

Selective Genotyping and Pooled DNA Analysis: An Innovative Use of an Old Concept

Yunbi Xu¹, Jiankang Wang², and Jonathan Crouch¹

¹Genetic Resources and Enhancement Unit, International Maize and Wheat Improvement Center (CIMMYT), Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico. www.cimmyt.org Email y.xu@cgiar.org

²Crop Research Informatics Laboratory, CIMMYT China Office/Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing 100081, China

Abstract

Selective genotyping and pooled DNA analysis using material from the two tails of the phenotypic distribution of a population provide an effective alternative to analysis of the entire population for genetic mapping. Past applications of this approach have been confounded by the use of small total and tail population sizes and insufficient marker density, which results in a high probability of false positive marker associations. Simulation studies indicate that when these issues are resolved selective genotyping and pooled DNA analysis can be effectively used for genetic mapping of QTL with relatively small effects, as well as linked and interacting QTL. Using diverse germplasm including cultivars, landraces, recombinant inbreds, doubled haploids, introgression lines and mutant libraries, it is theoretically possible that one 384-well plate could be designed to cover almost all major gene/QTL controlled agronomic traits of importance in a crop species. This “all-in-one plate” approach is feasible in all species where high density marker coverage is available. Selective genotyping can also be used for genomewide association mapping where it can be integrated with selective phenotyping approaches. In CIMMYT, over 1600 maize plants/lines have been collected from genetics and breeding programs worldwide. These represent phenotypic extremes for many traits including drought tolerance, disease and pest resistance, and grain quality traits. These collections have been genotyped using a maize 1536 SNP chip, allowing us for the first time to test the feasibility of a one-step simultaneous marker-trait association analysis for a large number of agronomic traits.

Media summary

Selective genotyping and pooled DNA analysis can be innovatively used for highly effective, efficient and precise genetic mapping of agronomic traits in crop plants

Key words

Molecular markers, selective genotyping, pooled DNA analysis, genetic mapping, marker-assisted selection

The original concept and issues associated with its application

Two contrasting approaches have been routinely used for marker-trait association analysis (i.e. genetic mapping); (i) testing the average phenotypic difference among groups of individuals with distinct marker genotypes and (ii) testing marker allele frequencies amongst groups of individuals with distinct phenotypes. The first approach is usually based on genotyping an entire segregating population with markers evenly covering the entire genome. It is extensive, time-consuming and expensive, while generating precision phenotype data at this scale may be logistically difficult or even impossible for some traits. The second approach involves selecting only the individuals from the high and low tails of the phenotypic distribution across the population (‘selective genotyping’; Lebowitz et al. 1987; Lander and Botstein 1989), followed by marker analysis of individuals or pools of individuals within those tails of the phenotypic distribution (‘pooled DNA analysis’). Marker-trait association is then inferred by finding the differences in allelic frequency (Stuber et al. 1980, 1982) or signal strength between the two DNA pools of individuals with contrasting phenotypes. This method has been referred to as ‘tail analysis’ (Hillel et al. 1990; Dunnington et al. 1992; Plotsky et al.

1993), ‘bulked segregant analysis’ (Giovannoni et al. 1991; Michelmore et al. 1991), or ‘selective DNA pooling’ (Darvasi and Soller 1994). It can be bidirectional if the two tails of the distribution are considered, or unidirectional if only one tail is considered. The latter is more suitable for traits that have been subjected to strong negative or lethal selection in unfavorable environments. Selective genotyping and pooled DNA analysis have been widely used in genetic mapping in plants with numerous reports for single major genes (e.g., Barua et al., 1993; Hormaza et al. 1994; Villar et al. 1996; van Treuren 2001; Zhang et al. 2002) and for detection and validation of quantitative trait loci (QTL) (Foolad and Jones 1993; Zhang et al. 2003; Wingbermuehle et al. 2004; Coque and Gallais 2006), including traits controlled by a few major-effect QTL (Quarrie et al. 1999). This approach has also been used to significantly change marker allele frequency through two cycles of recurrent selection for marker loci located in the vicinity of target QTL (Moreau et al. (2004).

