Review

Integrated genomics and molecular breeding approaches for dissecting the complex quantitative traits in crop plants

ALICE KUJUR, MANEESHA S SAXENA, DEEPAK BAJAJ, LAXMI and SWARUP K PARIDA*

Plant Genomics and Molecular Breeding Laboratory, National Institute of Plant Genome Research (NIPGR), Aruna Asaf Ali Marg, New Delhi 110 067, India

**Corresponding author (Fax, +91-11-26741658; Email, swarup@nipgr.ac.in; swarupdbt@gmail.com)*

The enormous population growth, climate change and global warming are now considered major threats to agriculture and world's food security. To improve the productivity and sustainability of agriculture, the development of high-yielding and durable abiotic and biotic stress-tolerant cultivars and/climate resilient crops is essential. Henceforth, understanding the molecular mechanism and dissection of complex quantitative yield and stress tolerance traits is the prime objective in current agricultural biotechnology research. In recent years, tremendous progress has been made in plant genomics and molecular breeding research pertaining to conventional and next-generation whole genome, transcriptome and epigenome sequencing efforts, generation of huge genomic, transcriptomic and epigenomic resources and development of modern genomics-assisted breeding approaches in diverse crop genotypes with contrasting yield and abiotic stress tolerance traits. Unfortunately, the detailed molecular mechanism and gene regulatory networks controlling such complex quantitative traits is not yet well understood in crop plants. Therefore, we propose an integrated strategies involving available enormous and diverse traditional and modern –omics (structural, functional, comparative and epigenomics) approaches/resources and genomics-assisted breeding methods which agricultural biotechnologist can adopt/utilize to dissect and decode the molecular and gene regulatory networks involved in the complex quantitative yield and stress tolerance traits in crop plants. This would provide clues and much needed inputs for rapid selection of novel functionally relevant molecular tags regulating such complex traits to expedite traditional and modern marker-assisted genetic enhancement studies in target crop species for developing high-yielding stress-tolerant varieties.

[Kujur A, Saxena MS, Bajaj D, Laxmi and Parida SK 2013 Integrated genomics and molecular breeding approaches for dissecting the complex quantitative traits in crop plants. *J. Biosci.* **38** 971–987] **DOI** 10.1007/s12038-013-9388-6

1. Introduction

The convergence of population growth, variable climatic conditions, water shortage and global warming is expected to threaten food security on a worldwide scale, imposing an immensely global challenge for agriculture today (IPCC 2007). It is anticipated that by 2030, a developing country like India will be most severely affected due to insufficient production of major food crops like cereals (rice, wheat and maize) and pulses as because of massive impact of climate change on the tropics and sub-tropics, huge population growth and dependency of more than half of this population on agriculture (Reynolds and Ortiz 2010). To meet the dietary demand of the fast-increasing population in India, it is thus imperative to develop high-yielding durable stress-tolerant cultivars, particularly of the staple cereal and pulse (chickpea and pigeonpea) crops, that feeds over half of the Indian population. However, most of these yield, and abiotic (drought, salinity, cold, high temperature, acidity and sodicity) and biotic (evolution of new races of pathogens due to climate change) stress component agronomic traits are complex and quantitative in nature and multiplicatively governed by many major and several small effect genes/QTLs (quantitative trait loci) (Wang *et al.* 2008; Tardieu and Tuberosa 2010). Henceforth, understanding the molecular mechanism and dissection of such complex quantitative traits is the prime objective in current plant genomics and molecular breeding research. This can be

Keywords. Complex quantitative traits; genomics; molecular breeding; SNP; SSR

Published online: 6 November 2013

effectively deciphered by decoding all the indispensable information of inheritance encoded in the nucleic acids and chromatin through integrated traditional and modern -omics approaches/resources (structural, functional, comparative and epigenomics) and subsequently their efficient utilization in genomics-assisted breeding for genetic improvement of crops for higher yield and stress tolerance.

The completion of genome and expressed sequence tag (EST) sequencing and gene discovery projects for several crop species like rice, *Arabidopsis*, sorghum, maize and soybean based on first-generation Sanger sequencing methods have generated a wealth of genomic and genic sequence information including fully characterized known and candidate genes, transcription factors and regulatory sequences. These sequence resources are now freely accessible at various web databases. Additionally, the recent advent of high-throughput next-generation whole genome and transcriptome sequencing, array-based genotyping and modern bioinformatics approaches have enabled to produce huge genomic and transcriptomic resources globally on a genome-wide scale in diverse crop genotypes. Further, the epigenetic modifications involving histone methylation and acetylation, and DNA methylation during transcriptomic activity and their regulation of genes involved in yield and stress responses have also been well studied (He *et al.* 2011; Kohler *et al.* 2012). Moreover, the integration of structural, functional and comparative genomics including epigenomics with marker-assisted breeding (MAB)/genomics-assisted breeding have been implicated to be an effective approach for identification of genes/QTLs and expressed QTLs (eQTLs) and their regulatory sequences involved in expression of an individual trait in crop plants (Emilsson *et al.* 2008; Kim *et al.* 2010). These integrated approaches have also been proved efficient for dissecting the complex quantitative traits like tuber quality in potato and diseases in *Arabidopsis thaliana* (Boerjan and Vuylsteke 2009; Terpstra *et al.* 2010). Collectively, all these integrated –omics and MAB approaches can identify functionally relevant molecular tags including informative molecular markers, novel gene-associated targets, genes/QTLs (eQTLs), alleles, epialleles, transcription factors and regulatory sequences which will be useful to unravel the molecular basis of complex quantitative yield and stress tolerance traits in crop species (Langridge and Fleury 2011). However, very limited efforts have been made till now in integrative usage of such genomic, transcriptomic and epigenomic resources in genomics-assisted breeding to decipher the molecular and gene regulatory networks and identify the potential genetic loci regulating the complex traits in crop plants.

Keeping all the above in view, we propose a strategy for integration of available traditional and modern –omics resources/approaches comprehensively with genomics-assisted breeding to decode the molecular and/or gene regulatory networks for identification of functionally relevant novel gene-associated targets and alleles controlling the complex quantitative yield and stress tolerance traits in crop plants. These proposed strategies are summarized under two broad research areas: (A) Integration and correlation of enormous structural, functional, comparative and epigenomics approaches/resources pertaining to complex yield and stress tolerance traits to identify functionally relevant molecular tags (traditional and modern next-generation molecular markers, novel genes/QTLs, eQTLs, alleles and epialleles) regulating the target traits and (B) use of superior functional gene/QTL combinations and favourable natural allelic variants in traditional and modern genomics-assisted breeding/MAB to develop novel high-yielding stress-tolerant crop varieties.

2. Impact of global climate change on abiotic and biotic stresses

The effect of climate change on potential abiotic stresses such as changes in CO_2 concentration, increased temperature, drought and changing rainfall patterns is now considered the major impediment to the crop productivity (Pachauri and Reisinger 2007; Vahdati and Leslie 2013). The continuous and combined impact of abiotic stress factors particularly the global water shortage and high temperature due to worldwide climate change impairs the photosynthetic activity during the day and night-time, which in turn increases the photorespiratory losses and ultimately lowering the crop productivity. This is clearly evident from the study documenting reduced major cereal production owing to the rising global temperature between the years 1981 to 2002 (Lobell and Field 2007). Additionally, the elevation of greenhouse CO_2 and/or O_2 gas concentration in the subtropics is assumed to create global drought stress and subsequent threat to sustainable crop farming in agriculture. In coming years, the drought, flooding, land degradation, high CO_2/O_2 concentration and extreme temperature (heat/cold) all associated with variable climates are likely to increase the problem of food crisis particularly in tropical and subtropical regions of world.

On the other hand, the effect of biotic stresses comprising of plant pathogens, pests and diseases on reducing crop productivity under changing climate scenario is quite evident from several lines of studies. These studies indicated that the effect of climate change possibly alters the stages and rates of development of the pathogen, and modify host resistance including changes in the physiology of host–pathogen interactions. This in turn shifts the geographical distribution of host and pathogen outside their historic ranges causing disease epidemics and substantial crop loss (Coakley *et al.* 1999;

Gregory *et al.* 2009). The findings on climatic influence on plant pathogens, pests and diseases revealed that climate change can either lead to the emergence of pre-existing pathogens as major disease agents or can provide the climatic conditions required for introduced pathogens to emerge. For instance, dry weather tends to favour growth of insect vectors and viruses, whereas wet weather favours fungal and bacterial pathogens. Higher winter temperature (−6 to −10°C) increases survival of overwintering rust fungi (*Puccinia graminis*) causing subsequent disease epidemics (Pfender and Vollmer 1999). Likewise, milder winters, higher nocturnal temperatures and higher overall temperatures enable increased winter survival of plant pathogens, accelerated vector and pathogen life cycles, and increased sporulation and infectiousness of foliar fungi (Harvell *et al.* 2002). The temperature also governs the rate of reproduction in many pathogens. For example, the root rot pathogen reproduces more quickly at higher temperatures (Waugh *et al.* 2003). The increasing level of atmospheric CO_2 reduces the ability of soybean to defend against bean leaf beetle resulting in increased beetle populations and subsequent yield loss (Zavala *et al.* 2008). Furthermore, new races may evolve rapidly under elevated temperature and CO_2 concentration (Chakraborty 2013). These races under evolutionary forces govern massive pathogen populations which are supported by a combination of increased fecundity and infection cycles under favourable microclimate. In addition to these, the changing geographic distribution brings diverse lineages/genotypes which do share the common ecological niche resulting in an increased pathogen diversity (Anderson *et al.* 2004; Garrett *et al.* 2006; Pautasso *et al.* 2012).

Henceforth, efforts have been made to mitigate the challenges posed by abiotic and biotic stresses by developing stress-tolerant high-yielding crop varieties (Mittler 2006; Redden 2013) that adapted to diverse climatic conditions through integrated genomics and molecular breeding approaches (Reynolds and Ortiz 2010). However, it seems much difficult particularly in case of biotic stresses because of rapid mutation of living organisms like pathogens and insects. Therefore, it becomes very tough to ascertain the similarity between strains and their pathogenicity in different environment and resistant phenotypes of resistant crop varieties in the near future.

3. Integration and use of genomic, transcriptomic and epigenomic resources/approaches for mining functionally relevant molecular tags controlling the complex traits

The use of traditional and modern structural, functional and comparative genomics including epigenomics approaches have made available the huge genomic, transcriptomic and epigenomic sequence resources. These sequence resources further expedited the discovery of a large number of annotated known cloned and candidate genes, microsatellite and single nucleotide polymorphism (SNP) markers, QTLs, alleles, epialleles, regulatory sequences and transcription factors in diverse crop genotypes targeting the yield and abiotic stress tolerance traits.

