Applications of molecular markers in breeding

C.T. Hash, D.A. Hoisington, R.K. Varshney, D. Kiambi, F. Sagnard, S. Senthilvel, T. Nepolean, S.P. Deshpande, Kassa Semegn, B.I.G. Haussmann, H.F.W. Rattunde, R. Bhattacharjee, S.L. Dwivedi, F.R. Bidinger, V. Vadez, R.P. Thakur, H.C. Sharma, and H.D. Upadhyaya

Summary

Molecular markers now provide appropriate complements to conventional breeding methods in most crops in ICRISAT's germplasm conservation and crop improvement mandates. Where appropriate molecular markers are available, they are effective and sometimes appropriate, tools for crop improvement research addressing biological components in agricultural production systems. Molecular markers offer specific advantages in the assessment of genetic diversity and in trait-specific crop improvement. Molecular markers are almost infinitely superior to conventional morphological marker genes for mapping or tagging gene blocks associated with economically important traits. Gene tagging and QTL mapping in turn permit marker-assisted selection (MAS) in backcross, pedigree and population improvement programs, which can be especially useful for crop traits that are otherwise difficult or impossible to deal with by conventional means. Near-isogenic products of marker-assisted backcrossing programs can in turn provide not only improved cultivars, but also useful genetic tools facilitating improved understanding of mechanisms of abiotic stress tolerance or mechanisms of host plant resistance to pests and diseases that are critical components of integrated crop management systems. Finally, when relatively high-density marker-based fingerprinting of elite breeding lines is possible, this can be combined with pedigree information and multi-environment performance data sets to greatly enhance the ability of conventional breeding programs to design new cultivars, identify desirable recombinants, and track factors controlling complex traits and trait combinations.

Introduction

Development and exploratory application of tools for marker-assisted breeding are much further advanced for sorghum than other ICRISAT crops (*e.g*., Feltus et al., 2006b; CCER document by Varshney *et al*.), but rapid progress is being made for chickpea and groundnut. Early progress in pearl millet provided the first applied product of marker-assisted selection in an ICRISAT crop, but a wider array of breeder-friendly markers are needed for technology adoption by pearl millet breeders. In finger millet, the recent publication of a marker-based genetic linkage map and SSR primer sequences provides an opportunity to initiate applications. In our other orphan crops (pigeonpea and other small millets) the very limited tool kits available do not currently allow more than limited use of markers in genetic diversity assessment.

Use of molecular markers in applied plant breeding programs

Molecular marker information, complemented by good quality phenotyping, can greatly facilitate the appropriate choice of parents for crosses for both inheritance studies and applied breeding. Molecular markers are also extremely useful for characterizing population structure of germplasm collections and sets of breeding lines. Molecular

markers are almost infinitely superior to conventional morphological marker genes for mapping or tagging gene blocks associated with economically important traits, whether these are inherited in a simple Mendelian fashion (*e.g.*, the d_2 dwarfing gene in pearl millet; Azhaguvel *et al*., 2003) or inherited in a more complex manner (*e.g.*, quantitative trait loci controlling the stay-green trait or *Striga* resistance in sorghum; Haussmann *et al*. 2002b, 2004). Gene tagging and QTL mapping in turn permit marker-assisted selection (MAS) in backcross, pedigree and population improvement programs, which can be especially useful for crop traits that are otherwise difficult or impossible to deal with by conventional means (*e.g.*, due to difficulties in obtaining repeatable field, greenhouse, or laboratory screening conditions "on-demand" as a result of natural variation in rainfall or pest pressure, or due to phytosanitary restrictions). Near-isogenic lines (NILs), the products of marker-assisted backcrossing programs in turn provide useful genetic tools facilitating improved understanding of mechanisms of abiotic stress tolerance (for reverse crop physiology studies) or mechanisms of host plant resistance to pests (for entomology, nematology and strigology studies) and diseases (for pathology studies). QTL mapping of yield and quality components can provide a better understanding of the basis for genetic correlations between economically important traits such as determining the role of linkage and/or pleiotropy for gene blocks controlling associated traits, *e.g.*,

- flowering time and biomass,
- inflorescence size and inflorescence number, or
- grain yield and grain protein content.