There are several issues that have confounded many applications of selective genotyping and pooled DNA analysis (Xu and Crouch (2008). (1) A relatively small number of markers has often been used to cover the whole genome with the assumption that genes can be readily identified using a marker density of 15-25cM, while frequently this is not feasible. (2) Contrasting individuals have been selected from a relatively small-size population (e.g., 100-300 individuals), this reduces the power of QTL detection such that only large-effect genes/QTL may be detected, which depending on the germplasm used may not exist for many complex agronomic traits. (3) When the allele signal is determined through a gel-based genotyping system, allele frequency in each pool cannot be quantified accurately and the signal generated by a rare allele present in only a small number of individuals of pool may not be detected which substantially reduces the power of the approach. (4) Pools are often based on a relatively small number of individuals (10-15) causing a high level of false positive marker-trait associations. These constraints have confounded many studies in this area leading to an apparent mixed success in the literature regarding the application of selective genotyping and pooled DNA analysis.

Revisiting the power of selective genotyping and pooled DNA analysis

Selective genotyping and pooled DNA analysis has significant advantages in terms of cost savings, compared to entire population analysis. Consider analysis of a large population with 1000 individuals where 30 individuals from each tail are selected as phenotypic extremes. Selective genotyping will only cost 6% of the total cost required for genotyping the entire population. Clearly, the bigger the original population size, the greater the power of this approach and the higher the savings compared to entire population genotyping. As whole genome sequences become available for more crop species, it will be increasingly common to carry out high throughput high precision genotyping for thousands of plants using single nucleotide polymorphism (SNP) markers. However, this relies on the existence of a high throughput precision phenotyping system for the target trait. Of course, pooled DNA analysis would provide a further magnitude saving in genotyping costs (0.2% of entire population individual genotype analysis). However, although readily achievable in many crops with SSR markers, this will require significant protocol advancements for SNP markers.

Reducing the size of a mapping population will, in general, decrease the power of QTL detection (Charcosset and Gallais, 1996), and increase the QTL confidence interval, as well as increasing the probability of detecting false positive QTL. A recent study on the optimization of selective genotyping has indicated that for a given population size of phenotyped plants the optimal proportion of individuals for selective genotyping is around 30% for each tail (Gallais et al. 2007). As the population size increases, the proportion of individuals required for a given power of QTL detection will decrease such that at a certain point an absolute number of plants from each tail will become the critical issue (see below for further details).

It can be inferred that plants with extreme phenotype chosen for selective genotyping would be those with an accumulation of favourable alleles from multiple loci with various additive effects. However, several significant questions remain to be answered before the full potential of selective genotyping and pooled DNA analysis can be achieved, including: (1) can selective genotyping be used to replace entire population analysis

for both qualitative and quantitative traits; (2) how many independent genes can be identified simultaneously; (3) can selective genotyping be used for mapping linked genes and/or genes with epistatic interactions; and (4) can selective genotyping be used for fine mapping required for map-based cloning?

Simulation studies have been carried out using QTL IciMapping (available at <http://www.isbreeding.net>), an integrated computing package by which common QTL mapping methods including single marker analysis, traditional interval mapping (Lander and Botstein 1989), inclusive composite interval mapping for additive (Li et al. 2007) and interacting (Li et al. 2008) QTL. Several parameters associated with selective genotyping were simulated based on the assumption that phenotypic extremes from two tails of a recombinant inbred population can be reliably selected and that they can be genotyped either individually so that the allele frequency in each tail can be inferred or genotyped using bulked DNA from each tail so that the allele frequencies can be estimated based on the relative signal strength of two DNA pools. Simulated parameters include total population size (ranged from 200 to 3000), tail population size (15 to 100 plants in each tail, equivalent to 13% to 50% of selection rate), number of QTL (1 to 5), marker density (1 to 15 cM), QTL effect (explaining 1 to 20% of phenotypic variation), two linked QTL, and two QTL with epistatic interaction. One hundred simulation runs were carried out for each scenario from which the power of QTL detection and the mean LOD score were then calculated.