3.1 *Genomic, transcriptomic and epigenomic sequence resource*

The whole genome sequencing efforts using the first-generation Sanger sequencing-based clone-by-clone and/or whole genome shotgun (WGS) approaches have been accomplished in crop species like rice, *Arabidopsis* sorghum, maize and soybean (table 1). Recently, utilizing the next-generation whole genome *de novo* sequencing approaches, the draft genomes of crop plants such as bread wheat, cucumber, foxtail millet and chickpea (*desi* and *kabuli*) have been released. Some of the crop genomes like barley, pigeonpea, potato, tomato, *Medicago* and *Brassica rapa* have also been sequenced integrating both Sanger sequencing and next-generation sequencing approaches (table 1). The completion of whole genome sequencing in diverse contrasting genotypes of many crop species have now enabled to generate enormous genomic sequence resources, and structurally and functionally annotated protein-coding and non-protein-coding genes and transcription factors (table 1). Additionally, the completely sequenced plant genomes have provided detail insight into their genomic constitution and structural organization including implication towards functional and comparative genomic analyses and phylogenetics during such crop plant domestication. For example, the completion of maize genome sequencing inferred the abundance (~85% of genome) of hundreds of transposable element families contributing to complexity and diversity of its genome (Schnable *et al.* 2009). The hexaploid wheat genome sequencing revealed the polyploidization and expansion of gene families like genes involved in energy harvesting, metabolism and growth associated with its crop productivity (Brenchley *et al.* 2012). The tomato genome sequencing gave insight into the fleshy fruit evolution due to its genome triplications resulting in neofunctionalization of genes controlling fruit characteristics like colour and fleshiness, which are absent in other sequenced Solanaceous crop species like potato (TGC 2012). The recent whole genome duplication events as observed in legume crop species like paleopolyploid soybean was absent in diploid legumes chickpea and pigeonpea that are clearly evident from their draft genome sequences (Schmutz *et al.* 2010; Varshney *et al.* 2012, 2013; Jain *et al.* 2013). The sorghum genome sequencing gave clues regarding retrotransposon accumulation in its recombinationally recalcitrant heterochromatic sequenced regions which might contribute to a larger genome size of sorghum (~75%) compared with rice (Paterson *et al.* 2009). This study also inferred that higher drought

tolerance in sorghum than other cereal species is possibly due to recent gene and microRNA duplications throughout its genome. The autotetraploid potato genome gave clues regarding gene family expansion, tissue-specific expression and recruitment of genes to novel pathways that contribute to the evolution of its tuber development (PGSC, 2011). The genome sequence of *Medicago* revealed perception towards evolution of endosymbiotic rhizobial nitrogen fixation by sub- and/or neo-functionalization of genes having specialized role in nodulation during its ancient whole genome duplication about 58 million years ago (Young *et al.* 2011).

To develop transcriptomics resources, the macro-array analysis such as suppression subtractive hybridization (SSH) and cDNA-AFLP (Amplified fragment length polymorphism) have been assayed in different vegetative and reproductive tissues during developmental stages of diverse crop genotypes under normal growth and stress-induced conditions. The integration of such approaches with low-throughput first generation conventional Sanger sequencing method and the recent next-generation long-read sequencing technology like Roche 454 Pyrosequencer have generated millions of EST (expressed sequence tags) and full-length cDNA sequences (NCBI GenBank, *http://www.ncbi.nlm.nih.gov*). For instance, about 400,000 ESTs and 60,045 full-length cDNA in rice (Kikuchi *et al.* 2003) and 38,097 ESTs representing 14,486 unigenes in chickpea (Varshney *et al.* 2009a, b) have been developed from various developmental stages/tissues and abiotic stress-induced cDNA libraries of diverse genotypes using the Sanger sequencing technology. The advent of array-based whole genome transcriptome profiling [such as microarray chips, serial analysis of gene expression (SAGE), massively parallel signature sequencing (MPSS)] and recently the whole genome next-generation transcriptome sequencing/RNA sequencing (RNA-seq) assayed in different tissues/developmental stages of diverse crop genotypes under stress-induced conditions have expedited the generation of numerous transcript sequences globally including differentially expressed transcripts encoding the known/candidate genes (NCBI, GEO database). Additionally, it also provide useful information regarding alternatively spliced genes as well as identification of rare and novel transcriptional active regions for diverse crop plants in a genome-wide scale. For example, about 37,265–53,409 *de novo* transcripts from diverse *desi*, *kabuli* and wild chickpea genotypes representing different developmental stages/tissues have been produced using the next-generation whole genome transcriptome sequencing technologies particularly the Roche 454/FLX Pyrosequencer and Illumina Solexa Genome Analyzer (Hiremath *et al.* 2011; Garg *et al.* 2011a, b; Jhanwar *et al.* 2012; Chickpea Transcriptome Database (CTDB) release 1.0, *http://www.nipgr.res.in/ctdb.html*).

The epigenome information like DNA methylation pattern and histone acetylation, methylation, phosphorylation and ubiquitination enrichment sites assayed during different developmental stages/tissues and stress induced conditions of diverse contrasting crop genotypes are available at present (He *et al.* 2011; Kohler *et al.* 2012). The documentation of all the epigenetic regulators (histone modifications and DNA methylation) across the whole plant genome forms a high-resolution map of plant epigenome (Bernstein *et al.* 2007). Such epigenome maps have been generated recently in plants using combination of chromatin profiling, genomic tiling microarrays and high-throughput next-generation sequencing technologies (Zhang 2008; Zhu 2008; Eckardt 2009; He *et al.* 2011). Such high-throughput profiling of plant epigenomes have provided about a deep insight regarding the diverse role of the epigenetic marks in regulating growth, development and stress responses. Decoding the genetic information on differential epigenetic modifications of chromatin (histone) structure and DNA methylation patterns have utility to provide clues regarding contribution of epialleles in controlling gene/promoter regulation and differential gene/transcript expression during developmental stages/tissues and stress responses in diverse contrasting crop genotypes (table 2). For instance, DNA methylation have shown to play a crucial role in rice endorsperm biogenesis during seed development by hypomethylating and thus activating the genes (encoding major storage proteins and starch synthesizing enzymes) preferentially expressed in endosperm (Zemach *et al.* 2010). Moreover, extensive epigenomic reprogramming and regulation comprising of RNA-directed DNA methylation and chromatin modifications have been observed during gametogenesis in plants (Baroux *et al.* 2011). The environmental signals also strongly trigger DNA methylation and nucleosome histone post-translational modifications leading to the establishment of stress-responsive epigenomes (Chinnusamy and Zhu 2009; Mirouze and Paszkowski 2011). Therefore, the heritable natural epigenetic variants (epialleles) and transposons mobility regulated during epigenetic mechanisms can be exploited to broaden the plant phenotypic and genotypic variations for improving its adaptation to abiotic and biotic stresses which in turn will increase the productivity (Mirouze and Paszkowski 2011). The significance of epiallele formation in the absence of genetic variation for controlling phenotypic variation is well demonstrated particularly in case of peloric and colourless fruit non-ripening/ripening epigenetic variants from *Linaria vulgaris* and *Solanum lycopersicum*, respectively (Cubas *et al.* 1999; Manning *et al.* 2006; Zhong et al. 2013). Henceforth, informative epialleles can potentially be utilized as epigenetic markers for identification and mapping of gene-associated targets regulating the traits of agricultural importance in diverse crop plants.

The enormous genomic and transcript sequences (differentially expressed transcripts/genes) generated from diverse crop genotypes using both traditional and modern –omics approaches have potential to develop a large number

Table 1. A short glimpse of plant genomes sequenced with traditional and modern next-generation sequencing approaches

Families	Common names of plant species	Scientific name of plant species (genotypes sequenced)	Ploidy level	Estimated genome size	Approaches used for genome sequencing	Genome size sequenced	chromosomes sequenced	Genetic information generated		References
								No. of genes	Molecular markers	
Poaceae	Rice	*Oryza sativa* (Nipponbare)	2n=2x=24	389 Mb	Map-based clone-by-clone approach (Sanger sequencing)	370 Mb	12	37,544	18,828 SSRs 80,127 SNPs	IRGSP (2005)
	Bread wheat	*Triticum aestivum* (Chinese Spring/ CS42)	2n=6x=42 (AABBDD)	17 Gb	Whole genome shotgun (Roche 454 pyrosequencing)	17 Gb	7	94,000- 96,000	132,000 SNPs	Brenchley *et al* (2012)
	Einkorn wheat	*Triticum urartu* (G1812/PI428198)	2n=2x=14 (AA)	4.94 Gb	Whole genome shotgun (Roche 454 and Illumina Next-generation sequencing)	4.66 Gb	7	34,879	-	Ling *et al* (2013)
	Tusch's goatgrass	*Aegilops tauschii* (AL8/78)	2n=2x=14 (DD)	4.36 Gb	Whole genome shotgun (Roche 454 and Illumina Next-generation sequencing)	4.23 Gb	7	43,150	860,126 SSRs 711,907 SNPs	Jia *et al* (2013)
	Maize	*Zea mays* (B73)	2n=2x=20	2.3 Gb	Map-based clone-by-clone approach (Sanger sequencing)	2.3 Gb	10	32,000	> 3.3 million SNPs and InDels	Schnable *et al* (2009)
	Sorghum	*Sorghum bicolor* (Moench)	2n=2x=20	730 Mb	Whole genome shotgun (Sanger sequencing)	~730 Mb	20	27,640	~71,000 SSRs	Paterson *et al* (2009)
	Barley	*Hordeum vulgare* (Morex)	2n=2x=14	5.1 Gb	Whole genome shotgun (Sanger sequencing and Illumina Next-generation sequencing)	4.98 Gb	7	26,159	> 15 million SNPs	IBGSC (2012)
Fabaceae	*Medicago*	*Medicago truncatula* (A17)	2n=2x=16	500 Mb	Map-based clone-by-clone approach (Sanger sequencing) and Whole genome shotgun (Illumina Next-generation sequencing)	375 Mb	8	62,338	> 3 million SNPs	Young *et al* (2011)
	Soybean	*Glycine max* (Williams82)	2n=4x=40	1.1 Gb	Whole genome shotgun (Sanger sequencing)	950 Mb	20	46,430	874 SSRs 4,991 SNPs	Schmutz *et al* (2010)
	Pigeonpea	*Cajanus cajan* (ICPL87119/Asha)	2n=2x=22	833.07 Mb	Whole genome shotgun (Sanger sequencing and Illumina Next-generation sequencing)	605.78 Mb	11	48,680	309,052 SSRs 28,104 SNPs	Varshney *et al* (2012)
	Chickpea	*Cicer arietinum* *Kabuli* (CDC Frontier)	2n=2x=16	740 Mb	Whole genome shotgun (Roche 454 and Illumina Next-generation sequencing)	~738 Mb	8	28,269	81,845 SSRs 76,084 SNPs	Varshney *et al* (2013)
		Desi (ICC4958)	2n=2x=16	740 Mb	Whole genome shotgun (Roche 454 and Illumina Next-generation sequencing)	520 Mb	8	27,571	30,000 SSRs 60,000 SNPs	Jain *et al* (2013)

Table 1 (continued)

Families	Common names of plant species	Scientific name of plant species (genotypes sequenced)	Ploidy level	Estimated genome size	Approaches used for genome sequencing	Genome size sequenced	chromosomes sequenced	Genetic information generated		References
								No. of genes	Molecular markers	
Solanaceae	Potato	*Solanum tuberosum* (Phureja/DM1-3 516 R44)	2n=4x=48	900 Mb	Whole genome shotgun (Sanger sequencing and Illumina Next-generation sequencing)	727 Mb	12	39,031	3.67 million SNPs	PGSC (2011)
	Tomato	*Solanum lycopersicum* (*S. pimpinellifolium*/ LA1589)	2n=2x=24	844 Mb	Map-based clone-by-clone approach (Sanger sequencing) and Whole genome shotgun (Illumina Next-generation sequencing)	760 Mb	12	34,727	5.4 million SNPs	TGC (2012)
Brassicaceae	Chinese cabbage	*Brassica rapa* (Chiifu-401-42)	2n=2x=20	475 Mbp	Map-based clone-by-clone approach (Sanger sequencing) and Whole genome shotgun (Illumina Next-generation sequencing)	283.8 Mb	10	41,174	-	BGSPC (2011)

of informative genomic and genic microsatellite and SNP markers *in silico* at a genome-wide scale.

