Research Review Article

Indian Journal of Agricultural Sciences **81** (9): 787–800, September 2011

Biotechnology and crop improvement

P ANANDA KUMAR[1], TRILOCHAN MOHAPATRA[2], TILAK R SHARMA[3], RAMCHARAN BHATTACHARYA[4], PRASANTA K DASH[5], NAVIN C GUPTA[6] and AMOL KUMAR U SOLANKE[7]

National Research Centre on Plant Biotechnology, New Delhi 110 012

Received: 12 January 2011; Revised accepted: 20 May 2011

ABSTRACT

Conventional plant breeding is the backbone of agricultural development. It has very significantly contributed in the past to genetic enhancement of crops, particularly for breeding high-yielding crop cultivars. The quantum jump in agricultural productivity which was achieved during late sixties and early seventies needs further enhancement to ensure food and nutritional security of the growing population. Advances in modern biology, especially biotechnology, offer many advantages over traditional techniques of plant breeding. The applications of biotechnology in crop improvement can be broadly grouped into three categories, viz precise isolation and deployment of genes, irrespective of source and target genome, marker-assisted selections and large throughput characterization of genome, transcriptome, proteome or metabolome. The most compelling advantage of plant biotechnology is the ability to transfer foreign genes to confer novel traits. An entire array of traits viz. insect pest and pathogen resistance, abiotic stress tolerance, herbicide tolerance, augmentation of nutritional qualities etc. have been successfully achieved by plant transformation. Another significant application of biotechnology in crop improvement has been 'marker-assisted selention (MAS). Development and integration of DNA-based molecular markers in the selection process has empowered the breeder to identify desired genotype without any interference of environmental effect of tissue specificity of expression. High throughout genomics emerged as a promising area in crop biotechnology programmes. This is because most of the commercially relevant plant traits are interaction of large number of genes, their positions on chromosomes and promoters controlling them. While structural genomics deals with sequence analysis of total genetic information in an organism, efforts in functional genomics are directed to unravel and understand the mechanism by which this information is used by an organism. Systematic study of complete repertoire of metabolites/chemicals of any organism has given birth to a new area of research 'metabolomics'. Integration of genomics and proteomics with metabolomics will enrich our understanding to gene-function relationship that can be utilized in achieving crop improvement towards higher productivity.

Key words: Abiotic stress, Bioinformatics, Biotechnology, Crop improvement, Gene cloning, Gene isolation, Insertional mutagenesis, Marker-assisted selection, Molecular markers, Transposon tagging

Progress in science and technology has laid the foundation of social and economic gains made in agriculture over the past 40 years. The world's population is expected to reach 8.0 billion by 2025, making food security the most important social issue. Food production will have to be doubled to meet the growing needs of the ever increasing population, 90% of which reside in the developing world. The dwindling water and land resources will further exacerbate the enormity of this challenge. In addition, crop losses due to insect pests, diseases and declining soil fertility will worsen because of climatic conditions that favour insect pests and disease vectors.

[1]Project Director (e mail: kumarpa@nrcpb.org); [2]Principal Scientist (e mail: tm@nrcpb.org); [3]Principal Scientist (e mail: trsharma@nrcpb.org); [4]Principal Scientist (e mail: rcb@nrcpb.org; [5]Scientist (Senior Scale) (email: prasanta01@yahoo.com); [6]Scientist (e mail: guptanc@nrcpb.org); [7]Scientist (e mail: amolsgene@gmail.com)

The most compelling case for biotechnology, and more specifically transgenic crops, is their capability to contribute for: (*i*) increasing crop productivity, and thus contribute to global food, feed and fibre security, (*ii*) lowering production costs, (*iii*) conserving biodiversity, as a land-saving technology capable of higher productivity, (*iv*) more efficient use of external inputs, for a more sustainable agriculture and environment, (*v*) increasing stability of production to lessen suffering during famines due to abiotic and biotic stresses, and (*vi*) provision of economic and social benefits and alleviation of abject poverty in developing countries.

Conventional breeding has very significantly contributed to genetic enhancement of crops, particularly for higher yield using the available variation in the germplasm. However, some of the limitations of conventional breeding are: (*i*) lack of germplasm resources possessing the desired traits within the crossable limits, (*ii*) problems of linkage drag while

sourcing genes from the wild relatives, (*iii*) a limited understanding of the biochemical/metabolic pathways for a systematic and precise improvement of nutritional qualities in major crops, (*iv*) phenotype-based selection that might not correctly reflect the genotypic worth, and (*v*) plant-environment interactions that negatively affect the conventional selection process. Advances in modern biology, especially biotechnology, offer many advantages over traditional techniques of plant breeding. Biotechnology offers possible solutions to previously intractable problems and difficult targets such as drought tolerance. Some of the above problems could be addressed through molecular marker-assisted breeding, while others essentially require isolation of genes and their transfer across sexual barriers through genetic engineering thereby making the whole living world a single gene pool. The applications of new tools of biotechnology will be one of the key factors to achieve the goals of crop improvement.

Plant biotechnology can be divided into three main areas: (*i*) plant genetic engineering (gene isolation and transgenic crops), (*ii*) molecular breeding (marker-aided selection), and (*iii*) genomics (genomics, metabolomics and bioinformatics). Plant genetic engineering deals with basic molecular biology, isolation of genes and their introduction into various plant species to develop novel genotypes. Molecular breeding encompasses the utilization of molecular tools in plant breeding programmes via construction of linkage maps, character tagging, marker-assisted selection and map-based cloning of genes. Genomics deals with the structural and functional dimensions of plant/crop genetic material.

Gene isolation

Gene isolation is an essential step to understand molecular control of fundamental plant processes that will facilitate their modification and use in crop improvement through genetic engineering. In plants, the genes which govern the genetic basis of development and stress-response will provide opportunities to optimize agricultural productivity in terms of increased yield, decreased negative environmental consequences and reduced ecological repercussions. Gene isolation can be followed by two broad strategies, the reverse genetics (from gene to phenotype) and the forward genetics (from phenotype to gene). Some of the commonly used approaches under these two broad strategies include screening of genomic/cDNA libraries, differential expression profiling, functional complementation, insertional mutagenesis and positional cloning of genes.

Genomic/cDNA library screening

Genomic library is made by cloning genomic DNA fragments obtained by either restriction endonuclease digestion or mechanical shearing. The library represents the genome regardless of the cell type or developmental stage from which the DNA is isolated. The cDNA (complementary DNA) libraries are generated by cloning DNA obtained by reverse transcribing the cellular mRNA. It thus represents the genes that are transcribed in a tissue from which mRNAs are derived. Plant genes are isolated from genomic and cDNA libraries by screening of clones with an appropriate hybridization probe derived from purified RNA, a predetermined/known sequence from a heterologous system or differential hybridization.

Besides, variations on normal hybridization-based library screening protocol include subtracted cDNA libraries preparations, which are enriched with differentially expressed cDNA clones. A recent resurgence of differential screening has come in the form of cDNA array, suppression subtractive hybridization (SSH) and representational difference analysis (RDA). RDA approaches have been recently used to tag and clone *Bph1* locus in rice for resistance to brown plant hopper *Nilaparvata lugens* Stål (Park *et al*. 2008). In cDNA microarrays, cDNA clones are transferred to a miniature solid support in a dense grid pattern (cDNA chip), and screened simultaneously with complex probes from two sources, which are labeled with two different fluorochromes.

Differential display

Identification of differentially expressed genes in various cells or under different environments is one of the core areas of molecular biology. Until 1992, subtractive hybridization was the only method by which differentially expressed genes were isolated. Although subtractive hybridization is a reliable method, it is tedious, time consuming and difficult to perform. It also requires large amounts of mRNA that can be limiting in many situations. Liang and Pardee (1992) developed a new PCR-based technique called Differential Display (DDRT-PCR). The method is based on detection of the differentially expressed cDNAs from two or more samples that are separated on adjacent lanes in sequencing gels. The differentially expressed bands can readily be cloned and used as probes in northern or southern (DNA) blots and to isolate genes from cDNA or genomic libraries.

A number of technical modifications have been made to reduce the problem of false positives and to increase the reproducibility of the technique (Huang *et al.* 1996, Jones *et al.* 1997, Zhao *et al.* 1995, Doss 1996). Modifications that allowed the display of longer cDNAs have also been reported (Averboukh *et al.* 1996). Therefore, with properly designed primers and controls, DDRT-PCR results will truly reflect gene expression patterns of different tissues.

Functional cloning or functional complementation

The alternative approach in isolating plant genes involves the exploitation of the biochemical or physiological activity of the gene product through production and analysis of mutants generated by ultraviolet light, ethyl methane sulphonate and X-rays. However, although a variety of

procedures have been used to produce mutant plants and despite a long history of mutant selection, there are relatively few examples where there is an understanding of the primary molecular defect from which a mutant phenotype arises. Where the molecular basis of a mutation is understood there is the potential of rescuing genes by complementation. However, in plants, this powerful technique has had only limited success and complementation has been reported only for genes that had previously been characterized. For example, an auxotrophic mutant of *Nicotiana plumbaginifolia* requiring isoleucine has been complemented by a yeast gene encoding threonine dehydratase (Colau *et al.* 1987). Another example has been provided by the complementation of a flower pigment mutation of *Petunia hybrida* with a maize gene previously cloned. In this case, the maize gene was able to catalyze the conversion of an intermediate metabolite in anthocyanin biosynthesis which accumulated in the mutant plant thus changing the flower colour of the plant from white to brick red (Meyer *et al.* 1987). Both these examples reveal the power of complementation but do not address the question of isolating genes by this method.

