The Molecularization of Public Sector Crop Breeding: Progress, Problems, and Prospects

Sangam L. Dwivedi,[1,*] Jonathan H. Crouch,[2] David J. Mackill,[3] Yunbi Xu,[2] Matthew W. Blair,[4] Michel Ragot,[5] Hari D. Upadhyaya[6] and Rodomiro Ortiz[2]

[1]Agricultural Science Center at Clovis, 2346, SR288, Clovis, New Mexico 88101
[2]International Maize and Wheat Improvement Center (CIMMYT), Apdo 0660 Mexico, D.F., Mexico
[3]International Rice Research Institute (IRRI), DAPO, Box 3777, Metro Manila, Philippines
[4]Centro Internacional de Agricultura Tropical (CIAT), AA6713, Cali, Colombia
[5]Syngenta Seeds Inc., Stanton, Minnesota 55018
[6]International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502324, Andhra Pradesh, India

I. Introduction to Global Food Production and Major Breeding Challenges
II. Development of Markers for Assisting Selection
  A. Genetic Resources
  B. Genomic Resources
  C. Genetic Linkage Map
  D. Marker-Trait Associations from Analysis of Diverse Germplasm
III. Marker Validation and Refinement
  A. Markers for Simply Inherited Traits
  B. QTL Marker for Complex Traits
IV. Successful Applications of Marker-Assisted Genetic Enhancement in Public Sector Breeding Programs
  A. Resistance to Biotic Stresses
  B. Tolerance to Abiotic Stresses
  C. Agronomic and Seed Quality Traits
  D. Specific Challenges for Alien Gene Introgression
V. Successful Application of Marker-Assisted Genetic Enhancement in Private Sector Breeding Programs
VI. Impact of Marker-Assisted Genetic Enhancement
  A. Enhanced Selection Power

*Present Address: International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502324, Andhra Pradesh, India.

Advances in Agronomy, Volume 95

0065-2113/07 $35.00
DOI: 10.1016/S0065-2113(07)95003-8

B. Reduced Cost, Increased Feasibility, Time Savings, and Parental Selection
C. Overview of Products from Molecular Breeding
VII. Approaches to Enhance the Efficiency and Scope of Molecular Breeding
A. Studying the Molecular Basis of Heterosis
B. Fine-Mapping, Cloning, and Pyramiding of QTL Associated with Improved Agronomic Traits
C. Expression QTL Mapping
D. Simulation and Modeling of MAS
VIII. The Role of Computational Systems in Molecular Breeding Programs
A. Germplasm Evaluation
B. Managing Breeding Populations
C. Genetic Map Construction
D. Identifying Marker-Trait Associations
E. Marker-Assisted Selection
F. GEI Analysis
G. Breeding Design and Simulation
H. Information Management and Integrated Tools
IX. Future Prospects for the Molecularization of Public Crop Improvement
Acknowledgments
References

Molecular markers and genetic maps are available for most important food crops. Marker-trait associations have been established for a diverse array of traits in these crops, and research on marker/quantitative trait loci (QTL) validation and refinement is increasingly common. Researchers are now routinely using candidate gene-based mapping and genome-wide linkage disequilibrium and association analysis in addition to classical QTL mapping to identify markers broadly applicable to breeding programs. Marker-assisted selection (MAS) is practiced for enhancing various host plant resistances, several quality traits, and a number of abiotic stress tolerances in many well-researched crops. Markers are also increasingly used to transfer yield or quality- enhancing QTL alleles from wild relatives to elite cultivars. Large-scale MAS-based breeding programs for crops such as rice, maize, wheat, barley, pearl millet, and common bean have already been initiated worldwide. Advances in ''omics'' technologies are now assisting researchers to address complex biological issues of significant agricultural importance: modeling genotype-by-environment interaction; fine-mapping, cloning, and pyramiding of QTL; gene expression analysis and gene function elucidation; dissecting the genetic structure of germplasm collections to mine novel alleles and develop genetically structured trait-based core collections; and understanding the molecular basis of heterosis. The challenge now is to translate and integrate this knowledge into appropriate tools and methodologies for plant breeding programs. The role of computational tools in achieving this is becoming increasingly important. It is expected that harnessing the outputs of genomics research will be an important component in successfully addressing the challenge of doubling world food production by 2050.

I. INTRODUCTION TO GLOBAL FOOD PRODUCTION AND MAJOR BREEDING CHALLENGES

Worldwide cereal, legume, oilseed, root and tuber, and plantain and banana crops are grown annually on 1068 million ha with a total production of 3238 million metric tons (Mt) (http://faostat.fao.org/site/340/default.aspx, February 2006); of which cereals contribute 68.6%, roots and tubers 22.0%, legumes 1.9%, oilseeds 4.2%, and plantain and banana 3.3%. Asia is the largest contributor to cereal production (45.9%) followed by North and Central America (21.0%) and Europe (20.5%), while Africa and South America each contributes about 5%. North and Central America (37.3%) and South America (34.9%) dominate legume production, while Asia contributes only 18.2%. Both Africa and Europe contribute about 3% of legume production. For oilseeds, Asia is the largest producer (48.8%) followed by Europe (21.3%), Africa (16.0%), and North and Central America (9.0%), while South America contributes 3.3%. Asia, Africa, and Europe together contribute about 88% to the world production of root and tuber crops, while Africa predominates in plantain and banana production (71.8%) followed by South America (18.1%) and North and Central America (6.9%). Significant trends in production during the period from 1961 to 2005 were noted (Table I). For example, maize has overtaken both wheat and rice; soybean maintains its predominant position among legume crops, although peanut (groundnut) production doubled while beans production slowly but steadily increased by 58%; and substantial increases in cassava and banana production were noted. In contrast, worldwide oat production declined substantially. Millet production remained stagnated, while sorghum production declined by 21% since its peak production in the first half of the 1980s. Across regions, wide variation exists in productivity of these crop commodity groups: cereals from 1.24 t ha^{-1} in Africa to 5.40 t ha^{-1} in North and Central America; legumes from 0.55 t ha^{-1} in Africa to 2.60 t ha^{-1} in North and Central America; oilseeds from 0.78 t ha^{-1} in Africa to 1.76 t ha^{-1} in Europe; root and tuber crops from 8.23 t ha^{-1} in Africa to 24.52 t ha^{-1} in North and Central America; and plantain and banana from 5.61 t ha^{-1} in Africa to 10.05 t ha^{-1} in North and Central America. Many factors have contributed to increased productivity of these food crops: the development of higher yielding cultivars, increased application of fertilizers, herbicides for weed control, insecticides and fungicides for the control of pests, and increases in irrigation.

Average increases in productivity vary considerably between crops: for example, maize (except for the period from 1986 to 1990), rice, and wheat productivity has increased steadily throughout the last 45 (1961–2005) years (Table II). In contrast, there were only marginal increases in barley and oat

Table I
World-Wide Average Production of the Major Cereal, Legume, Root and Tuber, and Banana and Plantain Crops[a]

Crop	Average production (million Mt) (1961–2005)								
	1961–1965	1966–1970	1971–1975	1976–1980	1981–1985	1986–1990	1991–1995	1996–2000	2001–2005
Banana and plantain									
Banana	23.3	28.8	31.8	34.6	38.0	44.3	52.8	61.2	70.1
Plantain	14.0	17.3	21.5	23.3	23.1	25.3	27.9	29.8	32.3
Cereal									
Barley	111.8	110.9	139.1	161.9	162.5	171.5	161.5	141.8	143.0
Maize	214.3	261.8	317.7	386.6	435.7	458.9	518.2	597.9	650.8
Millet	2.5	2.9	2.8	2.6	2.8	2.8	2.7	2.8	2.8
Oat	46.8	50.5	49.3	45.5	44.9	40.0	33.3	28.1	26.1
Rice	241.3	287.9	329.8	374.9	442.6	489.8	532.4	587.2	595.7
Sorghum	4.5	5.5	6.1	6.4	7.0	6.3	6.0	6.2	5.8
Wheat	247.7	308.9	354.9	421.8	485.6	532.9	549.2	593.0	594.5
Legume									
Beans	11.8	12.0	12.7	12.9	15.0	15.6	16.2	16.6	18.7
Broad bean	5.5	4.4	4.3	4.3	4.2	4.3	3.3	3.6	4.3
Chickpea	7.0	6.3	6.2	6.8	6.4	6.9	7.6	8.5	8.0
Cowpea	1.0	1.1	1.1	1.1	1.1	1.6	2.3	3.2	3.7
Lentil	0.9	1.0	1.1	1.3	1.7	2.5	2.4	2.9	3.4
Pea	10.7	9.0	8.9	9.2	10.5	14.8	13.3	11.4	10.9
Peanut	15.5	16.8	18.1	17.6	19.8	23.1	26.5	32.4	35.4
Pigeon pea	1.8	1.8	2.0	2.1	2.5	2.7	2.7	2.9	3.1
Soybean	28.6	40.3	53.8	75.3	90.4	100.7	119.3	150.8	192.6
Root and tuber									
Cassava	78.3	92.1	103.3	119.9	130.5	144.3	162.4	166.6	193.3
Potato	269.8	291.7	282.4	276.4	273.9	275.4	278.4	308.9	319.5
Sweet potato	100.6	123.8	136.0	140.8	129.8	124.5	128.0	136.7	131.4
Yam	9.4	14.4	13.5	12.0	11.8	15.9	30.6	35.9	39.2

[a](http://faostat.fao.org/faostat/collections?version=ext&hasbulk=0&subset=agriculture).

Table II
World-Wide Average Productivity of the Major Cereal, Legume, Root and Tuber, and Banana and Plantain Crops[a]

Crop	Average production (t ha^{-1}) (1961–2005)								
	1961–1965	1966–1970	1971–1975	1976–1980	1981–1985	1986–1990	1991–1995	1996–2000	2001–2005
Banana and plantain									
Banana	10.81	11.33	11.49	12.68	13.08	13.34	14.16	15.36	15.74
Plantain	5.42	5.98	6.26	5.91	5.67	5.93	5.97	6.23	6.27
Cereal									
Barley	1.48	1.75	1.87	2.00	2.05	2.26	2.21	2.41	2.54
Maize	2.01	2.34	2.69	3.10	3.46	3.50	3.82	4.29	4.56
Millet	0.58	0.66	0.66	0.68	0.76	0.76	0.73	0.77	0.80
Oat	1.45	1.67	1.67	1.70	1.76	1.79	1.75	1.98	2.13
Rice	1.99	2.22	2.41	2.63	3.08	3.36	3.61	3.84	3.93
Sorghum	0.96	1.10	1.27	1.38	1.50	1.39	1.36	1.41	1.33
Wheat	1.18	1.42	1.62	1.82	2.08	2.37	2.50	2.69	2.78
Legume									
Beans	0.49	0.51	0.54	0.54	0.59	0.60	0.65	0.66	0.71
Broad bean	1.04	0.93	1.05	1.14	1.25	1.42	1.47	1.53	1.61
Chickpea	0.59	0.61	0.62	0.65	0.66	0.70	0.72	0.76	0.78
Cowpea	0.31	0.21	0.25	0.34	0.32	0.35	0.34	0.36	0.38
Lentil	0.56	0.59	0.60	0.60	0.68	0.77	0.81	0.82	0.88
Pea	0.99	1.09	1.10	1.24	1.25	1.57	1.76	1.82	1.67
Peanut	0.85	0.87	0.90	0.95	1.06	1.17	1.24	1.40	1.42
Pigeon pea	0.65	0.63	0.68	0.70	0.73	0.74	0.67	0.70	0.70
Soybean	1.16	1.42	1.53	1.65	1.75	1.83	2.01	2.18	2.28
Root and tuber									
Cassava	7.68	8.22	8.34	9.00	9.41	9.85	9.81	10.13	10.83
Potato	12.34	13.82	14.03	14.51	14.70	15.35	15.37	16.12	16.81
Sweet potato	7.94	10.62	11.35	11.94	13.53	13.70	14.03	14.85	14.51
Yam	7.50	8.39	7.97	8.58	6.56	8.25	10.21	9.82	9.14

[a](http://faostat.fao.org/faostat/collections?version=ext&hasbulk=0&subset=agriculture).

productivity during the same period, while millet productivity has stagnated and average sorghum productivity declined. For the legumes, cowpea remained the lowest yielder, while lentil, chickpea, pigeon pea, and beans productivity remained stagnated for most part but broad bean yields steadily increased. In contrast, peanut productivity increased by 67%, while soybean yields consistently increased and remained the top yielder among the legumes. Three distinct patterns have emerged in the productivity of root and tuber and plantain and banana: plantain yield remained stagnant while cassava and yam yield moderately increased. In contrast, substantial increases in productivity were observed for potato, sweet potato, and banana, with potato being the highest yielder among these vegetatively propagated crops.

Both abiotic and biotic constraints limit the productivity of all food crops: for example, drought, salinity, temperature (both extreme high and low), phosphorous limitation, and aluminum toxicity in acidic soils among the abiotic stresses, and insect pests and fungal, bacterial, and virus diseases among the biotic stresses are the major constraints to sustainable production of these crops. The biotic constraints of greatest effect worldwide include bacterial blight (BB) and blast and several virus diseases in rice; rust in wheat, barley, soybean, and common bean; powdery mildew and *Fusarium* head blight (FHB) in wheat and barley; *Barley mild mosaic virus* (BaMMV) complex, *Barley yellow dwarf virus* (BYDV), and Russian wheat aphid in barley; stem borer in rice, corn, and sorghum; *Maize streak virus* in corn; downy mildew in corn, pearl millet, and sorghum; nematodes in soybean; rust and leaf spots in groundnut; common bacterial blight (CBB) and several virus diseases in common bean; anthracnose in common bean, cassava, and yam; *Ascochyta* blight in pea and chickpea; *Cassava mosaic virus* and *Cassava brown streak* virus in cassava; *Yam mosaic virus* (YMV) in yam; late blight and several virus diseases in potato; and Black Sigatoka in banana and plantain. Additionally, parasitic weeds, for example *Striga*, *Electra*, and *Orobanche*, seriously limit the production of cereal and legume crops in Africa and Asia. There are many documented cases where these constraints alone or in combination have caused havoc to production and famine in many parts of the world. Some fungal diseases of crop plants also produce mycotoxins that are detrimental to human and animal health. For example, aflatoxin (caused by *Aspergillus flavus*) in corn and peanut, and deoxynivalenol (DON) (caused by FHB) in wheat and barley pose serious risk to the safety of human food and livestock feed.

Conventional breeding is undoubtedly responsible for substantial gains in the productivity of the many food crops, for example, the introduction of dwarfing genes (*Sd1* in Dee Geo Woo Gen rice and *Rht1* and *Rht2 in* Norin 10 wheat) and hybrid maize tolerant to high crop density adapted these crops to intensive agriculture worldwide in what is collectively known as the Green

Revolution. The Green Revolution helped many developing countries to produce the needed food for their growing population. However, environmentalists, economists, and social scientists criticized this technology for what they assessed as its shortcomings (e.g., use of fertilizers and pesticides as well as monoculture of a few crop cultivars), or who benefited (Swaminathan, 2006). Additionally, only limited progress has been achieved through conventional breeding to address the production constraints with genetically more complex traits such as tolerance to drought and salinity, resistance to pathotypes (in the case of diseases) and biotypes (in the case of pests) with complex inheritance, low heritability, and high genotype-by-environment interaction (GEI).

From 5.66 billion in 1995, the world population will reach 7.5 billion in 2020, with developing and developed countries' share accounting for 97.5% and 2.5%, respectively (Pinstrup-Anderson *et al.*, 1999). The global demand for cereals during the same period will increase by 39% to 2466 Mt; meat by 58% to 313 Mt; and root and tuber crops by 37% to 864 Mt. The large increases in food demand will result not only from population growth but also from urbanization, income growth, and changes in lifestyles and food preferences. The developing countries will account for about 85% of the increase in global demand for cereal and meat. A demand-driven "livestock revolution" is under way in the developing world and the demand for meat in the developing world is projected to double between 1995 and 2020 (Pinstrup-Anderson *et al.*, 1999). In response to the strong demand for meat products, demands for cereals for feeding livestock will double in developing countries. Demand for maize in developing countries will increase much faster than for any other cereal and will overtake demand for rice and wheat by 2020. To meet this demand, the world's farmers will have to produce 40% more grain in 2020. Increases in cultivated area are expected to contribute only about one-fifth of the increase in global cereal production between 1995 and 2020, so substantial improvements in crop yields will be required to bring about the necessary production increases. This will need to be achieved through a combination of genetic improvements in cultivar and improved agronomic practices. However, without substantial and sustained additional investment in agricultural research and delivery mechanism, it will become more and more difficult to maintain, let alone increase, yields of these crops in the longer term. As gains from conventional breeding are gradually exhausted, further yield growth will be generated as conventional breeding is combined with wide-crossing, genomics, and transgenic technologies to tailor crop cultivars with multiple resistance to biotic and abiotic stresses and adapted to diverse agroecological niches (Rosegrant *et al.*, 1995).

Crop biomasses are potential raw materials for the production of agricultural biofuels (ethanol from sucrose or starch derived from vegetative

biomass or grains) or bio-diesel (from vegetable oils and animal fat). Preliminary work has already demonstrated that a great potential exists to develop cellulose-based bioenergy systems. This could lead to more demand for cereals (in terms of biomass and grains) for biofuel and oilseeds for biodiesel production that will compete with the demand of these crop commodity groups for food and feed purposes. Multipurpose crops combining food, feed, fiber, and biofuel traits are therefore needed to respond to these market changes (IFPRI 2020 vision for food, agriculture, and the environment).

Since the development of DNA marker technology in the 1980s, it has undergone tremendous advances in terms of marker development, genetic maps, functional and comparative genomic linkages, utilization of genome sequencing, and scale and cost of application technologies. As new developments unfold, the power of genomics to facilitate a more genetic-led approach to plant breeding will be one of the most important advances enabling crop improvement to solve some of the world's most difficult problems regarding sustainable agricultural production in many parts of the world. Molecular markers can now be routinely applied to assess and enhance diversity in germplasm collections, to identify genes that control key traits, and to introgress valuable traits from new sources. The ability to introgress beneficial genes under the control of specific promoters through transgenic approaches is another milestone on the path to targeted approaches to crop improvement for which genomic sciences have already identified a vast array of genes that have exciting potential for crop improvement (Delmer, 2005).

There are several generic reviews on plant genomics with respect to genetic mapping, quantitative trait loci (QTL) analysis, molecular breeding, and modeling genetic variability of plant responses to environmental stresses (Asíns, 2002; Dekkers and Hospital, 2002; Dwivedi *et al.*, 2005; Guo, 2000; Mohan *et al.*, 1997; Stuber *et al.*, 1999; Tardieu, 2003; Varshney *et al.*, 2005a). Similarly, there are a number of crop-specific reviews on applied genomics, including rice (Ashikari and Matsuoka, 2002; Mackill and McNally, 2004; Xu, 2003), wheat (Koebner *et al.*, 2001), barley (Koebner *et al.*, 2001; Thomas, 2003), common bean (Broughton *et al.*, 2003; Miklas *et al.*, 2006a), cowpea (Ortiz, 2003), peanut (Dwivedi *et al.*, 2003), plantain and banana (Crouch *et al.*, 1998b), yam (Mignouna *et al.*, 2003a), and potato (Barone, 2004). However, in this chapter, we focus on how progress in plant genomics has offered new opportunities for plant breeders and the extent to which these have been successfully applied in real breeding programs. We then go on to review the essential allied technologies that will be required for successful molecular breeding programs and synthesize the problems and prospects for a future technology-assisted crop improvement paradigm.

II. DEVELOPMENT OF MARKERS FOR ASSISTING SELECTION

A. Genetic Resources

Plant genetic resources (PGR) are the basic raw materials required to power current and future progress in crop improvement programs. The use of PGR in crop improvement is one of the most sustainable ways to conserve valuable genetic resources for the future, and simultaneously to increase agricultural production and food security. Key to successful crop improvement is a continued supply of genetic diversity including new or improved variability for target traits. The centers of the Consultative Group on International Agricultural Research (CGIAR) have the responsibility to collect, preserve, characterize, evaluate, and document the genetic resources of the cultivated and wild relatives of the cereals (barley, maize, millets, oat, rice, sorghum, and wheat), legumes (Bambara groundnuts, chickpea, common bean, cowpea, faba bean, grasspea, lentil, pea, peanut, pigeon pea, and soybean), roots and tubers (Andean root and tuber crops, cassava, potato, sweet potato, and yam), and *Musa* (both banana and plantain). Additionally, they have genetic improvement programs that integrate these genetic resources into elite breeding material for use in national cultivar development programs. These germplasm collections are under the aegis of FAO held in trust, and available to researchers globally for diverse use. Collectively, the CGIAR centers possess about 600,000 samples from about 370,000 cultivated accessions, 34,000 wild and weedy accessions, and nearly 177,000 accessions from an uncertain (unknown) category (Table III). The largest representation is of the cereals (64.65%) followed by legumes (30.28%), roots and tubers (4.82%), and *Musa* (0.25%). The CGIAR System-wide Information Network for Genetic Resources (SINGER) links the genetic resources information systems of individual CGIAR centers around the world, allowing them to be accessed and searched collectively. SINGER contains key data of more than half a million individual accessions of crops, forage, and agroforestry genetic resources held in the center genebanks (http://www.singer.cgiar.org/). The remaining germplasm are stored in other international, regional, and national genebanks, many of which collaborate closely with CGIAR centers.

Crop germplasm collections held in genebanks are the best genetic resources for detailed characterization of important traits such as tolerance to biotic and abiotic stresses, yield, nutrition, and grain quality. These existing diverse germplasm collections are "gold mines" for analysis of allelic diversity. The efficiency of crop improvement programs, whether conventional breeding alone or powered with marker-assisted selection (MAS), depends on the

Table III
Wild and Cultivated Accessions of the Andean Root and Tubers, Banana, Barley, Bean, Cassava, Chickpea, Faba Bean, Grasspea, Lentil, Maize, Minor Millets, *Musa*, Oat, Pea, Peanut, Pearl millet, Pigeon pea, Potato, Rice, Sorghum, Soybean, Sweet potato, Wheat, and Yam Preserved in CGIAR Gene Banks

	No. of accessions stored in CGIAR's gene bank			
Crop	Cultivated	Wild and weedy	Unknown	Total
Andean root and tuber crops	1042	58		1100
Banana[a]	979	178	283	1440
Barley	17,759	79	6382	24,220
Barley (wild *Hordeum*)	15	1817		1832
Barnyard millet	743			743
Cassava	3009	7137	679	10,825
Chickpea	30,748	419		31,167
Common bean	31,263	2272		33,535
Cowpea	11,268	1779	14,494	27,541
Faba bean BPL			5285	5285
Faba bean	2952	3025	6602	12,579
Finger millet	5844	105		5949
Foxtail millet	1481	54		1535
Grasspea	379	1116	1815	3310
Kodo millet	658			658
Lablab bean			42	42
Lentil	2646	498	6825	9969
Lima bean			40	40
Little millet	466			466
Maize	21,993	177		22,170
Mung bean			122	122
Oat	679	16		695
Pea	1658	176	4271	6105
Peanut	14,966	453		15,419
Pearl millet	20,844	750		21,594
Pigeon pea	13,077	555		13,632
Potato	4579	2108		6688
Proso millet	842			842
Rice (*indica* and *japonica*)	49,644	644	67047	11,7335
Rice (wild)	33	3789	4020	7842
Sorghum	36,975	418		37,393
Soybean	193		16985	17,178
Sweet potato	4717	1403		6120
Wheat (bread and durum)	85,152	1	41,469	126,622
Wheat (primitive)	525	5	84	614
Wheat (*Triticum* and *Aegilops*)	29	5126	12	5167
Yam	2897	17	362	3276
Total	370,055	34,175	176,819	581,050

[a]Also contains accessions from INIBAP.
(http://singer.grinfo.net/).

accuracy and precisions of evaluation techniques used to generate appropriate phenotyping data. However, the size of most crop-related global germplasm collections is simply too vast for systematic evaluation in replicated multilocational trials. Moreover, the diversity of adaptation and major phonological traits of such material highly confounds attempts to generate directly comparable agronomic performance data. Undoubtedly, the robustness of phenotyping is the single most important constraint for effective selection of appropriate new genetic resources, particularly for abiotic stress tolerance and yield potential. Genomic analysis will have a major role to play in helping to identify subsets of germplasm that are small enough to allow precision phenotyping of replicated multilocational trials for groups of accessions with sufficient homogeneity of phenological and adaptation backgrounds, yet maximum diversity for the target trait: *genetically structured trait-based core collections*.

The development of core collections has been shown to be a particularly powerful strategy for providing crop breeding programs with a systematic yet manageable entry point into global germplasm resources. Core collections are a cost-effective means of identifying accessions with desirable agronomic traits as well new sources of disease and pest resistance or abiotic stress tolerance. Core collections are usually constituted from the 10% of the entire germplasm collection that represents at least 70% of the collections variability in that collection (Brown, 1989). These representative accessions in these core collections are identified based on all available information, including passport data plus botanical and agronomical descriptors. In this way, the development of a core collection has the advantage of displaying much of the phenotypic variability conserved in the genebank in a limited number of accessions. This allows researchers to identify trait-based hot spots, for example, for new sources of resistance to new isolates or biotypes of diseases and pests at a substantially lower cost than systematically evaluating the entire collection. However, this approach can only be as good as the phenotypic data on which it is based, and thus may not be a more effective route for identifying the best genetic variability for new traits. In this case, it is hoped that a new generation of core collections based on combined phenotypic and genotyping analysis may be more effective. Conventional core collections are available in barley, cassava, cowpea, finger millet, maize, *Musa*, pearl millet, potato, quinova, rice, sorghum, sweet potato, West African yam, and wheat (Table IV), and for several legumes crops (Dwivedi *et al.*, 2005 and reference therein). However, in crops, such as rice, wheat, and maize, or even in legumes, such as chickpea, peanut, and cowpea with large number of accessions stored in the genebank, even a core collection could be unmanageably large so a further reduction is warranted provided it is not associated with losing too much of the spectrum of diversity. Thus, Upadhyaya and Ortiz (2001) developed a two-stage strategy for developing a mini-core collection, again based on selecting 10% of the accessions from the core collection representing 90% of the variability of the entire

Table IV
Description of Core Collection in Banana, Barley, Cassava, Cowpea, Finger Millet, Maize, Pearl Millet, Potato, Rice, Sorghum, Sweet potato, West African Yam, and Wheat

Crop	Description	No. of accessions	References
Banana	West African plantain core collection	25	Swennen and Vuylsteke, 1987
Barley	East Asian barley core collection	380	Liu *et al.*, 1999
	European barley core collection	79	Liu *et al.*, 2000a
	USDA-ARS barley core collection	2303	Bowman *et al.*, 2001
	American barley core collection	151	Liu *et al.*, 2001a
	Core collection	670	Fu *et al.*, 2005
Caribbean maize	Core collection	100	Taba *et al.*, 1998
Cassava	Core collection	630	Chavarriaga-Aguirre *et al.*, 1999
Cowpea	Core collection	2062	Mahalakshmi *et al.*, 2007a
Finger millet	Core collection	622	Upadhyaya *et al.*, 2006b
Maize	Chinese maize core collection	1193	Li *et al.*, 2004b
Pearl millet	Core collection	1600	http://icrtest:8080/Pearlmillet/Pearlmillet/coreMillet.html
Potato	Core collection	306	Huamán *et al.*, 2000
Rice	USDA core collection	1801	Yan *et al.*, 2004b
	IRRI core collection	11,200	Mackill and McNally, 2004
Sorghum	Core collection	3475	Rao and Rao, 1995
	Core collection	210	Deu *et al.*, 2006
Sweet potato	Core collection	85	Huamán *et al.*, 1999
Uruguayan maize	Core collection	720	Malosetti and Abadie, 2001
West African yam	Core collection	391	Mahalakshmi *et al.*, 2007b
Wheat	Novi Sad Core collection	710	Kobiljski *et al.*, 2002
	Chinese common wheat core collection	340	Dong *et al.*, 2003

collection. In this process, first a representative core collection is developed using all the available information on geographic origin, characterization, and evaluation data. In the second stage, the core collection is evaluated for various morphological, agronomic, and quality traits to select a subset of 10% accessions from this core subset (or 1% of the entire collection) that captures a large proportion (i.e., more than 80% of the entire collection) of the useful variation. At both stages in selection of core and mini-core collections, standard clustering procedures are used to separate groups of similar accessions combined with various statistical tests to identify the best representatives. Mini-core collections are reported for crops such as chickpea (Upadhyaya and Ortiz, 2001), peanut (Upadhyaya *et al.*, 2002), pigeon pea (Upadhyaya *et al.*, 2006c), and rice (1536 accessions, D. J. Mackill, IRRI, personal communication). Evaluation of core and mini-core collections has been suggested as the most efficient and reliable

means of carrying out an initial search of germplasm collections for desirable traits. Such efforts have led to the identification of diverse germplasm with beneficial traits in barley (Bowman *et al.*, 2001), quinoa (Ortiz *et al.*, 1999), and many legume crops of significant economic values (see Dwivedi *et al.*, 2005 and references therein; Brick *et al.*, 2006). It is appropriate to emphasize that the core or mini-core collections do not replace the need for evaluating large parts of the entire collection but simply offer a means of stratifying the process into more manageable batch sizes that can be evaluated more effectively. There is no doubt that this approach may still miss some useful alleles that are present at a very low frequency. In this case, for well-studied traits it may be possible to use genomics technologies to pursue allele mining and gene discovery approaches (Latha *et al.*, 2004; Maccaferri *et al.*, 2005).

The genomic revolution, including dramatic advances in molecular biology, bioinformatics, and information technology, provides the scientific community with tremendous opportunities for improving the pace and scale of plant breeding progress and thereby helping to solve some of the world's most serious agricultural and food security issues. For example, molecular markers can be used for (1) differentiating cultivars and constructing heterotic groups; (2) identifying germplasm redundancy, underrepresented alleles, and genetic gaps in current collections; (3) monitoring genetic shifts that occur during germplasm storage, regeneration, domestication, and breeding; (4) screening germplasm for novel genes or superior alleles; and (5) constructing a representative subset or core collection (Xu *et al.*, 2003). This realization led to the formation of the Generation Challenged Program (GCP) (www.generationcp.org). The GCP aims to utilize molecular tools and comparative biology to explore and exploit genetic diversity housed in existing germplasm collections, with a particular focus on improving the drought tolerance of various cereals, legumes, and clonal food crops. A primary goal of the GCP is extensive genomic characterization of global crop-related genetic resources (composite collections), initially using simple sequence repeat (SSR) markers to determine population structure and now moving onto whole-genome scans [including single nucleotide polymorphism (SNP) arrays and diversity arrays technology (DArT)] and functional genomic analysis of subsets of germplasm (mini-composite collections). Thus, the GCP has created composite collections to cover global diversity for most of the 20 CGIAR-mandated crops. These consists of 3000 accessions or no more than 10% of the total number of available accessions for inbreeding crops and 1500 accessions for outbreeding species (where each accession must be treated as a population). It is expected that this analysis will also lead to the development of genetically broad-based mapping and breeding populations. The results from these GCP-supported projects are already starting to flow for the benefit of the scientific community. For example, a global composite collection of 3000 accessions has been developed in chickpea (Upadhyaya *et al.*, 2006a), its genetic structure defined using

50 polymorphic microsatellites, and a reference collection of 300 accessions identified (ICRISAT/ICARDA unpublished). Further, GCP is supporting a project on allele diversity at orthologous candidate (ADOC) genes that will produce and deliver a public dataset of allelic diversity at orthologous candidate genes across eight important GCP crops and assess whole sequence polymorphism in a DNA bank of 300 reference accessions for each crop. This reference germplasm, which has already undergone genome scan, will be evaluated for traits associated with drought tolerance to test for association between observed polymorphism and trait variability (http://www.intl-pag.org/14/abstracts/PAG14_W264.html). The mini-composite collections and the associated marker technologies developed under GCP will be freely available to all those interested in using these genetic and genomic resources.

Eshed and Zamir (1994) proposed to exploit introgression lines (ILs), also known as chromosome segment substitution lines (CSSLs) or contig lines (CLs), which could be generated by systematic backcrossing and introgression of marker-defined exotic segments in elite genetic background. ILs have a high percentage of the recurrent parent genome and a low percentage of the donor parent genome. ILs offer several advantages over conventional populations: first, they provide useful stocks for highly efficient QTL or gene identification and fine-mapping of these; second, they can contribute to the detection of epistatic interactions between QTL; and third, they can be used to map new region-specific DNA markers (Eshed and Zamir, 1995; Fridman *et al.*, 2004). Several sets of ILs are now available in barley, maize, rice, soybean, and wheat (Table V) that contain beneficial alleles from wild relatives, thus enriching the genetic diversity in primary gene pools of these crops. These ILs when crossed produce progenies with enhanced trait values as demonstrated for increased yield in tomato and wheat (Gur and Zamir, 2004; Shubing *et al.*, 2006). Other useful genetic resources being developed in many crops include recombinant inbred lines (RILs) (Burr *et al.*, 1988), advanced backcross lines (Tanksley and Nelson, 1996), near isogenic lines (NILs) (Muehlbauer *et al.*, 1988), and double-haploid lines (DHL) (Kasha and Kao, 1970) that can be used to identify genes underlying traits by marker-phenotype correlations, dissecting the genetic structure of the complex traits, and for enhancing the trait performance.

In addition to naturally available and conventionally bred genetic resources preserved in genebanks, researchers are also creating new genetic variation by using novel technique such as Targeting Induced Local Lesions IN Genome (TILLING), which is a powerful reverse genetics technique that employs a mismatch-specific endonuclease to detect single base pair (bp) allelic variation in a target gene using high-throughput assay. Its advantages over other reverse genetic techniques include its applicability to virtually any organism, its facility for high throughput, and its independence of genome size, reproductive system, or generation time (Gilchrist and Haughn, 2005).

Table V
ILs (also known as Chromosome Substitution Lines, CSSLs) in Barley, Maize, Rice, Soybean, and Wheat

Description of genetic resources	References
Barley (*H. vulgare*)	
146 recombinant chromosome substitution lines, derived from BC_2F_6 of the cross Harrington and Caesarea (*H. vulgare* ssp. *spontaneum*), covering average *H. spontaneum* genome of 12.5%	Matus *et al.*, 2003
Two sets of ILs, containing 49 and 43 ILs, derived from BC_2DH populations of *H. vulgare* ssp. *spontaneum* (ISR42-8) crossed with German spring barley cultivar Scarlett and Thuringia, covering at least 98.1% and 93.0% of the exotic genome in overlapping introgressions and containing on average 1.5–2.0% additional nontarget introgressions	von Korff *et al.*, 2004
Maize (*Zea mays*)	
Maize chromosome disomic ($2n = 6x + 2 = 44$) addition lines for chromosomes 1–4, 6, 7, and 9 and monosomic ($2n = 6x + 1 = 43$) addition line for chromosome 8; and for monosomic ($n = 3x + 1 = 23$) addition lines for maize chromosome 5 and 10 to a haploid complement of oat isolated from oat × maize cross	Kynast *et al.*, 2001
Rice (*O. sativa*)	
147 ILs from *O. sativa* (Taichung 65) and *O. glumaepatula* reciprocal crosses containing *O. glumaepatula* or Taichung 65 cytoplasm but with entire chromosome segments of *O. glumaepatula* developed	Sobrizal *et al.*, 1999
140 near isogenic ILs derived from a cross between *japonica* cultivar Nipponbare, and an elite *indica* line Zhenshan 97B	Mu *et al.*, 2004
75 CSSLs, representing on average 97.6% background genome, carrying overlapping chromosome segments of Pai6S in a genetic background of elite cultivar 9311	Xiao *et al.*, 2005
20,000 ILs in three elite genetic backgrounds (IR64, Teqing, and IR68552-55-3-2) containing a significant portion of loci affecting complex phenotypes at which allelic diversity exists in the primary gene pool of rice	Li *et al.*, 2005a
25 monosomic alien addition lines (MAALs) containing the complete genome of *O. sativa* and individual chromosomes of *O. officinalis*	Tan *et al.*, 2005
159 ILs carrying variant introgressed segments from *O. rufipogon* Griff. in the background of *indica* cultivar, Guichao representing 67.5% of the *O. rufipogon* genome and recurrent parent genome ranging from 92.4% to 99.9%, with an average of 97.4%. The average proportion of donor genome was about 2.2%	Tian *et al.*, 2006b
Soybean (*G. max*)	
22 monosomic addition lines, containing an extra chromosome from *G. tomentella* to the 2*n* soybean complement, possess several modified plant characteristics such as flowering habit, plant height, degree of pubescence, seed fertility, number of seeds per pod and plant, pod and seed color, and seed yield	Singh *et al.*, 1998
Wheat (*T. aestivum*)	
36 homozygous lines carrying different segments of individual chromosomes of *Aegilops tauschii* genome	Pestsova *et al.*, 2001
84 ILs containing a single homozygous introgression from *A. tauschii* genome in "Chinese Spring" background	Pestsova *et al.*, 2006

As TILLING provides mutation in the target gene, it offers much greater prevision than previous random mutation techniques (using chemical or radioactive mutagens), and it has been successfully used for the detection of both induced and natural variation in several plant and animal species (Perry *et al.*, 2003; Smits *et al.*, 2004; Stemple, 2004; Till *et al.*, 2003, 2004; Wienholds *et al.*, 2003). For example, Slade *et al.* (2005) generated 246 alleles in the granule-bound starch synthase 1 (*GBSS1*) gene (*waxy*) in wheat using TILLING. Reduction or loss of *GBSS1* function results in starch with a decreased or absent amylase fraction, desired for its improved freeze-thaw stability and resistance to staling compared to conventional starch. Similarly in maize, Till *et al.* (2003, 2004) screened pools of DNA samples for mutations in 1-kb segments from 11 different genes, obtained 17 independently induced mutations from a population of 750 pollen-mutagenized maize plants, and established the public TILLING service for maize modeled on *Arabidopsis* TILLING project (Till *et al.*, 2003) at Purdue University (http://genome.purdue.edu/maizetilling). More recently, an EcoTILLING facility has been established at IRRI to identify putative SNPs in both cultivated and wild rice germplasm. EcoTILLING a set of 900 of the *Oryza sativa* lines for 1800 bp of coding and regulatory region of *ERF3* (a candidate gene associated with drought tolerance) identified 31 SNP and short indels that grouped into 9 haplotypes corresponding to the cultivar types (McNally *et al.*, 2006).

Powdery mildew is the devastating disease of barley. The genes *mlo* and *Mla* are involved in the host plant resistance of barley against the fungal pathogen causing powdery mildew. *Mla* has multiple alleles at its locus, while *mlo* is a single copy gene. Using EcoTILLING approach, Mejlhede *et al.* (2006) not only detected point mutations and deletions in each of the 11 *mlo* mutants tested but also identified most of the *Mla* alleles from 25 natural variants of *Hordeum vulgare* ssp. *spontaneum*, although the identification was complex due to the presence of highly similar paralogues of *Mla*.

Among the legumes, TILLING is being used to develop soybeans with better seeds (improved oil and protein content and allergen-free soybeans) (http://www.ars.usda.gov/is/pr/2005/050705.htm). TILLING has great potential to detect both induced and natural polymorphic variation, and as more DNA markers become available and the technological innovations advanced thus reducing the cost of high-throughput analysis, this technique has great potential for application in crop improvement. These structured mutant populations are also a valuable resource for forward genetic screens.

Natural biodiversity is an underexploited sustainable resource that can enrich the genetic basis of cultivated plants with novel alleles and genes to improve yield potential and stability adaptation and resilience. Wild relatives possess a high level of resistance to many biotic and abiotic stresses but are agronomically inferior to modern cultivars (albeit sometimes harboring masked genes of beneficial value for these traits). Tools developed for genetic

dissection of traits in cultivated germplasm can also be used to identify and assist the transfer of useful genes from wild relatives (Tanksley and Nelson, 1996) that has been effectively used for improving both yield and/or seed quality in barley, chickpea, common bean, oat, peanut, pearl millet, pigeon pea, rice, sorghum, soybean, and wheat (Dwivedi *et al.*, 2007).

For many crops, the level of genetic diversity in the primary gene pool is narrow. Expanding the genetic base of these crops is, therefore, important for continued crop improvement. Rapid developments in molecular genetic technologies have opened up the vast majority of plant genomes to investigation that in turn will enable the release of genetic variation not previously accessible through conventional crossing and selection.

B. Genomic Resources

1. Genetic Markers

Genetic markers were originally used in genetic mapping to determine the order of the genes along chromosomes, and evolved from morphological markers through isozyme markers to DNA markers which themselves evolved from hybridization-based detection to polymerase chain reaction (PCR) amplification and now to new sequence-based systems. Both morphological and isozyme markers are limited in number. Additionally, the morphological markers are affected by the environment, and a given marker can affect other morphological traits because of pleiotropic gene action. Consequently, genome-wide analysis is not feasible using both morphological and isozyme markers. DNA markers are typically derived from a small region of DNA that shows sequence polymorphism between individuals within a species, and may be classified into random DNA markers (RDM) (also known as anonymous or neutral markers), gene-targeted markers (GTM) (also known as candidate gene marker), and functional markers (FM) (Andersen and Lübberstedt, 2003). RDM are derived at random from polymorphic sites across the genome, whereas GTM are derived from polymorphisms within the gene. FM are derived from polymorphic sites within genes causally associated with phenotypic trait variation and are superior to RDM owing to complete linkage with trait locus alleles (Andersen and Lübberstedt, 2003). The major draw back of the RDM is that their predictive value depends on the known linkage phase between marker and target locus alleles (Lübberstedt *et al.*, 1998). In contrast, once genetic effects have been assigned to functional sequence motifs, FM derived from such motifs can be used for fixation of gene alleles in a number of genetic backgrounds without additional calibration. FM are superior to GTM and RDM owing to their association with genes of known function.

a. Random DNA Markers. Restriction fragment length polymorphisms (RFLPs) were the first DNA markers to be developed that have been widely and successfully used to construct linkage maps and detect QTL in many crop species. However, with the discovery of the polymerase chain reaction (PCR) (Saiki *et al.*, 1988), attention shifted to developing a wide range of PCR-based assays including random amplified polymorphic DNA (RAPD), amplified fragment length polymorphisms (AFLPs), and SSR (also known as microsatellites). RFLP, although providing high-quality codominant information, is labor intensive, time consuming, requires large amount of DNA, and is dependent on radioisotope-based protocols. While RAPD and AFLP only provide dominant information; the former suffers from reproducibility problems. However, it is possible to convert tightly linked RFLP markers into PCR-based sequence-tagged site (STS) markers (Olson *et al.*, 1989) and both RAPD and AFLP bands can be converted into sequence-characterized amplified region (SCAR) markers (Paran and Michelmore, 1993) or cleaved amplified polymorphic sequences (CAPs) markers (Konieczny and Ausubel, 1993). Microsatellite markers are ideal DNA markers for genetic mapping and population studies because of their abundance, high level of polymorphism, multiallelic nature, codominant inheritance and wide dispersion in genomes, ease of assay using PCR, and ease of dissemination among laboratories (Powell *et al.*, 1996). Barley has the largest collection of SSR markers followed by rice, wheat, maize, and sorghum (Table VI). Soybean, chickpea, pea, and peanut also have large well-assembled collections of SSR (Dwivedi *et al.*, 2005; Moretzsohn *et al.*, 2005; Sethy *et al.*, 2006). Other legume crops, such as cowpea and common bean, which are also globally important, are lagging behind in terms of SSR development, as is the case for *Musa* and many other clonal crops (Table VI).

DArT is microarray-based technique that detects genetic polymorphism, which can be used to construct medium-density genetic linkage maps in species with various genome sizes (Jaccoud *et al.*, 2001). DArT markers are biallelic and behave in a dominant (present vs absent) or codominant (two doses vs one dose vs absent) manner. DArT is a good alternative to currently used techniques (such as RFLP, AFLP, SSR, and SNP), in terms of cost and speed of marker discovery and analysis, for whole-genome fingerprinting. It is cost-effective, sequence-independent, nongel-based technology that is amenable to high-throughput automation, discover hundreds of high-quality markers in a single assay, and integration of DArT markers in genetic map is straightforward. An open source software package, DArTsoft, is available for automatic data extraction and analysis. DArT technology has been successfully developed for barley, cassava, rice, and wheat, while work is in progress to establish DArT in chickpea, pigeon pea, and sorghum (http://www.diversityarrays.com/pub/huttneretal2005.pdf). For example, a genetic map with 385 unique DArT markers spanning 1137-cM barley genome (Wenzl *et al.*, 2004) constructed, DArT markers with AFLP and SSR markers mapped on wheat genome

Table VI
SSR Markers Reported in Banana, Barley, Cassava, Maize, Oat, Pearl Millet, Potato, Rice, Sorghum, Sweet potato, Wheat, and Yam

Summary of the marker information reported	References
Banana	
24 SSRs from *M. acuminata* ssp. *malaccensis*	Crouch *et al.*, 1998a
44 B-genome-specific SSRs from enriched library of *M. balbisiana* cultivar Tani	Buhariwalla *et al.*, 2005a
9 B-genome-derived SSRs	Oriero *et al.*, 2006
Barley	
45 SSRs from genomic DNA library and from public databases	Liu *et al.*, 1996
568 SSRs from database sequences and small-insert genomic libraries	Ramsay *et al.*, 2000
1856 SSRs from 24,595 ESTs	Thiel *et al.*, 2003
127 SSRs from genomic DNA of barley cultivar Franka	Li *et al.*, 2003b
3530 SSRs from 170,746 ESTs	Nicot *et al.*, 2004
Cassava	
14 SSRs containing GA-repeats from cassava genome	Chavarriaga-Aguirre *et al.*, 1998
9 SSRs from genomic library of *Ipomoea batatas*	Buteler *et al.*, 1999
172 SSRs from 692 putative DNA clones from cassava	Mba *et al.*, 2001
Maize	
6 SSRs from sequences	Senior and Heun, 1993
200 SSRs from maize sequences	Chin *et al.*, 1996
655 indels from 8 maize inbreds	Bhattramakki *et al.*, 2002
1051 SSRs from maize microsatellite-enriched libraries and microsatellite-containing sequences from public and private databases	Sharopova *et al.*, 2002
200 SSRs from maize sequences	http://www.maizegdb.org/ssr.php
Oat	
34 SSRs from three oat microsatellite-enriched libraries	Li *et al.*, 2000
Pearl millet	
50 SSRs from pearl millet BAC clones	Qi *et al.*, 2001
18 SSRs from small-insert genomic library	Budak *et al.*, 2003
44 SSRs from a (CA)n-enriched small-insert library	Qi *et al.*, 2004
Potato	
42 SSRs from potato genomic libraries and SSR-containing sequences in the public databases	Ashkenazi *et al.*, 2001
Rice	
2414 SSRs representing 2240 unique marker loci, with majority from regions flanking perfect repeats $\geq$24 bp, corresponding to (GA) (36%), (AT) (15%), and (CCG) (8%) motifs. These SSRs along with previously mapped 500 SSRs total 2740 SSRs, 1 SSR every 157 kb	McCouch *et al.*, 2002

(*continued*)

Table VI (*continued*)

Summary of the marker information reported	References
Sorghum	
47 SSRs from sorghum genomic libraries and 2 SSRs from GenBank SSR-containing sequences	Brown *et al.*, 1996
10 SSRs from sorghum genomic libraries and 3 SSRs from database searches	Taramino *et al.*, 1997
313 SSRs from sorghum BAC and genomic-DNA libraries	Bhattramakki *et al.*, 2000
38 SSRs from sorghum genomic DNA libraries	Kong *et al.*, 2000
Sweet potato	
5 SSRs from size-fractionated genomic libraries	Jarret and Bowen, 1994
112 SSRs from EMBL database, cDNA, and selectively enriched small-insert DNA libraries	Milbourne *et al.*, 1998
102 SSRs from small-insert genomic library, microsatellite-enriched library, and mining EST-databases	Hu *et al.*, 2004a
15 SSRs from *Ipomoea trifida* sequences, closely related to sweet potato	Hu *et al.*, 2004b
Wheat	
230 SSRs from A, B, and D genomes	Röder *et al.*, 1998
22 EST-SSRs and 20 genomic-derived SSRs	Eujayl *et al.*, 2002
897 EST-derived SSRs	Gupta *et al.*, 2003
540 SSRs from A, B, and D genomes in addition to 570 previously reported SSRs	Song *et al.*, 2005
Yam	
20 SSRs identified from Gnidou parent	Mignouna *et al.*, 2003b; Scarcelli *et al.*, 2005

(Semagn *et al.*, 2006), and a cassava DArT genotyping array containing ~1000 polymorphic clones (Xia *et al.*, 2005) are available and display a high level of polymorphism that shows the genetic relationships among the samples consistent with the information available on them.

b. Gene-Targeted Markers. Expressed sequence tags (ESTs) are currently the most widely sequenced nucleotide element from the plant genomes with respect to the number of sequences and the total number of nucleotides available to researchers. EST provides a robust sequence resource that can be exploited for gene discovery, genome annotation, and comparative genomics. ESTs are typically unedited, automatically processed, single-read sequences produced from cDNA. Over 38 million sequences have been deposited in the publicly available plant EST sequence databases (dbEST-release 090806; http://www.ncbi.nlm.nih.gov/dbEST_summary.html). Many of these EST have been sequenced as an alternative to complete genome sequencing or as a substrate for cDNA array-based expression analysis.

Bioinformatics-based sequence analysis tools have extended the scope of EST analysis into the field of proteomics, marker development, and genome annotation. Although ESTs are no substitute for a whole-genome scaffold, this ''poor man's genome'' resource forms the core foundations for various genome-scale experiments for less well-funded crops or species with very large genomes (Rudd, 2003). EST constitutes a novel source of markers that are physically associated with coding regions of the genome. Moreover, ESTs are also a source of SSR in many crops. Kumpatla and Mukhopadhyay (2005) used this approach to examine the abundance of SSR in more than 1.54 million EST belonging to 55 dicotyledonous species. The frequency of EST-containing SSR among species ranged from 2.65% to 16.82%, with dinucleotide repeats most abundant followed by tri- or mononucleotide repeats, thus demonstrating the potential of *in silico* mining of EST for rapid development of SSR markers for genetic analysis and application in dicotyledonous crops. However, EST-SSR (also known as genic SSR) produce high-quality markers, but these are often less polymorphic than genomic SSR (Cho *et al.*, 2000; Eujayl *et al.*, 2002; Thiel *et al.*, 2003). SSR markers may also be transferable to related species and are useful for assaying the functional diversity in natural populations or germplasm collections and also as anchor markers for comparative mapping and evolutionary studies (Varshney *et al.*, 2005b). Tang *et al.* (2006) identified 428 UNI-SSR-EST from wheat genome homologous in rice, maize, and barley. They designed 243 SSR primers and when tested in each species 154 primers produced clear amplicons across the four species, demonstrating a high efficient transferability of wheat EST-SSR markers to the other cereal crops. Similarly, Choi *et al.* (2006) used 274 unigene sequences to develop PCR-based genetic markers across 15 legume genomes, representing 6 crops or model legume species from the phaseoloid and inverted repeat loss clades. They found 129 of these unigene sequence-amplified fragments representing single-copy loci across most target diploid genomes that 70.5% of these markers are intron spanning and 85.3% linked to legume genetic maps. EST resources are also being used to mine SNP (Kota *et al.*, 2003; Picoult-Newberg *et al.*, 1999). EST provides a quantitative method to measure specific transcripts within a cDNA library and represents a powerful tool for gene discovery, gene expression, gene mapping, and the generation of gene profiles. The National Center for Biotechnology Information (NCBI) database, dbEST 090806 (http://www.ncbi.nlm.nih.gov/dbEST_summary.html), contains the largest collection of EST in rice, wheat, barley, maize, soybean, sorghum, and potato (also see Table VII). Development of EST in cassava is catching up, while only a few hundred ESTs are reported in *Musa* and other clonal crops (Table VII) and legumes (except for soybean) (Dwivedi *et al.*, 2005; also see Table VII). Clearly, there is an urgent need to develop SSR in the legumes and clonal crops.

Table VII
Expressed Sequenced Tags (ESTs) Reported in Banana, Barley, Cassava, Chickpea, Common Bean, Maize, Oat, Potato, Rice, Sorghum, Soybean, Sweet Potato, and Wheat

Summary of the ESTs reported	References
Banana	
2286 ESTs from the leaves of *M. acuminata* ssp. *burmannicoides* variety Calcutta 4	Santos *et al.*, 2005
Barley	
13,109 ESTs from 3 cDNA libraries of barley cultivar, Barke, resulting 4,000 genes	Michalek *et al.*, 2002
271,630 ESTs from 23 barley varieties cDNA libraries resulting 56,302 tentative consensus sequences	Kota *et al.*, 2003
110,981 ESTs from 22 cDNA libraries resulting 25,224 unique sequences	Zhang *et al.*, 2004
437,321 ESTs reported in dbEST release 090806	http://www.ncbi.nlm.nih.gov/dbEST_summary.html
Cassava	
4000 ESTs from cassava mosaic disease resistant genotype	Fregene *et al.*, 2004
23,000 ESTs from various cassava tissues and genotypes identified 6000–7000 unigenes	Anderson *et al.*, 2004
5700 unigenes from ESTs of root tissues of cassava varieties with high and low starch contents and those challenged by cassava BB (*Xanthomonas axonopodis* pv. *manihotis*)	Lopez *et al.*, 2004
17,954 ESTs reported in dbEST release 090806	http://www.ncbi.nlm.nih.gov/dbEST_summary.html
Chickpea	
477 ESTs from root tissue of two closely related genotypes resulted 106 EST-based markers	Buhariwalla *et al.*, 2005b
Common bean	
5255 ESTs from 3 cDNA libraries resulting into 3126 unigenes	Melotto *et al.*, 2005
Maize	
73,000 ESTs from multiple organs and developmental stages resulting 22,000 tentative unique genes	Fernandes *et al.*, 2002
1,143,737 ESTs reported in dbEST release 090806	http://www.ncbi.nlm.nih.gov/dbEST_summary.html
Oat	
9792 EST from oat cDNA library detected 2800 cold-induced UniGene sets	Bräutigam *et al.*, 2005
7632 ESTs reported in dbEST release 090806	http://www.ncbi.nlm.nih.gov/dbEST_summary.html
Potato	
61,949 ESTs from aerial tissues, below ground tissues, and tissues challenged with late blight (*Phytophthora infestans*) identified 19,892 unique sequences	Ronning *et al.*, 2003
219,917 ESTs reported in dbEST release 090806	http://www.ncbi.nlm.nih.gov/dbEST_summary.html

Table VII (*continued*)

Summary of the ESTs reported	References
Rice	
1,188,881 ESTs reported in dbEST release 090806	http://www.ncbi.nlm.nih.gov/dbEST_summary.html
Sorghum	
204,208 ESTs reported in dbEST release 090806	http://www.ncbi.nlm.nih.gov/dbEST_summary.html
Soybean	
27,513 unigenes obtained from a variety of soybean cDNA libraries made from a wide array of source tissues and organ systems, developmental stages, and stress or pathogen-challenged plants	Vodkin *et al.*, 2004
2003 ESTs from full-length cDNA library of wild soybean (50,109) leaf treated with 150-mM NaCl	Ji *et al.*, 2006
359,158 ESTs reported in dbEST release 090806	http://www.ncbi.nlm.nih.gov/dbEST_summary.html
Sweet potato	
7841 ESTs reported in dbEST release 090806	http://www.ncbi.nlm.nih.gov/dbEST_summary.html
Wheat	
855,066 ESTs reported in dbEST release 090806	http://www.ncbi.nlm.nih.gov/dbEST_summary.html

Target region amplification polymorphisms (TRAP) are derived from a rapid and efficient PCR-based technique, which uses bioinformatic tools and EST database information to generate polymorphic markers around targeted candidate gene sequences (Hu and Vick, 2003). This TRAP technique uses two primers of 18 nucleotides to generate markers. TRAP are amplified by one fixed primer designed from a target EST sequence in the database and a second primer of arbitrary sequence except for AT- or GC-rich cores that anneal with introns and exons, respectively. The TRAP technique should be useful in genotyping germplasm collection and in tagging genes with beneficial traits in crop plants. TRAP markers are reported in mapping QTL in wheat (Liu *et al.*, 2005), mapping disease resistance genes in common bean (Miklas *et al.*, 2006b), and for nutritional quality of straw or tolerance to salinity and terminal drought in pearl millet (Mukhopadhyay, Senthilvel, and Hash, ICRISAT, personal communication).