3.2 *Traditional and next-generation modern microsatellite and SNP marker resources*

For the orphan crops for which no sequence information is available, the use of *de novo* microsatellite isolation methods, namely, size fractionated genomic and cDNA libraries and microsatellite-enriched genomic libraries based on first-generation Sanger sequencing methods, are indispensable prior to designing of traditional microsatellite markers (Parida *et al.* 2009a). The crop species for which enormous genomic (BAC-/PAC-end sequences, whole genome sequences/ psedomolecules) and genic sequence resources (annotated known/candidate genes, full-length cDNA, ESTs and unigenes) in diverse genotypes are available with on-line public databases [NCBI (*http://www.ncbi.nlm.nih.gov*), EMBL (*http://www.embl.de*), EBI (*http://www.ebi.ac.uk*), DDBJ (*http://www.ddbj.nig.ac.jp*), TIGR (*http://rice.plantbiology.msu.edu*), Phytozome (*http://www.phytozome.org*) and TAIR (*http://www.arabidopsis.org*)], the identification of large-scale genomic and genic microsatellite markers *in silico* (Varshney *et al.* 2005; Parida *et al.* 2006, 2009b, 2010) by employing different bioinformatics tools are found expedient (table 1). The development of such *in silico* microsatellite markers involve software tools like MISA (MIcroSAtellite, *http://pgrc.ipk-gatersleben.de/misa/*), SSRFinder, Tandem repeat Finder, Tandem Repeat Occurrence Locator, Simple Sequence Repeat Identification Tool (SSRIT, *www.gramene.org/db/searches/ssrtool*), Modified Sputnik, SSR Locator, BatchPrimer3 (*http://probes.pw.usda.gov/cgi-bin/batchprimer3/batchprimer3.cgi*) and Plant Genetics and Genomics (PGG) Bioinformatics (*http://hornbill.cspp.latrobe.edu.au*). Using above tools, a large-scale 2,240 rice microsatellite (RM) markers have been developed from the first phase draft whole genome sequences of *indica* and *japonica* (McCouch *et al.* 2002). In chickpea, a larger set of 250 microsatellite markers have been identified from the microsatellite-enriched genomic libraries using the Sanger sequencing (Sethy *et al.* 2006). Consequently, about 24,389 microsatellite markers from the whole genome sequences/pseudomolecules and annotated protein coding genes of 12 chromosomes in rice (*www.gramene.org*) and also more than thousands of genomic and genic microsatellite markers from the size fractionated and microsatellite-enriched genomic and cDNA libraries in chickpea (Choudhary *et al.* 2009, 2012; Nayak *et al.* 2010; Gujaria *et al.* 2011) have been developed. Moreover, the comparison of huge genomic/genic sequences generated from first-generation Sanger sequencing methods among diverse crop genotypes have enabled to produce a larger set of *in silico* polymorphic microsatellite markers for their immense use in genomics-assisted breeding. For example, 52,485 *in silico* polymorphic genomic and genic

Table 2. Epigenetics landmarks on yield and stress tolerance traits in crop plants

Families	Plant Species	Methods used for epigenome analysis	Gene expression/traits studied	References
Poaceae	Rice (*Oryza sativa*)	DNA methylation	Biotic stress resistance	Akimoto *et al* (2007)
		DNA methylation	Transposable element control	Kashkush and Khasdan (2007)
		DNA methylation and histone methylation (H3K4me2/H3K4me3)	Plant development	Li *et al* (2008)
		Histone acetylation (H3ac/H4ac) and Histone methylation (H3K4me/H3K4me2)	Salt stress response	Mukhopadhyay *et al* (2013)
		Histone modifications (H3K4me2, H3K4me3, H3K9ac and H3K27ac)	Genome-wide analysis	Du *et al* (2013)
		DNA methylation and histone modifications (H3K4me3, H3K27me3 and H3K9ac)	Global epigenetic analysis	He *et al* (2010)
		DNA hypomethylation	Endosperm development	Zemach *et al* (2010) Rodrigues *et al* (2013)
	Maize	DNA methylation and histone methylation (H3K27me3)	Genome-wide analysis	Eichten *et al* (2011) Wang *et al* (2009)
		DNA methylation and histone methylation (H3K9me2/H3K27me2)	Vegetative phase change	Li *et al* (2010)
		DNA methylation	Endosperm development	Waters *et al* (2011) Zhang *et al* (2011b)
	Hexaploid wheat	DNA methylation	Flowering development	Shitsukawa *et al* (2007)
	Amphiploid wheat	DNA methylation	Transposable element control	Kashkush *et al* (2003)
	Barley	Histone methylation (H3K27me3)	Vernalization induced flowering	Oliver *et al* (2009)
		DNA methylation	Seed development	Kapazoglou *et al* (2012)
Fabaceae	*Lotus*	DNA methylation	Plant development and transposable element control	Fukai *et al* (2013)
	Soybean	DNA methylation	Patterns and heritability of differentially methylated regions (DMR) and identification of methylQTLs controlling phenotypes	Schmitz *et al* (2013a)
Solanaceae	Tomato	DNA methylation	Fruit ripening regulation by modifying expression of SBP (SQUAMOSA binding protein) and RIN (ripening inhibitor) transcription factor genes and ripening genes	Manning *et al* (2006) Zhong *et al* (2013)
Brassicaceae	*Arabidopsis*	DNA methylation	Patterns of population epigenomic diversity in seed and pollen	Schmitz *et al* (2013b)

microsatellite loci have been identified by comparing the whole genome sequences/pseudomolecules of 12 chromosomes of *indica* 93-11 and *japonica* Nipponbare for their wider use in large-scale genotyping applications of rice (Zhang *et al.* 2007).

Likewise, the availability of whole genome sequences/pseudomolecules and transcriptomic sequence resources of multiple crop genotypes in public domain have enabled rapid identification and development of genomic and genic SNP markers *in silico* in a genome-wide scale (table 1). For example, the comparison of whole genome sequences of *indica* and *japonica* rice led to the discovery of more than 5 million SNP loci (*http://www.ncbi.nlm.nih.gov/snp/?term=oryza*) for high-throughput genotyping applications. Further, the continuously growing known/candidate gene encoding EST and/unigene databases have provided the opportunity to detect SNPs in the transcribed regions of the genome either by exploiting the redundancy of gene sequences/assembled contigs or the diversity of genotypes represented within database (Picoult-Newberg *et al.* 1999; Kota *et al.* 2007). To automate such process of SNP discovery, various SNP mining software tools like d2 cluster, CAP3, Auto SNP, SniPpER, SNPServer and SEAN have been developed. In rice, more than one million SNP loci and thousands of insertions/deletions (InDels) have been discovered through aligning about one lakh EST and full-length cDNA sequences of diverse rice genotypes (*http://cdna01.dna.affrc.go.jp/cDNA, www.cerealsdb.uk.net*). Moreover, acquiring and correlating the comparative genomic resources (high-density comparative genetic and physical maps) right from small diploid and model sequenced crop species to relatively large genome crops, the conserved orthologous genes, markers and sequences controlling the target yield and stress tolerance traits have been identified efficiently to utilize in genetic improvement of diverse large genome (not amenable to complete sequencing) and under-utilized/orphan crop genotypes. For example, the allele-specific resequencing of 220 orthologous

candidate genes (transcription factors) among diverse chickpea genotypes and reference legume and dicot crop species (*Medicago*, *Lotus*, *Glycine* and *Arabidopsis*) have identified more than two thousand marker loci including COS (conserved orthologous sequence), SNPs, CAPS (cleaved amplified polymorphic sequences) and CISR (conserved intron spanning region) markers to accelerate MAB in chickpea (Gujaria *et al.* 2011).

Recently, the next-generation high-throughput DNA sequencing technologies relied on Roche 454/FLX Pyrosequencer, ABI SOLiD, Illumina Solexa Genome Analyzer, Helicos, Ion Torrent and Polonator (Simon *et al.* 2009; Varshney *et al.* 2009c; Metzker 2010; Jain 2012) have been proved to be a powerful approach for mapping, quantifying and resolving the genomic, transcriptomic and epigenomic profiles at a genome-wide scale in diverse crop genotypes. The major utility of these next-generation sequencing approaches for generation of large-scale novel genomic and genic microsatellite and SNP markers is well understood at present. The generation of these large-scale markers is possible through different modified next-generation sequencing applications including long- and short-read amplicon sequencing, whole genome and transcriptome resequencing/*de novo* sequencing, targeted/gene-enriched genome sequencing, organelle (chloroplast and mitochondria) genome sequencing, reduced representation library (RRL) and genotyping-by-sequencing (GBS) in diverse crop genotypes. Availability of large-scale next-generation genomic and transcript sequence resources (NCBI, GEO database) have accelerated *in silico* identification and designing of genomic and genic functional microsatellite and SNP markers in crop plants by employing different bioinformatics tools like NextGene, CLC Genomics, Pyrobayes, POLYBAYES and MAQ. Using these strategies, a larger set of genomic and transcript sequences generated from diverse crop genotypes through array-based high-throughput genome/transcriptome sequencing approaches have been used to develop *in silico* genomic and genic microsatellite and SNP markers for individual genotypes. On the other hand, these sequences have been compared among genotypes to develop informative *in silico* polymorphic markers in a genome-wide scale. For example, the genome/transcript sequences generated for diverse contrasting *desi*, *kabuli* and wild chickpea genotypes using the 454/FLX Pyrosequencer, ABI SOLiD and Illumina Solexa Genome Analyzer have now generated more than 30,000 and 48,298 microsatellite markers from individual *desi* and *kabuli* genotypes, respectively (Jain *et al.* 2013; Varshney *et al.* 2013) and about 5,000 *in silico* microsatellite loci showing polymorphism among genotypes for large-scale genotyping applications in chickpea (Hiremath *et al.* 2011; Garg *et al.* 2011a, b; Jhanwar *et al.* 2012). More than 60,000 SNP loci among three diverse *indica* and *japonica* rice genotypes have been identified by comparing about 45,000 common transcripts generated in these genotypes through next-generation whole genome transcriptome sequencing using Illumina Solexa Genome Analyzer (Lu *et al.* 2010). The Perlegen hybridization-based genome/transcriptome resequencing in 100 Mb genic regions of 20 diverse *indica* and *japonica* rice genotypes contrasting for yield and stress tolerance traits have detected about 160,000 high-quality SNP loci differentiating these genotypes (McNally *et al.* 2009). Moreover, about 1.6 to 8 million SNPs have been mined recently by resequencing of 1,000 rice germplasm lines using the next-generation sequencing approaches (Xu *et al.* 2012; Huang *et al.* 2012b). In case of chickpea, about 4.4 million SNPs including ~76,084 SNPs in 15,526 genes have been identified through next-generation whole genome sequencing-based GBS analysis of 90 genotypes (Varshney *et al.* 2013). With more genome and transcriptome sequence information being generated, the trend is towards generation of a large number of informative traditional and modern genomic and genic microsatellite and SNP markers *in silico* (figure 1), cross-referencing the genes, nuclear and organellar genomes and epigenomes of crop plants by integrating/correlating various –omics approaches. Further, large-scale validation, high-throughput genotyping and efficient utilization of these developed markers in various applications of plant genetics and breeding have clear implications towards identification of novel potential genes (QTLs)/alleles regulating the qualitative and complex quantitative traits.

