Biotechnol Lett (2014) 36:417–426
DOI 10.1007/s10529-013-1377-1

REVIEW

From genomics to functional markers in the era of next-generation sequencing

R. K. Salgotra · B. B. Gupta · C. N. Stewart Jr.

Revised: 14 July 2013 / Accepted: 25 September 2013 / Published online: 16 October 2013
© Springer Science+Business Media Dordrecht 2013

Abstract The availability of complete genome sequences, along with other genomic resources for *Arabidopsis*, rice, pigeon pea, soybean and other crops, has revolutionized our understanding of the genetic make-up of plants. Next-generation DNA sequencing (NGS) has facilitated single nucleotide polymorphism discovery in plants. Functionally-characterized sequences can be identified and functional markers (FMs) for important traits can be developed at an ever-increasing ease. FMs are derived from sequence polymorphisms found in allelic variants of a functional gene. Linkage disequilibrium-based association mapping and homologous recombinants have been developed for identification of "perfect" markers for their use in crop improvement practices. Compared with many other molecular markers, FMs derived from the functionally characterized sequence genes using NGS techniques and their use provide opportunities to develop high-yielding plant genotypes resistant to various stresses at a fast pace.

Keywords Crop plants · Functional markers · Genomic selection · Next generation DNA sequencing · Plant biotechnology, plant breeding, polymorphisms

R. K. Salgotra (✉) · B. B. Gupta
Sher-e-Kashmir University of Agricultural Sciences & Technology of Jammu, Chatha, Jammu 190008, J & K, India
e-mail: rks_2959@rediffmail.com

B. B. Gupta
e-mail: dr_bbgupta@ymail.com

C. N. Stewart Jr.
Plant Molecular Genetics, Department of Plant Sciences, University of Tennessee, Knoxville, TN, USA
e-mail: nealstewart@utk.edu

Introduction

The recent progress in the area of plant molecular biology and genomics has the potential to initiate a new 'Green Revolution,' which is of vital importance for the development of drastically-improved crop germplasm (Gupta et al. 2005). Increasingly exact linkage of markers and genes to traits will lead to more efficient plant breeding in the future (Kulwal et al. 2011). Genomics technologies are being applied to the improvement of crop plants with encouraging results (Xu et al. 2005). The genomics revolution, which started in the 1990s, has greatly improved our understanding of the genetic make-up of a wide array of living organisms, including several plant species. Complete genome sequences of *Arabidopsis* (The *Arabidopsis* Genome Initiative 2000), rice (The Rice Chromosome 10 Sequencing Consortium 2003), soybean (Schmutz et al. 2010) and other species, together with high-throughput technology for the analysis of

transcripts, proteins, and mutants, provide the basis for understanding the relationships among genes, proteins and phenotypes. Characterization of plant genomes and the genes should translate to unprecedented crop improvement (Xu et al. 2005).

For over 20 years, DNA markers have been the most widely-used molecular markers in crop improvement, owing to their abundance and polymorphisms. Most of these markers are selectively neutral because they are usually located in non-coding and non-regulatory regions of DNA (McKay and Latta 2002). The first plant DNA markers were based on restriction fragment length polymorphisms (RFLPs) (Bernatsky and Tanksley 1986). Early hybridization-based, isotopically-labeled RFLP techniques were inherently challenging and time consuming, and were eventually replaced by less complex, more cost-effective PCR-based markers. Molecular markers include (i) RFLPs and other Southern blot-based markers (Botstein et al. 1980), (ii) PCR-based markers such as random amplification of polymorphic DNA (RAPD) (Williams et al. 1990), amplified fragment length polymorphism (AFLP) (Vos et al. 1995), microsatellite or simple sequence repeat (SSR) (Powell et al. 1996), sequence characterized amplified region (Paran and Michelmore 1993), cleaved amplified polymorphic sequences (Konieczny and Ausubel 1993), direct amplification of length polymorphism (Desmarais et al. 1998), and (iii) sequence-based markers, such as single nucleotide polymorphism (SNP) markers (Gupta et al. 2005) have been developed and applied to a large number of plant species. The majority of these molecular markers have been developed either from genomic DNA libraries (RFLPs and SSRs) or from random PCR amplification of genomic DNA (RAPDs) or both (AFLPs). However, when such markers are used for marker-assisted selection in plant breeding they may have some limitations owing to genetic recombination giving rise to false positives (Frisch et al. 1999). High throughput sequencing techniques in combination with the start of genome and expressed sequence tag (EST) sequencing programs in model plant species, led to the acceleration in the identification of variation at the single base pair resolution (Wang et al. 1998). Functional markers (FMs) are developed from polymorphic sites within genes that causally affect target trait variation i.e. based on functional characterization of the polymorphisms (Andersen and Lubberstedt 2003). Hence they are more meaningful in crop improvement. It is comparatively easier to develop functional markers in plants such as rice, tomato and soybean and medicago where either compete or nearly complete genome sequence information is available than in others in which little or no genomic information is available.