Specific genomic regions associated with QTLs of large effect for one target trait can be identified that have minimal effects on otherwise normally correlated traits, permitting improvement in the first trait that need not be accompanied by reduction(s) in the other(s). QTL mapping of economically important components of yield and quality can therefore be expected to facilitate more efficient incremental improvement of specific individual target traits in applied crop breeding programs. Finally, molecular markers can be used to more effectively discover and exploit evolutionary relationships between organisms, through comparative genomics and studies of synteny (*e.g.*, Devos *et al*., 2000; Ventelon *et al*., 2001; and Bowers *et al*., 2005).

Updated status of marker-assisted breeding for sorghum, pearl millet, finger millet, chickpea, groundnut, and pigeonpea The remainder of this document provides a brief update (compared to the more comprehensive 2006 in-house document on "Current status of marker-assisted breeding for chickpea, groundnut, pearl millet, pigeonpea and sorghum") on the application of marker-assisted breeding in terms of genetic diversity assessment, QTL mapping and marker-assisted selection for each of the six crops in ICRISAT's research mandate.

Sorghum applications

Diversity assessment: Diversity assessment of sorghum was among the first uses made of molecular markers in sorghum. One of the most interesting reports in this area is that by Jordan *et al*. (2003) on prediction of hybrid performance in grain sorghum using RFLP markers, in which it was found that markers associated with genomic regions governing

the stay-green trait and diversity at markers in a small number of other genomic regions, contribute to hybrid grain yield performance in Queensland, Australia.

SSR-based assessment of genetic diversity in sorghum has been a major activity in Subprogram 1 of the Generation Challenge Program (GCP) over the past three years, and ICRISAT contributed to this with partners at CIRAD (France) and CAAS (China). Our goal was to genotype a composite collection of 3000 accessions at 50 SSR loci that were well-distributed across nearly the full length of all 10 sorghum linkage groups. We have succeeded in genotyping some 3400 accessions across 39 loci, and diversity analysis of this data set has provided substantial information on population structure of sorghum germplasm globally and permitted selection of a highly informative reference set of 300 accessions for detailed study in future. Similar diversity studies are underway targeting sorghum germplasm from several NARS partners in Africa, with support from BMZ, USAID, the GCP, and as part of the African Biofortified Sorghum project. At Patancheru we are now initiating a similar diversity study of elite sorghum hybrid parental lines, with the intention of assessing opportunities to identify combining ability groups.

Collaborative efforts of CIRAD (France) with Diversity Arrays Pty Ltd (DArT, Australia) have resulted in DArT arrays for sorghum. DArT markers detected by these arrays have been demonstrated to possess the potential for undertaking the genome-wide diversity studies in sorghum at operational lower cost and in less time than would be needed for SSRs (J-F Rami, pers commun). DArT facilities are being established at ICRISAT as a part of the *Centre of Excellence in Genomics* (CEG), which will enhance our diversity assessment and utilization program.

QTL mapping and gene tagging: QTL mapping and other gene tagging procedures have identified flanking markers for several sorghum target traits, including a range of abiotic stress tolerances (primarily terminal drought tolerance and aluminum tolerance), host plant resistances to parastic weeds (*Striga* spp.), insect pests (aphids, shoot fly, midge and aphids), and diseases (of both foliage and panicles), and a range of traits related to grain quality, crop phenology (height and maturity), and yield components. Major genes controlling several traits of interest to sorghum breeders have been tagged with molecular markers and placed on the molecular marker-based linkage map of this crop. In the case of rust resistance, the gene responsible for one QTL detected has been identified and several sequence variants characterized (McIntyre *et al*., 2004). Similarly, the gene responsible for Al^{+++} tolerance has also been identified (J. Magalhaes, pers. comm.). Finally, *Rf1*, which is a major dominant gene required for restoration of pollen fertility in the *milo* (A_1) cytoplasmic-genetic male-sterility system used for most commercial sorghum hybrid seed multiplication globally, has recently been cloned (Klein *et al.*, 2005).