4. Use of functionally relevant molecular tags for traditional and modern genomics-assisted breeding

A large number of markers, genes/QTLs and alleles identified through integrated –omics approaches/resources can be targeted simultaneously and validated functionally using the traditional and modern genomics-assisted breeding approaches like genetic/association mapping, which is not possible through transgenics. To hasten the process of genomics-assisted breeding, the large-scale validation and high-throughput genotyping of genic and genomic microsatellite and SNP markers in natural germplasm collections and mapping/mutant populations including their precise phenotyping have been initiated recently in many crops using the novel high-throughput and cost-effective next-generation sequencing and array-based genotyping technologies, and modern phenotyping platforms.

4.1 *High-throughput genotyping and phenotyping*

At present, a larger set of cultivated varieties, breeding lines, landraces, wild accessions and germplasm lines representing diverse agro-climatic regions of the world including India are available for many crop species. Especially, in case of rice, a largest germplasm collection comprising of 102,547

accessions of *O. sativa*, 1,651 accessions of *O. glaberrima* and 4,508 accessions of 22 wild ancestors (McNally *et al.* 2006) are available at the International Rice Research Institute (IRRI) as International rice germplasm collections (IRGC). In chickpea, more than 20,000 germplasm lines at major global research institute ICRISAT (International Crops Research Institute for the Semi-Arid Tropics), India, ~14,000 accessions at NBPGR (National Bureau of Plant Genetic Resources), India, and ~13,000 accessions at ICARDA (International Centre for Agricultural Research in Dryland areas), Syrian Arab Republic, and many more at other national and international research institutes are available (Gaur *et al.* 2012a). The detailed phenotypic characterization of these readily available genetic resources of crop plants for diverse important agronomic traits including tolerance to abiotic and biotic stresses and yield component traits are underway. In recent years, efforts have also been made for constructing the core and mini-core collections in several crops by identifying the largest amount of genetic diversity with a minimum number of accessions. For instance, using the phenotypic and genotypic selection methods based on 36 genic microsatellite markers and 50 agronomic traits, Zhang *et al.* (2011a) constructed a core collection of 932 accessions comprising 1.7% of entire 55,908 accessions of the *O. sativa* germplasm in China. In case of chickpea, the phenotypically and genotypically diverse 1,956 core, 211 mini-core and 300 global reference germplasm collections representing diverse eco-geographical regions of 58 countries of the world are available (Upadhyaya and Ortiz 2001; Upadhyaya *et al.* 2001, 2002). A larger set of advanced generation bi-parental and back-cross mapping population, e.g. RILs (recombinant inbred lines), NILs (near isogenic lines) and DHs (double haploids), derived from the parental genotypes contrasting for yield and stress tolerance traits are available in many crop species. Besides, EMS (ethyl methanesulfonate) and γ-ray irradiated mutant lines of diverse genotypes contrasting for yield and stress tolerance traits have been generated in many crops to identify point mutations or SNP sites having functional relevance for trait regulation. For example, a total of about 66,891 EMS (ethyl methanesulfonate), MNU (N-methyl-N-nitroso urea), sodium azide and γ-ray irradiated mutant lines of IR64 (Wu *et al.* 2005) and Nipponbare (Till *et al.* 2007) have been generated in rice to identify point mutations in genome/genes having functional significance for qualitative and quantitative trait regulation (*http://tilling.ucdavis.edu; http://www.iris.irri.org*). The TILLING (targeting induced local lesions in genomes, Till *et al.* 2007) and EcoTILLING (Raghavan *et al.* 2007) using the diverse mutant populations and natural germplasm collections of crops has made it possible to mine novel functional allelic variants in the transcripts encoding the known/candidate genes, respectively. The recent development of high-throughput array-based next-generation sequencing and marker genotyping technologies such as automated fragment anlyzer, TILING array, Fluidigm dynamic array, Illumina GoldenGate and Infinium assays, KASP (KBioScience Allele-Specific Polymorphism) profiling and MALDI-TOF (matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry) mass spectrophotometry, GeneTitan, RRL and GBS assays have expedited the large-scale validation and evaluation of functional allelic variation/diversity level in the naturally occurring germplasm lines, mapping populations and mutant collections.

Considering the importance of high-throughput and precision phenotyping for gene isolation and MAB, the automated modern high-throughput phenotyping and E (environmental)-typing platforms have recently been developed. These platforms are efficient for precise phenotyping and multi-environmental assays of a larger set of natural/mutant and mapping populations of crop plants for complex yield, and abiotic and biotic stress component traits (Mir *et al.* 2012; Xu *et al.* 2012). An International Plant Phenomics Network (IPPN) has been initiated currently to provide novel technologies for high-throughput and precise phenotyping of complex quantitative traits in many crop plants (Clark *et al.* 2011).

4.2 *Construction of genome/transcript maps, genetic/association mapping and genetical genomics*

The construction of high-resolution genetic linkage and functional transcript maps, fine mapping and map-based cloning/positional cloning of genes/QTLs using the informative genomic and genic microsatellite and SNP markers have traditionally been proved to be the most powerful tools for gene isolation and dissection of the complex quantitative yield and stress tolerance traits in crop plants. Recently, in rice the high-throughput next-generation whole genome and transcriptome sequencing technology have been successfully applied for constructing SNP-based ultrahigh-density genetic linkage/transcript maps (Huang *et al.* 2009; Xie *et al.* 2010). These high-density genome/transcript maps have considerably improved the resolution and accuracy of genes/QTLs mapping in chromosomes (Wang *et al.* 2011) and thus significantly accelerated the process of fine mapping and map-based cloning in rice. For instance, utilizing such fine mapping and map-based cloning approaches, at present a total of 26 genes underlying major QTLs for agronomic traits including grain size, heading date, plant height, prostrate growth, cold, salinity, lodging and submergence tolerance and disease resistance have been cloned and characterized in rice (Miura *et al.* 2011). A high-throughput integrated method of whole-genome re-sequencing and genotyping of SNPs in recombinant populations have found to be an effective approach for identification of candidate genes/QTLs associated with complex traits in crop plants based on recombination break-points determination. Such strategy of trait-specific gene/QTL mapping

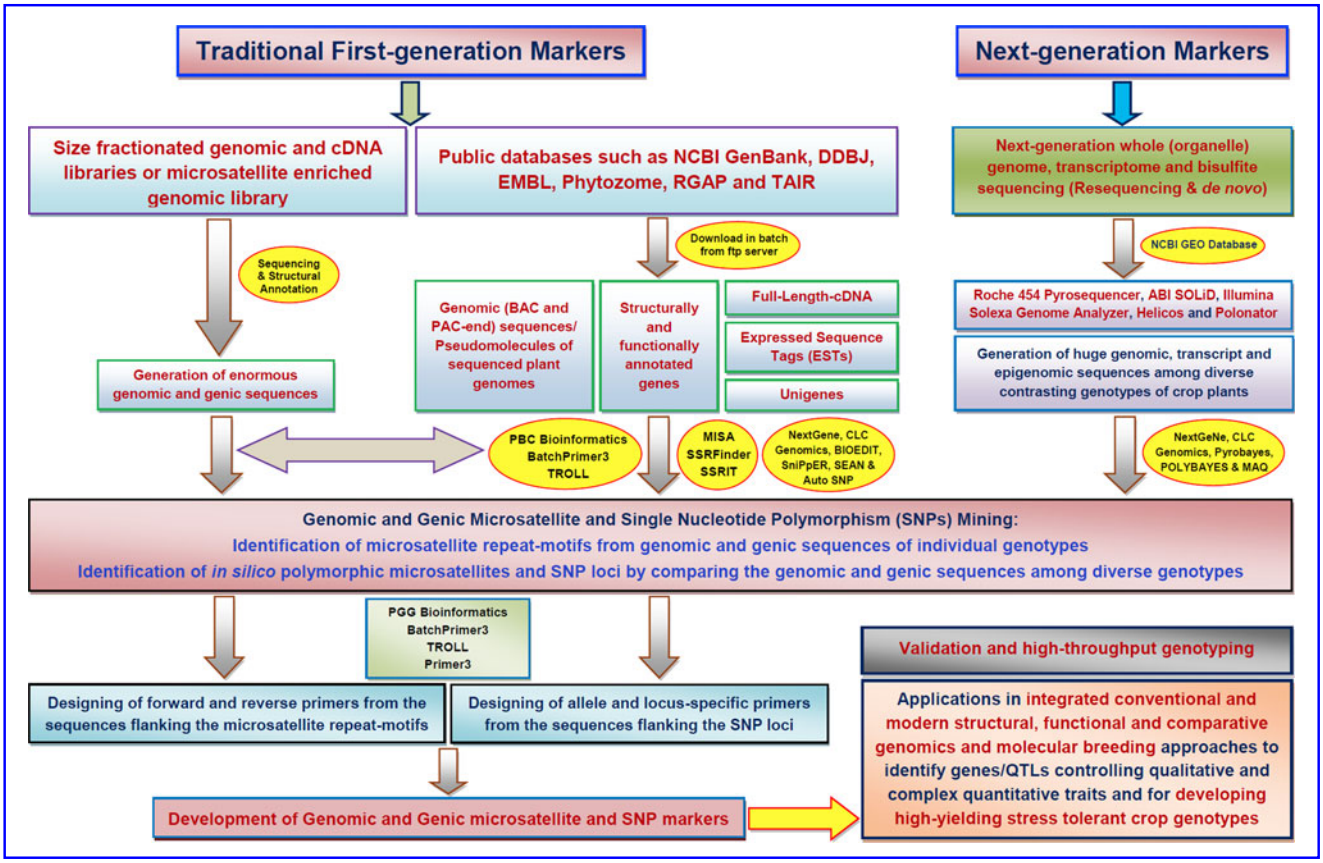

Figure 1. A schematic representation of development of traditional and next-generation microsatellite and SNP markers for large-scale genotyping applications in crop plants.