Comparative analysis of two selective genotyping strategies (Figure 1) indicates that conventional selective genotyping (Figure 1A, Strategy A, where relatively small total and tail population sizes were used with a low density of marker coverage), resulted in the detection of only one marker in the target region with an average LOD score of 3.94 and power of detection of 67%, which cannot be distinguished from a false positive without validation analysis through genotyping the entire population. In contrast, Strategy B (Figure 1B), where large total and tail population sizes are used along with a high density of marker coverage, resulted in the detection of multiple markers around the target region with the highest having a LOD score of 10.37 and a power of detection of 98%. Although the region with a LOD score larger than 6 spans about 15 cM (Figure 1B), there is a sharp peak for LOD scores within a 3 cM region that directly brackets the target QTL. This trend is consistent in all cases where a high density of marker coverage is used. This suggests that selective genotyping can be used for fine mapping when a high-density marker map is available.

When various QTL effects (responsible for 1% to 20% of the total phenotypic variation), tail sizes (15 to 100) and total population sizes (200, 500, 1000, and 3000) (Figure 2) are used in the simulation analysis, the power of QTL detection indicates the optimum total and tail population sizes required for detection of small QTL. To identify QTL explaining 15% of the phenotypic variation with a 95% or higher power of QTL detection, will require a population size of 200 or more with a minimum tail size of 15, which matches most reported cases of successful use of bulked DNA analysis. However, to detect QTL of small effect, ranging from 3 to 10% of the phenotypic variation, 50-100 individuals need to be selected from each tail of a population with 1000 individuals, in order to have a 95% power of QTL detection (Figure 2). Our simulation analysis also indicates that the power of detection will not change when multiple QTL (2 to 5) are involved but they are independent to each other (data not shown). Thus, the maximum number of QTL that can be simultaneously detected will be at least equal to the number of chromosomes of the crop species, assuming that the QTL are located on different chromosomes.

Genetic separation of two tightly linked QTL has been a challenge for most, if not all, statistical methods even through the entire population genotyping approach. Figure 3 provides the result for two QTL linked at a distance of 25 cM. The two target regions (spanning 4cM) are associated with two peaks for both power of detection and mean LOD score although the LOD scores reach the QTL threshold across the entire region containing the two QTL. Among the six scenarios studied through simulation analysis (Figure 3), those based on larger total and tail population sizes show not only stronger association but also more distinguishable peaks for two linked QTL, compared to those situations based on smaller total and tail population sizes.

For scenarios based on epistatic interactions between two QTL located on different chromosomes where each explains 5% of the phenotypic variation, the power of QTL detection of selective genotyping is very close to that seen for identifying independent QTL. As shown in Figure 4, however, population size should be larger than 500 combined with a tail size of 50 or more, in order to detect both QTL with a power of QTL detection of 95% or more. Even for a population with 3000 individuals, the tail size should be at least 30 in order to have a power of QTL detection of 95% or more for the two epistatic QTL. There is no difference in power of QTL detection for two different genetic models: one with positive additive effect for both QTL (Figure 4A) and the other with opposite additive effects for the two QTL (Figure 4B). In simulation, we also found that when individual QTL does not have any additive effects, the two interacting QTL cannot be readily detected by selective genotyping (results not shown).