In recent years, the field of agricultural genomics is in the midst of a technological revolution spurred by next-generation DNA sequencing (NGS) technologies used to complete the sequencing of many genomes (Young et al. 2011). NGS platforms, including Roche 454 FLX Titanium, Illumina MiSeq, Illumina HiSeq2500, Life Technologies 5500xl, SOLID and Ion Torrent PGM), rely on massively parallel sequencing and imaging techniques and can yield several hundreds of millions to several hundreds of billions of bases per run (Shendure and Ji 2008). By greatly reducing limitations in generating sequence information, these technological advances have facilitated the functional characterization of genes and genomes and have started to provide a more comprehensive view of diversity and gene function in plants. In the last decade, genome sequencing and gene expression studies, including transcriptome sequencing, have enabled the identification of genes responsible for many traits (Brady and Provart 2007). Gene-based or functional nucleotide polymorphism, if identifiable within the gene of interest, is a more powerful FM compared with anonymous markers. The term 'functional marker' for DNA markers has been derived from such functionally characterized sequence motifs (Andersen and Lubberstedt 2003). These FMs have been called "perfect markers" and random DNA markers (RDMs) "non-perfect markers," including RFLP, AFLP and, to a degree, SSR markers (Varshney et al. 2005).

Functional markers

Although genetic markers have been used for over 20 years, they were typically in the form of cDNA–RFLPs (Graner et al. 1991) and functions could not be assigned. However, the Sanger sequencing of individual cDNA clones has been performed to determine how genetic sequence relates to function (Michalek et al. 1999). The generation of even greater numbers of genetic markers has been realized in crops in recent history; thanks to genome and transcriptome sequencing,

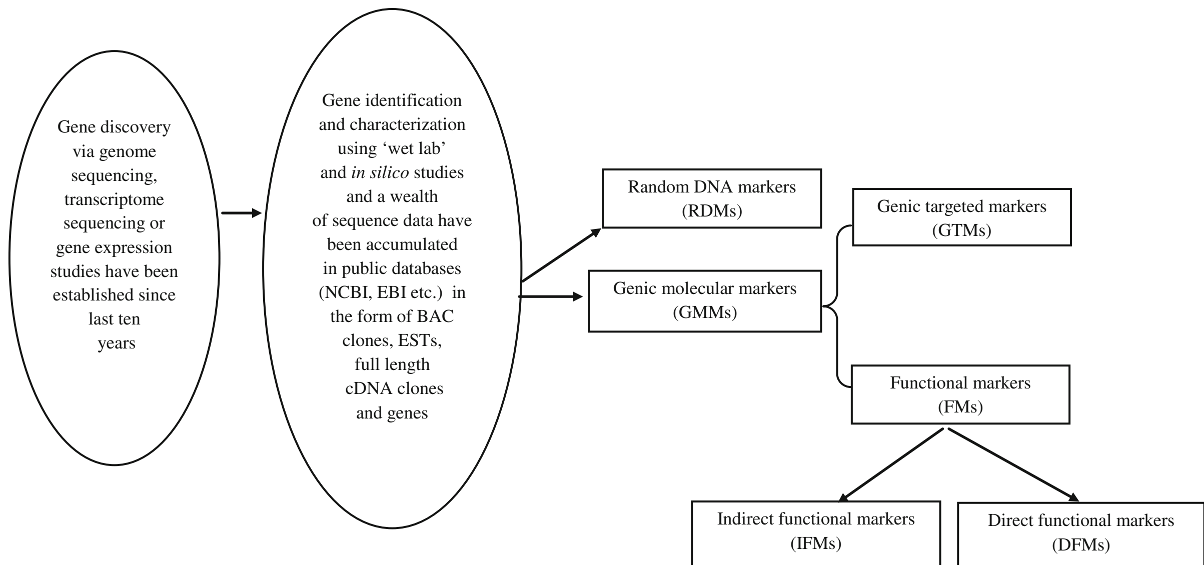

Fig. 1 Relationship and origins of various molecular- and functional markers

but also enabled by advances in bioinformatics (Gupta and Rustgi 2004). SNPs allow for high throughput analysis. However, first attempts at SNP discovery required further marker development and PCR optimization (Hazen and Kay 2003). As the numbers of genes were identified by new sequencing and in silico studies, SNP discovery became faster. Genetic molecular markers (GMMs) are constructed from genomic DNA, in coding or non-coding DNA, with or without known function (Andersen and Lubberstedt 2003), with or without phenotypic trait variation, (Aggarwal et al. 2007), and therefore they may or may not be FMs (Andersen and Lubberstedt 2003). FMs, associated with phenotypic trait variation, can be indirect (I) or direct (D) FMs. DFMs are well characterized markers closely associated for phenotypic trait variation, whereas IFMs are less characterized (Fig. 1).

FMs are unaffected by non-functional allelic variation and make it possible to identify an individual gene, which serve as good "translators" from genomic technologies into improved crop varieties (Andersen and Lubberstedt 2003). DNA markers that allow for identification of gene function and allelic variation for a particular trait are advantageous in breeding programs because of quick and powerful assays (Andersen et al. 2005). The major advantages of FMs compared with conventional markers are: (a) no recombination between marker and the gene of interest and thus no loss of information over time, (b) transfer of information to unknown materials resulting in a high predictive value, and (c) targeted exploitation of genetic diversity. When sequence motif to function has been assigned, plant breeders can use FMs to fix beneficial alleles independently of genetic background (Bagge et al. 2007). Useful target gene polymorphisms are prime FMs.

