGS technologies, genomes of hundreds or thousands of accessions of an individual crop like rice have been generated, thus providing estimates of genome wide diversity and making genome wide association studies (GWAS) more meaningful (Huang et al. 2010; Tian et al. 2011; Zhao et al. 2011; Chia et al. 2012; Hufford et al. 2012). All these developments also created a demand for computational tools to analyse the massive data that was generated at an unprecedented pace. This challenge was met successfully by parallel growth in the field of bioinformatics. These developments have been briefly described in this volume, which is appropriately titled as "*Cereal Genomics II*", so that it supplements our earlier edited volume "*Cereal Genomics*" (Gupta and Varshney 2004). The different aspects covered in this volume are briefly summarized in this introductory chapter.

1.2 Molecular Markers in Cereal Genomics

Although cereal genomics had its birth during 1980s with the development and use of restriction fragment length polymorphism (RFLP) markers, it gained momentum with the development and use of simple sequence repeat (SSR) and amplified fragment length polymorphism (AFLP) markers during 1990s and single nucleotide polymorphism (SNP) and diversity array technology (DArT) markers during the first decade of the present century. However, wide-spread use of these markers in crop breeding programs was not possible due to low throughput and expensive genotyping involved in using these markers. With the availability of microarray technology, during 1990s and early years of the present century, an increased use of microarray-based marker genotyping (particularly for SNPs) was witnessed (Gupta et al. 2008). An updated account of these array-based markers for cereal genomics research is presented by Pushpendra K Gupta (CCS University, Meerut, India) and his former students by (Sachin Rustgi and Reyaz Mir) in Chap. 2 of this volume.

In parallel and following the development of microarray technology, another major development has been the availability of a number of NGS platforms (as mentioned above), which facilitated development and use of high throughput and cost-effective markers like SNPs, SSRs, Insertion Site-Based Polymorphism (ISBPs), Restriction-site Associated DNAs (RADs), Copy-Number Variations (CNVs)/Presence-Absence Variations (PAVs), etc. Dave Edward of the University of Queensland, Australia and Pushpendra Gupta of CCS University, Meerut, India discussed these aspects in Chap. 3 of this volume. The different NGS technologies and their use for study of genetic diversity in cereals are discussed in Chap. 4 by

Eduard Akhunov and colleagues from Kansas State University, USA. This chapter partly overlaps the contents of Chap. 3 in dealing with markers developed using NGS, this overlap being unavoidable in an edited volume.

1.3 Organization and Evolution of Cereal Genomes

In true sense, plant genomics research had its beginning in December 2000, with the publication of the whole genome sequence of the model plant species *Arabidopsis thaliana* (AGI 2000). This was followed by publication of the whole genome sequences of more than a dozen plant species, which included some cereals such as rice (Yu et al. 2002; Goff et al. 2002; IRGSP 2005), maize (Schnable et al. 2009), sorghum (Paterson et al. 2009) and foxtail millet (Bennetzen et al. 2012; Zhang et al. 2012) and a model grass species, *Brachypodium distichum* (TIBI 2010). Available genome sequences of several other plant species (http://genomevolution.org/wiki/index.php/Sequenced_plant_genomes; Plant GDB), which also became available in parallel, were also compared with available cereal genome sequences, thus facilitating further progress in cereal genomics research. In Chap. 5 of this volume, Xi-Yin Wang and Andrew H. Paterson from University of Georgia, USA utilize this information and discuss comparative genomics in cereals. During the study of genomic sequences of cereals, in particular those of corn, it has been recognized that transposable elements constitute a major part of cereal genomes. In Chap. 6 of this volume, Beat Keller and his coworkers from University of Zurich, Switzerland, have discussed the role of transposable elements in shaping cereal genomes.

1.4 Functional Genomics of Cereals

Cereal have also been subjected to functional genomics research, which during the last two decades covered both basic and applied aspects. As a result, not only we understand better the genomes of major cereals and the mechanisms involved in the function of different cereal genes, but we have also utilized information generated from genomics research in producing better transgenic crops, which will give higher yields, sometimes with value addition. In Chap. 7 of this volume, Pankaj Jaiswal and colleagues from Oregon State University, USA have discussed the techniques and bioinformatics involved in functional annotation of cereal genomes. In Chap. 8, Nese Sreenivasulu and his coworkers from IPK, Gatersleben, Germany discussed the different ‘omics’ approaches involved in functional genomics and their implications for developing a system biology approach for study of the mechanism involved in tolerance against abiotic stress. In Chap. 9, Peter Langridge and his coworkers from Australian Centre for Plant Functional Genomics (ACPFG), Australia discuss the functional genomics of seed development and in Chap. 10, Bikram Gill and his coworkers from Kansas State University, USA discuss the genomics of cereal based functional foods.

