

# Discovery of SNPs in Soybean Genotypes Frequently Used as the Parents of Mapping Populations in the United States and Korea

K. VAN, E.-Y. HWANG, M. Y. KIM, H. J. PARK, S.-H. LEE, AND P. B. CREGAN

From the Department of Plant Science, Seoul National University, Seoul 151-921, Republic of Korea (Van, Hwang, Kim, and Lee); Biosafety Division, National Institute of Agricultural Biotechnology, Suwon 441-707, Republic of Korea (Park); and Soybean Genomics and Improvement Laboratory, USDA-ARS, Beltsville, MD 20705 (Hwang and Cregan).

Address correspondence to S.-H. Lee at the address above, or e-mail: sukhalee@snu.ac.kr.

## Abstract

Single nucleotide polymorphisms (SNPs) including insertion/deletions (indels) serve as useful and informative genetic markers. The availability of high-throughput and inexpensive SNP typing systems has increased interest in the development of SNP markers. After fragments of genes were amplified with primers derived from 110 soybean GenBank ESTs, sequencing data of PCR products from 15 soybean genotypes from Korea and the United States were analyzed by SeqScape software to find SNPs. Among 35 gene fragments with at least one SNP among the 15 genotypes, SNPs occurred at a frequency of 1 per 2,038 bp in 16,302 bp of coding sequence and 1 per 191 bp in 16,960 bp of noncoding regions. This corresponds to a nucleotide diversity ( $\theta$ ) of 0.00017 and 0.00186, respectively. Of the 97 SNPs discovered, 78 or 80.4% were present in the six North American soybean mapping parents. The addition of “Hwaeomputkong,” which originated from Japan, increased the number to 92, or 94.8% of the total number of SNPs present among the 15 genotypes. Thus, Hwaeomputkong and the six North American mapping parents provide a diverse set of soybean genotypes that can be successfully used for SNP discovery in coding DNA and closely associated introns and untranslated regions.

## Introduction

The development of DNA-based markers is important for selection and improvement of varieties and hybrids in plant breeding programs (Gupta et al. 2001; Kota et al. 2003). Single nucleotide polymorphisms (SNPs) including insertion/deletions (indels) can provide a rich source of useful molecular markers in genetic analysis. In the human genome, about 90% of sequence variants are SNPs. In some instances, these SNPs are actually the causal mutations of genetic diseases, such as cystic fibrosis (Kuppuswamy et al. 1991; Collins et al. 1998; Brookes 1999; Kwok and Gu 1999). SNPs are the most abundant type of sequence variation in plant genomes also (Cho et al. 1999). Much progress has been made in the discovery of sequence diversity in crop plants. The frequency of SNPs in maize (*Zea mays* ssp. *mays* L.) was reported as 1 SNP per 104 bp (Tenaillon et al. 2001). A total of 112 SNPs were found at 38 of 54 loci in barley (*Hordeum vulgare* L.) (Kanazin et al. 2002). Recently, a total of 280 SNPs were discovered among 25 diverse soybean (*Glycine max*

L. Merr.) genotypes in more than 76 kb of sequence of polymerase chain reaction (PCR) products amplified using primers designed for 116 genes and 27 nongenic regions (Zhu et al. 2003). Additionally, many SNP detection methods have been developed, such as denaturing high-performance liquid chromatography (DHPLC), oligonucleotide ligation, primer extension, DNA sequencing, PCR primer mismatch, pyrosequencing, and heteroduplex assays (Wu and Wallace 1989; Kuppuswamy et al. 1991; Ronaghi et al. 1998; Hoogendoorn et al. 1999; Pastinen et al. 2000; Wolford et al. 2000; Ye et al. 2001).

Because SNPs can be analyzed using high-throughput and inexpensive systems, they are useful for construction of high-density genetic maps as well as for genetic association studies (Cho et al. 1999; Picoult-Newberg et al. 1999; Nairz et al. 2002; Rafalski 2002; Kota et al. 2003). The relatively high level of linkage disequilibrium (LD) that would be anticipated in self-fertilizing plant species such as soybean (Zhu et al. 2003) may permit whole genome scans using SNPs for QTL discovery. In contrast, the lower levels of LD

in outcrossing species such as maize (*Zea mays* sp. *mays* L.) (Tenaillon et al. 2001) will require the use of the candidate gene approach to discover the specific gene(s) underlying phenotypic changes (Rafalski 2002).

Mutations in coding DNA sequences (cSNPs) may change amino acid sequences and affect gene function and could therefore be valuable as markers (Collins et al. 1998; Brookes 1999; Marth et al. 1999; Picoult-Newberg et al. 1999). Expressed sequence tag (EST) data serve as a useful source of DNA sequences in which SNPs can be discovered. The Soybean EST Project database contained more than 342,000 publicly available ESTs from 84 libraries ([www.ncbi.nlm.nih.gov/dbest/dbest\\_summary.html](http://www.ncbi.nlm.nih.gov/dbest/dbest_summary.html)) as of December 2004. This resource provides an excellent source for the development of gene-derived SNP markers.