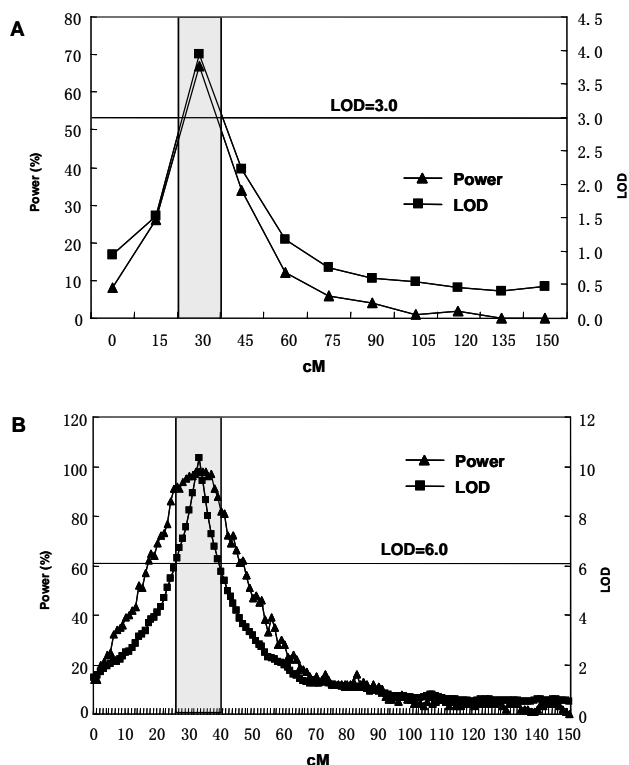


Figure 1. Effects of selective genotyping strategies on power of QTL detection and mean LOD score around the target region (15cM, grey area) assuming QTL explaining 10% of the phenotypic variation. Strategy A: population size =200, tail size=15, marker density=15 cM. B: population size=500, tail size=30, marker density=1 cM.

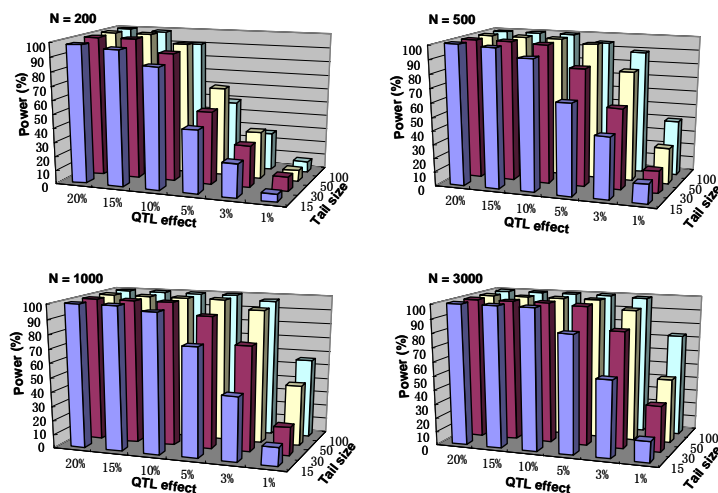


Figure 2. Power of QTL detection of selective genotyping for various QTL effects (percentage of the phenotypic variation explained) (1-20%) and tail sizes (15-100).