SNPs are the most abundant sequence variations encountered in most genomes (Cho *et al.*, 1999; Picoult-Newberg *et al.*, 1999). Their development costs are similar to those of SSR, but genotyping platforms are now available with very high-throughput potential and very low unit cost (Kanazin *et al.*, 2002). SNPs are especially useful for association studies because of their high

frequency across the genome and because they are genetically more stable than SSR. Thus, SNPs are ideally suited for the generation of high-density genetic maps (Cho *et al.*, 1999). However, currently there are only a few crops with large SNP marker resources; rice, maize, barley, and oat having the largest collection of SNPs (Table VIII). There are also a few hundred SNPs in soybean and common bean, and very few in peanut (Dwivedi *et al.*, 2006). For outbreeding crops, such as maize, polymorphic markers are highly abundant—1 SNP per 60.8 bp (Ching *et al.*, 2002) as compared to inbreeding species such as rice—3.0 SNP per kb in coding regions to 27.6 SNP per kb in transposable elements (Yu *et al.*, 2005)—or barley—1 SNP per 200 bp (Rostoks *et al.*, 2005). More research is needed to fully develop the potential of this class of marker, but this will surely rapidly occur due to the cost efficiencies gained during large-scale genotyping with SNPs.

c. Functional Markers. FM are derived from polymorphic sites within the genes known to be causally involved in phenotypic trait variation. The development of FM requires allele-specific sequences of functionally characterized genes from which polymorphic, functional motifs affecting plant phenotype can be identified.

Table VIII
Single Nucleotide Polymorphisms (SNP) Marker Reported in Barley, Cassava, Common Bean, Maize, Oat, Rice, and Wheat

Summary of the SNPs and indels reported	References
Barley	
3069 intervarietal and 3377 intravarietal SNP	Kota *et al.*, 2003
Cassava	
80 intercultivar and 146 intracultivar SNP	Lopez *et al.*, 2005
Common bean	
318 SNP and 68 indel	Melotto *et al.*, 2005
Maize	
169 SNP and indel from 36 maize inbreds	Ching *et al.*, 2002
14,832 SNP from 102,551 maize EST	Batley *et al.*, 2003
Oat	
>2000 genome-wide SNP	Rostoks *et al.*, 2005
Two SNP, SNP-REMAP and SNP-RAPD, linked with dwarfing gene, *Dw6*	Tanhuanpää *et al.*, 2006
Rice	
2800 SNP from 3 *Oryza* ssp. (*japonica*, *indica*, and wild rice)	Nasu *et al.*, 2002
384,431 SNP and 24,557 indels from two subspecies	Feltus *et al.*, 2004
Wheat	
20 SNP from 12 wheat genotypes	Somers *et al.*, 2003
40 SNP from 10 wheat cultivars	Ablett *et al.*, 2006

Dwarf8 in maize encodes a gibberellin response modulator from which FM can be developed for plant height and flowering time. For example, nine sequence motifs in the *Dwarf8* gene of maize were shown to be associated with variation in flowering time, and one particular 6-bp deletion accounted for 7–11 days difference in flowering time between inbreds (Thornsberry *et al.*, 2001). However, *Dwarf8* is a pleiotropic gene (also affecting plant height) and thus needs to identify FM from "additional flowering time genes" in addition to using *Dwarf8*-derived FM. Orthologues to *Dwarf8* have been identified in wheat (*Rht1*) (Peng *et al.*, 1999), rice (*SLR1*) (Ikeda *et al.*, 2001a), and barley (*sln1*) (Chandler *et al.*, 2002), and we know that such genes were bred into the high-yielding wheat and rice cultivars of the Green Revolution (Hedden, 2003). Altered function of alleles in these orthologous genes can reduce the response to gibberellin and consequently lead to decreased plant height. Thus, biallelic (gibberellin sensitive and insensitive) FM can be derived for targeted and rapid cultivar breeding aiming at increased lodging tolerance. Brown midrib (*bm*) mutants in maize have an increased digestibility but inferior agronomic performance (Barriere and Argillier, 1993). Two of the four *bm* genes (*bm1* and *bm3*) are involved in monolignol biosynthesis (Barrière *et al.*, 2003). These two genes and additional lignin biosynthesis genes have been isolated based on sequence homology. Candidate genes putatively affecting forage quality have been identified by expression profiling using isogenic *bm* lines, and detected association between a polymorphism at the caffeic acid *O*-methyltransferase (COMT) locus and digestible neutral detergent fiber (DNDF) in a collection of maize inbred lines (Lübberstedt *et al.*, 2005). Silage maize is a major source of forage for dairy cattle due to its high-energy content and good digestibility. Lignin structure and cross-linking between cell wall components influence digestibility (Barrière *et al.*, 2003). Analysis of allelic diversity in relation to cell wall digestibility revealed *ZmPox3* peroxidase, a candidate gene for silage maize digestibility improvement (Guillet-Claude *et al.*, 2004), as it is colocalized with a cell wall digestibility and lignification QTL (Barrière *et al.*, 2003). GBSS, starch branching enzymes 1 (SBE1) and 3 (SBE3), are major enzymes involved in starch biosynthesis in rice endosperm. Using variation in sequence diversity at *Sbe1* and *Sbe3* loci and *Wx* gene markers, Liu *et al.* (2004c) differentiated an *indica* allele from a *japonica* allele for both *Sbe1* and *Sbe3* loci. The same research team also showed that *Wx* and *Sbe3* loci had significant effects on the amylose content (AC) variation, and together account for 79% of the observed AC variation in a double-haploid population. The flavor and fragrance of Basmati and Jasmine rice is associated with increased levels of 2-acetyl-1-pyrroline (2AP) (Yoshihashi, 2002). Although various methods are employed to select for fragrant rice, such methods are difficult, labor intensive, time consuming, require more sampling, and are often unreliable (Reinke *et al.*, 1991). Fragrance in rice is a recessive trait and a deletion in

the gene encoding BAD2 on chromosome 8 that disables the BAD2 enzyme is the most likely cause of fragrance (Bradbury *et al.*, 2005). Bradbury *et al.* (2005) used a low-cost robust technique, allele-specific amplification (ASA), which allows discrimination between fragrant and nonfragrant rice cultivars and identifies homozygous fragrant, homozygous nonfragrant, and heterozygous nonfragrant individuals in populations segregating for fragrance. This test detects a 355-bp fragment from a nonfragrant allele and a 257-bp fragment from a fragrant allele, allowing simple analysis on agarose gels. In wheat, two candidate genes control a QTL for high-molecular-weight glutenin subunit (HMW-GS) GluBx: *Glu-B1-1* (structural gene coding for Glu1Bx) and *spa-B* (the B homoeologous gene coding for SPA) located on the 1BL chromosome at a distance of 1.3 cM from each other within the confidence interval of a QTL for the quantity of GluBx (Guillaumie *et al.*, 2004). In the absence of linkage disequilibrium (LD) between *Glu-B1-1* and *spa-B*, Ravel *et al.* (2006) conducted an association mapping (AM) study to identify the individual gene responsible for the QTL, and detected significant associations only between *Glu-B1-1* polymorphism and most of the traits (protein content, the quantity of HMW-GS, and protein fractions for each HMW-GS) evaluated. Malt from barley grains is the raw material for the production of beer. Genetic improvement of malting quality is impaired by the quantitative inheritance and the comparatively low heritability. By monitoring mRNA profiles during grain germination, Potokina *et al.* (2004) identified between 17 and 30 candidate genes for each of the 6 malting parameters, and 5 of the 8 mapped candidate genes display linkage to known QTL for malting-quality traits. Genes determining growth habit are well known in different species and all are recognized as *CEN/TFL1* homologous or *CEN/TFL1*-like genes (Avila *et al.*, 2006 and references therein). Avila *et al.* (2006) designed primers for conserved domains from sequences of *TFL1/CEN*-like genes and used Hind1II enzyme to produce a clear polymorphism between determinate and indeterminate genotypes in faba bean. This cleaved amplified polymorphism (CAP) marker showed 100% efficiency in discriminating determinate and nondeterminate individuals in an F_2 population segregating for growth habit. These examples demonstrate that gene-based markers are more robust than anonymous markers linked to the trait loci of interest.

2. Genome Sequencing

Plant genome sizes vary from the modest—54 million base pairs (Mb) in the bitter cress (*Cardamine amara*)—to the enormous—124,000 Mb in the lily *Fritillaria assyriaca*. Among the most important food crops, rice has the smallest genome (389 Mb) (IRGSP, 2005) and wheat has the largest genome (15,999 Mb). Other crops could be grouped into seven classes based on the

progressive increase in genome size: *Musa*, cowpea, and yam (555–613 Mb); sorghum, bean, chickpea, and pigeon pea (709–818 Mb); soybean (1115 Mb); potato and sweet potato (1597–1862 Mb); maize, pearl millet, and peanut (2352–2813Mb); pea and barley (4397–5361 Mb); and oat (11,315 Mb) (Arumuganathan and Earle, 1991). Although plant genomes vary substantially in size, the larger genomes do not necessarily have proportionally more genes, but instead the extra genome size is due to repetitive elements that have proliferated in the genomes of plant species across the plant kingdom (Bennetzen, 1998; Bennetzen *et al.*, 1994).

Genome sequencing in most plants is difficult because of the size and complexity of the genomes. Rice is the first cereal to be fully sequenced (Table IX) because of its importance as one of the major cereals in addition to its small genome size, small number of chromosomes ($n = 12$), well-characterized genetic and genomic resources, and availability of a large number of DNA markers and high-density genetic linkage map. The extremely large genome of other crops makes them difficult to sequence. Sequencing hexaploid wheat could yield a considerable amount of important new information about cereals and crop plant biology. The International Wheat Genome Sequencing Consortium (IWGSC) has been formed to advance agricultural research for wheat production and utilization by developing DNA-based tools and resources that result from the complete sequencing of the expressed genome of common (hexaploid) bread wheat and to ensure that these tools and the sequences are available for all to use without restriction and cost (Gill *et al.*, 2004; http://www.wheatgenome.org/). Sorghum is an important bridge to closely related large-genome crops in its own tribe such as maize and sugarcane and thus provides a better road map for study of these crops at the DNA level. Sorghum is also a C4 photosynthesis plant which uses a complex combination of biochemical and morphological specializations that result in more efficient carbon assimilation at high temperature. The genus *Sorghum* also includes one of the world's most noxious weeds, the Johnsongrass (*Sorghum halepense*). The rapid dispersal, high growth rate, and durability that make Johnsongrass such a troublesome weed are actually desirable in many forage, turf, and high-biomass crops that are genetically complex. Therefore, sorghum offers novel learning opportunities relevant to weed biology as well as to improvement of a wide range of forage crops.

The extremely large size of many cereal genomes makes it difficult to decode using the standard methods of genome sequencing such as clone-by-clone (Lander *et al.*, 2001) and whole-genome shotgun (Venter *et al.*, 2001). Determining their complete sequences is daunting and costly. In recent years, two genome filtration strategies, methylation filtration (MF) (Rabinowicz *et al.*, 1999) and C_0t-based cloning and sequencing (CBCS; Peterson *et al.*, 2002) or high C_0t (HC; Yuan *et al.*, 2003), have been

Table IX
Status of Genome Sequencing in Banana, Maize, Rice, and Sorghum

Summary of sequencing information	References
Banana	
Two BAC clones of *M. acuminata* sequenced: MuH9 is 82,723-bp long with an overall G + C content 38.2% and gene density of 1 per 6.9 kb while MuG9 73,268-bp long with an overall G + C content 38.5% and gene density of 1 per 10.5 kb	Aert *et al.*, 2004
Maize	
100,000 maize sequences reported using methylation filtration method of genome sequencing	Palmer *et al.*, 2003
One-eighth of the genome of maize inbred B73 sequenced (307 Mb) that contain large percentage of the genes and transposable elements: repeat sequences 58% and genic regions 7.5%, with ~59,000 predicted genes	Messing *et al.*, 2004
66% of the maize genome consists of repetitive elements; retrotransposons far more frequent than DNA transposons; full-length genes averaged 4 kb; 42,000–56,000 genes predicted	Haberer *et al.*, 2005
Rice	
A draft sequence of *indica* variety 93–11 contains 46,022–55,615 genes. 80% of *A. thaliana* genes had a homologue in rice but only 49.4% of rice genes had a homologue in *A. thaliana*	Yu *et al.*, 2002
A draft sequence of *japonica* variety Nipponbare consists of 32,000–50,000 predicted genes. 98% of the known maize, wheat, and barley proteins are homologues to proteins in rice. Extensive synteny and gene homology between rice and other cereals but limited synteny with *Arabidopsis*	Goff *et al.*, 2002
95% of the 389-Mb sequenced genome detected 37,544 nontransposable-element-related protein-coding genes of which 71% had a putative homologue in *Arabidopsis*. 29% of the 37,544 genes appear in clustered gene families. 2859 genes unique to rice and other cereals, and some might differentiate monocot and dicot lineages	IRGSP, 2005
Of the 38,000–40,000 genes, only 2–3% of these unique to the genomes of *indica* and *japonica* rice; 18 distinct pairs of duplicated segments cover 65.7% of the genome and 17 of these pairs date back to a common time before the divergence of the grasses	Yu *et al.*, 2005
Sorghum	
300 Mb of the 735-Mb of sorghum genome sequenced, tagging 96% of the genes with an average coverage of 65% across their length	Bedell *et al.*, 2005

suggested for selectively sequencing the gene space of large genomes. MF is based on the characteristics of plant genomes in which genes are largely hypomethylated but repeated sequences are highly methylated. Methylated DNA is cleaved when transferred into an *Mcr* + *Escherichia coli* strain and only hypomethylated DNA is recovered. CBCS/HC separates single- and

low-copy sequences, including most genes, from the repeated sequences on the basis of their differential renaturation characteristics. Using the MF strategy, Bedell *et al.* (2005) sequenced 96% of the genes in sorghum with an average coverage of 65% across their length. This strategy filtered away repetitive elements when sequencing the genome of sorghum that reduced the amount of sorghum DNA to be sequenced by two-third, from 735 Mb to ~250 Mb. Both MF and HC have been used for efficient characterization of maize gene space (Palmer *et al.*, 2003; Whitelaw *et al.*, 2003). Using HC and MF, Martienssen *et al.* (2004) generated up to twofold coverage of the gene space with less than 1 million sequencing reads and simulations using sequenced BAC clones predicted that 5× coverage of gene-rich regions, accompanied by less than 1× coverage of subclones from BAC contigs, will generate high-quality mapped sequence that meets the needs of geneticists while accommodating unusually high levels of structural polymorphism. Haberer *et al.* (2005) selected 100 random regions averaging 144 kb in size, representing about 0.6% of the genome, to define their content of genes and repeats for characterizing the structure and architecture of the maize genome (Table IX). Combining CBCS with genome filtration can greatly reduce the cost while retaining the high coverage of genic regions. An alternative approach is the identification of gene-rich regions on a detailed physical map and sequencing large-insert clones from these regions.

The banana genome is relatively small, 500- to 600-Mb (slightly bigger than rice) DNA across 11 chromosomes. A Global *Musa* Genomics Consortium (GMGC) is already in place to decode the *Musa* genome (http://www.newscientist.com/article.ns?id-dn1037); already two BAC clones of *Musa acuminata* Calcutta 4 have been sequenced (Table IX). The *Musa* genome has unique characteristics that will provide researchers with a powerful model for investigating fundamental questions with potentially widespread applications to agriculture. For example, comparing the genome of wild bananas that reproduce sexually with those of asexual crop bananas will provide insights into how quickly plant genomes evolve or comparing the genomes of wild Asian cultivars with those of African cultivars will provide an uncommon look at the effects of disease agents on genome evolution of the two species (*M. acuminata* and *M. balbisiana*), which gave rise to most cultivated bananas. A Global Cassava Partnership (GCP), an alliance of the world's leading cassava researchers and developers, has proposed that sequencing the cassava genome should be a priority (Fauquet and Tohme, 2004). The US Department of Energy's Joint Genome Institute (JGI) is providing fund and technical assistance to decode the cassava genome involving 10 institutes (http://www.ars.usda.gov/is/pr/2006/060830.htm). The benefits of deciphering cassava's genetic code include not only high-yielding pest- and disease-resistant cultivars with high protein content but also boosting its potential for fuel ethanol in developing countries. Genomic information from cassava could

also expedite research to reestablish castor bean, a close relative, as domestic source of industrial oil, minus the toxin ricin. Researchers from Purdue University and those from the JGI are sequencing the genome of soybean, *Glycine max*, the world's most valuable legume crop, to locate genes on the soybean chromosomes in order to create a physical map. Integrating the physical map with parts of the genetic map already available will ultimately allow sequencing of the entire soybean genome (http://www.csrees.usda.gov/newsroom/news/csrees_news/06news/soybean_dna.html).

Completed genome sequences provide templates for the design of genome analysis tools in orphan species lacking sequence information. For example, Feltus *et al.* (2006) designed 384 PCR primers to conserve exonic regions flanking introns, using sorghum and millet EST alignment to the rice genome. These conserved-intron scanning primers (CISPs) amplified single-copy loci at 37 to 80% success rates; that is, sampling most of the ~50 million years of divergence among grass species. When evaluating 124 CISPs across rice, sorghum, millet, Bermuda grass, tef, maize, wheat, and barley, about 18.5% of them seemed to be subject to rigid intron size constraints that were independent of per nucleotide DNA sequence variation. Likewise, about 487 conserved-noncoding sequence motifs were identified in 129 CISP loci. As pointed out by Feltus *et al.* (2006), CISP provides the means to effectively explore poorly characterized genomes for both polymorphism and noncoding sequence conservation on a genome-wide or candidate gene basis, and also provides anchor points for comparative genomics across a diverse range of species. After sequencing the whole genome of the major food crops, plant breeders may access new gene tools that will facilitate their ability to select outstanding individuals with resistance to biotic and abiotic stresses, possessing good seed quality, and producing more than the existing available cultivars.

C. Genetic Linkage Map

Genetic linkage mapping refers to determining the order and genetic distance between loci along chromosomes using recombination-based information in segregating populations. In contrast, physical mapping determines the absolute distance between different parts of the genome. Generally, researchers have started by producing a high-resolution genetic map populated with markers; then produced, fingerprinted, and assembled a deep-coverage library of bacterial artificial chromosomes (BACs); and then through comparative analysis of molecular markers, integrated the genetic and physical maps.

Marker-dense meiotic linkage maps serve multiple purposes ranging from dissection of simple and complex phenotypes to the isolation of genes by

map-based cloning (Tanksley *et al.*, 1995), facilitating for the construction of physical maps (Klein *et al.*, 2000), and for developing MAS of desirable genes in breeding programs (Burr *et al.*, 1983; Tanksley *et al.*, 1989). Meiotic linkage mapping uses the frequency of recombination events that occur during meiosis as a basis for calculating genetic distances between loci. The observed recombination frequency is commonly converted into map units (Centimorgan) by applying a mapping function (Kosambi, 1944), and by following the segregation of genetic markers in a meiotic mapping population, recombination events are linearly ordered along each chromosome, thus defining intervening segments of chromosomes, which vary in both physical and genetic size. The size of the mapping population, the number of markers, and the number of recombination events that occur during meiosis greatly influence the quality of resultant map. The genetic map provides a framework for anchoring the physical map. Deep-coverage large-insert genomic libraries, such as yeast artificial chromosomes (YAC) or BACs, are used for constructing the physical map. BACs are most preferred over YAC in plants for the construction of large-insert genomic libraries as they are easy to manipulate, produce low frequency of chimerism, and the clones are highly stable. By merging probe-to-BAC hybridization data with DNA fingerprint data, and using the BACRF method (Lin *et al.*, 2000) to resolve the chromosomal origin of BAC clones detected by multiple-DNA probes, the robustness of a physical map is improved over other methods that use arbitrary primer PCR-based fingerprinting of complex DNA populations resulting from pooling of low-coverage BAC libraries (Klein *et al.*, 2000). Cytogenetic stocks can also be used to generate a physical map by using genetically mapped DNA markers linked to specific chromosomal segments in cytogenetic stocks. However, isolation of a large number of cytogenetic stocks is a daunting task and not possible at all in some crops. For example, deletion stocks are generally not viable in diploid species. Additionally, the resolution of a physical map based on cytogenetic stocks is not only dependent on the number of stocks but also on the accuracy of their cytological characterization. A cytologically defined chromosomal fragment can include several megabases of DNA, which could significantly limit the power of such physical maps. The integrated genetic and physical genome maps are extremely valuable for map-based gene isolation, comparative genome analysis, and as sources of sequence-ready clones for genome sequencing.

Genetic linkage maps are reported for most of the legumes (Dwivedi *et al.*, 2005; Table X) and for cereals, and clonal crops (Table X), but with varying marker density and genomic coverage. For example, crops such as barley, maize, potato, rice, sorghum, and wheat have high-density genetic maps, while cassava, *Musa*, oat, pearl millet, sweet potato, and yam have less saturated genetic linkage maps. Soybean and common bean are the only

Table X
Overview of the Genetic and/or Physical Maps Reported in Azuki Bean, Banana, Barley, Black Gram, Cassava, Maize, Oat, Peanut, Pearl Millet, Potato, Rice, Sorghum, Sweet Potato, Wheat, and Yam

Marker and mapping population	Summary of the genetic and/or physical map	References
Azuki bean		
486 markers (SSR, RFLP, AFLP) and 187 BC_1F_1 (JP81481 × *Vigna nepalensis*)	486 markers mapped into 11 LGs spanning 832.1 cM with an average marker distance of 1.85 cM, 95% genome coverage, LGs length ranging from 54 to 124 cM and marker loci from 28 to 75 per LG	Han *et al.*, 2005
Banana		
90 markers (RFLP, RAPD, isozyme) on 92 F_2 (SF265 × Banksii)	77 of the 90 loci mapped on 15 LGs (ranging from 4 to 80 cM) with a total map length of 606 cM while 13 segregated independently	Fauré *et al.*, 1993
Barley		
252 SSR and 86 DHL (Lina × *H. spontaneum*)	242 markers on 7 LGs, with a total map length of 1173 cM that is comparable to those observed in DHLs using RFLPs (Heun *et al.*, 1996) but showing strong segregation distortion around the centromeric region of chromosome 2 H	Ramsay *et al.*, 2000
1172 markers (AFLP, SSR, STS, and *vrs1*) and 95 RIL (Russia 6 × H.E.S. 4)	The map consists of 7 LGs with a total distance of 1595.7 cM, and average marker density of 1.4 cM per locus. This map length longer than those of Ramsay *et al.* (2000) (1173 cM) or Costa *et al.* (2001) (1387 cM)	Hori *et al.*, 2003
1237 markers (SNP, SSR, RFLP, AFLP) and 3 DH populations	The integrated map based on 3 mapping populations consisted of 1237 loci, grouped into 7 LGs, with a total map length of 1211 cM and an average marker density of 1 locus per centimorgan	Rostoks *et al.*, 2005
Black gram		
145 markers (RFLP, AFLP, SSR, and morphological) and 180 BC_1F_1	The map consists of 11 LGs with a total distance of 783 cM, markers per LGs ranging from 6 to 23 and average distance between markers varying from 3.5 to 9.3 cM	Chaitieng *et al.*, 2006
Cassava		
168 markers (RFLP, RAPD, SSR, isozymes) and (TMS 30573 × CM 2177-2) F_1	The map consists of 20 LGs spanning 931.6 cM, with an average marker density 7.9 cM and covering 60% of the cassava genome. The male gametes-derived map contains 159 markers, 24 LGs, and 1220 cM map. Reduced recombination in gametes of the female parent resulted greater genetic distances on the male gamete-derived map between markers common to both parents	Fregene *et al.*, 1997

472 SSR and 286 F_2 (TMS 30572 × CM 2177-2)	The map has 100 markers spanning 1236.7 cM, distributed on 22 LGs with an average marker density of 12.36 cM, and markers uniformly distributed across cassava genome	Okogbenin *et al.*, 2006
Maize		
1736 markers (EST and STS, 90 core marker, and 237 from other grass species) and 54 F_2 (Tx303 × Co159)	The 1736 loci mapped on 10 LGs, with a total map length of 1727.4 cM and marker density of 0. 9 cM. 90 core markers with even spacing along chromosome delineate the 100 bins on the map with an average bin size of 17 cM. This map provides a more than fivefold increase in number of loci compared to previous map published in this population (Chao *et al.*, 1994) but slightly smaller than that of Matz *et al.*, 1995 (1883.6 cM) and Causse *et al.*, 1996 (1765 cM)	Davis *et al.*, 1999
184 RFLP and 748 SSR and 277 RIL (B73 × Mo17)	The 803 loci mapped on 10 LGs, with a total map length 4906 cM (347.7–714.5 cM per chromosome) of IBM map, with an average marker density of 6.6 cM	Sharopova *et al.*, 2002
954 cDNA probes and two RIL populations: IBM (B37 × Mo17) and LHRF (F2 × F252)	Framework maps consists of 237 and 271 loci in IBM and LHRF populations, that both maps contain 1454 loci (1056 on IBM_Gnp2004 and 398 on LHRF_Gnp2004) corresponding to 954 cDNA probes, and map size of 1825 cM for IBM_Gnp2004 and 1862 cM for LHRF_Gnp2004	Falque *et al.*, 2005
Oat		
441 markers (RFLP, AFLP, RAPD, STS, SSR, isozyme, morphological) and 136 $F_{6:7}$ RIL (Ogle × TAM O-301)	426 loci produced 34 LGs (with 2–43 loci each) spanning 2049 cM of the oat genome (from 4.2 to 174.0 cM per LG). Comparisons with other *Avena* maps revealed 35 genome regions syntenic between hexaploid maps and 16–34 regions conserved between diploid and hexaploid maps. 89–95% conservation of diploid genome regions on the hexaploid map; however, much lower colinearity at whole chromosome level	Portyanko *et al.*, 2001
510 markers (RFLP, AFLP, and SSR) and 152 $F_{2:6}$ RIL (Ogle × MAM17-5) (OM)	28 LGs, containing from 3 to 33 markers and varying in size from 5.2 to 123.0 cM, with a total distance of 1396.7 cM. Comparison with previously published hexaploid map from Kanota × Ogle (KO) (O'Donoughue *et al.*, 1995) revealed 9 OM LGs homologous to the LGs in the KO map	Zhu and Kaeppler, 2003

(*continued*)

Table X (*continued*)

Marker and mapping population	Summary of the genetic and/or physical map	References
Peanut		
204 SSR and 93 F_2 (*Arachis duranensis* × *Arachis stenosperma*)	SSR- and AA-genome-based map consists of 11 LGs covering 1230.89 cM, with an average marker density of 7.24 cM. This map is comparable to the 1063 cM in previously reported map from two AA-genome diploid species (Halward *et al.*, 1993) and to half of the 2210 cM reported for tetraploid map (Burow *et al.*, 2001)	Moretzsohn *et al.*, 2005
Pearl millet		
418 (RFLP and SSR) markers and four populations	A consensus genetic map of 353 RFLP and 65 SSR markers mapped on 7 LGs, with 85% of the markers clustered and occupy less than a third of the total map length; marker density in four maps ranged from 1.49 to 5.8 cM.	Qi *et al.*, 2004
Potato		
230 RFLP probes and two mapping populations	304 RFLP loci mapped on the 12 LGs with a total map length of 1034 cM and marker density of 3.4 cM. Comparisons between potato RFLP maps revealed conservation of marker order but differences in chromosome and total map length	Gebhardt *et al.*, 1991
RFLP (potato and tomato) and BC_1 [(*Solanum tuberosum* × *Solanum berthaultii*) *S. berthaultii*]	High-density map contains more than 1000 markers with an average marker density of ~1.2 cM, differentiating the tomato and potato genomes by 5 chromosomal inversions	Tanksley *et al.*, 1992
>10,000 AFLP markers and heterozygous diploid potato	An ultradense genetic linkage map with >10,000 AFLP loci, with marker density proportional to physical distance and independent of recombination frequency	van Os *et al.*, 2006
Rice		
726 markers and 113 BC_1 (BS125 × WL02) BS125	The map consists of 12 LGs with a total distance of 1491 cM and average marker density of 4.0 cM on the framework map, and 2.0 cM overall	Causse *et al.*, 1994
2275 markers and 186 (Nipponbare × Kasalath) F_2	The map consists of 12 LGs with a total distance of 1521.6 cM, and average marker density of 0. 67 cM per locus	Harushima *et al.*, 1998
703 markers and *japonica* cultivar Nipponbare	Physical map of rice chromosome 10 developed using FISH mapping of BAC clones on meiotic pachytene chromosomes that fully integrate with a genetic linkage map of rice chromosome 10 with uniform distribution of genetic recombination but with suppression in centromeric region	Cheng *et al.*, 2001
	BAC-based physical map of chromosome 4 consists of 11 contigs with a total length of 34.5 Mb, 94% of the chromosome size (36.8 Mb), long and short arm length 5.13 and 2.9 Mb, respectively	Zhao *et al.*, 2002

	BAC-based physical map of rice developed that represents ~90.6% of the rice genome, and its comparison with genetic map reveals that recombination is suppressed severely in centromeric regions as well on short arms of chromosomes 4 and 10	Chen *et al.*, 2002
6713 EST from 19 Nipponobare cDNA libraries screened on 4387 YAC clones	YAC-based transcript map consists of 6591 ESTs covering 80.8% of the genome, with chromosomes 1, 2, and 3 have relatively high EST densities, approximately twice those of chromosomes 11 and 12, and contain 41% of the total EST sites on the map. Most EST dense regions distributed on the distal regions of each chromosome arm	Wu *et al.*, 2002b
Sorghum		
470 loci (147 SSR, 323 RFLP) and 137 RIL (BTx623 × IS3620C)	The map consists of 470 loci that mapped into 10 LGs, with a total map distance of 1406 cM and average marker density of 2.99 cM	Bhattramakki *et al.*, 2000
2590 PCR-based markers and 137 RIL (BTx623 × IS3620C)	The 1713 cM map encompassed 2926 loci distributed on 10 LGs, and markers mapped between 121 and 243 on these LGs	Menz *et al.*, 2002
187 markers on 225 RIP 1 (IS9830 × E 36-1) and 228 markers on 226 RIP2 (N13 × E36-1)	The RIP 1 map consisted of 187 markers (AFLPs, SSRs, RFLPs, and RAPDs) distributed over 10 LGs with a total map length of 1265 cM while RIP 2 map had 228 markers spread into 12 LGs with a total map length of 1410 cM. The combined map contained 339 markers on 11 LGs with a map length of 1424 cM, comparing well with other maps except for few inversion, deletions, and additional distal regions	Haussmann *et al.*, 2002
2050 RFLP probes and 65 F_2 (*Sorghum bicolor* × *S. propinquum*)	The *S. bicolor* × *S. propinquum* map is composed of 2512 loci on 10 LGs that collectively span 1059.2 cM, with an average marker density of 0.4 cM	Bowers *et al.*, 2003
Sweet potato		
AFLP markers and (Tanzania × Bikilamaliya) F_2 population	632 (Tanzania) and 435 (Bikilamaliya) AFLP markers placed in 90 and 80 LGs, respectively. Total map lengths were 3655.6 and 3011.5 cM, respectively, with an average distance of 5.8 and 6.9 cM, respectively, between adjacent markers	Kriegner *et al.*, 2003

(*continued*)

Table X (*continued*)

Marker and mapping population	Summary of the genetic and/or physical map	References
Wheat		
230 SSR and ITMI population (Opata 85 × W7984)	279 loci amplified by 230 primers placed on to a genetic framework map composed of RFLPs previously mapped in ITMI population. 93 loci mapped to the A genome, 115 to the B genome, and 71 to the D genome. The markers randomly distributed along the linkage map, with clustering in several centromeric regions	Röder *et al.*, 1998
567 markers (RFLP, AFLP, SSR, and morphological and biochemical) and 96 DHL (CS × SQ1)	The genetic map consists of 567 markers assigned to 21 LGs, with a total map length of 3521.7 cM. Approximately similar map length for the A (1148.0 cM), B (1204.8 cM), and D (1168.9 cM) genomes but the D genome had only half the markers (115) of the other two genomes (A, 224; B, 228). This map is very similar in length to those reported for the ITMI map (3551 cM), CS × Synthetic map (2,830 cM), Arina × Forno map (3086 cM), and other 3 maps of 3164–4110 cM	Quarrie *et al.*, 2005 and references therein
478 SSR and 96 DHL (Kitamoe × Münstertaler)	The first SSR-based linkage map from intraspecific cross of common wheat consisted of 464 loci spread into 23 LGs, with a total map length of 3441 cM covering 86% wheat genome	Torada *et al.*, 2006
Yam		
341 AFLP markers and intraspecific F_1 population	The maternal map consists of 155 markers, 12 LGs, 891 cM map distance and 7.4 cM marker density while the paternal map contains 157 markers, 13 LGs, 852 cM map distance and 6.5 cM marker density	Mignouna *et al.*, 2002a

legume crops that have saturated maps (Dwivedi *et al.*, 2005). The large variation in map length results from differences in number of chromosomes and total size of the genomes as well as the use of different numbers of markers (increasing the number of markers will generally, until a certain threshold is reached, give a larger total map length), inclusion of skewed markers (that tend to exaggerate map distances), and use of different mapping software (which vary in estimates of genetic distances). In addition, many published maps report more linkage groups (LGs) than the basic chromosome number of that species. This is frequently the result of insufficient marker density, as most saturate maps can be directly aligned with the basic chromosome complement (Tekeoglu *et al.*, 2002).

The generation of integrated genetic and physical maps in many crops of significant economic importance is a daunting task because of large genome size, large amount of repetitive DNA, and polyploidy nature. However, genome-wide physical maps are reported in rice (Chen *et al.*, 2002; Cheng *et al.*, 2001), sorghum (Klein *et al.*, 2000), and maize (Coe *et al.*, 2002; Cone *et al.*, 2002; Yim *et al.*, 2002), which will be useful in genome sequencing, targeted marker development, efficient positional cloning, and high-throughput EST mapping in these and also closely related lesser studied crops wherein the genomic resources are not as developed as in these crops. For example, the sorghum genetic and physical map has been aligned to varying degrees with the genetic maps of wheat, rice, sugarcane, maize, and *Arabidopsis* and with the QTL mapped in these taxa.

There is a growing awareness that levels and patterns of allelic diversity are related to the chromosomal context of a locus. ''Diversity maps'' showing the distribution(s) of allelic diversity across the chromosomes and genomes of a variety of organisms are also related to structural features of chromosomes such as centromeres and telomeres and with the unique selection pressure specific to certain gene pools (Dvorak *et al.*, 1998; Gaut *et al.*, 2000; Hamblin and Aquadro, 1999). Diversity analysis of individual genes promises to shed new light on crop productivity and evolutionary processes underlying plant domestication (Wang *et al.*, 1999). When Draye *et al.* (2001) constructed diversity maps with genome-wide resolution based on neutral DNA markers for three gene pools in sorghum (*Sorghum propinquum*, *S. halepense*, and *S. almum*), they found a number of common features and also some key differences. Each gene pool showed low levels of variation near the central region of the LG ''G'' and both termini of the LG. The cultivated sorghum showed by far the lowest level of diversity of the three gene pools, the exotic diploid sorghum showed intermediate diversity, and the polyploids showed remarkably high levels of diversity. Similarly in one region near the marker *Psb347*, the tetraploid gene pool showed unusually high level of diversity, whereas the two diploid gene pools each showed unusually low levels of diversity. Crops with high resolution of genetic maps, such as rice,

maize, and sorghum, are ideal for developing diversity maps that promise new information about the consequences of natural selection, domestication, and polyploidy formation. Clearly, the approach of relating molecular level variation to phenotypic diversity is an essential precursor for diversity analysis studies using large populations of candidate genes. In this way, QTL information can be used together with association approaches to select a small number of candidates most likely to be directly related to a specific phenotype.

D. Marker-Trait Associations from Analysis of Diverse Germplasm

Conventional genetic linkage mapping approaches for polygenic traits are confounded by epistasis (adaptation and phenology traits influencing the target trait) and GEI (reducing the accuracy of phenotype data) that erodes the precision and power of QTL detection. In addition, linkage mapping has two other major constraints, particularly affecting practical applications: (1) marker-trait associations determined in genetic populations must be validated in target breeding populations before routine application can be considered which is time consuming and often introduces a major level of redundancy into the process, and (2) marker-trait associations identified in this way are based on genetic distance in the mapping population and tight linkage (and thus power of selection) may be eroded or lost entirely when the marker is applied to breeding populations with very different recombination patterns between the target loci and marker. Association mapping (AM), also known as linkage disequilibrium (LD) mapping, is a method that relies on LD to study the relationship between phenotypic variation and genetic polymorphism (Flint-Garcia *et al.*, 2003). LD refers to nonrandom association between two markers, or two genes, or between a gene and a marker locus. Mutation, population structure, epistasis, population perturbations like migration, inbreeding, and selection all influence LD, and some of these can lead to spurious associations (Jannink and Walsh, 2002). AM deals with unrelated individuals or members of a family with varying levels of phenotypic expression that are evaluated to detect and measure the degree of association between molecular markers and traits of interest. The principal advantage of this procedure lies in its ability to capture informative data stored in unrelated individuals who have undergone several rounds of gene shuffling over multiple generations. Significantly, it can be used on material offering better overall relevance to breeding programs and thus reduce the level of redundancy between marker identification and marker validation steps. AM can be investigated using candidate genes as well from randomly chosen molecular markers that are evenly distributed across genome. Indeed,

for outbreeding crops such as maize, the use of this type of marker in AM is highly desired.

There are many reviews describing the fundamentals of LD mapping (Boreck and Suarez, 2001; Flint-Garcia *et al.*, 2003; Gupta *et al.*, 2005a; Rafalski and Morgante, 2004). Both gene-based and genome-wide or chromosome-wide LD-based AM detected association of DNA markers with ecology, geography, disease resistance, and agronomic and seed quality traits in higher plants, thus being a viable alternative to classical QTL analyses (Dwivedi *et al.*, 2005 and references therein; Breseghello and Sorrells, 2006a; Gupta *et al.*, 2005a; Kraakman *et al.*, 2006; Maccaferri *et al.*, 2005; Malysheva-Otto *et al.*, 2006; Morrell *et al.*, 2005; Roy *et al.*, 2006; Stich *et al.*, 2006; Szalma *et al.*, 2005). In addition, many of the associated markers were located in chromosome regions previously identified as harboring QTL for yield and yield components, providing good validation that AM of diverse germplasm is a viable alternative to classical QTL analyses based on crosses between inbred lines (genetic populations), especially for complex traits (Breseghello and Sorrells, 2006a; Kraakman *et al.*, 2006; Szalma *et al.*, 2005). Large variation in LD estimates in different plant genomes or in different parts of the genome of an individual species is reported: 10–20 cM in barley and wheat, 100 kb in rice, <4 to ≤ 10 kb in sorghum, <50 kb in soybean (all self-pollinated species). The LD estimates in cross-pollinated crops ranged from 0.4 to 1.0 kb in maize, <3 cM in sugar beet, 0.3–1.0 cM in potato, and 10 cM in sugarcane (Gupta *et al.*, 2005a and references therein). Inbreeding drives lineages to homozygosity rendering recombinations ineffective in breaking down LD, while rapid decay of LD in outbreeding is probably because of increased crossover effects. Population-wide associations between loci due to LD can be used to map QTL with high resolution. However, spurious associations between markers and QTL can also arise as a consequence of population stratification and statistical methods that cannot differentiate between loci associations due to linkage disequilibria from those caused in other ways can render false-positive results (Deng *et al.*, 2001). The transmission-disequilibrium test (TDT) is a robust test for detecting QTL. TDT exploits within-family associations that are not affected by population stratification (Spielman *et al.*, 1993). It is used to check jointly for linkage and LD by testing whether alleles at a particular marker locus segregate randomly from parents to a specific subset of their offspring. TDTs have been developed for dichotomous and quantitative traits (Allison, 1997; Martin *et al.*, 2000; Rabinowitz, 1997; Zhao *et al.*, 2000). However, some TDTs are formulated in a rigid form, with reduced potential applications. Hernández-Sánchez *et al.* (2003) developed TDT that uses mixed linear models to allow greater statistical flexibility. In this test, allelic effects are estimated with two independent parameters: one exploiting the robust within-family information and the other the potentially biased

between-family information. Using this approach, they confirmed previous observations on effects of the fourth melanocortin receptor (MC4R) on production traits in pig that polymorphism is either causal or in very strong LD with the causal mutation, and provided no evidence for spurious associations.

Breseghello and Sorrells (2006b) compared the potentials and limitations of germplasm bank collections, synthetic populations, and elite germplasm as experimental materials for association analysis integrated with plant breeding practices and the application of AM differs among those populations in several aspects. They found that synthetics offer the most favorable balance of power and precision for association analysis and would allow mapping of quantitative traits with increasing resolution through cycles of intermating. Hence, Breseghello and Sorrells (2006b) proposed a model to describe the association between markers and genes as conditional probabilities in synthetic populations under recurrent selection, which can be computed on the basis of assumptions related to the history of the population. This model is useful for predicting the potential of different populations for association analysis and forecasting the response to MAS.

For efficient integration of AM with other methods currently in use, materials that are routinely generated and evaluated should be used for both purposes. For example, in case of germplasm, core collections (see Section II.A) are expected to represent a large proportion of the total genetic variability with a manageable number of accessions, and thus are suitable for genetic studies. Core collections representing the genetic diversity of a species are attractive for AM because of the wide allele diversity encompassed within a relatively small number of genotypes for which replicated multilocational precision phenotyping is feasible. The level of LD in a crop germplasm collection determines the scale at which AM will resolve the localization of favorable variation in the genome. The use of genome-wide survey for selecting a less-structured subsample of accession improves the significance of results and thus opens the door to genome-wide association studies and supports the identification of reference collection to integrate phenotypic and molecular characterization efforts (Deu and Glaszmann, 2004). The process of selection of a minimum sample with maximum variation has a normalizing effect that is expected to reduce population structure and LD between unlinked loci, thus creating a situation favorable for AM (Breseghello and Sorrells, 2006a). A difficulty likely to occur in this type of material is related to genetic heterogeneity within samples. Thus, it is not recommended at this time to use primary landraces and natural populations or any other mixture of genotypes, which will confound the genotyping and erode the precision of phenotyping. For elite materials, the sample could be composed of lines and checks evaluated in regional trials, whereas for synthetic populations, the evaluation unit should be largely homogeneous, whether it is an individual or

a family. Core collections are useful materials for AM for quality traits such as disease resistance, seed quality, and domestication-related traits. Conversely, the broad genetic variability of those collections normally makes them unsuitable for analysis of quantitative traits because part of accessions would be unadapted to the growing conditions and prevalent diseases of the test environment, resulting in poor precision of trait measurement. Similarly, phenological traits are likely to be highly variable in core collections which will highly confound attempts to measure traits such as abiotic stress tolerances. Elite lines are the most desirable materials for AM when attempting to analyze low heritability traits, including yield, yield components, and tolerance to abiotic stresses because elite lines are genetically stable and are well adapted to specific known growing conditions (Breseghello and Sorrells, 2006b). Synthetic populations are normally designed and maintained by random mating, and therefore population structure is expected to be mild or absent, which is an important advantage of synthetics for AM. The level of LD in synthetic populations is expected to be high in the initial generations, such that a genome scan could detect large chromosome segments associated with traits, and trace them back to parental haplotypes. In subsequent generations, the decay of LD by recombination would favor refined mapping. However, synthetic populations are often subjected to intense recurrent selection which could build up LD by favoring allelic combinations or by promoting genetic drift (Palaisa *et al.*, 2003). For this reason, populations subjected to mild or no selection would be preferred for AM. Alternatively, marker analysis of a large number of available genotypes can be used to define a subset of lines that represent the desired population structure for AM. AM in synthetic populations under selection will require intensive genotyping because in each cycle, new progenies have to be tested to reflect the current state of the population and for implementation of MAS. On the other hand, information about the population is cumulative over years, allowing a progressively refined genetic analysis of traits of interest to the breeding programs.

Both linkage analyses (LA) and LD mapping have their own limitations when used alone. Therefore, a joint linkage and LD mapping strategy has been devised for genetic mapping (Wu and Zeng, 2001; Wu *et al.*, 2002a) that has power to simultaneously capture the information about the linkage of the markers (as measured by recombination fraction) and the degree of LD created at historic time. This approach is based on the principle that during the transmission of genes from parents to progeny, linkage between marker and QTL is broken due to meiotic recombination. Thus, by combining the information about the linkage and LD, the joint mapping method displays increased power to detect LD compared to traditional methods of LD analyses. The use of this approach has also been suggested for multitrait fine-mapping of QTL (Lund *et al.*, 2003; Meuwissen and Goddard, 2004).

Like the genetic and physical maps developed in many plant genomes, LD maps can also be constructed in plants as is being done in humans using ALLASS and LDMAP VERSION 0.1 (University of Southampton, United Kingdom) softwares. These LD maps will make use of molecular markers that flank marker intervals delimited on the basis of estimations of LD, the distance being represented as LD units (Zhang *et al.*, 2002).

III. MARKER VALIDATION AND REFINEMENT

It is clear from Section II that there have been major advances that have occurred in the development of DNA markers, construction of genetic linkage maps, and the mapping of economic traits controlled by major genes and QTL. While the number of reports of mapped genes continues to grow rapidly, the literature on practical validation application of those markers in breeding populations remains relatively limited. One reason for this is that there are several scientific and logistical issues that must be resolved before a practical MAS strategy can flow from a mapping study, and at each step there will be a certain level of redundancy. Moreover, in some cases, researchers are more interested in understanding the genetic control of the trait and subsequent gene discovery, thus leave the validation and application to plant breeders who may be less interested in publishing their findings. Furthermore, once the mapping study is published, it may be difficult to publish the results of activities associated with validation, refinement, and application of those markers, particularly if the selective power of the marker lessened or lost when applied in breeding programs. This generally involves validation of the QTL or gene marker in a different set of germplasm or populations and development of markers assays suitable for high throughput, low cost, and MAS (Collard and Mackill, 2007; Langridge *et al.*, 2001). Marker validation step usually has some level of redundancy leading to the need to develop new markers or marker types around the target locus in order to find more polymorphic markers or develop gene-based markers for marker-trait associations that are shared across different breeding populations. The availability of thousands of SNP markers rather than several hundreds of SSR markers in some crops (Table VIII) that are currently being used makes it practical to validate marker-trait association through high-precision genotyping using the same set of markers for different parental lines and breeding populations. Thus, it is much more likely that the parents of any breeding population will be polymorphic for at least one of them, allowing breeders to track the alleles donated from each parent throughout the breeding process, speeding MAS and marker-assisted back crossing (MABC) in any cross. Marker validation can be also done through

selective genotyping and pooled DNA analysis, and development of gene-based markers and closely linked markers, as additions to testing marker-trait association in alternative or target populations. However, validation requirements can be minimized by MAS using large-effect QTL, precision phenotyping, identification of context independent QTL, mapping as we go, AM using large numbers of inbreds, genome-wide association scan, using breeding materials for mapping, and utilization of haplotype-based selection rather than single-marker based selection.

A. Markers for Simply Inherited Traits

For major gene traits such as many disease resistances, gene validation is fairly straightforward. In these cases, the effect of genetic background is usually minimal, and the ease of phenotyping makes fine-mapping of the gene simpler. In mapping studies, a gene for simply inherited trait can be mapped with adequate accuracy in a mapping population of 100–200 individuals. This can then be followed by fine-mapping involving larger populations of over 500 individuals. The fine-mapping will allow identification of tightly linked markers that will not suffer recombination between marker and target gene in segregating breeding populations. An alternative to use a tightly linked gene in MAS is to use flanking markers on either side of the gene. Use of both flanking markers ensures that the gene is accurately detected in segregating populations, but it can also result in the transfer of large chromosomal fragments along with the target gene (linkage drag) if the interval between the two markers is large. If the donor of the gene contains deleterious alleles that are linked to the target gene, it will be necessary to identify more tightly linked flanking markers (Frisch *et al.*, 1999a; Tanksley *et al.*, 1989). The process of fine-mapping can be carried forward to positional cloning of the target gene. Plant populations of several thousands are commonly used even in species with small genomes where recombination rates might be around 250 kb cM^{-1} (Durrett *et al.*, 2002).

The marker or markers identified during the process of fine-mapping may be suitable for direct application in breeding programs following some level of validation. However, in many cases, these markers may not be polymorphic in all breeding populations of interest, thus requiring the identification of alternative markers for those populations. For well-characterized genomes, this is straightforward. In rice, for example, any one of the 2414 SSR markers can be quickly identified from the dense public maps or located using the genome sequence in online databases. In addition to identifying markers tightly linked to the gene of interest, it is also useful to identify a similar set of around 10 markers 3–10 cM either side of the target gene (Langridge *et al.*, 2001). These markers can then be used to reduce the effects

of linkage drag if recombinant selection is practiced (Collard and Mackill, 2007).

An ideal marker for selection of the target gene would be one that provides 100% accurate prediction of the phenotype. Except for the traits in alien gene introgression regions, this usually requires a marker associated with the sequence change in the gene associated with the favorable allele. These are so-called "FM" (Andersen and Lübberstedt, 2003) or "perfect markers" (see Section II.B). These markers provide sufficient benefits for MAS application to justify cloning of important economic genes and QTL aside from the other benefits that gene discovery can bring (see Section VII.B).

B. QTL Marker for Complex Traits

The difficulty for phenotypic selection of many quantitative traits in plant breeding gave rise to an optimistic view of the prospects of MAS for QTL. However, to date very few studies have demonstrated the usefulness of marker-QTL information vis-à-vis conventional phenotypic selection for the development of genetically enhanced breeding populations. Many studies reported that no substantial genetic progress was achieved or only a fraction of putative QTL actually contributed to the improvement of the trait when validated through MAS (Bohn *et al.*, 2001; Bouchez *et al.*, 2002; Flint-Garcia *et al.*, 2003; Schneider *et al.*, 1997; Stromberg *et al.*, 1994; Yousef and Juvik, 2001a). Several factors contribute to false positive (Type I errors) in QTL mapping studies, including population structure and size, parental selection and genetic background effects, epistasis and inaccurate phenotyping, QTL × environment interaction and inappropriate evaluation conditions, and finally inappropriate logarithm of odds (LOD) thresholds or low statistical rigor (Beavis, 1998; Moreau *et al.*, 1998). Additionally, inaccurate phenotyping data in the mapping populations further reduce the capacity to detect real QTL.

In a literature search conducted for the crops under review from 1991 to 2005 in journals with high-impact factor, over 500 articles reported QTL contributing to phenotypic variance for several agronomic and seed quality traits as well resistance to biotic and abiotic stresses, predominantly in cereal crops such as barley, maize, rice, and wheat. In contrast, during the same period, there were only 80 articles that dealt with validation of the reported QTL (Tables XI and XII), concentrating mostly in wheat, barley, rice, maize, and few in common bean, soybean, pea, yam, and potato. However, the community has become more concerned about reporting false QTL discovery, with a resultant increase in the number of reports regarding validation of QTL.