1.5 QTL Analysis, Domestication and Molecular Breeding

Another important development in cereal genetics and genomics during the last two decades is the availability of approaches for genetic dissection of complex quantitative traits. This became possible due to the availability of DNA-based molecular markers and statistical tools for analysis of complex traits. These aspects have been discussed in Chap. 11 by Pushpendra K Gupta and his two former students (Pawan Kulwal and Reyaz Mir). QTL analysis also facilitated the study of domestication process, so that domestication syndromes involving the selection of a set of genes have been discovered in all major cereals. These aspects have been discussed in Chap. 12 by Tao Sang from Key Laboratory of Plant Resources, China and Beijing Botanical Garden, China and Jiayang Li from National Center for Plant Gene Research, China. It has also been recognized that precision in phenotyping has been a limitation in studying the genetic architecture of cereal crops, and is absolutely necessary in order to improve the power and resolution of genetic approaches available for genetic dissection of complex traits. This has led to the development of a new discipline called phenomics, which is gaining momentum, so that phenotyping platforms are being established in several countries to facilitate precision in phenotyping. BM Prasanna and his coworkers from International Maize and Wheat Improvement Center (CIMMYT), Mexico discussed the subject of high throughput precision phenotyping for cereal breeding in Chap. 13 of this volume. Chapter 14 of this volume is devoted to molecular breeding written by Yunbi Xu from CIMMYT and his co workers from Chinese Academy of Agricultural Sciences, China and CIMMYT.

1.6 Summary and Outlook

In summary, this volume *Cereal Genomics II* with 14 chapters (including this introductory chapter) provides a glimpse of the advances in cereals genomics research made during the last eight years, that elapsed between now and the year 2004, when our earlier volume, *Cereal Genomics* was published. This volume presents state-of-art of cereal genomics and its utilization in both basic studies such as comparative genomics and functional genomics as well as applied aspects like QTL mapping and molecular breeding. Keeping in view the information that became available during the last one decade, one can certainly foresee an exciting period that lies ahead for cereal researchers globally, particularly because the large and complex cereal genomes of barley and wheat will also be fully sequenced within the next 2–3 years. Molecular mapping and breeding approaches will be shifting from marker-based genotyping to sequencing-based genotyping. Comparative genomics will be moving from comparison of genomes of two or more species to comparison of genomes of hundreds to thousands accessions of the same species. While data generation for even complex genomes of cereal species is expected to become routine, analysis and interpretation of data will be a challenge for both cereal biologists as well as for those involved in applied cereal genomics research. This will be facilitated through advances in bioinformatics including high-throughput data analysis and cloud computing, which will help make further advances in this fascinating area of cereal genomics.

References

Arabidopsis Genome Initiative (AGI) (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815

Bennett MD, Leitch (1995) Nuclear DNA amounts in angiosperms. Ann Bot 76:113–176

Bennetzen JL, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli AC, Estep M, Feng L, Vaughn JN, Grimwood J, Jenkins J, Barry K, Lindquist E, Hellsten U, Deshpande S, Wang X, Wu X, Mitros T, Triplett J, Yang X, Ye CY, Mauro-Herrera M, Wang L, Li P, Sharma M, Sharma R, Ronald PC, Panaud O, Kellogg EA, Brutnell TP, Doust AN, Tuskan GA, Rokhsar D, Devos KM (2012) Reference genome sequence of the model plant *Setaria*. Nat Biotechnol 30:555–561

Berkman BJ, Skarshewski A, Lorenc MT, Lai K, Duran C, Ling EYS, Stiller J, Smits L, Imelfort M, Manoli S, McKenzie M, Kubalakova M, Simkova H, Batley J, Fleury D, Dolezel J, Edwards D (2011) Sequencing and assembly of low copy and genic regions of isolated *Triticum aestivum* chromosome arm 7DS. Plant Biotechnol J 9:768–775