Discovery and development of FMs

The ability to develop an FM requires critical knowledge about the identity and sequence of the gene underlying the trait of interest. FM polymorphisms among alleles for genes of interest may be SNPs, insertions or deletions (INDELs), including partial or complete loss of the gene and on some occasions, different numbers of repeat motifs within SSRs (Gupta and Rustgi 2004; Varshney et al. 2005). After FMs are discovered and validated, then should be developed into functional assays, e.g., PCR for SNP or INDELs detection technologies (Syvanen 2001). Identifying the DNA sequences that demonstrate large effects on adaptive plant characteristics is fundamental for FM development towards breeding for traits such as biotic and abiotic stress tolerance (Andersen and Lubberstedt 2003). However, sequences may contribute to FM development only if they are polymorphic and correspond with a trait of interest. FMs can only be considered robust if they are validated against

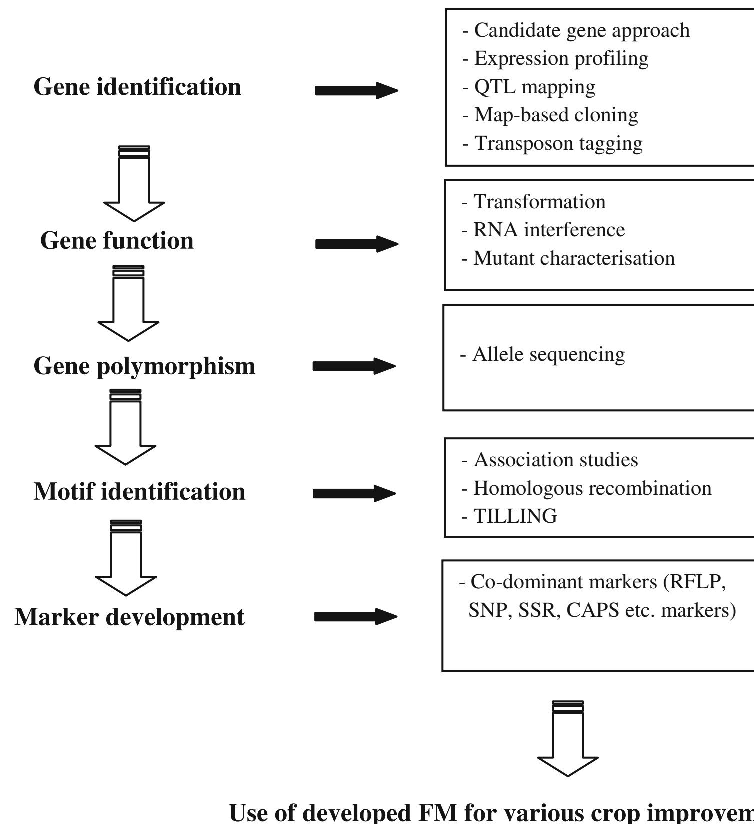

Fig. 2 Schematic outline for development of functional marker (FM) based on functional motif(s)

phenotyped populations in association studies (Andersen and Lubberstedt 2003).

Forward and reverse genetic approaches have also been used to discover genes underlying phenotypic variation in various crop plants. The methods of choice are indirect linkage disequilibrium (LD)-mapping or a direct approach such as 'EcoTILLING' (Comai et al. 2004). EcoTILLING, a variant of TILLING is an efficient method to reveal polymorphisms between different alleles (Mejlhede et al. 2006). Gene identification is achieved by map-based gene isolation, expression profiling, use of sequence homology to characterize genes from other species, or other methods such as transposon tagging (Borevitz and Nordborg 2003) (Fig. 2). Plant transformation for overexpression or knockdown analysis is necessary for gene functional characterization (Ingvardsen et al. 2008).

Challenges in functional marker development

In mapping populations with substantial intragenetic LD decay, adjacent polymorphic sites might still be incomplete LD, obscuring the unambiguous identification of causal quantitative trait nucleotides (QTN) or quantitative trait insertion-deletion mutations (QTINDEL) polymorphisms (Flint-Garcia et al. 2003). LD decay can lead to an overestimation of trait-associated polymorphisms. Another LD-related issue is the identification or development of optimal QTN or QTINDEL haplotypes, if several polymorphisms within the target gene affect the trait of interest. If not available in the characterized population, development of optimal QTN or QTINDEL allele combinations based on intragenetic recombination events might be difficult to achieve, even by use of large populations and intragenetic markers. Alternatively, exotic germplasm might provide a source for novel intragenetic combinations of QTN and QTINDEL alleles.

Comparable to QTL studies, a major concern after detection of QTN or QTINDELs in association studies is the transferability of information gained in one study to other situations. Other factors with potential impact on the detection of QTN or QTINDELs are epistatic and dominance as well as environment and genotype

by environment effects. The genetic effects of QTN or QTINDELs are background-, population-, and environment dependent. Thus, presence of a beneficial QTN or QTINDEL allele merely reflects a certain potential of trait expression, analogous to the risk concept in human genetic diseases, depending on the genetic effect and penetrance of the respective allele.