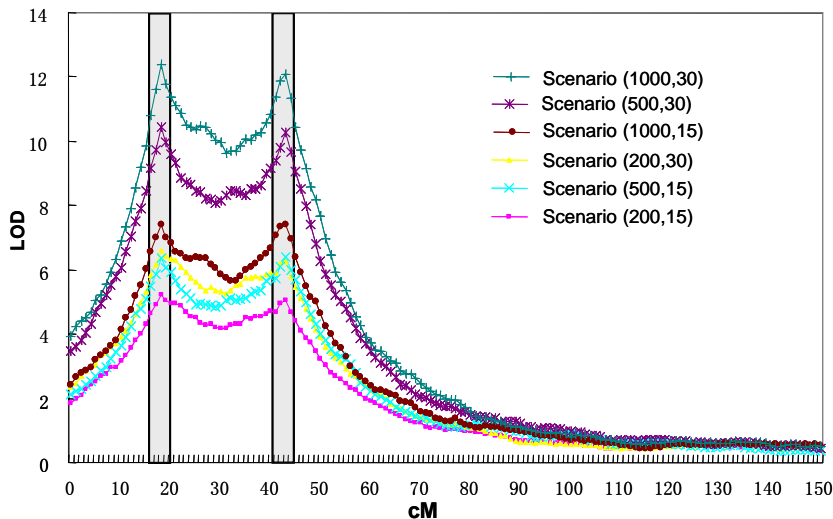


Figure 3. Mean LOD scores of selective genotyping for two linked QTL (located at 18 cM and 43 cM points on the same chromosome, each explaining 5% of the phenotypic variation). Six scenarios are compared, and the two numbers in parenthesis indicate population and tail sizes. The grey areas indicate two target QTL regions, each spanning four cM.

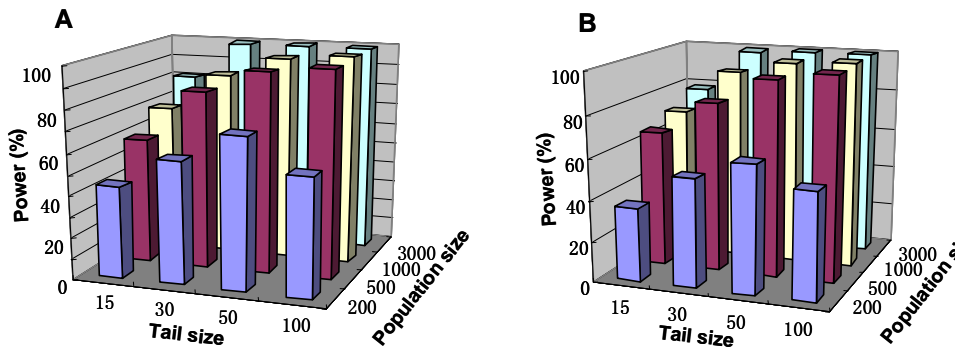


Figure 4. Power of QTL detection of selective genotyping for two epistatic QTL located on chromosomes 1 (18 cM) and 2 (33 cM). **A**: two interacting QTL with $a_1=a_2=aa=0.2236$; **B**: two interacting QTL with $a_1=aa=0.2236$, $a_2=-0.2236$. a_1 , a_2 and aa are additive effects for QTL 1 (chromosome 1) and QTL 2 (chromosome 2) and additive by additive interaction between the two QTL. The power for each case was presented by the mean value averaged from the two QTL.

Innovative uses of selective genotyping and pooled DNA analysis

Replacement of entire population genotyping

A general conclusion we can draw from the simulation analyses described above is that selective genotyping can be used to replace the entire population genotyping approach in almost all cases we discussed here, if the total and tail population sizes are large enough and a high-density of marker coverage is used. In addition, there is no need to eliminate false positive markers by entire population genotyping as the power of QTL detection under these conditions is extremely high. Selective genotyping can be used for effective genetic mapping of QTL with relatively small effects as well as for QTL with epistatic interactions or link QTL. In addition, selective genotyping can be used for fine mapping to narrow down associated genetic regions to less than 1cM or even few candidate genes. Our recommendation for selective genotyping for QTL of different effects would be: 20 individuals from each tail of a population with 200 individuals for large QTL (15% or bigger), 50 individuals from each tail of a population with 500-1000 individuals for medium-size QTL (3-10%), and 100 individuals from each tail of a population with 3000-5000 individuals for small QTL (0.2 to 3%). The numbers are nearly

equivalent to those required for entire population genotyping. As the number of QTL and their effects are unknown in most cases, the total and tail population sizes required for a specific experiment will depend on the objectives of the study, i.e. how much genetic variation and how large QTL to be detected.