has been most widely studied in case of rice by resequencing and genotyping of 2.8 lakhs to 1.2 million SNPs in at least 150–200 individuals of RIL mapping populations using the Illumina Solexa Genome Analyzer (Huang *et al.* 2009; Yu *et al.* 2011). Moreover, the EMS-induced TILLING-based positional cloning method for identification of genes/QTLs regulating the complex agronomic traits has been demonstrated successfully in rice (Abe *et al.* 2012). The construction of high-density intra- and inter-specific genetic linkage/transcript map using more than thousands of traditional and modern next-generation microsatellite and SNP markers are underway in chickpea for identification and fine mapping of genes/QTLs controlling traits of agricultural importance (Nayak *et al.* 2010; Gujaria *et al.* 2011; Hiremath *et al.* 2012; Gaur *et al.* 2012b). However, such strategies are time consuming and provide difficulties in isolating casual genes underlying the QTLs, especially with minor effects. Remarkably, the integration of RIL population-based QTL mapping with microarray-based global transcript profiling and/or next-generation whole genome transcriptome sequencing involving the parents and bulks of homozygous RILs is found to be a powerful approach to narrow down the number of candidate genes underlying the QTLs regulating the complex and quantitative yield and stress tolerance traits in crop plants. More recently, a rapid method called 'QTL-seq' has been developed for mapping of genes/QTLs associated with blast disease resistance and seedling vigor in rice by whole genome resequencing of DNA from their two bulked RIL populations (Takagi *et al.* 2013).

Currently, the 'genetical genomics'/ 'expression genetics' integrating the genetic or QTL mapping with transcript profiling have been corroborated to be an effective approach for identification of candidate genes encoding transcripts and its regulatory sequences (transcription factors) involved in expression of quantitative traits in crop plants (Emilsson *et al.* 2008). Using such integrated approaches, several recent studies have mapped the transcripts showing differential expression either by traditional macro-/micro-arrays or next-generation transcriptome sequencing to the whole genome and further correlated that differential transcript profiling with QTL mapping to identify 'expression QTLs' (eQTLs) involved in the *cis*- and *trans*-trait regulation.

With the recent discovery of low cost next-generation whole genome and transcriptome sequencing and high-throughput genotyping technologies and novel advanced

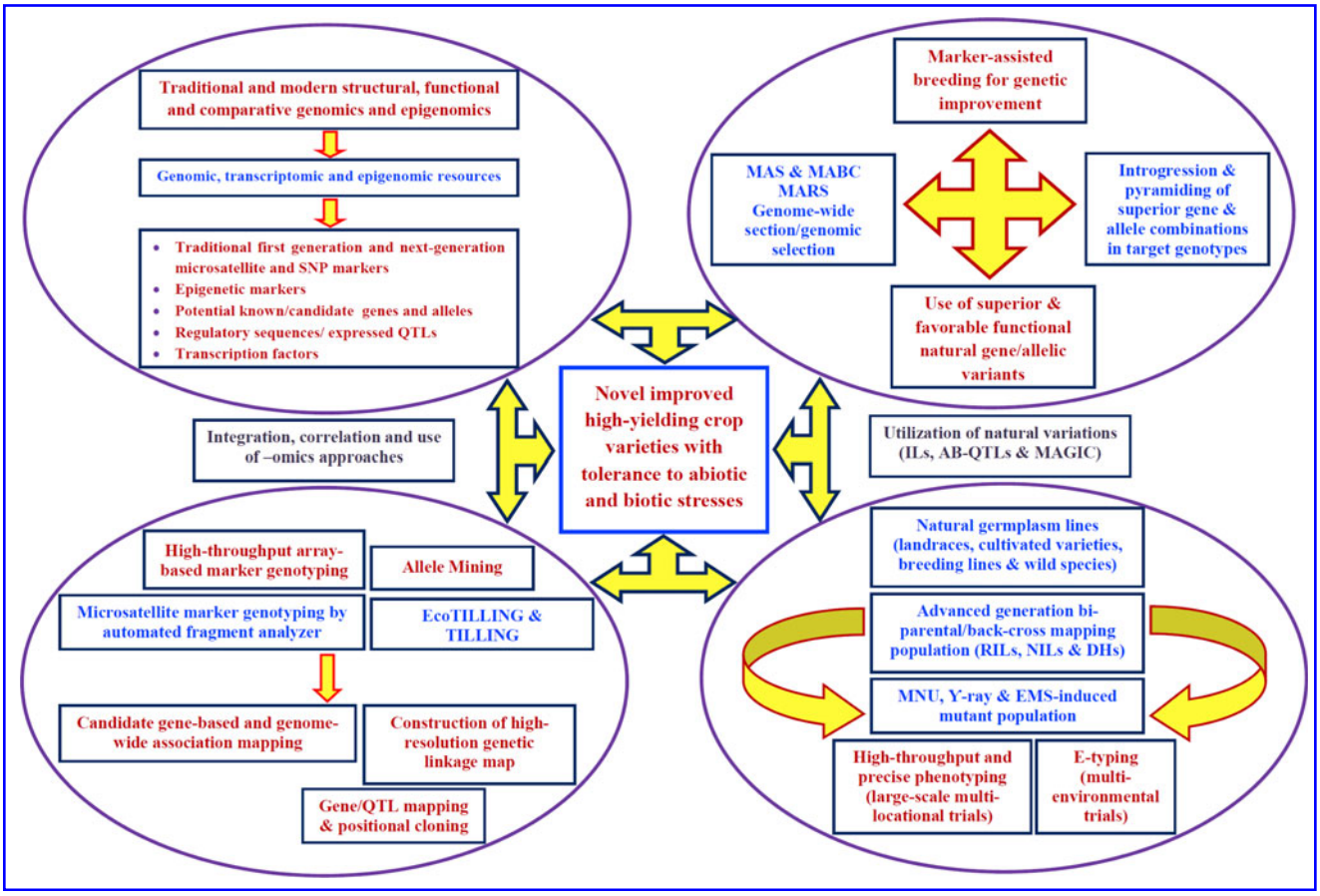

Figure 2. A schematic representation of integrated traditional and modern –omics resources/approaches and genomic-assisted breeding methods for dissecting the complex quantitative traits in crop plants.

structural, functional and comparative genomic tools, currently the large-scale validation and genotyping of genomic and genic functional microsatellite and SNP markers is not a great matter of concern. Therefore, the candidate gene-based and genome-wide association mapping relying on the large-scale genotyping information of such markers have proven to be an effective approach for identification of genes/QTLs and alleles regulating the complex quantitative yield and stress tolerance traits in crop species (Zhao *et al.* 2011; Li *et al.* 2011). In rice, genome-wide association (GWAS) mapping by high-throughput genotyping of thousands to millions of SNPs across 400–900 germplasm lines have been performed to identify target genomic regions including genes/QTLs associated with flowering time and grain yield traits (Huang et al. 2010, 2012c; Zhao *et al.* 2011). The candidate gene-based association mapping by targeting hundreds of SNPs in different coding and regulatory sequence components of genes also have significance to identify genes/QTLs controlling abiotic tolerance, biotic resistance, grain size, pericarp colour, seed shattering and starch synthesis in rice (Sweeney *et al.* 2007; Fan *et al.* 2006; Mao *et al.* 2010; Kharabian-Masouleh *et al.* 2012; Parida *et al.* 2012; Negrão *et al.* 2013). Notably, the integrated approach of genome-wide and candidate gene-based association mapping and traditional bi-parental genetic linkage (transcript) mapping has now clearly been established as an efficient approach for precise identification of functionally relevant robust genes/QTLs particularly those regulating the complex quantitative yield and stress component traits in crop plants. This is quite evident in case of *GS3* (Wang *et al.* 2011) and *GS5* (Li *et al.* 2011) for grain size trait regulation in rice. The integration of GWAS (44,000 SNPs genotyped across 373 gerpmlasm lines) with QTL mapping (78 back-cross inbred lines derived from Azucena x IR64) have enabled to identify a metal transporter gene regulating aluminium tolerance in rice (Famoso *et al.* 2011). Currently, the candidate gene-based association mapping, QTL mapping, differential transcript profiling and LD (linkage disequilibrium)-based haplotype gene evolution studies have been integrated to identify functionally relevant transcription factor genes and QTLs controlling 100-seed weight/seed size in chickpea (Kujur *et al.* 2013).

4.3 *Traditional and modern MAB*

The much wider functional natural genetic variation scanned in a larger set of germplasm lines including landraces and wild species particularly for yield and stress component traits have been transferred into the cultivated genetic background for crop improvement through introgression of favourable genes/QTLs/chromosomal segments. It involves traditional approaches like introgression lines (ILs) and advanced-backcross QTL (AB-QTL) analysis as well as modern methodologies such as association genetics and multi-parent advanced generation intercross (MAGIC) population (Tan *et al.* 2008). The transfer of functional natural allelic variants for yield component traits and starch quality has been successfully demonstrated in rice using ILs (Tian *et al.* 2006) and AB-QTLs (McCouch *et al.* 2007; Tan *et al.* 2008). Recently, the QTLs controlling plant height and hectolitre weight have been identified and mapped using 1,579 MAGIC population derived from four elite wheat cultivars (Huang et al. 2012a). The trait-specific favourable genomic regions, superior functional genes/QTLs with major as well as minor effects and natural allelic variants once identified or validated through genetic and association mapping and/or genetical genomics approaches (eQTLs) can now be easily combined and pyramided in the selected crop genotypes of interest through the conventional and modern breeding methods. The traditional breeding strategies include approaches like marker-assisted selection (MAS) and marker-assisted back-crossing (MABC)/marker-assisted foreground and background selection, while the novel modern breeding approaches involve methods such as marker-assisted recurrent selection (MARS) and genomic/genome-wide (haplotype) selection for genetic enhancement of crop plants for yield and stress tolerance. This is clearly evident from the studies involving genetic improvement of Basmati rice for yield, quality and resistance to bacterial leaf blight and blast diseases by pyramiding the multiple genes through MAS and MABC (Joseph *et al.* 2004; Sundaram *et al.* 2008; Gopalakrishnan *et al.* 2008; Singh *et al.* 2011). Moreover, the improvement of hybrid performance and bacterial blight, blast and brown plant hopper disease resistance of Indian rice varieties have been well executed through MAS using fertility restorer and disease resistant genes, respectively (Basavaraj *et al.* 2010; Singh *et al.* 2011). The sub-mergence tolerance in swarna through MAS using the *Sub1* QTL (Septiningsih *et al.* 2009) and MAS for grain length using functional markers developed from multiple loci of *GS3* gene (Wang *et al.* 2011) have also suggested the implications of traditional breeding methods for crop genetic enhancement. However, the genetic enhancement of crops for complex quantitative traits using traditional MAS (QTL-MAS) is not always feasible due to complications in genetic background effects/epistasis and linkage drag of QTLs. Moreover, the minor and major QTLs/genes showing minor effects on complex trait regulation cannot be transferred or combined together by traditional MAS for crop genetic improvement. The MARS, GWAS-MAS and genome-wide haplotype selection/genomic selection; other emerging approaches for pyramiding the favourable alleles of minor effect genes/QTLs controlling the complex quantitative traits at whole genome level, have also been initiated recently in crop plants for genetic enhancement (Meuwissen *et al.* 2001; Chia and Ware 2011). Once the suitable training population is selected and estimated breeding value (EBV) assigned to the training population, the selection of additional large-scale germplasm lines can be done based on EBV (alleles at associated markers) of training population (Jannink *et al.* 2010), which will expedite the process of genomic selection for complex trait genetic enhancement studies in diverse crop species.