Insertional mutagenesis

The use of insertional mutagenesis in principle provides a more rapid way to clone a mutated gene. DNA elements such as transposons or the T-DNA of *Agrobacterium tumefacians* (Azpiroz-Leehan and Feldmann 1997) that are able to insert in random manner within chromosomes, can be used as mutagens to create loss of function mutations in plants. T-DNA and transposon tagging (Gierl and Saedler 1992) have been used widely for isolation of genes. This approach has the additional advantage that an altered phenotype might provide clues to the function of the product of the gene in question as well as facilitating its isolation. In this tagging approach, first the mutation caused by the insertion sequence has to be identified through phenotypic screening. Second, cosegregation of the insertion sequence with the mutant phenotype has to be verified. Subsequently, the gene can be isolated molecularly by cloning the DNA sequences flanking the insertion element. The most commonly used techniques to clone the sequences flanking the insertion element from the mutant plants are thermal asymmetric interlaced PCR (TAIL) (Liu *et al.* 1995), plasmid rescue (Yanofsky *et al.* 1990, Nakazawa *et al.* 2001) and inverse PCR (IPCR) (Ochman *et al.* 1988).

Transposon tagging

Transposable element systems have been used as tags in their host plants (*Ac/Ds*, *En/Spm*, *Mu* in maize; *Tam* elements in *Antirrhinum majus*), and in transgenic heterologous plant species (Kunze *et al.* 1997). Since the sequence of the inserted element is known, the gene in which the insertion has occurred can be recovered, using various cloning or PCR-based strategies. The maize Ac/Ds and En/Spm transposable elements have been developed for use in heterologous species. The behaviour of these elements has been extensively studied, and they have been modified for efficient transposition in plants such as tobacco, tomato and *Arabidopsis* (Osborne *et al.* 1995). Recently, the endogenous retrotransposon Tos17 has also been shown to be an efficient insertional mutagen in rice. Considering the ease of mutagenesis with Tos17 and its multiple-copy nature, saturation mutagenesis with this retrotransposon clearly demonstrated that the Tos17 system can contribute to the functional genomics of rice. Significantly, several important genes have been cloned using Tos17 including a new class of gene involved in brassinosteroid-signalling (Yamazaki *et al.* 2001), the cellulose synthase gene (Tanaka *et al.* 2003), a GH3-like gene (Komatsu *et al.* 2003) and the zeaxanthin epoxidase gene (Agrawal *et al.* 2001), which were responsible for the semi-dwarf, brittle culm, open glume, and viviparous phenotypes, respectively.

T-DNA tagging

The T-DNA segment of the Ti-plasmid of the soil bacterium *Agrobacterium tumefacience* when transferred to plant cell, not only disrupts the expression of the gene into which it is inserted, but also acts as a marker for subsequent identification of the mutation. If a large population of insertion mutant lines is generated, there are reasonably good chances of finding a plant carrying a T-DNA insert within any gene of interest. One of the advantages of using this approach is their low copy number and random nature of insertions. The original root explant method of Valvekens *et al.* (1988) allowed one to isolate a few transformed plants, via a laborious tissue culture process. Tens of thousands of transformed plants were beyond reach, until Feldmann and Marks (1987) devised a method for producing independent T-DNA transgenic lines via seed transformation. The development of transformation methods based on dipping floral bud into *Agrobacterium* suspensions has made it possible to generate hundreds of thousands of insertional mutations necessary for saturation of the genome while minimizing the effect of somaclonal variation associated with the process of *in vitro* culture and regeneration (Clough and Bent 1998).

Different methods that use T-DNA as tag for gene isolation differ depending on the reporter gene construct that is used. For example, activation trap, enhancer trap and gene/promoter trap vectors are specialized versions of insertional mutagens. Reporter genes can be used to construct three basic types of traps: gene trap, enhancer trap and promoter trap. Each type is able to respond to *cis*-acting regulatory sequences at the site of insertion. These gene trap reporter constructs can be delivered into the plant genomes either by T-DNAs or transposable elements and their expression not affected if reporter gene positioned at either the left (Lindsey *et al.* 1993) or right (Campisi *et al.* 1999) border of the T-DNA.

Table 1 *Bt* cotton events approved for cultivation

Event	Event no.	Company/ institution	Genes	Year
Bollgard I	Mon 531	Monsanto	*cry1Ac*	2002
Bollgard II	Mon 15985	Monsanto	*cry1Ac+cry2Ab*	2006
Event 1	Event 1	IIT, Kharagpur	*cry1Ac*	2006
GFM Cry1A	GFM Cry1A	Chinese Acad. Sci., China	*cry1Ab::cry1Ac*	2006
BNLA 601	BNLA 601	UAS, Dharwad / CICR, Nagpur	*cry1Ac*	2008
Event 9124	Event 9124	Metahelix, Bangalore	*cry1C*	2009

Map-based/positional gene cloning

This is another forward genetics approach of the gene isolation where markers on the genetic map are used for gene cloning. The identification of a gene affected by natural variation or through chemically or radiation-induced mutation requires map-based cloning (MBC) approach in which markers linked to the mutated gene are used to define the region containing the gene of interest. The rapid advances in sequencing projects, the abundance of various marker systems and the progress made in methods to detect DNA polymorphisms facilitate fast map-based cloning (MBC) of genes, although this is presently true only for a few of the plant species where sequence information and markers are available such as *Arabidopsis* and rice. For example, in case of rice, six bacterial blight resistance genes against *Xantomonas oryzae* pv. oryzae, *Xa1*, *xa5*, *xa13*, *Xa21*, *Xa26* and *Xa27* (Chu *et al.* 2006, Iyer and McCouch 2004, Sun *et al.* 2004, Gu *et al.* 2005) have been successfully cloned by using the map base cloning approach. With the recent advancement in genome sequencing projects, we will have more sequence and marker availability in future which will facilitate gene cloning by this method in other crop plants.

Transgenic crops: Empowering the geneticists and plant breeders to address the urgency of productivity enhancement, the techniques of plant transformation have witnessed a phenomenal development over the years and now transformation protocols are available for all most all the major crops. The global area of genetically improved agricultural crops has seen a tremendous increase since its first adoption in 1996 to 148 million hectares in 2010 (James 2010). A variety of traits have been introduced in plant species which include: herbicide resistance, pest resistance, viral resistance, slow-ripening, fungal and bacterial resistance, abiotic stress tolerance (drought, salinity, temperature etc), quality improvement (starch, protein and oil), value-addition (vitamins, micro- and macro-elements), pharmaceutical and therapeutic proteins, edible vaccines, industrial importance and phytoremediation.

The first commercial transgenic crop variety 'FlavrSavr' tomato was released in 1994 and was engineered for slow-ripening character. The first generation commercial transgenic crops and traits were mostly relevant to the developed world and were driven mostly by profit motive. A list of transgenic plant species that have been commercialized in the past 16 years is given below (http://www.isb.vt.edu/):

(i) Herbicide resistance (canola, soybean, cotton, rice, wheat, carnation, chicory, corn, sunflower, tobacco and sugarbeet)
(ii) Insect Pest resistance (cotton, corn, rice, tomato and potato)
(iii) Viral resistance (papaya, squash, plum and potato)
(iv) Slow-ripening and softening (tomato and melon)
(v) Improved oil quality (canola and soybean)
(vi) Male sterility (canola and corn)
(vii) Pigmentation pattern (carnation)
(viii) Reduced nicotine content (tobacco)

Status of transgenic crops in India: Efforts are being made in India since early eighties to develop transgenic crops. The first transgenic crop commercialized in India was *Bt* cotton in 2002. Currently, six transgenic events of *Bt* cotton (>1 000 hybrids) are available and the success story of Bt cotton cultivation has been reviewed (Karihaloo and Kumar 2009). Fruit borer resistant *Bt* brinjal hybrids developed by Mahyco and varieties developed by UAS, Dharwad and TNAU, Coimbatore have been cleared by GEAC with respect to environmental safety and biosafety. The *Bt* cotton events approved for cultivation in India is listed in Table 1.

Virus resistance

Incorporation of resistance against plant viruses is another important area in crop improvement for which we are heavily dependent on biotechnological intervention. Expression of viral genes encoding coat protein, non structural proteins (replicase, movement protein), and use of antisense technology are some of the strategies that have been effectively used in plants to confer resistance against viral diseases (Beachy *et al.* 1990). The transgenic expression of viral structural protein stops replication and spread of the infecting virus and the plant shows resistance reaction. The biggest success story of transgenic mediated virus resistance is the cultivation of transgenic papaya expressing capsid protein in Hawaii, which virtually saved the papaya industry from dreaded threat of ring spot disease (Yeh *et al.* 1998).