ICRISAT has initiated development of QTL mapping population(s) for sorghum salinity tolerance, sweet sorghum productivity traits, and grain mold resistance, and has six RIL populations for other target traits (*Striga* resistance + stay-green drought tolerance; shoot fly resistance; and stem borer resistance + midge resistance) that are well advanced in terms of both genotyping and phenotyping.

Marker-assisted selection: Harris *et al.* (2007) validated several stay-green QTLs from donor B35 by detailed phenotypic assessment of introgression lines developed by six generations of marker-assisted backcrossing. ICRISAT has initiated marker-assisted

backcrossing for several traits in sorghum including *Striga hermonthica* resistance from donor N 13 (BMZ funding); shoot fly resistance from donor IS 18551 (with ADB support, which ended in 2004) including PhD thesis research of SP Mehtre (2006) and T Jyothi (on-going); and the stay-green component of terminal drought tolerance from donors B35 and E 36-1 (initiated with core funding in 2001 and more recently supported in part by the Generation Challenge Program, the Water for Food Challenge Program and the Cereals Comparative Genomics Initiative), including MSc thesis research of Ramu (2003), Chandra Mouli (2004) and Venkateswararao (2005), and PhD thesis research of Kassahun Bannte (2007).

For shoot fly resistance components, BC3/4F3 lines with introgressed QTLs for seedling glossiness and leaf blade trichome density from donor IS 18551, show clear advantages compared to their recurrent parent BTx623 when screened against shoot fly using the fish meal interlard technique (PhD research of T Jyothi), but the improvement in resistance is not enough to protect sorghum seedlings from shoot fly damage when infestation levels are high. ICRISAT's most interesting sorghum MAB products to date are advanced generation backcross derivatives of RSG 03123 (PhD thesis research of Kassahun Bannte), which carries several stay-green QTLs from donor parent B35 in the genetic background of fully-senscent *rabi*-adapted breeding line R 16. Perhaps the most favorable results obtained to date from ICRISAT's evaluation of early generation marker-assisted breeding products in sorghum has been the prediction (by NIRS) of substantial improvements in stover digestibility for stay-green introgression lines compared to their senescent recurrent parent R 16.

Clearly there are now opportunities for application of marker-assisted breeding in sorghum improvement, including introgression of the stay-green trait into diverse tropically adapted elite farmer-preferred cultivars for Africa, Asia and Latin America; **however**, both larger numbers of polymorphic markers and further backcrossing and/or pedigree selection may be needed to obtain breeding products that can be directly useful to farmers. With the establishment of the DArT facilities at the CEG, we anticipate the acceleration of marker genotyping to support marker-assisted backcrossing in sorghum in real time and at relatively low cost.

Pearl millet applications

Diversity assessment: Marker-assisted diversity assessment in pearl millet began with the isozyme studies conducted by the ORSTOM team [Tostain (1994); Tostain *et al*. (1987); Tostain and Marchais (1989)]. This was followed by limited studies based on RAPDs (Chowdari *et al*., 1998a, 1998b), RFLPs (Busso *et al*., 2000; Bhattacharjee *et al*., 2002), AFLPs (vom Brocke *et al*., 2003) and SSRs (Budak *et al*., 2003; Mariac *et al.*, 2006). ICRISAT is now expanding SSR-based studies of genetic diversity in pearl millet in a Generation Challenge Program sponsored project on t SSR genotyping at 20 loci of up to 1000 accessions. This study will provide allele frequency estimation in heterogenous landrace germplasm accessions and improved open-pollinated varieties, as well as a range of genetic stocks and elite inbred lines. In addition, we are involved with national program partners in Africa in assessing diversity within and between important improved and landrace cultivars with support from BMZ and the GCP, and have initiated a project

at Patancheru to characterize elite B- and R-lines on the basis of morphological and SSR-based diversity.