The low resolution of most QTL mapping studies reduces the likelihood of successful QTL marker validate (Holland, 2004). In a milestone publication by Beavis (1998), the power, precision, and accuracy of QTL mapping was

Table XI
Validation of Marker/QTL Associated with Resistance to Biotic and Abiotic Stresses in Barley, Common Bean, Maize, Pea, Potato. Rice, Soybean, Wheat, and Yam

Trait	Gene	Validated marker/QTL	References
		Biotic stresses	
Barley			
Barley leaf scald (*Rhynchosporium secalis*)	*Rrs.B87*	Closest marker 2.2 cM from the resistance locus *Rrs.B87*	Williams *et al.*, 2001
BaMMV and BaYMV	*ym4*	OP-ZO4H$_{660}$	Ordon *et al.*, 1995
Barley stripe rust (BSR) (*Puccinia striiformis* Westend. f. sp. *hordei*)	QTL4, QTL5, and QTL7	QTL4 and QTL5 linked with BSR resistance at seedling stage; three QTL linked with BSR resistance at adult plant stage	Castro *et al.*, 2003a,b
BYDV	*Yd2*	YLM	Paltridge *et al.*, 1998
BaYMV	*rym1* and *ryn5*	A CAPS marker from an RFLP probe MWG2134	Okada *et al.*, 2003a
FHB (*Fusarium graminearum* Schwabe) and Kernel discoloration (KD)		Two major QTL (near *HVBKasi* and the *Vrs1* locus); a major QTL for KD and a QTL for FHB	Canci *et al.*, 2004; Mesfin *et al.*, 2003
Leaf rust (*Puccinia hordei*)	13 QTL	Six QTL (Rphq1–6)	van Berloo *et al.*, 2001
Leaf stripe (*Pyrenophora graminea*)	*Rdg2a*	MWG2018	Arru *et al.*, 2003
Net form of net blotch (NFNB) [*Drechslera teres* (Sacc.) Shoem. f. *teres* Smedeg]	7–12 QTL	EBmac0906 and Bmac0181	Raman *et al.*, 2003
Net type net blotch (NTNB) (*Pyrenophora teres* f. *teres*)	1–6 genes	M61P12K116, M55P13T311, Bmag0173, and Ebmac0874 l	Cakir *et al.*, 2003
Powdery mildew (*Erysiphe graminis* f. sp. *hordei*)	*Ml(La)*	MWG097-R,L and MWG097	Mohler and Jahoor, 1996
Russian wheat aphid (RWA) [*Diuraphis noxia* (Mordvilko)]	Two genes	ABG8 and KV1/KV2	Raman and Read, 2000
Spot blotch (SB), NTNB, Septoria speckled leaf blotch (SSLB), and leaf scald (LS)	2 QTL each for SB, NTNB, and SSLB and one QTL for LS	*Rcs-qtl-7H-2-4* and *Rcs-qtl-4H-4-6* for SB; *Rpt-qtl-3H-4* and *Rpt-qtl-4H-5-7* for NTNB; *Rsp-qtl-2H-7-11* and *Rsp-qtl-6H-10-14* for SSLB, and *Rrs-qtl-1H-1-4* for LS	Yun *et al.*, 2006

(*continued*)

Table XI (*continued*)

Trait	Gene	Validated marker/QTL	References
Common bean			
BCMV	A dominant gene, *I*, and six recessive genes	$SW13_{690}$	Melotto *et al.*, 1996
CBB (*Xanthomonas campestris* pv. *Phaseoli*)	A major and few minor genes	R7313 and R4865	Tar'an *et al.*, 1998
	bc-I^2	$SBD5_{1300}$	Miklas *et al.*, 2000
Maize			
Sorghum downy mildew (SDM) (*Peronosclerospora sorghi*) and Rajasthan downy mildew (RDM) (*Peronosclerospora heteropogoni*)	5 QTL	Three QTL for SDM and two QTL for RDM, a major QTL confers resistance to SDM and RDM	Nair *et al.*, 2005
Pea			
Ascochyta blight (*Mycosphaerella pinnodes*, *Phoma medicaginis* variety *pinodella*, *Ascochyta pisi*)	Many QTL	Six QTL on LG II, III, IV, V, and VI (two QTL)	Timmerman-Vaughan *et al.*, 2004
Potato			
PVY	*Ny* and *Ry*	CD17, GP125, CT168, and TG508 linked with Ry_{adg}	Hämäläinen *et al.*, 1997
Rice			
Blast [*Pyricularia grisea* (Cooke) Sacc.]	*Pi44(t)*	$AFLP_{348}$	Chen *et al.*, 1999
	Pi-z	MRG5836	Conaway-Bormans *et al.*, 2003
	Pi-ta^2, *Pi-*k^h, *Pi-*k^s, and *Pi-b*	SSRs *Pi-b* (RM138, RM166, RM208), *Pi-*k^h (RM144, RM224), and *Pi-*ta^2 (OSM89, RM155, RM7102)	Fjellstrom *et al.*, 2004
	Pi-z	SSRs AP5659-1, AP5659-3 and AP5659-5	Fjellstrom *et al.*, 2006
Gall midge (*Orseolia oryzae* Wood-Mason)	*Gm2*	$F10_{600}$ and $F8_{1700}$	Nair *et al.*, 1995
Sheath blight (*Rhizoctonia solani* Kuhn)	–	Six QTL	Pinson *et al.*, 2005

Soybean			
Brown stem rot (BSR) (*P. gregata*)	*Rbs1*, *Rbs2*, and *Rbs3*	BSR3.sp1, K375.sp1, 14H13.sp1, 21E22.sp1, 21E22.sp2, 30L19.sp1, 35E22.sp1, 98P22.sp2, and Satt244	Klos *et al.*, 2000
Root knot nematode [*Meloidogyne incognita* (Kofoid and White) Chitwood]	Few genes	Satt012, Satt358, Satt492, and Satt505	Li *et al.*, 2001a
Soybean cyst nematode (SCN) (*H. glycines* Ichinohe)		Two major QTL against resistance to SCN race 3	Wang *et al.*, 2001b
	rhg1, *rhg2*, *rhg3*, *rhg4* and *rhg5*	QTL containing *rhg1* on LG G and QTL *rhg4* on LG A2	Concibido *et al.*, 2004
Wheat			
Common bunt [*Tilletia tritici* (Bjrk.) Wint. and *T. laevis* Kuhn]	*Bt-10*	$UBC196_{590}$	Demeke *et al.*, 1996
FHB (*F. graminearum*)		Sumai 3-derived QTL on 3BS and 6BS	Anderson *et al.*, 2001
		gwm389, gwm493, gwm533, and gwm644	Yang *et al.*, 2003
		SSRs linked to the major QTL on chromosome 3BS	Zhou *et al.*, 2003b
Hessian fly [*Mayetiola destructor* (Say)]	*H1* to *H25*	OpA01 and OpA17	Dweikat *et al.*, 1994
Leaf rust (*Puccinia recondita* f. sp. *tritici*)	*Lr9*	$OPA\text{-}07_{1500}$, $OPR15_{950}$, and J13/1 + 2	Schachermayr *et al.*, 1994
	Lr19	*Ep-D1c*	Winzeler *et al.*, 1995
	Lr10	*Lrk10–6*	Schachermayr *et al.*, 1997
	Lr28	$OPJ01_{378}$	Naik *et al.*, 1998
	Lr28, *Lr35*, and *Lr39*	Puc19/HpaII900	Sharp *et al.*, 2001
	Lr19 and *Lr24*	STS	Singh *et al.*, 2004
	Lr9, *Lr10*, *Lr19*, *Lr24*, *Lr28*, *Lr29*, *Lr35*, and *Lr39*	RFLP and AFLP markers 1100 bp, 310 bp, 130 bp, 310 bp, 850/900 bp, 900 bp, and 100 bp	Blaszczyk *et al.*, 2004
		$SCS5_{550}$	Gupta *et al.*, 2005b

(*continued*)

Table XI (*continued*)

Trait	Gene	Validated marker/QTL	References
Leaf rust and leaf tip necrosis (LTN)	*Lr34* and *Ltn*	Major QTL for leaf rust (*QLr.sfr-7DS*) and *QLtn.sfr-7DS* for LTN located within the Xgwm1220-Xgwm130 interval	Schnurbusch *et al.*, 2004
Powdery mildew [*E. graminis* DM f. sp. tritici (Em. Marchal)]	*Pm1* and *Pm2*	Whs350–1,2	Mohler and Jahoor, 1996
	Pm1 to *Pm25*	Xgwm337	Huang *et al.*, 2000
		QPm.vt-1B, QPm.vt-2A, and QPm.vt-2B	Liu *et al.*, 2001c
	Pm1	Xsts638-7A, XE39M58-77-7A, and Xgwm344-7A	Stepien *et al.*, 2004
RWA [(*D. noxia* (Mordvilko)]	*Dn4*	X_{gwm106} and X_{gwm337}	Arzani *et al.*, 2004
Stem rust (*Puccinia graminis*)	*Sr2*	$gwm533_{120}$	Spielmeyer *et al.*, 2003
Stem rust and leaf rust	*Sr39* and *Lr35*	$Sr39F2/R3_{900}$	Gold *et al.*, 1999
Stem rust, leaf rust, and yellow rust	*Yr17*, *Lr37*, and *Sr38*	$VPM1_{383}$, $scar15_{550}$, and $Xgwm636_{104}$	Sharp *et al.*, 2001
Yam (*Dioscorea* spp.)			
YMV in white yam (*Dioscorea rotundata*)	*Ymv-1*	OPW_{850} and OPX_{850}	Mignouna *et al.*, 2002b
Anthracnose (*Colletotrichum gloeosporioides*) in water yam (*Dioscorea alata*)	More than one dominant gene	$OP17_{1700}$ and $OPE6_{950}$	Mignouna *et al.*, 2002c
Abiotic stresses			
Barley			
Aluminum (Al) toxicity	*Alp*	Bmag353	Raman *et al.*, 2001
Frost tolerance	*Fr1*	*OPA17* and *Psr637*	Toth *et al.*, 2004
Maize			
Abscisic acid (ABA)		Major QTL for leaf ABA	Landi *et al.*, 2005
Rice			
Submergence tolerance	*Sub1*	RM219 and RM464A linked to *Sub1*	Xu *et al.*, 2004b
Soybean			
Salt tolerance	*Ncl*	Sat_091 and Sat237	Lee *et al.*, 2004
Wheat			
Boron (B) toxicity	*Bo1*, *Bo2*, and *Bo3*	Xpsr680-7B and Xpsr160-7D	Jefferies *et al.*, 2000

Table XII
Validation of Marker/QTL Associated with Agronomic and/or Seed Quality Traits in Barley, Pea, Rice, Soybean, and Wheat

Trait	Gene	Validated marker/QTL	References
Barley			
Agronomic traits (grain yield, plant height, maturity, and lodging severity)	Two QTL on chromosome 3	aABG396 and aCDO113 loci on chromosome 3 flanked by aABG057 and aABG37	Larson *et al.*, 1996
	Many QTL with small to large effects	A QTL on "plus" arm of chromosome 7(5H)	Spaner *et al.*, 1999
		QTL1 and QTL6 on chromosome 3 and 6, respectively	Romagosa *et al.*, 1999
Diastatic power (DP) affecting malt quality	Nine QTL	Xabg057, Bmy1, and XEBmac501	Coventry *et al.*, 2003
Malt extract	25 chromosome regions	Two alleles each from chromosome 2H and 2 regions chromosome 5H	Collins *et al.*, 2003
Pea			
Lodging	Two genes	A001 and A004	Warkentin *et al.*, 2004
Rice			
Fragrance	*fgr*	SCU015RM and RSP04	Christopher *et al.*, 2004
Regeneration ability		RZ474 and RZ575	Kwon *et al.*, 2001
Semidwarf stature	*sd-1* to *sd-60*	*sd-1* linked with RG220 and RG109	Cho *et al.*, 1994
Soybean			
Seed weight, protein, and oil content	Many QTL	*cqProt-001* and *cqProt-002* for seed protein; *cqOil-001*, *cqOil-002*, and *cqOil-003* for oil content; *cqSd wt-001* and *cqSd wt-002* for seed weight	Fasoula *et al.*, 2004

(*continued*)

Table XII (*continued*)

Trait	Gene	Validated marker/QTL	References
Wheat			
Bread-making quality (HMW glutenins)	Six genes at Glu-1 loci on 1A, 1B, and 1D	A 15 bp in-frame insertion in Glu-B1–1d(B-x6) discriminate genotypes with good or bad bread-making quality	Schwarz *et al.*, 2004
Doughs	HMW glutenin subunits 1Dx5 + 1Dy10 linked with high dough strength/good bread while 1Dx2 + 1Dy12 with poor bread quality	Oligonucleotide primers: P1 and P2 (Dx2 and Dx5 alleles), P3 and P4 (Dy10 and Dy12 alleles), and P5 and P6 (Bx7 allele)	Ahmad, 2000
Flour color	QTL on chromosome 7A	$Xcdo347_{52}$	Sharp *et al.*, 2001
Grain protein content (GPC)	Six QTL	WMC41 and WMC415	Singh *et al.*, 2001b
HMW glutenins	Glu-1 and Glu-B1 locus	Ax2*F2543, Ax2*R3605, Bx7F-428, Bx7R693, Bx7F-572, Bx7R693, Dx5F384, and DxR655	Radovanovic and Cloutier, 2003
Noodle quality	GBSS locus	GBSS-4A null mutation	Zhao *et al.*, 1998
	null GBSS 4A allele	440 bp from GBSS4A	Briney *et al.*, 1998
	13 QTL	42 SSRs	Prasad *et al.*, 2003
Seed dormancy	Major QTL	*Xhbe03*	Torada *et al.*, 2005
Semidwarf	*Rht-B1b (Rht1)* and *Rht-D1b (Rht2)*	*Rht-B1b* and *Rht-D1b*	Ellis *et al.*, 2002
Storage protein (Gliadines and glutenins)	Alleles in Gli-B1 and Glu-B3 locus associated with variation in HMW and LMW, respectively	PCR product of genotypes with LMW-2 glutenin has 50-bp longer fragment than those with LMW-1 glutenin	D'Ovidio, 1993

clearly shown to be highly dependent on sample sizes (n). When populations of less than 500 individuals are used for QTL mapping (irrespective of marker density), the power to detect true QTL is low and the estimated proportion of the genetic variance explained by mapped QTL is overestimated (see below), and it is very unlikely that QTL with small effects will be identified.

Number of true QTL	h^2	Sample size	Power (%)	Bias (σ_g^2) (%)
10	0.30	100	9	+559
10	0.30	500	57	+144
40	0.30	100	3	+2104
40	0.30	500	11	+423
10	0.95	100	39	+197
10	0.95	500	94	+106
40	0.95	100	6	+690
40	0.95	500	46	+165

Bias in the estimated genetic variance occurs mainly due to sampling of small populations, where the true QTL that are not detected (most of them in small sample sizes) tend to enhance the apparent effects of those QTL that are detected, through what is often referred to as the "Beavis effect" (Beavis, 1998; Melchinger *et al.*, 1998). Using a large population composed of 976 F_5 maize testcross progenies evaluated in 19 environments, Schön *et al.* (2004) also detected large effect of sample size on the power of QTL detection as well as on the accuracy and precision of QTL detection. The number of detected QTL and the proportion of genotypic variance explained by QTL generally increased more with increasing population size than with increasing the number of test environments, although the average bias of QTL estimates and their range are reduced by increasing population size and by increasing the number of test environments. Cross-validation performed well with respect to yielding asymptotically unbiased estimates of the genotypic variance explained by the QTL. However, by increasing the population size from 478 to 976, the increase in the proportion of genetic variance explained by QTL per additionally tested genotype is smaller as compared to increasing the population size from 244 to 488. This diminishing returns relationship (as the population size is increased) is expected due to the nonlinear relationship between sample size and power of QTL detection beyond a certain threshold (Lynch and Walsh, 1998). Genetic factors, such as enzyme variation in metabolic pathways, lead to an L-shaped distribution of QTL effects for a given quantitative trait (Bost *et al.*, 2001). For example, this trend was reported for grain moisture in maize with the result that the distribution was skewed toward smaller values (L-shaped) (Schön *et al.*, 2004).

Care should also be taken to report QTL-trait associations only at higher significance thresholds to avoid false identification of QTL when in fact a QTL is not present (Type I error). For example, Bernardo (2004) suggested that to prevent false QTL from confusing the literature and databases, a detected QTL should, in general, be reported as a QTL only if it is identified at a stringent significance level (Type I error probability or $\alpha_c = 0.0001$). Increasing the size of the mapping population leads to both increased power (Beavis, 1998) and a lower rate of false-positive QTL. However, the breeders in general like to work on many populations with small sample size rather than concentrating on few populations with large sample size. This trend needs to be reversed in order to exploit the QTL information in crop breeding programs or otherwise to deploy statistical methods for combining QTL analysis from related populations. Also more efforts should be directed toward accurate evaluation of progenies (both at the genotypic and phenotypic level) in order to avoid application failures. Benjamini and Yekutieli (2005) suggested using a false discovery rate (FDR) estimate in QTL analysis. The FDR is the expected proportion of Type I errors. FDR-controlling procedures ensure reproducible results with few false positives and offering increased power of QTL discovery. The two advantages of the FDR approach, which make it particularly suitable for QTL analysis, are its flexibility regarding the amount of information in the data and its scalability. Controlling the FDR for multiple traits may result in no loss of power to detect QTL. However, a renewed optimism regarding QTL mapping has emerged based on analysis of cloned QTL, which indicates that the original low-density map positions are relatively accurate (Price, 2006). Clearly, marker validation should be carried out after initial QTL mapping in order to determine whether fine-mapping is required.

When traits are controlled by multiple QTL of small effect, the confidence intervals for their location are wide (Visscher *et al.*, 1996). For these QTL, flanking markers may be widely spaced (>20 cM) and a large chromosomal fragment will be transferred during MAS. Thus, QTL of relatively large effect are the most appropriate targets for MAS. These QTL are easier to validate, are more likely to be effective in different genetic backgrounds, and less likely to suffer from confounding linkage drag problems during MAS (Holland, 2004; Mackill, 2006). They are also easier to fine-map, a process which requires accurate differentiation between the phenotypes resulting from the two alleles of the QTL. QTL of large effect may also be readily detected even in populations of smaller size (Vales *et al.*, 2005).

The genetic background of parental genotypes of the mapping population has a profound effect on the number, location, and effects of the identified QTL. For example, if a QTL allele with beneficial effect is identified in population A, its introgression by means of MAS in population B will not necessarily lead to tangible benefits. This is because population B may

already have alleles of similar or even greater value at this QTL and/or because of different interactions between the QTL and the two genetic backgrounds. Campos *et al.* (2004) estimated that most drought-tolerant QTL detected in maize would have limited utility for applied breeding, partially due to the prevalence of genetic background and environment effects.

Use of MAS for transferring QTL is more suitable when a trait is being introduced from an exotic source into elite germplasm, thus ensuring higher levels of polymorphism and higher probability of expression of the gene/QTL in the new genetic background (i.e., more likely that the allele is different to the recipient). A mapping study involving an exotic donor crossed with an elite line lacking the trait will increase the chance that the identified markers will be useful in the targeted cultivars. Large-effect QTL are also more likely to be expressed in different genetic backgrounds. For traits controlled by smaller QTL, the effect of the background can be extreme. However, it is currently impossible to predict these interaction effects in most crops, thus field evaluation must be used to validate the expression of introgressed QTL.

Epistasis, as detected by identification of different QTL when the same donor is crossed to different parents, is often observed. In *Arabidopsis*, significant effects of epistasis were observed for two QTL found in a 210-kb interval controlling growth rate, with gene effects depending on genetic background (Kroymann and Mitchell-Olds, 2005). Li *et al.* (2006a) provided an example of complex interactions among QTL for partial resistance to bacterial blight (BB) in rice, and it is suggested that this results from genetic networks of the underlying genes. Clearly, even for QTL that are observed in multiple populations, their robustness for applications in breeding must still be validated in relevant populations. Development of reciprocal introgression lines is useful for estimating the effects of the genetic background. For many traits, the overlap of QTL detected in reciprocal genetic backgrounds is low, showing the large effect of background on trait expression.

QTL $\times$ E effects are another factor that must be considered during validation studies. There are many reports of the lack of consistency between QTL detected in different environments. For examples, when Paterson *et al.* (1991) evaluated F_2 and $F_{2:3}$ progenies in 3 environments, they detected 29 putative QTL distributed over 11 of the 12 chromosomes, accounting for 4.7–42% of the phenotypic variation for fruit size, soluble solids concentration, and pH in tomato. Of these, 4 were detected in all the 3 environments, 10 in 2 environments, and 15 only in a single environment. QTL mapping using the same rice population for analysis of seedling vigor revealed major differences for QTL detected at different temperature regimes (Redoña and Mackill, 1996). Experiments conducted with the same mapping population in nine environments showed that rice QTL detection for plant height and heading date was markedly affected by environment (Li *et al.*, 2003a). Drought stress at flowering adversely affects grain yield in maize that causes a delay in silking, an increase in anthesis silking interval (ASI), thus decrease in grain yield. Vargas

et al. (2006) identified QTL for ASI that are stable across the eight environments and corresponded well with those reported by Ribaut *et al.* (1996). For grain yield, Vargas *et al.* (2006) detected a much larger GEI than for ASI; however, a couple of QTL consistent across environments identified, thus confirming the previous report of the QTL for grain yield and yield components on chromosomes 1 and 10 (Ribaut *et al.*, 1997a).

Cross-validation of QTL in independent samples and in different genetic backgrounds and environments is necessary to obtain unbiased estimates of QTL effects and the proportion of the genetic variance explained by the detected marker-QTL association before using them in MAS breeding programs. In general, QTL detected in multiple mapping studies using different populations would be considered as the most important targets for MAS application. For example, a grain length and weight QTL near the centromere of rice chromosome 3 was identified in at least eight independent mapping studies and has been identified as a putative transmembrane protein (Fan *et al.*, 2006).

In some cases, mapping the QTL in multiple generations from the same cross can be used to confirm the presence of QTL, as was observed for sheath blight in a rice RIL population (Pinson *et al.*, 2005). Similarly, an advanced backcross population ($BC_2F_{6:8}$) validated all QTL for resistance to *Septoria* speckled leaf blotch of barley that had been identified in an RIL population with the same parents (Yun *et al.*, 2006). QTL detected in a rice RIL population were validated in NIL developed for the two major plant-type QTL (Kobayashi *et al.*, 2006). However, usually it is only the successful validations that are reported in the literature. A rare exception to this is Steele *et al.* (2006) who attempted to validate four root QTL during the three backcrosses aimed at transferring root QTL from the upland rice cultivar Azucena into the variety Kalinga III. While all four root QTL were successfully introduced, only one showed a significant effect when transferred into the Kalinga III background. Where recurrent selection is used in breeding programs, QTL effects can change over time in subsequent selection cycles. This led to the development of the "Mapping As You Go" (MAYG) approach (Podlich *et al.*, 2004), where QTL effects are estimated in each cycle before selection and intermating are performed.

Fine-mapping of QTL is very useful for identifying tightly linked markers that will not suffer from loss of linkage due to recombination between marker and QTL during applications in different breeding populations. This will also serve to minimize the size of the introgressed fragment during backcrossing. Few QTL with major effects on traits of agricultural importance have been fine-mapped and successfully delimited their position on the chromosome in tomato, rice, wheat, and maize (see Section VII.B).

IV. SUCCESSFUL APPLICATIONS OF MARKER-ASSISTED GENETIC ENHANCEMENT IN PUBLIC SECTOR BREEDING PROGRAMS

MAS is most useful for traits where phenotypic evaluation is expensive or difficult, particularly for those polygenic traits with low heritability that are highly affected by the environment. It is also useful to break linkages between the target traits and undesirable genes in so-called marker-accelerated backcross breeding. MAS may also offer the opportunity to address goals not possible through conventional breeding, such as pyramiding different sources of disease resistance that have similar phenotypes. Indirect selection based on marker genotype rather than phenotype can be used to accelerate the speed and increase the precision of genetic progress, reduce the number of generations, and when integrated into optimized molecular breeding strategies, it can also lower the costs of selection. The efficiency of MAS depends on many factors associated with how the underlying marker-trait associations were identified, including the size of the mapping population, the nature of the phenotyping, the design and analysis of the experiment, the number of markers used, the distance between marker loci, the genomic region containing the desired QTL, and the proportion of additive genetic variance explained by the marker, the selection method, and the experimental design (Dwivedi *et al.*, 2005 and references therein). The efficiency of MAS also depends on many factors associated with its application, including the crop and breeding system, the molecular breeding process, and the nature of the genotyping pipeline. In this section, we briefly summarize the cases where MAS has been used to incorporate beneficial traits into improved genetic backgrounds of major food crops.

A. Resistance to Biotic Stresses

1. Single Gene Introgression

a. Cereals. MAS coupled with backcross and pedigree breeding methods and field evaluation has led reports in the literature of genetic enhancement for resistance to bacterial blight (BB) (*Xa21*), gall midge (*Gm-6t*), and brown plant hopper (BPH) (*Bph1* and *Bph2*) in rice; to leaf rust (*Lr19*, *Lr51*, and *Yr15*) in wheat; to yellow dwarf virus (*Yd2*), stripe rust (*Yr4*), and powdery mildew (*mlo-9*) in barley; and to downy mildew (major QTL) in pearl millet (Table XIII). The progenies showed same resistance level as the donor parental lines both in greenhouse and field evaluations.

Table XIII
Examples of Single Gene Transfer for Resistance to Biotic Stresses Using Marker-Assisted Selection in Barley, Common Bean, Maize, Pearl Millet, Potato, Rice, Soybean, and Wheat

Gene	Breeding scheme	Marker	Marker-assisted product	References
		Barley		
Barley yellow dwarf virus				
Yd2	Two backcrosses	YLM	Lines with *Yd2* had few leaf symptoms but no adverse effect on agronomic traits	Jefferies *et al.*, 2003
Powdery mildew [*Blumeria graminis* f. sp. *hordei* (Bgh.)]				
mlo9	Double-haploid breeding	SNPs	DHLs carrying *mlo9mlo9* completely resistant to powdery mildew	Paris *et al.*, 2003
Stripe rust (*P. striiformis* f. sp. *hordei*)				
Yr4	Double-haploids from BC_1F_1	RFLPs	DHLs carrying *Yr4* less susceptible to stripe rust	Toojinda *et al.*, 1998
		Common bean		
Comman bacterial blight (CBB) [*Xanthomonas campestris* pv. *phaseoli* (Xcp)]				
Quantitative	Pedigree breeding	$BC420_{900}$ and $C7_{900}$	Marker-based selected RILs resistant to CBB	Yu *et al.*, 2000
		Maize		
Southwestern corn borer (SWCB) (*Diatraea grandiosella* Dyar)				
6–9 QTL	Two backcrosses	89 RFLPs and a morphological marker, grain color (*y1*)	Progenies with improved resistance to SWCB leaf feeding damage selected	Willcox *et al.*, 2002
		Pearl millet		
Downey mildew (*Sclerospora graminicola*)				
Major gene	Backcross breeding	Xpsm464, Xpsm716, Xpsm265, and Xpsm416	HHB 67-2 with improved downy mildew resistance	Hash, 2005

Potato				
Late blight [*P. infestans* (Mont.) de Bary]				
RB	Two backcrosses	*RGA1/rga1*, *RGA2/rga2*, *RGA3/rga3*, and *RGA4/rga4*	Several marker-positive breeding lines showed resistance to late blight	Colton *et al.*, 2006
Rice				
Bacterial blight (BB) [*Xanthomonas oryzae* pv. *Oryzae* (*Xoo*)]				
Xa21	Three backcrosses	PCR-based markers close to *Xa21*, and 128 RFLPs	Lines with high yield and BB resistance selected	Chen *et al.*, 2000
Xa21	Three backcrosses	21, C189, and AB9 for foreground and AFLPs for background selections	6078(*Xa21*) performed well under heavy disease pressure	Chen *et al.*, 2001
Gall midge (*Orseolia oryzae*)				
Gm-6t	Pedigree breeding	RAPD and STS	*Gm-6t* successfully transferred to hybrid rice parents	Katiyar and Bennett, 2001
Soybean				
Soybean cyst nematode (*H. glycines* Ichinohe)				
Quantitative	Pedigree breeding	98 RFLPs	MAS-selected lines comparable to phenotypic selection	Concibido *et al.*, 1996
Wheat				
Leaf rust (*Puccinia triticina*)				
Lr19	Pedigree breeding	*Ep-D1c*	Families with *Ep-D1c* allele resistant to leaf rust	Slikova *et al.*, 2003
Lr51	Six backcrosses	XAga7 and Xmwg710	*Lr51* transferred into three cultivars	Helguera *et al.*, 2005
Yr15	Two backcrosses	1000 SSRs	*Yr15* transferred into Zak	http://wheatlifemagzine.com/0105/pg68_0105.pdf

b. Legumes. In contrast to the cereals, there are very few reports in the literature of success stories for single gene transfer by MAS in legumes, and only in two crops: soybean and common bean. However, this is proportional to the relative stage of development of genomics in these crops and the number of trait mapping studies that has been completed. Loci for resistance to common bacterial blight in common bean and cyst nematode in soybean have been transferred into improved breeding lines using MAS (Table XIII).

c. Roots and Tubers. Late blight is the most devastating disease in potato and has received much research attention across the world (Ojiambo *et al.*, 2000). However, resistance breeding has been a challenge because of the short period during which race-specific resistance genes remain effective, while breeding for "horizontal" or race-nonspecific resistance has achieved only moderate successes. *Solanum bulbocastanum* ($2n = 24$), a diploid species native to Mexico, has been characterized as possessing durable resistance to all known races of late blight (van Soest *et al.*, 1984), and mapped to a single locus on chromosome 8 (Naess *et al.*, 2000). Using PCR-based DNA markers for tracking the *RB* gene in breeding populations, several marker-positive selected lines showed resistance to late blight (Table XIII). *RB* has also been cloned and transformed into Katahdin, a highly susceptible potato cultivar. The Katahdin-transformed plants with *RB* showed broad-spectrum resistance against a wide range of late blight isolates (Lozoya-Saldana *et al.*, 2005; Song *et al.*, 2003). Clearly, by having the full sequence of the target gene, it should be possible to develop a highly efficient low-cost assay system for this trait.

2. Gene Pyramiding

Gene pyramiding is a useful approach to the durability or level of pest and disease resistances, or to increase the level of abiotic stress tolerance. Genes controlling resistance to different races or biotypes of a pest or pathogen and genes contributing to agronomic or seed quality traits can be pyramided together to maximize the benefit of MAS through simultaneous improvement of several traits in an improved genetic background.

a. Cereals. Many major genes (recessive or dominant) and QTL conferring resistance to pests and diseases have reported in major cereals. Using MAS coupled with field evaluation, researchers were able to combine multiple resistances to these pests and diseases in many cereal crops. Successful examples include improved pyramided lines and cultivars containing gene combinations for bacterial blight (BB) (*xa3*, *xa4*, *xa5*, *xa7*, *Xa10*, *xa13*, *Xa21*, and *Om*); blast (Bl) (*Pi1*, *Piz-5*, and *Pita*); brown plant hopper (BPH) (*Bph1* and *Bph2*);

Bl (*Piz-5*) and BB (*Xa21*); BB (*Xa21*) and yellow stem borer (YSB) (*Bt*); BB (*Xa21*), YSB (*Bt*), and sheath blight (ShB) (*RC7 chitinase*); and BB (*Xa21* and *Xa7*), YSB (*Bt*), Bl (*Pi1*, *Pi2*, and *Pi3*), and BPH (*Qbph1* and *QBph2*) in rice (Table XIV). In wheat, powdery mildew (*Pm2*, *Pm4a*, *Pm6*, *Pm8*, and *Pm21*) pyramided lines and those with resistance to *Fusarium* head blight (FHB) (six QTL), orange blossom midge (*Sm1*), and leaf rust (*Lr21*) were bred through MAS. Resistance to *Barley mild mosaic virus* (BaMMV) and *Barley yellow mosaic virus* (BaYMV and BaYMV-2) complex (*rym4*, *rym5*, *rym9*, and *rym11*) and stripe rust (QTL: 1H, 4H, and 5H or their combination: 1H and 4H, 1H and 5H, 4H and 5H, or 1H, 4H, and 5H) has been separately incorporated through MAS in barley. Many of these pyramided lines showed enhanced resistance to pests and diseases, some even outyielded the controls under high disease or pest pressure in field conditions (Table XIV).

b. Legumes. Reports of gene pyramiding in legumes include combining QTL for resistance to corn earworm and *Pseudoplusia includens* (soybean looper) with *cry1Ac* resistance in soybean; while resistances to rust and anthracnose (QTL) or to CBB, *Bean common mosaic virus* (BCMV), and anthracnose have been combined in common bean (Table XIV). The pyramided lines in soybean showed improved resistance to defoliators, while common bean lines showed multiple resistances to these diseases.

c. Roots and Tubers. A single dominant gene for extreme resistance to *Potato virus Y* (PVY, genus *Potyvirus*), Ry_{adg}, was mapped to a distal position on potato chromosome 11 (Hämäläinen *et al.*, 1997). For *Potato virus X* (PVX, genus *Potexvirus*), dominant genes, *Rx1* and *Rx2*, were mapped to potato chromosomes 12 and 5, respectively (Ritter *et al.*, 1991). The dominant gene *Gro1* for resistance to all known pathotypes of the root cyst nematode (*Globodera rostochiensis*) was mapped to potato chromosome 7 (Barone *et al.*, 1990). A single dominant gene *Sen1* for resistance to potato wart (*Synchytrium endobioticum*) pathotype 1 was mapped to a similar position on potato chromosome 11 as the Ry_{adg} (Hehl *et al.*, 1999). Using four PCR-based diagnostic assays, tetraploid progeny from tetraploid–diploid crosses combining the Ry_{adg} for extreme resistance to PVY with *Gro1* for nematode resistance and with *Rx1* for extreme resistance to PVX, or with *Sen1* for wart resistance were selected (Table XIV).

B. Tolerance to Abiotic Stresses

1. Drought Tolerance

Rice: selection for a well-developed root system with long thick roots should improve the drought tolerance of upland rice because the plant would avoid water stress by absorbing water stored in the deep soil layers

Table XIV
Examples of Gene Pyramiding for Resistance to Biotic Stresses Using MAS in Barley, Common Bean, Potato, Rice, Soybean, and Wheat

Gene	Breeding scheme	Marker	Marker-assisted product	References
		Barley		
BaYMV-I, BaYMV-II, and BaYMV-III; BaMMV-Ka1 and Na1				
rym1	One backcross	RFLPs	Mokkei 01530 with *rym1* resistant to BaYMV-1 and BaYMV-II, and similar in malt quality as of Haruna Nijo	Okada *et al.*, 2003b
rym4, *rym5*, *rym9*, and *rym11*	Simple and complex crosses using double-haploids	RAPDs and SSRs	DHLs carrying *rym4*, *rym9*, and *rym11* and those with *rym5*, *rym9*, and *rym11* selected	Werner *et al.*, 2005
Barley stripe rust				
QTL (1H, 4H, and 5H)	Backcross-derived ILs	SSRs	ILs in susceptible genetic background carrying 1H, 4H, or 5H individually or in combinations were resistant to barley stripe rust	Richardson *et al.*, 2006
		Common bean		
Common bacterial blight (CBB), BCMV, and anthracnose				
Several loci for BCMV and anthracnose	Complex crossing and pedigree breeding	UBC420, BC73, SW13-*I*, and *Co-4*^{2}	Marker-based selected progenies resistant to CBB, BCM, and anthracnose	http://www.Ontariobeans.on.ca/liu5thcapsulemsapaperfinal.pdf
Rust (*Uromyces appendiculatus*) and anthracnose (*Colletotrichum lindemuthianum*)				
Nine major genes each for rust and anthracnose	Three backcrosses	RAPDs	Lines combining resistance to rust and anthracnose developed	Faleiro *et al.*, 2004

		Potato		
***Potato virus* Y (PVY), *Potato virus* X (PVX), nematode, and wart (*S. endobioticum*)**				
Ry_{adg}(PVY), *Rx1* (PVX), *Gro1* (nematode), and *Sen1* (wart)	F1 hybrids (2 × 4 cross)	RYSC3 (Ry_{adg}), Gro1–4 (*Gro1*), CP60 (*Rx1*), and N125 (*Sen1*)	Marker-based selection of tetraploid potato clones showed multiple resistance to four diseases, all with monogenic resistance	Gebhardt *et al.*, 2006
		Rice		
Bacterial blight (BB)				
Xa3, *Xa4*, *xa5*, and *Xa10*	Pedigree breeding	RZ390, RG556, RG207, XNpb181, and $Oo7_{2000}$	Lines carrying multiple genes provided broader spectra of resistance to BB	Yoshimura *et al.*, 1995
Xa4, *xa5*, *xa13*, and *Xa21*	Pedigree breeding	Npb181, Npb78, RG103, RG136, RG556, RZ28, RZ207, pTA248, and pTA818	Pyramided lines showed broader spectrum of resistance to BB	Huang *et al.*, 1997
xa5, *xa13*, and *Xa21*	Three backcrosses	RG556, RG207, RG136, and pTA248	Lines with *Xa21* had increased resistance than *xa5*, *xa13*, or both	Sanchez *et al.*, 2000
	Two backcrosses	RG556, RG136, and pTA248	Lines with gene combinations provided broader spectrum of resistance to BB	Singh *et al.*, 2001a
xa5, *xa13*, and *Xa21*	Pedigree breeding	pTA248, RG136, and RM122	Lines carrying multiple genes showed greater resistance than those with single gene(s)	Swamy *et al.*, 2004
xa5, *xa7*, *Xa21*, and *Om*	Three backcrosses	RG556a (*xa5*), OPL13 (*Om*), pTAta258 (*Xa21*), and 10 RAPD markers	Angke (*xa5*) and Conder (*xa7*) released, and few other lines combining yield and resistance in advance trials in Indonesia	http://www.isuagcenter.com/inst/research/stations/rice/proceedings.pdf

(*continued*)

Table XIV (*continued*)

Gene	Breeding scheme	Marker	Marker-assisted product	References
BB, leaf folders, yellow stem borer (YSB) (*Scirpophaga incertulas*)				
Xa21 and *Bt*	Pedigree breeding	21, 248, C189, AB9 for *Xa21* and pFHBT1(1.8 kb) for *Bt*	Minghui63 containing *Bt* and *Xa21* and its hybrids showed multiple resistance and produced two to three times more grain yield under natural infestation	Jiang *et al*., 2004
BB, stem borer (SB), blast, and BPH				
Xa21 and *Xa7* (BB); *Bt* (SB); *Pi1*, *Pi2*, *Pi3* (blast); and *Qbph1* and *Qbph2* (BPH)	Pedigree breeding	AFLP 1415, STS P3, M5, 248, RM144, RM224, and *Pi2*	Minghui 63(*Xa21* and *Xa7*) showed broader resistance to BB; Minghui 63(*Xa21* and *Bt*) showed combined resistance to BB and SB; Zhenshan97(*Qbph1* and Qbph2) showed better resistance to BPH	Yuqing *et al.*, 2004
BB, YSB, sheath blight (ShB) (*R. solani*)				
Xa21, *Bt*, and *RC7 chitinase* (Shb)	Pedigree breeding	Pc822 (*Xa21*), *Bt*, and *RC7 chitinase*	Lines carrying three genes were resistant to BB, YSB, and ShB	Datta *et al.*, 2002
Blast (Bl) [*Magnaporthae grisea* (Herbert) Borr. (ananmorphe *Pyricularia oryza* Cav.)				
Pi1, *Piz-5*, and *Pita*	Pedigree breeding	Npb181, RZ536, RZ64, RZ612, RG456, RG64-SAP, RG869, RZ397, and RG241	The pyramided lines showed better resistance to blast	Hittalmani *et al*., 2000
Bl and BB				
Piz-5 and *Xa21*	Four backcrosses (*Piz-5*) and transgenic (*Xa21*)	$RG64_{750}$ (*Piz-5*) and 1.4-kb fragment (*Xa21*)	Lines showed combined resistance to Bl and BB	Narayanan *et al*., 2002
*Piz-1*and *Piz-5* (blast) and *Xa21* (BB)	Pedigree breeding	RZ536 and r10 (blast) and *Xa21* (1.4-kb fragment of pC822)	The pyramids showed enhanced resistance to blast and BB	Narayanan *et al.*, 2004

Brown plant hopper (BPH) (*Nilaparvata lugens* Stal) (*Bph1* and *Bph2*)				
Several major genes and QTL	Pedigree breeding	em24G, EM5814N, em32G, KPM1, KPM2, KPM3, KPM4, KPM5, and KPM8	Pyramided lines showed similar resistance as to those with single gene	Sharma *et al.*, 2004
***Rice yellow mottle virus* (RYMV)**				
Many QTL	Three backcrosses	RG869 and BNL 16–06 for foreground and RFLPs and SSRs for background selections	Lines containing QTL 12 and QTL 7 alleles showed partial resistance to RYMV	Ahmadi *et al.*, 2001
		Soybean		
Corn earworm (CEW) (*Helicoverpa zea* Boddie)				
QTL and Bt (*cry1Ac*)	Three backcrosses	Nine SSRs	The pyramid lines had a detrimental effect on larval weights and on defoliation by CEB	Walker *et al.*, 2002
CEW and soybean looper (SBL) (*P. includens*)				
cry1Ac and QTL (PI 229358)	Two backcrosses	Six SSRs and sequence-specific primers *cry1Ac*	Lines carrying *cry1Ac* and QTL alleles resistant to three lepidopteran pests	Walker *et al.*, 2004
		Wheat		
***Fusarium* head blight (FHB) (*F. graminearum*), orange blossom midge (*Sitodiplosis mosellana*), and leaf rust (*Lr21*)**				
Six FHB QTL, *Sm1* for midge and *Lr21* for leaf rust	Two backcrosses	gwm533, gwm493, and wmc808	Resistant progenies containing chromosome segments FHB, *Sm1* and *Lr21* identified	Somers *et al.*, 2005
Powdery mildew (*E. graminis* DC. *F. tritici* Em. Marchal)				
Pm2, *Pm4a*, and *Pm21*	Pedigree breeding	Xbcd1871-5D-*Eco*RV, Xwhs350-5D-*Eco*RV, Xbcd1231-2A-*Eco*RI, pHv62, and psr113	Gene combinations (*Pm2* + *Pm4a*, *Pm2* + *Pm21*, and *Pm4a* + *Pm21*) integrated into Yang158 that showed resistance to powdery mildew	Liu *et al.*, 2000b
Pm2, *Pm4a*, *Pm6*, *Pm8*, and *Pm21*	Pedigree breeding	RAPD and SCAR markers	Lines with *Pm2* and *Pm4a* immune to powdery mildew	Wang *et al.*, 2001a

(Yoshida and Hasegawa, 1982). However, phenotypic selection for root morphological traits in conventional breeding is not feasible. The tropical *japonica* rice cultivars are reported to have thicker and deeper roots than *indica* cultivars (Courtois *et al.*, 1996). Using four QTL (QTL2, QTL7, QTL9, and QTL11) from Azucena (a *japonica* cultivar), which each contributing between 5% and 30% phenotypic variance for root traits (root length and thickness), Steele *et al.* (2006) initiated marker-assisted backcrossing (MABC) to improve drought tolerance into Kalinga III, an upland *indica* cultivar. After five backcrosses and conducting over 3000 marker assays (2548 RFLPs and 700 SSRs) on 323 plants, the NILs were developed and evaluated for root traits. The target segment on chromosome 9 (RM242-RM201) significantly increased root length under both irrigated and drought stress environments. Azucena alleles at the locus RM248 (below the target root QTL on chromosome 7) delayed flowering. However, selection for the recurrent parent allele at this locus produced early flowering NILs that are suited to upland environments in eastern India. Other target regions had no significant effects on root length in Kalinga III genetic background. In a similar study, Shen *et al.* (2001) also demonstrated the effectiveness of MAS to transfer QTL from three of the four target regions (chromosomes 1, 2, 7, and 9) associated with root traits (root length and root mass) from Azucena to NIL in IR64 genetic background. NIL carrying the QTL from chromosomes 1, 7, and 9 had shown significantly improved root traits over IR64, while none of the NIL containing QTL from chromosome 2 had root phenotype significantly different from that of IR64. In both the studies, progenies containing QTL from chromosome 7 confer improved root characteristics that are now being tested under field conditions to assess their performance under water-limited conditions.

Maize: Anthesis silking interval (ASI) is an important trait associated with drought tolerance in maize. Ribaut *et al.* (1996, 1997b) initiated a major marker-assisted breeding program to transfer five genomic regions involved in the expression of a short ASI from Ac7643 (a drought-tolerant line) to CML247 (an elite tropical breeding line). Five genomic regions were transferred using flanking PCR-based markers. Seventy of the best BC_2F_3 (i.e., S_2 lines) lines were crossed with two testers, CML254 and CML 274. These hybrids and the BC_2F_4 families derived from selected BC_2F_3 plants were evaluated for 3 years under drought stress conditions. Results show that stress conditions induced a yield reduction of at least 80%, but the mean of the 70 selected genotypes performed better than the control (all evaluated as testcross products). In addition, the best genotypes among 70 selected (BC_2F_3 $\times$ testers) performed two to four times better than the control. However, this difference became less marked when the intensity of stress decreased: for a stress inducing less than 40% yield reduction, performance of testcross hybrids resulting from MAS was no better than the "original" version of CML274.

Pearl millet: a major QTL on LG2 is associated with increased grain yield and harvest index under terminal stress in PRLT 2/89-33 (Yadav *et al.*, 2002). PRLT 2/89-33 is a drought-tolerant, low-tillering, and large-panicle landrace from West Africa (Andrews and Anand Kumar, 1996). In contrast, H77/833-2 is a drought-sensitive, high-tillering, and small-panicle landrace from India (Kapoor *et al.*, 1989). The performance of QTL MAS-derived topcross hybrids (TCH) was compared with that of field-based TCH. Progenies with the best overall ability to maintain under terminal stress environments were used to generate the TCH, and these were compared with randomly mated TCH made from randomly selected progenies from the entire population (irrespective of performance under terminal drought stress). In both the cases, progenies were selected irrespecitve of the presence or absence of favorable alleles at the putative drought-tolerant QTL and evaluated across 21 environments (nonstress, terminal stress, and gradient stress). The QTL MAS-derived hybrids were significantly, but only modestly, higher yielding both in full and partial terminal stress environments. However, this advantage under stress was at the cost of lower yield of the same hybrids under non-stressed environments. The QTL MAS-derived hybrids flowered earlier and had limited effective basal tillers, low biomass, and high harvest index. All these traits are similar to that of the drought-tolerant parent PRLT-2/89-33, thus confirming the effectiveness of the putative drought-tolerant QTL on LG2 (Bidinger *et al.*, 2005). A number of marker-assisted backcross progenies have been generated from the cross between H77/833-2 (drought sensitive) and PRLT 2/89-33 (LG2 drought-tolerant QTL). Initial results indicate that it has been possible to improve grain yield under terminal stress in these lines without a biomass penalty under stress conditions or a grain yield penalty under well-watered conditions (Hash *et al.*, 2004).

Common bean: Schneider *et al.* (1997) identified four to five RAPD markers in two mapping populations that were consistently and significantly associated with yield under stress, yield under optimum irrigation, and geometric mean yield across a broad range of environments. To examine the effectiveness of these markers, they selected genotypes from either extremes and evaluated them in three locations. MAS in the Sierra/AC1028 population was effective in Michigan under severe stress but ineffective in Mexico under moderate stress. The Sierra/Lef-2RB population showed improved performance by 11% in stress and 8% in nonstress environments.

2. Submergence Tolerance

In many parts of the lowlands of south, southeast, and eastern Asia, rice cropping during the rainy season is completely submerged for varying periods of time, resulting in substantial losses to rice production in these regions.

Genetic variation for submergence tolerance has been reported in rice, for example, FR13A, a landrace from India, can survive up to 2 weeks of complete submergence owing to a major QTL, *submergence 1* (*Sub1*) on chromosome 9 (Xu and Mackill, 1996; Xu *et al.*, 2000). Further, Xu *et al.* (2006) identified a cluster of three genes related to the ethylene-response-factor (ERF) at the *Sub1* locus. A variant of *Sub1A-1* is found only in submergence-tolerant rice, FR13A. Overexpression of *Sub1A-1* in submergence-intolerant *O. sativa* ssp. *japonica* (cultivar Liaogeng) conferred enhanced tolerance. The same research group used marker-assisted backcross breeding to introgress the *Sub1A-1* gene into a widely grown Indian cultivar, Swarna. The introgressed progenies showed strong submergence tolerance and maintained high yield and other agronomic properties of the recurrent parent, Swarna. Submergence tolerance has also been introduced into a Thai Jasmine rice, KDML105 following marker-assisted breeding (Siangliw *et al.*, 2003).

C. Agronomic and Seed Quality Traits

Many agronomic or seed quality traits are conferred by QTL each with varying contributions and different interaction with each other (epistasis) and the environment thus greatly complicating cultivar development. Unlike many success stories of pests and disease-resistance transfer by MAS in many crops, there are few reports of successful transfer of beneficial alleles associated with improved yield or seed quality traits into improved genetic background. The foremost among them include yield-enhancing QTL alleles from wild relatives of rice and soybean and grain quality in rice, wheat, and maize, and malt quality in barley.

Rice: Using marker-assisted backcross breeding, the two yield-enhancing QTL alleles, *yld1.1* and *yld2.1* from wild rice *Oryza rufipogon*, have been successfully transferred into an improved agronomic background, whose progenies out-yield the controls by 24–42%. Most of this improvement was accounted for by increases in two yield components: grains per panicle and 1000-grain weight (Liang *et al.*, 2004). In another marker-assisted backcross breeding program, Yue-guang *et al.* (2004) selected progenies in BC_3 generation that produced more than 30% greater grain yield over Minghui 63, a restorer line of the many commercially grown hybrids in China.

Grain quality represents a major problem, particularly in hybrid rice which are now commercially grown in substantial acreage worldwide. The most serious grain quality problems in hybrid rice are eating and cooking qualities, and to some extent milling quality. Both eating and cooking qualities are largely determined by three characters, specific to the physical and chemical properties of the starch in the endosperm, that is,

amylose content (AC) (Juliano, 1985; Webb, 1980), gel consistency (GC) (Cagampang *et al.*, 1973), and gelatinization temperature (GT) (Little *et al.*, 1958). The chalkiness, or opacity, of the endosperm of the grains is another important grain quality trait that not only affects the appearance of the grains but also the resistance to grain breakage during milling. Medium AC/soft GC/high GT together with a translucent endosperm represent good grain quality, while high AC/hard GC/low GT together with chalky endosperm represent poor grain quality (Tan *et al.*, 1999, 2000).

Shanyou 63, a hybrid between the male-sterile line Zhenshan 97A and the restorer line Minghui 63, was the most widely grown hybrid rice in the 1990s, accounting for ~25% of the rice production in China (Lin and Min, 1991). However, in recent years, the area declined as this hybrid became susceptible to bacterial blight and because of greater consumer awareness about its relatively poor cooking and eating qualities. AC, GC, and GT cosegregate and are controlled by the *waxy* locus and other genes tightly linked to this locus (Tan *et al.*, 1999). It should be, therefore, possible to simultaneously improve all three traits. Chalkiness, or opacity, of the grains is controlled by 6 QTL located on 5 of the 12 rice chromosomes (Tan *et al.*, 2000). Using MAS in three generations of backcrossing followed by one generation of selfing, Zhou *et al.* (2003a) successfully introduced the wx-MH fragment from Minghui 63 into Zhenshan 97B, which was subsequently transferred to Zhenshan 97A. The improved version of the male-sterile and maintainer lines, Zhenshan 97A (wx-MH) and Zhenshan 97B (wx-MH), contained a fragment less than 6.1 cM in length around the *waxy* gene region from the donor parent, with the rest of the genome being from the original Zhenshan 97. The introduction of this fragment has greatly improved the cooking and eating quality of inbred lines and their resultant hybrids, with the agronomic performance essentially the same as the original maintainer line and resultant hybrid. Additionally, the selected lines and their hybrids showed reduction in opacity (a change that is highly preferred from consumer's view point) and grain weight. However, the hybrids yielded at a similar level to the original hybrid (Shanyou 63), presumably because of phenotypic plasticity as a result of strong heterosis (Zhang *et al.*, 1994). Long-te-fu (LTF) and Zhan-shan 97 (ZS) are the two key female parents widely used for the generation of *indica* hybrid rice in China. However, both have poor cooking and eating qualities because of high AC. Liu *et al.* (2006) used MAS to introgress *Wx-T* allele (conferring intermediate AC and thus good quality) into the maintainer (LTF-B and ZS-B) and their relevant male-sterile lines (LTF-A and ZS-A) to generate improved *indica* hybrids. The resulting maintainer lines (LTF(tt)-B and ZS(tt)-B) and hybrids showed improved cooking and eating qualities with no significant alterations in their agronomic traits.

Rice with low glutelin content is suitable for patients affected by diabetes and kidney failure. The *Lgc-1* locus confers low glutelin in the rice grain,

located on chromosome 2 between flanking markers (Miyahara, 1999). This trait has been successfully incorporated into *japonica* rice with 93–97% selection efficiency using SSR2-004 and RM358 markers (Wang *et al.*, 2005a). Additionally, grain quality traits such as 1000-seed weight, kernel length/breadth ratio, basmati type aroma, and high AC have been combined with resistance to bacterial blight using marker-assisted backcross breeding (Joseph *et al.*, 2003; Ramalingam *et al.*, 2002).

Wheat: the major grain quality traits in wheat are protein content and composition and grain color that influence bread- and noodle-making qualities. Gliadins and glutenins determine physical quality of wheat flour dough (Payne, 1987). Dough with high elasticity and reasonable extensibility is ideal for bread making, while highly extensible dough is good for making biscuits, and dough with intermediate properties is good for flat bread or noodles. Most of these quality traits are genetically highly complex, conferred by many genes showing considerable GEI. Moreover, evaluation of these traits requires well-developed laboratory procedures and equipments and a large sample size for evaluation. These factors force most wheat breeders to only evaluate quality traits in advanced generations of their breeding programs. Thus, it is surprising that although for many of these traits markers have been identified and validated (see Section III), their use in breeding has been limited. Exploiting allelic variation at the Glu-1 (endosperm storage protein subunit) locus to improve bread-making quality has been one of the early examples in which markers were used to improve wheat quality traits (de Bustos *et al.*, 2001; Koebner, 2003).

Sun *et al.* (2005) used a novel STS marker for improving polyphenol oxidase (PPO) activity in bread wheat. Breeding wheat cultivars with low PPO activity is the best approach to reduce undesirable darkening of bread wheat-based end-products, particularly for Asian noodles. Based on the sequences of genes conditioning PPO activity during kernel development, 28 pairs of primers were developed. One of these markers designated as *PPO18*, mapped to chromosome 2AL, can amplify a 685 and an 876-bp fragment in the cultivars with high- and low-PPO activity, respectively. QTL analysis indicated that the PPO gene cosegregated with the STS marker *PPO18* and is closely linked to *Xgwm312* and *Xgwm294* on chromosome 2AL, explaining 28–43% of phenotypic variance for PPO activity across three environments. A total of 233 Chinese wheat cultivars and advanced lines were used to validate the correlation between the polymorphic fragments of *PPO18* and grain PPO activity. The results showed that *PPO18* is a codominant, efficient, and reliable molecular marker for PPO activity and can be used in wheat breeding programs targeting noodle quality improvement.

Maize: maize plays a very important role in human and animal nutrition. The endosperm of the maize seed has several distinct regions that have different physical properties. The aleurone is the outer layer of the

endosperm, composed of specialized cells that secrete hydrolytic enzymes during germination. Beneath the aleurone are starchy endosperm cells filled with starch and storage proteins, thus creating two distinct regions—the ''vitreous'' or glassy endosperm and the ''starchy'' endosperm. The vitreous endosperm transmits light, whereas the starchy endosperm does not. Typically, the endosperm is ~90% starch and 10% protein (Gibbon and Larkins, 2005). Normal maize protein is deficient in two essential amino acids (lysine and tryptophan) and has a high leucine:isoleucine ratio and biological value (Babu *et al.*, 2004). A naturally occurring recessive mutant gene *opaque-2*, observed first in a Peruvian maize landrace, gives a chalky appearance to the kernels and has improved protein quality due to increased levels of lysine and tryptophan in the endosperm (Mertz *et al.*, 1964). However, this trait appears to be associated with inferior agronomic traits such as brittleness and increased susceptibility to insect pests. With the discovery of ''modifier genes'' (*mo2*) that alter the soft, starchy texture of the endosperm, maize breeders developed hard endosperm *o2* mutants designated as ''quality protein maize'' (QPM) (Nelson, 2001; Prasanna *et al.*, 2001), which have the phenotypes and yield potential of normal maize but maintain the increased lysine content of *o2*. *Opaque 2* is a recessive trait but due to the effect of the modifiers, QPM behaves as a quantitative trait. Using SSRs and backcross breeding, Babu *et al.* (2004) developed maize lines that had twice the amount of lysine and tryptophan as compared to local cultivars and recovered up to 95% of the recurrent parent genome.

Sweet corn is another class of edible-grade maize, which is highly preferred as roasted/or boiled cobs. In sweet corn, breeding for improved seedling emergence and eating quality is complicated because of the inverse relationship between these traits. High kernel sugar content is one of the reasons for poor seedling emergence (Douglass *et al.*, 1993), influenced by many kernel characteristics that are under the control of many genes (Azanza *et al.*, 1996a,b). Evaluation of these traits requires difficult and expensive characterization in the laboratory. However, using marker-assisted backcross or population breeding, it has been possible to select progenies with improved seedling emergence that also has high sucrose content (Yousef and Juvik, 2001a, 2002).

Barley: malt is a major raw material for the production of beer. Characters that affect malting quality include malt extract content, α- and β-amylase activity, diastatic power, malt β-glucan content, malt β-glucanase activity, grain protein content, kernel plumpness, and dormancy, all are quantitatively inherited variously influenced by the environment (Zale *et al.*, 2000). There are few barley cultivars with good malt quality that brewers are reluctant to change from due to their concerns about the resultant changes in flavor and brewing procedures. For example, the goal of US Pacific Northwest barley breeding program is to produce high-yielding NILs that maintain traditional

malting quality characteristics but transfer QTL associated with yield, via marker-assisted backcrossing, from the high-yielding cultivar Baronesse to the North American two-row malting barley industry standard cultivar Harrington. Schmierer *et al.* (2004) targeted Baronesse chromosome 2HL and 3HL fragments presumed to contain QTL that affect yield. Using backcross breeding and QTL/marker information, they identified a NIL (00–170) that when evaluated for yield over 22 environments and for malt quality over 6 environments produced yield equal to Baronesse while maintaining a Harrington-like malt quality profile. Other studies have also reported the development of lines with improved malt quality: white aleurone color and high α-amylase content (Ayoub *et al.*, 2003), and high in β-glucan and fine-coarse difference (Igartua *et al.*, 2000).

Soybean: Concibido *et al.* (2003) introgressed yield-enhancing QTL from exotic soybean germplasm *Glycine soja* (PI 407305). They detected yield-enhancing QTL located on LG B2 (U26). In a 2-year multilocation trial, individuals carrying the PI407305 haplotype at the QTL locus demonstrated 8–9% yield advantage over individuals that did not contain the exotic haplotype. When assessing the QTL effect in various elite genetic backgrounds, they found that this QTL conferred enhanced yield in only two of the six genetic backgrounds, although individuals carrying the PI407305 haplotype at the QTL locus always had an average 9% yield advantage in yield trials across locations.

Common bean: Tar'an *et al.* (2003) used an index based on QTL-linked markers and ultrametric genetic distances between progeny lines and a target parent to select for increased yield in their breeding program. Lines with a combination of phenotypic performance and high QTL-based index produced greater yield over those developed by using high QTL-based index, conventional phenotypic selection, and a low QTL-based index. They also demonstrated that the use of the QTL-based index in conjunction with the ultrametric genetic distance to the target parent would enable a plant breeder to select lines that retain important QTL in a desirable genetic background.

Pea: Resistance to lodging, a key objective in many pea breeding programs, is controlled by two genes that markers A001 (in coupling phase) and A004 (in repulsion phase) are associated with resistance to lodging (Warkentin *et al.*, 2004). Zhang *et al.* (2006a) evaluated the effectiveness of these markers in F_2 population of eight crosses. The lowest lodging score for each population was obtained from plants with the combination of A001 (presence) and A004 (absence). They detected a higher proportion of lodging resistant F_3 families from this marker combination as compared with phenotypic selection in F_3 generation. Thus, A001 and A004 are useful for MAS for lodging resistance in early generation pea breeding populations.

The preceding examples demonstrate that marker-assisted breeding is a viable option to supplement conventional breeding programs for certain

traits and where robust markers are available. To date, MAS has been frequently used to transfer simply inherited traits or to pyramiding genes with major effects but much less for improving polygenic traits. However, a good knowledge of the trait genetics, interaction effects (epistasis, genetic background, and environment), population size limitations, accurate phenotyping, user-friendly PCR-based marker assays, marker-trait association, and genetic recombination (closer the distance between marker and the gene/QTL, lesser the chance of recombination and loss of selective power), and the ability to timely manage and interpret the voluminous marker data largely influence our ability to successfully integrate MAS into crop breeding programs. In addition, many breeders still consider the use of marker technology as prohibitorily expensive for routine use in breeding programs. However, it is encouraging to note that high-throughput genotyping platforms for large-scale, low-cost applications are rapidly advancing, largely driven by the human diagnostics community. In turn, this is encouraging the development of a genotyping service industry, thus disconnecting breeding programs from the need to establish and maintain capital-intensive in-house facilities, although many of these companies struggle to provide a speed of service in-line with the often very short breeders' decision window. Hence, the cost for MAS genotyping will become more affordable to breeding programs but probably only for those who can embrace SNP markers.