Disease resistance

Plant produces assorted set of PR proteins as defense response to fungal pathogens. Several reports of bioassay at the laboratory level indicate that over-expression of PR

proteins leads to enhanced resistance with reduced severity of symptoms in transgenic plants. However, any of these researches did not culminate into any commercial release. Besides, there are resistance (R) genes in plants which interact with avirulence gene of pathogen in a 'gene to gene' fashion. Researchers at National Research Centre on Plant Biotechnology, New Delhi have identified a blast resistance allele (Pikh from rice variety 'Tetep' and its corresponding avirulence counterpart from the fungal pathogen (Sharma *et al.* 2005) and validated its efficacy in transgenic rice.

Tolerance to abiotic stresses

Crop productivity suffers due to a variety of environmental stresses, viz drought, salinity, high/low temperature etc. Tolerance to these environmental stresses is known to be under polygenic control. However, a variety of genes isolated from bacteria, animal and plant source have been demonstrated for their ability to confer tolerance against these stresses. For a variety of stresses like water deficit, salinity and cold, the mechanism of damage to the physiology of the crop is overlapping, which ultimately causes water deficit at the cellular level. Therefore, genes leading to the biosysnthesis of osmolytes, which can retain intracellular water, viz glycinebetaine, trehalose, proline etc have been extensively used in developing transgenic lines tolerant to salinity, drought and cold. Cellular water deficit leads to generation of reactive oxygen species (ROS) which are harmful to the cell. Hence, another group of genes whose products can scavenge and thus protect the cell from ROS have been explored for developing stress-tolerant transgenic crops. Recently, a variety of transcription factors have been identified which are involved in stress acclimation process. The genes encoding these transcription factors are potential targets for genetic engineering in crop cultivars towards improved stress tolerance (Mahajan and Tuteja 2005).

Nutrition rich crops

Along with food security it is also important to achieve nutritional security, which is particularly relevant to developing nations. There are several examples of biotechnological attempts to develop nutritionally rich crop varieties. The most celebrated example is the development of 'Golden rice'. Three different genes phytoene synthase (*psy*), lycopene cyclase (*lyc*) and phytoene desaturase (*crtI*) have been introduced into japonica rice through *Agrobacterium* mediated transformation that has resulted in synthesis of beta carotene which is the precursor of vitamin A in human body (Ye *et al.* 2000). Directorate of Rice Research, Hyderabad, Tamil Nadu Agricultural University, Cimbatore and Indian Agricultural Research Institute, New Delhi are involved in transferring the genes to elite indica varieties. To increase nutrititive value of potato in terms of its protein content a single albumin protein *AmA1*, rich in essential amino acids has been isolated from *Amaranthus* (Chakraborty *et al.* 2000) and introduced in potato. National Institute of Plant Genome Research, New Delhi in collaboration with Central Potato Research Institute, Shimla has introduced the *AmA1* gene into seven genotypic backgrounds suitable for cultivation in different agro-climatic regions (Chakraborty *et al.* 2010).

Molecular breeding: An array of tools and techniques in the field of molecular biology has become available for supplementing the conventional genetic approaches for improving crop plants. Consequently, new integrated approaches are being designed. One of the approaches, popularly known as 'Marker Assisted Breeding' or 'Molecular Breeding' employs molecular markers for genome mapping, gene tagging and marker-assisted selection (MAS). With the complete sequencing of whole genome and a large number of random cDNAs in many different crop species, newer opportunities are emerging.

Molecular markers

Molecular markers are heritable differences in nucleotide sequences of DNA, which are used to mark genomic regions, differentiate and identify individuals, and tag genes of interest for marker-assisted breeding. These differences are detected by employing several molecular tools and techniques. The commonly used molecular markers are: (*i*) restriction fragment length polymorphism (RFLP), (*ii*) random amplified polymorphic DNA (RAPD), (*iii*) sequence tagged sites (STS), (*iv*) sequence tagged microsatellite sites (STMS), (*v*) amplified fragment length polymorphism (AFLP), and (*vi*) single nucleotide polymorphism (SNP).

Advantages with molecular markers

The molecular markers offer several advantages over the other genetic markers. These include: (*i*) abundance, (*ii*) co-dominance and, and (*iii*) developmental stage, tissue and environment independent expression. Large number of molecular markers can be obtained by using DNA probes from a variety of sources in combination with any of the commercially available restriction enzymes, by designing and using a variety of DNA primers in PCR and also by direct sequencing. This enables study of several markers in a single population for construction of high-density linkage maps of plant genomes. The above attributes make molecular markers ideal for characterization of quantitative traits.

Marker-assisted selection

Crop improvement involves creation of genetic variation and selection of desirable variants/recombinants, which are subsequently evaluated and released for commercial cultivation. Selection of desirable variants/recombinants has been based on the phenotype. Success of such selection largely depends upon the skill of the concerned plant breeder. Selection of a genotype carrying desirable gene or gene

combination via linked marker(s) is called marker-assisted selection. Breeders sometimes practice marker-assisted selection when an important trait, that is difficult to assess phenotypically, is tightly linked to another Mendelian trait, which can be easily scored. The molecular markers can increase the screening efficiency in breeding programmes in a number of ways: (*i*) the segregants can be scored at the seedling stage, (*ii*) it is possible to screen for traits that are extremely difficult, expensive or time consuming to score and measure such as tolerance to drought, salt, mineral deficiencies and toxicity, root morphology, resistance to nematodes or to specific races or biotypes of diseases or insects; (*iii*) selection can be practised for several traits simultaneously, which is difficult or even impossible by conventional means, (*iv*) heterozygotes are easily identified and distinguished from either homozygotes without resorting to progeny testing, and (*v*) creation of selection environment is not essential for identifying the desirable segregants. This saves time and effort. Because of these advantages offered by MAS over the conventional phenotype-based selection, markers are becoming increasingly important for the plant breeders. Some of the areas, where the markers are currently being used are outlined here.

Introgression of useful genes from the unadapted germplasm

Wild relatives of crops serve as rich reservoir of agronomically important genes. In wheat for instance, most of the genes for rust resistance have been derived from wild relatives. It has been reported that introgression of a specific locus from a donor genome using a backcross breeding programme relying on traditional method of selection requires a minimum of six generations of backcrosses. At this junction, the resultant progeny constitutes nearly 99% of recipient genome and the rest from the donor genome. In comparison, marker-assisted selection (MAS) in a backcross-breeding programme would aid in attaining the same desired objective in only three backcrosses. The use of MAS thereby significantly reduces the time needed to develop crop varieties with desired traits. Besides, it is possible to eliminate linked undesirable regions introgressed from the wild along with the gene of interest by exercising allele-specific selection against the wild species. Simultaneous mapping and introgression of QTLs present in the wild relatives for yield and its components in rice and tomato exemplifies the application of molecular markers in breaking yield barriers in crop plants. At the Directorate of Rice Research, Hyderabad, many loci for grain yield have been identified, and transferred to cultivated rice from its wild relative, *Oryza rufipogon* using molecular markers (Marri *et al.* 2005). The lines carrying the useful loci would now serve as donors of yield genes in rice breeding programmes.

Pyramiding of multiple genes for stress tolerance

Gene pyramiding refers to combining two or more major genes for a trait of interest in a single plant genotype. For instance, genes conferring resistance against different races of a particular pathogen are combined together in the genetic background of an agronomically superior genotype or popular variety. A combined effect of the resistance genes is thought to provide a broad spectrum of resistance by both individual gene action and additive gene action. However, it may be very difficult and a lengthy procedure to ascertain the number of genes that have been pyramided using a traditional breeding and selection programme. Therefore, molecular markers that are linked or co-inherited with the individual resistance genes could be used for marker-assisted selection. A typical example of gene pyramiding is in rice in India. Two genes for bacterial blight (BB) resistance namely, *Xa*13 and *Xa*21 have been combined in the background of popular Basmati rice variety Pusa Basmati 1 through marker-assisted selection programme jointly carried out by the National Research Centre on Plant Biotechnology and Indian Agricultural Research Institute, New Delhi. This has resulted in the development of a new variety called Improved Pusa Basmati1, which has been released for commercial cultivation in the country (Joseph *et al.* 2004, Gopalakrishnan *et al.* 2008). Similar efforts at the Directorate of Rice Research, Hyderabad has led to combining three genes for BB resistance, *xa5*, *xa13* and *Xa21* in the background of a popular rice variety BPT5204 (Sundaram *et al.* 2008).