This study reports results from the screening of 110 soybean EST-derived PCR primer pairs to identify and compare SNPs in nine cultivated soybean genotypes from Korea (Van et al. 2004) and six parents of North American mapping populations that were previously analyzed for their sequence diversity (Zhu et al. 2003). Thus, an important objective of this research was the identification of a subset of soybean genotypes that can maximize the discovery of SNPs in coding and noncoding perigenic DNA.

## Materials and Methods

### Plant Materials and Genomic DNA Extraction

A set of 15 soybean genotypes was used for SNP discovery. Nine Korean cultivated genotypes were included: Sinpaldalkong 2, SS2-2, Danbaekkong, Taekwangkong, Jinpumkong 2, Pureunkong, Daewonkong, Dongsan 163, and Hwaecomputkong (Van et al. 2004). These cultivated genotypes not only these possess interesting phenotypes (Van et al. 2004) but also are the parents of various recombinant inbred line mapping populations in Korea.

Six additional lines, including "Archer," PI 209332, "Peking," "Minsoy," Noir 1, and "Evans" (Table 1) that have been used as the parents of mapping populations in the United States (Mansur et al. 1996; Concibido et al. 1997; Mudge et al. 1997; Cregan et al. 1999; Orf et al. 1999; Qiu et al. 1999) were also included. The six North American genotypes had previously been identified by Zhu et al. (2003) as a good subset of genotypes for SNP discovery because the sequence analysis of these six genotypes identified 85% of the total SNPs and 93% of the common SNPs (frequency > 0.1) discovered in a larger set of 25 genotypes. Van et al. (2004) described method of genomic DNA extractions from the 15 homozygous soybean genotypes.

### Designing and Testing of PCR Primers

After a total of 110 soybean ESTs were selected from GenBank, designing primers were followed by Van et al. (2004). Lists of selected soybean ESTs and their primer information are available in Van et al. (2004). Each PCR primer set was used to amplify genomic DNA of Sinpaldalkong 2

**Table 1.** Soybean genotypes used for surveying SNPs

Genotypes	Pedigree
Sinpaldalkong 2 SS2-2	Dokyukong (Korea) × SS79168 (Korea) Developed by EMS mutagenesis from Sinpaldalkong 2
Danbaekkong	Dongsan 69 (Korea) × D76-8070 (USA)
Taekwangkong Jinpumkong 2	SS77011 (Korea) × Dongsan 53 (Korea) Selected from pure line of IT 167163 (Japan) after local adaptability test Chungsaeknamulkong (Local landrace, Korea) × L78-378 (USA)
Pureunkong	Suwon 133 (Korea) × Milyang 18 (Korea)
Daewonkong	Collected from Korea, local land race Renamed IT 179871 after introduced from Japan
Dongsan 163 Hwaecomputkong	Developed by transferring <i>Rps1-k</i> , ob- tained from Williams 82 and <i>Rps6</i> , obtained from PRX54-54 for resistance to phytophthora rot into BSR 101 by backcrossing. BC <sub>4</sub> F <sub>2</sub> plants were inter- crossed to obtain F <sub>1</sub> plants with both genes
Archer	Collected from Japan
PI 209332	Collected from China
Peking	Variety name assigned directly to or variety selected from PI 27890
Minsoy	Donated from France
Noir 1	Merit × Harosoy
Evans	

for testing. PCRs were performed in a 50 µl volume with 2 U *Taq* DNA polymerase (Vivagen, Sungnam, Korea) following the manufacturer's recommended protocols and cycling conditions. Gel electrophoresis on an ethidium bromide-stained 1.0% agarose gel confirmed the presence of amplified products. The primer sets that produced a single amplicon with Sinpaldalkong 2 genomic DNA were selected and used in identical amplification reactions with the other 14 soybean genotypes using the same conditions just described.

### Purification and Sequence Analysis of PCR Products

After PCR amplicons were purified using NucleoSpin Extract (Machery-Nagel, Düren, Germany), one of the primers used in the PCR amplification was used as the primer in the sequencing reaction. Sequence analysis was performed for all 15 cultivars using BigDye Terminator Cycle Sequencing (Applied Biosystems, Foster City, CA) as described by Kim et al. (2004). The sequencing reaction mixture was ethanol-precipitated and resuspended in 10 µl water, and an ABI 3700 sequencer (Applied Biosystems) was used for the sequence analysis.