Chia JM, Song C, Bradbury PJ, Costich D, de Leon N, Doebley J, Elshire RJ, Gaut B, Geller L, Glaubitz JC, Gore M, Guill KE, Holland J, Hufford MB, Lai J, Li M, Liu X, Lu Y, McCombie R, Nelson R, Poland J, Prasanna BM, Pyhäjärvi T, Rong T, Sekhon RS, Sun Q, Tenaillon MI, Tian F, Wang J, Xu X, Zhang Z, Kaeppler SM, Ross-Ibarra J, McMullen MD, Buckler ES, Zhang G, Xu Y, Ware D (2012) Maize hapmap2 identifies extant variation from a genome in flux. Nat Genet 44:803–807

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun WL, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalma T, Oliphant A, Briggs S (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica). Science 296:92–100

Gupta PK, Varshney RK (2004) *Cereal Genomics*. Kluwer Academic Publishers, The Netherlands, pp 1–369

Gupta PK, Rustgi S, Mir RR (2008) Array-based high-throughput DNA markers for crop improvement. Heredity 101:5–18

Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y, Li C, Zhu C, Lu T, Zhang Z, Li M, Fan D, Guo Y, Wang A, Wang L, Deng L, Li W, Lu Y, Weng Q, Liu K, Huang T, Zhou T, Jing Y, Li W, Lin Z, Buckler ES, Qian Q, Zhang QF, Li J, Han B (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. Nat Genet 42:961–967

Hufford MB, Xun X, van Heerwaarden J, Pyhäjärvi T, Chia JM, Cartwright RA, Elshire RJ, Glaubitz JC, Guill KE, Kaeppler SM, Lai J, Morrell PL, Shannon LM, Song C, Springer NM, Swanson-Wagner RA, Tiffin P, Wang J, Zhang G, Doebley J, McMullen MD, Ware D, Buckler ES, Yang S, Ross-Ibarra J (2012) Comparative population genomics of maize domestication and improvement. Nat Genet 44:808–811

International Rice Genome Sequencing Project (IRGSP) (2005) The map-based sequence of the rice genome. Nature 436:793–800

Mayer KF, Martis M, Hedley PE, Simková H, Liu H, Morris JA, Steuernagel B, Taudien S, Roessner S, Gundlach H, Kubaláková M, Suchánková P, Murat F, Felder M, Nussbaumer T, Graner A, Salse J, Endo T, Sakai H, Tanaka T, Itoh T, Sato K, Platzer M, Matsumoto T, Scholz U, Dolezel J, Waugh R, Stein N (2011) Unlocking the barley genome by chromosomal and comparative genomics. Plant Cell 23:1249–1263

Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, Schmutz J, Spannagl M, Tang H, Wang X, Wicker T, Bharti AK, Chapman J, Feltus FA, Gowik U, Grigoriev IV, Lyons E, Maher CA, Martis M, Narechania A, Otillar RP, Penning BW, Salamov AA, Wang Y, Zhang L, Carpita NC, Freeling M, Gingle AR, Hash CT, Keller B, Klein P, Kresovich S, McCann MC, Ming R, Peterson DG, Mehboob-ur-Rahman Ware D, Westhoff P, Mayer KF, Messing J, Rokhsar DS (2009) The *Sorghum bicolor* genome and the diversification of grasses. Nature 457:551–556

Paux E, Sourdille P, Salse J, Saintenac C, Choulet F, Leroy P, Korol A, Michalak M, Kianian S, Spielmeyer W, Lagudah E, Somers D, Kilian A, Alaux M, Vautrin S, Bergès H, Eversole K, Appels R, Safar J, Simkova H, Dolezel J, Bernard M, Feuillet C (2008) A physical map of the 1-gigabase bread wheat chromosome 3B. Science 322:101–104