Development of improved varieties/hybrids

The MAS is very effective in transferring genes for simply inherited traits from one variety to another. It increases the efficiency of selection and enables selection at early stages of plant growth. In maize, a microsatellite marker present in the gene *opaque* 2, which a regulator of zein protein synthesis in maize kernels, has been used at the Vivekanand Parvatiya Krishi Anushandhan Sansthan, Almora to breed high lysine maize lines, popularly known as quality protein maize (QPM). These lines, being the parental lines of hybrids, have been used to develop single cross QPM hybrids (Gupta *et al.* 2009). Besides, markers linked to loci affecting complex quantitative traits such as yield have been used in MAS to develop improved hybrids. Through marker-aided transfer of yield QTLs from inbred line T303 to B73 and from inbred line Oh43 to Mo17, improved inbreds have been generated in maize at the North Carolina State University, USA. Fifteen enhanced B73 lines were crossed with 18 enhanced Mo17 lines to produce 93 hybrids that were evaluated in replicated field trials at North Carolina, USA. Six of the hybrids were found to exceed the national check hybrids by two standard deviations or more and two highest yielding enhanced hybrids gave 15% more yield than the checks. These efforts need to be replicated in other crops for complex traits such as drought tolerance.

Genomics

Transgenic approaches along with advanced breeding

Table 2 Different types of sequences of important crops available in public domain

Type of database in public domain	Plant species
Whole genome	*Oryza sativa*, *Arabidopsis thaliana, Vitis vinifera, Populus trichocarpa, Carica papaya*
Partial genome	*T. aestivum, Z. mays, S. bicolor, B. oleracea, B. rapa, G. max, S. tuberosum, L. esculentum, V. vinifera, Poncirus trifoliate, Medicago truncatula, Lotus corniculatus etc*
EST	*Aegilops tauschii, Allium cepa, Arabidopsis thaliana, Avena sativa, Beta vulgaris subsp. vulgaris, Brassica napus, Brassica oleracea, Brassica rapa, Capsicum annuum, Coffea arabica, Glycine max, Gossypium arboreum, Gossypium hirsutum, Helianthus annuus, Hordeum vulgare, Lactuca sativa, Lolium perenne, Lotus corniculatus, Lycopersicon esculentum, Malus domestica, Medicago sativa, Medicago truncatula, Nicotiana benthamiana, Nicotiana tabacum, Oryza sativa, Phaseolus coccineus, Phaseolus vulgaris, Saccharum officinarum, Secale cereale, Solanum melongena, Solanum tuberosum, Sorghum bicolor, Triticum monococcum, Vitis vinifera, Zea mays etc.*
mRNA	*T. aestivum, Z. mays, S. bicolor, B. oleracea, B. rapa, G. max, S. tuberosum, L. esculentum, V. vinifera, Medicgo truncatula, L. corniculatus, O. sativa, A. thaliana etc.*
Protein	*Z. mays, S. bicolor, B. oleracea, B. rapa, G. max, S. tuberosum, V. vinifera, C. sinensis, M. truncatula, E. globulus, O. sativa, A. thaliana*
BAC end	*Oryza australiensis, O. brachyantha, O. glaberrima, O. granulata, O. latifolia, O. minuta, O. officinalis, O. punctata, O. ridleyi, O. rufipogon, O. schlechteri, G. hirsutum etc.*

Source: National Centre for Biotechnology Information

techniques and tissue culture methods are integral parts of today's crop improvement programmes. High throughput genomic study of crop plants is an emerging and promising area in current development of crop improvement. This is because most of the commercially relevant plant traits are interaction of large number of genes, their positions on chromosomes and promoters controlling them. As per Flavell (2010), any plant breeder would like to know about chromosome segment, gene and their alleles, alone or in combination with others contributing for traits and further want to alter the genome to manipulate traits. Here plant genomics plays an important role. It has a great potential to revolutionize the crop improvement programme. With advances in new sequencing technologies (Mardis 2008, Morozova and Marra 2008, Morozova *et al.* 2009, Edwards and Batley 2010), bioinformatics and genomic resources (Mochida and Shinozaki 2010) available for crop plants, breeders in collaboration with plant molecular biologist are able to manipulate crop plants to increase yield with better nutritional qualities even in adverse environmental conditions.

Genomics is the research strategy that uses molecular characterization and cloning of whole genomes to understand the structure, function and evolution of genes and to answer fundamental biological questions (http://www.fao.org/docrep/003/x3910e/X3910E13.htm). Thus, genomics not only deals with the unraveling of the genetic information present in an organism, but also with the understanding the mechanism by which this information is used by an organism. Genomics is often divided into structural and functional genomics. Structural genomics deals with the determination of the complete genome sequence or the complete set of proteins produced by an organism. This also includes construction of high-resolution genetic and physical maps. The functional genomics, on the other hand, studies the functioning of genes and pathways, i e the gene expression patterns in organisms. This also includes gene expression profiling, interactions among proteins, and between proteins and other molecules. Functional genomics characterizes all the genes present in the genome in one go. Therefore, the techniques used in functional genomics enables high throughput analyses that leads to very rapid data accumulation. Thus, the information generated in genomics is enormous. Interpretation and management of this information requires the use of powerful computers and specific bioinformatics tools.

Arabidopsis is the first sequenced plant genome which is playing the role of model plant for characterization of many genes for crop traits. With the next generation sequencing technologies now more than 30 crop plants are in genome sequencing pipeline. Sequencing of many of them has been already finished. The list of these crop plants is given in Table 2, which includes almost all type of crops like cereals, pulses, oilseeds, fibre, fruit and vegetables. A large amount of ESTs are also available from many crop plants on NCBI (http://www.ncbi.nlm.nih.gov/genomes/PLANTS/PlantList.html#EST) and DFCI-Plant gene indices (http://compbio.dfci.Harvard.edu/tgi/plant.html) databases. Thus increasing DNA sequence information is helping discovery of genes and identification of molecular markers associated with agronomically important traits.

Now-a-days many techniques are available which can integrate knowledge of genome sequence with cytogenetics that in turn would have implications to marker-assisted selection. Szinay *et al.* (2010) utilized fluorescent *in situ* hybridization (FISH) to improve breeding strategies in tomato

and other Solanaceaous crops. Microarray is a functional genomics tool, which has been successfully used to study expression profiles in many crop plants (Rensink and Buell 2005). Along with over-expression and RNAi studies, loss-of-function and gain-of-function mutants are also playing important role in functional genomics (Jiang and Ramachandran 2010, Kondou *et al.* 2010). Allele mining is a promising approach to dissect naturally occurring allelic variation at candidate genes controlling key agronomic traits, which has potential applications in crop improvement programmes (Kumar *et al.* 2010). The data generated by these high-throughput genomic techniques enhances breeding efficiency in crops like rice (Tyagi *et al*. 2004), wheat (Varshney *et al*. 2007, Gupta *et al*. 2008), tomato (Barone *et al.* 2009), soybean (Tran and Mochida 2010) etc.

Generation of genomic resources for complex and large crop genomes requires huge amount of financial and technical inputs. Comparative genomics plays an important role in these situations. In this branch of genomics, genes related to agronomically important traits are first characterized in model or related plants, genomes of which have been already sequenced. Using this information, similar genes are identified using bioinformatics tools from other crop plants having more complex genomes which are yet to be sequenced completely. Since cereals are the unavoidable part of our food, large amount of genomic data is generated in these plant species, which are being utilized further for comparative genomics. Similarly, the genome sequences of soybean (*Glycine max*), barrel medic (*Medicago truncatula*), and birdsfoot trefoil (*Lotus japonicus*) are being used as the platforms for comparative genomics of other orphan legumes (Cannon *et al.* 2009, Varshney *et al.* 2009, Young and Udvardi 2009). Integrated resources for comparative genomics are available in some databases like PlantGDB (Dong *et al.* 2005) and GreenPhylDB (Conte *et al.* 2008) for all plants, Gramene for cereals (http://www.gramene.org/), RoBuST for root and bulb crop families Apiaceae and Alliaceae (Bhasi *et al.* 2010) and GRASSIUS for grasses (Yilmaz *et al.* 2009).

Genomics has several practical applications in crop improvement. Genomics is useful in mining of genes of agronomic importance, gene identification in orphan or large complex crop species, development of DNA markers, tracing evolution of crop plants, marker-assisted selection, transgenic breeding and QTL mapping. Advances in our understanding of gene function and the availability of genomic resources along with a better understanding of genetic variation will alter the way that plant breeders identify genes underlying traits and then manipulate those traits (Flavell 2010). The combination of genome sequences and the other methods for defining functions of genes are providing a wealth of information about the molecular basis of plant phenotypes. The ultimate goal of genomics is to improve our ability to identify the genotypes with optimal agronomic traits in order to improve crop yield, its stability under varying climate and enhance nutritional quality, which are necessary for meeting the demands of the world population.

Metabolomics

Metabolomics, a new branch of science, is nascent by definition but quite old in use. With the advent of new biological era, a revolution in research witnessed a move from study of chromosomes to studies that encompass entire genomes, transcriptomes, proteomes, and metabolomes. 'Metabolomics' defined as systematic study of complete repertoire of metabolome/ chemicals of any organism refers to study of all small-molecular metabolites (metabolic intermediates, hormones and other signaling molecules, and secondary metabolites) found within a biological sample such as a single cell, tissue, organ or organism, which are the end products of cellular processes (Adams 2003). Thus, while gene expression data and other 'Omics' analyses do not tell the whole story of what might be happening in a cell, metabolic profiling gives an instantaneous snapshot of the physiology of that cell. Study of metabolites is critical in biology and plants are especially rich in diverse metabolic compounds. Moreover, within a plant species substantial quantitative and qualitative variation exists in metabolite composition. It has been estimated that plants contain over 100,000 metabolites with each species having its own chemotypic expression pattern. Within the context of metabolomics, with few exceptions, a metabolite is usually defined as any molecule less than 1 kDa in size. In plant-based metabolomics, there are 'primary' and 'secondary' metabolites. Primary metabolite is directly involved in the normal growth and development while secondary metabolite is not directly involved in those processes, but usually has important ecological function such as antibiotics and pigments.