In order to exploit the potential of SNP markers for diversity assessment, a pilot study involving screening of 24 inbred lines with hundreds of SSR, CISP (conserved intron spanning primers), and SSCP-based SNP assays demonstrated the high polymorphism information content of SNP assays in pearl millet. Thus now it is possible to undertake SNP assays in pearl millet. although the cost per data point is high compared to SSRs.

QTL mapping and gene tagging: Development of pearl millet mapping populations for various purposes was summarized by Hash and Witcombe (1994, 2001), and several new populations (targeting grain mineral content, salinity tolerance, *Striga* resistance and drought tolerance) are under development. QTL mapping has been reported in pearl millet for downy mildew resistance (Jones *et al*., 1995, 2002; Breese *et al*., 2002; Thakur and Hash, 2004), resistance to foliar diseases (Morgan *et al*., 1998); traits related to ruminant nutritional quality of grain crop residues (Hash *et al*., 2001, 2003); seedling heat tolerance (Howarth *et al*., 1994); terminal drought tolerance (Yadav *et al*., 2002, 2004; Bidinger *et al*., 2007); grain and stover yield potential and genotype x environment interaction of these that is associated with flowering time variation (Yadav *et al*., 2003); the domestication syndrome and its component traits (Poncet *et al*., 1998, 2000, 2002); apomixis (Ozias-Akins *et al*. 1998, 2003).

Marker-assisted selection: ICRISAT was the first to undertake molecular MAS in pearl millet, and this bore fruit in 2005 with the gazetted notification by the Government of India of release of pearl millet hybrid "HHB 67 Improved", which is the first non-genetically modified product of marker-assisted selection in any crop to be released for cultivation in India. We are presently conducting marker-assisted selection for downy mildew resistance (with national programs in India and Nigeria) as well as drought tolerance (India) and stover quality (India).

Results from ICRISAT's marker-assisted backcrossing

Bidinger *et al*. (2005) showed MAS for terminal drought tolerance was more effective than field selection when both selection systems were based on the same set of testcross hybrid field screening evaluations. Hash (2005) reported MAS has been effective in moving both improved levels of downy mildew resistance and improved levels of terminal drought tolerance into the genetic background of elite inbred pollinator line H 77/833-2 and these improvements are expressed in the hybrids of the QTL introgression lines. Serraj *et al*. (2005) indicated MAS has been effective in introgressing a major QTL for terminal drought tolerance from Iniadi landrace-derived breeding line PRLT 2/89-33 to the genetic background of elite early maturing, high tillering inbred pollinator H 77/833-2. Hybrids of the best introgression line, ICMR 01029, show consistent grain and stover yield advantages over their counterparts based on recurrent parent H 77/833-2, over a grain yield range of 800 kg/ha to 3500 kg/ha.

Chickpea applications

Diversity assessment: Working together in the GCP, ICRISAT and ICARDA have completed genotyping of a composite collection of 3000 accessions (Upadhyaya *et al*., 2006) at 50 SSR loci distributed across the chickpea genome. The information obtained

from this study has improved our understanding of the structure of genetic variation in the chickpea. We have selected a reference set of 300 accessions consisting of chickpea mini core (211 accessions) (Upadhyaya and Ortiz, 2001) and 89 other accessions capturing maximum diversity of composite collection and would be used. The reference collection will be used for extensive phenotyping for different traits and for genotyping using large number of markers for association studies. This reference collection is to be used for screening with about 15 candidate drought responsive genes for exploring the potential of candidate gene sequence-based association studies for drought tolerance.