5. Conclusion

Considering the availability of largely diverse traditional (first-generation) and modern (next-generation) –omics approaches/ resources and genomics-assisted breeding methods in many crop species as discussed above, we would like to propose an integrated strategies (figures 1 and 2). These integrated strategies can be adopted/utilized by agricultural biotechnologist for understanding the genetic basis of complex quantitative yield and stress tolerance traits in crop plants.

- Integrate/correlate the genomic, transcriptomic and epigenomic sequence resources in different developmental stages/tissues of diverse crop genotypes under normal growth and stress-induced conditions using different traditional and modern –omics approaches: Develop a large-scale traditional and next-generation genomic and genic microsatellite and SNP marker resource and identify gene-associated targets (differentially expressed and alternatively spliced known/candidate genes), alleles, QTLs, epialleles, regulatory sequences and transcription factors controlling the complex quantitative yield and abiotic stress tolerance traits.
- Use such functionally relevant molecular tags regulating the complex quantitative traits in individual or integrated approach of genomics-assisted breeding (traditional bi-parental linkage/ QTL mapping, fine mapping/positional cloning, whole genome and candidate gene-based association mapping and genetical genomics/eQTLs): Obtain significant inputs and utilize these in traditional and modern MAB to develop superior high-yielding stress-tolerant crop varieties.

Acknowledgements

The authors gratefully acknowledge the financial support by the Department of Biotechnology (DBT), India. AK acknowledges the Council of Scientific and Industrial Research (CSIR) for the award of Junior Research Fellowship (JRF). We wish to thank the anonymous referees for critically evaluating the manuscript and providing constructive comments/suggestions to improve its quality.

References

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

Akimoto K, Katakami H, Kim HJ, Ogawa E, Sano CM, Wada Y and Sano H 2007 Epigenetic inheritance in rice plants. *Ann. Bot.* **100** 205–217

Anderson PK, Cunningham AA, Patel NG, Morales FJ, Epstein PR and Daszak P 2004 Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends Ecol. Evol.* **19** 535–544

Baroux C, Raissig MT and Grossniklaus U 2011 Epigenetic regulation and reprogramming during gamete formation in plants. *Curr. Opin. Genet. Dev.* **21** 124–133

Basavaraj SH, Singh VK, Singh A, Singh A, Singh A, Yadav S, Ellur RK, Singh D, *et al.* 2010 Marker-assisted improvement of bacterial blight resistance in parental lines of Pusa RH10, a superfine grain aromatic rice hybrid. *Mol. Breed.* **2** 293–305

Bernstein BE, Meissner A and Lander ES 2007 The mammalian epigenome. *Cell* **128** 669–681

Boerjan W and Vuylsteke M 2009 Integrative genetical genomics in *Arabidopsis*. *Nat. Genet.* **41** 144–145

Brenchley R, Spannagl M, Pfeifer M, Barker GL, D'Amore R, Allen AM, McKenzie N, Kramer M, *et al.* 2012 Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature* **491** 705–710

Chakraborty S 2013 Migrate or evolve: options for plant pathogens under climate change. *Glob. Chang. Biol.* **19** 1985–2000

Chia JM and Ware D 2011Sequencing for the cream of the crop. *Nat Biotechnol*. **29** 138–139

Chinnusamy V and Zhu JK 2009 Epigenetic regulation of stress responses in plants. *Curr. Opin. Plant Biol.* **12** 133–139

Choudhary S, Gaur R and Gupta S 2012 EST-derived genic molecular markers: development and utilization for generating an advanced transcript map of chickpea. *Theor. Appl. Genet*. **124** 1449–1462

Choudhary S, Sethy NK, Shokeen B and Bhatia S 2009 Development of chickpea EST-SSR markers and analysis of allelic variation across related species. *Theor. Appl. Genet*. **118** 591–608

Clark R, Maccurdy R, Jung J, Shaff J, McCouch SR, Aneshansley D and Kochian L 2011 3-dimensional root phenotyping with a novel imaging and software platform. *Plant Physiol.* **156** 455–465

Coakley SM, Scherm H and Chakraborty S 1999 Climate change and plant disease management. *Annu. Rev. Phytopathol.* **37** 399–426

Cubas P, Vincent C and Coen E 1999 An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* **401** 157–161

Du Z, Li H, Wei Q, Zhao X, Wang C, Zhu Q, Yi X, Xu W, *et al.* 2013 Genome-wide analysis of histone modifications: H3K4me2, H3K4me3, H3K9ac, and H3K27ac in *Oryza sativa* L. *Japonica*. *Mol. Plant* **6** 1463–1472

Eckardt NA 2009 Deep sequencing maps the maize epigenomic landscape. *Plant Cell* **21** 1024–1026

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

Emilsson V, Thorleifsson G, Zhang B, Leonardson AS, Zink F, Zhu J, Carlson S, Helgason A, *et al.* 2008 Genetics of gene expression and its effect on disease. *Nature* **452** 423–428

Famoso AN, Zhao K, Clark RT, Tung CW, Wright MH, Bustamante C, Kochian LV and McCouch SR 2011 Genetic architecture of aluminum tolerance in rice (*Oryza sativa*) determined through genome-wide association analysis and QTL mapping. *PLoS Genet.* **7** e1002221

Fan C, Xing Y, Mao H, Lu T, Han B, Xu C, Li X and Zhang Q 2006 *GS3*, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theor. Appl. Genet.* **112** 1164–1171

Fukai E, Stougaard J and Hayashi M 2013 Activation of an endogenous retrotransposon associated with epigenetic changes in *Lotus japonicus*: A tool for functional genomics in legumes. *Plant Genome* **doi** 10.3835/plantgenome2013.04.0009

Garg R, Patel RK, Jhanwar S, Priya P, Bhattacharjee A, Yadav G, Bhatia S, Chattopadhyay D, *et al.* 2011b Gene discovery and tissue-specific transcriptome analysis in chickpea with massively parallel pyrosequencing and web resource development. *Plant Physiol.* **156** 1661–1678

Garg R, Patel RK, Tyagi AK and Jain M 2011a *De novo* assembly of chickpea transcriptome using short reads for gene discovery and marker identification. *DNA Res.* **18** 53–63

Garrett KA, Dendy SP, Frank EE, Rouse MN and Travers SE 2006 Climate change effects on plant disease: genomes to ecosystems. *Annu. Rev. Phytopathol.* **44** 489–509

Gaur PM, Jukanti AK and Varshney RK 2012a Impact of genomic technologies on chickpea breeding strategies. *Agronomy* **2** 199–221

Gaur R, Azam S, Jeena G, Khan AW, Choudhary S, Jain M, Yadav G, Tyagi AK, *et al.* 2012b High-throughput SNP discovery and genotyping for constructing a saturated linkage map of chickpea (*Cicer arietinum* L.). *DNA Res.* **19** 357–373

Gopalakrishnan S, Sharma RK, Rajkumar KA, Joseph M, Singh VP, Singh AK, Bhat KV, Singh NK, *et al*. 2008 Integrating marker-assisted background analysis with foreground selection for identification of superior bacterial blight resistant recombinants in Basmati rice. *Plant Breed.* **127** 131–139

Gregory PJ, Johnson SN, Newton AC and Ingram JSI 2009 Integrating pests and pathogens into the climate change/food security debate. *J. Exp. Bot.* **60** 2827–2838.

Gujaria N, Kumar A, Dauthal P, Dubey A, Hiremath P, Bhanu Prakash A, Farmer A, Bhide M, *et al*. 2011 Development and use of genic molecular markers (GMMs) for construction of a transcript map of chickpea (*Cicer arietinum* L.). *Theor. Appl. Genet.* **122** 1577–1589

Harvell CD, Mitchell CE, Ward JR, Altizer S, Dobson AP, Ostfeld RS and Samuel MD 2002 Climate warming and disease risks for terrestrial and marine biota. *Science* **296** 2158–2162

He G, Elling AA and Deng XW 2011 The epigenome and plant development. *Annu. Rev. Plant Biol.* **62** 411–435

He G, Zhu X, Elling AA, Chen L, Wang X, Guo L, Liang M, He H, *et al.* 2010 Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids. *Plant Cell* **22** 17–33

Hiremath PJ, Farmer A, Cannon SB, Woodward J, Kudapa H, Tuteja R, Kumar A, Bhanuprakash A, *et al*. 2011 Large-scale transcriptome analysis in chickpea (*Cicer arietinum* L.), an orphan legume crop of the semi-arid tropics of Asia and Africa. *Plant Biotechnol. J.* **9** 922–931

Hiremath PJ, Kumar A, Penmetsa RV, Farmer A, Schlueter JA, Chamarthi SK, Whaley AM, Carrasquilla-Garcia N, *et al*. 2012 Large-scale development of cost-effective SNP marker assays for diversity assessment and genetic mapping in chickpea and comparative mapping in legumes. *Plant Biotechnol. J.* **10** 716–732

Huang BE, George AW, Forrest KL, Kilian A, Hayden MJ, Morell MK and Cavanagh CR 2012a A multiparent advanced generation inter-cross population for genetic analysis in wheat. *Plant Biotechnol J.* **10** 826–839

Huang X, Feng Q, Qian Q, Zhao Q, Wang L, Wang A, Guan J, Fan D, *et al*. 2009 High-throughput genotyping by whole-genome resequencing. *Genome Res.* **19** 1068–1076

Huang X, Kurata N, Wei X, Wang ZX, Wang A, Zhao Q, Zhao Y, Liu K, *et al*. 2012b A map of rice genome variation reveals the origin of cultivated rice. *Nature* **490** 497–501

Huang X, Wei X, Sang T, Zhao Q, Feng Q, Zhao Y, Li C, Zhu C, *et al.* 2010 Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat Genet.* **42** 961–967