D. Specific Challenges for Alien Gene Introgression

Wild crop relatives are traditionally looked on as potential sources of gene(s) for resistance to many pests and diseases that are not available in cultigens, thus making them a valuable resource for gene transfer in cultivated species. Both conventional crossing and selection, and molecular breeding (MAS and transgenics) have been used to transfer pest and disease resistances from wild relatives to cultivated crop species (Dwivedi *et al.*, 2007 and references therein). Resistance gene(s) from wild relatives have facilitated large-scale cultivation of crops in disease or pest endemic regions of the world, that is, bacterial blight (BB) and grassy stunt virus in rice, BB in maize and potato, and nematodes in many crops. Wild relatives are usually inferior to modern cultivars with respect to yield and seed quality. However, the successful transfer of improved fruit yield and processing quality in tomato (Bernacchi *et al.*, 1998a,b; de Vicente and Tanksley, 1993; Fridman *et al.*, 2000; Fulton *et al.*, 1997; Rick, 1974; Yousef and Juvik, 2001b) led to the realization that wild relatives can contain beneficial genes (in addition to resistance to biotic stresses) associated with yield and seed quality, although these are often phenotypically masked by deleterious genes and are thus difficult to identify and transfer through conventional selection and breeding.

Using advanced backcross and QTL analysis (Tanksley and Nelson, 1996), yield and grain quality enhancing alleles from wild relatives have been successfully introgressed in rice, wheat, barley, sorghum, common bean, and soybean (Dwivedi *et al.*, 2007 and references therein). Dramatic yield advantages have been reported in rice, for example, through the introduction of two yield-enhancing QTL alleles (*yld1.1* and *yld2.1*) from *O. rufipogon* (AA genome) into 9311 (one of the top performing parental lines used in the production of super hybrid rice in China) contributed in excess of 20% yield increases in rice; that is, about 1 t ha^{-1} gain in yield in some of the newly bred cultivars, largely because of increases in panicle length, panicles per plant, grains per plant, and grain weight. These improved lines with 9311-type genetic backgrounds are being used to raise the existing yield potential of super hybrid rice in China (Liang *et al.*, 2004). *Oryza grandiglumis* (allotetraploid, CCDD genome species) is another wild relative contributing positive alleles for increased grain yield in rice. In contrast, only 6–8% increase in grain yield was reported when positive alleles from *Hordeum spontaneum* were introgressed into barley. Wild relatives also contributed positive alleles for improved grain characteristics in rice (long, slender, and translucent grains, and grain weight), wheat (grain weight and hardness), and barley (grain weight, protein content, and some malt quality traits). Of particular interest is a locus for grain weight, *tgw2*, which contributed positive alleles from *O. grandiglumis* that are independent from undesirable effects of height and maturity (Yoon *et al.*, 2006). In a similar study, Ishimaru (2003) identified a grain weight QTL, *tgw6*, responsible for increased yield potential without any adverse effects on plant type, or grain quality in the Nipponbare genetic background. Similarly, alleles from *G. soja* conveyed 8–9% increased in grain yield and improved the protein content in soybean (Concibido *et al.*, 2003).

Development of exotic genetic libraries (also known as CSSL, IL, or CL) is another approach to enhance utilization of wild relatives to expand crop gene pools (see Section II.A). These genetic stocks provide a well-characterized potential resource for uplifting the yield barriers through pyramiding beneficial loci and fixing of positive heterosis. For example, when tomato ILs carrying three independent yield-promoting genomic regions were pyramided, the progenies produced more than 50% greater yield compared to controls (Gur and Zamir, 2004). In a report (Yoon *et al.*, 2006), several rice lines outperformed Hwaseongbyeo ($\sim$1 t ha^{-1} increase in grain yield). Several grain characteristics, including grain weight, were improved after crossing an advanced IL containing *O. grandiglumis* segments, HG101 (very similar to Hwaseongbyeo) with Hwaseongbyeo. The above examples demonstrate that wild relatives contain desirable alleles for agronomic traits, even though their effect is phenotypically not evident in wild relatives. It is important that more emphasis should be given to exploit

wild relatives to identify yield enhancing alleles to further raise the yield potential of crop cultivars. This is now an achievable goal as we progress toward saturating the genetic linkage maps of many crops with user-friendly markers, and the technological cost of applying marker technology is substantially reduced.

V. SUCCESSFUL APPLICATION OF MARKER-ASSISTED GENETIC ENHANCEMENT IN PRIVATE SECTOR BREEDING PROGRAMS

During the 1990s, MAS was often presented as holding the potential to replace phenotypic selection and dramatically reduce the time required to breed new cultivars (Mazur, 1995). Multinational seed companies have made large investments in genomic technologies and are now routinely using applied genomic tools to (1) dissect the genetic structure of the germplasm to understand gene pools and germplasm (heterotic) groups, (2) provide insights into allelic content of potential germplasm for use in breeding, (3) screen early generation breeding populations in order to select segregants with desired combinations of marker alleles associated with beneficial traits (especially where this avoids the costly phenotypic evaluations), (4) for accelerating the introgression and backcrossing of transgenes into diverse elite breeding lines, and (5) establish genetic identity (through DNA fingerprinting) of their products (Cooper *et al.*, 2004; Crosbie *et al.*, 2006; Fu and Dooner, 2002; Niebur *et al.*, 2004).

MAS has been successfully applied in cultivar development for maize (Crosbie *et al.*, 2006; Eathington, 2005; Johnson, 2004; Niebur *et al.*, 2004). Private sector soybean breeders have also made extensive use of MAS to select for resistance to soybean cyst nematode (SCN, *Heterodera glycines*), phytophthora root rot (*Phytophthora sojae*), and brown stem rot (*Phialophora gregata*). Using MAS breeders have been able to fix these resistance traits in their breeding materials before proceeding to yield trials (Cahill and Schmidt, 2004; Cregan *et al.*, 1999; Crosbie *et al.*, 2006). It is reported that MAS has allowed Pioneer to double their rate of genetic improvement for yield among SCN-resistant cultivars (https://www.pioneer.com/pioneer_news/press_releases/products/marker_assisted_selection). More recently, Monsanto breeders used MAS in the development of soybean cultivar Vistive that has low levels of linolenic fatty acid, thus reducing the need for postharvest processing to lower or eliminate the presence of unhealthy *trans* fats from foods. Vistive soybeans meet processor's growing demand for low-linolenic oils, which attract premiums for growers. Other upcoming products from Monsanto are Vistive mid-oleic (increase shelf life and flavor), Vistive low

saturates (combining lower saturated fats, lower *trans* fats, and improved stability), and Vistive omega-3 (providing consumers new options for omega-rich foods) products (http://www.monsanto.com/monsanto/layout/products/seeds_genomics/oilseeds.asp). Despite these successes, many private sector breeding programs still rely heavily or solely on phenotypic selection and most agree that MAS will never entirely replace phenotypic evaluation.

Introgression breeding, also referred to as MABC, has been one of the most, if not the most, successful form of MAS in private breeding programs to date. The use of MABC to introgress transgenes into elite maize or soybean inbred lines (Crosbie *et al.*, 2006; Ragot *et al.*, 1995) has permitted the rapid deployment of transgenic insect and herbicide resistance traits across regions, creating tremendous value for seed companies, farmers, and other downstream actors. MABC is also very effective for introgressing specific genes or QTL from donor genotypes (nonadapted materials or related species) into elite breeding lines reducing both the time needed to produce commercial cultivars and the risk of undesirable linkage drag with deleterious donor attributes. Reports of successful use of MABC in private breeding programs are scarce in spite of positive outcomes from a variety of public programs on tomato, rice, barley, and soybean (Dwivedi *et al.*, 2007). Financial cost-benefit considerations will usually determine whether introgression breeding should be conducted with or without the assistance of molecular markers.

In public breeding programs, marker-assisted recurrent selection (MARS) has often been used in the context of population improvement (Gallais *et al.*, 1997; Hospital *et al.*, 1997; Knapp, 1998; Moreau *et al.*, 1998; Xie and Xu, 1998), based on breeding schemes where selected individuals are random-mated. In contrast, private breeding programs, in particular for maize, have often implemented MARS schemes focused more on directed recombination (Crosbie *et al.*, 2006; Eathington, 2005; Ragot *et al.*, 2000) in order to recover an ideal genotype through the creation of a mosaic of favorable chromosomal segments from the parental genotypes. This approach is referred to as genotype construction and is based on simultaneous selection for multiple traits (often using marker information only) such as yield, biotic and abiotic stress resistance, and quality attributes (Eathington, 2005; Ragot *et al.*, 2000). Although several of these target traits have complex inheritance, the commercial breeding programs report dramatic increases in the rate of genetic gain over phenotypic selection in maize (Crosbie *et al.*, 2006; Eathington, 2005). The specific molecular breeding systems used by commercial breeding programs are often trade secrets, but it is likely that there are several critical factors in their success including: (1) simultaneous marker-only selection for several traits involving probably 10 to more than 50 QTL or genes, (2) multiple cycles of MARS per year using markers flanking

QTL, (3) use of off-season nursery facilities for generation advance, and (5) genotyping large populations and use marker information to select plants prior to flowering to enable directed recombination. In these breeding systems, phenotypic selection is not applied at every generation. For example, the cycle length in MARS can be as short as 3 months, while that of phenotypic recurrent selection can span from 1 to several years. Such substantial differences in cycle length are expected to have significant impacts on the rate of genetic gain over the entire breeding system. Commercial breeding programs have also put great efforts into reducing costs, not only for genotyping data but also for phenotypic data. It is likely that cost ratio between marker data points to experimental field plot data points is lower in large private breeding programs than in most public research laboratories or small private programs. These are important factors for the economic efficiency of MARS applications.

Successful application of MAS in the private sector has been featured by its crops. For example, rice, as an autogamous crop, is very hard to make its hybrid vigor utilized compared to open-pollinated crops such as maize. Hybrid rice breeding has been depending on using either male sterility and its fertility restoration or environment-induced genic male sterility for hybrid seed production. The former needs a large number of testcrosses and progeny tests to identify the genes for male sterility and fertility restoration during the breeding process, while the latter depends on specific environments and multiple location or season trials to select for the related genes, both of which are extremely time consuming and labor intensive. MAS in hybrid rice breeding for the traits requiring testcrossing or progeny testing and for environment-dependent traits has been intensively discussed elsewhere (Xu, 2003), and now has become routine in hybrid breeding using both cytoplasmic male sterility and environment-induced genic male sterility. In addition, MAS has been widely used in the private sector for seed quality assurance. One of the examples is to identify and remove the false hybrids produced because the temperature during flowering time goes abnormal and down below the critical level that is required for conversion of environment-induced male sterility lines from sterility to fertility, which would not happen under normal temperature conditions.

The international seed companies have invested heavily in the assembly, modification, and integration of new methods and tools for the detection of DNA polymorphisms, the continuous operation of nurseries, and the optimization of data management, analysis, and interpretation. The development of PCR technology and the large-scale identification of SNPs (Lindblad-Toh *et al.*, 2000) have facilitated the development of molecular marker systems amenable to the levels of miniaturization and automation. This has in turn allowed the development of genotyping pipelines capable of rapidly and cost effectively generating millions of data points a year. It is only

at this level and timeliness of throughput that large breeding programs can realize true benefits of MAS. The allelic diversity at SNP loci is low (usually limited to two alleles, although generally providing codominant information), and the level of polymorphism at any given SNP loci may also be low in breeding populations. However, this is generally considered to be more than offset by the very high abundance and random distribution of SNP loci which can be combined and analyzed as haplotypes (Ching *et al.*, 2002). Thus, highly dense genetic maps can be developed with thousands of SNP markers, and marker-trait associations can be readily identified that are very close or inside the target gene. For these reasons, SNP-based genotyping is becoming the assay of choice for private MAS programs for well-studied crops.

The ability to select plants without their being phenotypically characterized is one of the main advantages of MAS. Many private breeding programs have upgraded or are upgrading their continuous nurseries (greenhouses, screenhouses, or open fields) so that they can be managed, equipped, and staffed in such a way that the plants complete their life cycle as quickly as possible and that tissue samples be collected efficiently at each generation for genotyping. Efficient MAS programs require access to and synthesis of very large amounts of data of different types (phenotypes, genotypes, pedigrees, environmental characteristics) and from various sources into useful genetic information. The rapidly increasing amounts of data generated in crop research and breeding programs driving dramatic advances in supporting computational sciences. Modern molecular breeding requires a range of complex large-scale data analyses to be carried out very rapidly. In particular, the development of computer software to track, manipulate, and comparatively analyze data for major genes, QTL, background haplotypes, and phenotypes across germplasm, pedigrees and cycles of the breeding process. Most of the computational tools used in private sector molecular breeding programs have been developed internally and remained under proprietary protection. Some large private breeding programs had established large research and support groups of dedicated data managers prior to the advent of MAS and genomics. Today, there is a fundamental dependence on dedicated specialists, systematically integrated into breeding programs, genotyping pipelines, and repositories of internal and external genetic information.

Many private breeding programs have invested heavily in the implementation of MAS. While there are no public reports of the cost-benefit ratio of the commercialization of MAS-derived cultivars in private sector, the growing portfolio of patent applications associated with MAS technologies (e.g., US5,492,547 1996; US5,746,023P 1998; US6,368,806B1 2002; US6,399,855B1 2002; US6,455,758B1 2002; US2005/0144664A1 2005; WO2005/000006A2 2005; WO2005/014858A2 2005) clearly suggests that commercial breeding programs see significant comparative advantage from the use of such approaches. Moreover, the likely scale of the investment

suggests that commercial seed companies are much more convinced of the benefits of MAS than most public breeding programs.

Small- to medium-sized seed companies without access to technology and with limited resources are forming alliances with multinational companies, universities, and CGIAR institutions to enable access to the necessary infrastructure, core competencies, and marker technologies without the prohibitively high-capital investment normally associated with such endeavors, for example, the "Agribiotech Park" at ICRISAT in India (http://www.agrisciencepark.icrisat.org/amenities.htm), the BecA at ILRI in Kenya (http://www.biosciencesafrica.org/BecA%20home.htm/), the Agronatura at CIAT in Colombia (http://www.ciat.cgiar.org/agronatura/index.htm), and CRIL of the IRRI-CIMMYT alliance (http://www:iita.org/cms/articlefiles/490-Genonics%20Taskforce%20Report%20March%202006.doc).

VI. IMPACT OF MARKER-ASSISTED GENETIC ENHANCEMENT

A. Enhanced Selection Power

The enhanced selection power of DNA markers resides in their ability to precisely identify a plant's genotype for a specific target trait without the confounding effects of the environment (Ribaut and Hoisington, 1998). The selection of genotypes based on genetic values predicted by molecular marker data can increase the rate of genetic gain by enhancing the precision of selection and by shortening selection cycles (Meuwissen *et al.*, 2001). MAS may also be valuable for pyramiding genes of similar phenotypic effect or selecting for resistance to pests and diseases not present in the breeding location. The high heritability of genetic markers (in theory being 1.0, although in practice rarely achieving this absolute level) compared to the trait for which they have been developed make them useful for MAS. Improvements in marker techniques have increasingly added to the selection power of MAS, both by providing more reliable types of markers and a rapidly increasing list of trait-associated loci. A critical improvement was the move from time-consuming hybridization-based assay (RFLP) to PCR-based assays (initially RAPD) for which amplification is dependent on DNA concentration and quality, annealing temperature and thermocycling conditions, *Taq* polymerase concentrations, and the relative proportion of all components in the PCR cocktail. Unfortunately, RAPD suffers many reproducibility and transferability problems, thus considerable efforts have been made to develop more robust PCR-based marker systems such as SCAR markers and other single-copy markers which have proven more reliable and repeatable and therefore of higher heritability. However, most recently two

new classes of PCR-based marker have emerged that have the added advantage of being highlight polymorphic in most breeding populations (SSR markers) or highly abundant across most plant species genes (SNP markers). SSR and SNP markers offer greater precision, power of selection, and perhaps most importantly, ease of scale-up, and thus, have become the markers of choice for molecular breeding programs of most crops. Thus, the type of marker has become an important determinant of the power of MAS to enhance selection. The selection power of molecular markers also resides in their good genome coverage and capacity to provide complete genome information, a characteristic that has also improved with newer marker technologies.

The enhanced selection power of MAS in addition to being related to the reliability and ease of applying a given type of marker also depends on proximity of linkage between markers and the gene(s) of interest (Ribaut *et al.*, 1997b). In addition, the level of phenotypic variance explained by the marker compared to the total genetic variance for the trait is also a critically important criterion (Bearzoti and Vencovsky, 1998). Greater distance between a marker and the gene(s) of interest underlying the target trait reduces the power of selection. In terms of linkage, the nature of the cross, particularly in terms of how closely are the parents related to each other and to the pedigree of target breeding populations, affects the frequency of recombination around target genes within the mapping populations versus the target breeding populations. The choice of parental genotypes for mapping populations also determines the level of polymorphism and whether the marker will facilitate the positive selection for the desirable or undesirable alleles. The potential risk that recombination will decouple the linkage between marker locus and gene of interest can be addressed by using flanking markers, which have greater power to counteract the effects of recombination around loci of interest by providing a diagnostic for the introgression of an entire genomic segment. MAS is most effective when there is a high level of polymorphism in the crosses being screened, and this is also the breeding situation in which gene introgression is most difficult, time consuming, and plagued by linkage drag. Not surprisingly, therefore, marker-assisted introgression and marker-accelerated backcross breeding are the areas where genomic applications have had their widest application and greatest success. Thus, there is a range of successful reports of using flanking markers for introgression of new traits through interspecific crosses with wild relatives or crosses between gene pools within the cultivated species, where markers are often more effective.

In the case of markers linked to the QTL, the proportion of the total phenotypic variance conveyed by each QTL is a key to the value of that marker in enhancing the breeding gain for the target trait. Similarly, there should be a high level of confidence in the existence of a QTL associated with

the target trait, as determined by the use of high LOD likelihood threshold during the identification of QTL markers (Tanksley, 1993). Simulation studies have shown that when a moderate-to-large number of QTL are influencing the target trait, a whole-genome scanning approach is often necessary and that the efficiency of MAS is substantially affected by population size and heritability of the target trait (Bearzoti and Vencovsky, 2002; Lande and Thompson, 1990). Enhanced power of selection through MAS can come not only from the power to make positive selection for a single gene but also from its power to assert negative and positive selection for a suite of genes or QTL across the entire genome (Hospital and Charcosset, 1997). It is in this transition from single point interventions of MAS to holistic molecular breeding strategies that we expect to see an exponential gain from the application of genomics in plant breeding programs. In this case, marker genotypes at various loci (associated with several mono-, oligo, and/or polygenic traits) are used within the context of an index for eliminating part of a breeding population, thus reducing nursery growout space and costs (Bearzoti and Vencovsky, 1998; Gimelfarb and Lande, 1994, 1995).

The most common application of MAS is in marker-assisted/accelerated backcross breeding. Optimally, this is based on positive foreground selection for donor trait, positive background selection for the recurrent parent genome, and negative background selection against undesirable donor parent alleles (Frisch *et al.*, 1999b; Ribaut *et al.*, 2002). Marker-assisted introgression can dramatically reduce the number of generations of backcrossing required to recover the elite parent background (Hospital *et al.*, 1992), although the number of generations saved depends on the size of the genome, level of recombination in the cross, size of the progeny population, and number of available markers. Genomic map length, population size, and duration of backcrossing also influence on the attainable rate of donor genome substitution. For example, larger genome requires larger population as well as more markers to attain a given rate of donor genome substitution (Stam, 2003). Meanwhile, partial or whole chromosome selection can be used when introgressing from an exotic genome where recombination with the cultivated genome is very low or nonexistent (Wittaker *et al.*, 1995). MAS can also be a great assistance in the selection of favorable recombinants during inbreeding and/or crossbreeding cycles using backcross products, thus increasing the speed with which advanced lines are generated (Frisch *et al.*, 2000). Furthermore, MABC can reduce the effects of linkage drag by selecting for fewer and smaller donor genome fragments. In this case, increasing selection power and breeding gain is obtained by use of a greater number of background markers combined with closer flanking markers for the target trait gene(s).

Using computer simulations and additive, dominance, and epistasis genetic model, Liu *et al.* (2004a) demonstrated that combining MAS in

early generations with phenotypic selection in later generations is the most efficient breeding strategy for self-fertilizing crops. Investigation on different crossing strategies and consideration of when to screen, what proportion to retain, and the impacts of dominant versus codominant marker expression revealed important choices in the design of MAS programs that can produce large efficiency gains. F_2 enrichment, increasing homozygosity through inbreeding or DH, and backcrossing to increase the frequency of recurrent parent alleles are effective strategies for improving the efficiency of MAS that will allow either smaller populations to be screened or selection at more loci. However, fixation of alleles in early generation requires larger populations and is undesirable in most instances (Bonnett *et al.*, 2005).

B. Reduced Cost, Increased Feasibility, Time Savings, and Parental Selection

MAS can be useful for the selection of traits that are difficult or impossible to breed through phenotypic selection due to logistical, biological, or quantitative-based constraints. In terms of genetic associations, codominant markers for recessive genes are especially valuable since phenotypic selection will be highly inefficient as it is likely to discard all heterozygous progeny during early generations of the breeding cycle. While recessive genes can be selected with progeny testing or testcrosses, this clearly adds substantial time and effort to the breeding process. Thus, MAS has the advantages of obviating these time-consuming steps and facilitating precise and efficient early generation selection. Dominant markers in coupling phase with target trait can also be of value in such breeding systems. However, if only a dominant marker in repulsion is available, then early generation MAS would be limited to negative selection against homozygous dominant and heterozygous plants, which would be inefficient since potentially useful allele-carrying genotypes would be eliminated. This type of marker is most useful in advanced generations of self-pollinated crops when a recessive gene has already been fixed by inbreeding. However, MAS with this type of marker is impossible in generations where no homozygous recessive plants exist at all such as the BC_1F_1 to the dominant allele-containing parent.

MAS scenarios with the greatest cost-benefit ratio include traits that would otherwise require highly expensive phenotypic or biochemical evaluation procedures (Ribaut and Hoisington, 1998). This is the case for traits that require extensive field testing at specific locations or times of the year. Likewise, many phytochemical traits analyzed in reproductive or vegetative tissues at various growth stages are expensive to carry out. For example, the analysis of seed quality, secondary metabolites, and micronutrients remains expensive and time-consuming and MAS can replace more costly and

difficult assays with more standardized DNA-based technologies. Molecular markers are proving more efficient, rapid, and simple to implement on a large scale for seed protein traits since they are based on DNA extracted at any growth stage from a small amount of expendable tissue. For example, in the selection of quality protein maize, MAS is cost-effective when a visual marker is not available (Dreher *et al.*, 2002, 2003; Morris *et al.*, 2003). Similarly, for the evaluation of mineral content in seed tissue, MAS might be less expensive than traditional quality evaluations, a process that sometimes requires dissected seed organs or collecting several grams of seed tissue. The advantage of MAS resides in the small amount of template DNA required for carrying out a large number of assays. Thus, MAS efficiency can be dramatically increased by using a single DNA extraction for the evaluation of several to many markers.

After the development of molecular markers and validation of their power of indirect selection for the trait (see Section III), it is then often necessary to optimize the assay for scale-up to large-scale application (Young, 1999). Sometimes this involves changes in breeding program logistics, PCR protocols, marker detection technique, or even complete redesign of the markers themselves. In all cases, the driving criteria being to reduce unit costs and turn around times while increasing throughput and minimizing errors, and ultimately optimizing the cost-benefit advantage of MAS over phenotypic selection. Marker redesign has been a common element of scaling-up exercises and can involve something as simple as optimizing the size or genomic position of the PCR amplification fragment. Technologies that speed up the implementation process, reduce laboratory requirements or errors, and lower the costs associated with scaling-up are crucial to the success of MAS (Gu *et al.*, 1995). For example, techniques have been developed which reduce the cost of DNA extraction and result in large time-savings (Dellaporta *et al.*, 1983; Ikeda *et al.*, 2001b; Klimyuk *et al.*, 1993). Kuchel *et al.* (2005) designed a genetically effective and economically efficient marker-assisted breeding strategy aimed at selecting for favorable alleles in wheat breeding. Although incorporating MAS for allele enrichment in the BC_1F_1 population, gene selection at the haploid stage, and the selection of recurrent parent background of DH prior to field testing was effective to select for a high frequency of desired alleles, the incorporation of marker selection at the BC_1F_1 and haploid stage was the most effective as it not only increased genetic gain over the phenotypic selection but also reduced cost by 40%.

Furthermore, MAS can be used in conditions that are not favorable for phenotypic screening, for example selection of resistance genes in regions where quarantine restriction prevents introduction of an exotic pathogen or pathogen strain or where a pathogen does not occur at a sufficiently high level to perform effective field screening and selection (Ribaut and Hoisington, 1998). Markers for disease resistance have the advantage of

obviating the need for field or greenhouse inoculations that sometimes are ineffective or unreliable if environmental conditions are not propitious and can result in savings in time and cost compared to phenotypic selection. A further advantage of MAS is that it can be implemented in any generation of the breeding process and under both field or greenhouse conditions, while phenotypic selection often requires planting a separate trial and provision of specialized labor for inoculation, agronomic management, and evaluations or scoring. In addition, phenotypic screening of fixed lines or segregating populations often requires replicated testing to minimize the effect of GEI, whereas MAS can be evaluated on a single plant basis as long as the marker is associated with a locus which contributes a large percentage of the genetic variance of the target trait. A potential disadvantage of relying on MAS over phenotypic selection is that it commits a breeder to a unique gene or set of genes for a given trait. Thus, where a breeder relies solely on MAS for selection, this can exclude other possible genes and the use of other potentially useful parents that do not possess the allele(s) being targeted by the MAS. Of course, this is rarely the recommended approach, and most molecular breeding programs will involve at least one or two cycles of phenotypic evaluation during the overall breeding process. In this way, the results of the MAS can be validated, while other alleles and genes positively contributing to the target trait can be selected. A refined model for this approach has been proposed by Ribaut and Betrán (2000) in maize for fixing valuable genes in a population improvement breeding program (that includes a large number of parents) through the application of single large-scale (SLS) MAS and then intercrossing to recreate diverse populations for further selection.

MAS can also help in situations where timeliness is a major constraint since DNA can be obtained at the seedling stage or depending on the crop, even from the seed itself. Timeliness is an especially important issue in the case of perennial crops where many economically important traits are only expressed at the reproductive stage which may take one or more years. Therefore, MAS for late cycle traits in long-duration crops provides a much greater cost-benefit ratio than in annual crops (Morris *et al.*, 2003).

When breeding complex traits with low heritability and high GEI, selection based on phenotypic evaluation can become very difficult. In these situations, the dissection of complex traits into component traits can increase the chances of effective selection as each component can be selected separately. Then, in turn, MAS for major QTL underlying each component trait may provide the best breeding gains. Selection of just the QTL that account for the largest proportion of phenotypic variance is advisable under these conditions (Asíns, 2002; Tanksley, 1993). In the case of polygenic traits, MAS has the potential for pyramiding different sources of genes for a given trait, whether it be to create durable disease resistance through simultaneous deployment of multiple R gene combinations or to create superior cultivars

through the accumulation of positive alleles for different components of a given trait such as drought or low soil fertility tolerance. There are a number of good examples of successful pyramiding of pest or disease-resistance genes (see Section IV.A). However, there are very few reports of successful applications of MAS for complex abiotic stress tolerance traits (see Section IV.B). Thus, the long held belief that MAS would have its greatest impact on trait with low heritability and high GEI interaction still awaits widespread practical demonstration. However, experience has shown that the ability to manipulate even one important component trait with confidence can make a breeding program more efficient if that gene is highly desirable and valuable for advanced materials.

MAS can also be useful in the selection of parental genotypes, especially in the breeding of crops where heterosis is expressed. In this case, parental selection can benefit from marker assessments of genetic distance between individuals in crops where genetic distance has been shown to be predictive of heterotic pools or combining ability. Finally, MAS can also be used to determine heterozygosity during the creation of inbred lines for allogamous crops.

C. Overview of Products from Molecular Breeding

To date, polymorphic DNA markers and genetic maps are available for virtually all crops, albeit in varying numbers and levels of genomic saturation (see Sections II.B and C). Similarly, the genetics of many agronomic traits is well understood in many crops, and the marker-trait linkages have been reported for many traits in a large number of crops, although reports of validation in different genetic backgrounds and environments are naturally only beginning to emerge (see Section III). MAS is now being practiced in most well-studied crops (see Section IV.A–C), yet in the private sector MAS applications are dominated by transgene introgression and backcross programs with only limited reports of their use for complex traits. In this section, we provide an overview of the products (cultivars and breeding lines) developed using MAS in combination with conventional breeding. Eighteen MAS-derived cultivars and several advanced lines combining resistance to biotic and abiotic stresses or improved grain quality have been reported in rice, wheat, barley, pearl millet, common bean, and soybean (Table XV). To date, MAS has been most successful in the selection of resistance to diseases and for improving grain quality. For example, rice cultivars resistant to blast in United States and to bacterial blight in Indonesia, wheat cultivars resistant to rust in Canada, and common bean cultivars resistant to anthracnose and *Bean golden yellow mosaic virus* in United States, and those with resistance to *Sclerotinia* white mold in Canada have been developed using MAS and

Table XV
List of Cultivars and Hybrids, Advanced Lines and Improved Germplasm Developed by MAS in Barley, Common Bean, Pearl millet, Rice, Soybean, and Wheat

Advanced lines and cultivars developed by marker-aided breeding	References
Barley	
Aluminum	
Advanced lines including WB259, possessing good malt quality and aluminum tolerance developed in Australia	http://www.cdesign.com.au/bts2005/pages/papers_2003/papers/134venkatanagappaS.pdf
Grain yield and malt quality	
An isogenic line 00-170 consistently produced high yield and good malt quality in Australia	Schmierer *et al.*, 2004
Common bean	
Angular leaf spot	
Resistance to angular leaf spot transferred into Carioca type bean, Rudá in Brazil	(M. Blair, CIAT, personal communication)
Anthracnose	
Co-4^2 allele transferred into pinto beans (highly susceptible to Durango race) grown in North America	Miklas and Kelly, 2002
Resistance to anthracnose incorporated in Pinto bean cultivar, USPT-ANT-1 containing *Co-4*^{2} gene that confers resistance to all known North American races of anthracnose in United States	Miklas *et al.*, 2003
Resistance to anthracnose transferred in cultivar Perola in Brazil	Ragagnin *et al.*, 2003
***Bean common mosaic necrosis virus* (BCMV)**	
Red bean with resistance to BCMV, containing *I* and *bc-3*, developed for central America	Beaver *et al.*, 1998
BCMV and anthracnose	
1800 breeding lines of climbing beans, containing *bc-3*, *I*, *C0-4*, and *Co-5*, with combined resistance to BCMV and anthracnose selected in Colombia	http://www.african crops.net/abstracts2/bean/blair.htm
***Bean golden yellow mosaic virus* (BGYMV)**	
A pole bean cultivar, Genuine, resistant to BGYMV developed in Central America	Stavely *et al.*, 1997
A pole garden bean cultivar, Genuine, with moderate resistance to BGYMV developed in United States	Stavely *et al.*, 2001
Comman bacterial blight (CBB)	
Pinto bean germplasm, ABCP-8, resistant to CBB developed in United States	Mutlu *et al.*, 2005
USDK-CBB-15, dark red kidney bean, highly resistant to CBB released in United States	(M. Blair, CIAT, personal communication)

Table XV (*continued*)

Advanced lines and cultivars developed by marker-aided breeding	References
CBB, anthracnose, and BCMV	
Advanced lines with multiple resistance to CBB, BCMV, and anthracnose developed in Canada	http://www.ontariobeans.on.ca.ppyramidingDiseaseResistanceGenes.html
Rust	
Rust resistant genes, *Ur-4* and *Ur-5*, combined in the BARC-rust resistant green and waxy bean germplasm lines in Honduras	Stavely and Steinke, 1990; Stavely and McMillan, 1992
Rust resistant genes, *Ur-4* and *Ur-11*, introgressed into navy bean lines BelMiDak-RR-1 to 7 in Honduras	Stavely *et al.*, 1994
Rust and anthracnose	
Five lines resistant to rust and anthracnose developed, with Vi0699 and Vi2599 significantly outyielding controls in Brazil	Faleiro *et al.*, 2004
Rust, anthracnose, and angular leaf spot	
Resistance to anthracnose in TO and AB136; to angular leaf spot in AND277; to rust in Ouro Negro; and to rust and anthracnose in Ouro Negro transferred in Brazil	(M. Blair, CIAT, personal communication)
Rust and *Bean golden yellow mosaic virus* (BGYMV)	
White-seeded Snap bean cultivars, BELDADE-RGMR 4, 5, and 6, possessing resistance to rust and BGYMV released United States	(M. Blair, CIAT, personal communication)
***Sclerotinia* white mold**	
QTL B7 and B8 QTL linked with resistance to white mold transferred into Winchester and Maverick that yielded at par with controls in Canada	Miklas *et al.*, 2004 (http://www.whitemoldresearch.com/presentation2004/Miklas.pdf)
Pearl millet	
Downy mildew	
The parental lines of the original hybrid (HHB 67) improved for downy mildew resistance through MAS and conventional backcross breeding, and new hybrid HHB 67-2 with improved resistance to downy mildew released in India	http://www.secheresse.info/article.php3?id-article=1919
Rice	
Amylose content	
Cadet and Jacinto with unique cooking and processing quality traits released in United States	http://usda-ars-beaumont.tamu.edu/marker.html
Bacterial blight (BB)	
Angke and Conde, possessing resistance to BB, produced 20% greater yield over IR64 and released in Indonesia	Bustamam *et al.*, 2002

(*continued*)

Table XV (*continued*)

Advanced lines and cultivars developed by marker-aided breeding	References
Resistance to BB transferred in R8006 and R1176 and when crossed to Zhong 9A, the hybrids (Zhongyou 6 and Zhongyou 1176) produced high yield, resistant to BB, and good grain quality in China	Cao *et al.*, 2003
AR32-19-3-3, AR32-19-3-4, AR32-4-3-1, and AR32-4-58-2, all resistant to BB, showed 18–31% yield advantage over PSB Rc28 in Philippines	Leung *et al.*, 2004
BB resistant hybrids, Guofeng No 2 and Hybrid II You 218 released in China, produced 11–19% greater yield over Shanyou	Leung *et al.*, 2004
PR 106-P2 and PR 106-P9, both resistant to BB, showed 18–22% yield increase over PR 106 in India	Leung *et al.*, 2004
Blast	
CS 2, CS 11, CS 18, CS 35, CS 36, CS 62, and CS 67 combining resistance to blast and good agronomic traits developed, with potential to replace CR 203 in Vietnam	http://www.fftc.agnet.org/library/article/rh2003013a.html
Soybean	
Oil quality	
Vistive low-linolenic soybean developed by Monsanto and released for cultivation in United States	http://www.monsanto.com
Wheat	
Aluminum toxicity	
Advanced backcross lines tolerant to aluminum developed	http://www.dfid-psp.org/ccstudio/publications/annualreport/2004_aluminium.pdf
Bread-making quality	
A wheat cultivar, Burnside, with CWES (Canadian Western Extra Strong) traits developed in Canada	Radovanovic and Cloutier, 2003)
***Fusarium* head blight (FHB)**	
NILs containing major 3BS QTL and resistant to FHB developed in United States	Zhou *et al.*, 2003b
Rust	
Resistance to stem (*Sr39*) and leaf rust (*Lr35*) incorporated into "Canada Prairie Spring" and "Canada Western Extra Strong" classes of wheat lines in Canada	Gold *et al.*, 1999
Multiple resistance to pest, fungal and viral diseases + grain quality	
Several germplasm lines possessing resistance to pest, fungal, and viral diseases, and those with improved grain quality developed in United States	http://maswheat.ucdavis.edu

released for commercial cultivation. Two rice cultivars with MAS-derived improvements in amylose content are grown in United States. MAS has also been successful in the development of disease-resistant hybrids. For example, superior rice hybrids with resistance to bacterial blight in China and pearl millet hybrid with resistance to downy mildew have been released for cultivation. In addition, many advanced lines and improved germplasm combining multiple resistances to diseases or with improved seed quality have been bred, which are now being evaluated in several countries prior to their release as new cultivars (Table XV). Marker-assisted backcross breeding and marker-aided gene pyramiding have been the most frequently used molecular breeding methods to aid the introgression of disease resistance or quality traits into improved genetic backgrounds. MAS has also been used in wide crosses to minimize the linkage drag associated with beneficial traits (see Section IV.D).

Although there are only small numbers of reports regarding successful use of MAS in plant breeding, the technology has nevertheless demonstrated its potential as a tool to support conventional genetic enhancement of crops. Large-scale adoption of MAS technology has already begun for incorporating disease resistance or grain quality in rice (http://www.uark.edu/ua/ricecap/index.htm), wheat (http://maswheat.ucdavis.edu), barley (http://www.barleycap.org/), and common bean (Kelly *et al.*, 2003; Miklas *et al.*, 2006a) in United States. For example, MAS wheat consortium has developed protocols for more than 40 molecular markers for resistance genes and quality traits and used MABC to incorporate 27 different disease- and pest-resistance genes and 20 alleles with beneficial effects on bread making and pasta quality into ~180 lines adapted to the primary US production regions (http://maswheat.ucdavis.edu/). Rice researchers in China are using MAS to combine resistance to diseases and improved grain quality in some of their best-performing hybrids (Leung *et al.*, 2004). MAS is being used to combine disease resistance and/or grain quality in wheat and common bean in Canada (Radovanovic and Cloutier, 2003; http://www.ontariobeans.on.ca.ppyramidingDiseaseResistanceGenes.html) and for improving wheat, barley, and rice in Australia (Christopher *et al.*, 2004; Eagles *et al.*, 2001; McLauchlan *et al.*, 2001; Ogbonnaya *et al.*, 2001; Paris *et al.*, 2003; Schmierer *et al.*, 2004; http://www.cdesign.com.au/bts2005/pages/papers/134venkatangappaS.pdf). CIMMYT wheat breeding program has already initiated marker-assisted breeding to introgress gene(s) for resistance to cereal cyst and root lesion nematodes, boron toxicity, *Barley yellow dwarf virus*, scab, rust, and crown rot as well using *Ph1b* to promote pairing between alien and wheat chromosomes to accelerate gene transfer from alien species to wheat. Moreover, it is expected that many more successful applications do exist but remain within the confidentiality restriction of commercial breeding companies around the world.

Developing countries are not left behind in the use of MAS in crop breeding programs. For example, researchers from the Indian Council of

Agricultural Research are collaborating with their colleagues at IRRI, CIMMYT, and ICRISAT on the use of MAS in cereal and legume breeding. In fact, the first downy mildew resistant pearl millet hybrid (HHB 67-2) released in India was bred using MAS by improving the male parent with improved resistance to downy mildew (Hash, 2005). India is testing marker-derived submergence-tolerant lines (Xu *et al.*, 2006), developed through collaboration with IRRI, for their adaptation to deepwater paddy cultivation in eastern India. Development of submergence-tolerant cultivars using MAS has already been reported from Thailand (Siangliw *et al.*, 2003), and work is in progress to introduce this trait in cultivars adapted in Bangladesh, Laos, the Philippines, and Vietnam. The ultimate goal of this collaboration with IRRI is the development of improved rice inbred and hybrid cultivars with good grain quality and multiple resistances to pests and diseases. MAS-derived rice cultivars are already being grown in Indonesia. These marker-aided rice cultivars and hybrids have produced on average 11–34% increased yield over popular inbred and hybrid cultivars in Asian countries. This has led to an estimated increase in grain harvest of 0.8 million Mt (worth US $20.5 million) of paddy rice per cropping season in India, Indonesia, the Philippines, and China as a result of the growing bacterial blight resistance present in these inbred and hybrid cultivars (Leung *et al.*, 2004). Many national programs from South America are cooperating with CIAT and advanced research institutes in United States to improve the genetic potential of common bean, the most widely grown pulse crop in that region, by using MAS (Miklas *et al.*, 2006a).

VII. APPROACHES TO ENHANCE THE EFFICIENCY AND SCOPE OF MOLECULAR BREEDING

A. Studying the Molecular Basis of Heterosis

Heterosis is defined as the superior performance of an F_1 hybrid as compared with its parents. Hybrid cultivars have made significant contribution to world food supply (Duvick, 1999). In the literature, dominance, overdominance, and epistasis have been implicated as the genetic basis of superior hybrid performance. The dominance model attributes increased vigor to the action of favorable dominant alleles from both parents combined in the hybrid, whereas the overdominance model postulates the existence of loci where the heterozygous state is superior to either homozygote (Xiao *et al.*, 1995; Xu, 2003; Yu *et al.*, 1997). Evidence for the role of epistasis (interaction of the favorable alleles at different loci contributed by the two parents) in hybrid vigor have also been reported (Li *et al.*, 2001b; Luo *et al.*, 2001;

Stuber *et al.*, 1992; Xu, 2003). The genetic basis of heterosis, heterotic groups, hybrid prediction and hybrid performance, relationships between heterozygosity and genetic distance with hybrid performance and heterosis, and use of MAS in hybrid breeding have been discussed elsewhere (Xu, 2003).

The complex nature of heterosis makes it difficult to partition into individual components because of the epistatic interactions among segregating loci throughout the genome (Li *et al.*, 2001b). To assess the importance of loci with overdominant (ODO) effects in expression of heterosis, Semel *et al.* (2006) employed NIL, carrying single marker-defined chromosome segments from distantly related wild species *Solanum pennellii* to partition heterosis into defined genomic regions, eliminating a major part of the genome-wide epistasis. They detected 841 QTL for 35 diverse traits. NILs showing greater reproductive fitness are characterized by the prevalence of ODO QTL, which were virtually absent for the nonreproductive traits. Overdominance results from true overdominance due to allelic interactions of a single gene or from pseudo-overdominance involving linked loci with dominant alleles in repulsion. In their study, although they detected dominant and recessive QTL for all phenotypic traits, overdominance only for the reproductive traits indicates that pseudo-overdominance is unlikely to explain heterosis in NIL, thus they favor the true ODO model, a single functional Mendelian locus involved in heterosis.

Milborrow (1998) proposed a mechanistic, biochemical interpretation of the superior performance of F_1 hybrids in comparison to their homozygous parents. Their interpretation is based on the concept that growth is restricted below the potential maximum by internal genetic factors. In this model, the hybrid vigor is caused by a slight reduction in the strictness of this control mechanism in heterozygotes compared with homozygotes, particularly with respect to metabolism and growth processes. This effect is believed to be mediated by the presence of changes in regulatory features of certain loci when in the heterozygote state.

Among the cereals, heterosis has been exploited in maize, rice, sorghum, and pearl millet to produce superior yielding hybrids that by far dominate the global acreage for each crop. For example, about 95% of US maize acreage is planted to hybrids that exhibit a 15% yield advantage relative to the best open-pollinated cultivars (Duvick, 1999). A popular hybrid rice cultivar in China (*LYP9*) produces 20–30% more grains per hectare than other hybrids or inbred rice cultivars (Lu and Zhou, 2000). More recently, an "immortalized F_2" population was generated by randomly permutated intermating of 240 RILs from a cross between the parents of Shanyou 63, another widely cultivated hybrid rice cultivar in China. These lines were field evaluated over 2 years and genotyped using 231 polymorphic molecular markers covering the entire rice genome. From this analysis, 33 loci were detected that

contributed to heterotic effects in grain yield, tillers per plant, grains per panicle, and 1000-grain weight (Hua *et al.*, 2003). The heterotic loci showed little overlap with QTL previously identified for the same traits. Thus, in contrast to the Milborrow model (Milborrow, 1998), it appears that in rice there are unique loci conditioning heterosis. Moreover, all kinds of genetic effects were observed in this study to contribute to heterosis, including partial-, full-, and overdominance at the single-locus level and all three forms of digenic epistatic interactions (additive by additive, dominance by dominance, and additive by dominance). Heterosis effects at the single-locus level, in combination with the marginal advantages of double heterozygotes caused by dominance interaction at the two-loci level, adequately explain the genetic basis of heterosis in Shanyou 3. Using serial analysis of gene expression (SAGE), Bao *et al.* (2005) surveyed transcriptomes in panicles, leaves, and roots of a super-hybrid rice (*LYP9*) in comparison to its parental inbred cultivar genotypes (*93-11* and *PA64s*). They identified 595 upregulated and 25 downregulated tags in *LYP9* that were related to enhancing carbon- and nitrogen-assimilation, including photosynthesis in leaves, nitrogen uptake in roots, and rapid growth in both roots and panicles. This adds a crucial new set of observations for understanding the molecular mechanisms of heterosis and gene regulation networks in rice. In this study, they found massive complementation at the transcript level that further suggests that the underlying mechanisms of heterosis may not be as simple as have been reported from studies of a small number of genes (Birchler *et al.*, 2003).

Previous studies using multiple hybrids and their corresponding parents revealed that some differential gene expression patterns are significantly correlated with heterosis in wheat (Ni *et al.*, 2000, 2002; Sun *et al.*, 1999, 2004; Wu *et al.*, 2003). However, information on systematic identification and on characterization of differentially expressed genes is limited. Yao *et al.* (2005) used an interspecific hybrid between common wheat (*Triticum aestivum* L., $2n = 42$, AABBDD) line 3338 and spelt (*Triticum spelta* L., $2n = 42$, AABBDD) line 2463, which is highly heterotic both for aerial growth and root related traits. In their research, they included an expression assay using modified suppression subtractive hybridization (SSH) to generate four subtracted cDNA libraries between the wheat hybrid and its parental genotypes. Of the 748 nonredundant cDNAs obtained, 465 cDNAs had high sequence similarity to GenBank entries in diverse functional categories, such as metabolism, cell growth and maintenance, signal transduction, photosynthesis, response to stress, transcription regulation, and others. They further confirmed the expression patterns of 68.2% SSH-derived cDNAs by reverse Northern blot, while semiquantitative RT-PCR exhibited similar results (72.2%). This suggests that the genes differentially expressed between hybrids and their parents are involved in diverse physiological pathways, which may contribute to heterosis in wheat.

Maize inbred lines B73 and Mo17 produce a heterotic F_1 hybrid. Based on analysis with 13,999 cDNA microarrays, Swanson-Wagner *et al.* (2006) compared global patterns of gene expression in seedlings of the hybrid (B73 × Mo17) with those of its parental genotypes. A total of 1367 ESTs were observed to be significantly differentially expressed, using an estimated 15% FDR as cutoff. All possible modes of gene action were observed, including additivity, high- and low-parent dominance, underdominance, and overdominance. A total of 1062 of the 1367 ESTs exhibited expression patterns that are not statistically distinguishable from additivity, while the remaining 305 ESTs exhibited nonadditive gene expression. About 181 of the 305 nonadditive ESTs exhibited high parent dominance, 23 ESTs showed low parent dominance, while 44 ESTs displayed underdominance or overdominance. These results suggest that multiple genetic mechanisms, including overdominance, contribute to heterosis. This contrasts with previous studies that reported heterosis was due to gene action of only a small set of maize genes (Auger *et al.*, 2005; Guo *et al.*, 2004; Song and Messing, 2003). Further analysis of allelic variation in gene expression in the maize hybrid and its parental lines (B73 and Mo17) identified a subset of 27 genes that are differentially expressed in parental lines. When the transcriptional contribution of each allele from the inbred line was analyzed in the hybrid, the majority of the differential expression was observed to be due to *cis*-regulatory variation, and not due to differences in *trans*-acting regulatory factors. This suggests a predominance of additive expression and a lack of epistatic effects, as genes subject to *cis*-regulatory variation are expected to be expressed at mid-parent, or additive, levels in the hybrids (Stuper and Springer, 2006). Scheuring *et al.* (2006) used a 57,000 maize gene-specific long-oligonucleotide microarray containing about 32,000 genes to study the differential gene expression between a maize hybrid and its parental genotypes (B73 and Mo17). Preliminary analysis revealed that at least 800 genes were expressed at two- to ten-fold higher levels in the hybrid than the parent genotypes. Using Massively Parallel Signature Sequencing (MPSS), an open-ended mRNA profiling technology, of nearly 400 allelic signature tag pairs, Yang *et al.* (2006) found that 60% of the genes expressed in meristems of hybrid were significantly different in allele-specific transcript level as compared to the parental genotypes. This suggests an abundance of *cis*-regulatory polymorphisms affecting hybrid meristem gene expression. Furthermore, when comparing the expression of the same allele in the hybrid versus inbred parents, they found 50% of the genes expressed at a significantly different level. Such differences in expression are likely attributed to the effect of *trans*-acting factors that differ between the hybrid and inbreds. While *cis*-regulatory variation predicts additive expression, *trans*-regulation may result in nonadditive expression in the hybrid. Thus, studying the effect of transcript regulation at an allele-specific level provides a different level of understanding of gene regulation than focusing on overall expression in the hybrid.

With the vast genomic and technological resources available in *Arabidopsis thaliana* and the occurrence of heterosis in many traits (Meyer *et al.*, 2004 and references therein; Syed and Chen, 2005), *Arabidopsis* may be the best model for investigating the genetic basis of heterosis (Jansen and Nap, 2001). However, it is heterosis in yield which holds the greatest promise in plant breeding; thus, efforts must also be focused on validating and/or translating findings in *Arabidopsis* for greater understanding, and ultimately ability to manipulate, the genetic basis of heterosis in crop plants.

B. Fine-Mapping, Cloning, and Pyramiding of QTL Associated with Improved Agronomic Traits

Many agronomically important traits including yield are controlled by a few to a large number of genes (QTL), each with varying effects and different levels of GEI, which together confer a trait with continuous phenotypic variation. With the development of high-density genetic linkage maps based on DNA markers, it is possible to map QTL of large effect with a high level of resolution (Paterson *et al.*, 1988). However, it is difficult to identify all genes underlying QTL because the effects of many are relatively small and easily confounded by environmental conditions. Selfed lines from backcrosses (advanced backcross lines) are a common method of fine-mapping of QTL, where phenotypic differences can be more readily identified without the confounding effects of diverse segregating backgrounds (Darvasi and Soller, 1995; Graham *et al.*, 1997; Saito *et al.*, 2001; Yamamoto *et al.*, 1998). Alternately, NIL provides the means to dissect complex traits into simple Mendelian factors. Each NIL varies for a defined genomic segment containing a target QTL in an otherwise uniform genetic background. NILs are produced by repeatedly backcrossing a donor parent with a recurrent parent in combination with MAS. Comparing the phenotypes of NIL with those of the recurrent and donor parents permits an accurate evaluation of the effects of the target QTL in an adapted background without the confound factor of interaction with other segregating loci. Developing NIL has the added advantage of providing QTL ILs (with elite agronomic backgrounds) with the minimum of deleterious alleles in the vicinity of target QTL (linkage drag) which can then be used in marker-assisted pyramiding of QTL with different beneficial effects. NILs are also useful resources for developing large mapping populations for fine-mapping and map-based cloning of specific QTL. Thus, NILs are a uniquely powerful means of linking marker identification, QTL gene isolation, and advanced product development. ILs can also be used for fine-mapping of QTL (Eshed and Zamir, 1995). Peleman *et al.* (2005) proposed a method to fine-map multiple QTL in a single population: QTL are mapped in a relatively small population, and a large population of

1000 plants or more is used to derive QTL isogenic recombinants (QIRs). This reduces the number of lines required for phenotyping. LD methods for fine-mapping may also offer improved accuracy of QTL detection (Bink and Meuwissen, 2004; Grapes *et al.*, 2004).

There a very large number of reports in the literature regarding the identification of putative QTL for traits of agricultural importance in many crops. However, only a few studies have succeeded in fine-mapping and cloning of those QTL. The earliest examples of successful QTL cloning include a major fruit-weight QTL of tomato (*fw2.2*), delimited to a segment of cloned DNA (<150 kb) (Alpert and Tanksley, 1996), and QTL for tomato sugar content (*Brix9-2-5*) to a 484-bp region within an invertase gene (*Lin-5*) (Fridman *et al.*, 2000). With advances made in rice genomics, several QTL associated with agronomic traits have now also been cloned, for example, four QTL for heading date—*Hd1*, *Hd3a*, *Hd6*, and *Ehd1* (Doi *et al.*, 2004; Kojima *et al.*, 2002; Takahashi *et al.*, 2001; Yano *et al.*, 2000); QTL for grain number (*Gn1a*) and grain size (*GS3*) (Ashikari *et al.*, 2005; *Fan et al.*, 2006); QTL for salt tolerance (*SKC1*) (Ren *et al.*, 2005); QTL for regeneration ability (*PSR1*) (Nishimura *et al.*, 2005); and QTL for shattering (*Sh4* and *qSH1*) (Konishi *et al.*, 2006; Li *et al.*, 2006b). *Hd1*, *Hd3a*, and *Hd6* encode orthologues of *CONSTANS* (*CO*) and *Flowering locus T* (*FT*) and the α-subunit of casein kinase 2 (*CK2*), which are well-characterized factors for flowering or the circadian clock in *Arabidopsis* (Hayama and Coupland, 2004; Izawa *et al.*, 2003). However, rice *Hd1* promotes flowering under short-day lengths, while *Arabidopsis CO* promotes flowering in long-day conditions (Izawa *et al.*, 2003). *Gn1a* encodes a cytokinin oxidase/dehydrogenase (*OsCKX2*), an enzyme that degrades the phytohormone cytokinin. Reduced expression of *OsCKX2* causes cytokinin accumulation in inflorescence meristems, which increases the number of reproductive organs, resulting in higher grain yield (Ashikari *et al.*, 2005). *GS3* encodes a putative transmembrane protein, and a mutation in this gene induces large grain size, suggesting that *GS3* might function as a negative regulator for grain development (Fan *et al.*, 2006). *SKC1* encodes a sodium transporter involved in regulating K^+/Na^+ homeostasis under salt stress (Ren *et al.*, 2005). *Sh4* encodes an unknown protein that when mutated inhibits the normal development of an abscission layer, necessary for shattering (Li *et al.*, 2006b), similarly an SNP in the 5́ regulatory region of the *qSH1* gene causes loss of shattering owing to the absence of abscission layer formation in *japonica* rice (Konishi *et al.*, 2006).

The QTL for grain weight, *gw3.1* and *gw8.1*, have been fine-mapped in rice, the former in the pericentromeric region of chromosome 3 (93.8-kb region) (Li *et al.*, 2004a) while the latter on chromosome 8 to about 306.4-kb region between markers RM23201.CNR151 and RM30000.CNR99 (Xie *et al.*, 2006). The former locus has also been fine-mapped simultaneously by

three other groups, and it has been cloned using map-based cloning (Fan *et al.*, 2006). Similarly, another QTL influencing the number of grains per panicle (*gpa7*) has been successfully delimited to a 35-kb genome region on rice chromosome 7 (Tian *et al.*, 2006a). Andaya and Tai (2006) have fine-mapped a major QTL, *qCTS12*, for seedling cold tolerance in rice and successfully delimited it to a region of about 55 kb on the short arm of chromosome 12, with *OsGSTZ1* and *OsGSTZ2* the most likely candidates gene(s) for *qCTS12*.

VRN1 and *VRN2* are the main genes involved in the vernalization response in diploid wheat *T. monococcum* (Dubcovsky *et al.*, 1998; Tranquilli and Dubcovsky, 1999). However, vernalization in hexaploid wheat (*T. aestivum*) is controlled by the *VRN1* locus (Law *et al.*, 1975; Tranquilli and Dubcovsky, 1999). *VRN1* is closely linked to MADS-box genes *AP1* and *AGLG1* (similar to *Arabidopsis* meristem genes *AP1* and *AGL2*, respectively) in a 0.3-cM interval flanked by genes *Cysteine* and *Cytochrome B5*. *AP1* is a more likely candidate for *VRN1* than *AGLG1* (Yan *et al.*, 2003). *VRN2* has expression patterns opposite to that of *VRN1*, and is located 0.04 cM from *ZCCT1*, the most likely candidate gene for *VRN2* (Yan *et al.*, 2004a). *Fusarium* head blight (FHB) is a devastating disease of wheat worldwide. Waldron *et al.* (1999) detected a major QTL, *Qfhs.ndsu-3BS*, contributing to FHB resistance in Sumai 3 and located in the deletion bin 3BS (Liu and Anderson, 2003). When constructing a fine genetic map of the *Qfhs.ndsu-3BS* region that spanned 6.3 cM, Liu *et al.* (2006) placed *Qfhs.ndsu-3BS* into a 1.2-cM region flanked by STS3B-189 and STS3B-206, and redesignated it as *Fhb1*.