QTL mapping and gene tagging: Major efforts on QTL mapping in chickpea have concentrated on disease resistance (Ascochyta blight and Fusarium wilt) and root traits. Two major QTLs for Ascochyta blight resistance are common in several reports and have also been validated at ICRISAT, so are ready to be used in MAS. Mapping of Fusarium wilt (FW) resistance in chickpea has identified four resistance genes organized in two clusters on LG 2, and another located on LG 5. SSR markers closely linked with these FW resistance genes have been identified and can be used to pyramid resistance genes for various races.

ICRISAT developed a recombinant inbred line (RIL) population from the cross (Annigeri x ICC 4958) and characterized these for root traits and SSR markers. An SSR marker (TAA 170) was identified for a major QTL that accounted for 33% of the variation for both root weight and root length (Chandra *et al*., 2004). Recently, the chickpea mini-core collection was evaluated for root traits and accessions showing larger variation than that found between Annigeri and ICC 4958 were selected for development of new mapping populations. Chickpea SSR markers, available in public domain as well as the ones developed at ICRISAT recently, are being screened to identify the polymorphic SSRs in these mapping populations. These mapping populations are expected to facilitate identification of markers for additional QTLs for root traits.

Marker-assisted selection: ICRISAT has only recently initiated backcrossing to move the putative major QTL for root mass and there are limited reports for other target traits from McKnight-supported collaboration involving workers in the USA and in India. However, to date there have been no results reported from MAS in chickpea.

Groundnut applications

Diversity assessment: Most reports to date have used molecular markers to assess relationships between cultivated groundnut (an allotetraploid) and its wild relatives (predominantly diploid) in order to identify progenitors of the cultigen. RAPD markers were used for this purpose by Hilu and Stalker (1995), whereas Dwivedi *et al*. (2001) used these to assess variation in the cultigen. AFLP markers (Gimenes *et al*., 2002), ISSR fingerprints (Raina *et al*., 2001) and SSRs(Ferguson et al., 2004a; Moretzsohn et al., 2004) have also been used successfully to assess diversity within cultivated groundnut. We have largely completed a major SSR-based diversity assessment in groundnut, with support from the Generation Challenge Program, which has targeted genotyping of 1000 accessions at 20 SSR loci, in collaboration with EMBRAPA, Brazil. Data analysis is in progress and results will be used to to select a reference set of 300 most diverse accessions.

QTL mapping and gene tagging: There has been limited work to date by ICRISAT and our partners on mapping resistance to rust, early leaf spot and bacterial wilt, in a project supported by the Asian Development Bank. Others have mapped resistance to the aphid vector of groundnut rosette disease (Herselman *et al*., 2004). Burow *et al*. (1996) reported RAPD markers associated with root-knot nematode resistance. Progress towards identification of markers associated with water use efficiency and drought tolerance has been made at ICRISAT in recent years. However, all of this work continues to be limited by scanty molecular marker polymorphism in cultivated groundnut. Nevertheless, preliminary analysis after genotyping one mapping population with about 80 polymorphic SSR markers has revealed a few molecular markers associated with the water-use efficiency and drought tolerance-related parameters.

Marker-assisted selection: Garcia *et al*. (1995) reported *ex-post* introgression analysis of an interspecific hybrid population, using RFLP and RAPD markers. Church *et al*. (2000) reported use of RFLP markers to select individuals homozygous for resistance to root-knot nematode. Otherwise there have been no reports of attempts to implement marker-assisted selection in groundnut to date, and ICRISAT has yet to establish marker-trait associations for traits of our interest.

Pigeonpea applications

Diversity assessment: A genotyping project aimed at understanding the heterogeneity, population structure as well as relationships among 1000 pigeonpea lines at 20 SSR loci has been completed and analysis of the dataset is underway at ICRISAT, with support from the Generation Challenge Program.

Due to extremely limited molecular marker polymorphism in cultivated pigeonpea, regardless of marker platform used to date, medium-term application of these tools for pigeonpea enhancement will likely be limited to tracking segregation of superior transgenic events and/or wild species introgressions in backcross and pedigree breeding.