For application of FMs in crop improvement or breeding programs, it will be essential to test for negative pleiotropic side-effects (Thornsberry et al. 2001; Andersen et al. 2005) thereby yielding a better understanding of the nature of trait correlations or pleiotropic effects described for major genes. For example, various studies have found close genetic correlations between plant height and flowering time. Interestingly, the flowering time associated polymorphisms in *Dwarf 8*, a gene initially identified by its mutant allele leading to dwarfing, had no effects on plant height (Thornsberry et al. 2001; Andersen et al. 2005). In conclusion, composition of optimal haplotypes for genes shown to affect one or more traits of interest needs to take multiple traits into consideration.

Validation of FMs

FM validation also can include functional genetic maps for genes whose function has been characterized, their underlying QTL, and/or, closely linked to mutations. Numerous functional maps have been constructed in different crops, using genes from specific pathways or from ESTs of specific organs. The most useful polymorphisms are those that are causative for phenotypic trait variation, e.g., by association studies (Thornsberry et al. 2001). A more direct approach is the use of isogenetic genotypes produced by homologous recombination. Only a few FMs have been developed for targeted traits and have relevance in crop improvement programmes (Table 1). For successful incorporation of FMs in crop improvement programs, flexible and multi-parallel detection methods need to be available at low costs (Varshney et al. 2005).

Comparison of FMs with genetic molecular markers (GMMs), random DNA markers (RDMs) and genome selection (GS)

The fundamental difference between FMs and GMMs, RDMs and GS is their impact on the effectiveness of selection. FMs, GMMs and RDMs are limited to predicting the breeding values based on limited well-associated markers. In contrast, genomic selection is based on a dense set of markers from across the genome. Meuwissen et al. (2001) made a first step toward predicting a total genetic value using a genome-wide dense map of highly informative markers. GS uses the genome as the selective unit instead of using individual genetic loci that are associated with a trait. One obvious difference between GS and FM is the greater number of GS markers required for genotyping in a breeding population. In most GS studies, the entire set of markers used in the training population is applied to the breeding or validating populations. FMs are powerful in trait-by-trait selection owing to complete linkage with trait locus alleles, which reduces the amount of linkage drag when used in combination with closely linked markers (Frisch 2005).

FMs have several advantages over RDMs and GMMs (Table 2; Gupta and Varshney 2004). FMs allow for efficient selection of recombinants between the target gene and closely linked markers in large seedling populations (Frisch 2005), that could significantly reduce the number of backcross (BC) generations needed. The use of RDMs also bears the risk of being lost through genetic recombination even in the presence of flanking markers. Even GMMs can be lost through recombination (Varshney et al. 2007). Hence, the use of FMs is more efficient for gene identification and selection in breeding programs compared to RDMs and GMMs (Andersen and Lubberstedt 2003).

Application of FMs in plant breeding

As noted above, gene cloning and correct annotation facilitate FM development. FMs that reside within target genes themselves can be used with great reliability and efficiency to identify favorable alleles in breeding programs. The availability of FMs for agronomic traits would give plant breeders the possibility to select rare recombinants without the need for screening large number of plants (Xu et al. 2005). More generally, FMs are useful for: (i) more efficient fixation of alleles in populations; (ii) controlled balancing selection; (iii) screening for alleles in natural as well as breeding populations; (iv) combining of FM alleles affecting identical or different traits

Table 1 Candidate genes for functional marker (FM) development

Trait/resistance	Gene (s)	Crop	References
Low glutenin content	*Lgc1*	Rice	Chen et al. (2010)
Forage quality for digestibility	*Bm3*	Maize	Lubberstedt et al. (2005)
Lipoxygenase gene	*Talox-B1*	Wheat	Geng et al. (2012)
Powdery mildew	*Pm3*	Wheat	Tommasini et al. (2006)
Bacterial blight resistance	*xa5*	Rice	Iyer-Pascuzzi and McCouch (2007)
Bacterial blight resistance	*Xa21*	Rice	Song et al. (1997)
Grain quality	*Waxy*	Rice	Ayres et al. (1997)
Plant stature	*tb1*	Maize	Doebley et al. (1995)
Fragrance	*Badh2*	Rice	Shi et al. (2008)
Blast resistance	*Pit*	Rice	Hayashi et al. (2010)
Plant height	*Dwarf8 orthologs*	Several	Ikeda et al. (2001)
Powdery mildew	*NBS–LRR*	Barley	Madsen et al. (2003)
Leaf rust resistance	*Rph7*	Barley	Brunner et al. (2003)
Photoperiod response	*Phd-H1*	Barley	Dunford et al. (2002)
Fruit size	*w2.2*	Tomato	Nesbitt and Tanksley (2002)
Wide-compatible gene	*S5*	Rice	Sundaram et al. (2008
Low molecular glutenin weight	*Glu-D3 and Glu-B3*	Wheat	Zhao et al. (2007)
Stress response	*ERF transcription factors*	Several crops	Singh et al. (2002)
Vernalization requirements	*VFR2*	Oilseed	Kole et al. (2001)
Food quality	*GBSS*	Rice	Ayres et al. (1997)

in plant breeding; and (v) construction of linked FM haplotypes (Andersen and Lubberstedt 2003).