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, Minx P, Reily AD, Courtney L, Kruchowski SS, Tomlinson C, Strong C, Delehaunty K, Fronick C, Courtney B, Rock SM, Belter E, Du F, Kim K, Abbott RM, Cotton M, Levy A, Marchetto P, Ochoa K, Jackson SM, Gillam B, Chen W, Yan L, Higginbotham J, Cardenas M, Waligorski J, Applebaum E, Phelps L, Falcone J, Kanchi K, Thane T, Scimone A, Thane N, Henke J, Wang T, Ruppert J, Shah N, Rotter K, Hodges J, Ingenthron E, Cordes M, Kohlberg S, Sgro J, Delgado B, Mead K, Chinwalla A, Leonard S, Crouse K, Collura K, Kudrna D, Currie J, He R, Angelova A, Rajasekar S, Mueller T, Lomeli R, Scara G, Ko A, Delaney K, Wissotski M, Lopez G, Campos D, Braidotti M, Ashley E, Golser W, Kim H, Lee S, Lin J, Dujmic Z, Kim W, Talag J, Zuccolo A, Fan C, Sebastian A, Kramer M, Spiegel L, Nascimento L, Zutavern T, Miller B, Ambroise C, Muller S, Spooner W, Narechania A, Ren L, Wei S, Kumari S, Faga B, Levy MJ, McMahan L, Van Buren P, Vaughn MW, Ying K, Yeh CT, Emrich SJ, Jia Y, Kalyanaraman A, Hsia AP, Barbazuk WB, Baucom RS, Brutnell TP, Carpita NC, Chaparro C, Chia JM, Deragon JM, Estill JC, Fu Y, Jeddeloh JA, Han Y, Lee H, Li P, Lisch DR, Liu S, Liu Z, Nagel DH, McCann MC, SanMiguel P, Myers AM, Nettleton D, Nguyen J, Penning BW, Ponnala L, Schneider KL, Schwartz DC, Sharma A, Soderlund C, Springer NM, Sun Q, Wang H, Waterman M, Westerman R, Wolfgruber TK, Yang L, Yu Y, Zhang L, Zhou S, Zhu Q, Bennetzen JL, Dawe RK, Jiang J, Jiang N, Presting GG, Wessler SR, Aluru S, Martienssen RA, Clifton SW, McCombie WR, Wing RA, Wilson RK (2009) The B73 maize genome: complexity, diversity, and dynamics. Science 326:1112–1115

The International Barley Genome Sequencing Consortium (IBGSC) (2012) A physical, genetic and functional sequence assembly of the barley genome. Nature doi:10.1038/nature11543

The International Brachypodium Initiative (TIBI) (2010) Genome sequencing and analysis of the model grass *Brachypodium distachyon*. Nature 463:763–768

Thudi M, Li Y, Jackson SA, May GD, Varshney RK (2012) Current state-of-art of sequencing technologies for plant genomics research. Brief Funct Genomics 11:3–11

Tian F, Bradbury PJ, Brown PJ, Hung H, Sun Q, Flint-Garcia S, Rocheford TR, McMullen MD, Holland JB, Buckler ES (2011) Genome-wide association study of leaf architecture in the maize nested association mapping population. Nat Genet 43:159–162

Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, Cao M, Liu J, Sun J, Tang J, Chen Y, Huang X, Lin W, Ye C, Tong W, Cong L, Geng J, Han Y, Li L, Li W, Hu G, Huang X, Li W, Li J, Liu Z, Li L, Liu J, Qi Q, Liu J, Li L, Li T, Wang X, Lu H, Wu T, Zhu M, Ni P, Han H, Dong W, Ren X, Feng X, Cui P, Li X, Wang H, Xu X, Zhai W, Xu Z, Zhang J, He S, Zhang J, Xu J, Zhang K, Zheng X, Dong J, Zeng W, Tao L, Ye J, Tan J, Ren X, Chen X, He J, Liu D, Tian W, Tian C, Xia H, Bao Q, Li G, Gao H, Cao T, Wang J, Zhao W, Li P, Chen W, Wang X, Zhang Y, Hu J, Wang J, Liu S, Yang J, Zhang G, Xiong Y, Li Z, Mao L, Zhou C, Zhu Z, Chen R, Hao B, Zheng W, Chen S, Guo W, Li G, Liu S, Tao M, Wang J, Zhu L, Yuan L, Yang H (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. indica). Science 296:79–92

Zhang G, Liu X, Quan Z, Cheng S, Xu X, Pan S, Xie M, Zeng P, Yue Z, Wang W, Tao Y, Bian C, Han C, Xia Q, Peng X, Cao R, Yang X, Zhan D, Hu J, Zhang Y, Li H, Li H, Li N, Wang J, Wang C, Wang R, Guo T, Cai Y, Liu C, Xiang H, Shi Q, Huang P, Chen Q, Li Y, Wang J, Zhao Z, Wang J (2012) Genome sequence of foxtail millet (*Setaria italica*) provides insights into grass evolution and biofuel potential. Nat Biotechnol 30:549–554

Zhao K, Tung CW, Eizenga GC, Wright MH, Ali ML, Price AH, Norton GJ, Islam MR, Reynolds A, Mezey J, McClung AM, Bustamante CD, McCouch SR (2011) Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. Nat Commun 2:467