Only five major QTL differentiate maize from teosinte (Doebley and Stec, 1993). Just two QTL confer the major morphological differences between maize and teosinte, which have been dissected into single Mendelian loci: *teosinte branched1* (*tb1*) (Doebley *et al.*, 1995, 1997; Wang *et al.*, 1999) and *teosinte glume architecture* (*tga1*) (Dorweiler *et al.*, 1993; Wang *et al.*, 2005b). The gene *tb1* suppresses lateral branching (leading to apical dominance), whereas *tga1* affects the hardness of the seed testa (hard casing that envelops the seed in its ancestor teosinte); both the genes were important in the evolution of teosinte to the agronomically suitable maize crop. *Vgt1* is a QTL involved in the control of the transition of the apical meristem from the vegetative to the reproductive phase (flowering) that was initially mapped to a region of 5 cM on chromosome bin 8.05 (Vladutu *et al.*, 1999). Using PCR-based assays for markers flanking *Vgt1* and screening of NIL homozygous for independent crossovers near the QTL, Salvi *et al.* (2002) conclude that *Vgt1* is in a 1.3-cM region between AFLP13 and AFLP14, ca. 0.3 cM away from AFLP 14.

For QTL with small effects, fine-scale mapping and positional cloning will be very difficult in the absence of whole-genome sequence. However, in these cases, reverse genetics may offer a solution, through functional genomic

analysis of candidate genes that underlie QTL. For example, Liu *et al.* (2004b) identified five candidate defense response (DR) genes that colocated with QTL for resistance to blast disease and were associated with level of blast resistance.

QTL pyramiding is an important strategy for rebuilding the outputs from reductionist genomic research into whole traits of value for crop improvement. Once the desirable QTL have been detected, NIL are generated for each QTL in a common elite genetic background, and the effect of each QTL individually evaluated. The selected NIL containing the most important QTL for the target trait are subjected to pair-wise crosses to pyramid two or more QTL for one or more target traits. For example, in rice QTL for increased grain number (*Gn1*) and QTL for reduced plant height [*Ph1*(*sd1*)] were pyramided in the Koshihikari background producing a 23% increase in grain yield while reducing the plant height by 20% compared with Koshihikari (Ashikari *et al.*, 2005).

Dissecting QTL to simple Mendelian factors, often through reduction to component traits, and developing NIL for evaluation, selection, and subsequent use in marker-assisted pyramiding present an effective strategy for molecular breeding of complex traits.

C. Expression QTL Mapping

Traditional genetic mapping has largely focused on the identification of loci affecting one, or at most a few, complex traits. Dissection of the genetics underlying gene expression combines large-scale microarray analyses of expression profiles and conventional QTL mapping of the same segregating population. In this analysis, the expression profiling is considered a quantitative phenotype affected by multiple genes and environmental factors (Jansen and Nap, 2001). This approach has facilitated the identification of genomic regions [gene expression QTL (eQTL)] associated with transcript variation in coregulated genes and, when correlated with phenotypic data from a quantitative character, has successfully identified candidate genes by colocalizing gene eQTL and trait QTL (Brem *et al.*, 2002; Klose *et al.*, 2002; Rockman and Kruglyak, 2006; Schadt *et al.*, 2003; Wayne and Mclntyre, 2002).

The power of a genetic mapping study depends on the heritability of the trait, the number of individuals included in the analysis, and the genetic dissimilarity among them. In experiments involving microarrays and complex physiological assays, phenotyping can be expensive and time consuming and may impose limits on the sample size. A random selection of individuals may not provide sufficient power to detect linkage until a large sample size is reached. Jin *et al.* (2004) developed an algorithm for selecting a subset of

individuals solely on the basis of genotype data that can substantially improve sensitivity compared to a random sample of the same size. The selective phenotyping method involves preferentially selecting individuals to maximize their genotypic dissimilarity while also representing phenotyping extremes. Selective phenotyping is most effective when prior knowledge of the genetic architecture allows us to focus on specific genetic regions. However, it can also provide modest improvements in efficiency when applied on a whole-genome basis. Selective phenotyping does not reduce the efficiency of mapping as compared to a random sample in regions that have not been exposed to strong selection pressure. In contrast to selective genotyping, inferences based solely on a selectively phenotyped population of individuals are representative of the whole population.

Kendziorski *et al.* (2006) demonstrated the deficiencies of using conventional single or multiple QTL analyses for the eQTL approach. Instead, they proposed a mixture over markers (MOM) model that shares information across both markers and transcripts. Results from simulation studies indicate that the MOM model is the best at controlling false-positive associations without sacrificing power of detection. Plants exhibit massive changes in gene expression during morphophysiological and reproductive development as well when exposed to a range of biotic and abiotic stresses. These have been observed as differences in transcriptional profiles in rice (Bao *et al.*, 2005; Matsumura *et al.*, 2003; Rabbani *et al.*, 2003; Tang *et al.*, 2005; Wasaki *et al.*, 2006; Yang *et al.*, 2004; Zhu *et al.*, 2003), maize (Kollipara *et al.*, 2002; Yu and Setter, 2003; Zinselmeier *et al.*, 2002), wheat (Gulick *et al.*, 2005; Wilson *et al.*, 2004), barley (Ozturk *et al.*, 2002; Ueda *et al.*, 2002, 2004; Walia *et al.*, 2006), chickpea (Boominathan *et al.*, 2004), potato (Nielsen *et al.*, 2005; Rensink *et al.*, 2005), banana (Coemans *et al.*, 2005), and cassava (Fregene *et al.*, 2004). Variation in transcript abundance is now being associated with gene expression using eQTL analysis in an increasing number of crops. For example, Kirst *et al.* (2004) dissected the genetic and metabolic network underlying variation in growth in an interspecific backcross population of eucalyptus. QTL analysis of transcript levels of lignin-related genes showed that their mRNA abundance is regulated by two genetic loci, coordinating genetic control of lignin biosynthesis. These two loci colocalize with QTL for growth, suggesting that the same genomic regions are regulating growth, and lignin content and composition. Using a high-density oligonucleotide array and phenotypically divergent rice accessions and their transgressive segregants, Hazen *et al.* (2005) measured the expression of approximately half of the genes in rice (~21,000) to associate changes in stress-regulated gene expression with QTL for osmotic adjustment (OA), which is a known mechanism of drought tolerance. A total of 662 transcripts were observed to be expressed differentially between the parental lines. Only 12 genes were induced in the low OA parent (CT9993) at moderate

dehydration stress levels, while over 200 genes were induced in the high-OA parent (IR62266). Sixty-nine genes were upregulated in all high-OA lines and nine of those genes were not induced in any of the low-OA lines, of which four could be annotated as followings: sucrose synthase, a pore protein, a heat shock protein, and an LEA protein. Previous conventional QTL mapping using the same two rice accessions showed that the parental genotypes differed for five of the OA QTL, that two of these QTL are syntenic with other cereal drought stress QTL (Zhang *et al.*, 2001), and a major OA QTL in the same genomic region on rice chromosome 7 is also reported in a different cross (Lilley *et al.*, 1996). Of the 3954 probes that correspond to this part of the chromosome, few showed a differential expression pattern between the high- and low-OA lines. Thus, these preliminary results demonstrate the power of integrating quantitative analysis of gene expression data with genetic map information to identify genetic and metabolic networks that would not have been identified through conventional QTL analysis.

D. Simulation and Modeling of MAS

Some of the most agronomically and economically important traits in most crops have quantitative phenotypic variation, are under polygenic control, and are significantly affected by the environment. Whole-plant physiology modeling is becoming an increasingly important tool for partitioning complex traits into their components and understanding how those components interact with each other and contribute to the overall trait expression in different environmental conditions. With a commitment to genomic analysis of component traits, whole-plant physiology modeling provides a critical link between molecular genetics and crop improvement. Crop models with generic approaches to underlying physiological processes (Wang *et al.*, 2002) provide a means to link phenotype and genotype, through simulation analysis, of an *in silico* or virtual plant (Tardieu, 2003). In this way, it is possible to dissect the physiological basis of adaptive traits and determine their control at whole-plant level through modeling, and then to use simulation analysis as a predictive decision-support tool for molecular breeders. The substantial progress in ''omics'' technologies for high-throughput data generation allows researchers to create comprehensive datasets on the mechanisms underlying plant growth and plant responses to perturbation. A plant requires information about its environment and interaction with that environment and uses that information to dictate its adaptive responses that result in the plant phenotype. Significant endeavors in the field of whole-plant modeling are now being directed at understanding genetic regulation and aiding crop improvement (Chapman *et al.*, 2002, 2003; Cooper *et al.*, 2002; Hammer *et al.*, 2002; Wang *et al.*, 2004, 2005c; Yin *et al.*, 2003, 2004).

QTL mapping allows the dissection of a phenotype into underlying genetic factors, but it has limited ability to predict how QTL detected in one set of environmental factors or management practices will behave in a new set of conditions (Stratton, 1998). Ecophysiological modeling provides an insight into the factors influencing GEI (Tardieu, 2003), but it does help define the genetic basis for differences in response to environmental changes. Combining ecophysiological modeling with genetic mapping provides the opportunity for creating a QTL-based crop physiology model that could be powerful tool for resolving the genetic basis of complex environment-dependent yield-related traits. For example, using this approach, researchers predicted specific leaf area in barley (Yin *et al.*, 1999), stay-green response to nitrogen in sorghum (Borrell *et al.*, 2001), leaf-growth response to temperature and water deficit in maize (Reymond *et al.*, 2003), and preflowering duration in barley (Yin *et al.*, 2005). Hammer *et al.* (2005) explored whether physiological dissection and integrative modeling of complex traits could link the complexity of the phenotype to underlying genetic systems in a way that could enhance the power of molecular breeding strategies in sorghum. This approach was applied to four key adaptive traits (phenology, osmotic adjustment, transpiration efficiency, and stay-green) using 547 location-season combinations and 4235 genotypic expression states derived from allelic variation at 15 loci for each of the 547 environments. The environmental characterization and physiological knowledge helped to dissect and explain gene and environment context dependencies in the data and based on estimated gene effects to simulate a range of MAS breeding strategies. By removing gene and environment context dependencies, it was possible to devise breeding strategies that generated an enhanced rate of yield improvement over several cycles of selection. Similarly, Messina *et al.* (2006) combined an ecophysiological model (CROPGRO-Soybean) with a linear model that predicted cultivar-specific parameters as function of E-loci. This approach predicted 75% of the variance in time to maturity and 54% of the variance in yield. This demonstrates that agricultural genomics data can be effectively used for predicting cultivar performance and refining crop breeding systems.

Innovative simulation models bridge the gap between molecular and conventional plant breeding and will inform both strategic research and tactical breeding decisions (www.generationcp.org/sccv10/sccv10_upload/modelling_links.pdf). The CGIAR Generation Challenge Program (GCP) is supporting several projects on whole-plant physiology modeling, QTL $\times$ E analysis, and simulation of molecular breeding programs that will collectively link physiological and genetic models toward the optimization of marker-assisted breeding systems for drought tolerance in cereals. Simulation models integrate molecular information about interaction between genes and simpler traits to allow realistic predictions for more complex traits such as drought tolerance and yield. QuGene software platform (Podlich and Cooper, 1998;

http://www.uniquest.com.au) defines gene effects and builds breeding modules to compare breeding efficiencies. For example, using QuGene software, researchers in Australia developed a breeding module for sorghum incorporating physiological constraints that were implemented by linking QuGene to the Agricultural Production System Simulator (APSIM) cropping systems model (Keating *et al.*, 2003; http://www.apsru.gov.au), thus providing a powerful set of programs that can simulate crop breeding line performance in a given environment and extrapolate the effects of long-term selection over many breeding cycles and seasons. Another GCP supported project links QuGene/APSIM with QTL data on maize leaf growth under drought. These projects aim to deliver modeling tools into the hands of molecular breeders and other researchers to extend the scope and impact of their use, particularly with respect to molecular breeding of complex traits such as drought tolerance.

Developing and implementing a design-led breeding system for complex traits require enhanced attention to precision phenotyping, ecophysiological modeling, and marker validation to ensure robustness and selective power. These approaches require the iterative and systemic integration of a range of scientific disciplines, including modelers, physiologists, geneticists, breeders, and molecular biologists. Nevertheless, the first preliminary studies reviewed in this section suggest that a new paradigm in knowledge-led design-driven plant breeding is a feasible option and that for the first time genomics may finally realize its potential impact on breeding complex traits is increasingly likely.

VIII. THE ROLE OF COMPUTATIONAL SYSTEMS IN MOLECULAR BREEDING PROGRAMS

Effective marker-aided breeding requires the balance of many diverse elements in order to provide the best compromise between time, cost, and genetic gain:

- Identify beneficial genetic variation and develop robust marker-trait associations
- Effectively manage and manipulate large amounts of genotype, pedigree, and phenotype data
- Select desirable recombinants through an optimum combination (in time and space) of phenotypic and genotypic data
- Develop breeding systems that minimize population sizes, number of generations, and overall costs while maximizing genetic gain for traditional and novel target traits

In general, MAS works best with simply inherited markers that are inside or flanking markers that are in proximity to the genetic factors affecting monogenic, oligogenic, and polygenic traits. The journey from the phenotyping-and-genotyping of individuals from genetic populations to the identification of marker-trait associations and onto the application of markers in molecular breeding depends on the sequential use of a number of decision-support tools that facilitate communication between genomics scientists, geneticists, bioinformaticians, trait specialists, and breeders. In this section, we provide an overview of key decision-support tools for assisting germplasm evaluation, breeding population management, GEI, genetic map construction, marker-trait linkage and association analysis, marker-assisted application, breeding system design and simulation, information management, and other integrated tools needed to support molecular breeding programs (Table XVI).

A. Germplasm Evaluation

Marker-assisted germplasm evaluation (MAGE) aims to complement phenotypic evaluation by helping define the architecture of genetic resources and by identifying germplasm that contains alleles associated with traits of economic importance. Molecular markers can be used to for characterization based on genes, genotypes, or genomes, which provide more accurate and detailed information than classical phenotypic or passport data. Many features revealed by molecular markers, such as unique alleles, allele frequencies, and heterozygosity at marker loci, mirror the genetic structure of germplasm resources and will lead to the identification of useful genes and their transfer into well-adapted cultivars. MAGE will play an important role in acquisition, distribution, maintenance, and use of germplasm (Bretting and Widrlechner, 1995; Xu, 2003). During germplasm evaluation, molecular markers can be used to (1) differentiate cultivars and construct heterotic groups; (2) identify germplasm redundancy, underrepresented alleles, and genetic gaps in current germplasm collections; (3) monitor genetic shifts that occur during germplasm domestication, storage, regeneration, and breeding; (4) screen germplasm for novel and/or superior genes or alleles; and (5) construct a representative subset or core collection (Xu *et al.*, 2003, 2004a). Although computational programs are available for all relevant analyses including computer simulation and resampling (Xu *et al.*, 2004a), a fully integrated, user-friendly graphical program is needed to bring all these functions together to facilitate decisions through all aspects of germplasm evaluation.

Several software packages, such as Statistica, JMAP, SAS, NTSYS, GeneFlow, can be used for the analysis of germplasm evaluation data. This

Table XVI
List of Decision Support Tools to Support Molecular Breeding Programs

Tool	Function	References
Germplasm evaluation		
JMAP/SAS	Clustering, PCA	http://www.sas.com/
Structure	Identify distinct populations and estimate allele frequencies	Pritchard *et al.*, 2000a
GGT (Graphical GenoTypes)	Transform marker data into simple colorful chromosome drawings	van Berloo, 1999
GERMPLASM	Classify cultivars and construct heterotic groups; identify germplasm redundancy, underrepresented alleles, and genetic gaps; monitor genetic shifts; screen for novel/superior genes (alleles); construct a representative subset or core collection	Xu *et al.*, 2004a
Breeding population management		
Hybrid performance prediction	BLUP-based methods	Bernardo, 1994, 1996
Genetic map construction		
MAPMAKER/EXP	Build linkage map from molecular marker data	Lander *et al.*, 1987
MAPDISTO	Build linkage map from molecular marker data with distorted segregation	http://mapdisto.free.fr/
MAP MANAGER CLASSIC	A graphic, interactive program for linkage map construction	Manly, 1993
JOINMAP	Combine data derived from several sources into an integrated map	Van Ooijen and Voorrips, 2001
GMendel	Linkage mapping using simulated annealing and multiple pair-wise methods for F_2, BC, DH, RIL, and any generations of SSD	http://www.maizegdb.org/mnl/66/45echt.html
Genotype-phenotype association		
MAPMAKER/QTL	Map QTL using interval mapping, dealing with simple QTL and several standard populations	Lander *et al.*, 1987
MAP MANAGER QT	A graphic, interactive program for QTL mapping by regression methods	Manly and Olsen, 1999
MAP MANAGER QTX	A graphic, interactive program for QTL mapping using intercross, BC or RIL in plants or animals	Manly *et al.*, 2001

(*continued*)

Table XVI (*continued*)

Tool	Function	References
QTL Cartographer	QTL mapping using several interval mapping methods with permutation tests to estimate QTL thresholds	http://statgen.ncsu.edu/qtlcart/cartographer.html
PLABQTL	Identifying QTL using composite interval mapping and QTL × environment interaction analysis	Utz and Melchinger, 1996
QTL EXPRESS	QTL mapping in outbred populations including line crosses, half-sib families, nuclear families and sib-pairs, with permutation tests to determine empirical significance levels and boots-trapping to estimate empirical confidence intervals of QTL locations	Seaton *et al*., 2002
MapPop	Identify QTL using selective and bin mapping by choosing good samples from mapping populations and for locating new markers on preexisting maps	Vision *et al*., 2000
MCQTL	QTL mapping using multicross designs	Jourjon *et al*., 2005
EPISTACY	A SAS program to test for all possible two-locus interaction effects on a QTL using least squares methods	Holland, 1998
STRAT	Association mapping with incorporated function for structure analysis	Pritchard *et al*., 2000b
TASSEL	A comprehensive software for trait analysis by association, evolution, and linkage, including association mapping, diversity estimation and calculating linkage disequilibrium	Zhang *et al*., 2006b
BQTL (Bayesian QTL mapping)	Maximum likelihood estimation of multigene models; Bayesian estimation of multigene models via Laplace Approximations; and interval mapping and composite interval mapping of genetic loci	Borevitz *et al*., 2002
MAS		
Plabsim	MAS simulation for all common breeding methods. Selection can be carried out at defined loci or for selection indices calculated from allele frequencies at several loci. The simulated data can be analyzed for genetic parameters such as population size, marker density and positions, and selection strategies	Frisch *et al*., 2000
Popmin	Numerical optimization of population sizes in marker-assisted backcross programs	Hospital and Decoux, 2002

BCSIM	Simulation for evaluation of marker-assisted backcross programs	http://www.dpw.wau.nl/pv/pub/bcsim/index.htm
Breeding design and simulation		
QU-GENE (QUantitative GENEtics)	Simulation platform for quantitative analysis of genetic models including genotype by environment interaction analysis	Podlich and Cooper, 1998
QuCim	Identify the best crosses and breeding strategies from mass selection, pedigree system, bulk population system, backcross breeding, top cross (or three-way cross) breeding, DH breeding, MAS, and many combinations and modifications of these methods	Wang *et al.*, 2004
QuLine	Define genetic models from simple to complex based on simulation experiments to optimize breeding programs and improve breeding efficiency	Wang *et al.*, 2004
Information management and integrated tools		
CMTV	Display syntenic regions across taxa, combine maps from separate experiments into a consensus map, or project data from different maps into a common coordinate framework using dynamic coordinate translations between source and target maps	Sawkins *et al.*, 2004
QTLFinder	Integrate QTL and linkage maps into a consensus map; do QTL meta-analysis and show colocations; construct comparative map of interspecies (or intraspecies) genomes; and compare collinearity of same or similar traits across genomes	Yan *et al.*, personal communication
ICIS (International Crop Information System)	Kink the gene, gene value, and target environment data with the uniquely identified germplasm units used and manipulated in breeding programs. It has ICIS as the Genealogy Management System (GMS) to manage data on nomenclature, origin, development and deployment of germplasm and the Data Management System (DMS) to manage and document characterization and evaluation data	http://www.icics.cgiar.org:8080/
iMAS (integrated decision support system for marker-assisted plant breeding)	Facilitate an integrated, error-free, and appropriate data analysis from the beginning to end of the molecular breeding pathway, including experimental design, biometric analysis of phenotypic data, linkage and association mapping, linkage map construction, and MAS	http://www.generationcp.org/vw/Download/Commisioned_Research_2005/33_SP4_MAS.pdf#search ='iMAS%20marker%20assisted%20selection'

includes the use of principal component or coordinate analysis to identify distinct groups or populations, and for cluster or structure analysis to define population structure. For example, STRUCTURE (Pritchard *et al.*, 2000a) uses multilocus genotype data to investigate population structure, assign individuals to populations, study hybrid zones, identify migrants and admixed individuals, and estimate population allele frequencies in situations where many individuals are migrants or admixed. It can be applied to datasets from most of the commonly used genetic markers, including SSR, RFLP, and SNP.

B. Managing Breeding Populations

Decision-support tools to help the management of breeding populations are needed to assist in the choice of parental lines, types of crosses, and nature of breeding system. Computational tools may also assist in the establishment and maintenance of heterotic groups, the selection of lines for creation of a synthetic cultivar, the prediction of progeny and hybrid performance, and the monitoring of genomic profiles during population improvement.

Genotyping parental lines on a genome-wide scale, especially when gene-based markers are available, provide an opportunity for establishing parent–hybrid performance relationships at the molecular level. Genome-wide heterozygosity and specific combinations of alleles (linkats) may be useful determinants in some crops for maximizing heterosis and hybrid vigor. Melchinger and Gumber (1998) used a multistage procedure to identify heterotic groups, which consists of the following steps: (1) grouping the germplasm based on genetic similarity, (2) selection of representative genotypes (e.g., two or four lines or one population) from each subgroup for producing diallel crosses, (3) evaluation of diallel crosses among the subgroups together with parents, and (4) selection of the most promising cross combinations as potential heterotic patterns.

The ability to use molecular markers to predict hybrid performance would greatly enhance the efficiency of hybrid breeding programs. Development of a reliable method for predicting hybrid performance or heterosis without generating and testing hundreds or thousands of single cross combinations has been the goal of numerous studies using marker data and combinations of marker and phenotypic data, particularly in maize and rice. The best linear unbiased prediction (BLUP) procedure has been used for decades for evaluating the genetic merit of animals, especially dairy cattle. Intrapopulation, additive genetic models have traditionally been used for BLUP in animal breeding (Henderson, 1975). Bernardo (1994, 1996) used BLUP in maize breeding with interpopulation genetic models that involve both

general combining ability and specific combining ability and found that BLUP is useful for routine prediction of single-cross performance. The predicted performance of single crosses may subsequently be used to predict the performance of F_2 × tester combinations, three-way crosses, or double crosses. Along with the pedigree relationship, BLUP can use trait data, or both trait and marker data, for prediction.

A synthetic cultivar is developed by intercrossing selected clones or inbred lines, with seed production of the cultivar through open-pollination. MAS can be used to develop synthetic cultivars by mixing inbred lines that have been bred by MAS or by mixing individual plants derived from any stage of MAS. With genotypic information available across the whole genome for all the selected individuals or inbred lines, synthetic cultivars can be created to contain complementary genotypes, fixed heterozygosity, and the best combinations of genetic structure.

C. Genetic Map Construction

Genetic maps can be constructed using segregating populations of different types, which have different advantages depending on the species and level of polyploidy. MAPMAKER/EXP is the most frequently used software for map construction (Lander *et al*., 1987). Various maps can be generated based on populations derived from different crosses or the same population evaluated in different environments. These maps can be integrated into a single or consensus map. JOINMAP is used to construct genetic linkage maps for several types of mapping populations. It can combine (join) data from several sources into an integrated map, with several other functions, including LG determination, automatic phase determination for outbred full-sib family, several diagnostics, and map charts (Van Ooijen and Voorrips, 2001). GMendel uses simulated annealing and multiple pair-wise methods for locus ordering. All markers within an LG are used simultaneously to estimate a locus order that provides maps equivalent to those found by MAPMAKER and JOINMAP. It can be used to build maps using F_2, backcross, DHL, RIL, and in any generation of SSD lines. Other software packages in use are MAPDISTO (http://mapdisto.free.fr/) and MAP MANAGER CLASSIC (Manly, 1993) that perform specific functions.

D. Identifying Marker-Trait Associations

Establishing a highly significant genotype–phenotype association is one of the prerequisites for MAS. Linkages or associations between target traits or genes and molecular markers are detected based on genetic linkage or

association mapping experiments. Decision-support tools required for genotype–phenotype association include (1) statistical methods and tools to establish, validate, and compare genotype–phenotype associations through linkage mapping, LD, or AM, and *in silico* mapping, using single or multiple genetic populations, genetic resources, or breeding populations; (2) statistical methods and tools for identification of genetic background effects, QTL alleles at multiple loci, and multiple alleles at a single locus; (3) tools facilitating the validation of candidate gene markers with linked markers in order to generate functional markers; and (4) tools facilitating management of genetic populations, linkage maps, and related data. A widely used QTL mapping software is QTL Cartographer (http://statgen.ncsu.edu/qtlcart/cartographer.html), which implements several statistical approaches to analysis of multiple marker data including composite interval mapping (CIM) and multiple interval mapping. The interaction between different QTL can also be estimated. Another populated QTL mapping software is PLABQTL that uses CIM with many functions common to those of QTL Cartographer. QTL can be localized and characterized in populations derived from a biparental cross. Simple interval mapping (SIM) and CIM are performed using a fast multiple regression procedure. PLABQTL can also be used to analyze QTL × environment interactions (Utz and Melchinger, 1996).

For mapping with populations from outbreeding species, QTL EXPRESS can be used to map QTL using line crosses, half-sib families, nuclear families, and sib-pairs (Seaton *et al.*, 2002). EPISTACY is a SAS-based program which can test pair-wise epistatic (interaction) effects on a quantitative trait using QTL-mapping datasets (Holland, 1998). Other softwares for mapping QTL include MAPMAKER/QTL (Lander *et al.*, 1987), MAP MANAGER QT (Manly and Olsen, 1999), MAP MANAGER QTX (Manly *et al.*, 2001), MapPop (Vision *et al.*, 2000), and MCQTL (Jourjon *et al.*, 2005).

Software packages are also now available for mapping genetic traits using Bayesian approaches. For example, BQTL performs (1) maximum likelihood estimation of multigene models, (2) Bayesian estimation of multigene models via Laplace approximations, and (3) interval mapping and CIM of genetic loci (Borevitz *et al.*, 2002), while BLADE is used for Bayesian analysis of haplotypes for LD mapping (Liu *et al.*, 2001b; Lu *et al.*, 2003).

AM or LD mapping, using unstructured populations, is gaining increasing credibility over traditional QTL mapping using genetic populations (see Section II.D). However, softwares are needed that analyze and remove the effect of population structure. STRAT uses a structured association method for AM, enabling valid case-control studies even in the presence of population structure (Pritchard *et al.*, 2000b). The software TASSEL has been released, which performs a variety of genetic analyses, including AM, diversity estimation, and LD analysis (Zhang *et al.*, 2006b). The association

analysis between genotypes and phenotypes can be performed by either a general linear model or a mixed linear model. The general linear model allows users to analyze complex field designs, environmental interactions, and epistatic interactions. The mixed model is especially designed to handle polygenic effects at multiple levels of relatedness, including pedigree information. These new analyses should permit association analysis in a wide-range plant and animal species.

E. Marker-Assisted Selection

Many factors influence the efficiency of MAS in plant breeding programs (see Section VI.A and B). Decision-support tools are needed to determine sample size for foreground and background selection, for estimation of genetic gains (response to selection), for construction of selection indices for multiple traits and whole-genome selection, for estimation and graphical display of RGC of selected individuals at each generation of introgression, for identification of desirable plants based on both phenotype and genotype information, for cost-benefit analysis, and for marker-aided simulations studies.

There has been much interest in the development of software that simulates MAS using genetic models. Early efforts had somewhat limited results, for example, GREGOR simulates MAS based only on predefined genetic linkage maps, and is thus restricted in its value for simulation of MAS in breeding programs (Tinker and Mather, 1993). More recently, Plabsim was developed for the simulation of MAS programs, with the following features: (1) simulations can be made for any diploid genome with an arbitrary number of loci at arbitrary positions on an arbitrary number of chromosomes; (2) the implemented reproduction schemes include all common breeding methods; (3) an arbitrary number of selection steps can be combined with a specified selection strategy and selection can be carried out for genotypes at defined loci, or for selection indices calculated from allele frequencies at several loci; and (4) the simulated data can be analyzed for a broad range of genetic parameters including population size, marker density and positions, and selection strategies on the genetic composition of the breeding product and on the required number of marker data points (Frisch *et al.*, 2000). Other software packages related to MAS include Popmin for the numerical optimization of population sizes in marker-assisted backcross programs (Hospital and Decoux, 2002), GGT for displaying molecular marker data into simple colorful graphical representations of chromosome haplotypes (van Berloo, 1999), and BCSIM for evaluation of marker-assisted backcross programs (http://www.dpw.wau.nl/pv/pub/bcsim/index.htm).

F. GEI Analysis

Computational tools are needed to assist in dealing with many complex issues related to the effect of the environment, particularly regarding complex traits, including:

- To separate genetic (G) effects from the environment (E) and GEI interaction
- To incorporate environmental and genotypic variables into statistical models to explain GEI
- To define target populations and genotypes for a given environment
- To determine subsets of genotypes and sites with negligible crossover effects to identify subgroups of sites and genotypes with similar response to maximize response to selection
- To develop selection indices using phenotypic and marker data to select the best genotypes
- To study genetic diversity of crop genotypes associated with the target traits and perform AM
- To study gene expression under target conditions using microarray technology

Podlich and Cooper (1998) developed QU-GENE software for carrying out quantitative genetic analyses of GEI in crop breeding and this has become an increasingly widely utilized decision-support tool in breeding programs. Statistical models have been refined in order to incorporate pedigree information (or coefficient of parentage) among genotypes when modeling GEI (Crossa *et al.*, 2006). It is likely that these will soon be further refined using whole-plant physiology models.

G. Breeding Design and Simulation

The major objective of plant breeding programs is to develop new cultivars superior to those currently available in a given target production environment (TPE). Designing effective breeding systems requires information about target genes, donor germplasm, and proposed elite recurrent parents. This can then be combined with evaluation data on the target biological characteristics, breeding objectives for the TPE, in order to optimize the breeding procedure and selection methods through modeling and simulation analysis. This type of analysis will also predict the desirable target genotype and the probability of successfully generating new cultivars through the proposed breeding system. QU-GENE, a simulation platform based on quantitative genetic models, facilitates the simulation of actual breeding programs through its two-stage process (Podlich and Cooper, 1998).

The genetics and breeding simulation tool (QuLine and QuCim) has the potential to utilize vast and varied genetic information. QuLine is capable of defining genetic models ranging from simple to complex inheritance. QuCim can be used to identify the best crosses and breeding strategies by predicting cross performance and comparing different selection methods. Using simulation experiments, breeders may optimize their breeding programs and thereby greatly improve the breeding efficiency (Wang *et al.*, 2004). Almost all efforts in this field have been focused on genetic models, thus none provides the facility to carry out such as cost benefit analysis or integrate whole-plant physiology models.

H. Information Management and Integrated Tools

Crop informatics has become a prerequisite in molecular breeding because breeding-related information is increasing at such a high rate that collecting, storing, mining, and manipulating such a large amount of information for selection decisions would not be possible without appropriate statistical, biometrical, and informatics tools. An integrated breeding tool is therefore needed to rapidly collect, analyze, and represent breeding-related data in the short-time window available for most selection decisions. In addition, computational tools are required to translate and integrate research outputs into a usable form for plant breeding programs.

International Crop Information System (ICIS) is open-source community developed software that has been evolving over many years. ICIS can link gene, gene value, and target environment data with the uniquely identified germplasm units used and manipulated in breeding programs (http://www.icis.cgiar.org:8080/). ICIS has a modular structure with a core consisting of Genealogy Management System (GMS) that manages data on nomenclature, origin, development, and deployment of germplasm and the Data Management System (DMS) that manages and documents characterization and evaluation data. Specialized user interfaces deliver data views and decision-support tools to crop scientists from different disciplines, which can access common data resources leading to efficient use and reuse of research data. ICIS databases tailored to different crops are also being developed for separate ICIS implementations. ICIS has also embedded a parallel structure of central and local versions that provides local read/write capabilities, allowing data generated locally to be merged and harmonized with the central database at the local user's discretion. Some of the issues that need to be further integrated into ICIS to meet breeding requirements include: (1) a database for all environmental characterization data such as climate, soil, and abiotic stress information; (2) data-mining tools for all breeding purposes such as GEI and identification of novel alleles

and genetic variation; (3) modeling breeding processes and selection schemes using multiple sources of breeding information to eliminate some field and laboratory tests required for making selection decision, which may be critical for complex traits; and (4) linkage to major public databases with appropriate data comparison and mining tools to enable extraction of useful information through comparative analysis of the specific breeding program data with global research outputs.

Researchers need efficient and intuitive tools to help identify common genomic regions, and, where possible, specific genes involved in influencing the expression of target traits across diverse germplasm and growing conditions. Sawkins *et al.* (2004) developed the comparative map and trait viewer (CMTV) that can help integrate various kinds of genomic maps. Its major strength is in the comparative display of LGs or chromosomes across different species, populations, or evaluation environments and link information associated with different objects on the maps. These correspondences could then be displayed as graphical lines linking corresponding loci between maps in order to illustrate syntenic relationships. Alternatively, they could be used to construct a consensus map using these common markers as anchors from which the positions of other markers could be interpolated. However, the current version of this software stops short of being able to carry out combined analysis across the maps to be compared. In contrast, QTLFinder can carry out this type of QTL meta-analysis. This software integrates QTL from separate experiments and linkage maps into a consensus map. QTLFinder can also construct comparative maps across species using sequence similarity, and compare the colinearity of same or similar traits across genomes (Jianbing Yan, China Agricultural University, Beijing, personal communication).

An integrated decision support system for marker-assisted plant breeding iMAS (a GCP-supported software) will be released by the end of 2006 (Subhash Chandra, ICRISAT, personal communication), is expected to assist the development and application of marker-assisted plant breeding by integrating the best freely available quality software required for the journey from phenotyping-and-genotyping of individuals to identification and application of trait-linked markers. iMAS will provide simple-to-understand-and-use online decision-support guidelines to help the user correctly use this software, and correctly interpret the outputs. Software identified for inclusion in iMAS includes IRRISTAT (for experimental design, biometric analysis of phenotypic data, and AM), GMendel and MapDisto (for linkage map construction), PlabQTL and QTL Cartographer (for QTL analysis), PopMin (for estimating sample size for foreground and background selection), GGT (for estimation and display of RGC of selected individuals), and TASSEL (for AM).

Many support tools are available for use with functional genomic data, but these are yet to be fully explored for direct application in breeding

programs. These tools are for sequence comparison, handling and analysis, microarray data treatment and analysis, motif alignment and search, and comparative genomics. Various softwares such as EHAP (http://wpicr.wpic.pitt.edu/WPICCompGen/ehap_v1.htm), DPPH (Bafna *et al.*, 2003), HAPLOVIEW (Barrett *et al.*, 2005), HAPLOT (Gu *et al.*, 2005), HAP (1) (http://research.calit2.net/hap/), and HAP (2) (Zhao, 2004) have been developed for haplotype analysis using SNP data. It is likely that these approaches will soon be widely used by molecular breeders across diverse crops as sequence and expressional data become increasingly available.

IX. FUTURE PROSPECTS FOR THE MOLECULARIZATION OF PUBLIC CROP IMPROVEMENT

Plant breeding is the science, art, and business of improving plants for human benefit
(Bernardo, 2002)

The rate, scale, and scope of uptake of genomics in crop breeding programs have continually lagged behind expectations. This is little different to the adoption of quantitative genetics, mechanization, and computerization during the last century. This is partly due to the long product development cycle in plant breeding and in turn the long-term nature of feedback from the market regarding the impact of any changes in the cultivar development pipeline. Thus, although molecularization of plant breeding is the fourth natural paradigm shift for crop improvement programs, we must assume that the introduction of MAS and the breeding with transgenic germplasm will be a gradual stepwise process. At the same time, there is considerable and immediate need for computational tools to help breeders more effectively translate and integrate the outputs from bioscience research and to help efficiently select the best technology interventions and associated breeding systems for their target traits and markets. With the availability of comprehensive and robust facilitating and decision-support tools, it is expected that plant breeders will become much more responsive to the emergence of new technologies.

Polymorphic DNA markers and genetic maps are now available for most important food crops, albeit in varying numbers and levels of genomic saturation (Tables VI–VIII and X; Dwivedi *et al.*, 2005). Similarly, the genetic control of many agronomic traits is well characterized in many crops, and marker-trait linkages have been reported for a diverse array of traits in a large number of crops (Section II). A critical mass of reports of

validation in different genetic backgrounds and environments is naturally only just beginning to emerge (Section III; Tables XI and XII). Nevertheless, MAS is now being practiced in most well-researched crops (Section IV.A–D; Tables XIII and XIV). In the private sector, molecular breeding applications are still dominated by MAS for transgene introgression and to a lesser extent for backcross programs for simple traits (Section V). Thus, only a very small proportion of marker genotyping is currently being used for complex traits where it has been long since highlighted that MAS will have its greatest impact. In the short term, we expect the greatest growth in MAS of mono- and oligogenic traits that are difficult or expensive to screen using conventional phenotyping methods (Section VI). In the medium term, we envisage that a number of emerging technologies will facilitate a gradual shift from MAS for individual simply inherited traits to more holistic molecular breeding strategies (Section VII). It is only at this point that we expect to see a significant increase of interest in the application of MAS for polygenic traits. However, there are a number of technical and logistical hurdles that must be overcome before genomic tools can assist the breeding of such complex targets.

Traditionally, the heritability of quantitative traits was the most common predictor of genetic gains for different plant breeding methods. DNA markers may be used today to accelerate and enhance overall breeding methods by combining DNA marker and phenotyping data in a selection index. The best current success stories of MAS in plant breeding tend to focus on traits that are difficult to screen and controlled by one or few genes. However, more recently there have been a number of successes in pyramiding a range of difference sources of biotic and abiotic stress resistances (Table XIV). This engenders hope for the potential of MAS to improve important quantitative traits, particularly when accelerating the use of new sources of variation in elite germplasm. DNA markers will also be useful tools for early testing. However, geneticists and plant breeders will still need to deal with LD while using MAS in recurrent selection, especially when using polymorphic markers arising from mapping populations, which tend to be from diverse parents, and thus may not be relevant for target breeding materials. The power of MAS will also continue to rely heavily on the accuracy and precision of phenotyping, and the characterization and evaluation of germplasm in the field. Issues such as the error term to test for the significance of a QTL, detecting small effects with narrow genetic variance, or the number of QTL not related to genetic variance or divergence of parents are all under-researched areas that need priority attention by geneticists. Addressing these issues will allow plant breeders to define the optimum number of individuals/lines and markers to be used in their MAS programs.

Plant breeders are ready to apply MAS for quantitative traits when the genetic gain and time or cost efficiency from doing so are clearly higher than

through conventional selection methods. Initial emphasis in this area should be on traits for which a robust cost-effective phenotyping system is not available for the target trait. To quickly reach this stage requires a paradigm shift in strategy among the marker-trait identification community: from efforts to identify all QTL influencing the target trait to a focus on identification of a few QTL having the largest effect on the target trait. QTL of major effect may be easier to detect (in the right genetic material), and be less influenced by GEIs and genetic background effects. Of great importance will be a shift away from analysis of entire genetic populations to an emphasis on selected individuals with extreme phenotypes from relevant breeding populations and genetic stocks and likely, pooled DNA analysis using the selected individuals. Of equal importance will be a shift from linked markers to diagnostic gene-based markers, which will generally be SNP based and thus readily scalable for high-throughput haplotyping. Detailed cost-benefit analysis of various elements of DNA marker development and application, including the cost of the required genotyping platforms and professional expertise, needs to be assessed at the earliest possible stage. This is particularly important at this time when most public plant breeding programs are not adequately funded or poorly equipped to reach a critical threshold of marker assay throughput. Molecular breeding consortia accessing joint venture genotyping hubs or commercial service providers appear to be an increasingly realistic option where those facilities can provide the right quality, quantity, and timelines of service to fit the given breeding system.

In the last decade, computational tools have rapidly evolved to provide solutions for the data acquisition, management, analysis, and visualization needs arising from the development and widespread use of high-throughput genomics technologies. Plant breeders expect that informatics will assist with the development of diagnostic tools for identification of the best breeding systems, optimization of the best crosses, and selection of the best ensuing segregating progeny. Likewise, bioinformatic research should identify causative alleles and estimate breeding values or relative risks in the context of breeding populations. Moreover, besides assisting with candidate genes, bioinformatics should provide plant breeders with information regarding LD and epistatic and pleiotropic effects of the allele in the target breeding population. Statistical methods will assist in estimating and predicting allele effects which should be updated as the alleles are assessed in distinct breeding backgrounds and across other environments. Information on breeding values provided by DNA markers may enable identification of DNA markers for further use in a more robust MAS system.

Geneticists can use DNA markers to dissect complex epitasis effects, which may arise as an outcome of selection-induced variation. For example, a minor or neutral QTL may become a major QTL when selection brings changes that create the most appropriate genetic background for interaction

with that target QTL. The release of genetic variability through capitalizing on epistasis may allow a more extended response to selection than that currently resulting solely from additive variance.

Genotype-by-environment interaction (GEI) occurs when the effects of the environment, the genotype or both, are nonadditive. GEI may lead to divergence, convergence, or crossover performance of genotypes across the environments; that is, the distinct performance among genotypes depends on the environment (location, year, cropping season). Linear mixed models are used for modeling GEI and assisting the grouping of environments and genotypes. Factorial and partial least squares regressions incorporate external environmental and genotypic covariables directly into the model. These are useful tools for gaining more insight into the genetics of the target trait by adding molecular marker data associated with quantitative trait variation in the model for interpreting GEI. With more and more information accumulating from genotyping and phenotyping, integration of these diverse datasets with environmental characterization data will help establish genetic models for GEI and apply them to crop improvement. Molecular markers could further explain some of the GEI variabilities and assist in breeding for low-heritability traits. For example, Paterson *et al.* (1991) suggested that, for a low-heritability trait such as soluble solids in tomato, the phenotype of F_3 progeny could be predicted more accurately from the QTL genotype of the F_3 parent than from the phenotype of the F_2 individual.

Applied genomic tools are being used to unravel the molecular mechanism of heterosis, classifying germplasm with distinct heterotic groups, predicting hybrid performance, understanding the relationships between heterozygosity and genetic distance with hybrid performance and heterosis. All these will lead to a better understanding of the genes regulating the network of diverse physiological pathways that control the expression of hybrid vigor. This will undoubtedly lead to enhanced use of MAS for the development of superior yielding hybrids. So far, various hypotheses have been proposed to explain the genetic mechanisms of heterosis, each being supported to some extent by different experimental data. Considering that heterosis may mediate its effect at various levels and developmental stages for different traits, it is feasible that there is no single genetic model or hypothesis that can be used to explain all heterotic effects observed in hybrids across traits, crops, and breeding systems (Xu, 2003). Molecular markers will provide new insights into heterosis as it becomes feasible to carry out genome-wide analysis of parental lines across large numbers of hybrids, germplasm accessions, and breeding materials.

Plants exhibit massive changes in gene expression during morphophysiological and reproductive development as well as when exposed to a range of biotic and abiotic stresses (Section VII.C). A new field of genetics of global gene expression has emerged based on the application of traditional

techniques of linkage and association analysis for the thousands of transcripts measured by microarrays. Dissecting the architecture of quantitative traits in this way connects DNA sequence variation with phenotypic variation, and is improving our understanding of transcriptional regulation and regulatory variation (Rockman and Kruglyak, 2006).

A range of decision-support tools are needed to facilitate communication among scientists involved in different elements of the crop improvement product development pipeline. While there are a number of computational tools to carry out various functions in the research domain, it is essential that these tools are integrated into a common platform to assist their effective deployment in crop improvement. iMAS (www.generationcp.org), an integrated decision support systems for marker-assisted plant breeding, is a preliminary attempt to create a publicly available computational platform to assist the development and application of marker-assisted plant breeding. iMAS currently integrates freely available software for the journey from phenotyping-and-genotyping of individuals to identification and application of trait-linked markers. iMAS also provides simple-to-understand-and-use online decision-support guidelines to help the user correctly operate these softwares, and correctly interpret the outputs.

It has been argued that genetically modified food is the next great scientific and technological revolution in agriculture and the only efficient and cheap way to feed a growing population in a shrinking world. Genetic transformation is particularly important for transfer of genes from distant species. In many cases, genetic transformation will be the only mechanism for harnessing the outputs of large-scale whole-genome research, particularly in model systems. At the same time, rapidly accumulating information about crop genomes is allowing scientists to identify genes associated with beneficial traits in ''crop relatives.'' Marker-assisted introgression of these beneficial alleles into existing cultivars will be increasingly critical for efficient use of exotic genetic variation in breeding programs. Thus, the intimate integration of MAS and genetic transformation approaches in field breeding programs will be an important challenge for the future success of public sector crop improvement. Using molecular biology tools and outputs, researchers will be able to broaden the scope of breeding goals, improve the rate and precision of genetic gain toward specific trait targets, and significantly reduce the time needed to breed new cultivars. However, there is still much work to be done in understanding the ''choreography'' of molecular breeding to the extent required to reach a knowledge-led design-based plant breeding paradigm. For example, the relationship between single genetic loci, complex genetic traits, and environmental factors all differentially interact to affect the development of the plant, its response to biotic and abiotic stresses, and ultimately the yield. Over the next decade, MAS technologies will become cheaper and easier to apply at large scale, and knowledge from genomics

research will become more readily translated into breeding tools and integrated into breeding systems. These advances will empower plant breeders around the world to use molecular breeding approaches as part of a much larger systemic and holistic approach to sustainable agricultural development (http://www.washingtonpost.com/wp-dyn/content/article/2006/07/03/AR2006070300922.html).

Plant breeders in the twentieth century accomplished improvements in crop performance through knowledge and application of scientific advances in genetics research. However, a substantial proportion of genetic progress also resulted from pragmatic practice of the art of plant breeding. The crop genetic enhancers in this twenty-first century will harness the outputs of bioscience research (especially genomics) in order to address the challenge of doubling food production sustainable on same land area (1.5 billion ha) by 2050. To substantially contribute to achieving this goal, it will be necessary to build holistic knowledge and implementation systems to understand, predict, and manipulate the interaction of genes and gene networks. This should lead to the efficient improvement of a wide range of important agronomic traits that will be introduced into commercial cultivars by an increasingly controlled and targeted coordination of recombination throughout the breeding system. DNA markers will therefore play a dual role through aiding genetic analysis of the underlying basis of important traits, and for assisting in the selection of promising progeny that after validation through field testing may become new cultivars in farmers' fields.

ACKNOWLEDGMENTS

Sangam Dwivedi is grateful to the GCP for financial support during the first stages of manuscript development, ICRISAT library staff for assistance in literature searches and sourcing reprints; to Naveen Puppala of New Mexico State University (NMSU), United States for his support during the later stages of manuscript development; and to Valerie Pipkin (NMSU) for help in preparing the manuscript for final submission.

REFERENCES

Ablett, G., Hill, H., and Henry, R. J. (2006). Sequence polymorphism discovery in wheat microsatellite flanking regions using pyrophosphate sequencing. *Mol. Breed.* **17**, 281–289.

Aert, R., Sági, L., and Volckaert, G. (2004). Gene content and density in banana (*Musa acuminata*) as revealed by genomic sequencing of BAC clones. *Theor. Appl. Genet.* **109**, 129–139.

Ahmad, M. (2000). Molecular marker-assisted selection of HMW glutenin alleles related to wheat bread quality by PCR-generated DNA markers. *Theor. Appl. Genet.* **101**, 892–896.

Ahmadi, N., Albar, L., Pressoir, G., Pinel, A., Fargette, D., and Ghesquiere, A. (2001). Genetic basis and mapping of the resistance to rice yellow mottle virus. III. Analysis of QTL efficiency in introgressed progenies confirmed the hypothesis of complementary epistasis between two resistance QTL. *Theor. Appl. Genet.* **103**, 1084–1092.

Allison, D. B. (1997). Transmission-disequilibrium tests for quantitative traits. *Am. J. Hum. Genet.* **60**, 676–690.

Alpert, K. B., and Tanksley, S. D. (1996). High-resolution mapping and isolation of a yeast artificial chromosome contig containing *fw2.2*: A major fruit weight quantitative trait locus in tomato. *Proc. Natl. Acad. Sci. USA* **93**, 15503–15507.

Andaya, V. C., and Tai, T. H. (2006). Fine mapping of the *qCTS12* locus, a major QTL for seedling cold tolerance in rice. *Theor. Appl. Genet.* **113**, 467–475.

Andersen, J. R., and Lübberstedt, T. (2003). Functional markers in plants. *Trends Plant Sci.* **8**, 554–560.

Anderson, J. A., Stack, R. W., Liu, S., Waldron, B. L., Fjeld, A. D., Coyne, C., Moreno-Sevilla, B., Fetch, J. M., Song, Q. J., Cregan, P. B., and Frohberg, R. C. (2001). DNA markers for fusarium head blight resistance QTL in two wheat populations. *Theor. Appl. Genet.* **102**, 1164–1168.

Anderson, J. V., Delseny, M., Fregene, M. A., Jorge, V., Mba, C., Lopez, C., Restrepo, S., Soto, M., Piegu, B., Verdier, V., Cooke, R., Tohme, J., *et al.* (2004). An EST resource for cassava and other species of Euphorbiaceae. *Plant Mol. Biol.* **56**, 527–539.

Andrews, D. J., and Anand Kumar, K. (1996). Use of West African pearl millet landrace *Iniadi* in cultivar development. *Plant Genet. Resour. Newsl.* **105**, 15–22.

Arru, L., Faccini, N., Govoni, C., Cattivelli, L., Pecchioni, N., Delogu, G., Stanca, A. M., and Vale, G. (2003). The PCR-based marker MWG2018 linked to the *RDG2A* leaf stripe resistance gene is a useful tool for assessing barley resistance in breeding programs. *Crop Sci.* **43**, 1036–1042.

Arumuganathan, K., and Earle, E. D. (1991). Nuclear DNA content of some important plant species. *Plant Mol. Biol. Rep.* **9**, 208–219.

Arzani, A., Peng, J. H., and Lapitan, N. L. V. (2004). DNA and morphological markers for a Russian wheat aphid resistance gene. *Euphytica* **139**, 167–172.

Ashikari, M., and Matsuoka, M. (2002). Application of rice genomics to plant biology and breeding. *Bot. Bull. Acad. Sin.* **43**, 1–11.

Ashikari, M., Sakakibara, H., Lin, S., Yamamoto, T., Takashi, T., Nishimura, A., Angeles, R. E., Qian, Q., Kitano, H., and Matsuoka, M. (2005). Cytokinin oxidase regulates rice grain production. *Science* **309**, 741–745.

Ashkenazi, V., Chani, E., Lavi, U., Levy, D., Hillel, J., and Veilleux, R. E. (2001). Development of microsatellite markers in potato and their use in phylogenetic and fingerprinting analyses. *Genome* **44**, 50–62.

Asíns, M. J. (2002). Present and future of quantitative trait locus analysis in plant breeding. *Plant Breed.* **121**, 281–291.

Auger, D. L., Gray, A. D., Ream, T. S., Kato, A., Coe, E. H., Jr., and Birchler, J. A. (2005). Nonadditive gene expression in diploid and triploid hybrids of maize. *Genetics* **169**, 389–397.

Avila, C. M., Nadal, S., Moreno, M. T., and Torres, A. M. (2006). Development of a simple PCR-based marker for the determination of growth habits in *Vicia faba* L. using a candidate gene approach. *Mol. Breed.* **17**, 185–190.

Ayoub, M., Armstrong, E., Bridger, G., Fortin, M. G., and Mather, D. E. (2003). Marker-based selection in barley for a QTL region affecting α-Amylase activity of malt. *Crop Sci.* **43**, 556–561.

Azanza, F., Barzur, A., and Juvik, J. A. (1996a). Variation in sweet corn kernel characteristics associated with stand establishment and eating quality. *Euphytica* **87**, 7–18.

Azanza, F., Tadmor, Y., Klein, B. P., Rocheford, T. R., and Juvik, J. A. (1996b). Quantitative trait loci influencing chemical and sensory characteristics of eating quality in sweet corn. *Genome* **39**, 40–50.

Babu, E. R., Mani, V. P., and Gupta, H. S. (2004). Combining high quality protein and hard endosperm traits through phenotypic and marker assisted selection.Fourth International Crop Science Congress, Australia(http://www.regional.org.au/au/cs/2004/symposia/5/2/1321_babu.htm).

Bafna, V., Gusfield, D., Lancia, G., and Yooseph, S. (2003). Haplotyping as perfect phylogeny: A direct approach. *J. Comput. Biol.* **10**, 323–340.

Bao, J. B., Lee, S., Chen, C., Zhang, X.-Q., Zhang, Y., Liu, S.-Q., Clark, T., Wang, J., Cao, M.-L., Yang, H.-M., Wang, S. M., and Yu, J. (2005). Serial analysis of gene expression study of a hybrid rice strain (*LYP9*) and its parental cultivars. *Plant Physiol.* **138**, 1216–1231.

Barone, A. (2004). Molecular marker-assisted selection for potato breeding. *Am. J. Potato Res.* **81**, 111–117.

Barone, A., Ritter, E., Schachtschabel, U., Debener, T., Salamini, F., and Gebhardt, C. (1990). Localization by restriction fragment length polymorphism mapping in potato of a major dominant gene conferring resistance to the potato cyst nematode *Globodera rostochiensis*. *Mol. Gen. Genet.* **224**, 177–182.

Barrett, J. C., Fry, B., Maller, J., and Daly, M. J. (2005). Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265.

Barriere, Y., and Argillier, O. (1993). Brown-midrib genes of maize: A review. *Agronomie* **13**, 865–876.

Barrière, Y., Guillet, C., Goffner, D., and Pichon, M. (2003). Genetic variation and breeding strategies for improved cell digestibility in annual forage crops: A review. *Anim. Res.* **52**, 193–228.

Batley, J., Barker, G., O'Sullivan, H., Edwards, K. J., and Edwards, D. (2003). Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data. *Plant Physiol.* **132**, 84–91.

Bearzoti, E., and Vencovsky, R. (1998). Methodology estimation of the proportion of genetic variance explained by molecular markers. *Genet. Mol. Biol.* **21**, 557–566.

Bearzoti, E., and Vencovsky, R. (2002). Simulation of marker-assisted recurrent selection in autogamous species. *Crop Breed. Appl. Biotechnol.* **2**, 1–10.

Beaver, J. S., Miklas, P. N., Kelly, J. D., Steadman, J. R., and Rosas, J. C. (1998). Registration of PR9357- 107 small red dry bean germplasm resistant to BCMV, BCMNV, and rust. *Crop Sci.* **38**, 1408–1409.

Beavis, W. D. (1998). QTL analysis: Power, precision and accuracy. *In* "Molecular Dissection of Complex Traits" (A. H. Paterson, Ed.), pp. 145–162. CRC Press, Boca Raton, FL.

Bedell, J. A., Budiman, M. A., Nunberg, A., Citek, R. W., Robbins, D., Jones, J., Flick, E., Rohlfing, T., Fries, J., Bradford, K., McMenamy, J., Smith, M., *et al.* (2005). Sorghum genome sequencing by methylation filtration. *PLoS Biol.* **3**, 0103–0115.