“All-in-one plate” - genetic mapping of all target traits in one step

Selective genotyping and pooled DNA analysis can be explored for large-scale and effective genetic mapping. A large number of trait-specific genetic and breeding materials, with novel properties including inbreds/cultivars with extreme phenotypes, eternal/fixed segregating populations (e.g., recombinant inbred lines), doubled haploids, near isogenic lines, introgression lines, genetic stocks (e.g. single segment substitution lines) and mutant libraries, have been developed and maintained across the world for most crops. These are valuable directly for the purpose they were developed but also offer a novel resource for genetic mapping and gene discovery when used collectively. These materials have often been phenotyped in multiple environments due to their permanently fixed genetic composition. By collecting phenotypic extremes from currently available genetic and breeding materials, and utilizing selective genotyping and pooled DNA analysis (once a suitable protocol is developed for SNP markers), it is theoretically possible that one 384-well plate could be designed to cover almost all major gene/QTL controlled agronomic traits of importance in a crop species. In CIMMYT, over 1600 maize plants/lines have been collected from genetics and breeding programs worldwide with phenotypic extremes for many traits including drought tolerance, disease and pest resistance, and grain quality traits. These collections have been genotyped using a maize SNP chip, allowing us for the first time to test the feasibility of a one-step simultaneous marker-trait association analysis for a large number of agronomic traits (see below for further details).

Genomewide association mapping

Developments in SNP genotyping technologies and methodologies recently reported in human genomics have made it possible now to carry out genomewide linkage-disequilibrium based association mapping in human beings by using an integrated technology package including selective genotyping, pooled DNA analysis and microarray-based SNP genotyping with 100,000 markers (Sham et al. 2002; Meaburn et al. 2006; Yang et al. 2006). This system has the power to estimate allele frequencies and identify unique alleles from a pooled DNA sample of several hundreds of individuals. If this approach is successfully translated to plants it will resolve many of the constraints of pooled DNA analysis. The high frequency of false positive markers that would be detected when substantially fewer plants are used in each pool could be avoided if a pooled DNA can be formed using many more plants selected from a large population, as described in the previous section. However, optimizing SNP genotyping systems for pooled DNA analysis is considerably more complicated than for SSR markers and suffers a much higher level of redundancy. Where this has been achieved in human genomics, it required at least half a million SNPs as a starting point in order to identify 100,000 optimized SNPs suitable for pooled analysis. This density of SNP markers will soon be available in rice and maize and in due course other crops when whole genome sequences are generated. However, SNP genotyping systems using pooled DNA from heterogeneous individuals may have to be established and optimized specifically for each crop species and set of markers.

Genomewide association mapping may provide a shortcut to discovering functional alleles and allelic variations that are associated with agronomic traits of interest. Selective genotyping and pooled DNA analysis can be extended to using inbred lines with extreme phenotypes selected from various collections of germplasm such as the mini-composite collections developed by Generation Challenge Program. This is in principal similar to linkage disequilibrium-based association mapping but using selected phenotypic extremes. For association mapping of quantitative traits governed by a large number of minor genes which interact with each other and the environment, selective genotyping will face the same challenges as experienced with linkage-based QTL mapping using entire population genotyping.

Integration with selective phenotyping

The selective phenotyping method involves preferentially selecting individuals to maximize their genotypic dissimilarity. Selective phenotyping is most effective when prior knowledge of genetic architecture allows focus on specific genetic regions (Jannink 2005; Jin et al., 2004) and specific allele combinations. As genotyping becomes cheaper, it may be more efficient to first carry out low density genotyping of the whole population in order to identify the most informative subset of individuals in terms of minimum level of relatedness between individuals plus optimum subpopulation structure and allele representativeness. Then carry out precision phenotyping of this subset, particularly for the traits that are difficult or expensive to evaluate. And then finally carry out dense whole genome genotyping of the individuals from the tails of the phenotypic distribution. In this way, the total number of individuals to be phenotyped and genotyped may not change, but the power of the analysis will be dramatically increased. This approach could also be achieved for traits where phenotypic extremes can be easily identified by using a simple screening method, for example abiotic stress tolerance where a large number of plants/families can be eliminated easily under stress conditions through visual scoring. As the original population can be selected under a strong environmental stress to eliminate a large proportion of the plants, only the most stress tolerant, and probably the most stress sensitive plants too, are selected for genotyping. Following selective genotyping of the individuals with extreme phenotypes, precision phenotyping of the resultant subset of individuals can be carried out using physiological component and surrogate traits. High-density planting and selection at early stages of plant development, combined with selective phenotyping and genotyping should also be investigated as a potential option for some traits in order to allow one to work with more plants/families at the same cost (Xu and Crouch 2008). Where the target trait is influenced by planting density or strong selection pressure this will clearly confound the ability to make genetic gain. However, many major-gene controlled traits can be investigated in this way without much disturbance.