The concept of 'metabolomics' was introduced by Roger Williams in the late 1940s. With technological advancements in the 1960s and 1970s, the term 'metabolic profile' was introduced by Horning in 1971. Recently, various metabolomic research areas have used differing terminology for the definition of metabolic approaches. However, the core of metabolomic research is 'metabolite fingerprinting' and 'metabolite profiling' approach. Fingerprinting involves the detection of all metabolites within a sample without regard to their identification while metabolic profiling involves the detection, quantification and identification of metabolites within an extract, commonly by employing chromatographic separation.

Instrumentation for metabolomic analysis

Central to metabolomics is precise detection and measurement of metabolite. As compared with genomics (nucleic acids) and proteomics (peptides and proteins), metabolomics deals with a range of chemically more diverse

compounds. Large variations in the relative concentrations of metabolites make their analysis more complicated. Therefore, comprehensive coverage can only be achieved by using multiparallel complementary extraction and detection technologies with careful experimental design (Saito and Matsuda 2010). Since there is no single technology currently available to detect all compounds found in plants or any other organism, a combination of multiple analytical techniques, such as gas chromatography (GC), liquid chromatography (LC), capillary electrophoresis (CE)-MS and nuclear magnetic resonance (NMR) are generally performed. A plant sample is subjected to chemical analysis using chromatography-mass spectrometry (MS), NMR, or Fourier transform-infrared spectrometer (FT-IR) to produce metabolite profiles covering a range of metabolome.

It is noteworthy that metabolome analysis can be performed by all chromatographic and spectroscopic techniques. Hyphenated mass spectrometries such as GC- and LC-MS are effective for detailed profiling of plant metabolites. Additionally suitable analytical method such as LC equipped with a photodiode array detector (PDA) for UV active metabolites and MALDI-TOF/MS for carotenoids is useful for focused metabolomics. Metabolic fingerprinting using a single spectrometer such as FT-IR enables fast and high-throughput analyses, suitable for screening metabolic mutants from large bio-resources.

Application to plant biology

Nutritional quality of crop plants is a direct function of metabolite content. It is therefore important to use metabolomic approach to understand better what especially has occurred during crop domestication in order to design new concepts for more targeted crop improvement. The potential for crop improvement is enormous as has recently been demonstrated in studies on genetical metabolomics (linking metabolic profiles with genomic information). These extensively new chemical signatures, resulting from different intra-specific allelic combinations (transgressive segregation), reveal the potential of metabolomics to detect transgressive segregants before a trait is expressed. This hidden biochemical variation, now revealed to plant breeders through metabolomics will open the door to general (biochemical) crop improvement and fine tuning of breeding specifically for nutritional traits in crop species' related to macro- and micronutrition and other aspects of food quality. Specific case of application of metabolomics in plant species is discussed below.

Biofortification and genetic modification: There is considerable interest in biofortification of foodstuffs where one or a small number of nutritional components have been particularly enhanced to supplement/ compensate for shortages elsewhere in the diet. Golden Rice with enhanced Provitamin A levels and lycopene-rich tomatoes are well-known examples. Other new examples are lysine-rich corn, folate-enriched tomatoes and ferritin-rich lettuce. Many consumers remain skeptical about modified products because of socio-psychological factors associated with potential food risks. Currently, metabolomic approaches are now being used to determine the extent and effectiveness of the modification introduced by GM products (Hall 2008). It has been a useful tool to identify biochemical differences between GM and non-GM crops using metabolomics.

Similarly, targeted profiling of carbohydrate metabolism of rice under chilling, salt and osmotic stress in different genotypes differing in chilling tolerance was carried out using a quantitative HPLC assay. It was found the chilling-tolerant genotype accumulated galactose and raffinose under stress, while these sugars declined in chilling-sensitive genotype. Based on the carbohydrate profiling results combined with the measurements of oxidative products and antioxidative enzymes, it was concluded that the chilling-tolerant genotype possess a more effective ROS-scavenging system. Metabolic fingerprinting of salt stress using FT-IR spectroscopy revealed presence of saturated and unsaturated nitriles, cyanide-containing compounds, a broad peak of NH2 and other nitrogen-containing compounds in resistant populations. More detailed metabolic analysis of salt stress response on salt-stressed *Arabidopsis* cell cultures using GC-MS and LC-MS profiling revealed that short-term responses to salt stress included the induction of the methylation cycle for the supply of methyl groups, the phenylpropanoid pathway for lignin production and glycine betaine production (Hall *et al.* 2008). The long-term effects were the coinduction of glycolysis and sucrose metabolism and coreduction of the methylation cycle.

Nutrient stress: Metabolomics of sulfur deficiency in *Arabidopsis* using GC-MS and LC-MS profiling studied response of 134 known metabolites and 6023 unknown non-redundant ions. Based on profiling data, coupled with transcriptomics data, a gene-metabolite correlation networks was reconstructed in *Arabidopsis* in response to sulfur deprivation. Similarly, combination of transcriptomics and metabolomics approaches was used to investigate transcriptional and metabolic responses of bean plants growing under P-deficient and P-sufficient conditions. GC-TOF-MS profiling identified a set of metabolites significantly changed in P-deficient roots. Most metabolites, including amino acids, polyols and sugars, were increased in P-stressed plants (Hernandez *et al.* 2007).

Oxidative stress: Dynamics of metabolites rapidly change in response to oxidative stress in plants. GC-MS profiling measured the levels of 50 polar metabolites following stress treatment and correlated metabolic changes to changes in mRNA levels in pants. It was observed that oxidative stress initially had dramatic inhibition of the TCA cycle and large sectors of amino acid metabolism followed by backing up of glycolysis and diversion of carbon into the oxidative pentose phosphate pathway (Baxter *et al.* 2007). Transcriptomics

analysis of the same samples also revealed a coordinated transcriptional response to cope with the stress with a major switch from anabolic to catabolic metabolism.

Heavy metal stress: Metabolic consequences of stress induced by heavy metals in plants have been studied using NMR-based metabolic fingerprinting and metabolite profiling. Compounds that showed an increase to cadmium toxicity were identified as malic acid and acetate, while concentration of glutamate and branched chain amino acids decreased (Le Lay *et al.* 2006).

Stress combination: Traditionally, abiotic stress conditions are studied in plants by applying a single stress condition such as drought, salinity or heat. This type of analysis is in sharp contrast to the conditions that occur in nature, in which plants are routinely subjected to a combination of different abiotic stresses. Drought and heat stress represent an excellent example of two different abiotic stress conditions that occur in the field simultaneously. Metabolite profiling of plants subjected to a combination of drought and heat stress revealed that sucrose and other sugars such as maltose and glucose accumulated in cells. In contrast, proline that accumulated in plants subjected to drought did not accumulate in plants during a combination of drought and heat stress. Heat stress was found to exaggerate the toxicity of proline to cells, suggesting that during a combination of stress, sucrose replaces proline in plants as the major osmoprotectant (Mittler 2006). These findings of different metabolic responses to stress combination in comparison with each individual stress highlight the need for further studies of different stress combinations at the metabolic level.

Bioinformatics

Bioinformatics is the development and application of computational tools to acquire, store, organize, archive, analyze and visualize biological data. Rapid advances and development in genome research in the recent past have resulted in the generation of large data set of DNA and protein sequences from different prokaryotes and eukaryotes. Storage, management and retrieval of these sequences at one place and their accessibility to the biologists working world over have essentially required automatic information storage and retrieval systems. Automation of Sanger`s dideoxy method of DNA sequencing in the last decade laid the foundation of genome research in different organisms including plants. The genomes of large number of microorganisms and other species including human, have already been sequenced, though the progress is rather slow in case of sequencing complete genomes of plants. The Genomes of *Arabidopsis thaliana* (The *Arabidopsis thaliana* Genome Sequencing Project 2000), *Oryza sativa* (International Rice Genome Sequencing Project 2005), grape (Jaillon *et al.* 2007), poplar (Tuskan *et al.* 2006), papaya (Ming *et al.* 2008), sorghum (Paterson *et al.* 2009) and Soybean (Schmutz *et al.* 2010) have been completely sequenced. The genome or EST sequences of the important crops like wheat, maize, sorghum, brassica and sugarcane are available in the public databases (Table 2). In vegetables, tomato and capsicum are being sequenced. Among the pulses, *Phaseolus* is being sequenced at international level. The sequence of *Medicago sativa* is also available in the public domain, which can be used in various comparative genome analysis in related pulses.