Marker-assisted selection

In recent years, marker-assisted selection (MAS), the use of genetic markers to facilitate the identification of favorable (or deleterious) alleles in a collection of diverse genotypes (Dubcovsky 2004), has been a boon to plant breeding as it allows the effective screening of large populations with less progeny testing (Francia et al. 2005). MAS has, more often than not, used non-genetic markers. Even when using tightly linked markers in MAS, the process loses efficacy when marker(s) recombine away from the gene of interest, thus resulting in misdiagnosis of traits of interest (Iyer and McCouch 2007). With the advances in the area of genomics and gene-derived markers, genetical genomics, and LD-based association mapping, FMs are becoming available for plant breeding in many crop species (Azhaguve et al. 2006). The genes then become the markers of interest and the process becomes more accurate and efficient (Perumalsamy et al. 2010). The location and sequence of candidate genes makes it possible to design allele-specific markers that readily lend themselves to automation (Collard and Mackill 2008). The false selection in MAS from recombination can be overcome by the use of FMs (Ingvardsen et al. 2008).

FMs have been developed within resistance genes and used in MAS, as exemplified by the L locus, conferring rust resistance in flax (Hausner et al. 1999), the *pvr1* gene for potyvirus resistance in *Capsicum sp.* (Yeam et al. 2005) and the *Pm3* gene for powdery mildew resistance in bread wheat (*Triticum aestivum*) (Tommasini et al. 2006). Cloning of bacterial blight resistance genes in rice (*Xa1*, *xa5, xa13, Xa21, Xa26* and *Xa27*) (Iyer and McCouch 2004;) made it possible to develop FMs. An FM for a recessive resistance gene, *xa5*, has been developed by Iyer and McCouch (2007). In addition, FMs for promoter-level differences between resistant and susceptible biotypes have been identified (Chu et al. 2007).

Genetic diversity

Assessing genetic diversity is essential in both conservation and optimizing germplasm for agriculture.

Table 2 Comparison between functional markers (FMs), genic molecular markers (GMMs), random DNA markers (RDMs) and genome selection (GS)

Feature	FMs	GMMs	RDMs	GS
Function of markers	Known	Known majority of times	Unknown majority of times	Unknown majority of times
Requirement of sequence data	Genes/ESTs data essential	Genes/ESTs data essential	Required for SSRs, SNPs; not required for RFLPs, RAPDs, AFLPs etc.	Sequence for SNP required
Selection of markers	Limited	Limited	Limited	Entire genomic markers
Function of polymorphic site	Functional motif	Not known	Not known	Not known
Utility in marker assisted selection	Great, as FMs derived from polymorphic sequences or sites within genes involved in phenotypic trait variation	Great, if the marker is derived from the gene, involved in expression of trait	High for SSRs, SNPs; moderately low for RFLPs, RAPDs, AFLPs etc.	Less effective in plant breeding
Labour involved	Less	Less	Moderately more	Moderately more in statistical analysis
Number of markers required	Low	Low	High for SSRs, SNPs; moderately low for RFLPs, RAPDs, AFLPs etc.	High
Costs of generation	Low	Low	Moderately high	High as more number of markers required
Utility of markers to exploit the functional diversity of genetic resources	High	Moderately low	Moderately low	High

As resources for conservation are limited, prioritization is often necessary. Molecular tools are helpful in identifying genes involved in a number of traits, including adaptive traits, and polymorphisms causing functional genetic variation (Beaumont and Balding 2004). Genetic diversity within and among closely related genotypes is essential for a rational use of genetic resources. Genetic resources can be defined as all materials that are available for the improvement of a cultivated plant species (Azhaguve et al. 2006). An understanding of germplasm diversity and genetic relationships among breeding materials is an invaluable aid for crop improvement strategies. RDMs are typically used for assessing genetic diversity at or below the species level (Tanksley and McCouch 1997) and are frequently used in MAS-based breeding (Xu 2003). Conservation biologists are less interested in random variation but in adaptive variation that allows for species survival and evolution. Neutral molecular markers can be used in genetic and LD-mapping approaches for this purpose and breeding (Varshney et al. 2007). However, FMs can evaluate germplasm for genetic diversity-based functionality and also by plant breeders. Alternatively, DNA-profiling can also be used to select genomic regions that imbue specific functions. However, with advent of DNA sequencing techniques, genotyping-by-sequencing (GBS) has emerged as a new concept, where SNPs in a large population allow the rapid and direct study of genetic diversity (Deschamp et al. 2012).

Germplasm evaluation

Evaluation of germplasm resources is required for the continuous improvement of crop plants, including the analysis of variation within and among germplasm (Hodgkin et al. 2001). To date vast genetic resources are available for crop species but are not well characterized at the molecular level. However, next generation sequencing is rapidly changing this situation for many crops (Kaisoon et al. 2008). Yet, the costs for DNA extraction, complexity reduction and bar-coding need to be decreased for systematic resequencing of germplasm collections. Keeping in view the above-mentioned costs and non-availability of techniques, FM can be used effectively to evaluate germplasm, construct heterotic groups, identify rare alleles, identify potential gaps in germplasm collections, monitor evolutionary changes, and find superior alleles (Tanksley and McCouch 1997).