It can be expected that selective genotyping and pooled DNA analysis, which have been widely used with mixed success in genetic mapping, will become increasingly important in genetic mapping and marker-assisted selection, and will gradually replace entire population genotyping in many cases. Selective genotyping will greatly facilitate and improve genetic mapping and marker-assisted breeding procedures in general. As genomewide selective genotyping become possible, an effective information management and data analysis system will be required to make full use of the potentialities of selective genotyping in genetics, genomics and plant breeding.

References

- Barua UM, Chalmers KJ, Hackett CA, Thomas WT, Powell W, Waugh R. 1993. Identification of RAPD markers linked to a *Rhynchosporium secalis* resistance locus in barley using near-isogenic lines and bulked segregant analysis. *Heredity* 71:177-184.
- Charcosset A, Gallais A. 1996. Estimation of the contribution of quantitative trait loci (QTL) to the variance of quantitative trait by means of genetic markers. *Theor. Appl. Genet.* 93:1193-1201.
- Coque M, Gallais A. 2006. Genomic regions involved in response to grain yield selection at high and low nitrogen fertilization in maize. *Theor. Appl. Genet.* 112:1205-1220.
- Darvasi A, Soller M. 1994. Selective DNA pooling for determination of linkage between a molecular marker and a quantitative trait. *Genetics* 138:1365-1373.
- Dunninton EA, Haberefeld A, Stallard LG, Siegel PB, Hillel J. 1992. Deoxyribonucleic-acid fingerprint bands linked to loci coding for quantitative traits in chicken. *Poultry Science* 71:1251-1258.
- Foolad MR, Jones RA. 1993. Mapping salt-tolerance genes in tomato (*Lycopersicon esculentum*) using trait-based marker analysis. *Theor. Appl. Genet.* 87:184-192.
- Gallais A, Moreau L, Charcosset A. 2007. Detection of marker-QTL associations by studying change in marker frequencies with selection. *Theor. Appl. Genet.* 114: 669-681.
- Giovannoni JJ, Wing RA, Ganai MW, Tanksley SD. 1991. Isolation of molecular markers from specific chromosomal interval using DNA pools from existing mapping populations. *Nucl. Acid Res.* 19:6553-6558.