Huang X, Zhao Y, Wei X, Li C, Wang A, Zhao Q, Li W, Guo Y, *et al.* 2012c Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat Genet.* **44** 32–39

International Barley Genome Sequencing Consortium (IBGSC) 2012 A physical, genetic and functional sequence assembly of the barley genome. *Nature* **491** 711–716

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

Jain M 2012 Next-generation sequencing technologies for gene expression profiling in plants. *Brief. Funct. Genomics* **11** 63–70

Jain M, Misra G, Patel RK, Priya P, Jhanwar S, Khan AW, Shah N, Singh VK, *et al.* 2013 A draft genome sequence of the pulse crop chickpea (*Cicer arietinum* L.). *Plant J.* **74** 715–729

Jannink JL, Lorenz AJ and Iwata H 2010 Genomic selection in plant breeding: from theory to practice. *Brief Funct. Genomics* **9** 166–177

Jhanwar S, Priya P, Garg R, Parida SK, Tyagi AK and Jain M 2012 Transcriptome sequencing of wild chickpea as a rich resource for marker development. *Plant Biotechnol. J.* **10** 690–702

Jia J, Zhao S, Kong X, Li Y, Zhao G, He W, Appels R, Pfeifer M, *et al*. 2013 *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. *Nature* **496** 91–95

Joseph M, Gopala Krishnan S, Sharma RK, Singh VP, Singh AK, Singh NK and Mohapatra T 2004 Combining bacterial blight resistance and Basmati quality characteristics by phenotypic and molecular marker-assisted selection in rice. *Mol. Breed.* **13** 377–387

Kapazoglou A, Engineer C, Drosou V, Kalloniati C, Tani E, Tsaballa A, Kouri ED, Ganopoulos I, *et al.* 2012 The study of two barley type I-like MADS-box genes as potential targets of epigenetic regulation during seed development. *BMC Plant Biol.* **12** 166

Kashkush K, Feldman M, Levy AA 2003 Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nat. Genet.* **33** 102–106

Kashkush K, Khasdan V 2007 Large-scale survey of cytosine methylation of retrotransposons and the impact of readout transcription from long terminal repeats on expression of adjacent rice genes. *Genetics* **177** 1975–1985

Kharabian-Masouleh A, Waters DLE, Reinke RF, Ward R and Henry RJ 2012 SNP in starch biosynthesis genes associated with nutritional and functional properties of rice. *Sci. Rep.* **2** 557

Kikuchi S, Satoh K, Nagata T, Kawagashira N, Doi K, Kishimoto N, Yazaki J, Ishikawa M, *et al*. 2003 Rice Full-Length cDNA Consortium; Collection, mapping, and annotation of over 28,000 cDNA clones from *japonica* rice. *Science* **301** 376–379

Kim TY, Kim HU and Lee SY 2010 Data integration and analysis of biological networks. *Curr. Opin Biotechnol.* **21** 78–84

Kohler C, Wolff P and Spillane C 2012 Epigenetic mechanisms underlying genomic imprinting in plants. *Annu. Rev. Plant Biol.* **63** 18.1–18.22

Kota R, Varshney RK, Prasad M, Zhang H, Stein N and Graner A 2007 EST-derived single nucleotide polymorphism markers for assembling genetic and physical maps of the barley genome. *Funct. Integr. Genomics* **8** 223–233

Kujur A, Bajaj D, Saxena MS, Tripathi S, Upadhyaya HD, Gowda CL, Singh S, Jain M, *et al.* 2013 Functionally relevant microsatellite markers from chickpea transcription factor genes for efficient genotyping applications and trait association mapping. *DNA Res.* **20** 355–374

Langridge P and Fleury D 2011 Making the most of 'omics' for crop breeding. *Trends Biotechnol.* **29** 33–40

Li H, Freeling M and Lisch D 2010 Epigenetic reprogramming during vegetative phase change in maize. *Proc. Natl. Acad. Sci. USA* **107** 22184–22189

Li X, Wang X, He K, Ma Y, Su N, He H, Stolc V, Tongprasit W, *et al.* 2008 High-resolution mapping of epigenetic modifications of the rice genome uncovers interplay between DNA methylation, histone methylation, and gene expression. *Plant Cell* **20** 259–276

Li Y, Fan C, Xing Y, Jiang Y, Luo L, Sun L, Shao D, Xu C, *et al*. 2011 Natural variation in *GS5* plays an important role in regulating grain size and yield in rice. *Nat Genet.* **43** 1266–1269

Ling HQ, Zhao S, Liu D, Wang J, Sun H, Zhang C, Fan H, Li D, *et al.* 2013 Draft genome of the wheat A-genome progenitor *Triticum urartu*. *Nature* **496** 87–90

Lobell DB and Field CB 2007 Global scale climate-crop yield relationships and the impacts of recent warming. *Environ. Res. Lett. 2* **004000** 7 pp

Lu T, Lu G, Fan D, Zhu C, Li W, Zhao Q, Feng Q, Zhao Y, *et al*. 2010 Function annotation of the rice transcriptome at single-nucleotide resolution by RNA-seq. *Genome Res.* **20** 1238–1249

Manning K, Tör M, Poole M, Hong Y, Thompson AJ, King GJ, Giovannoni JJ and Seymour GB 2006 A naturally occurring epigenetic mutation in a gene encoding an *SBP*-box transcription factor inhibits tomato fruit ripening. *Nat Genet.* **38** 948–952

Mao H, Sun S, Yao J, Wang C, Yu S, Xu C, Li X and Zhang Q 2010 Linking differential domain functions of the *GS3* protein to natural variation of grain size in rice. *Proc. Natl. Acad. Sci. USA* **107** 19579–19584

McCouch SR, Sweeney M, Li J, Jiang H, Thomson M, Septiningsih E, Edwards J, Moncada P, *et al*. 2007 Through the genetic bottleneck: *O. rufipogon* as a source of trait-enhancing alleles for *O. sativa*. *Euphytica* **154** 317–339

McCouch SR, Teytelman L, Xu Y, Lobos KB, Clare K, Walton M, Fu B, Maghirang R, *et al.* 2002 Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). *DNA Res.* **9** 199–207

McNally KL, Bruskiewich R, Mackill D, Buell CR, Leach JE and Leung H 2006 Sequencing multiple and diverse rice varieties: connecting whole-genome variation with phenotypes. *Plant Physiol.* **141** 26–31

McNally KL, Childs KL, Bohnert R, Davidson RM, Zhao K, Ulat VJ, Zeller G, Clark RM, *et al*. 2009 Genome-wide SNP variation reveals relationships among landraces and modern varieties of rice. *Proc. Natl. Acad. Sci. USA* **106** 12273–12278

Metzker ML 2010 Sequencing technologies - the next generation. *Nat. Rev. Genet.* **11** 31–46

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

Mir RR, Zaman-Allah M, Sreenivasulu N, Trethowan R and Varshney RK 2012 Integrated genomics, physiology and breeding approaches for improving drought tolerance in crops. *Theor. Appl. Genet.* **125** 625–645

Mirouze M and Paszkowski J 2011 Epigenetic contribution to stress adaptation in plants. *Curr. Opin. Plant Biol.* **14** 267–274

Mittler R 2006 Abiotic stress, the field environment and stress combination. *Trends Plant Sci.* **11** 15–19

Miura K, Ashikari M and Matsuoka M 2011 The role of QTLs in the breeding of high-yielding rice. *Trends Plant Sc*i. **16** 319–326

Mukhopadhyay P, Singla-Pareek SL, Reddy MK and Sopory SK 2013 Stress-mediated alterations in chromatin architecture correlate with down-regulation of a gene encoding 60S rpL32 in rice. *Plant Cell Physiol.* **54** 528–540

Nayak SN, Zhu H, Varghese N, Datta S, Choi HK, Horres R, Jüngling R, Singh J, *et al*. 2010 Integration of novel SSR and gene-based SNP marker loci in the chickpea genetic map and establishment of new anchor points with *Medicago truncatula* genome. *Theor. Appl. Genet.* **120** 1415–1441

Negrão S, Almadanim MC, Pires IS, Abreu IA, Maroco J, Courtois B, Gregorio GB, McNally KL, *et al.* 2013 New allelic variants found in key rice salt-tolerance genes: an association study. *Plant Biotechnol. J.* **11** 87–100

Oliver SN, Finnegan EJ, Dennis ES, Peacock WJ and Trevaskis B 2009 Vernalization-induced flowering in cereals is associated with changes in histone methylation at the *VERNALIZATION1* gene. *Proc. Natl. Acad. Sci. U S A* **106** 8386–8391

Pachauri RK and Reisinger A 2007 Intergovernmental Panel on Climate Change (IPCC); in *Climate change synthesis report* (eds) IPCC, Geneva

Parida SK, Dalal V, Singh NK and Mohapatra T 2009b Genic non-coding microsatellites in the rice genome: characterization, marker design and use in assessing genetic and evolutionary relationships among domesticated groups. *BMC Genomics* **10** 140

Parida SK, Kalia SK, Kaul S, Dalal V, Hemaprabha G, Selvi A, Pandit A, Singh A, *et al*. 2009a Informative genomic microsatellite markers for efficient genotyping applications in sugarcane. *Theor. Appl. Genet.* **118** 327–338

Parida SK, Mukerji M, Singh AK, Singh NK and Mohapatra T 2012 SNPs in stress-responsive rice genes: validation, genotyping, functional relevance and population structure. *BMC Genomics* **13** 426

Parida SK, Rajkumar KA, Dalal V, Singh NK and Mohapatra T 2006 Unigene derived microsatellite markers for the cereal genomes. *Theor. Appl. Genet.* **112** 808–817

Parida SK, Yadava DK and Mohapatra T 2010 Microsatellites in *Brassica* unigenes: Relative abundance, marker design and use in comparative physical mapping and genome analysis. *Genome* **53** 55–67

Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, *et al.* 2009 The *Sorghum bicolor* genome and the diversification of grasses *Nature* **457** 551–556

Pautasso M, Döring TF, Garbelotto M, Pellis L and Jeger MJ 2012 Impacts of climate change on plant diseases opinions and trends. *Eur. J. Plant Pathol.* **133** 295–313

Pfender WF and Vollmer SS 1999 Freezing temperature effect on survival of *Puccinia graminis* subsp. *Graminicola* in *Festuca arundinacea* and *Lolium perenne*. *Plant Disease* **83** 1058–1062

Picoult-Newberg L, Ideker TE, Pohl MG, Taylor SL, Donaldson MA, Nickerson DA and Boyce-Jacino M 1999 Mining SNPs from EST databases. *Genome Res.* **9** 167–174

Potato Genome Sequencing Consortium (PGSC) 2011 Genome sequence and analysis of the tuber crop potato. *Nature* **475** 189–195

Raghavan C, Naredo MEB, Wang HH, Atienza G, Liu B, Qiu FL, McNally KL and Leung H 2007 Rapid method for detecting SNPs on agarose gels and its application in candidate gene mapping. Mol Breed. **19** 87–101