High throughput analysis

NGS platforms and experiments now commonly yield several hundreds of billions of bases per run (Shendure and Ji 2008). Technological advances in genomics have facilitated the functionally characterization of genes with a comprehensive view of gene function in plants. If gene-based or functional nucleotide polymorphism FMs within a gene of interest are identified, they can be used for screening of a vast genetic/breeding material. Although NGS platforms are available for sequence discovery, capillary electrophoresis (CE) techniques can be used in high throughput screening (HTS). CE can be efficiently used with FMs for identification of gene of interests in large number of F_2 and F_3 individuals and other advanced breeding lines (Frazier 2004). FMs can give better results in multiplex PCR assays in foreground and background selections for simultaneously introgression and identification of genes of interest in marker assisted backcross breeding (MABB) programs (Salgotra et al. 2011).

Future perspectives and conclusions

Realistically, genomics is applied to plant breeding when it becomes easy and cost-effective to do so. RDMs and GMMs in MAS have become more commonplace and we can expect FMs to be used more frequently in the near future as their costs drop. In the post-genomics era when sequence data have already become available through genome or EST sequencing projects for some plant species and similar efforts are underway for many other plant species, it has been possible to develop the molecular markers directly from gene sequences deposited in public databases. Development of such FMs may speed up in coming years as these markers will prove promising in MAS and a useful resource for assessment of functional diversity in germplasm collection. Genomic selection has still not become a popular methodology in the field of plant breeding. However, functional genomics approaches such as transcriptomics, genetical genomics/expression genetics, HR, TILLING, association mapping and allele mining possess the potential to facilitate plant breeding practices. Exploitation of these functional genomics approaches to delimit the genes for a trait of interest may lead us towards genomics-assisted breeding for increasingly more crops in the near future.

Acknowledgments R. K. Salgotra is thankful to the Department of Science and Technology (DST) of Government of India for providing BOYSCAST Fellowship to carry out research at Plant Molecular Genetics Laboratory, University of Tennessee, USA.

References

Aggarwal RK, Hendre PS, Varshney RK, Bhat PR, Krishna KV, Singh L (2007) Identification, characterization and utilization of EST-derived genetic microsatellite markers for genome analyses of coffee and related species. Theor Appl Genet 114:359–372

Andersen JR, Lubberstedt T (2003) Functional markers in plants. Trends Plant Sci 8:554–560

Andersen JR, Schrag T, Zein I, Melchinger AE, Lubberstedt T (2005) Functional marker validation of polymorphisms in the maize *Dwarf8* gene affecting flowering time in European elite materials. Theor Appl Genet 111:206–217

Ayres NM, McClung AM, Larkin PD et al (1997) Microsatellites and a single-nucleotide polymorphism differentiate apparent amylose classes in an extended pedigree of US rice germ plasm. Theor Appl Genet 94:773–781

Azhaguve P, Saraswathi D, Sharma A, Varshney RK et al (2006) Methodological advancement in molecular markers to delimit the gene(s) for crop improvement. In: Azhaguvel P (ed) Advances in molecular markers for crop improvement. Global Science Books, Isleworthpp, pp 460–469

Bagge M, Xia X, Lubberstedt T (2007) Functional markers in wheat. Curr Opin Plant Biol 10:211–216

Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. Mol Ecology 13:969–980

Bernatsky R, Tanksley S (1986) Towards a saturated linkage map in tomato based on isozymes and random cDNA sequences. Genet 112:887–898

Bligh HFJ, Till RI, Jones CA (1995) A microsatellite sequence closely linked to the *waxy* gene of *Oryza sativa*. Euphytica 86:83–85

Borevitz JO, Nordborg M (2003) The impact of genomics on the study of natural variation in *Arabidopsis*. Plant Physiol 132:718–725

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphism. Am J Hum Genet 32: 314–331

Brady SM, Provart NJ (2007) Extreme breeding: leveraging genomics for crop improvement. J Sci Food Agric 87: 925–929

Brunner S, Keller B, Feuillet C (2003) A large rearrangement involving genes and low copy DNA interrupts the microcollinearity between rice and barley at the *Rph7* locus. Genet 164:673–683

Chen T, Meng-xiang T, Zhang Y, Zhen Z, Zhao L, Zhao Q, Lin J, Zhou L, Wang C (2010) Development of simple functional markers for low glutelin content gene 1 (*Lgc1*) in rice (*Oryza sativa*). Rice Sci 17:173–178

Chu Z, Yuan M, Yao J et al (2007) Promoter mutations of an essential gene for pollen development result in disease resistance in rice. Genes Dev 20:1250–1255

Collard BCY, Mackill DJ (2008) Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Phil Trans Royal Soc B: Biol Sci 363:557–572

Comai L, Young K, Till BJ et al (2004) Efficient discovery of DNA polymorphisms in natural populations by Eco-tilling. Plant J 37:778–786