- Hillel J, Avner R, Baxter-Jones C, Dunnington EA, Cahaner A, Siegel PB. 1990. DNA fingerprints from blood mixes in chickens and turkeys. *Animal Biotech.* 2:201-204.
- Hormaza JI, Dollo L, Polito VS. 1994. Identification of a RAPD marker linked to sex determination in *Pistacia vera* using bulked segregant analysis. *Theor. Appl. Genet.* 89: 9-13.
- Jannink JL. 2005. Selective phenotyping to accurately mapping quantitative trait loci. *Crop Sci.* 45: 901-908.
- Jin C, Lan H, Attie AD, Churchill GA, Bulutuglo D, Yandell BY. 2004. Selective phenotyping for increased efficiency in genetic mapping study. *Genetics* 168:2285–2293.
- Lander ES, Botstein D. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185-199.
- Lebowitz RL, Soller M, Beckmann JS. 1987. Trait-based analysis for the detection of linkage between marker loci and quantitative trait loci in cross between inbred lines. *Theor. Appl. Genet.* 73:556-562.
- Li H, Ye G, Wang J. 2007. A modified algorithm for the improvement of composite interval mapping. *Genetics* 175: 361-374.
- Li H, Ribaut, JM, Li Z, Wang J. 2008. Inclusive composite interval mapping (ICIM) for digenic epistasis of quantitative traits in biparental populations. *Theor. Appl. Genet.* 116: 243-260.
- Meaburn E, Butcher LM, Schalkwyk LC, Plomin R. 2006. Genotyping pooled DNA using 100K SNP microarrays: a step towards genomewide association scans. *Nucl. Acids Res.* 34(4), e28.
- Michelmore RW, Paran I, Kesseli RV. 1991. Identification of markers linked to disease resistance gene by bulked segregant analysis: a rapid method to detect markers in specific genomic regions using segregating populations. *Proc. Natl. Acad. Sci. USA* 88: 9828-9832.
- Moreau L, Charcosset A, Gallais A. 2004. Experimental evaluation of several cycles of marker-assisted selection in maize. *Euphytica* 137: 111–118.
- Plotsky Y, Cahaner A, Haberfeld A, Lavi U, Lamont SJ, Hillel J. 1993. DNA fingerprint bands applied to linkage analysis with quantitative trait loci in chickens. *Animal Genet.* 24: 105-110.
- Quarrie SA, Lazić-Jančić V, Kovačević D, Steed A, Pekić S. 1999. Bulk segregant analysis with molecular markers and its use for improving drought resistance in maize. *J. Exp. Bot.* 50:1299-1306.
- Sham P, Bader JS, Craig I, O'Donovan M, Owen M. 2002. DNA pooling: A tool for large-scale association studies. *Nat. Rev. Genet.* 3:862-871.
- Stuber CW, Moll RH, Goodman MM, Schaffer HE, Weir BS. 1980. Allozyme frequency changes associated with selection for increased grain yield in maize (*Zea mays*). *Genetics* 95:225–336.
- Stuber CW, Goodman MM, Moll RH. 1982. Improvement of yield and ear number resulting from selection at allozyme loci in a maize population. *Crop Sci.* 22:737–740.
- van Treuren R. 2001. Efficiency of reduced primer selectivity and bulked DNA analysis for the rapid detection of AFLP polymorphisms in a range of crop species. *Euphytica* 117:27-37.
- Villar M, Lefevre F, Bradshaw HD Jr, du-Cros ET. 1996. Molecular genetics of rust resistance in Poplars (*Melampsora larici-populina* Kleb/*Populus* sp.) by bulked segregant analysis in a 2 X 2 factorial mating design. *Genetics* 143:531-536.
- Wingbermuehle WJ, Gustus C, Smith KP. 2004. Exploiting selective genotyping to study genetic diversity of resistance to Fusarium head blight in barley. *Theor. Appl. Genet.* 109: 1160–1168.
- Xu Y, Crouch JH. 2008. Marker-assisted selection in plant breeding: from publications to practice. *Crop Sci.* 48: 391-407.
- Xu, Z., Zou, F., and Vision, T. (2005) Improving quantitative trait loci mapping resolution in experimental crosses by the use of genotypically selected samples. *Genetics* 170: 401–408.
- Yang HC, Liang YJ, Huang MC, Li LH, Lin CH, Wu JY, Chen YT, Fann CSJ. 2006. A genome-wide study of preferential amplification/hybridization in microarray-based pooled DNA experiments. *Nucl. Acids Res.* 34 (15) e106
- Zhang J, Xu Y, Wu X, Zhu L. 2002. A bentazon and sulfonylurea sensitive mutant: breeding, genetics and potential application in seed production of hybrid rice. *Theor. Appl. Genet.* 105:16-22.
- Zhang LP, Lin GY, Niño-Liu D, Foolad MR. (2003) Mapping QTLs conferring early blight (*Alternaria solani*) resistance in a *Lycopersicon esculentum* × *L. hirsutum* cross by selective genotyping. *Mol. Breed.* 12:3–19.