Benjamini, Y., and Yekutieli, D. (2005). Quantitative trait loci analysis using the false discovery rate. *Genetics* **171**, 783–790.

Bennetzen, J. L. (1998). The evolution of grass genome organization and function. *Symp. Soc. Exp. Biol.* **51**, 123–126.

Bennetzen, J. L., Schrick, K., Springer, P. S., Brown, W. E., and SanMiguel, P (1994). Active maize genes are unmodified and flanked by diverse classes of modified, highly repetitive DNA. *Genome* **37**, 565–576.

Bernacchi, D., Beck-Bunn, T., Emmatty, D., Eshed, Y., Inai, S., Lopez, J., Petiard, V., Sayama, H., Uhlig, J., Zamir, D., and Tanksley, S. D. (1998a). Advanced backcross QTL

analysis of tomato: II. Evaluation of near-isogenic lines carrying single-donor introgressions for desirable wild QTL- alleles derived from *Lycopersicon hirsutum* and *L. pimpinellifolium*. *Theor. Appl. Genet.* **97**, 170–180.

Bernacchi, D., Beck-Bunn, T., Eshed, Y., Lopez, J., Petiard, V., Uhlig, J., Zamir, D., and Tanksley, S. D. (1998b). Advanced backcross QTL analysis in tomato. I. Identification of QTLs for traits of agronomic importance from *Lycopersicon hirsutum*. *Theor. Appl. Genet.* **97**, 381–397.

Bernardo, R. (1994). Prediction of maize single-cross performance using RFLPs and information from related hybrids. *Crop Sci.* **34**, 20–25.

Bernardo, R. (1996). Best linear unbiased prediction of maize single-cross performance. *Crop Sci.* **36**, 50–56.

Bernardo, R. (2002). "Breeding for Quantitative Traits in Plants." Stemma Press, Woodbury, Minnesota.

Bernardo, R. (2004). What proportion of declared QTL in plants are false? *Theor. Appl. Genet.* **109**, 419–424.

Bhattramakki, D., Dong, J., Chhabra, A. K., and Hart, G. E. (2000). An integrated SSR and RFLP linkage map of *Sorghum bicolor* (L.) Moench. *Genome* **43**, 988–1002.

Bhattramakki, D., Dolan, M., Hanafey, M., Wineland, R., Vaske, D., Register, J. C., III, Tingey, S. V., and Rafalski, A. (2002). Insertion-delition polymorphisms in 3′ regions of maize genes occur frequently and can be used as highly informative genetic markers. *Plant. Mol. Biol.* **48**, 539–547.

Bidinger, F. R., Serraj, R., Rizvi, S. M. H., Howarth, C., Yadav, R. S., and Hash, C. T. (2005). Field evaluation of drought tolerance QTL effects on phenotype and adaptation in pearl millet (*Pennisetum glaucum* (L.) R. Br.) top cross hybrids. *Field Crop Res.* **94**, 14–32.

Bink, M., and Meuwissen, T. H. E. (2004). Fine mapping of quantitative trait loci using linkage disequilibrium in inbred plant populations. *Euphytica* **137**, 95–99.

Birchler, J. A., Auger, D. L., and Riddle, N. C. (2003). In search of the molecular basis of heterosis. *Plant Cell* **15**, 2236–2239.

Blaszczyk, L., Chelkowski, J., Korzun, V., Kraic, J., Ordon, F., Ovesna, J., Purnhauser, L., Tar, M., and Vida, G. (2004). Verification of STS markers for leaf rust resistance genes by seven European laboratories. *Cell. Mol. Biol. Lett.* **9**, 805–817.

Bohn, M., Groh, S., Khairallah, M. M., Hoisington, D. A., Utz, H. F., and Melchinger, A. E. (2001). Re-evaluation of the prospects of marker-assisted selection for improving insect resistance against *Diatraea* spp. In tropical maize by cross validation and independent validation. *Theor. Appl. Genet.* **103**, 1059–1067.

Bonnett, D. G., Rebetzke, G. J., and Spielmeyer, W. (2005). Strategies for efficient implementation of molecular markers in wheat breeding. *Mol. Breed.* **15**, 75–85.

Boominathan, P., Shukla, R., Kumar, A., Manna, D., Negi, D., Praveen, K. V., and Chattopadhyay, D. (2004). Long term transcript accumulation during the development of dehydration adaptation in *Cicer arietinum*. *Plant Physiol.* **135**, 1608–1620.

Boreck, I. B., and Suarez, B. K. (2001). Linkage and association: Basic concepts. *Adv. Genet.* **42**, 45–66.

Borevitz, J. O., Maloof, J. N., Lutes, J., Dabi, T., Redfern, J. L., Trainer, G. T., Werner, J. D., Asami, T., Berry, C. C., Weigel, D., and Chory, J. (2002). Quantitative trait loci controlling light and hormone response in two accessions of *Arabidopsis thaliana*. *Genetics* **160**, 683–696.

Borrell, A. K., Hammer, G. L., and van Oosterom, E. (2001). Staygreen: A consequence of the balance between supply and demand for nitrogen during grain filling? *Ann. Appl. Biol.* **138**, 91–95.

Bost, B., de Vienne, D., Hospital, F., Moreau, L., and Dillmann, C. (2001). Genetic and nongenetic basis for the L-shaped distribution of quantitative trait loci effects. *Genetics* **157**, 1773–1787.

Bouchez, A., Hospital, F., Causse, M., Gallais, A., and Charcosset, A. (2002). Marker-assisted introgression of favorable alleles at quantitative trait loci between maize elite lines. *Genetics* **162**, 1945–1959.

Bowers, J. E., Abbey, C., Anderson, S., Chang, C., Draye, X., Hoppe, A. H., Jessup, R., Lemke, C., Lennington, J., Li, Z., Lin, Y.-R., Liu, S.-C., *et al.* (2003). A high-density genetic recombination map of sequence-tagged site for *Sorghum*, as a framework for comparative structural and evolutionary genomics of tropical grains and grasses. *Genetics* **165**, 367–386.

Bowman, J. G. P., Blake, T. K., Surber, L. M. M., Habernicht, D. K., and Bockelman, H. (2001). Feed-quality variation in the barley core collection of the USDA National small grain collection. *Theor. Appl. Genet.* **41**, 863–870.

Bradbury, L. M. T., Henry, R. J., Jin, Q., Reinke, R. F., and Waters, D. L. E. (2005). A perfect marker for fragrance genotyping in rice. *Mol. Breed.* **16**, 279–283.

Bräutigam, M., Lindlöf, A., Zakhrabekova, S., Gharti-Chhetri, G., Olsson, B., and Olsson, O. (2005). Generation and analysis of 9,792 EST sequences from cold acclimated oat, *Avena sativa*. *BMC Plant Biol.* **5**, 18. doi:10.1186/1471-2229-5-18.

Brem, R. B., Yvert, G., Clinton, R., and Kruglyak, L. (2002). Genetic dissection of transcriptional regulation in budding yeast. *Science* **296**, 752–755.

Breseghello, F., and Sorrells, M. (2006a). Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics* **172**, 1165–1177.

Breseghello, F., and Sorrells, M. (2006b). Association analysis as a strategy for improvement of quantitative traits in plants. *Crop Sci.* **46**, 1323–1330.

Bretting, P. K., and Widrlechner, M. P. (1995). Genetic markers and plant genetic resource management. *Plant Breed. Rev.* **13**, 11–86.

Brick, M. A., Byrne, P. F., Schwartz, H. F., Ogg, J. B., Otto, K., Fall, A. L., and Gilbert, J. (2006). Reaction to three races of *Fusarium* wilt in the *Phaseolus vulgaris* core collection. *Crop Sci.* **46**, 1245–1252.

Briney, A., Wilson, R., Potter, R. H., Barclay, I., Crosbie, G., Appels, R., and Jones, M. G. K. (1998). A PCR-based marker for selection of starch and potential noodle quality in wheat. *Mol. Breed.* **4**, 427–433.

Broughton, W. J., Hernández, G., Blair, M., Beebe, S., Gepts, P., and Vanderleyden, J. (2003). Beans (*Phaseolus* spp.)—model food legumes. *Plant Soil* **252**, 55–128.

Brown, A. H. D. (1989). Core collections: A practical approach to genetic resources management. *Genome* **31**, 818–824.

Brown, S. M., Hopkins, M. S., Mitchell, S. E., Senior, M. L., Wang, T. Y., Duncan, R. R., Gonzalez-Candelas, F., and Kresovich, S. (1996). Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [*Sorghum bicolor* (L.) Moench]. *Theor. Appl. Genet.* **93**, 190–198.

Budak, H., Pedraza, F., Cregan, P. B., Baenziger, P. S., and Dweikat, I. (2003). Development and utilization of SSRs to estimate the degree of genetic relationships in a collection of pearlmillet germplasm. *Crop Sci.* **43**, 2284–2290.

Buhariwalla, H. K., Jarret, R. L., Jayashree, B., Crouch, J. H., and Ortiz, R. (2005a). Isolation and characterization of microsatellite markers from *Musa balbisiana*. *Mol. Ecol. Notes* **5**, 327–330.

Buhariwalla, H. K., Jayashree, B., Eeshwar, K., and Crouch, J. H. (2005b). Development of ESTs from chickpea roots and their use in diversity analysis of the *Cicer* genus. *BMC Plant Biol.* **5**, 16. DOI:10.1186/1471-2229-5-16.

Burow, M. D., Simpson, E. C., Starr, J. L., and Paterson, A. H. (2001). Transmission genetics of chromatin from a synthetic amphidiploid to cultivated peanut (*Arachis hypogaea* L.): Broadening the gene pool of a monophyletic polyploidy species. *Genetics* **159**, 823–837.

Burr, B., Evola, S. V., Burr, F. A., and Backemann, J. S. (1983). The application of restriction fragment length polymorphism to plant breeding. *In* "Genetic Engineering" (J. K. Setlow and A. Hollaender, Eds.), Vol. 5, pp. 45–49. Plenum Press, New York.

Burr, B., Burr, F. A., Thompson, K. H., Alberston, M. C., and Stubber, C. W. (1988). Gene mapping with recombinant inbreds in maize. *Genetics* **118**, 519–526.

Bustamam, M., Tabien, R. E., Suwarmo, A., Abalos, M. C., Kadir, T. S., Ona, I., Bernardo, M., VeraCruz, C. M., and Leung, H. (2002). Asian rice biotechnology network: Improving popular cultivars through marker-assisted backcrossing by the NARES. Abstract of the Int. Rice Congress, 16–22 September 2002, Beijing China, (www.irri.org/irc2002/index.htm).

Buteler, M. I., Jarret, R. L., and LaBonte, D. R. (1999). Sequence Characterization of microsatellites in diploid and polyploidy *Ipomoea*. *Theor. Appl. Genet.* **99**, 123–132.

Cagampang, G. B., Perez, C. M., and Juliano, B. O. (1973). A gel consistency test for eating quality of rice. *J. Sci. Food Agric.* **24**, 1589–1594.

Cahill, D. J., and Schmidt, D. H. (2004). Use of marker-assisted selection in a product development breeding program. "Fourth International Crop Science Congress, Brisban, Australia, 26 September–1 October 2004" (T. Fischer, Ed.). (www.cropscience.org.au).

Cakir, M., Gupta, S., Platz, G. J., Ablett, G. A., Loughman, R., Emebiri, L. C., Poulsen, D., Li, C. D., Lance, R. C. M., Galwey, N. W., Jones, M. G. K., and Appels, R. (2003). Mapping and validation of the genes for resistance to *Pyrenophora teres* f. *teres* in barley (*Hordium vulgare* L.). *Aust. J. Agric. Res.* **54**, 1369–1377.

Campos, H., Cooper, A., Habben, J. E., Edmeades, G. O., and Schussler, J. R. (2004). Improving drought tolerance in maize: A view from industry. *Field Crop Res.* **90**, 19–34.

Canci, P. C., Nduulu, L. M., Muehlbauer, G. J., Dill-Macky, R., Rasmussoon, D. C., and Smith, K. P. (2004). Validation of quantitative trait loci for fusarium head blight and kernel discoloration in barley. *Mol. Breed.* **14**, 91–204.

Cao, L.-Y., Zhuang, J.-Y., Zhan, X.-D., Zheng, K.-L., and Cheng, S.-H. (2003). Hybrid rice resistant to bacterial blight developed by marker assisted selection. *Chinese J. Rice Sci.* **17**, 184–186.

Castro, A. J., Chen, X., Hayes, P. M., and Johnston, M. (2003a). Pyramiding quantitative trait locus (QTL) alleles determining resistance to barley stripe rust: Effects on resistance at seedling stage. *Crop Sci.* **43**, 651–659.

Castro, A. J., Chen, X., Corey, A., Filichkina, T., Hayes, P. M., Mundt, C., Richardson, K., Sandoval-Islas, S., and Vivar, H. (2003b). Pyramiding and validation of quantitative trait locus (QTL) alleles determining resistance to barley stripe rust: Effects on adult plant resistance. *Crop Sci.* **43**, 2234–2239.

Causse, M., Santoni, S., Damerval, C., Maurice, A., Charcosset, A., Deatrick, J., and de Vienne, D. (1996). A composite map of expressed sequences in maize. *Genome* **39**, 418–432.

Causse, M. A., Fulton, T. M., Cho, Y. G., Ahn, S. N., Chunwongse, J., Wu, K., Xiao, J., Yu, Z., Ronald, P. C., Harrington, S. E., Second, G., McCouch, S. R., *et al.* (1994). Saturated molecular map of the rice genome based on an interspecific backcross population. *Genetics* **138**, 1251–1274.

Chaitieng, B., Kaga, A., Tomooka, N., Isemura, T., Kuroda, Y., and Vaughan, D. A. (2006). Development of a black gram [*Vigna mungo* (L.) Hepper] linkage map and its comparison with an azuki bean [*Vigna angularis* (Willd.) Ohwi and Ohashi] linkage map. *Theor. Appl. Genet.* **113**, 1261–1269.

Chandler, P. M., Marrion-Poll, A., Ellis, M., and Gubler, F. (2002). Mutants at the *Slender1* locus of barley cv Himalaya. Molecular and physical characterization. *Plant Physiol.* **129**, 181–190.

Chao, S., Baysdorfer, C., Heredia-Diaz, O., Musket, T., Xu, G., and Coe, E. H., Jr. (1994). RFLP mapping of partially sequenced leaf cDNA clones in maize. *Theor. Appl. Genet.* **88**, 717–721.

Chapman, S. C., Hammer, G. L., Podlich, D. W., and Cooper, M. (2002). Linking biophysical and genetic models to integrate physiology, molecular biology and plant breeding. *In* "Quantitative Genetics, Genomics and Plant Breeding" (M. S. Kong, Ed.), pp. 167–187. CABI, Wallingford, UK.

Chapman, S. C., Cooper, M., Podlich, D. W., and Hammer, G. L. (2003). Evaluating plant breeding strategies by simulating gene action and dryland environment effects. *Agron. J.* **95**, 99–113.

Chavarriaga-Aguirre, P., Maya, M. M., Bonierbale, M. W., Kresovich, S., Fergene, M. A., Tohme, J., and Kochert, G. (1998). Microsatellites in cassava (*Manihot esculenta* Crantz): Discovery, inheritance and variability. *Theor. Appl. Genet.* **97**, 493–501.

Chavarriaga-Aguirre, P., Maya, M. M., Tohme, J., Duque, M. C., Carlos, I., Bonierbale, M. W., Kresovich, S., and Kochert, G. (1999). Using microsatellites, isozymes and AFLPs to evaluate genetic diversity and redundancy in the cassava core collection and to assess the usefulness of DNA-based markers to maintain germplasm collections. *Mol. Breed.* **5**, 263–273.

Chen, D.-H., dela Vina, M., Inukai, T., Mackill, D. J., Ronald, P. C., and Nelson, R. J. (1999). Molecular mapping of the blast resistance gene, *Pi44*(*t*), in a line derived from a durably resistant rice cultivar. *Theor. Appl. Genet.* **98**, 1046–1053.

Chen, M., Presting, G., Barbazuk, W. B., Goicoechea, J. L., Blackmon, B., Fang, G., Kim, H., Frisch, D., Yu, Y., Higingbottom, S., Phimphilai, J., Phimphilai, D., *et al.* (2002). An integrated physical and genetic map of the rice genome. *Plant Cell* **14**, 537–545.

Chen, S., Lin, X. H., Xu, C. G., and Zhang, A. (2000). Improvement of bacterial blight resistance of Minghui 63, an elite restorer line of hybrid rice by marker-assisted selection. *Crop Sci.* **40**, 239–244.

Chen, S., Xu, C. G., Lin, X. H., and Zhang, Q. (2001). Improving bacterial blight resistance of 6078, an elite restorer line of hybrid rice, by molecular marker-assisted selection. *Plant Breed.* **120**, 133–137.

Cheng, Z., Presting, G. G., Buell, C. R., Wing, R. A., and Jiang, J. (2001). High resolution pachytene chromosome mapping of bacterial artificial chromosomes anchored by genetic markers reveals the chromosome location and the distribution of genetic recombination along chromosome 10 of rice. *Genetics* **157**, 1749–1757.

Chin, E. C. L., Senior, M. L., Shu, H., and Smith, J. S. C. (1996). Maize simple repetitive DNA sequences: Abundance and allelic variation. *Genome* **39**, 866–873.

Ching, A., Caldwell, K. S., Jung, M., Dolan, M., Smith, O. S., Tingey, S., Morgante, M., and Rafalski, A. J. (2002). SNP frequency, haplotype structure and linakge disequilibrium in elite maize inbred lines. *BMC Genet.* **3**, 19–33.

Cho, R. J., Mindrinos, M., Richards, D. R., Sapolsky, R. J., Anderson, M., Drenkard, E., Dewdney, J., Reuber, T. L., Stammers, M., Federspiel, N., Theologis, A., Yang, W.-H., *et al.* (1999). Genome-wide mapping with biallelic markers in *Arabidposis thaliana*. *Nat. Genet.* **23**, 203–207.

Cho, Y. G., Eun, M. Y., McCouch, S. R., and Chae, Y. A. (1994). The semidwarf gene, *sd-1*, of rice (*Oryza sativa* L.). II. Molecular mapping and marker-assisted selection. *Theor. Appl. Genet.* **89**, 54–59.

Cho, Y. G., Ishii, T., Temnykh, S., Chen, X., Lipovich, L., McCouch, S. R., Park, W. D., Ayres, N., and Cartinhour, S. (2000). Diversity of microsatellites derived from genomic libraries and GeneBank sequences in rice. *Theor. Appl. Genet.* **100**, 713–722.

Choi, H.-K., Melissa, A. L., Doyle, J., and Cook, D. R. (2006). Development of nuclear gene-derived markers linked to legume species. *Mol. Gen. Genom.* **276**, 56–70.

Christopher, M., Cordeiro, G., Waters, D., and Henry, R. (2004). Marker assisted selection in rice improvement. RIRDC Publication No 04/011 (http://www.rirdc.gov.au/reports/ric/04-011sum.html).

Coe, E., Cone, K., McMullen, M., Chen, S.-S., Davis, G., Gardiner, J., Liscum, E., Polacco, M., Paterson, A., Sanchez-Villeda, H., Soderlund, C., and Wing, R. (2002). Access to the maize genome: An integrated physical and genetic map. *Plant Physiol.* **128**, 9–12.

Coemans, B., Matsumura, H., Terauchi, R., Remy, S., Swennen, R., and Sági, L. (2005). SuperSAGE combined with PCR walking allows global gene expression profiling of banana (*Musa acuminata*), a non-model organism. *Theor. Appl. Genet.* **111**, 1118–1126.

Collard, B. C. Y., and Mackill, D. J. (2007). Marker-assisted selection: An approach for precision plant breeding in the 21st century. *Phil. Trans. R. Soc. B. Rev.* (in press).

Collins, H. M., Panozzo, J. F., Logue, S. J., Jefferies, S. P., and Barr, A. R. (2003). Mapping and validation of chromosome regions associated with high malt extract in barley (*Hordeum vulgare* L.). *Aust. J. Agric. Res.* **54**, 1223–1240.

Colton, L. M., Groza, H. I., Wielgus, S. M., and Jiang, J. (2006). Marker-assisted selection for the broad- spectrum potato late blight resistance conferred by gene *RB* derived from a wild potato species. *Crop Sci.* **46**, 589–594.

Conaway-Bormans, C. A., Marchetti, M. A., Johnson, C. W., McClung, A. M., and Park, W. D. (2003). Molecular markers linked to the rice blast resistance gene *Pi-z* in rice for use in marker-assisted selection. *Theor. Appl. Genet.* **107**, 1014–1020.

Concibido, V. C., Denny, R. L., Lange, D. A., Orf, J. H., and Young, N. D. (1996). RFLP mapping and molecular marker-assisted selection of soybean cyst nematode resistance in PI 209332. *Crop Sci.* **36**, 1643–1650.

Concibido, V. C., La Vallee, B., Mclaird, P., Pineda, N., Meyer, J., Hummel, L., Yang, J., Wu, K., and Delannay, X. (2003). Introgression of a quantitative trait locus for yield from *Glycine soja* into commercial soybean cultivars. *Theor. Appl. Genet.* **106**, 575–582.

Concibido, V. C., Diers, B. W., and Arelli, P. R. (2004). A decade of QTL mapping for cyst nematode resistance in soybean. *Crop Sci.* **44**, 1121–1131.

Cone, K. C., McMullen, M. D., Bi, I. V., Davis, G. L., Yim, Y.-S., Gardiner, J. M., Polacco, M. L., Sanchez-Villeda, H., Fang, Z., Schroeder, S. G., Havermann, S. A., Bowers, J. E., *et al.* (2002). Genetic, physical, and informatics resources for maize. On the road to integrated map. *Plant Physiol.* **130**, 1598–1605.

Cooper, M., Chapman, S. C., Podlich, D. W., and Hammer, G. L. (2002). The GP problem: Quantifying gene-to-phenotype relationships. *In Silico Biol.* **2**, 151–164.

Cooper, M., Smith, O. S., Graham, G., Arthur, L., Feng, L., and Podlich, D. W. (2004). Genomics, Genetics, and Plant Breedings: A private sector perspective. *Crop Sci.* **44**, 1907–1913.

Costa, J. M., Corey, A., Hayes, P. M., Jobet, C., Kleinhofs, A., Kopisch-Obusch, A., Kramer, S. F., Kudrna, D., Li, M., Riera-Lizarazu, O., Sato, K., Szucs, P., *et al.* (2001). Molecular mapping of the Oregon Wolfe barleys: A phenotypically polymorphic double-haploid population. *Theor. Appl. Genet.* **103**, 415–424.

Courtois, B., Chaitep, W., Moolsri, S., Sinha, P. K., Trebuil, G., and Yadav, R. (1996). Drought resistance and germplasm improvement. Ongoing research in the Upland rice research consortium. *In* "Upland Rice Research in Partnership IRRI" (C. Piggin, B. Curtois, and V. Schmit, Eds.), pp. 154–175. IRRI Discussion Paper Series 16, Los Baños, the Philippines.

Coventry, S. J., Collins, H. M., Barr, A. R., Jefferues, S. P., Chalmers, K. J., Logue, S. J., and Langridge, P. (2003). Use of putative QTLs and structural genes in marker-assisted selection for diastatic power in malting barley (*Hordeum vulgare* L.). *Aust. J. Agric. Res.* **54**, 1241–1250.

Cregan, P. B., Mudge, J., Fickus, E. W., Danesh, D., Denny, R., and Young, N. D. (1999). Two simple sequence repeat markers to select for soybean cyst nematode resistance conditioned by the *rhg1* locus. *Theor. Appl. Genet.* **99**, 811–818.

Crosbie, T. M., Eathington, S. R., Johnson, G. R., Edwards, M., Reiter, R., Stark, S., Mohanty, R. G., Oyervides, M., Buehler, R. E., Walker, A. K., Dobert, R., Delannay, X., *et al.* (2006). Plant breeding: Past, present, and future. *In* "Plant Breeding: The Arnel R. Hallauer International Symposium" (K. R. Lamkey and M. Lee, Eds.). Blackwell Publishing, Ames, Iowa.

Crossa, J., Burgueño, J., Cornelius, P. L., McLaren, G., Trethowan, R., and Krischnamachari, A. (2006). Modeling genotype x environment interaction using additive genetic covariance of relatives for predicting breeding values of wheat genotypes. *Crop Sci.* **46**, 1722–1733.

Crouch, H. K., Crouch, J. H., Jarret, R. L., Cregan, P. B., and Ortiz, R. (1998a). Segregation at microsatellite loci in haploid and diploid gametes of *Musa*. *Crop Sci.* **98**, 211–217.

Crouch, J. H., Vuylsteke, D., and Ortiz, R. (1998b). Perspectives on the application of biotechnology to assist the genetic enhancement of plantain and banana (*Musa* spp.). *Electronic J. Biotechnol.* **1**, 1–12(http://www.ejb.org/content/vol1/issue1/full/2/).

Darvasi, A., and Soller, M. (1995). Advanced intercross lines, an experimental population for fine genetic mapping. *Genetics* **141**, 1199–1207.

Datta, K., Baisakh, N., Thet, K. M., Tu, J., and Datta, S. K. (2002). Pyramiding transgenes for multiple resistance in rice against bacterial blight, yellow stem borer, and sheath blight, *Theor. Appl. Genet.* **106**, 1–8.

Davis, G. L., McMullen, M. D., Baysdorfer, C., Musket, T., Grant, D., Staebell, M., Xu, G., Polacco, M., Koster, L., Melia-Hancock, S., Houchins, K., Chao, S., *et al.* (1999). A maize map standard with sequenced core markers, grass genome reference points and 932 expressed sequence tagged sites (ESTs) in a 1736-locus map. *Genetics* **152**, 1137–1172.

de Bustos, A., Rubio, P., Soler, C., Garcia, P., and Jouve, N. (2001). Marker-assisted selection to improve HMW-glutenins in wheat. *Euphytica* **119**, 69–73.

Dekkers, J. C. M., and Hospital, F. (2002). The use of molecular genetics in the improvement of agricultural populations. *Nat. Rev. Genet.* **3**, 22–32.

Dellaporta, S. L., Wood, J., and Hicks, J. B. (1983). A plant DNA minipreparation: Version II. *Plant Mol. Biol. Rep.* **1**, 19–21.

Delmer, D. P. (2005). Agriculture in the developing world: Connecting innovations in plant research to downstream application. *Proc. Natl. Acad. Sci. USA* **102**, 15739–15746.

Demeke, T., Laroche, A., and Gaudet, D. A. (1996). A DNA marker for the *Bt-10* common bunt resistant gene in wheat. *Genome* **39**, 51–56.

Deng, H. W., Chen, W. M., and Reskker, R. R. (2001). Population admixture: Detection by Hardy- Weinberg test and its quantitative effects on linkage disequilibrium methods for localizing genes underlying complex traits. *Genetics* **157**, 885–897.

Deu, M., and Glaszmann, J. C. (2004). Linkagedisequilibrium in sorghum. Plant and Animal Genomes XII Conference, PAG-XII: (W10).

Deu, M., Rattunde, F., and Chantereau, J. (2006). A global view of genetic diversity in cultivated sorghums using a core collection. *Genome* **49**, 168–180.

de Vicente, M. C., and Tanksley, S. D. (1993). QTL analysis of transgressive segregation in an interspecific tomato cross. *Genetics* **134**, 585–596.

Doebley, J., and Stec, A. (1993). Inheritance of the morphological differences between maize and teosinte: Comparison of results for two F_2 population. *Genetics* **134**, 559–570.

Doebley, J., Stec, A., and Gustus, C. (1995). *Teosinte branched 1* and the origin of maize: Evidence for epistasis and the evolution of dominance. *Genetics* **141**, 333–346.

Doebley, J., Stec, A., and Hubbard, L. (1997). The evolution of apical dominance in maize. *Nature* **386**, 485–488.

Doi, K., Izawa, T., Fuse, T., Yamanouchi, U., Kubo, T., Shimatani, Z., Yano, M., and Yoshimura, A. (2004). *Ehd1*, a B-type response regulator in rice, confers short-day promotion of flowering and controls *FT-like* gene expression independently of *Hd1*. *Genes Dev.* **18**, 926–936.

Dong, Y. S., Cao, Y. S., Zhang, X. Y., Liu, S. C., Wang, L. F., You, G. X., Pang, B. S., Li, L. H., and Jia, J. Z. (2003). Establishment of candidate core collections in Chinese common wheat germplasm. *J. Plant Genet. Resour.* **4**, 1–8.

Dorweiler, J., Stec, A., Kermicle, J., and Doebley, J. (1993). *Teosinte glume architecture 1*: A genetic locus controlling a key step in maize evolution. *Science* **262**, 233–235.

Douglass, S. K., Juvik, J. A., and Spilltstoesser, W. E. (1993). Sweet corn seedling emergence and variation in kernel carbohydrate reserves. *Seed Sci. Technol.* **21**, 433–445.

D'Ovidio, R. (1993). Single-seed PCR of LMW glutenin genes to distinguish between durum wheat cultivars with good and poor technological properties. *Plant Mol. Biol.* **22**, 1173–1176.

Draye, X., Lin, Y.-R., Qian, X.-Y., Bowers, J. E., Burow, G. B., Morell, P. L., Peterson, D. G., Presting, G. G., Ren, S.-X., Wing, R. A., and Paterson, A. H. (2001). Toward integration of comparative genetic, physical, diversity, and cytomolecular maps for grasses and grains, using the sorghum genome as a foundation. *Plant Physiol.* **125**, 1325–1341.

Dreher, K., Morris, M., Khairallah, M., Ribaut, J. M., Pandey, S., and Srinivasan, G. (2002). Is marker-assisted selection cost-effective compared with conventional plant breeding methods? The case of quality protein Maize. *In* "Economic and Social Issues in Agricultural Biotechnology" (R. E. Evenson, V. Santaniello, and D. Zilberman, Eds.), pp. 203–236. CABI, Wellingford, New York, USA.

Dreher, K., Khairallah, M., Ribaut, J. M., and Morris, M. (2003). Money matters (I): Costs of field and laboratory procedures associated with conventional and marker-assisted maize breeding at CIMMYT. *Mol. Breed.* **11**, 221–234.

Dubcovsky, J., Lijavetzky, D., Appendino, L., and Tranquilli, G. (1998). Comparative RFLP mapping of *Triticum monococcum* genes controlling vernalization requirement. *Theor. Appl. Genet.* **97**, 968–974.

Durrett, R. T., Chen, K. Y., and Tanksley, S. D. (2002). A simple formula useful for positional cloning. *Genetics* **160**, 353–355.

Duvick, D. N. (1999). Heterosis: Feeding the people and protecting natural resources. *In* "Genetics and Exploitation of Heterosis in Crops" (J. G. Coors, and S. Pandey, Eds.), pp. 19–29. American Society of Agronomy, Madison, Wisconsin, USA.

Dvorak, J., Luo, M. C., and Yang, Z. L. (1998). Restriction fragment length polymorphism and divergence in the genomic regions of high and low recombination in self-fertilizing and cross-fertilizing *Aegilops* species. *Genetics* **148**, 423–434.

Dweikat, I., Ohm, H., Mackenzie, S., Paterson, F., Cambron, S., and Ratcliffe, R. (1994). Association of DNA marker with Hessian fly resistance gene *H9* in wheat. *Theor. Appl. Genet.* **89**, 964–968.

Dwivedi, S. L., Crouch, J. H., Nigam, S. N., Ferguson, M. E., and Paterson, A. H. (2003). Molecular breeding of groundnut for enhanced productivity and food security in the semi-arid tropics: Opportunities and challenges. *Adv. Agron.* **80**, 154–221.

Dwivedi, S. L., Blair, M., Upadhyaya, H. D., Serraj, R., Balaji, J., Buhariwalla, H. K., Ortiz, R., and Crouch, J. H. (2005). Using genomics to exploit grain legume biodiversity in crop improvement. *Plant Breed. Rev.* **26**, 176–357.

Dwivedi, S. L., Bertioli, D. J., Crouch, J. H., Valls, J. F., Upadhyaya, H. D., Fávero, A., Moretzsohn, M., and Paterson, A. H. (2006). Peanut. *In* "Genome Mapping and Molecular Breeding in Plants, Volume 2: Oiseeds" (C. Kole, Ed.). Springler-Verlag, Berlin Heidelberg, Germany.

Dwivedi, S. L., Stalker, H. T., Blair, M. W., Bertioli, D., Upadhyaya, H. D., Nielen, S., and Ortiz, R. (2007). Enhancing crop gene pools with beneficial traits using wild relatives. *Plant Breed. Rev.* **30** (in press).

Eagles, H. A., Bariana, H. S., Ogbonnaya, F. C., Rebetzke, G. J., Hollamby, G. J., Henry, R. J., Henschke, P. H., and Carter, M. (2001). Implementation of markers in Australian wheat breeding. *Aust. J. Agric. Res.* **52**, 1349–1356.

Eathington, S. R. (2005). Practical applications of molecular technology in the development of commercial maize hybrids. *In* "Proceedings of the 60th Annual Corn and Sorghum Seed Research Conferences." American Seed Trade Association, Washington DC.

Ellis, M. H., Spielmeyer, W., Gale, K. R., Rebetzke, G. J., and Richards, R. A. (2002). 'Perfect' markers for the *Rht-B1b* and *Rht-D1b* dwarfing genes in wheat. *Theor. Appl. Genet.* **105**, 1038–1042.

Eshed, Y., and Zamir, D. (1994). A genomic library of *Lycopersicon pennellii* in *L. esculentum*: A tool for fine mapping of genes. *Euphytica* **79**, 175–179.

Eshed, Y., and Zamir, D. (1995). An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. *Genetics* **141**, 1147–1162.

Eujayl, I., Sorrels, M. E., Baum, M., Wolters, P., and Powell, W. (2002). Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. *Theor. Appl. Genet.* **104**, 399–407.

Faleiro, F. G., Ragagnin, V. A., Moreira, M. A., and de Barros, E. G. (2004). Use of molecular markers to accelerate the breeding of common bean lines resistant to rust and anthracnose. *Euphytica* **138**, 213–218.

Falque, M., Decousset, L., Dervins, D., Jacob, A.-M., Joets, J., Martinant, J.-P., Raffoux, X., Ribière, N., Ridel, C., Samson, D., Charcosset, A., and Murigneux, A. (2005). Linkage mapping of 1454 new maize candidate gene loci. *Genetics* **170**, 1957–1966.

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

Fasoula, V. A., Harris, D. K., and Boerma, H. R. (2004). Validation and designation of quantitative trait loci for seed protein, seed oil, and seed weight from two soybean populations. *Crop Sci.* **44**, 1218–1225.

Fauquet, C. M., and Tohme, J. (2004). The global cassava partnership for genetic improvement. *Plant Mol. Biol.* **86**, v–x(editorial).

Fauré, S., Noyer, J. L., Horry, J. P., Bakry, F., Lanaud, C., and de León, D. G. (1993). A molecular marker-based linkage map of diploid bananas (*Musa acuminata*). *Theor. Appl. Genet.* **87**, 517–526.

Feltus, F. A., 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 subspecies *indica* and *japonica* genome alignments. *Genome Res.* **14**, 1812–1819.

Feltus, F. A., Singh, H. P., Lohithaswa, H. C., Schulze, S. R., Silva, T. D., and Paterson, A. H. (2006). A comparative genomic strategy for targeted discovery of single-nucleotide polymorphisms and conserved-noncoding sequences in orphan crops. *Plant Physiol.* **140**, 1183–1191.

Fernandes, J., Brendel, V., Gai, X., Lal, S., Chandler, V. L., Elumalai, R. P., Galbraith, D. W., Pierson, E. A., and Walbot, V. (2002). Comparison of RNA profiles based on maize expressed sequence tag frequency analysis and micro-array hybridization. *Plant Physiol.* **128**, 896–910.

Fjellstrom, R., Conaway-Bormans, C. A., McClung, A. M., Marchetti, M. A., Shank, A. R., and Park, W. D. (2004). Development of DNA markers suitable for marker-assisted selection of three *Pi* genes conferring resistance to multiple *Pyricularia grisea* pathotypes. *Crop Sci.* **44**, 1790–1798.

Fjellstrom, R., McClung, A. M., and Shank, A. R. (2006). SSR markers closely linked to the *Pi-z* locus are useful for selection of blast resistance in a broad array of rice germplasm. *Mol. Breed.* **17**, 149–157.

Flint-Garcia, S. A., Thornsberry, J. M., and Buckler, E. S., IV. (2003). Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* **54**, 357–374.

Fregene, M., Angel, F., Gomez, R., Rodriguez, F., Chavarriaga, P., Roca, W., Tohme, J., and Bonierbale, M. (1997). A molecular genetic map of cassava (*Manihot esculenta* Crantz). *Theor. Appl. Genet.* **95**, 431–441.

Fregene, M., Matsumura, H., Akano, A., Dixon, A., and Terauchi, R. (2004). Serial analysis of gene expression (SAGE) of host-plant resistance to the cassava mosaic disease (CMD). *Plant Mol. Biol.* **56**, 563–571.

Fridman, E., Pleban, T., and Zamir, D. (2000). A recombination hotspot delimits a wild species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. *Proc. Natl. Acad. Sci. USA* **97**, 4718–4723.

Fridman, E. F., Carrari, F., Liu, Y.-S., Fernie, A. R., and Zamir, D. (2004). Zooming in on a quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. *Proc. Natl. Acad. Sci. USA* **97**, 4718–4723.

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

Frisch, M., Bohn, M., and Melchinger, A. E. (1999b). Comparison of selection strategies for marker-assisted backcrossing of a gene. *Crop Sci.* **39**, 1295–1301.

Frisch, M., Bohn, M., and Melchinger, A. E. (2000). PLABSIM: Software for simulation of marker-assisted backcrossing. *J. Hered.* **91**, 86–87.

Fu, H., and Dooner, H. K. (2002). Intraspecific violation of genetic colinearity and its implications in maize. *Proc. Natl. Acad. Sci. USA* **99**, 9573–9578.

Fu, Y. B., Peterson, G. W., Williams, D., Richards, K. W., and Fetch, J. M. (2005). Patterns of AFLP variation in a core subset of cultivated hexaploid oat germplasm. *Theor. Appl. Genet.* **111**, 530–539.

Fulton, T. M., Beck-Bunn, T., Emmatty, D., Eshed, Y., Lopez, J., Petiard, V., Uhlig, J., Zamir, D., and Tanksley, S. D. (1997). QTL analysis of an advanced backcross of *Lycopersicon peruvianum* to the cultivated tomato and comparisons with QTLs found in other wild species. *Theor. Appl. Genet.* **95**, 881–894.

Gallais, A., Dillmann, C., and Hospital, F. (1997). An analytical approach of marker assisted selection with selection on markers only. *In* "Advances in Biometrical Genetics. Proceedings of the Tenth Meeting of the EUCARPIA Section Biometrics in Plant Breeding" (R. Krajewski and Z. Kaczmarek, Eds.), pp. 111–116. Institute of Plant Genetics, Polish Academy of Sciences, Poznan.

Gaut, B. S., Le Thierry I′, E. M., Peek, A. S., and Saukins, M. C. (2000). Maize as a model for the evolution of plant nuclear genomes. *Proc. Natl. Acad. Sci. USA* **97**, 7008–7015.

Gebhardt, C., Ritter, E., Barone, A., Debener, T., Walkemeier, B., Schachtschabel, U., Kaufmann, H., Thompson, R. D., Bonierbale, M. W., Ganal, M. W., Tanksley, S. D., and Salamini, F. (1991). RFLP map of potato and their alignment with the homoeologous tomato genome. *Theor. Appl. Genet.* **83**, 49–57.

Gebhardt, C., Bellin, D., Henselewski, H., Lehmann, W., Schwarzfischer, J., and Valkonen, J. P. T. (2006). Marker-assisted combination of major genes for pathogen resistance in potato. *Theor. Appl. Genet.* **112**, 1458–1464.

Gibbon, B. C., and Larkins, B. A. (2005). Molecular genetic approaches to developing quality protein maize. *Trends Genet.* **21**, 227–233.

Gilchrist, E. J., and Haughn, G. W. (2005). TILLING without a plough: A new method with applications for reverse genetics. *Curr. Opin. Plant Biol.* **8**, 1–5.

Gill, B. S., Appels, R., Botha-Oberholster, A.-M., Buell, C. R., Bennetzen, J. L., Chalhoub, B., Chumley, F., Dvořák, J., Iwanaga, M., Keller, B., Li, W., McCombie, W. R., *et al.* (2004). A workshop report on wheat genome sequencing: International genome research on wheat consortium. *Genetics* **168**, 1087–1096.

Gimelfarb, A., and Lande, R. (1994). Simulation of marker-assisted selection in hybrid populations. *Genet. Res.* **63**, 39–47.

Gimelfarb, A., and Lande, R. (1995). Marker-assisted selection and marker-QTL associations in hybrid populations. *Theor. Appl. Genet.* **91**, 522–528.

Goff, S. A., Ricke, D., Lan, T.-H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H., Hadley, D., Hutchinson, D., *et al.* (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* **296**, 92–100.

Gold, J., Harder, D., Townley-Smith, F., Aung, T., and Procunier, J. (1999). Development of a molecular marker for rust resistance gene *Sr39* and *Lr35* in wheat breeding lines. *Electron. J. Biotechnol.* **2**(1), (http://www.ejb.org/content/vol2/issue1/full/1).

Graham, G. I., Wolff, D. W., and Stuber, C. W. (1997). Characterization of a yield quantitative trait locus on chromosome five of maize by fine mapping. *Crop Sci.* **37**, 1601–1610.

Grapes, L., Dekkers, J. C. M., Rothschild, M. F., and Fernando, R. L. (2004). Comparing linkage disequilibrium-based methods for fine mapping quantitative trait loci. *Genetics* **166**, 1561–1570.

Gu, S., Pakstis, A. J., and Kidd, K. K. (2005). HAPLOT: A graphical comparison of haplotype blocks, tagSNP sets and SNP variation for multiple populations. *Bioinformatics* **21**, 3938–3939.

Gu, W., Weeden, N., Yu, J., and Wallace, D. (1995). Large-scale, cost-effective screening of PCR products in marker-assisted selection. *Theor. Appl. Genet.* **91**, 465–470.

Guillaumie, S., Linossier, L., Torney, V., Robert, N., and Ravel, S. (2004). Co-location between a gene encoding for the bZip factor SPA and an eQTL for high molecular weight glutenin subunit in wheat (*Triticum aestivum*). *Genome* **47**, 705–713.

Guillet-Claude, C., Birolleau-Touchard, C., Manicacci, D., Rogowsky, P. M., Rigau, J., Murigneux, A., Martinant, J.-P., and Barrière, Y. (2004). Nucleotide diversity of the *ZmPox3* maize peroxidase gene: Relationships between a MITE insertion in exon 2 and variation in forage maize digestibility. *BMC Genet.* **5**, 19, doi: 10.1186/1471-2156-5-19.

Gulick, P. J., Drouin, S., Yu, Z., Danyluk, J., Poisson, G., Monroy, A. F., and Sarhan, F. (2005). Transcriptome comparison of winter and spring wheat responding to low temperature. *Genome* **48**, 913–923.

Guo, M., Rupe, M. A., Zinselmeier, C., Habben, J., Bowen, B. A., and Smith, O. S. (2004). Allelic variation of gene expression in maize hybrids. *Plant Cell* **16**, 1707–1716.

Guo, S.-W. (2000). Genetic mapping of complex traits: Promises, problems, and prospects. *Theor. Popul. Biol.* **57**, 1–11.

Gupta, P. K., Rustgi, S., Sharma, S., Singh, R., Kumar, N., and Balyan, H. S. (2003). Transferable EST- SSR markers for the study of polymorphism and genetic diversity in bread wheat. *Mol. Genet. Genomics* **270**, 315–323.

Gupta, P. K., Rustgi, S., and Kulwal, P. L. (2005a). Linkage disequilibrium and association studies in higher plants: Present status and future prospects. *Plant Mol. Biol.* **57**, 461–485.

Gupta, S. K., Charpe, A., Koul, S., Prabhu, K. V., and Haq, Q. M. R. (2005b). Development and validation of molecular markers linked to an *Aegilops umbellulata*-derived leaf rust-resistance gene, *Lr9*, for marker-assisted selection in bread wheat. *Genome* **48**, 823–830.

Gur, A., and Zamir, D. (2004). Unused natural variation can lift yield barriers in plant breeding. *PLoS Biol.* **2**, 1610–1615.

Haberer, G., Young, S., Bharati, A. K., Gundlach, H., Raymond, C., Fuks, G., Butler, E., Wing, R. A., Rounsley, S., Birren, B., Nusbaum, C., Mayer, K. F. X., *et al.* (2005). Structure and architecture of the maize genome. *Plant Physiol.* **139**, 1612–1624.

Halward, T. M., Stalker, H. T., and Kochert, G. (1993). Development of an RFLP linkage map in diploid peanut species. *Theor. Appl. Genet.* **87**, 379–384.

Hämäläinen, J. H., Watanabe, K. N., Valkonen, J. P. T., Arihira, A., Plaisted, R. L., Pehu, E., Miller, L., and Slack, L. A. (1997). Mapping and marker-assisted selection for a gene for extreme resistance to potato virus Y. *Theor. Appl. Genet.* **94**, 192–197.

Hamblin, M. T., and Aquadro, C. F. (1999). DNA sequence variation and the recombinational landscape in *Drosophila pseudoobscura*: A study of the second chromosome. *Genetics* **153**, 859–869.

Hammer, G. L., Kropff, M. J., Sinclair, T. R., and Porter, J. R. (2002). Future contribution of crop modeling: From heuristics and supporting decision making to understanding genetic regulation and aiding crop improvement. *Eur. J. Agron.* **18**, 15–31.

Hammer, G. L., Chapman, S., van Oosterom, E., and Podlich, D. W. (2005). Trait physiology and crop modeling as a framework to link phenotypic complexity to underlying genetic systems. *Aust. J. Agric. Res.* **56**, 947–960.

Han, O. K., Kaga, A., Isemura, T., Wang, X. W., Tomooka, N., and Vaughan, D. A. (2005). A genetic linkage map for azuki bean [*Vigna angularis* (Willd.) Ohwi & Ohashi]. *Theor. Appl. Genet.* **111**, 1278–1287.

Harushima, Y., Yano, M., Shomura, A., Sato, M., Shimano, T., Kuboki, Y., Yamamoto, T., Lin, S. Y., Antonio, B. A., Parco, A., Kajiya, H., Huang, N., *et al.* (1998). A high-density rice genetic linkage map with 2275 markers using a single F_2 population. *Genetics* **148**, 479–494.

Hash, C. T. (2005). Opportunities for application of molecular markers for sustainable crop production in stress environments: Sorghum and pearlmillet. *In* "International Conference on Sustainable Crop Production in Stress Environments: Management and Genetic Options," pp. 113 (abstract). Jawaharlal Nehru Krishi Vishwa Vidyalaya, Jabalpur, India.

Hash, C. T., Rizivi, S. M. H., Serraj, S., Bidinger, F., Vadez, V., Sharma, A., Howarth, C. J., and Yadav, R. S. (2004). Field assessment of back-cross derived hybrids validates a major pearlmillet drought tolerance QTL. *In* "49th Annual Meeting of Crop Science Society of America," p. 211 (abstract). Seattle, Washington, USA.

Haussmann, B. I. G., Hess, D. E., Seetharama, N., Welz, H. G., and Geiger, H. H. (2002). Construction of a combined sorghum linkage map from two recombinant inbred populations using AFLP, SSR, RFLP, and RAPD markers, and comparison with other sorghum maps. *Theor. Appl. Genet.* **105**, 629–637.

Hayama, R., and Coupland, G. (2004). The molecular basis of diversity in the photoperiodic flowering responses of *Arabidopsis* and rice. *Plant Physiol.* **135**, 677–684.

Hazen, S. P., Pathan, M. S., Sanchez, A., Baxter, I., Dunn, M., Estes, B., Chang, H.-S., Zhu, T., Kreps, J. A., and Nguyen, H. T. (2005). Expression profiling of rice segregating for drought tolerance QTL using a rice genome array. *Funct. Integr. Genomics* **5**, 104–116.

Hedden, P. (2003). The genes of the green revolution. *Trends Genet.* **19**, 5–19.

Hehl, R., Faurie, E., Hesselbach, J., Salamini, F., Whitham, S., Baker, B., and Gebhardt, C. (1999). TMV resistance gene *N* homologues are linked to *Synchytrium endobioticum* resistance in potato. *Theor. Appl. Genet.* **98**, 379–386.

Helguera, M., Soria, V. M., Khan, I. A., Kolmer, J., and Dubcovsky, J. (2005). PCR markers for *Triticum speltoides* leaf rust resistance gene *Lr51* and their use to develop isogenic hard red spring wheat lines. *Crop Sci.* **45**, 728–734.

Henderson, C. R. (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics* **31**, 423–447.

Hernández-Sánchez, J., Visscher, P., Plastow, G., and Haley, C. (2003). Candidate gene analysis for quantitative traits using the transmission disequilibrium test: The example of the Melanocortin 4- receptor in pigs. *Genetics* **164**, 637–644.

Heun, M., Kennedy, A. E., Anderson, J. A., Lapitan, N. L. V., Sorrels, M. E., and Tamksley, S. D. (1996). Construction of a restriction fragment length polymorphism map of barley (*Hordeum vulgare*). *Genome* **34**, 437–447.

Hittalmani, S., Parco, A., Mew, T. V., Zeigler, R. S., and Huang, N. (2000). Fine mapping and DNA marker-assisted pyramiding of the three major genes for blast resistance in rice. *Theor. Appl. Genet.* **100**, 1121–1128.

Holland, J. B. (1998). EPISTACY: A SAS program for detecting tow-locus epistasis interactions using genetic marker information. *J. Hered.* **89**, 374–375.

Holland, J. B. (2004). Implementation of molecular markers for quantitative traits in breeding programs—challenges and opportunities. *In* "New Directions for a Diverse Planet." Proceedings of Fourth International Crop Science Congress, Brisbane, Australia.

Hori, K., Kobayashi, T., Shimizu, A., Sato, K., Takeda, K., and Kawasaki, S. (2003). Efficient construction of high-density linkage map and its application to QTL analysis in barley. *Theor. Appl. Genet.* **107**, 806–813.

Hospital, F., and Charcosset, M. (1997). Marker-assisted introgression of quantitative loci. *Genetics* **147**, 1469–1485.

Hospital, F., and Decoux, G. (2002). Popmin: A program for the numerical optimization of population sizes in marker-assisted programs. *J. Hered.* **93**, 383–384.

Hospital, F., Chevalet, C., and Mulsant, P. (1992). Using markers in gene introgression breeding programs. *Genetics* **132**, 1119–1210.

Hospital, F., Moreau, L., Lacoudre, F., Charcosset, A., and Gallais, A. (1997). More on the efficiency of marker-assisted selection. *Theor. Appl. Genet.* **95**, 1181–1189.

Hu, J., and Vick, B. A. (2003). Target region amplification polymorphism: A novel marker technique for plant genotyping. *Plant Mol. Biol. Rep.* **21**, 289–294.

Hu, J.-J., Nakatani, M., Mizuno, K., and Fujimura, T. (2004a). Development and characterization of microsatellite markers in sweetpotato. *Breed. Sci.* **54**, 177–188.

Hu, J.-J., Nakatani, M., Lalusin, A.-G., and Fujimura, T. (2004b). New microsatellite markers developed from reported *Ipomoea trifida* sequences and their application to sweetpotato and its related wild species. *Sci. Hortic.* **102**, 375–386.

Hua, J., Xing, Y., Wu, W., Xu, C., Sun, X., Yu, S., and Zhang, Q. (2003). Single-locus heterotic effects and dominance by dominance interactions can adequately explain the genetic basis of heterosis in an elite rice hybrid. *Proc. Natl. Acad. Sci. USA* **100**, 2574–2579.

Huamán, Z., Aguilar, C., and Ortiz, R. (1999). Selecting a Peruvian sweetpotato core collection on the basis of morphological, eco-geographical, and disease and pest reaction data. *Theor. Appl. Genet.* **98**, 840–844.

Huamán, Z., Ortiz, R., Zhang, D., and Rodríguez, F. (2000). Isozyme analysis of entire and core collection of *Solanum tuberosum* subsp. *andigena* potato cultivars. *Crop Sci.* **40**, 273–276.

Huang, N., Angeles, E. R., Domingo, J., Magpantay, G., Singh, S., Zhang, G., Kumaravadivel, N., Bennet, J., and Khush, G. S. (1997). Pyramiding of bacterial blight resistance genes in rice: Marker- assisted selection using RFLP and PCR. *Theor. Appl. Genet.* **95**, 313–320.

Huang, X. Q., Hsam, S. L. K., Zeller, F. J., Wenzel, G., and Mohler, V. (2000). Molecular mapping of the wheat powdery mildew resistance gene *Pm24* and marker validation for molecular breeding. *Theor. Appl. Genet.* **101**, 407–414.

Igartua, E., Edney, M., Rossnagel, B. G., Spaner, D., Legge, W. G., Scoles, G. J., Eckstein, P. E., Penner, G. A., Tinker, N. A., Briggs, K. G., Falk, D. E., and Mather, D. E. (2000). Marker-based selection of QTL affecting grain and malt quality in two-row barley. *Crop Sci.* **40**, 1426–1433.

Ikeda, A., Ueguchi-Tanaka, M., Sonoda, Y., Kitano, H., Koshioka, M., Futsuhara, Y., Matsuoka, M., and Yamaguchi, J. (2001a). *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.

Ikeda, N., Bautista, N. S., Yamada, T., Kamijima, O., and Ishii, T. (2001b). Ultra-simple DNA extraction method for marker-assisted selection using microsatellite markers in rice. *Plant Mol. Biol. Rep.* **19**, 27–32.

IRGSP (2005). The map-based sequence of the rice genome. *Nature* **436**, 793–800.

Ishimaru, K. (2003). Identification of a locus increasing rice yield and physiological analysis of its function. *Plant Physiol.* **122**, 1083–1090.

Izawa, T., Takahashi, Y., and Yano, M. (2003). Comparative biology comes into bloom: Genomic and genetic comparison of flowering pathways in rice and *Arabidopsis*. *Curr. Opin. Plant Biol.* **6**, 113–120.

Jaccoud, D., Peng, K., Feinstein, D., and Kilian, A. (2001). Diversity arrays: A solid state technology for sequence information independent genotyping. *Nucleic Acids Res.* **29**, e25.

Jannink, J. L., and Walsh, B. (2002). Association mapping in plant populations. *In* "Quantitative Genetics, Genomics and Plant Breeding" (M. S. Kong, Ed.), pp. 59–68. CABI, Wallingford, UK.

Jansen, R. C., and Nap, J.-P. (2001). Genetical genomics: The added value from segregation. *Trends Genet.* **17**, 388–391.

Jarret, R. L., and Bowen, N. (1994). Simple sequence repeats (SSRs) for sweetpotato germplasm characterization. *Plant Genet. Resour. Newsl.* **100**, 9–11.

Jefferies, S. P., Pallotta, M. A., Paull, J. G., Karakousis, A., Kretschmer, J. M., Manning, S., Islam, A. K. M. R., Langridge, P., and Chalmers, K. J. (2000). Mapping and validation of chromosome regions conferring boron toxicity tolerance in wheat (*Triticum aestivum*). *Theor. Appl. Genet.* **101**, 767–777.

Jefferies, S. P., King, B. J., Barr, A. R., Warner, P., Logue, S. J., and Langridge, P. (2003). Marker-assisted backcross introgression of the *Yd2* gene conferring resistance to barley yellow dwarf virus in barley. *Plant Breed.* **122**, 52–56.

Ji, W., Li, Y., Li, J., Dai, C.-H., Wang, X., Bai, X., Cai, H., Wang, L., and Zhu, Y.-M. (2006). Generation and analysis of expressed sequence tags from NaCL-treated *Glycine soja*. *BMC Plant Biol.* **6**, 4.

Jiang, G. H., Xu, C. G., Tu, J. M., Li, X. H., He, Y. Q., and Zhang, Q. F. (2004). Pyramiding of insect- and disease-resistance genes into an elite *indica*, cytoplasm male sterile restorer line of rice, Minghui 63. *Plant Breed.* **123**, 112–116.