DNA sequence information is becoming an indispensable tool in modern biology. Therefore, the huge wealth of DNA and protein sequence resources are being stored in the data banks. Major public data banks which take care of the DNA and protein sequences are GenBank in USA (http://www.ncbi.nlm.nih.gov), EMBL (European Molecular Biology Laboratory) in Europe (http://www.ebi.ac.uk/embl) and DDBJ (DNA Data Bank) in Japan (http://www.ddbj.nig.ac.jp). This rapid growth in DNA sequence data is because of various collaborative international programmes that were implemented during the past few years to sequence complete genomes of various organisms. The whole genomes of various microorganisms have also been sequenced by The Institute of Genome Research (TIGR) and are available from their website www.tigr.org. Now, these DNA sequences have to be used in a meaningful way for the welfare of mankind.

Uses of bioinformatics tools in agriculture

DNA sequence analysis: Bioinformatics tools are now easily available to the biologists with the advent of internet and various Web Browsers on World Wide Web. These tools are indispensable for any Genome Sequencing Centre. Most of these softwares are available free of charge to the public institutions. Beside these softwares, small scripts written in Perl and Java languages are also used to help biologists in handling large genome sequences at various stages of data generation, assembly and annotation. Similarly, bioinformatics tools can be used for protein function analysis. Finding SSR and SNP markers from the EST or genome sequences can be performed in silico by using different algorithms (Baxevanis 2001).

Searching for sequence similarity: Once high quality sequence is obtained it is required to determine if the sequence is similar to other DNA sequences available in the databases. All sequence searching methods rely on the basic concepts of alignment and distance between the sequences. There are different algorithms to perform global and local alignments. This type of sequence comparison is generally performed with BLAST (Basic Local Alignment Search Tools), which compares unknown sequence against all the sequences available in the database (http://www.ncbi.nlm.nih.gov/). In BLAST, the best local alignment between the unknown sequences and the database is found by using an approach based on matching short sequence fragments and a powerful statistical model.

Gene prediction and annotation: Simply determining the order of four alphabets (ATGC) in a DNA sequence of any organism has little value unless some meaning is derived from this by gene prediction. Gene prediction is complex and there is no algorithm, which can exactly predict the true exons in a DNA sequence (Burge and Karlin 1998). Basically two major considerations are taken into account while predicting a gene: (*i*) Identification of structural elements such a start/ stop codon and splice sites of the unknown sequence, and (*ii*) performing homology search against protein, EST and cDNA database to identify potential coding regions. For gene prediction, the software GENSCAN developed by MIT, USA (http://www.genes.mit.edu/ GENSCAN.html) is commonly used. This is freely available on Web for online analysis of DNA sequences. The output obtained from the GENSCAN is then used for gene annotation by using BLAST to search the public or private DNA sequence databases to find out the matches to the unknown query sequence with millions of sequences available in the GenBank. A very popular website, http:// www.ncbi.nlm.nih.gov is available for BLAST at NCBI`s Home page, which performs searches by using various criteria and options.

Primer design for allele mining and development of DNA markers: Another important aspect in the use of genome sequence information after gene prediction is to design primers either for PCR or for sequencing specific regions of different genomes. Such primers are used for amplification of genes or its alleles from the known sources and making best use of it in plant breeding programmes. Though PRIME software within GCG package is mainly used for this purpose, PRIMER 3 web-based software (www-genoem.wi.mit.edu/ genome software/ other/ primer3.html) is being commonly used for designing primers.

Phylogenetic analysis: Once similarity search is performed between an unknown sequence and the database sequence to find per cent homology between them, it is obvious to know how these sequences are related to each other. The sequences derived from two closely related organisms show more similarity at DNA level than the sequences from distantly related organisms. To find an evolutionary relationship among sequences derived from different organisms, a phylogenetic tree is constructed. Such evolutionary trees can also be constructed using phenotypic markers, molecular markers and protein sequence information. A typical phylogentic tree comprises nodes, branches and termini of the branches. The DNA STAR software (www.dnastar.com) has options to construct tree from different DNA or protein sequences. However, web-based tools like MacClade (//www. phylogeny. arizona.edu/ macclade/) can also be used for evolutionary studies of different organisms based on their DNA sequences.

Similarly, bioinformatic tools can be used for protein function analysis by database search. Finding SSR markers and SNP markers from the EST or genome sequences can be performed *in silico* by using different algorithms. The SNP resource of rice is now available in the public domain (Feltus *et al.* 2004). Many of the steps, which are required in the wet lab experiments, can be hastened with the use of *in silico* methods. However, technical know how to use these genomic resources and biotools is very essential before embarking upon any programme in functional and comparative genomics.

Perspectives

Plant biotechnology has made significant strides in the past 20 years encompassing within its fold the spectacular developments in plant molecular biology, genetic engineering, genomics, molecular crop breeding and bioinformatics. Development of efficient procedures of genetic transformation led to the introduction of several transgenes in all major crop species. A variety of traits have been introduced in many crops. More than a dozen GM crops are cultivated globally. Marker-assisted selection has given rise to commercial products such as blight-resistant rice, downy mildew-resistant *bajra*, submergence tolerant rice etc. Unraveling of genomes aided by innovations and discoveries in bioinformatics would hasten such efforts. Intensive and all round efforts are needed to meet the challenges of feeding the future populations. Integration of modern biology, plant breeding, computational boiology, phenomics and related tools will make the future crop improvement endeavours highly successful.

REFERENCES

Adams A. 2003. Metabolomics: Small-molecule 'omics'. *The Scientist* **17**: 38–40.

Agrawal G K, Yamazaki M, Kobayashi M, Hirochika R, Miyao A and Hirochika H. 2001. Screening of the rice viviparous mutants generated by endogenous retrotransposon Tos17 insertion: tagging of a zeaxanthin epoxidase gene and a novel OsTATC gene. *Plant Physiology* **125**: 1248–57.

Averboukh L, Douglas S A, Lowe K, Maher J and Pardee A B.1996. Better gel resolution and longer cDNAs increase the precision of differential display. *BioTechniques* **20**: 918–21.

Azpiroz-leehan R and Feldmann K A. 1997. T-DNA insertion mutagenesis in Arabidopsis: going back and forth. *Trends in Genetics* **13**: 152–6.

Barone A, Di Matteo A, Carputo D and Frusciante L. 2009. High-throughput genomics enhances tomato breeding efficiency. *Current Genomics* **10**: 1–9.

Baxevanis A D. 2001. Predictive methods using DNA sequences (*in*) *Bioinformatics: a Practical Guide to The Analysis of Genes and Proteins*, pp 233–52. Baxevanis A D and Ouellette B F F (Eds). John Wiley & sons, New York.

Baxter C J, Redestig H, Schauer N, Repsilber D, Patil K R, Nielsen J, Selbig J, Liu J, Fernie A R and Sweetlove L J. 2007. The metabolic response of heterotrophic *Arabidopsis* cells to oxidative stress. *Plant Physiology* **143**: 312–25.

Beachy RN, Loesch-Fries S and Tumer N E. 1990. Coat protein-

mediated resistance against virus infection. *Annual Review of Phytopathology* **28**: 451–74.

Bhasi A, Senalik D, Simon P W, Kumar B, Manikandan V, Philip P and Senapathy P. 2010. RoBuST: an integrated genomics resource for the root and bulb crop families Apiaceae and Alliaceae. *BioMed Central Plant Biology* **10**: 161.

Burge C B and Karlin S. 1998. Finding the genes in genomic DNA. *Current Opinion Structural Biology* **8**: 346–54.

Campisi L, Yi Y, Heilig E, Herman B, Cassista A J, Allen D W, Xiang H and Jack T. 1999. Generation of enhancer trap lines in Arabidopsis and characterization of expression patterns in the inflorescence. *The Plant Journal* **17**: 699–707.

Cannon S B, May G D and Jackson S A. 2009. Three sequenced legume genomes and many crop species: rich opportunities for translational genomics. *Plant Physiology* **151**: 970–7.

Chakraborty S, Chakraborty N, Agrawal L, Ghosh S, Narula K, Shekhar S, Naik P S, Pande P C, Chakrborti S K and Datta A. 2010. Next-generation protein-rich potato expressing the seed protein gene AmA1 is a result of proteome rebalancing in transgenic tuber. *Proceedings of the National Academy of Sciences*, USA **107**: 17533–8.

Chakraborty S, Chakraborty N and Datta A. 2000. Increased nutritive value of transgenic potato by expressing a nonallergenic seed albumin gene from *Amaranthus hypochondriacus*. *Proceedings of the National Academy of Sciences*, USA **97**: 3724–9.

Chu Z, Fu B, Yang H, Xu C, Li Z, Sanchez A, Park Y J, Bennetzen J L, Zhang Q and Wang S. 2006. Targeting xa13, a recessive gene for bacterial blight resistance in rice. *Theoretical Applied Genetics* **112**: 455–61.

Clough S J and Bent A F. 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal* **16**: 735–43.

Colau D, Negrutiu I, Montagu M and Hernalsteens J P. 1987. Complementation of a threonine dehydratase-deficient mutant of *Nicotiana plumbaginifoh* mutant after *Agrobacterium tumefaciens* mediated transfer of *Saccharomyces cerevisiae* ILVl gene. *Molecular and Cellular Biology* **l7**: 2552–7.