Deschamp S, Llaca V, May GD (2012) Genotyping-by-sequencing in plants. Biology 1:460–483

Desmarais E, Lanneluc I, Lagnel J (1998) Direct amplification of length polymorphisms (DALP), or how to get and characterize new genetic markers in many species. Nucleic Acids Res 26:1458–1465

Doebley J, Stec A, Gustus C (1995) *Teosinte branched1* and the origin of maize: evidence for epistasis and the evolution of dominance. Genetics 141:333–346

Dubcovsky J (2004) Marker-assisted selection in public breeding programs: the wheat experience. Crop Sci 44:1895–1898

Dunford RP, Yano M, Kurata N, Sasaki T, Huestis G, Rocheford T, Laurie DA (2002) Comparative mapping of the barley *Phd-H1* photoperiod response gene region, which lies close to a junction between two rice linkage segments. Genetics 161:825–834

Flint-Garcia M, Thornsberry JM, Buckler ES (2003) Structure of linkage disequilibrium in plants. Annu Rev Plant Biol 54:357–374

Francia E, Tacconi G, Crosatti C et al (2005) Marker assisted selection in crop plants. Plant Cell, Tissue Organ Cult 82:317–342

Frazier R (2004) Capillary electrophoresis in food analysis. Labtech, Business Briefing, pp 42–45

Frisch M (2005) Optimum design of marker-assisted backcross programs. In: Loerz H, Wenzel G (eds) Biotechnology in agriculture and forestry, vol 55. Springer, Berlin, pp 319–334

Frisch M, Bohn M, Melchinger TE (1999) Minimum sample size and optimal positioning of flanking markers in marker-assisted backcrossing for transfer of a target gene. Crop Sci 39:967–975

Gale KR (2005) Diagnostic DNA markers for quality traits in wheat. J Cereal Sci 41:181–192

Geng H, Xia X, Zhang L, Qu Y, He Z (2012) Development of functional markers for lipoxygenase gene *Talox-B1* on chromosome 4 BS in common wheat. Crop Sci 52:568–576

Graner A, Jahoor A, Schondelmaier J, Siedler H et al (1991) Construction of an RFLP map of barley. Theor Appl Genet 83:250–256

Gupta PK, Rustgi S (2004) Molecular markers from the transcribed/expressed region of the genome in higher plants. Funct Integr Genomics 4:139–162

Gupta PK, Varshney RK (2004) Cereal genomics: an overview. In: Gupta PK, Varshney RK (eds) Cereal genomics. Kluwer Academic Press, Dordrecht, pp 639–650

Gupta PK, Varshney RK, Sharma PC, Ramesh B (1999) Molecular markers and their application in wheat breeding. Plant Breed 118:369–390

Gupta PK, Rustgi S, Kulwal PL (2005) Linkage disequilibrium and association studies in higher plants: present status and future prospects. Plant Mol Biol 57:461–485

Hausner G, Rashid KY, Kenaschuk EO, Procunier JD (1999) The development of codominant PCR/RFLP based markers for the flax rust resistance alleles at the L locus. Genome 42:1–8

Hayashi K, Yasuda N, Fujita Y, Koizumi S, Yoshida H (2010) Identification of the blast resistance gene *Pit* in rice cultivars using functional markers. Theor Appl Genet 121: 1357–1367

Hazen SP, Kay SA (2003) Gene arrays are not just for measuring gene expression. Trends Plant Sci 8:413–416

Hodgkin T, Roviglioni R, De Vicente MC, Dudnik N (2001) Molecular methods in the conservation and use of plant genetic resources. Acta Horticult 546:107–118

Ikeda A, Ueguchi-Tanaka M, Sonoda Y et al (2001) Slender rice, a constitutive gibberellin response mutant, is caused by a null mutation of the *SLR1* gene, an ortholog of the height-regulating gene GAI/RGA/RHT/D8. Plant Cell 13:999–1010

Ingvardsen CR, Schejbel B, Lubberstedt T (2008) Functional markers in resistance breeding. In: Luttge U, Beyschlag W, Murata J (eds) Progress in Botany. Springer, Berlin, pp 61–87

Iyer AS, McCouch SR (2004) The rice bacterial blight resistance gene *xa5* encodes a novel form of disease resistance. Mol Plant Microbe Interact 17:1348–1354

Iyer AS, McCouch SR (2007) Functional markers for *xa5* mediated resistance in rice (*Oryza sativa* L.). Mol Breed 19:291–296

Kaisoon O, Siriamornpun S, Meeso N (2008) Distinction between cereal genotypes based on the protein and DNA composition of the grain by capillary electrophoresis. World Applied Sci J 4:384–395

Kole C, Quijada P, Michaels SD, Amasino RM, Osborn TC (2001) Evidence for homology of flowering-time genes *VFR2* from *Brassica rapa* and FLC from *Arabidopsis thaliana*. Theor Appl Genet 102:425–430

Konieczny A, Ausubel FM (1993) A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. Plant J 4:403–410

Lubberstedt T, Zein I, Andersen J, Wenzel G, Krutzfeldt B, Eder J, Ouzunova M, Chun S (2005) Development and application of functional markers in maize. Euphytica 146:101–108