Redden R 2013 New approaches for crop genetic adaptation to the abiotic stresses predicted with climate change. *Agronomy* **3** 419–432

Reynolds MP and Ortiz R 2010 Adapting crops to climate change: a summary; in *Climate change and crop production* (eds) MP Reynolds (CAB international) 1–8 pp

Rodrigues JA, Ruan R, Nishimura T, Sharma MK, Sharma R, Ronald PC, Fischer RL and Zilberman D 2013 Imprinted expression of genes and small RNA is associated with localized hypomethylation of the maternal genome in rice endosperm. *Proc. Natl. Acad. Sci. U S A* **110** 7934–7939

Schmitz RJ, He Y, Valdés-López O, Khan SM, Joshi T, Urich MA, Nery JR, Diers B, *et al.* 2013 Epigenome-wide inheritance of cytosine methylation variants in a recombinant inbred population. *Genome Res.* **23** 1663–1674

Schmitz RJ, Schultz MD, Urich MA, Nery JR, Pelizzola M, Libiger O, Alix A, McCosh RB, *et al.* 2013 Patterns of population epigenomic diversity. *Nature* **495** 193–198

Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, *et al.* 2010 Genome sequence of the palaeopolyploid soybean. *Nature* **463** 178–183

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

Septiningsih EM, Pamplona AM, Sanchez DL, Neeraja CN, Vergara GV, Heuer S, Ismail AM and Mackill DJ 2009 Development of submergence-tolerant rice cultivars: the *Sub1* locus and beyond. *Ann. Bot.* **103** 151–160

Sethy NK, Shokeen B, Edwards KJ and Bhatia S 2006 Development of microsatellite markers and analysis of intra-specific genetic variability in chickpea (*Cicer arietinum* L.). *Theor. Appl. Genet.* **112** 1416–1428

Shitsukawa N, Tahira C, Kassai K, Hirabayashi C, Shimizu T, Takumi S, Mochida K, Kawaura K, *et al.* 2007 Genetic and epigenetic alteration among three homoeologous genes of a class E *MADS* box gene in hexaploid wheat. *Plant Cell* **19** 1723–1737

Simon SA, Zhai J, Nandety RS, McCormick KP, Zeng J, Mejia D and Meyers BC 2009 Short-read sequencing technologies for transcriptional analyses. *Annu. Rev. Plant Biol.* **60** 305–333

Singh AK, Gopalakrishnan S, Singh VP, Prabhu KV, Mohapatra T, Singh NK, SharmaTR, Nagarajan M, *et al.* 2011 Marker assisted selection: a paradigm shift in Basmati breeding. *Indian J. Genet.* **71** 120–128

Sundaram RM, Vishnupriya MR, Biradar SK, Laha GS, Reddy GA, Shoba Rani N, Sarma NP and Sonti RV 2008 Marker assisted introgression of bacterial blight resistance in Samba Mahsuri, an elite *indica* rice variety. *Euphytica* **160** 411–422

Sweeney MT, Thomson MJ, Cho YG, Park YJ, Williamson SH, Bustamante CD and McCouch SR 2007 Global dissemination of a single mutation conferring white pericarp in rice. *PLoS Genet.* **3** e133

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

Tan L, Zhang P, Liu F, Wang G, Ye S, Zhu Z, Fu Y, Cai H, *et al.* 2008 Quantitative trait loci underlying domestication- and yield-related traits in an *Oryza sativa* x *Oryza rufipogon* advanced backcross population. *Genome* **51** 692–704

Tardieu F and Tuberosa R 2010 Dissection and modelling of abiotic stress tolerance in plants. *Curr. Opin. Plant Biol.* **13** 206–212

Terpstra IR, Snoek LB, Keurentjes JJ, Peeters AJ and van den Ackerveken G 2010 Regulatory network identification by genetical genomics: signaling downstream of the *Arabidopsis* receptor-like kinase ERECTA. *Plant Physiol.* **154** 1067–1078

The *Brassica rapa* Genome Sequencing Project Consortium (BGSPC) 2011 The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* **43** 1035–1039

Tian F, Li DJ, Fu Q, Zhu ZF, Fu YC, Wang XK and Sun CQ 2006 Construction of introgression lines carrying wild rice (*Oryza rufipogon* Griff.) segments in cultivated rice (*Oryza sativa* L.) background and characterization of introgressed segments associated with yield-related traits. *Theor. Appl. Genet.* **112** 570–580

Till BJ, Cooper J, Tai TH, Colowit P, Greene EA, Henikoff S and Comai L 2007 Mutant resources in rice discovery of chemically induced mutations in rice by TILLING. *BMC Plant Biol.* **7** 19

Tomato Genome Consortium (TGC) 2012 The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* **485** 635–641

Upadhyaya HD and Ortiz R 2001 A mini core subset for capturing diversity and promoting utilization of chickpea genetic resources in crop improvement. *Theor. Appl. Genet.* **102** 1292–1298

Upadhyaya HD, Bramel PJ and Singh S 2001 Development of a chickpea core subset using geographic distribution and quantitative traits. *Crop Sci.* **41** 206–210

Upadhyaya HD, Ortiz R, Bramel PJ and Singh S 2002 Phenotypic diversity for morphological and agronomic characteristics in chickpea core collection. *Euphytica* **123** 333–342

Vahdati K and Leslie C 2013 Abiotic stress - plant responses and applications in agriculture (eds) InTech **doi** 10.5772/45842

Varshney RK, Chen W, Li Y, Bharti AK, Saxena RK, Schlueter JA, Donoghue MT, Azam S, *et al.* 2012 Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. *Nat. Biotechnol.* **30** 83–89

Varshney RK, Close TJ, Singh NK, Hoisington DA and Cook DR 2009a Orphan legume crops enter the genomics era. *Curr. Opin. Plant Biol.* **12** 202–210

Varshney RK, Graner A and Sorrells ME 2005 Genic microsatellite markers in plants: features and applications. *Trends Biotechnol.* **23** 48–55

Varshney RK, Hiremath PJ, Lekha PT, Kashiwagi J, Jayasree B, Deokar AA, Vadez V, Xiao Y, *et al.* 2009b A comprehensive resource of drought-and salinity-responsive ESTs for gene discovery and marker development in chickpea (*Cicer arietinum* L). *BMC Genomics* **10** 523

Varshney RK, Nayak SN, May GD and Jackson SA 2009c Next-generation sequencing technologies and their implications for crop genetics and breeding. *Trends Biotechnol.* **27** 522–530

Varshney RK, Song C, Saxena RK, Azam S, Yu S, Sharpe AG, Cannon S, Baek J, *et al.* 2013 Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. *Nat. Biotechnol.* **31** 240–246

Wang C, Chen S and Yu S 2011 Functional markers developed from multiple loci in *GS3* for fine marker-assisted selection of grain length in rice. *Theor. Appl. Genet.* **122** 905–913

Wang E, Wang J, Zhu X, Hao W, Wang L, Li Q, Zhang L, He W, *et al.* 2008 Control of rice grain-filling and yield by a gene with a potential signature of domestication. *Nat Genet.* **40** 1370–1374

Wang X, Elling AA, Li X, Li N, Peng Z, He G, Sun H, Qi Y, *et al.* 2009 Genome-wide and organ-specific landscapes of epigenetic modifications and their relationships to mRNA and small RNA transcriptomes in maize. *Plant Cell* **21** 1053–1069

Waters AJ, Makarevitch I, Eichten SR, Swanson-Wagner RA, Yeh CT, Xu W, Schnable PS, Vaughn MW, *et al.* 2011 Parent-of-origin effects on gene expression and DNA methylation in the maize endosperm. *Plant Cell* **23** 4221–4233

Waugh MM, Kim DH, Ferrin DM and Stanghellini ME 2003 Reproductive potential of *Monosporascus cannonballus*. *Plant Disease* **87** 45–50

Wu JL, Wu C, Lei C, Baraoidan M, Bordeos A, Madamba MR, Ramos-Pamplona M, Mauleon R, *et al.* 2005 Chemical and irradiation-induced mutants of *indica* rice IR64 for forward and reverse genetics. *Plant Mol. Biol.* **59** 85–97

Xie W, Feng Q, Yu H, Huang X, Zhao Q, Xing Y, Yu S, Han B, *et al.* 2010 Parent-independent genotyping for constructing an ultrahigh-density linkage map based on population sequencing. *Proc. Natl. Acad. Sci. USA* **107** 10578–10583

Xu Y, Lu Y, Xie C, Gao S, Wan J and Prasanna B 2012 Whole-genome strategies for marker-assisted plant breeding. *Mol. Breed.* **29** 833–854

Young ND, Debellé F, Oldroyd GE, Geurts R, Cannon SB, Udvardi MK, Benedito VA, Mayer KF, *et al.* 2011 The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature* **480** 520–524

Yu H, Xie W, Wang J, Xing Y, Xu C, Li X, Xiao J and Zhang Q 2011 Gains in QTL detection using an ultra-high density SNP map based on population sequencing relative to traditional RFLP/SSR markers. *PLoS One* **6** e17595

Zavala JA, Casteel CL, Delucia EH and Berenbaum MR 2008 Anthropogenic increase in carbon dioxide compromises plant defense against invasive insects. *Proc. Natl. Acad. Sci. USA* **105** 5129–5133

Zemach A, Kim MY, Silva P, Rodrigues JA, Dotson B, Brooks MD and Zilberman D 2010 Local DNA hypomethylation activates genes in rice endosperm. *Proc. Natl. Acad. Sci. U S A* **107** 18729–18734

Zhang H, Zhang D, Wang M, Sun J, Qi Y, Li J, Wei X, Han L, *et al.* 2011a A core collection and mini core collection of *Oryza sativa* L. in China. *Theor. Appl. Genet.* **122** 49–61

Zhang M, Zhao H, Xie S, Chen J, Xu Y, Wang K, Zhao H, Guan H, *et al.* 2011b Extensive, clustered parental imprinting of protein-coding and noncoding RNAs in developing maize endosperm. *Proc. Natl. Acad. Sci. U S A* **108** 20042–20047

Zhang X 2008 The epigenetic landscape of plants. *Science* **320** 489–492

Zhang Z, Deng Y, Tan J, Hu S, Yu J and Xue Q 2007 A genome-wide microsatellite polymorphism database for the *indica* and *japonica* rice. *DNA Res*. **14** 37–45

Zhao K, Tung CW, Eizenga GC, Wright MH, Ali ML, Price AH, Norton GJ, Islam MR, *et al.* 2011 Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Commun.* **2** 467

Zhong S, Fei Z, Chen YR, Zheng Y, Huang M, Vrebalov J, McQuinn R, Gapper N, *et al.* 2013 Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. *Nat. Biotechnol.* **31** 154

Zhu JK 2008 Epigenome sequencing comes of age. *Cell* **133** 395–397

MS received 27 December 2012; accepted 07 October 2013

Corresponding editor: RENU KHANNA-CHOPRA