Conte M G, Gaillard S, Lanau N, Rouard M and Perin C. 2008. GreenPhylDB: a database for plant comparative genomics. *Nucleic Acid Research* **36**: 991–8.

Dong Q, Lawrence C J, Schlueter S D, Wilkerson M D, Kurtz S, Lushbough C and Brendel V. 2005. Comparative plant genomics resources at Plant GDB. *Plant Physiology* **139**: 610–18.

Doss R P. 1996. Differential display without radioactivity-a modified procedure. *BioTechniques* **21**: 408–412.

Edwards D and Batley J. 2010. Plant genome sequencing: applications for crop improvement. *Plant Biotechnology Journal* **8**: 2–9.

Feldmann K A and Marks M D. 1987. *Agrobacterium* mediated transformation of germinating seeds of *Arabidopsis thaliana*. *Molecular and General Genetics* **208**: 1–9.

Feltus A F, Wan J, Schulze S R, Estill J C, Jiang N and Paterson A H. 2004. An SNP resource for rice genetics and breeding based on sub-species *indica* and *japonica* genome alignments. *Genome Research* **14**: 1812–19.

Flavell R. 2010. From genomics to crop breeding. *Nature Biotechnology* **28**: 144–5.

Gierl A and Saedler H. 1992. Plant-transposable elements and gene tagging. *Plant Molecular Biology* **19**: 39–49.

Gopalakrishnan S, Sharma R K, Rajkumar K A, Joseph M, Singh V P, Singh A K, Bhat K V, Singh N K and Mohapatra T. 2008. Integrating marker-assisted background analysis with foreground selection for identification of superior bacterial blight resistant recombinants in Basmati rice. *Plant Breeding* **127**: 131–9.

Gu K, Yang B, Tian D, Wu L, Wang D, Sreekala C, Yang F, Chu Z, Wang G L, White F F and Yin Z. 2005. R gene expression induced by a type-III effector triggers disease resistance in rice. *Nature* **435**: 1122–5.

Guo L, Ma L, Jiang H, Zeng D, Hu J, Wu L, Gao Z, Zhang G and Qian Q. 2009. Genetic Analysis and Fine Mapping of Two Genes for Grain Shape and Weight in Rice. *Journal of Integrative Plant Biology* **51**: 45–51.

Gupta H S, Agrawal P K, Mahajan V, Bisht G S, Kumar A, Verma P, Srivastava A, Saha S, Babu R, Pant M C and Mani V P. 2009. Quality protein maize for nutritional security: rapid development of short duration hybrids through molecular marker assisted breeding. *Current Science* **96**: 230–7.

Gupta P K, Mir R R, Mohan A and Kumar J. 2008. Wheat genomics: present status and future prospects. *International Journal of Plant Genomics* **896451**: 1–36.

Hall R D, Brouwer I D and Fitzgerald M A. 2008. Plant metabolomics and its potential application for human nutrition. *Physiologia Plantarum* **132**(2): 162–75.

Hernandez G, Ramirez M, Valdes-Lopez O, Tesfaye M, Graham M A, Czechowski T, Schlereth A, Wandrey M, Erban A, Cheung F, Wu H C, Lara M, Town C D, Kopka J, Udvardi M K and Vance C P. 2007. Phosphorus stress in common bean: root transcript and metabolic responses. *Plant Physiology* **144**: 752–67.

Huang Z, Fasco M J and Kaminsky L S. 1996. Optimization of Dnase I removal of contaminating DNA from RNA for use in quantitative RNAPCR. *BioTechniques* **20**: 1012–20.

IRGSP. 2005. The map based sequence of rice genome. *Nature* **436**: 793–800.

Iyer A S and McCouch S R. 2004. The rice bacterial blight resistance gene *xa5* encodes a novel form of disease resistance. *Molecular Plant Microbe Interaction* **17**: 1348–54.

Jaillon O, Aury J M, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, Vezzi A, Legeai F, Hugueney P, Dasilva C, Horner D, Mica E, Jublot D, Poulain J, Bruyère C, Billault A, Segurens B, Gouyvenoux M, Ugarte E, Cattonaro F, Anthouard V, Vico V, Del Fabbro C, Alaux M, Di Gaspero G, Dumas V, Felice N, Paillard S, Juman I, Moroldo M, Scalabrin S, Canaguier A, Le Clainche I, Malacrida G, Durand E, Pesole G, Laucou V, Chatelet P, Merdinoglu D, Delledonne M, Pezzotti M, Lecharny A, Scarpelli C, Artiguenave F, Pè M E, Valle G, Morgante M, Caboche M, Adam-Blondon A F, Weissenbach J, Quétier F and Wincker P. 2007. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* **449**: 463–7.

James C. 2010. Global status of commercialized biotech/GM crops: 2010. International Service for the *Acquisition of Agri-biotech Applications (ISAAA)*, Ithaca, NY, USA: 42

Jiang S Y and Ramachandran S. 2010. Natural and artificial mutants as valuable resources for functional genomics and molecular breeding. *International Journal of Biological Science* **6**: 228–51.

Jones S W, Cai D, Weislow O S and Babak E A. 1997. Generation of Multiple mRNA fingerprints using fluorescence-based differential display and automated DNA sequencer. *BioTechniques* **22**: 536–43.

Joseph M, Gopalakrishnan S, Sharma R K, Singh V P, Singh A K, Singh N K and Mohapatra T. 2004. Combining bacterial blight resistance and Basmati quality characteristics by phenotypic and molecular marker-assisted selection in rice. *Molecular Breeding* **13**: 377–87.

Karihaloo J L and Kumar P A. 2009. *Bt* cotton in India: A status report. Asia-Pacific Consortium on Agricultural Biotechnology, New Delhi.

Komatsu M, Chujo A, Nagato Y, Shimamoto K and Kyozuka J. 2003. Frizzy panicle is required to prevent the formation of axillary meristems and to establish floral meristem identity in rice spikelets. *Development* **130**: 3841–50.

Kondou Y, Higuchi M and Matsui M. 2010. High-throughput characterization of plant gene functions by using gain-of-function technology. *Annual Review of Plant Biology* **61**: 373–93.

Kumar G R, Sakthivel K, Sundaram R M, Neeraja C N, Balachandran S M, Rani N S, Viraktamath B C and Madhav M S. 2010. Allele mining in crops: prospects and potentials. *Biotechnology Advances* **28**: 451–61.

Kunze R, Saedler H and Lonning W E.1997. Plant transposable elements. *Advances in Botanical Research* **27**: 321–470.

Le Lay P, Isaure M P, Sarry J E, Kuhn L, Fayard B, Le Bail J L, Bastien O, Garin J, Roby C and Bourguignon J. 2006. Metabolomic, proteomic and biophysical analyses of *Arabidopsis thaliana* cells exposed to a caesium stress. *Biochimie* **88**: 1533–47.

Liang P and Pardee A B. 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* **257**: 967–71.

Lindsey K, Wei W, Clarke M C, Mcardle H F, Rooke L M and Topping J F. 1993. Tagging genomic sequences that direct transfer expression by activation of a promoter trap in plants. *Transgenic Research* **2**: 33–47.

Liu Y, Mitsukawa N, Oosumi T and Whittier R. 1995. Efficient isolation and mapping of *Arabidopsisthaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *The Plant Journal* **8**: 457–63.

Mahajan S and Tuteja N. 2005. Cold, salinity and drought stresses: an overview. *Archives of Biochemistry and Biophysics* **444**: 139–58.

Mardis E R. 2008. Next-generation DNA sequencing methods. *Annual Review of Genomics Human Genetics* **9**: 387–402.

Marri P R, Sarla N, Reddy L V and Siddiq E A. 2005. Identification and mapping of yield and yield related QTLs from an Indian accession of *Oryza rufipogon*. *BioMedical Central Genetics* **6**: 33.

Meyer P, Heidmanii I, Forkmann G and Saedler H. 1987. A new petunia flower colour generated by transformation of a mutant with a maize gene. *Nature* **330**: 677–78.

Ming R, Hou S, Feng Y, Yu Q, Dionne-Laporte A, Saw J H, Senin P, Wang W, Ly B V, Lewis K L T, Salzberg S L, Feng L, Jones M R, Skelton R L, Murray J E, Chen C, Qian W, Shen J, Du P, Eustice M, Tong E, Tang H, Lyons E, Paull R E, Michael T P, Wall K, Rice D W, Albert H, Wang M L, Zhu Y J, Schatz M, Nagarajan N, Acob R A, Guan P, Blas A, Wai C M, Ackerman C M, Ren Y, Liu C, Wang J, Wang J, Na J K, Shakirov E V, Haas B, Thimmapuram J, Nelson D, Wang X, Bowers J E, Gschwend A R, Delcher A L, Singh R, Suzuki J Y, Tripathi S, Neupane K, Wei H, Irikura B, Paidi M, Jiang N, Zhang W, Presting G, Windsor A, Navajas P R, Torres M J, Feltus F A, Porter B, Li Y, Burroughs A M, Luo M C, Liu L, Christopher D A, Mount S M, Moore P H, Sugimura T, Jiang J, Schuler M A, Friedman V, Mitchell O T, Shippen D E, Depamphilis C W, Palmer J D, Freeling M, Paterson A H, Gonsalves D, Wang L and Alam M. 2008. The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nature* **452**: 991–96.