Madsen LH, Collins NC, Rakwalska M et al (2003) Barley disease resistance gene analogs of the NBSLRR class: identification and mapping. Mol Genet Genomics 269:150–161

McKay JR, Latta RG (2002) Adaptive divergence population: markers, QTLs and traits. Trends Ecol Evol 17:285–291

Mejlhede N, Kyjovska Z, Backes G, Burhenne K, Rasmussen SK, Jahoor A (2006) EcoTILLING for the identification of allelic variation in the powdery mildew resistance genes *mlo* and *Mla* in barley. Plant Breed 125:461–467

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829

Michalek W, Kunzel G, Graner A (1999) Sequence analysis and gene identification in a set of mapped RFLP markers in barley (*Hordeum vulgare*). Genome 42:849–853

Nesbitt TC, Tanksley SD (2002) Comparative sequencing in the genus *Lycopersicon*. Implications for the evolution of fruit size in the domestication of cultivated tomatoes. Genetics 162:365–379

Paran I, Michelmore RW (1993) Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theor Appl Genet 85:985–993

Perumalsamy S, Bharani M, Sudha M et al (2010) Functional marker-assisted selection for bacterial leaf blight resistance genes in rice (*Oryza sativa* L.). Plant Breed 129:400–406

Powell W, Machray G, Provan J (1996) Polymorphism revealed by simple sequence repeats. Trends Plant Sci 1:215–222

Salgotra RK, Millwood RJ, Agarwal S, Stewart CN Jr (2011) High-throughput functional marker assay for detection of *Xa/xa* and *fgr* genes in rice (*Oryza sativa* L.). Electrophoresis 32:2216–2222

Schmutz J, Steven B, Cannon Schlueter J et al (2010) Genome sequence of the palaeopolyploid soybean. Nature 463:178–183

Shendure J, Ji H (2008) Next-generation DNA sequencing. Nat Biotechnol 26:1135–1145

Shi W, Yang Y, Chen S, Xu M (2008) Discovery of a new fragrance allele and the development of functional markers for the breeding of fragrant rice varieties. Mol Breed 22:185–192

Singh KB, Foley RC, Sánchez LO (2002) Transcription factors in plant defense and stress responses. Curr Opin Plant Biol 5:430–436

Song WY, Pi LY, Wang GL, Gardner J, Holsten T, Ronald PC (1997) Evolution of the rice *Xa21* disease resistance gene family. Plant Cell 9:1279–1287

Sundaram RM, Vishnupriya MR, Biradar SK et al (2008) Marker-assisted introgression of bacterial blight resistance in Samba Mahsuri, an elite *indica* rice variety. Euphytica 80:411–422

Syvanen AC (2001) Accessing genetic variation: genotyping SNPs. Nat Rev Genet 2:930–941

Tanksley SD, McCouch SR (1997) Seed banks and molecular maps: unlocking genetic potential from the wild. Science 277:1063–1066

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

The Rice Chromosome 10 Sequencing Consortium (2003) In-depth view of structure, activity, and evolution of rice chromosome 10. Science 300:1566–1569

Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, Buckler IVES (2001) *Dwarf8* polymorphisms associate with variation in flowering time. Nat Genet 28:286–289

Tommasini L, Yahiaoui N, Srichumpa P, Keller B (2006) Development of functional markers specific for seven *Pm3* resistance alleles and their validation in the bread wheat gene pool. Theor Appl Genet 114:165–175

Varshney RK, Graner A, Sorrells ME (2005) Genomics-assisted breeding for crop improvement. Trends Plant Sci 10:621–630

Varshney RK, Mahendar T, Aggarwal RK, Borner A (2007) Genetic molecular markers in plants: development and applications. In: Varshney RK, Tuberosa R (eds) Genomics-assisted crop improvement: genomics approaches and platforms. Springer, Berlin, pp 13–29

Vos P, Hogers R, Bleeker M et al (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Wang DG, Fan JB, Siao CJ et al (1998) Large-scale identification, mapping and genotyping of single nucleotide polymorphisms in the human genome. Science 280:1077–1082

Williams J, Kubelik A, Livak K et al (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535

Xu Y (2003) Developing marker-assisted selection strategies for breeding hybrid rice. Plant Breed Rev 23:73–174

Xu Y, McCouch SR, Zhang Q (2005) How can we use genomics to improve cereals with rice as a reference genome? Plant Mol Biol 59:7–26

Yeam I, Kang BC, Lindeman W et al (2005) Allele specific CAPS markers based on point mutations in resistance alleles at the *pvr1* locus encoding eIF4E in *Capsicum sp*. Theor Appl Genet 112:178–186

Young ND, Debellé F, Oldroyd GE et al (2011) The *Medicago* genome provides insight into the evolution of rhizobial symbioses. Nature 480:520–524

Zhao XL, Ma EW, Gale EKR et al (2007) Identification of SNPs and development of functional markers for LMW-GS genes at *Glu-D3* and *Glu-B3* loci in bread wheat (*Triticum aestivum* L.). Mol Breed 20:223–231