Jin, C., Lan, H., Attie, A. D., Bulutuglo, D., Churchill, G. A., and Yandell, B. S. (2004). Selective phenotyping for increased efficiency in genetic mapping studies. *Genetics* **168**, 2285–2293.

Johnson, G. R. (2004). Marker assisted selection. *Plant Breed. Rev.* **24**(Pt. 1), 293–310.

Joseph, M., Gopalakrishnan, S., Sharma, R. K., Singh, V. P., Singh, A. K., Singh, N. K., and Mohapatra, T. (2003). Combining bacterial blight resistance and Basmati quality characteristics by phenotypic and molecular marker-assisted selection in rice. *Mol. Breed.* **13**, 1–11.

Jourjon, M. F., Jasson, S., Marcel, J., Ngom, B., and Mangin, B. (2005). MCQTL: Multi-allelic QTL mapping in multi-cross design. *Bioinformatics* **21**, 128–130.

Juliano, B. O. (1985). "Rice Chemistry and Technology," 2nd ed. American Association of Cereal Chemists, Incorporated, St. Paul, Minnesota, USA.

Kanazin, V., Talbert, H., See, D., DeCamp, P., Nevo, E., and Glake, T. (2002). Discovery and single- nucleotide polymorphisms in barley (*Hordeum vulgare* L.). *Plant Mol. Biol.* **48**, 529–537.

Kapoor, R. L., Karkar, P. S., Khairwal, I. S., Bainiwal, C. R., Nijhawan, D. C., and Yadav, H. P. (1989). Bajra hybrid HHB67: A major breakthrough. *Haryana Farm.* **18**, 17–21.

Kasha, K. J., and Kao, K. N. (1970). High frequency haploid production in barley (*Hordeum vulgare* L.). *Nature* **225**, 874–876.

Katiyar, S. K., and Bennett, J. (2001). Biotechnology for gall midge resistance: From molecular tagging to gene pyramiding. *In* "Rice Research for Food Security and Poverty Alleviation" (S. Peng and B. Hardy, Eds.), pp. 692. Proceedings of International Rice Research Conference 31 March–3 April 2000.

Keating, B. A., Carberry, P. S., Hammer, G. L., Probert, M. E., Robertson, M. J., Holzworth, D., Huth, N. I., Hargreaves, J. N. G., Meinke, H., Hochman, Z., McLean, G., Verburg, K., *et al.* (2003). An overview of APSIM, a model designed for farming system simulation. *Eur. J. Agron.* **18**, 267–288.

Kelly, J. D., Gepts, P., Miklas, P. N., and Coyne, D. P. (2003). Tagging and mapping of genes and QTL molecular marker-assisted selection for traits of economic importance in bean and cowpea. *Field Crop Res.* **82**, 135–154.

Kendziorski, C. M., Chen, M., Yuan, M., Lan, H., and Attie, A. D. (2006). Statistical methods for expression quantitative trait loci (eQTL) mapping. *Biometrics* **62**, 19–27.

Kirst, M., Myburg, A. A., De León, J. P. G., Kirst, M. E., Scott, J., and Sederoff, R. (2004). Coordinated genetic regulation of growth and lignin revealed by quantitative trait locus analysis of cDNA microarray data in an interspecific backcross of eucalyptus. *Plant Physiol.* **135**, 2368–2378.

Klein, P. E., Klein, R. R., Cartinhour, S. W., Ulanch, P. E., Dong, J., Obert, J. A., Morishige, D. T., Schlueter, S. D., Childs, K. L., Ale, M., and Mullet, J. E. (2000). A high-throughput AFLP-based method for constructing integrated genetic and physical maps: Progress toward a sorghum genome map. *Genome Res.* **10**, 789–807.

Klimyuk, V. I., Carroll, B. J., Thomas, C. M., and Jones, J. D. G. (1993). Alkali treatment for rapid preparation of plant material for reliable PCR analysis. *Plant J.* **3**, 493–494.

Klos, K. L. E., Paz, M. M., Marek, L. P., Cregan, P. B., and Shoemaker, R. C. (2000). Molecular markers useful for detecting resistance to brown stem rot in soybean. *Crop Sci.* **40**, 1445–1452.

Klose, J., Nock, C., Herrmann, M., Stühler, K., Marcus, K., Blüggel, M., Krause, E., Schalkwyk, L. C., Rastan, S., Brown, S. D. M., Büssow, K., Himmelbauer, H., *et al.* (2002). Genetic analysis of mouse brain proteome. *Nat. Genet.* **30**, 385–393.

Knapp, S. J. (1998). Marker-assisted selection as a strategy for increasing the probability of selecting superior genotypes. *Crop Sci.* **38**, 1164–1174.

Kobayashi, S., Araki, E., Osaki, M., Khush, G. S., and Fukuta, Y. (2006). Localization, validation and characterization of plant-type QTLs on chromosomes 4 and 6 in rice (*Oryza sativa* L.). *Field Crop Res.* **96**, 106–112.

Kobiljski, B., Quarrie, S., Denčić, S., Kirby, J., and Ivegeš, M. (2002). Genetic diversity of the Novi Sad wheat core collection revealed by microsatellites. *Cell. Mol. Biol.* **7**, 685–694.

Koebner, R. (2003). MAS in cereals: Green for maize, amber for rice, still red for wheat and barley. Marker Assisted Seletion: A fast track to increase genetic gain in plant and animal breeding?Session 1: MAS in plants (http://www.fao.org/biotech/docs/koebner.pdf).

Koebner, R. M. D., Powell, W., and Donini, P. (2001). Contributions of DNA molecular marker technologies to the genetics and breeding of wheat and barley. *Plant Breed. Rev.* **21**, 181–220.

Kojima, S., Takahashi, Y., Kobayashi, Y., Monna, L., Sasaki, T., Araki, T., and Yano, M. (2002). *Hd3a*, a rice ortholog of the *Arabidopsis FT* gene, promotes transition to flowering downstream of *Hd1* under short-day conditions. *Plant Cell Physiol.* **43**, 1096–1105.

Kollipara, K. P., Saab, I. N., Wych, R. D., Lauer, M. J., and Singletary, G. W. (2002). Expression profiling of reciprocal maize hybrids divergent for cold germination and desiccation tolerance. *Plant Physiol.* **129**, 974–992.

Kong, L., Dong, J., and Hart, G. E. (2000). Characteristics, linkage map positions, and allelic differentiation of *Sorghum bicolor* (L.) Moench DNA simple-sequence repeats (SSRs). *Theor. Appl. Genet.* **101**, 438–448.

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

Konishi, S., Izawa, T., Lin, S. Y., Ebana, K., Fukuta, Y., Sasaki, T., and Yano, M. (2006). An SNP caused loss of seed shattering during rice domestication. *Science* **312**, 1392–1396.

Kosambi, D. D. (1944). The estimation of map distances from recombination values. *Ann. Eugen.* **12**, 172–175.

Kota, R., Rudd, S., Facius, A., Kolesov, G., Theil, T., Zhang, H., Stein, N., Mayer, K., and Graner, A. (2003). Snipping polymorphisms from large EST collections in barley (*Hordeum vulgare* L.). *Mol. Genet. Genomics* **270**, 24–33.

Kraakman, A. T. W., Martínez, F., Mussiraliev, B., van Eeuwijk, F. A., and Niks, R. E. (2006). Linkage disequilibrium mapping of morphological, resistance, and other agronomically relevant traits in modern spring barley cultivars. *Mol. Breed.* **17**, 41–58.

Kriegner, A., Cervantes, J. C., Burg, K., Mwanga, R. O. M., and Zhang, D. P. (2003). A genetic linkage map of sweetpotato (*Ipomoea batatas* (L) Lam) based on AFLP markers. *Mol. Breed.* **11**, 169–185.

Kroymann, J., and Mitchell-Olds, T. (2005). Epistasis and balanced polymorphism influencing complex trait variation. *Nature* **435**, 95–98.

Kuchel, H., Ye, G.-Y., Fox, R., and Jefferies, S. (2005). Genetic and economic analysis of a targeted marker-assisted wheat breeding strategy. *Mol. Breed.* **16**, 67–78.

Kumpatla, S. P., and Mukhopadhyay, S. (2005). Mining and survey of simple sequence repeats in expressed sequence tags of dicotyledonous species. *Genome* **48**, 985–998.

Kwon, Y.-S., Eun, M. Y., and Sohn, J.-K. (2001). Marker-assisted selection for identification of plant regeneration ability of seed-derived calli in rice (*Oryza sativa* L.). *Mol. Cells* **12**, 103–106.

Kynast, R. G., Riera-Lizarazu, O., Vales, M. I., Okagaki, R. J., Maquieira, S. B., Chen, G., Ananiev, E. V., Odland, W. E., Russell, C. D., Stec, A. O., Livingston, S. M., Zaia, H. A., *et al.* (2001). A complete set of monogenic individual chromosome additions to the oat genome. *Plant Physiol.* **125**, 1216–1127.

Lande, R., and Thompson, R. (1990). Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* **124**, 743–756.

Lander, E. S., Green, P., Abrahamson, J., Barlow, A., Daly, M. J., Lincoln, S. E., and Newburg, L. (1987). MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* **1**, 174–181.

Lander, E. S., Linton, M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., *et al.* (2001). Initial sequencing and analysis of human genome. *Nature* **409**, 860–921.

Landi, P., Sanguineti, M. C., Salvi, S., Giuliani, S., Billotti, M., Maccaferri, M., Conti, S., and Tuberosa, R. (2005). Validation and characterization of a major QTL affecting leaf ABA concentration in maize. *Mol. Breed.* **15**, 291–303.

Langridge, P., Lagudah, E. S., Holton, T. A., Appels, R., Sharp, P. J., and Chalmers, K. J. (2001). Trends in genetic and genome analyses in wheat: A review. *Aust. J. Agric. Res.* **52**, 1043–1077.

Larson, S. R., Kadyrzhanova, D., McDonald, C., Sorrells, M., and Blake, T. K. (1996). Evaluation of barley chromosome-3 yield QTL in a backcross F_2 population using STS-PCR. *Theor. Appl. Genet.* **93**, 618–625.

Latha, R., Rubia, L., Bennett, J., and Swaminathan, M. S. (2004). Allele mining for stress tolerance genes in *Oryza* species and related germplasm. *Mol. Biotechnol.* **27**, 101–108.

Law, C. N., Worland, A. J., and Giorgi, B. (1975). The genetic control of ear emergence time by chromosomes 5A and 5D of wheat. *Heredity* **36**, 49–58.

Lee, G. J., Boerma, H. R., Villagarcia, M. R., Zhou, X., Carter, T. E., Jr., Li, Z., and Gibbs, M. O. (2004). A major QTL conditioning salt tolerance in S-100 soybean and descendent cultivars. *Theor. Appl. Genet.* **109**, 1610–1619.

Leung, H., Wu, J., Liu, B., Bustaman, M., Sridhar, R., Singh, K., Redona, E., Quang, V. D., Zhang, K., Bernardo, M., Wang, G., Leach, J., *et al.* (2004). Suscetainable disease resistance in rice: Current and future strategies. (http://www.cropscience org.au/icsc2004/symposia/3/7/1022_leungh.htm).

Li, C., Zhou, A., and Sang, T. (2006b). Rice domestication by reducing shattering. *Science* **311**, 1936–1939.

Li, C. D., Rossnagel, B. G., and Scoles, G. J. (2000). The development of oat microsatellite markers and their use in identifying relationships among *Avena* species and oat cultivars. *Theor. Appl. Genet.* **101**, 1259–1268.

Li, J., Thomson, M., and McCouch, S. R. (2004a). Fine mapping of a grain-weight quantitative trait locus in the pericentromeric region of rice chromosome 3. *Genetics* **168**, 2187–2195.

Li, J. Z., Sjakste, T. G., Röder, M. S., and Ganal, M. W. (2003b). Development and genetic mapping of 127 new microsatellite markers in barley. *Theor. Appl. Genet.* **107**, 1021–1027.

Li, Y., Shi, Y., Cao, Y., and Wang, T. (2004b). Establishment of a core collection for maize germplasm preserved in Chinese national gene bank using geographic distribution and characterization data. *Genet. Resour. Crop Evol.* **51**, 845–852.

Li, Z., Jakkula, L., Hussey, R. S., Tamulonis, J. P., and Boerma, H. R. (2001a). SSR mapping and confirmation of the QTL from PI96354 conditioning soybean resistance to southern root-knot nematode. *Theor. Appl. Genet.* **103**, 1167–1173.

Li, Z. K., Luo, L. J., Mei, H. W., Wang, D. L., Shu, Q. Y., Tabien, R., Zhong, D. B., Ying, C. S., Stansel, J. W., Khush, G. S., and Patrson, A. H. (2001b). Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. I. Biomass and grain yield. *Genetics* **158**, 1737–1753.

Li, Z. K., Yu, S. B., Lafitte, H. R., Huang, N., Courtois, B., Hittalmani, S., Vijayakumar, C. H. M., Liu, G. F., Wang, G. C., Shashidhar, H. E., Zhuang, J. Y., Zheng, K. L., *et al.* (2003a). QTL x environment interactions in rice. I. Heading date and plant height. *Theor. Appl. Genet.* **108**, 141–153.

Li, Z.-K., Fu, B.-Y., Gao, Y.-M., Xu, J.-L., Ali, J., Lafitte, H. R., Jiang, Y.-Z., Rey, D., Vijayakumar, C. H. M., Maghirang, R., Zheng, T.-Q., and Zhu, L.-H (2005). Genome-wide introgression lines and their use in genetic and molecular dissection of complex phenotypes in rice (*Oryza sativa* L.). *Plant Mol. Biol.* **59**, 33–52.

Li, Z. K., Arif, M., Zhong, D. B., Fu, B. Y., Xu, J. L., Domingo-Rey, J., Ali, J., Vijayakumar, C. H. M., Yu, S. B., and Khush, G. S. (2006a). Complex genetic networks underlying the defensive system of rice (*Oryza sativa* L.) to *Xanthomonas oryzae* pv. oryzae. *Proc. Natl. Acad. Sci. USA* **103**, 7994–7999.

Liang, F., Deng, Q., Wang, Y., Xiong, Y., Jin, D., Li, J., and Wang, B. (2004). Molecular marker-assisted selection for yield-enhancing genes in the progeny of "9311 x *O. rufipogon*" using SSR. *Euphytica* **139**, 159–165.

Lilley, J. M., Ludlow, M. M., McCouch, S. R., and O'Toole, J. C. (1996). Locating QTL for osmotic adjustment and dehydration tolerance in rice. *J. Exp. Bot.* **47**, 1427–1436.

Lin, S. C., and Min, S. K. (1991). "Rice Varieties and Their Genealogy in China (in Chinese)." Sanghai Science and Technolgy Press, Sanghaai.

Lin, Y.-R., Draye, X., Qian, X., Ren, S., Zhu, L.-H., Tomkins, J., Wing, R., Li, Z., and Paterson, A. H. (2000). Locus-specific contig assembly in highly-duplicated genomes, using the BAC-RF method. *Nucleic Acid Res.* **28**, e23.

Lindblad-Toh, K., Winchester, E., Daly, M., Wang, D., Hirschhorn, J. N., Laviolette, J. P., Ardie, K., Reich, D. E., Robinson, E., Sklar, P., Shah, N., Thomas, D., *et al.* (2000). Large-scale discovery and genotyping of single-nucleotide polymorphisms in the mouse. *Nat. Genet.* **24**, 381–386.

Little, R. R., Hilder, G. B., and Dawson, E. H. (1958). Differential effects of dilute alkali on 25 varieties of milled white rice. *Cereal Chem.* **35**, 111–126.

Liu, S., and Anderson, J. A. (2003). Marker-assisted evaluation of *Fusarium* head blight resistant wheat germplasm. *Crop Sci.* **43**, 760–766.

Liu, B., Zhang, S. H., Zhu, X. Y., Yang, Q. Y., Wu, S. Z., Mei, M. T., Mauleon, R., Leach, J., Mew, T., and Leung, H. (2004b). Candidate defense genes as predictors of quantitative blast resistance in rice. *Mol. Plant Microbe. Interact.* **17**, 1146–1152.

Liu, F., von Bothmer, R., and Salomon, B. (1999). Genetic diversity among East Asian accessions of the barley core collection as revealed by six isozyme loci. *Theor. Appl. Genet.* **98**, 1226–1233.

Liu, F., von Bothmer, R., and Salomon, B. (2000a). Genetic diversity in Eurpoean accessions of the barley core collection as detected by isozyme electrophoresis. *Genet. Resour. Crop Evol.* **47**, 571–581.

Liu, F., Sun, G.-L., Salomon, B., and von Bothmer, R. (2001a). Distirbution of allozymic alleles and genetic diversity in the American barley core collection. *Theor. Appl. Genet.* **102**, 606–615.

Liu, J., Liu, D., Tao, W., Li, W., Wang, S., Chen, P., Cheng, S., and Gao, D. (2000b). Molecular marker-facilitated pyramiding of different genes for powdery mildew resistance in wheat. *Plant Breed.* **119**, 21–24.

Liu, J. S., Sabatti, C., Teng, J., Keats, B. J. B., and Risch, K. (2001b). Bayesian analysis of haplotypes for linkage disequilibrium mapping. *Genome Res.* **11**, 1716–1724.

Liu, P. Y., Zhu, J., and Lu, Y. (2004a). Marker-assisted selection in segregating generations of self- fertilizing crops. *Theor. Appl. Genet.* **109**, 370–376.

Liu, Q.-Q., Li, Q.-F., Cai, X.-L., Wang, H.-M., Tang, S.-Z., Yu, H.-X., Wang, Z.-Y., and Gu, M.-H. (2006). Molecular marker-assisted selection for improved cooking and eating quality of two elite parents of hybrid rice. *Crop Sci.* **46**, 2354–2360, doi:10.2135/cropsci2006.03.0180.

Liu, S., Griffey, C. A., and Saghai Maroof, M. A. (2001c). Identification of molecular markers associated with adult plant resistance to powdery mildew in common wheat cultivar Massey. *Crop Sci.* **41**, 1268–1275.

Liu, X., Gu, M., Han, Y., Ji, Q., Gu, S., Zhang, R., Li, X., Chen, J., Korban, S. S., and Xu, M. (2004c). Developing gene tagged markers for functional analysis of starch-synthesizing genes in rice (*Oryza sativa* L.). *Euphytica* **135**, 345–353.

Liu, Z., Anderson, J. A., Hu, J., Friesen, T. L., Rasmussen, J. B., and Faris, J. D. (2005). A wheat intervarietal linkage map based on microsatellite and target region amplified polymorphism markers and its utility for detecting quantitative trait loci. *Theor. Appl. Genet.* **111**, 782–794.

Liu, Z.-W., Biyashev, R. M., and Saghai Maroof, M. A. (1996). Development of simple sequence repeat DNA markers and their integration into a barley linkage map. *Theor. Appl. Genet.* **93**, 869–876.

Lopez, C., Jorge, V., Piégu, B., Mba, C., Cortes, D., Restrepo, S., Soto, M., Laudié, M., Berger, C., Cooke, R., Delseny, M., Tohme, J., *et al.* (2004). A unique catalogue of 5700 expressed genes in cassava. *Plant Mol. Biol.* **56**, 541–554.

Lopez, C., Piégu, B., Cooke, R., Delseny, M., Tohme, J., and Verdier, V. (2005). Using cDNA and genomic sequences as tools to develop SNP strategies in cassava (*Manihot esculenta* Crantz). *Theor. Appl. Genet.* **110**, 425–431.

Lozoya-Saldana, H., Belmar-Diaz, C., Bradeen, J. M., and Helgeson, J. P. (2005). Chatacterization of *Phytophthora infestans* isolates infecting transgenic and somatic hybrid potatoes resistant to pathogen in the Toluca Valley, Mexico. *Am. J. Potato Res.* **82**, 79(abstract).

Lu, C., and Zhou, J. (2000). Breeding and utilization of two-line interspecific hybrid rice Liangyou Peijiu. *Hybrid Rice* **15**, 4–5.

Lu, X., Niu, T., and Liu, J. S. (2003). Haplotype information and linkage disequilibrium mapping for single nucleotide polymorphisms. *Genome Res.* **13**, 2112–2117.

Lübberstedt, T., Melchenger, A. E., Fähr, S., Klein, D., Dally, A., and Westhoff, P. (1998). QTL mapping in test crosses of flint lines of maize: III. Comparison across populations for forage traits. *Crop Sci.* **38**, 1278–1289.

Lübberstedt, T. L., Zein, I., Andersen, J. R., Wenzel, G., Krützfeldt, B., Eder, J., Quzunova, M., and Chun, S. (2005). Development and application of functional markers in maize. *Euphytica* **146**, 101–108.

Lund, M. S., Sorensen, P., Gulbrandsten, B., and Sorensen, D. A. (2003). Multitrate fine mapping of quantitative trait loci using combined linkage disequilibria and linkage analysis. *Genetics* **163**, 405–410.

Luo, L. J., Li, Z. K., Mei, H. W., Shu, Q. Y., Tabien, R., Zhong, D. B., Ying, C. S., Stansel, J. W., Khush, G. S., and Paterson, A. H. (2001). Overdominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. II. Grain yield components. *Genetics* **158**, 1755–1771.

Lynch, M. N., and Walsh, B. (1998). "Genetics Analysis of Quantitative Traits." Sinauer Associates, Sunderland, MA.

Maccaferri, M., Sanguineti, M. C., Noli, E., and Tuberosa, R. (2005). Population structure and long-range linkage disequilibrium in a durum wheat elite collection. *Mol. Breed.* **15**, 271–289.

Mackill, D. J. (2006). Breeding for resistance to abiotic stresses in rice: The value of quantitative trait loci. *In* "Plant Breeding: The Arnel R. Hallauer International Symposium" (K. R. Lamkey and M. Lee, Eds.), pp. 201–212. Blackwell Pub, Ames, IA.

Mackill, D. J., and McNally, K. L. (2004). A model crop species: Molecular markers in rice. *In* "Molecular Marker Systems in Plant Preeding and Crop Improvement" (H. Lörz and G. Wenzel, Eds.), Vol. 55, pp. 39–54. Springer Verlag, Heidelberg.

Mahalakshmi, V., Ng, N., Lawson, M., and Ortiz, R. (2007a). Cowpea [*Vigna unguiculata* (L.) Walp.] core collection defined by geographical and agro-botanical discriptors. *Plant Genet. Resour. :Characterization and Utilization.* (in press).

Mahalakshmi, V., Ng, Q., Atalobhor, J., Ogunsola, D., Lawson, M., and Ortiz, R. (2007b). Development of a West African Yam *Dioscorea* spp. core collection. *Genet. Resour. Crop Evol.,* doi: 10.1007/S10722-006-92034. (in press).

Malosetti, M., and Abadie, T. (2001). Sampling strategy to develop a core collection of Uruguyan maize landraces based on morphological traits. *Genet. Resour. Crop Evol.* **48**, 381–390.

Malysheva-Otto, L. V., Ganal, M. W., and Röder, M. S. (2006). Analysis of molecular diversity, population structure and linkage disequilibrium in a worldwide survey of cultivated barley germplasm (*Hordeum vulgare* L.). *BMC Genet.* **7**, 6.

Manly, K. F. (1993). A Macintosh program for storage and analysis of experimental genetic mapping data. *Mamm. Genome* **4**, 303–313.

Manly, K. F., and Olsen, J. M. (1999). Overview of QTL mapping software and introduction to Map Manager QT. *Mamm. Genome* **10**, 327–334.

Manly, K. F., Cudmore, R. H., and Meer, J. M. (2001). Map Manager QTX, cross-paltform software for genetic mapping. *Mamm. Genome* **12**, 930–932.

Martienssen, R. A., Rabinowicz, P. D., O'Shaughnessy, A., and McCombie, W. R. (2004). Sequencing the maize genome. *Curr. Opin. Plant Biol.* **7**, 102–107.

Martin, E. R., Monks, S. A., Warren, L. L., and Kaplan, N. L. (2000). A test for linkage and association in general pedigrees: The pedigree disequilibrium test. *Am. J. Hum. Genet.* **67**, 146–154.

Matsumura, H., Reich, S., Ito, A., Saitoh, H., Kamoun, S., Winter, P., Kahl, G., Reuter, M., Kruger, D. H., and Terauchi, R. (2003). Gene-expression analysis of plant host-pathogen interactions by SuperSAGE. *Proc. Natl. Acad. Sci. USA* **100**, 15718–15723.

Matus, I., Corey, A., Filichkin, T., Hayes, P. M., Vales, M. I., Kling, J., Riera-Lizarazu, O., Sato, K., Powell, W., and Waugh, R. (2003). Development and characterization of recombinant chromosome substitution lines (RCSLs) using *Hordeum vulgare* subsp. *spontaneum* as a source of donor alleles in a *Hordeum vulgare* subsp. *vulgare* background. *Genome* **46**, 1010–1023.

Matz, E. C., Burr, F. A., and Burr, B. (1995). Molecular map based on TxCM and COxTx recombinant inbred families. *Maize Genet. Coop. Newsl.* **69**, 257–267.

Mazur, B. J. (1995). Commercializing the products of plant biotechnology. *Trends Biotechnol.* **13**, 319–323.

Mba, R. E. C., Stephenson, P., Edwards, K., Melzer, S., Nkumbira, J., Gullberg, U., Apel, K., Gale, M., Tohme, J., and Fregene, M. (2001). Simple sequence repeat (SSR) markers survey of the cassava (*Manihot esculenta* Crantz) genome: Towards an SSR-based molecular genetic map of cassava. *Theor. Appl. Genet.* **102**, 21–31.

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

McLauchlan, A., Ogbonnaya, F. C., Hollingsworth, B., Carter, M., Gale, K. R., Henry, R. J., Holton, T. A., Morell, M. K., Rampling, L. R., Sharp, P. J., Shariflou, M. R., Jones, M. G. K., *et al.* (2001). Development of robust PCR-based DNA markers for each homoeo-allele of granule- bound starch synthase and their application in wheat breeding programs. *Aust. J. Agric. Res.* **52**, 1409–1416.

McNally, K. L., Wang, H., Naredo, E. B., Raghavan, C., Atienza, G., and Leung, H. (2006). "EcoTILLING in rice. Plant and Animal Genomes XIV Conference. W 258 (Allele mining)". San Diego, CA.

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

Melchinger, A. E., and Gumber, P. K. (1998). Overview of heterosis and heterotic groups in agronomic crops. *In* "Concepts and Breeding of Heterosis in Crop Plants" (K. R. Lamkey and J. E. Staub, Eds.), pp. 29–44. CSSA, Madison, WI.

Melchinger, A. E., Utz, H. F., and Schön, C. C. (1998). Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. *Genetics* **149**, 383–403.

Melotto, M., Afanador, L., and Kelly, J. D. (1996). Development of a SCAR marker linked to the *I* gene in common bean. *Genome* **39**, 1216–1219.

Melotto, M., Monteiro-Vitorello, C. B., Bruschi, A. G., and Camargo, L. E. A. (2005). Competitive bioinformatics analysis of genes expressed in common bean (*Phaseolus vulgaris* L.) seedlings. *Genome* **48**, 562–570.

Menz, M. A., Klein, R. R., Mullet, J. E., Obert, J. A., Unruh, N. C., and Klein, P. E. (2002). A high-density genetic map of *Sorghum bicolor* (L.) Moench based on 2926 AFLP, RFLP, and SSR markers. *Plant Mol. Biol.* **48**, 483–499.

Mertz, E. T., Bates, L. S., and Nelson, O. E. (1964). Mutant gene that changes protein composition and increases lysine content of maize endosperm. *Science* **145**, 279–280.

Mesfin, A., Smith, K. P., Dill-Macky, R., Evans, C. K., Waugh, R., Gustus, C. D., and Muehlbauer, G. J. (2003). Quantitative trait loci for Fusarium head blight resistance in barley detected in a two- rowed by six-rowed population. *Crop Sci.* **43**, 307–318.

Messina, C. D., Jones, J. W., Boote, K. J., and Vallejos, C. E. (2006). A gene-based model to simulate soybean development and yield response to environment. *Crop Sci.* **46**, 456–466.

Messing, J., Bharti, A. K., Karlowski, W. M., Gundblach, H., Kim, H. R., Yu, Y., Wei, F., Fuks, G., Soderlund, C. A., Mayer, K. F. X., and Wing, R. A. (2004). Sequence composition and genome organization of maize. *Proc. Natl. Acad. Sci. USA* **101**, 14349–14354.

Meuwissen, T. H. E., and Goddard, M. E. (2004). Mapping multiple QTL using linkage disequilibrium and linkage analysis information and multitrait data. *Genet. Sel. Evol.* **36**, 261–269.

Meuwissen, T. H. E., Hayes, B. J., and Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**, 1819–1829.

Meyer, R. C., Törjék, O., Becher, M., and Altmann, T. (2004). Heterosis of biomass production in *Arabidopsis*: Establishment during early development. *Plant Physiol.* **134**, 1813–1823.

Michalek, W., Weschke, W., Pleissner, K.-P., and Graner, A. (2002). EST analysis in barley defines a unigene set comprising 4,000 genes. *Theor. Appl. Genet.* **104**, 97–103.

Mignouna, H. D., Mank, R. A., Ellis, T. H. N., van den Bosch, N., Asiedu, R., Ng, S. Y. C., and Peleman, J. (2002a). A genetic linkage map of Guinaea yam (*Dioscorea rotundata* Poir.) based on AFLP markers. *Theor. Appl. Genet.* **105**, 716–725.

Mignouna, H. D., Abang, M. M., Onasanya, A., Agindotan, B., and Asiedu, R. (2002b). Identification and potential use of RAPD markers linked to *Yam mosaic virus* resistance in white yam *(Dioscorea rotundata)*. *Ann. Appl. Biol.* **140**, 163–169.

Mignouna, H. D., Abang, M. M., Onasanya, A., and Asiedu, R. (2002c). Identification and application of RAPD markers for anthracnose resistance in wateryam *(Dioscorea alata)*. *Ann. Appl. Biol.* **141**, 61–66.

Mignouna, H. D., Abang, M. M., and Asiedu, R. (2003a). Harnessing modern biotechnology for tropical tuber crop improvement: Yam (*Dioscorea* spp.) molecular breeding. *Afr. J. Biotechnol.* **2**, 478–485.

Mignouna, H. D., Abang, M. M., and Fagbemi, S. A. (2003b). A comparative assessment of molecular marker assays (AFLP, RAPD, and SSR) for white yam (*Discorea rotundata* Poir.) germplasm characterization. *Ann. Appl. Biol.* **142**, 269–276.

Miklas, P. N., and Kelly, J. D. (2002). The use of MAS to develop pinto bean germplasm possessing *Co-4*^{2} gene for anthracnose resistance. *Annu. Rep. Bean Improv. Coop.* **45**, 68–69.

Miklas, P. N., Larsen, R. C., Riley, R., and Kelly, J. D. (2000). Potential marker-assisted selection for *bc-1*^{2} resistance to bean common mosaic potyvirus in common bean. *Euphytica* **116**, 211–219.

Miklas, P. N., Kelly, J. D., and Singh, S. P. (2003). Registration of anthracnose-resistant pinto bean germplasm line USPT-ANT-1. *Crop Sci.* **43**, 1889–1890.

Miklas, P., Bosak, K., and Grafton, K. (2004). Marker-assisted backcrossing of QTL for resistance to Sclerotinia white mold in pinto bean (http://www.whitemoldresearch.com/presentation2004/Miklas.pdf).

Miklas, P. N., Kelly, J. D., Beebe, S. E., and Blair, M. W. (2006a). Common bean breeding for resistance against biotic and abiotic stresses: From classical to MAS breeding. *Euphytica* **147**, 105–131.

Miklas, P. N., Hu, J., Niklas, J. G., and Larsen, K. M. (2006b). Potential application of TRAP (Targeted Region Amplified Polymorphism) markers for mapping and tagging disease resistance traits in common bean. *Crop Sci.* **46**, 910–916.

Milborrow, B. V. (1998). A biochemical mechanism for hybrid vigour. *J. Exp. Bot.* **49**, 1063–1071.

Milbourne, D., Meyer, R. C., Collins, A. J., Ramsay, L. D., Gebhardt, C., and Waugh, R. (1998). Isolation, characterization and mapping of simple sequence repeat loci in potato. *Mol. Gen. Genet.* **259**, 233–245.

Miyahara, K. (1999). Analysis of LGC1, low glutelin mutant of rice. *Gamma Field Symp.* **38**, 43–52.

Mohan, M., Nair, S., Bhagwat, A., Krishna, T. G., Yano, M., Bhatia, C. R., and Sasaki, T. (1997). Genome mapping, molecular markers and marker-assisted selection in crop plants. *Mol. Breed.* **3**, 87–103.

Mohler, V., and Jahoor, A. (1996). Allele-specific amplification of polymorphic sites for the detection of powdery mildew resistance loci in cereals. *Theor. Appl. Genet.* **93**, 1078–1082.

Moreau, L., Charcosset, A., Hospital, F., and Gallais, A. (1998). Marker-assisted selection efficiency in populations of finite size. *Genetics* **148**, 1353–1365.

Moretzsohn, M. C., Leoi, L., Proite, K., Guimarães, P. M., Leal-Bertioli, S. C. M., Gimenes, M. A., Martins, W. S., Valls, J. F. M., Grattapadaglia, D., and Bertioli, D. J. (2005). A microsatellite-based, gene-rich linkage map for the AA genome of *Arachis* (Fabaceae). *Theor. Appl. Genet.* **111**, 1060–1071.

Morrell, P. L., Toleno, D. M., Lundy, K. E., and Clegg, M. T. (2005). Low levels of linkage disequilibrium in wild barley (*Hordeum vulgare* ssp. *spontaneum*) despite high rate of self fertilization. *Proc. Natl. Acad. Sci. USA* **102**, 2442–2447.

Morris, M., Dreher, K., Ribaut, J. M., and Khairallah, M. (2003). Money matters (II): Costs of maize inbred line conversion schemes at CIMMYT using conventional and marker-assisted selection. *Mol. Breed.* **11**, 235–247.

Mu, J., Zhou, H., Zhao, S., Xu, C., Yu, S., and Zhang, Q. (2004). Development of contiguous introgression lines covering entire genome of the sequenced J*aponica* rice. *In* "Proceedings of the Fourth International Crop Science Congress." Brisbane, Australia (http://www.cropscience.org.au/icsc2004).

Muehlbauer, G. J., Specht, J. E., Thomas-Compton, M. A., Staswick, P. E., and Bernard, R. L. (1988). Near-isogenic lines—A potential source in the integration of conventional and molecular genetic linkage maps. *Crop Sci.* **28**, 729–735.

Mutlu, N., Miklas, P. N., Steadman, J. R., Vidaver, A. K., Lindgren, D. T., Reiser, J., Coyne, D. P., and Paster-Corrales, M. A. (2005). Registration of common bacterial blight resistant pinto bean germplasm line ABCP-8. *Crop Sci.* **45**, 805–806.

Naess, S. K., Bradeen, J. M., Wielgus, S. M., Haberlach, G. T., McGrath, J. M., and Helgeson, J. P. (2000). Resistance to late blight in *Solanum bulbocastanum* is mapped to chromosome 8. *Theor. Appl. Genet.* **101**, 697–704.

Naik, S., Gill, K. S., Rao, V. S. P., Gupta, V. S., Tamhankar, S. A., Pujar, S., Gill, B. S., and Ranjekar, P. K. (1998). Identification of a STS marker linked to the *Aegilops speltoides*-derived leaf rust resistance gene *Lr28* in wheat. *Theor. Appl. Genet.* **97**, 535–540.

Nair, S., Bentur, J. S., Rao, U. P., and Mohan, M. (1995). DNA markers tightly linked to a gall midge resistance gene (*Gm2*) are potentially useful for marker-aided selection in rice breeding. *Theor. Appl. Genet.* **91**, 68–73.

Nair, S. K., Prasanna, B. M., Garg, A., Rathore, R. S., Setty, T. A. S., and Singh, N. N. (2005). Identification and validation of QTL conferring resistance to sorghum downy mildew (*Peronosclerospora sorghi*) and Rajasthan downy mildew (*P. heteropogoni*) in maize. *Theor. Appl. Genet.* **110**, 1384–1392.

Narayanan, N. N., Baisakh, N., Cruz, C. M. V., Gnanamanickam, S. S., Datta, K., and Datta, S. K. (2002). Molecular breeding for the development of blast and bacterial blight resistance in rice cv. IR50. *Crop Sci.* **42**, 2072–2079.

Narayanan, N. N., Baisakh, N., Oliva, N. P., VeraCruz, C. M., Gnanamanickam, S. S., Datta, K., and Datta, S. K. (2004). Molecular breeding: Marker-assisted selection combined with biolistic transformation for blast and bacterial blight resistance in Indica rice (cv. CO39). *Mol. Breed.* **14**, 61–71.

Nasu, S., Suzuki, J., Ohta, R., Hasegawa, K., Yui, R., Kitazawa, N., Monna, L., and Minobe, Y. (2002). Search for and analysis of single nucleotide polymorphisms (SNPs) in rice (*Oryza sativa*, *O. rufipogon*) and establishment of SNP markers. *DNA Res.* **9**, 163–171.

Nelson, O. E. (2001). Maize: The long trail to QTM. *In* "Encyclopedia of Genetics" (E. C. R. Reeve and I. Black, Eds.), pp. 657–660. Fitzroy Dearborn.

Ni, Z. F., Sun, Q. X., and Liu, Z. Y. (2000). Identification of a hybrid-specific expressed gene encoding novel RNA-binding protein in wheat seedling leaves using differential display of mRNA. *Mol. Gen. Genet.* **263**, 934–938.

Ni, Z. F., Sun, Q. X., Wu, L. M., and Xie, C. J. (2002). Differential gene expression between wheat hybrids and their parental inbreds in primary roots. *Acta Bot. Sin.* **44**, 457–462.

Nicot, N., Chiquet, V., Gandon, B., Amilhat, L., Legeai, F., Leroy, P., Bernard, M., and Sourdille, P. (2004). Study of simple sequence repeat (SSR) markers from wheat expressed sequence tags (ESTs). *Theor. Appl. Genet.* **109**, 800–805.

Niebur, W. S., Rafalski, J. A., Smith, O. S., and Cooper, M. (2004). Applications of genomics technologies to enhance rate of genetic progress for yield of maize within a commercial breeding program. *In* "New Directions for a Diverse Planet: Proceedings of the Fourth International Crop Science Congress" (T. Fischer, Ed.). Brisbane, Australia (www.cropscience.org.au).

Nielsen, K. L., Grønkjaer, K., Welinder, K. G., and Emmersen, J. (2005). Global transcript profiling of potato tuber using LongSAGE. *Plant Biotechnol. J.* **3**, 175–185.

Nishimura, A., Ashikari, M., Lin, S., Takashi, T., Angeles, E. R., Yamamoto, T., and Matsuoka, M. (2005). Isolation of a rice regeneration quantitative trait loci gene and its application to transformation. *Proc. Natl. Acad. Sci. USA* **102**, 11940–11944.

O'Donoughue, L. S., Kianian, S. F., Rayapati, P. J., Penner, G. A., Sorrells, M. E., Tanksley, S. D., Phillips, R. L., Rines, H. W., Lee, M., Fedak, G., Molnar, S. J., Hoffman, D., *et al.* (1995). A molecular linkage map of cultivated oat. *Genome* **38**, 368–380.

Ogbonnaya, F. E., Subrahmanyam, N. C., Moullet, O., De Majnik, J., Eagles, H. A., Brown, J. S., Eastwood, R. F., Kollmorgen, J., Appels, R., and Lagudah, E. S. (2001). Diagnostic DNA markers for cereal cyst nematode resistance in bread wheat. *Aust. J. Agric. Res.* **52**, 1367–1374.

Ojiambo, P. S., Nyanapah, J. O., Lung'aho, C., Karinga, J. K., and Kidanemariam, H. M. (2000). Comparing different ecological models in field evaluations of selected genotypes from *Solanum tuberosum* CIP population A for resistance to *Phytophthora infestans* (Mont.) de Bary in Kenya. *Euphytica* **111**, 211–218.

Okada, Y., Kanatani, R., Arai, S., and Ito, K. (2003a). A CAPS marker that distinguishes the barley yellow mosaic disease resistance locus *rym1* derived from Chinese landrace Mokusekko 3. *J. Inst. Brew.* **109**, 103–105.

Okada, Y., Kanatani, R., Arai, S., Asakura, T., and Ito, K. (2003b). Production of a novel virus-resistant barley line introgression to the *rym1* locus with high malting quality using DNA marker assisted selection. *J. Inst. Brew.* **109**, 99–102.

Okogbenin, E., Marin, J., and Fergene, M. (2006). An SSR-based molecular genetic map of cassava. *Euphytica* **147**, 433–440.

Olson, M., Hood, L., Cantor, C., and Botsetin, D. (1989). A common language for physical mapping of the human genome. *Science* **245**, 1434–1435.

Ordon, F., Bauer, E., Friedt, W., and Graner, A. (1995). Marker-based selection for the *ym4* BaMMV-resistance gene in barley using RAPDs. *Agronomie* **15**, 481–485.

Oriero, C. E., Odunola, O. A., Lokko, Y., and Ingelbrecht, I. (2006). Analysis of B-genome derived simple sequence repeat (SSR) markers in *Musa* spp. *Afr. J. Biotechnol.* **5**, 126–128.

Ortiz, R. (2003). An international public partnership for genetic enhancement of cowpea using a holistic approach to biotechnology. *Genomic Proteomic Technol.* **3**, 45–47.

Ortiz, R., Madsen, S., Ruiz-Tapia, E. N., Jacobsen, S. E., Mujica-Sánchez, A., Christiansen, J. L., and Stølen, O. (1999). Validating a core collection of Peruvian quinoa germplasm. *Genetic Resour. Crop Evol.* **46**, 285–290.

Ozturk, Z. N., Talame, V., Deyhoyos, M., Maichalowski, C. B., Galbriath, D. W., Gozukirmizi, N., Tuberosa, R., and Bohnert, H. J. (2002). Monitoring large-scale changes in transcript abundance in drought- and salt-stressed barley. *Plant Mol. Biol.* **48**, 551–573.

Palaisa, K. A., Morgante, M., Williams, M., and Rafalski, A. (2003). Contrasting effect of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. *Plant Cell* **15**, 1795–1806.

Palmer, L. E., Rabinowicz, P. D., O'Shaughnessy, A. L., Balija, V. S., Nascimento, L. U., Dike, S., de la Bastide, M., Martienssen, R. A., and McCombie, W. R. (2003). Maize genome sequencing by methylation filtration. *Science* **302**, 2115–2117.

Paltridge, N. G., Collins, N. C., Bendahmane, A., and Symons, R. H. (1998). Development of YLM, a co-dominant PCR marker closely linked to the *Yd2* gene for resistance to barley yellow dwarf disease. *Theor. Appl. Genet.* **96**, 1170–1177.

Paran, I., and Michelmore, R. W. (1993). Development of reliable PCR-based markers linked to downy mildew resistance gene in lettuce. *Theor. Appl. Genet.* **85**, 985–993.

Paris, M., Potter, R. H., Lance, R. C. M., Li, C. D., and Jones, M. G. K. (2003). Typing *Mlo* alleles for powdery mildew resistance in barley by single nucleotide polymorphism analysis using MALDI-ToF mass spectrometry. *Aust. J. Argic. Res.* **54**, 1343–1349.

Paterson, A. H., Lander, E. S., Hewitt, J. D., Paterson, S., Lincoln, S. E., and Tanksley, S. D. (1988). Resolution of quantitative traits into Mendelian factors, using a complete linkage map of restriction fragment length polymorphism. *Nature* **335**, 721–726.

Paterson, A. H., Damon, S., Hewitt, J. D., Zamir, D., Rabinowitch, H. D., Lincoln, S. E., Lander, E. S., and Tanksley, S. D. (1991). Mendelian factors underlying quantitative trait in tomato: Comparison across species, generations, and environments. *Genetics* **127**, 181–197.

Payne, P. I. (1987). Genetics of wheat storage proteins and their effect of allelic variation on bread-making quality. *Annu. Rev. Plant Physiol.* **38**, 141–153.

Peleman, J. D., Wye, C., Zethof, J., Sorensen, A. P., Verbakel, H., van Oeveren, J., Gerats, T., and van der Voort, J. R. (2005). Quantitative trait locus (QTL) isogenic recombinant analysis: A method for high-resolution mapping of QTL within a single population. *Genetics* **171**, 1341–1352.

Peng, J., Richard, D. E., Hartley, N. M., Murphy, G. P., Devos, K. M., Flintham, J. E., Beales, J., Fish, L. J., Worland, A. J., Pelica, F., Sudhakar, D., Christou, P., *et al.* (1999). 'Green revolution' genes encode mutant gibberellin response modulators. *Nature* **400**, 256–261.

Perry, J. A., Wang, T. L., Welham, T. J., Gardner, S., Pike, J. M., Yoshida, S., and Parniske, M. (2003). A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume *Lotus japonicus*. *Plant Physiol.* **131**, 866–871.

Pestsova, E. G., Borner, A., and Röder, M. S. (2001). Development of a set of *Triticum aestivum—Aegilops tauschii* introgression lines. *Hereditas* **135**, 139–143.

Pestsova, E. G., Borner, A., and Röder, M. S. (2006). Development and QTL assessment of *Triticum aestivum-Aegilops tauschii* introgression lines. *Theor. Appl. Genet.* **112**, 634–647.

Peterson, D. G., Schulze, S. R., Sciara, E. B., Lee, S. A., Bowers, J. E., Nagel, A., Jiang, N., Tibbitts, D. C., Wessler, S. R., and Paterson, A. H. (2002). Integration of Co*t* analysis, DNA cloning, and high throughput sequencing facilitate genome characterization and gene discovery. *Genome Res.* **12**, 795–807.

Picoult-Newberg, L., Ideker, T. E., Pohl, M. G., Taylor, S. L., Donaldson, M. A., Nickerson, D. A., and Boyce-Jacino, M. (1999). Mining SNPs from EST databases. *Genome Res.* **9**, 167–174.

Pinson, S. R. M., Capdevielle, F. M., and Ord, J. H. (2005). Confirming QTL and finding additional loci conditioning sheat blight resistance in rice using recombinant inbred lines. *Crop Sci.* **45**, 503–510.

Pinstrup-Anderson, P., Pandya-Lorch, R., and Rosegrant, M. W. (1999). World food prospects: Critical issues for the early twenty-first century. 2020 Vision. Food policy report. International Food Policy Research Institute, Washington, DC, October 1999, 32pp.

Podlich, D. W., and Cooper, M. (1998). QU-GENE: A simulation platform for quantitative analysis of genetic models. *Bioinformatics* **14**, 632–653.

Podlich, D. W., Winkler, C. R., and Cooper, M. (2004). Mapping As You Go (MAYG): An effective approach for marker-assisted selection of complex traits. *Crop Sci.* **44**, 1560–1571.

Portyanko, V. A., Hoffman, D. L., Lee, M., and Holland, J. B. (2001). A linkage map of hexaploid oat based on grass anchor DNA clones and its relationship to other oat maps. *Genome* **44**, 249–265.

Potokina, E., Caspers, M., Prasad, M., Kota, R., Zhang, H., Sreenivasulu, N., Wang, M., and Graner, A. (2004). Functional association between malting quality trait components and cDNA array based on expression patterns in barley (*Hordeum vulgare* L.). *Mol. Breed.* **14**, 153–170.

Powell, W., Machray, G. C., and Provan, J. (1996). Polymorphism revealed by simple sequence repeats. *Trends Plant Sci.* **7**, 215–222.

Prasad, M., Kumar, N., Kulwal, P. L., Roder, M. S., Balyan, H. S., Dhaliwal, H. S., and Gupta, P. K. (2003). QTL analysis for grain protein content using SSR markers and validation studies using NILs in bread wheat. *Theor. Appl. Genet.* **106**, 659–667.

Prasanna, B. M., Vasal, S. K., Kassahun, B., and Singh, N. N. (2001). Quality protein maize. *Curr. Sci.* **81**, 1308–1319.

Price, A. H. (2006). Believe it or not, QTLs are accurate. *Trends Plant Sci.* **11**, 213–216.

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000a). Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.

Pritchard, J. K., Stephens, M., Rosenberg, N. A., and Donnelly, P. (2000b). Association mapping in structured populations. *Am. J. Human Genet.* **67**, 170–181.

Qi, X., Lindup, S., Pittaway, T. S., Allouis, S., Gale, M. D., and Devos, K. M. (2001). Development of simple sequence repeat markers from bacterial artificial chromosomes without subcloning. *BioTechniques* **31**, 355–362.

Qi, X., Pittaway, T. S., Lindup, S., Liu, H., Waterman, E., Padi, F. K., Hash, C. T., Zhu, J., Gale, M. D., and Devos, K. M. (2004). An integrated genetic map and a new set of simple sequence repeat markers for pearlmillet, *Pennisetum glaucum*. *Theor. Appl. Genet.* **109**, 1485–1493.

Quarrie, S. A., Steed, A., Calestani, C., Semikhodskii, A., Lebreton, C., Chinoy, C., Steele, N., Pljevljakusić, D., Waterman, E., Weyen, J., Schondelmaier, J., Habash, D. Z., *et al.* (2005).

A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese Spring x SQ1 and its use to compare QTLs for grain yield across a range of environments. *Theor. Appl. Genet.* **110**, 865–880.

Rabbani, M. A., Maruyama, K., Abe, H., Khan, M. Y., Katsura, K., Ito, Y., Yoshiwara, K., Seki, M., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2003). Monitoring expression profiles of rice gene under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA gel-blot analyses. *Plant Physiol.* **133**, 1755–1767.

Rabinowitz, D. (1997). A transmission disequilibrium test for quantitative trait loci. *Hum. Hered.* **47**, 342–350.

Rabinowicz, P. D., Schulz, K., Dedhia, N., Yordan, C., Parnemm, L. D., Parnell, L. D., Stein, L., McCombie, R., and Martienssen, R. A. (1999). Differential methylation of genes and retrotransposons facilitates shot gun sequencing of maize genome. *Nat. Genet.* **23**, 305–308.

Radovanovic, N., and Cloutier, S. (2003). Gene-assisted selection for high molecular weight glutenin subunits in wheat double haploid breeding programs. *Mol. Breed.* **12**, 51–59.

Rafalski, A., and Morgante, M. (2004). Corns and humans: Recombination and linkage disequilibrium in two genomes of similar size. *Trends Genet.* **20**, 103–111.

Ragagnin, V. A., Sanglard, D. A., de Souza, T. L. P. O., Moreira, M. A., and de Barros, E. G. (2003). Simultaneous transfer of resistance genes for rust, anthracnose, and angular leaf spot to cultivar Perola assisted by molecular markers. *Annu. Rep. Bean Improv. Coop.* **46**, 159–160.

Ragot, M., Biasiolli, M., Delbut, M. F., Dell'Orco, A., Malgarini, L., Thevenin, P., Vernoy, J., Vivant, J., Zimmermann, R., and Gay, G. (1995). Marker-assisted backcrossing: A practical example. *In* "Techniques et utilisations des marqueurs moléculaires" (A. Bervillé and M. Tersac, Eds.), pp. 45–56. Les Colloques, No. 72, INRA, Paris, France.

Ragot, M., Gay, G., Muller, J.-P., and Durovray, J. (2000). Efficient selection for the adaptation to the environment through QTL mapping and manipulation in maize. *In* "Molecular Approaches for the Genetic Improvement of Cereals for Stable Production in Water-Limited Environments" (J.-M. Ribaut and D. Poland, Eds.), pp. 128–130. CIMMYT, México, DF.

Raman, H., and Read, B. J. (2000). Molecular breeding for resistance against Russian wheat aphid in Australian barley. *J. Agric. Genomics* **5**, 1–5.

Raman, H., Karakousis, A., Moroni, S., Raman, R., and Read, B. (2001). Development, validation and allele diversity of microsatellite markers closely linked with the *Alp* locus with aluminium tolerance in barley. (http://www.regional.org.au/au/abts/2001/t4/raman2.htm)

Raman, H., Platz, G. J., Chalmer, K. J., Raman, R., Read, B. J., Barr, A. R., and Moody, D. B. (2003). Mapping of genomic regions associated with net form of net blotch resistance in barley. *Aust. J. Agric. Res.* **54**, 1359–1367.

Ramalingam, J., Basharat, H. S., and Zhang, G. (2002). STS and microsatellite marker-assisted selection for bacterial blight resistance and waxy gene in rice, *Oryza sativa* L. *Euphytica* **127**, 255–260.

Ramsay, L., Macaulay, M., degli Ivanissevich, S., MacLean, K., Cardle, L., Fuller, J., Edwards, K. J., Tuvesson, S., Morgante, M., Massari, A., Maestri, E., Marmiroli, N., *et al.* (2000). A simple sequence repeat-based linkage map of barley. *Genetics* **156**, 1997–2005.

Rao, K. E. P., and Rao, V. R. (1995). The use of characterization data in developing a core collection of sorghum. *In* "Core Collections of Plant Genetic Resources" (T. Hodgkin, A. H. D. Brown, Th. J. L. van Hintum, and E. A. V. Morales, Eds.), pp. 109–115. John Wiley & Sons, Chichester, West Sussex, United Kingdom.

Ravel, C., Praud, S., Murigneux, A., Linossier, L., Dardevet, M., Balfourier, F., Dufour, P., Brumel, D., and Charmet, G. (2006). Identification of *Glu-B1-1* as a candidate gene for the

quantity of high-molecular-weight glutenin in bread wheat (*Triticum aestivum* L.) by means of an association study. *Theor. Appl. Genet.* **112**, 738–743.

Redoña, E. D., and Mackill, D. J. (1996). Mapping quantitative trait loci for seedling vigor in rice using RFLPs. *Theor. Appl. Genet.* **92**, 395–402.

Reinke, R. F., Welsh, L. A., Reece, J. E., Lewin, L. G., and Blakenet, A. B. (1991). Procedures for quality selection in aromatic rice varieties. *Int. Rice Res. Newsl.* **16**, 10–11.

Ren, Z. H., Gao, J. P., Li, L. G., Cai, X. L., Huang, W., Chao, D. Y., Zhu, M. Z., Wang, Z. Y., Luan, S., and Lin, H. X. (2005). A rice quantitative trait locus for salt tolerance encodes a sodium transporter. *Nat. Genet.* **37**, 1141–1146.

Rensink, W., Hart, A., Liu, J., Ouyang, S., Zismann, V., and Buell, C. R. (2005). Analyzing the potato abiotic stress transcriptome using expressed sequence tags. *Genome* **48**, 598–605.

Reymond, M., Muller, B., Leonardi, A., Charcosset, A., and Tardieu, F. (2003). Combining quantitative trait loci analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and water deficit. *Plant Physiol.* **131**, 664–675.

Ribaut, J. M., and Betrán, J. (2000). Single large-scale marker-assisted selection (SLS-MAS). *Mol. Breed.* **5**, 531–541.

Ribaut, J. M., and Hoisington, D. (1998). Marker-assisted selection: New tools and strategies. *Trends Plant Sci.* **3**, 236–239.

Ribaut, J. M., Hoisington, D. A., Deutsch, J. A., Jiang, C., and González-de-León, D. (1996). Identification of quantitative trait loci under drought conditions in tropical maize. I. Flowering parameters and the anthesis-silking interval. *Theor. Appl. Genet.* **92**, 905–914.

Ribaut, J.-M., Jiang, C., González-de-León, D., Edmeades, G. O., and Hoisington, D. A. (1997a). Identification of quantitative trait loci under drought conditions in tropical maize. II. Yield components and marker-assisted selection strategies. *Theor. Appl. Genet.* **94**, 887–896.

Ribaut, J. M., Huu, X., Hoisington, D., and Gonzales de Leon, D. (1997b). Use of STSs and SSRs as rapid and reliable preselection tools in marker-assisted selection backcross scheme. *Plant Mol. Biol. Rep.* **15**, 156–164.

Ribaut, J. M., Jiang, C., and Hoisington, D. (2002). Simulation experiments on efficiencies of gene introgression by backcrossing. *Crop Sci.* **42**, 557–565.

Richardson, K. L., Vales, M. I., Kling, J. G., Mundt, C. C., and Hayes, P. M. (2006). Pyramiding and dissecting disease resistance QTL to barley stripe rust. *Theor. Appl. Genet.* **113**, 485–495.