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

Mochida K and Shinozaki K. 2010. Genomics and bioinformatics resources for crop improvement. *Plant Cell Physiology* **51**: 497–523.

Morozova O and Marra M A. 2008. Applications of next-generation sequencing technologies in functional genomics. *Genomics* **92**: 255–64.

Morozova O, Hirst M and Marra M A.2009. Applications of new sequencing technologies for transcriptome analysis. *Annual Review of Genomics and Human Genetics* **10**: 135–51.

Nakazawa M, Yabe N, Ichikawa T, Yamamoto Y Y, Yoshizumi T, Hasunuma K and Matsui M. 2001. Dfl1, an auxin-responsive GH3 gene homologue, negatively regulates shoot cell elongation and lateral root formation, and positively regulates the light response of hypocotyl length. *Plant Journal* **25**: 213–21.

Ochman H, Gerber A S and Hartl D L. 1988. Genetic applications of an inverse polymerase chain reaction. *Genetics* **120**: 621–3.

Osborne B I, Wirtz U and Baker B. 1995. A system for insertional mutagenesis and chromosomal rearrangement using the Ds transposon and Cre-lox. *Plant Journal* **7**: 687–701.

Park D S, Song M Y, Park S K, Lee S K, Lee J H, Song S Y, Eun M Y, Hahn T R, Sohn J K, Yi G, Nam M H and Jeon J S. 2008. Molecular tagging of the Bph1 locus for resistance to brown planthopper (*Nilaparvata lugens* Stål) through representational difference analysis. *Molecular Genetics and Genomics* **280**: 163–72.

Paterson A H, Bowers J E, 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 A K, Chapman J, Feltus F A, Gowik U, Grigoriev I V, Lyons E, Maher C A, Martis M, Narechania A, Otillar R P, Penning B W, Salamov A A, Wang Y, Zhang L, Carpita N C, Freeling M, Gingle A R, Hash C T, Keller B, Klein P, Kresovich S, McCann M C, Ming R, Peterson D G, Rahman M, Ware D, Westhoff P, Mayer K F X, Messing J and Rokhsar D S.2009. The *Sorghum bicolor* genome and the diversification of grasses. *Nature* **457**: 551–6.

Rensink W A and Buell C R. 2005. Microarray expression profiling resources for plant genomics. *Trends in Plant Science* **10**: 603–9.

Saito K and Matsuda F. 2010. Metabolomics for functional genomics, systems biology and biotechnology. *Annual Review of Plant Biology* **61**: 463–89

Schmutz J, Cannon S B, Schlueter J, Ma J, Mitros T, Nelson W, Hyten D L, Song Q, Thelen J J, Cheng J, Xu D, Hellsten U, May G D, Yu Y, Sakurai T, Umezawa T, Bhattacharyya M K, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu S, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du J,

Tian Z, Zhu L, Gill N, Joshi T, Libault M, Sethuraman A, Zhang X C, Shinozaki K, Nguyen H T, Wing R A, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker R C and Jackson S A. 2010. Genome sequence of the palaeopolyploid soybean. *Nature* **463**: 178–83.

Sharma T R, Madhav M S, Singh B K, Shanker, P, Jana T K, Dalal V, Pandit A, Singh A, Gaikwad K, Upreti H C and Singh N K. 2005. High-resolution mapping, cloning and molecular characterization of the gene of rice, which confers resistance to rice blast. *Molecular Genetics Genomics* **274**: 569–78.

Sun X, Cao Y, Yang Z, Xu C, Li X, Wang S and Zhang Q. 2004. Xa26, a gene conferring resistance to *Xanthomonas oryzae* pv. *oryzae* in rice, encodes an LRR receptor kinase-like protein. *Plant Journal* **37**: 517–27.

Sundaram R M, Vishnupriya M R, Biradar S K, Laha G S, Reddy G A, Rani N S, Sarma N P and Sonti R V. 2008. Marker assisted introgression of bacterial blight resistance in Samba Mahsuri, an elite indica rice variety. *Euphytica* **160**: 411–22.

Szinay D, Bai Y, Visser R and de Jong H. 2010. FISH applications for genomics and plant breeding strategies in tomato and other solanaceous crops. *Cytogenetics and Genome Research* **129**: 199–210.

Tanaka K, Murata K, Yamazaki M, Onosato K, Miyao A and Hirochika H. 2003. Three distinct rice cellulose synthase catalytic subunit genes required for cellulose synthesis in the secondary wall. *Plant Physiology* **133** : 73–83.

Tran L S and Mochida K. 2010. Functional genomics of soybean for improvement of productivity in adverse conditions. *Functional and Integrative Genomics* **10**(4): 447–62.

Tuskan G A, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, Schein J, Sterck L, Aerts A, Bhalerao R R, Bhalerao R P, Blaudez D, Boerjan W, Brun A, Brunner A, Busov V, Campbell M, Carlson J, Chalot M, Chapman J, Chen G L, Cooper D, Coutinho P M, Couturier J, Covert S, Cronk Q, Cunningham R, Davis J, Degroeve S, Déjardin A, dePamphilis C, Detter J, Dirks B, Dubchak I, Duplessis S, Ehlting J, Ellis B,. Gendler K, Goodstein D, Gribskov M, Grimwood J, Groover A, Gunter L, Hamberger B, Heinze B, Helariutta Y, Henrissat B, Holligan D, Holt R, Huang W, Islam-Faridi N, Jones S, Jones-Rhoades M, Jorgensen R, Joshi C, Kangasjärvi J, Karlsson J, Kelleher C, Kirkpatrick R, Kirst M, Kohler A, Kalluri U, Larimer F, Leebens-Mack J, Leplé J C, Locascio P, Lou Y, Lucas S, Martin F, Montanini B, Napoli C, Nelson D R, Nelson C, Nieminen K, Nilsson O, Pereda V, Peter G, Philippe R, Pilate G, Poliakov A, Razumovskaya J, Richardson P, Rinaldi C, Ritland K, Rouzé P, Ryaboy D, Schmutz J, Schrader J, Segerman B, Shin H, Siddiqui A, Sterky F, Terry A, Tsai C J, Uberbacher E, Unneberg P, Vahala J, Wall K, Wessler S, Yang G, Yin T, Douglas C, Marra M, Sandberg G, de Peer Y V and Rokhsar D. 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**: 1596–1604.

Tyagi A K, Khurana J P, Khurana P, Raghuvanshi S, Gaur A, Kapur A, Gupta V, Kumar D, Ravi V, Vij S and Sharma S. 2004. Structural and functional analysis of rice genome. *Journal of Genetics and Genomics* **83**: 79–99.

Valvekens D, Van montagu M and Van Lijse Bettens M.1988. A. tumefaciencs mediated transformation of *Arabidopsisthaliana* root explants by using kanamycin selection. *Proceedings of the National Academy of Sciences USA* **85**: 5536–46.

Varshney R K, Close T J, Singh N K, Hoisington D A and Cook D R. 2009. Orphan legume crops enter the genomics era. *Current Opinion in Plant Biology* **12**: 202–10.

Varshney R K, Langridge P and Graner A. 2007. Application of genomics to molecular breeding of wheat and barley. *Advances in Genetics* **58**: 121–55.

Yamazaki M, Tsugawa H, Miyao A, Yano M, Wu J, Yamamoto Y, Matsumoto T, Sasaki T and Hirochika H. 2001. The rice retrotransposon Tos17 prefers low copy sequences as integration targets. *Molecular and General Genetics* **265**: 336–44.

Yanofsky M F, Ma H, Bowman J L, Drews G N, Feldmann K A and Meyerowitz E M. 1990. The protein encoded by the *Arabidopsis* homeotic gene agamous resembles transcription factors. *Nature* 5: 35–9.

Ye X, Al-Babili S, Klöti A, Zhang J, Lucca P, Beyer P and Potrykus I. 2000. Engineering the provitamin A (b-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science* **287**: 303–05.

Yeh S D, Bau H J, Cheng Y H, Yu T A, Yang J S and Drew R A.1998. Greenhouse and field evaluations of coat-protein transgenic papaya resistant to papaya ring spot virus. Acta *Horticulture* **461**: 321–8.

Yilmaz A, Nishiyama M Y Jr, Fuentes B G, Souza G M, Janies D, Gray J and Grotewold E. 2009. GRASSIUS: a platform for comparative regulatory genomics across the grasses. *Plant Physiology* **149**: 171–80.

Young N D and Udvardi M. 2009. Translating *Medicago truncatula* genomics to crop legumes. *Current Opinion in Plant Biology* **12**: 193–201.

Zhao S, Ooi S L and Pardee A B. 1995. New primer strategy improves precision of differential display. *BioTechniques* **18**: 842–50.