Finger millet applications

Diversity assessment: A genotyping project aimed at understanding the population structure among 1000 finger millet lines at 20 recently mapped (Dida *et al*., 2007) SSR loci has been completed and analysis of the dataset is underway at ICRISAT, with support from the GCP.

Other minor millets

A genotyping project to assess population structure and relationships between 500 foxtail millet accessions is underway at ICRISAT with support from the GCP. Since SSR markers have not been isolated in foxtail millet, we have made use of primers for finger millet and pearl millet SSR markers. For example, after screening a set of 93 pearl millet and 31 finger millet SSR primer pairs on two foxtail genotypes (ISC 31, ISC 746), 39 SSR markers (26 pearl millet and 13 finger millet) that yielded amplicons from the foxtail millet genotypes were identified. Subsequently, a set of 20 polymorphic SSR markers was identified after screening these 39 SSR markers on 8 diverse foxtail millet genotypes.

Genotyping of the foxtail millet composite collection with the selected SSR markers is currently underway.

Looking to the future

Establishment of DArT platforms for ICRISAT crops at the CEG is expected to dramatically reduce the data point cost for molecular marker data, stimulating the use of this tool in our strategic and applied crop breeding programs. Databases of high-density marker fingerprints of elite breeding lines, combined with pedigree information and multilocational field performance data sets, will enable "breeding by design" (Peleman and van der Voort, 2003) to enhance the productivity, nutritive value and resilience of farmer- and market-preferred cultivars of the crops are most competitive in the SAT, and permit association mapping of genomic regions controlling economically important traits in the course of these breeding programs.

For additional information/clarification, contact Dr C T Hash (c.hash@cgiar.org)

References:

Bidinger *et al*. 2007. Identification of QTLs for grain yield of pearl millet [*Pennisetum glaucum* (L.) R. Br.] in environments with variable moisture during grain filling. Crop Sci. 47 (in press)

Dida *et al*. 2007. The genetic map of finger millet, *Eleusine coracana*. Theor. Appl. Genet. 114: 321-332.

Feltus *et al*. 2006b. Alignment of genetic maps and QTLs between inter- and intra-specific sorghum populations. Theor. Appl. Genet. 112: 1295-1305.

Harris *et al*. 2007. Sorghum stay-green QTL individually reduce post-flowering drought-induced leaf senescence. J. Exp. Bot. 58: 327-338.

Klein *et al*. 2005. Fertility restorer locus *Rf1* of sorghum (*Sorghum bicolor* L.) encodes a pentatricopeptide repeat protein not present in the collinear region of rice chromosome 12. Theor. Appl. Genet. 111: 994-1012.

Mariac *et al*. 2006. Diversity of wild and cultivated pearl millet accessions (*Pennisetum glaucum* [L.] R. Br.) in Niger assessed by microsatellite markers. Theor. Appl. Genet. 114: 49-58.

Ozais-Akins *et al*. 1998. Tight clustering and hemizygosity of apomixes-linked molecular markers in *Pennisetum squamulatum* implies genetic control of apospory by a divergent locus which may have no allelic form in sexual types. Proc. Natl. Acad. Sci. USA 85: 5127-5132.

Ozias-Akins *et al*. 2003. Molecular characterization of the genomic region linked with apomixis in *Pennisetum*/*Cenchrus*. Funct. Integr. Genomics 3: 94-104.

Peleman and van der Voort. 2003. Breeding by design. Trends Plant Sci. 8: 330-334.

Upadhyaya *et al*. 2006. Development of a composite collection for mining germplasm possessing allelic variation for beneficial traits in chickpea. Plant Genetic Resources 4: 13-19.

Upadhyaya and Ortiz. 2001. A mini core subset for capturing diversity and promoting utilization of chickpea genetic resources in crop improvement. Theor. App. Genet. 102: 1292-1298.

ICRISAT MAS Product Timeline