Rick, C. M. (1974). High soluble-solids content in large-fruited tomato lines derived from a wild green-fruited species. *Hilgardia* **42**, 493–510.

Ritter, E., Debener, T., Barone, A., Salamini, F., and Gebhardt, C. (1991). RFLP mapping on potato chromosomes of two genes controlling extreme resistance to potato virus X (PVX). *Mol. Gen. Genet.* **227**, 81–85.

Rockman, M. V., and Kruglyak, L. (2006). Genetics of global gene expression. *Nat. Rev. Genet.* **7**, 862–872.

Röder, M. S., Korzum, V., Wendehake, K., Plaschke, J., Tixier, M.-H., Leroy, P., and Ganal, M. W. (1998). A microsatellite map of wheat. *Genetics* **149**, 2007–2023.

Romagosa, I., Han, F., Ullrich, S. E., Hays, P. M., and Wesenberg, D. M. (1999). Verification of yield QTL through realized molecular marker-assisted selection responses in a barley cross. *Mol. Breed.* **5**, 143–152.

Ronning, C. M., Stegalkina, S. S., Ascenzi, R. A., Bougri, O., Hart, A. L., Utterbach, T. R., Vanaken, S. E., Riedmuller, S. B., White, J. A., Cho, J., Pertea, G. M., Lee, Y., *et al.* (2003). Comparative analysis of potato expressed sequence tag libraries. *Plant Physiol.* **131**, 419–429.

Rosegrant, M. W., Agcaoili-Sombilla, M., and Perez, N. D. (1995). Global food projection to 2020: Implications for investment. Food, Agriculture, and the Environment Discussion Paper 5. International Food Policy Research Institute, Washington DC, October 1995, 54pp.

Rostoks, N., Mudie, S., Cardle, L., Russell, J., Ramsay, L., Booth, A., Svensson, J. T., Wanamaker, S. I., Walia, H., Rodriguez, E. M., Hedley, P. E., Liu, H., *et al.* (2005). Genome-wide SNP discovery and linkage analysis in barley based on genes responsive to abiotic stress. *Mol. Gen. Genom.* **274**, 515–527.

Roy, J. K., Bandopadhyay, R., Rustgi, S., Balyan, H. S., and Gupta, P. K. (2006). Association analysis of agronomically important traits using SSR, SAMPL and AFLP markers in bread wheat. *Curr. Sci.* **90**, 683–689.

Rudd, S. (2003). Expressed sequence tags: Alternative or complement to whole genome sequencing. *Trends Plant Sci.* **8**, 321–329.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S., Higuchi, R. H., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487–491.

Saito, K., Miura, K., Nagano, K., Hayano-Saito, Y., Araki, H., and Kato, A. (2001). Identification of two closely linked quantitative trait loci for cold tolerance on chromosome 4 of rice and their association with anther length. *Theor. Appl. Genet.* **103**, 862–868.

Salvi, S., Tuberosa, R., Chiapparino, E., Maccaferri, M., Veillet, S., van Beuningen, L., Isaac, P., Edwards, K., and Phillips, R. L. (2002). Toward positional cloning of *Vgt1*, a QTL controlling the transition from the vegetative to the reproductive phase in maize. *Plant Mol. Biol.* **48**, 601–613.

Sanchez, A. C., Brar, D. S., Huang, N., Li, Z., and Khush, G. S. (2000). Sequence tagged site marker- assisted selection for three bacterial blight resistance genes in rice. *Crop Sci.* **40**, 792–797.

Santos, C. M. R., Martins, N. F., Hörberg, H. M., de Almeida, E. R. P., Coelho, M. C. F., Togawa, R. C., da Silva, F. R., Caetano, A. R., Miller, R. N. G., and Souza, M. T., Jr. (2005). Analysis of expressed sequence tags from *Musa acuminata* ssp. *burmannicoides*, var. Calcutta 4 (AA) leaves submitted to temperature stresses. *Theor. Appl. Genet.* **110**, 1517–1522.

Sawkins, M. C., Farmer, A. D., Hoisington, D., Sullivan, J., Tolopko, A., Jiang, Z., and Ribaut, J.-M. (2004). Comparative map and trait viewer (CMTV): An integrated bioinformatic tool to construct consensus maps and compare QTL and functional genomics data across genomes and experiments. *Plant Mol. Biol.* **56**, 465–480.

Scarcelli, N., Dainou, O., Agbangla, C., Tostain, S., and Pham, J.-L. (2005). Segregation patterns of isozyme loci and microsatellite markers show the diploidy of African yam *Dioscorea rotundata* (2n = 40). *Theor. Appl. Genet.* **111**, 226–232.

Schachermayr, G., Sielder, H., Gale, M. D., Winzeler, H., Winzeler, M., and Keller, B. (1994). Identification and localization of molecular markers linked to the *Lr9* leaf rust resistance gene of wheat. *Theor. Appl. Genet.* **88**, 110–115.

Schachermayr, G., Feuillet, C., and Keller, B. (1997). Molecular markers for the detection of the wheat leaf rust resistance gene *Lr10* in diverse genetic backgrounds. *Mol. Breed.* **3**, 65–74.

Schadt, E. E., Monks, S. A., Drake, T. A., Lusis, A. J., Che, N., Colinayo, V., Ruff, T. G., Milligan, S. B., Lamb, J. R., Cavet, G., Linsley, P. S., Mao, M., *et al.* (2003). Genetics of gene expression surveyed in maize, mouse and man. *Nature* **422**, 297–302.

Scheuring, C., Barthelson, R., Gailbraith, D., Betran, J., Cothren, J. T., Zeng, Z.-B., and Zhang, H.-B (2006). Preliminary analysis of differential gene expression between a maize superior hybrid and its parents using the 57K maize gene-specific long-oligonucleotide microarray. *In* "48th Annual Maize Genetic Conference," 9–12 March 2006. California. 132pp.

Schmierer, D. A., Kandemir, N., Kudrna, D. A., Jones, B. L., Ullrich, S. E., and Kleinhofs, A. (2004). Molecular marker-assisted selection for enhanced yield in malting barley. *Mol. Breed.* **14**, 463–473.

Schneider, K. A., Brothers, M. E., and Kelly, J. D. (1997). Marker-assisted selection to improve drought resistance in common bean. *Crop Sci.* **37**, 51–60.

Schnurbusch, T., Bossolini, E., Messmer, M., and Keller, B. (2004). Tagging and validation of a major quantitative trait locus for leaf rust resistance and leaf tip necrosis in winter wheat cultivar Forno. *Phytopathology* **94**, 1036–1041.

Schön, C. C., Utz, H. F., Groh, S., Truberg, B., Openshaw, S., and Melchinger, A. E. (2004). Quantitative trait locus mapping based on resampling in a vast maize testcross experiment and its relevance to quantitative genetics for complex traits. *Genetics* **167**, 485–498.

Schwarz, G., Felsenstein, F. G., and Wenzel, G. (2004). Development and validation of a PCR-based marker assay for negative selection of the HMW glutenin allele *Glu-B1-1d* (*Bx-6*) in wheat. *Theor. Appl. Genet.* **109**, 1064–1069.

Seaton, G., Haley, C. S., Knott, S. A., Kearsey, M., and Visscher, P. M. (2002). QTL Express: Mapping quantitative trait loci in simple and complex pedigrees. *Bioinformatics* **18**, 339–340.

Semagn, K., Bjørnstad, Å., Skinnes, H., Marøy, A. G., Tarkegne, Y., and William, M. (2006). Distribution of DArT, AFLP, and SSR markers in a genetic linkage map of a doubled-haploid hexaploid wheat population. *Genome* **49**, 545–555.

Semel, Y., Nissembaum, J., Menda, N., Zinder, M., Krieger, U., Issman, N., Pleban, T., Lippman, Z., Gur, A., and Zamir, D. (2006). Overdominant quantitative trait loci for yield and fitness in tomato. *Proc. Natl. Acad. Sci. USA* **103**, 12981–12986.

Senior, M. L., and Heun, M. (1993). Mapping maize microsatellites and polymerase chain reaction confirmation of the targeted repeats using a CT primer. *Genome* **36**, 884–889.

Sethy, N. K., Shokeen, B., Edwards, K. J., and Bhatia, S. (2006). Development of microsatellite markers and analysis of interspecific genetic variability in chickpea (*Cicer arietinum* L.). *Theor. Appl. Genet.* **112**, 1416–1428.

Sharma, P. N., Torii, A., Takumi, S., Mori, N., and Nakamura, C. (2004). Marker-assisted pyramiding of brown planthopper (*Nilaparvata lugens* Stal) resistance genes *Bph1* and *Bph2* on rice chromosome 12. *Hereditas* **140**, 61–69.

Sharopova, N., McMullen, M. D., Schultz, L., Schroeder, S., Sanchez-Villeda, H., Gardiner, J., Bergstrom, D., Houchins, K., Melia-Hancock, S., Musket, T., Duru, N., Polacco, M., *et al.* (2002). Development and mapping of SSR markers for maize. *Plant Mol. Biol.* **48**, 463–481.

Sharp, P. J., Johnston, S., Brown, G., McIntosh, R. A., Pallotta, M., Carter, M., Bariana, H. S., Khatkar, S., Lagudah, E.S, Singh, R. P., Khairallah, M., Potter, R., *et al.* (2001). Validation of molecular markers for wheat breeding. *Aust. J. Agric. Res.* **52**, 1357–1366.

Shen, L., Courtois, B., McNally, K. L., Robin, S., and Li, Z. (2001). Evaluation of near-isogenic lines of rice introgressed with QTLs for root depth through marker-aided selection. *Theor. Appl. Genet.* **103**, 75–83.

Shubing, L., Zhou, R., Dong, Y., Li, P., and Jia, J. (2006). Development, utilization of introgression lines using a synthetic wheat as donor. *Theor. Appl. Genet.* **112**, 1360–1373.

Siangliw, M., Toojinda, T., Tragoonrung, S., and Vanavichit, A. (2003). Thai Jasmine rice carrying QTLch9 (*Sub*QTL) is submergence tolerant. *Ann. Bot.* **91**, 255–261.

Singh, H., Prasad, M., Varshney, R. K., Roy, J. K., Balyan, H. S., Dhaliwal, H. S., and Gupta, P. K. (2001b). STMS marker for grain protein content and their validation using near-isogenic lines in bread wheat. *Plant Breed.* **120**, 273–278.

Singh, R., Datta, D., Priyamvad, Singh, S., and Tiwari, R. (2004). Marker-assisted selection for leaf rust resistance genes *Lr19* and *Lr24* in wheat (*Triticum aestivum* L.). *J. Appl. Genet.* **45**, 399–403.

Singh, R. J., Kollipara, K. K., and Hymowtiz, T. (1998). Monosomic alien addition lines derived from *Glycine max* (L.) Merr. and *G. tomentella* Hayata: Production, characterization, and breeding behavior. *Crop Sci.* **38**, 1483–1489.

Singh, S., Sindhu, J. S., Huang, N., Vikal, Y., Li, Z., Brar, D. S., Dhaliwal, H. S., and Khush, G. S. (2001a). Pyramiding three bacterial blight resistance genes (*xa5*, *xa13*, and *Xa21*) using marker-assisted selection into *indica* rice cultivar PR106. *Theor. Appl. Genet.* **102**, 1011–1015.

Slade, A. J., Fuerstenberg, S. I., Loeffler, D., Steine, M. N., and Facciotti, D. (2005). A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. *Nat. Biotechnol.* **23**, 75–81.

Slikova, S., Gregova, E., Bartos, P., and Kraic, J. (2003). Marker-assisted selection for leaf rust resistance in wheat by transfer of gene *Lr19*. *Plant Protect. Sci.* **39**, 13–17.

Smits, B. M., Mudde, J., Cuppen, E., and Plasterk, R. H. (2004). Target-selected mutagenesis of the rat. *Genomics* **83**, 332–334.

Sobrizal, K., Ikeda, K., Sanchez, P. L., Doi, K., Angeles, E. R., Khush, G. S., and Yoshimura, A. (1999). Development of *Oryza glumaepatulla* introgression lines in rice, *O. sativa* L. *Rice Genet. Newsl.* **16**, 107.

Somers, D. J., Kirckpatrick, R., Moniwa, M., and Walsh, A. (2003). Mining single-nucleotide polymorphisms from hexaploid ESTs. *Genome* **49**, 431–437.

Somers, D. J., Thomas, J., DePauw, R., Fox, S., Humphreys, G., and Fedak, G. (2005). Assembling complex genotypes to resist *Fusarium* in wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* **111**, 1623–1631.

Song, J., Bradeen, J. M., Naess, S. K., Raasch, J. A., Wielgus, S. M., Haberlach, G. T., Liu, J., Austin-Phillips, S., Buell, C. R., Helgeson, J. P., and Jiang, J. (2003). Gene *RB* cloned from *Solanum bulbocastanum* confers broad spectrum resistance to potato late blight. *Proc. Natl. Acad. Sci. USA* **100**, 9128–9133.

Song, Q. J., Shi, J. R., Singh, S., Fickus, E. W., Costa, J. M., Lewis, J., Gill, B. S., Ward, R., and Cregan, P. B. (2005). Development and mapping of microsatellite (SSR) markers in wheat. *Theor. Appl. Genet.* **110**, 550–560.

Song, R., and Messing, J. (2003). Gene expression of a gene family in maize based on noncollinear haplotypes. *Proc. Natl. Acad. Sci. USA* **100**, 9055–9060.

Spaner, D., Rossnagel, B. G., Legge, W. G., Scoles, G. J., Eckstein, P. E., Penner, G. A., Tinker, N. A., Briggs, K. G., Falk, D. E., Afele, J. C., Hayes, P. M., and Mather, D. E. (1999). Verification of a quantitative trait locus affecting agronomic traits in two-row barley. *Crop Sci.* **39**, 248–252.

Spielman, R. S., McGinnis, R. E., and Ewens, W. J. (1993). Transmission test for linkage disequilibrium: The insulin gene region and insulin-dependent *diabetes mellitus* (IDDM). *Am. J. Hum. Genet.* **52**, 506–516.

Spielmeyer, W., Sharp, P. J., and Lagudah, E. S. (2003). Identification and validation of markers linked to broad-spectrum stem rust resistance gene by *Sr2* in wheat (*Triticum aestivum* L.). *Crop Sci.* **43**, 333–336.

Stam, P. (2003). Marker-assisted introgression: Speed at any cost? *In "EUCARPIA Leafy Vegetables 2003"* (Th. J. L. van Hintum, A. Lebeda, D. Pink, and J. W. Schut, Eds.), pp. 117–124. EUCARPIA, Valencia, Spain.

Stavely, J. R., and McMillan, R. T., Jr. (1992). BARC-rust resistant bush, fresh-market green bean germplasm lines. *HortScience* **27**, 1052–1054.

Stavely, J. R., and Steinke, J. (1990). BARC-rust resistant bush waxy bean germplasm lines. *HortScience* **25**, 1451–1452.

Stavely, J. R., Kelly, J. D., and Grafton, K. F. (1994). BelMiDak rust resistant navy dry beans germplasm lines. *HortScience* **29**, 709–710.

Stavely, J. R., Kelly, J. D., Grafton, K. F., Mullins, C. A., Straw, A., McMillan, R. T., Jr., Beaver, J. S., Miklas, P. N., Steinke, J., Steadman, J. R., Coyne, D. P., Lindgren, D. T., *et al.* (1997). Rust resistant bean germplasm releases, 1994–96. *Annu. Rep. Bean Improv. Coop.* **40**, 120–121.

Stavely, J. R., McMillan, R. T., Beaver, J. S., and Miklas, P. N. (2001). Release of three McCaslan type, indeterminate, rust and golden mosaic resistant snap bean germplasm lines BelDade RGMR 4, 5, and 6. *Annu. Rep. Bean Improv. Coop.* **44**, 197–198.

Steele, K. A., Price, A. H., Shashidhar, H. E., and Witcombe, J. R. (2006). Marker-assisted selection to introgress rice QTL controlling root traits into an Indian upland rice variety. *Theor. Appl. Genet.* **112**, 208–221.

Stemple, D. L. (2004). TILLING—a high-throughput harvest for functional genomics. *Nat. Rev. Genet.* **5**, 145–150.

Stepien, L., Chelkowski, J., Wenzel, G., and Mohler, V. (2004). Combined use of linked markers for genotyping the *Pm1* locus in common wheat. *Cell. Mol. Biol. Lett.* **9**, 819–827.

Stich, B., Maurer, H. P., Melchinger, A. E., Frisch, M., Heckenberger, M., van der Voort, J. R., Peleman, J., Sørensen, A. P., and Reif, J. C. (2006). Comparison of linkage disequilibrium in elite European maize inbred lines using AFLP and SSR markers. *Mol. Breed.* **17**, 217–226.

Stratton, D. A. (1998). Reaction norm functions and QTL-environments for flowering time in *Arabidopsis thaliana*. *Heredity* **81**, 144–155.

Stromberg, L. D., Dudley, J. W., and Rufener, G. K. (1994). Comparing conventional early generation selection with molecular marker assisted selection in maize. *Crop Sci.* **34**, 1221–1225.

Stuber, C. W., Lincoln, S. E., Wolff, D. W., Helentjaris, T., and Lander, E. S. (1992). Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. *Genetics* **132**, 823–839.

Stuber, C. W., Polacco, M., and Senior, M. L. (1999). Strategy of empirical breeding, marker-assisted selection, and genomics to increase crop yield potential. *Crop Sci.* **39**, 1571–1583.

Stuper, R., and Springer, N. (2006). Analysis of allelic variation in gene expression in B73 and Mo17 and implications for hybrid expression patterns. *In* "48th Annual Maize Genetic Conference," 9–12 March 2006. California. p. 115.

Sun, D. J., He, Z. H., Xia, X. C., Zhang, L. P., Morris, C., Appels, R., Ma, W., and Wang, H. (2005). A novel STS marker for polyphenol oxidase activities in bread wheat. *Mol. Breed.* **16**, 209–218.

Sun, Q. X., Ni, Z. F., and Liu, Z. Y. (1999). Differential gene expression between wheat hybrids and their parental inbreds in seedling leaves. *Euphytica* **106**, 117–123.

Sun, Q. X., Wu, L. M., Ni, Z. F., Meng, F. R., Wang, Z. K., and Lin, Z. (2004). Differential gene expression patterns in leaves between hybrids and their parental inbreds are correlated with heterosis in a wheat diallel cross. *Plant Sci.* **166**, 651–657.

Swaminathan, M. S. (2006). An evergreen revolution. *Crop Sci.* **46**, 2293–2303.

Swamy, P., Panchbhai, A. N., Dodiya, P., Naik, V., Panchbhai, S. D., Zehr, U. B., Azhakanandam, K., and Bharat, R. (2004). Evaluation of bacterial blight resistance in rice lines carrying multiple resistance genes and *Xa21* transgenic lines. *In* "Proceedings of the Fourth International Crop Science Congress." Australia (http://www.regional.org.au/au/cs/2004/poster/3/7/4/251_charbr.htm).

Swanson-Wagner, R. A., Jia, Y., DeCook, R., Borsuk, L. A., Nettleton, D., and Schnable, P. S. (2006). All possible modes of gene action are observed in a global camparison of gene expression in a maize F_1 hybrid and its inbred parents. *Proc. Natl. Acad. Sci. USA* **103**, 6805–6810.

Swennen, R., and Vuylsteke, D. (1987). Morphological taxonomy of plantains (*Musa* cultivars AAB) in West Africa. *In* "Banana and Plantain Breeding Strategies" (G. J. Persley and E. A. de Langhe, Eds.), pp. 165–171. Proc. Workshop held at Cairns, Australia. ACIAR Proc. No. 26, pp. 13–17.

Syed, N. H., and Chen, Z. J. (2005). Molecular marker genotypes, heterozygosity and genetic interactions explain heterosis in *Arabidopsis thaliana*. *Heredity* **94**, 295–304.

Szalma, S. J., Buckler, E. S., IV, Snook, M. E., and McMullen, M. D. (2005). Association analysis of candidate genes for maysin and chlorogenic acid accumulation in maize silks. *Theor. Appl. Genet.* **110**, 1324–1333.

Taba, S., Diaz, J., Franco, J., and Cross, J. (1998). Evalution of Caribbean maize accessions to develop a core subset. *Crop Sci.* **38**, 1378–1386.

Takahashi, Y., Shomura, A., Sasaki, T., and Yano, M. (2001). *Hd6*, a rice quantitative trait locus involved in photoperiod sensitivity, encodes the α subunit of protein kinase CK2. *Proc. Natl. Acad. Sci. USA* **98**, 7922–7927.

Tan, G., Jin, H., Li, G., He, R., Zhu, L., and He, G. (2005). Production and characterization of a complete set of individual chromosome additions from *Oryza sativa* using RFLP and GISH analysis. *Theor. Appl. Genet.* **111**, 1585–1595.

Tang, J., Xia, H., Li, D., Cao, M., Tao, Y., Tong, W., Zhang, X., Hu, S., Wang, J., Yu, J., Yang, H., and Zhu, L. (2005). Gene expression profiling in rice young panicle and vegetative organs and identification of panicle-specific genes through known gene functions. *Mol. Gen. Genom.* **274**, 467–476.

Tang, J., Gao, L., Cao, Y., and Jia, J. (2006). Homologous analysis of SSR-ESTs and transferability of wheat SSR-EST markers across barley, rice and maize. *Euphytica* **151**, 87–93.

Tan, Y. F., Li, J. X., Yu, S. B., Xing, Y. Z., Xu, C. G., and Zhang, Q. (1999). The three important traits for cooking and eating quality of rice grains are controlled by a single locus in an elite rice hybrid, Shanyou 63. *Theor. Appl. Genet.* **99**, 642–648.

Tan, Y. F., Xing, Y. Z., Li, J. X., Yu, S. B., Xu, C. G., and Zhang, Q. (2000). Genetic basis of the appearance quality of rice grains in Shanyou 63, an elite rice hybrid. *Theor. Appl. Genet.* **101**, 823–829.

Tanhuanpää, P., Kalendar, R., Laurila, J., Schulman, A. H., Manninen, O., and Kiviharju, E. (2006). Generation of SNP markers for short straw in oat (*Avena sativa* L.). *Genome* **49**, 282–287.

Tanksley, S. D. (1993). Mapping polygenes. *Annu. Rev. Genet.* **27**, 205–233.

Tanksley, S. D., and Nelson, J. C. (1996). Advanced backcross QTL analysis: A method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theor. Appl. Genet.* **92**, 191–203.

Tanksley, S. D., Young, N. D., Paterson, A. H., and Bonierbale, M. W. (1989). RFLP mapping in plant breeding: New tools for an old science. *Biotechnology* **7**, 257–264.

Tanksley, S. D., Ganal, M. W., Prince, J. P., de Vincente, M. C., Bonierbale, M. W., Broun, P., Fulton, T. M., Giovannoni, J. J., Grandillo, S., Martin, G. B., Messeguer, R., Miller, J. C., *et al.* (1992). High density molecular linkage maps of the tomato and potato genomes. *Genetics* **132**, 1141–1160.

Tanksley, S. D., Ganal, M. W., and Martin, G. B. (1995). Chromosome landing: A paradigm for map based gene cloning in plants with large genomes. *Trends Genet.* **11**, 63–68.

Taramino, G., Tarchini, R., Ferrario, S., Lee, M., and Pe, M. E. (1997). Characterization and mapping of simple sequence repeats (SSRs) in *Sorghum bicolor*. *Theor. Appl. Genet.* **95**, 66–72.

Tar'an, B., Michaels, T. E., and Pauls, K. P. (1998). Stability of association of molecular markers with common bacterial blight resistance in common bean (*Phaseolus vulgaris* L.). *Plant Breed.* **117**, 553–558.

Tar'an, B., Michaels, T. E., and Pauls, K. P. (2003). Marker-assisted selection for complex trait in common bean (*Phaseolus vulgaris* L.) using QTL-based index. *Euphytica* **130**, 423–432.

Tardieu, F. (2003). Virtual plants: Modeling as a tool for the genomics of tolerance to water deficit. *Trends Plant Sci.* **8**, 9–14.

Tekeoglu, M., Rajesh, P. N., and Muehlbauer, F. J. (2002). Integration of sequence tagged microsatellites to the chickpea genetic map. *Theor. Appl. Genet.* **105**, 847–854.

Thiel, T., Michalek, W., Varshney, R. K., and Graner, A. (2003). Exploiting EST data bases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* **106**, 411–422.

Thomas, W. T. B. (2003). Prospects for molecular breeding of barley. *Ann. Appl. Genet.* **142**, 1–12.

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

Tian, F., Zhu, Z., Zhang, B., Tan, L., Fu, Y., Wang, X., and Sun, C. Q. (2006a). Fine mapping of a quantitative trait locus for grain number per panicle from wild rice (*Oryza rufipogon* Griff.). *Theor. Appl. Genet.* **113**, 619–629.

Tian, F., Li, D. J., Fu, Q., Zhu, Z. F., Fu, Y. C., Wang, X. K., and Sun, C. Q. (2006b). Construction of introgression lines carrying wild rice (*Oryza rufipogon* Griff.) segments in cultivated rice (*Oryza sativa* L.) background and characterization of introgressed segments associated with yield-related traits. *Theor. Appl. Genet.* **112**, 570–580.

Till, B. J., Reynolds, S. H., Grean, E. A., Codomo, C. A., Enns, L. C., Johnson, J. E., Burtner, C., Odden, A. R., Young, K., Taylor, N. E., Henikoff, J. G., Comai, L., *et al.* (2003). Large-scale discovery of induced point mutations with high throughput TILLING. *Genome Res.* **13**, 524–530.

Till, B. J., Reynolds, S. H., Weil, C., Springer, N., Burtner, C., Young, K., Bowers, E., Codomo, C. A., Enns, L. C., Odden, A. R., Greene, E. A., Comai, L., *et al.* (2004). Discovery of induced point mutations in maize by TILLING. *BMC Plant Biol.* **4**, 12.

Timmerman-Vaughan, G. M., Frew, T. J., Butler, R., Murray, S., Gilpin, M., Falloon, K., Johnston, P., Lakeman, M. B., Russell, A., and Khan, T. (2004). Validation of quantitative trait loci for *Ascochyta* blight resistance in pea (*Pisum sativum* L.), using populations from two crosses. *Theor. Appl. Genet.* **109**, 1620–1631.

Tinker, N. A., and Mather, D. E. (1993). GREGOR: Software for genetic simulation. *J. Hered.* **84**, 237.

Toojinda, T., Baird, E., Booth, A., Broers, L., Hayes, P., Powell, W., Thomas, W., Vivar, H., and Young, G. (1998). Introgression of quantitative trait loci (QTL) determining stripe rust resistance in barley: An example of marker-assisted line development. *Theor. Appl. Genet.* **96**, 123–131.

Torada, A., Ikeguchi, S., and Koike, M. (2005). Mapping and validation of PCR-based markers associated with a major QTL for seed dormancy in wheat. *Euphytica* **143**, 251–255.

Torada, A., Koike, M., Mochida, K., and Ogihara, Y. (2006). SSR-based linkage map with new markers using an intraspecific population of common wheat. *Theor. Appl. Genet.* **112**, 1042–1051.

Toth, B., Francia, E., Rizza, F., Stanca, A. M., Galiba, G., and Pecchioni, N. (2004). Development of PCR-based markers on chromosome 5H for assisted selection of frost-tolerant genotypes in barley. *Mol. Breed.* **14**, 265–273.

Tranquilli, G. E., and Dubcovsky, J. (1999). Epistatic interaction between vernalization genes *Vrn-A^{m}1* and *Vrn-A^{m}2* in diploid wheat. *J. Hered.* **91**, 304–306.

Ueda, A., Shi, W., Nakamura, T., and Takabe, T. (2002). Analysis of salt-inducible genes in barley roots by differential display. *J. Plant Res.* **115**, 119–130.

Ueda, A., Kathiresan, A., Inada, M., Narita, Y., Nakamura, T., Shi, W., Takabe, T., and Bennett, J. (2004). Osmotic stress in barley regulates expression of a different set of genes than salt stress does. *J. Exp. Bot.* **55**, 2213–2218.

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

Upadhyaya, H. D., Bramel, P. J., Ortiz, R., and Singh, S. (2002). Developing a mini core of peanut for utilization of genetic resources. *Crop Sci.* **42**, 2150–2156.

Upadhyaya, H. D., Furman, B. J., Dwivedi, S. L., Udupa, S. M., Gowda, C. L. L., Baum, M., Crouch, J. H., Buhariwalla, H. K., and Singh, S. (2006a). Development of a composite collection for mining germplasm possessing allelic variation for beneficial traits in chickpea. *Plant Genet. Resour.* **4**, 13–19.

Upadhyaya, H. D., Gowda, C. L. L., Pundir, R. P. S., Reddy, V. G., and Singh, S. (2006b). Development of core subset of fingermillet germplasm using geographical origin and data on 14 quantitative traits. *Genet. Resour. Crop Evol.* **53**, 679–685.

Upadhyaya, H. D., Reddy, L. J., Gowda, C. L. L., Reddy, K. N., and Singh, S. (2006c). Development of a mini core subset for enhanced and diversified utilization of pigeonpea germplasm resources. *Crop Sci.* **46**, 2127–2132.

Utz, H. F., and Melchinger, A. E. (1996). PLABQTL: A program for composite interval mapping of QTL. *J. Quant. Trait Loci* **2**, 1–5.

Vales, M. I., Schon, C. C., Capettini, F., Chen, X. M., Corey, A. E., Mather, D. E., Mundt, C. C., Richardson, K. L., Sandoval-Islas, J. S., Utz, H. F., and Hayes, P. M. (2005). Effect of population size on the estimation of QTL: A test using resistance to barley stripe rust. *Theor. Appl. Genet.* **111**, 1260–1270.

van Berloo, R. (1999). GGT: Software for the display of graphical genotypes. *J. Hered.* **90**, 328–329.

van Berloo, R., Aalbers, H., Werkman, A., and Niks, R. E. (2001). Resistance QTL confirmed through development of QTL-NILs for barley leaf rust resistance. *Mol. Breed.* **8**, 187–195.

Van Ooijen, A. J., and Voorrips, R. E. (2001). "JoinMap (tm) 3.0: Software for the Calculation of Genetic Linkage Maps," Plant Research International, Wageningen, the Netherlands.

van Os, H., Andrzejewski, S., Bakker, E., Barrena, I., Bryan, G. J., Caromel, B., Ghareeb, B., Ishidore, E., de Jong, W., van Koert, P., Lefebvre, V., Milbourne, D., *et al.* (2006). Construction of a 10,000-marker ultradense genetic recombination map of potato: Providing a framework for accelerated gene isolation and a genomewide physical map. *Genetics* **173**, 1075–1087.

van Soest, L. J. M., Schober, B., and Tazelaar, M. F. (1984). Resistance to *Phytophthora infestans* in tuber bearing species of *Solanum* and its geographic distribution. *Potato Res.* **27**, 393–411.

Vargas, M., van Eeuwijk, F. A., Crossa, J., and Ribaut, J.-M. (2006). Mapping QTLs and QTL × environment interaction for CIMMYT maize drought stress program using factorial regression and partial least squares methods. *Theor. Appl. Genet.* **112**, 1009–1023.

Varshney, R. K., Graner, A., and Sorrells, M. E. (2005a). Genomics-assisted breeding for crop improvement. *Trends Plant Sci.* **10**, 621–630.

Varshney, R. K., Graner, A., and Sorrells, M. E. (2005b). Genic micorsatellite markers in plants: Features and applications. *Trends Biotechnol.* **23**, 48–55.

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., *et al.* (2001). The sequence of the human genome. *Science* **291**, 1304–1351.

Visscher, P. M., Thompson, R., and Haley, C. S. (1996). Confidence intervals in QTL mapping by bootstrapping. *Genetics* **143**, 1013–1020.

Vision, T. J., Brown, D. G., Shmoys, D. B., Durrett, R. T., and Tanksley, S. D. (2000). Selective mapping: A strategy for optimizing the construction of high-density linkage maps. *Genetics* **155**, 407–420.

Vladutu, C., McLaughlin, J., and Phillips, R. L. (1999). Fine mapping and characterization of linked quantitative trait loci involved in the transition of the maize apical meristem from vegetative to generative structures. *Genetics* **153**, 993–1007.

Vodkin, L. O., Khanna, A., Shealy, R., Clough, S. J., Gonzalez, D. O., Philip, R., Zabala, G., Thibaud-Nissen, F., Sidarous, M., Strömvik, M. V., Shoop, E., *et al.* (2004). Microarrays for global expression constructed with a low redundancy set of 27,500 sequenced cDNAs representing an array of developmental stages and physiological conditions of the soybean plant. *BMC Genomics* **4**, 73.

von Korff, M., Wang, H., Léon, J., and Pillen, K. (2004). Development of candidate introgression lines using an exotic barley accession (*Hordeum vulgare* ssp. *spontaneum*) as donor. *Theor. Appl. Genet.* **109**, 1736–1745.

Waldron, B. L., Moreno, S. B., Anderson, J. A., Stack, R. W., and Frohberg, R. C. (1999). RFLP mapping of QTL for *Fusarium* head blight resistance in wheat. *Crop Sci.* **39**, 805–811.

Walia, H., Wilson, C., Wahid, A., Condamine, P., Cui, X., and Close, T. J. (2006). Expression analysis of barley (*Hordeum vulgare* L.) during salinity stress. *Funct. Integr. Genomics* **6**, 143–156.

Walker, D., Boerma, H. R., All, J., and Parrott, W. (2002). Combining *cry1Ac* with QTL alleles from PI 229358 to improve soybean resistance to lepidopteran pests. *Mol. Breed.* **9**, 43–51.

Walker, D. R., Narvel, J. M., Boerma, H. R., All, J. N., and Parrott, W. A. (2004). A QTL that enhances and broadens Bt insect resistance in soybean. *Theor. Appl. Genet.* **109**, 1051–1957.

Wang, D., Arelli, P. R., Shoemaker, R. C., and Diers, B. W. (2001b). Loci underlying resistance to race 3 of soybean cyst nematode in *Glycine soja* plant introduction 468916. *Theor. Appl. Genet.* **103**, 561–566.

Wang, E., Robertson, M. J., Hammer, G. L., Carberry, P. S., Holzworth, D., Meinke, H., Chapman, S. C., Hargreaves, J. N. G., Huth, N. I., and McLean, G. (2002). Development of a generic crop model template in the cropping system model APSIM. *Eur. J. Agron.* **18**, 121–140.

Wang, H., Nussbaum-Wagler, T., Li, B., Zhao, Q., Vigouroux, Y., Faller, M., Bomblies, K., Lukens, L., and Doebley, J. F. (2005b). The origin of naked grains of maize. *Nature* **436**, 714–719.

Wang, J., van Ginkel, M., Trethowan, R., Ye, G., DeLacy, I., Podlich, D., and Cooper, M. (2004). Simulating the effects of dominance and epistasis on selection response in the CIMMYT wheat breeding program using QuCim. *Crop Sci.* **44**, 2006–2018.

Wang, J., Eagles, H. A., Trethowan, R., and van Ginkel, M. (2005c). Using computer simulation of the selection process and known gene information to assist in parental selection in wheat quality breeding. *Aust. J. Agric. Res.* **56**, 465–473.

Wang, R. L., Stec, A., Hey, J., Lukens, L., and Doebley, J. (1999). The limits of selection during maize domestication. *Nature* **398**, 236–239.

Wang, X. Y., Chen, P. D., and Zhang, S. Z. (2001a). Pyramiding and marker-assisted selection for powdery mildew resistance genes in common wheat. *Acta Genetica Sinica* **28**, 640–646 (in Chinese; summary in English).

Wang, Y. H., Liu, S. J., Ji, S. L., Zhang, W. W., Wang, C. M., Jiang, L., and Wan, J. M. (2005a). Fine mapping and marker-assisted selection (MAS) of a low glutelin content gene in rice. *Cell Res.* **15**, 622–630.

Warkentin, T., Tar'an, B., Zhang, C., Bett, K., Tullu, A., Somers, D., and Vandenberg, A. (2004). Marker-assisted selection for lodging resistance in pea. *In* "Fifth European Conference on Grain Legumes and Second International Conference on Legume Genomics and Genetics," pp. 149–150. Dijon, France.

Wasaki, J., Shinano, T., Onishi, K., Yonetani, R., Yazaki, J., Fujii, F., Shimbo, K., Ishikawa, M., Shimatani, Z., Nagata, Y., Hashimoto, A., Ohta, T., *et al.* (2006). Transcriptomic analysis indicates putative metabolic changes caused by manipulation of phosphorus availability in rice leaves. *J. Exp. Bot.* **57**, 2049–2059.

Wayne, M. L., and Mclntyre, L. M. (2002). Combining mapping and arraying: An approach to candidate gene identification. *Proc. Natl. Acad. Sci. USA* **99**, 14903–14906.

Webb, B. D. (1980). Rice quality and grades. *In* "Rice: Production and Utilization" (B. S. Luh, Ed.), pp. 543–565. Avi Publishing Co., Inc., Westport, CT.

Wenzl, P., Carling, J., Kudrna, D., Jaccoud, D., Huttner, E., Kleinhofs, A., and Kilian, A. (2004). Diversity array technology (DArT) for whole-genome profiling of barley. *Proc. Natl. Acad. Sci. USA* **101**, 9915–9920.

Werner, K., Friedt, W., and Ordon, F. (2005). Strategies for pyramiding resistance genes against the barley yellow mosaic virus complex (BaMMV, BaYMV, BaYMV-2). *Mol. Breed.* **16**, 45–55.

Whitelaw, C. A., Barbazuk, W. B., Pertea, G., Chan, A. P., Cheung, F., Lee, Y., Zheng, L., van Heeringen, S., Karamycheva, S., Bennetzen, J. L., SanMiguel, P., Lakey, N., *et al.* (2003). Enrichment of gene-coding sequences in maize by genome filtration. *Science* **302**, 2118–2120.

Wienholds, E., van Eden, F., Kosters, M., Mudde, J., Plasterk, R. H., and Cuppen, E. (2003). Efficient target-selected mutagenesis in Zebrafish. *Genome Res.* **13**, 2700–2707.

Willcox, M. C., Khairallah, M. M., Bergvinson, D., Crossa, J., Deutsch, J. A., Edmeades, G. O., Gonzalez-de-Leon, D., Jiang, C., Jewell, D. C., Mihm, J. A., Williams, W. P., and Hoisington, D. (2002). Selection for resistance to Southwestern corn borer using marker-assisted and conventional backcrossing. *Crop Sci.* **42**, 1516–1528.

Williams, K., Bogacki, P., Scott, L., Karakousis, A., and Wallwork, H. (2001). Mapping of a gene for leaf scald resistance in barley line 'B87/14' and validation of microsatellite and RFLP markers for marker-assisted selection. *Plant Breed.* **120**, 301–304.

Wilson, I. D., Barker, G. L. A., Beswick, R. W., Shepherd, S. K., Lu, C., Coghill, J. A., Edwards, D., Owen, P., Lyons, R., Parker, J. S., Lenton, J. R., Holdsworth, M. J., *et al.* (2004). A transcriptomics resource for wheat functional genomics. *Plant Biotechnol. J.* **2**, 495–506.

Winzeler, W., Winzeler, H., and Keller, B. (1995). Endopeptidase polymorphism and linkage of the *Ep-D1c* null allele with the *Lr19* leaf rust resistance gene in hexaploid wheat. *Plant Breed.* **114**, 24–28.

Wittaker, J. C., Curnow, R. N., Haley, C. S., and Thompson, R. (1995). Using marker-maps in marker-assisted selection. *Genet. Res.* **66**, 255–265.

Wu, R., and Zeng, Z.-B (2001). Joint linkage and linkage disequilibrium mapping in natural populations. *Genetics* **157**, 899–909.

Wu, J., Maehara, T., Shimokawa, T., Yamamoto, S., Harada, C., Takazaki, Y., Ono, N., Mukai, Y., Koike, K., Yazaki, J., Fujii, F., Shomura, A., *et al.* (2002b). A comprehensive rice transcript map containing 6591 expressed sequence tag sites. *Plant Cell* **14**, 525–535.

Wu, L. M., Ni, Z. F., Meng, F. R., Lin, Z., and Sun, Q. X. (2003). Cloning and characterization of leaf cDNAs that are differentially expressed between wheat hybrids and their parents. *Mol. Gen. Genet.* **270**, 281–286.

Wu, R., Ma, C.-S., and Casella, G. (2002a). Joint linkage and linkage disequilibrium mapping of qualtitative trait loci in natural mapping populations. *Genetics* **160**, 779–792.

Xia, L., Peng, K., Yang, S., Wenzl, P., de Vincente, M. C., Fregene, M., and Kilian, A. (2005). DarT for high-throughput genotyping of cassava (*Manihot esculenta*) and its wild relatives. *Theor. Appl. Genet.* **110**, 1092–1098.

Xiao, J., Li, J., Yuan, L., and Tanksley, S. D. (1995). Dominance is the major genetic basis of heterosis in rice as revealed by QTL analysis using molecular markers. *Genetics* **140**, 745–754.

Xiao, Y. H., Liu, L. L., Jiang, L., Lu, C. G., Yu, C. Y., Zhang, W. W., Zuo, J. S., Zhai, H. Q., and Wan, J. M. (2005). Development of chromosome segment substitution lines (CSSLs) derived from a hybrid rice cross, Pai64S/9311 with super high yield potential. *Rice Genet. Newsl.* **22**, 17–19.

Xie, C., and Xu, S. (1998). Efficiency of multistage marker-assisted selection in the improvement of multiple quantitative traits. *Heredity* **80**, 489–498.

Xie, X., Song, M.-H., Jin, F., Ahn, S.-N., Suh, J.-P., Hwang, H.-G., and McCouch, S. R. (2006). Fine mapping of a grain weight quantitative trait locus on rice chromosome 8 using near-isogenic lines derived from a cross between *Oryza sativa* and *Oryza rufipogon*. *Theor. Appl. Genet.* **113**, 885–894.

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

Xu, K., and Mackill, D. J. (1996). A major locus for submergence tolerance mapped on rice chromosome 9. *Mol. Breed.* **2**, 219–224.

Xu, K., Xu, X., Ronald, P. C., and Mackill, D. J. (2000). A high resolution genetic linkage map in the vicinity of the rice submergence tolerance locus *Sub1*. *Mol. Gen. Genet.* **263**, 681–689.

Xu, K., Deb, R., and Mackill, D. J. (2004b). A microsatellite marker and a codominant PCR-based marker for marker-assisted selection of submergence tolerance in rice. *Crop Sci.* **44**, 248–253.

Xu, K., Xia, X., Fukao, T., Canlas, P., Maghirang-Rodriguez, R., Heuer, S., Ismail, A. I., Bailey-Serres, J., Ronald, P. C., and Mackill, D. J. (2006). *Sub1A* is an ethylene response factor-like gene that confers submergence tolerance to rice. *Nature* **442**, 705–708.

Xu, Y., Ishii, T., and McCouch, S. R. (2003). Marker-assisted evaluation of germplasm resources for plant breeding. *In* "Rice Science: Innovations and Impact for Livelihood" (T. W. Mew, D. S. Brar, S. Peng, and B. Hardy, Eds.), pp. 213–229. Proceedings of the 24th International Rice Res. Conference, 16–19 September 2002, Beijing, China. IRRI, CAE, and CAAS, Manila, Philippines.

Xu, Y., Beachell, H., and McCouch, S. R. (2004a). A marker-based approach to broadening the genetic base of rice (*Oryza sativa* L.) in the U.S. *Crop Sci.* **44**, 1947–1959.

Yadav, R. S., Hash, C. T., Bidinger, F. R., Cavan, G. P., and Howarth, C. J. (2002). Quantitative trait loci associated with traits determining grain and stover yield in pearlmillet under terminal drought stress conditions. *Theor. Appl. Genet.* **104**, 67–83.

Yamamoto, T., Kuboki, Y., Lin, S. Y., Sasaki, T., and Yano, M. (1998). Fine mapping of quantitative trait loci *Hd-1*, *Hd-2* and *Hd-3*, controlling heading date of rice, as single Mendelian factors. *Theor. Appl. Genet.* **97**, 37–44.

Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T., and Dubcovsky, J. (2003). Positional cloning of the wheat vernalization gene *VRN1*. *Proc. Natl. Acad. Sci. USA* **100**, 6263–6268.

Yan, L., Loukoianov, A., Blechl, A., Tranquilli, G., Ramakrishna, W., SanMiguel, P., Bennetzen, J. L., Echenique, V., and Dubcovsky, J. (2004a). The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. *Science* **303**, 1640–1644.

Yan, W., Rutger, J. N., Bockelman, H. E., and Tai, T. (2004b). Development of a core collection from the USDA rice germplasm collection. *In* "B. R. Wells Rice Research Studies 2003" (R. J. Norman, R. J. Meullenet, and J.-F. Moldenhauer, Eds.), pp. 88–96. Arkansas Agric. Expt. Res. St. Series 517 (http://www.uark.edu/depts/agripub/publications/research).

Yang, L., Zheng, B., Mao, C., Qi, X., Liu, F., and Wu, P. (2004). Analysis of transcript that are differentially expressed in three sectors of the rice root system under water deficit. *Mol. Gen. Genom.* **272**, 433–442.

Yang, X., Rupe, M., Bickel, D., Arthur, L., Smith, O., and Guo, M. (2006). Effects of cis-trans-regulation on allele-specific transcript expression in the meristems of maize hybrids. *In* "48th Annual Maize Genetic Conference," 9–12 March 2006. California. p. 118.

Yang, Z. P., Gilbert, J., Somers, D. J., Fedak, G., Procunier, J. D., and McKenzie, I. H. (2003). Marker-assisted selection of fusarium head blight resistance genes in two double haploid populations of wheat. *Mol. Breed.* **12**, 309–317.

Yano, M., Katayose, Y., Ashikari, M., Yamanouchi, U., Monna, L., Fuse, T., Baba, T., Yamamoto, K., Umehara, Y., Nagamura, Y., and Sasaki, T. (2000). *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. *Plant Cell* **12**, 2473–2483.

Yao, Y., Ni, Z., Zhang, Y., Chen, Y., Ding, Y., Han, Z., Liu, Z., and Sun, Q. (2005). Identification of differentially expressed genes in leaf and root between wheat hybrid and its parental inbreds using PCR-based cDNA subtraction. *Plant Mol. Biol.* **58**, 367–384.

Yim, Y.-S., Davis, G. L., Duru, N. A., Musket, T. A., Linton, E. W., Messing, J. W., McMullen, M. D., Soderlund, C. A., Polacco, M. L., Gardiner, J. M., and Coe, E. H., Jr. (2002). Characterization of three maize bacterial artificial chromosome libraries toward

anchoring of the physical map to the genetic map using high-density bacterial artificial chromosome filter hybridization. *Plant Physiol.* **130**, 1686–1696.

Yin, X., Kropff, M. J., and Stam, P. (1999). The role of ecophysiological models in QTL analysis: The example of specific leaf area in barley. *Heredity* **82**, 415–421.

Yin, X., Stam, P., Kropff, M. J., and Schapendonk, A. H. C. M. (2003). Crop modeling, QTL mapping, and their complementary role in plant breeding. *Agron. J.* **95**, 90–98.

Yin, X., Struik, P. C., and Kropff, M. J. (2004). Role of crop physiology in predicting gene-to-phenotype relationships. *Trends Plant Sci.* **9**, 426–432.

Yin, X., Struik, P. C., Tang, J., Qi, C., and Liu, T. (2005). Model analysis of flowering phenology in recombinant inbred lines of barley. *J. Exp. Bot.* **56**, 959–965.

Yoon, D.-B., Kang, K.-H., Kim, H.-J., Ju, H.-G., Kwon, S.-J., Suh, J.-P., Jeong, O.-Y., and Ahu, S.-N. (2006). Mapping quantitative trait loci for yield components and morphological traits in an advanced backcross population between *Oryza grandiglumis* and the *O. japonica* cultivar Hwaseongbyeo. *Theor. Appl. Genet.* **112**, 1052–1062.

Yoshida, S., and Hasegawa, S. (1982). The rice root system: Its development and function. *In* "Drought Resistance in Crops with Emphasis on Rice," pp. 97–114. International Rice Research Institute, Los Baños, Laguna, Philippines.

Yoshihashi, T. (2002). Quantitative analysis on 2-acetyl-1-pyrroline of an aromatic rice by stable isotope dilution method and on model studies on its formation during cooking. *J. Food Sci.* **67**, 619–622.

Yoshimura, S., Yoshimura, A., Iwata, N., McCouch, S. R., Abenes, M. L., Baraoidan, M. R., Mew, T. W., and Nelson, R. J. (1995). Tagging and combining bacterial blight resistance genes in rice using RAPD and RFLP markers. *Mol. Breed.* **1**, 375–387.

Young, N. D. (1999). A cautiously optimistic vision for marker assisted breeding. *Mol. Breed.* **5**, 505–510.

Yousef, G. G., and Juvik, J. A. (2001a). Comparison of phenotypic and marker-assisted selection for quantitative traits in sweet corn. *Crop Sci.* **41**, 645–655.

Yousef, G. G., and Juvik, J. A. (2001b). Evaluation of breeding utility of a chromosomal segment from *Lycopersicon chmielewskii* that enhances cultivated tomato soluble solids. *Theor. Appl. Genet.* **103**, 1022–1027.

Yousef, G. G., and Juvik, J. A. (2002). Enhancement of seedling resistance in sweet corn by marker-assisted backcrossing of beneficial QTL. *Crop Sci.* **42**, 96–104.

Yu, J., Hu, S., Wang, J., Wong, G. K.-S., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., Zhang, X., Cao, M., Liu, J., *et al.* (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* **296**, 79–92.

Yu, J., Wang, J., Lin, W., Li, S., Li, H., Zhou, J., Ni, P., Dong, W., Hu, S., Zeng, C., Zhang, J., Zhang, Y., *et al.* (2005). The genome of *Oryza sativa*: A history of duplications. *PLoS Biol.* **3**, 0266–0281.

Yu, K., Park, S. J., and Poysa, V. (2000). Marker-assisted selection of common beans for resistance to common bacterial blight: Efficiency and economics. *Plant Breed.* **119**, 411–415.

Yu, L.-X., and Setter, T. L. (2003). Comparative transcriptional profiling of placenta and endosperm in developing maize kernels in response to water deficit. *Plant Physiol.* **131**, 568–582.

Yu, S. B., Li, J. X., Tan, Y. F., Gao, Y. J., Li, X. H., Zhang, Q. F., and Maroof, M. A. S. (1997). Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. *Proc. Natl. Acad. Sci. USA* **94**, 9226–9231.

Yuan, Y., SanMiguel, P. J., and Bennetzen, J. L. (2003). High Co*t* sequence analysis of the maize genome. *Plant J.* **34**, 249–255 (Erratum: *Plant J.* **36**, 430).

Yue-guang, W., Qi-yun, D., Feng-shan, L., Quan-hua, X., Ji-ming, L., Yue-dong, X., Shi-mong, S., Bao-tai, G., Long-ping, Y., and Bin, W. (2004). Molecular marker-assisted selection for yield-enhancing genes in the progeny of Minghui63 x *Oryza rufipogon*.

Agric. Sci. (in Chinese) **3**, 89–93 (http://www.wanfangdata.com.cn/qikan/periodical.articles/zgnykx-e/zgny2004/0402/04020).

Yun, S. J., Geynis, L., Bossolini, E., Hayes, P. M., Matus, I., Smith, K. P., Steffenson, B. J., Tuberosa, R., and Muehlbauer, G. J. (2006). Validation of quantitative trait loci for multiple disease resistance in barley using advanced backcross lines developed with a wild barley. *Crop Sci.* **46**, 1179–1186.

Yuqing, H., Li, X., Zhang, J., Jiang, G., Liu, S., Chen, S., Tu, J., Xu, C., and Zhang, Q. (2004). Gene pyramiding to improve hybrid rice by molecular marker technique. *In* "Proceedings of the Fourth International Crop Science Congress" (http://www.regional.org.au/au/cs/2004/poster/3/4/4/1041-heyuqing.htm).

Zale, J. M., Clancy, J. A., Ullrich, S. E., Jones, B. L., Hays, P. M. and the North American Barley Genome Mapping Project. (2000). Summary of barley malting QTL mapped in various mapping populations. *Barley Genet. Newsl.* **30**, 1–4.

Zhang, C., Bunyamin, T., Warkentin, T., Tullu, A., Bett, K. E., Vandenberg, B., and Somers, D. J. (2006a). Selection for lodging resistance in early generations of field pea by molecular markers. *Crop Sci.* **46**, 321–329.

Zhang, H., Sreenivasulu, N., Weschke, W., Stein, N., Rudd, S., Radchuk, V., Potokina, E., Scholz, U., Schweizer, P., Zeirold, U., Langridge, P., Varshney, R. K., *et al.* (2004). Large-scale analysis of the barley transcriptome based on expressed sequence tags. *Plant J.* **40**, 276–290.

Zhang, J., Zheng, H. G., Aarti, A., Pantuwan, G., Nguyen, T. T., Tripathi, J. N., Sarial, A. K., Robin, S., Babu, R. C., Nguyen, B. D., Sarkarung, S., Blum, A., *et al.* (2001). Locating genomic regions associated with components of drought resistance in rice: Comparative mapping within and across species. *Theor. Appl. Genet.* **103**, 19–29.

Zhang, Q., Gao, Y. J., Yang, S. H., Ragab, R. A., Saghai Maroof, M. A., and Li, Z. B. (1994). A diallel analysis of heterosis in elite hybrid rice based on RFLPs and microsatellites. *Theor. Appl. Genet.* **89**, 185–192.

Zhang, W., Collins, A., Maniatis, N., Tapper, W., and Morton, N. E. (2002). Properties of linkage disequilibrium LD maps. *Proc. Natl. Acad. Sci. USA* **99**, 17004–17007.

Zhang, Z., Bradbury, P. J., Kroon, D. E., Casstevens, T. M., and Buckler, E. S. (2006b). TASSEL 2.0: A software package for association and diversity analyses in plants and animals, Plant & Animal Genomes XIV Conference.

Zhao, J. H. (2004). 2LD, GENECOUNTING and HAP: Computer programs for linkage disequilibrium analysis. *Bioinformatics* **20**, 1325–1326.

Zhao, H., Zhang, S., Merikangas, K. R., Trixler, M., Wildenauer, D. B., Sun, F., and Kidd, K. K. (2000). Transmission/disequilibrium tests using multiple tightly linked markers. *Am. J. Hum. Genet.* **67**, 936–946.

Zhao, Q., Zhang, Y., Cheng, Z., Chen, M., Wang, S., Feng, Q., Huang, Y., Li, Y., Tang, Y., Zhou, B., Chen, Z., Yu, S., *et al.* (2002). A fine physical map of the rice chromosome 4. *Genome Res.* **12**, 817–823.

Zhao, X. C., Batey, I. L., Sharp, P. J., Crosbie, G., Barclay, I., Wilson, R., Morell, M. K., and Appels, R. (1998). A single genetic locus associated with starch granule properties and noodle quality in wheat. *J. Cereal Sci.* **27**, 7–13.

Zhou, P. H., Tan, Y. F., He, Y. A., Xu, C. G., and Zhang, A. (2003a). Simultaneous improvement of four quality traits of Zhenshan 97, an elite parent of hybrid rice, by molecular marker-assisted selection. *Theor. Appl. Genet.* **106**, 326–331.

Zhou, W.-C., Kolb, F. L., Bai, G.-H., Domier, L. L., Boze, L. K., and Smith, N. J. (2003b). Validation of a major QTL for scab resistance with SSR markers and use of marker-assisted selection in wheat. *Plant Breed.* **122**, 40–46.

Zhu, S., and Kaeppler, H. F. (2003). A genetic linkage map for hexaploid, cultivated oat (*Avena sativa* L.) based on an intraspecific cross Ogle/MAM17–5. *Theor. Appl. Genet.* **107**, 26–35.

Zhu, T., Budworth, P., Chen, W., Provart, N., Chang, H.-S., Guimil, S., Su, W., Estes, B., Zou, G., and Wang, X. (2003). Transcriptional control of nutrient partitioning during rice grain filling. *Plant Biotechnol. J.* **1**, 59–70.

Zinselmeier, C., Sun, Y., Helentjaris, T., Beatty, M., Yang, S., Smith, H., and Habben, J. (2002). The use of gene expression profiling to dissect the stress sensitivity of reproductive development in maize. *Field Crop Res.* **73**, 111–121.