### SNP Survey and Nucleotide Diversity ( $\theta$ )

With default conditions for "basecaller" and ending base, mixed-base settings, clear range methods, and filter settings,

**Table 2.** Summary of SNP analysis

Parameters	Korea	USA	Korea + USA
Number of gene fragments screened	110	110	110
Total length of sequenced amplicons (bp)	33,262	33,262	33,262
Number of SNPs (SNPs and indels)	66	78	97
Number of nucleotide substitutions	62	71	89
Transition/transversion ratio	2.05	2.25	2.12
Frequency of polymorphic sites per bp	1/504	1/426	1/343
Frequency of polymorphic sites per bp (coding)	1/3260	1/2038	1/2038
Frequency of polymorphic sites per bp (noncoding)	1/278	1/242	1/191
Number of indels	4	7	8
Overall indel frequency	1/8316	1/4752	1/4158
Frequency of indels per bp (coding)	—	—	—
Frequency of indels per bp (noncoding)	1/4240	1/2423	1/2120
Mean nucleotide diversity ( $\theta$ )	0.00070	0.00083	0.00103

ABI trace files were aligned and mutations were identified using ABI Prism SeqScape Software version 2.0 (Applied Biosystems), for detection of SNPs among the 15 soybean cultivars (Van et al. 2004).

Calculations of nucleotide diversity ( $\theta$ ) were followed by Halushka et al. (1999):

$$\theta = K / \sum_{i=2}^n (i-1)^{-1} L$$

where  $K$  is defined as the number of SNPs identified in  $n$  chromosomes or sequences with a total sequence length,  $L$ .

## Results and Discussion

Single nucleotide changes in coding regions can lead to alteration of amino acid sequences or early termination of translation and can therefore affect gene function (Brookes 1999; Collins et al. 1998; Marth et al. 1999; Picoult-Newberg et al. 1999). These functional SNPs would be valuable markers if the altered gene function or altered phenotype was of value for breeding purposes. However, even SNPs that do not alter the amino acid sequence but that are in or near a gene can be very useful as genetic markers. Thus, our SNP discovery research in soybean was focused on gene fragments amplified using PCR primers designed to EST sequences using 15 different genotypes (Table 1). These genotypes are the parents of soybean recombinant inbred line mapping populations that have been used in Korea or the United States.

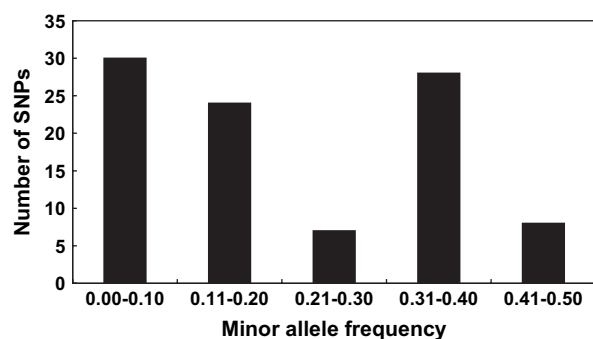
Among 110 randomly chosen soybean ESTs from GenBank for which primers were designed (Van et al. 2004), only 70% of them amplified a single PCR product and high-quality sequence data were obtained from only 60% of the 110 primer sets. This was likely the result of two or more amplicons of similar size but with slightly different sequences (Van et al. 2004).

Characterizations of SNPs discovered in all fifteen genotypes as well as in the subsets of mapping parents from Korea and the United States are presented in Table 2. A total of 16,302 bp of coding sequence, 7,372 bp of 5' UTR,

6,659 bp of intron and 2,929 bp of 3' UTR were surveyed. At least 1 SNP was discovered in 35 of the 66 gene fragments for which data were obtained. The analysis of the complete set of 15 genotypes indicated that that SNPs occurred at a frequency of 1 per 2,038 bp ( $\theta = 0.00017$ ) and 1 per 191 bp ( $\theta = 0.00186$ ), in coding and noncoding regions, respectively.

A total of 97 SNPs, including 8 indels, were discovered in the 15 genotypes, but the number of SNPs discovered was less within either the Korean or North American groups of mapping parents (Table 2). The SNP frequency was higher in noncoding regions in both the Korean and U.S. mapping parents. Among the Korean lines, 1 SNP occurred every 3,260 bp in coding sequence and every 278 bp in noncoding sequence. A total of 66 polymorphisms were discovered, and the overall frequency of SNPs was 1 every 504 bp (Van et al. 2004). Among the U.S. soybean mapping parents, frequencies of SNPs in both coding and noncoding regions were greater than among the Korean cultivars. The distribution of minor allele frequencies in all 15 genotypes is given in Figure 1. Of 97 SNPs, 67 SNPs (69%) could be classified as common, occurring at a frequency greater than 0.10. In contrast to the nucleotide diversity ( $\theta$ ) reported in maize (Tenaillon et al. 2001), the 15 cultivars had about a ninefold lower diversity ( $\theta = 0.00103$ ) in the 33,262 bp of sequenced amplicons. A similar result was reported by Zhu et al. (2003). As would be anticipated, the frequency of single base substitutions or indels was higher in the 15 cultivars versus within either the Korean or North American subpopulations. The Korean soybean lines had slightly lower nucleotide diversity than the six U.S. mapping parents.

In humans, about 2/3 of the SNP types are transitions (purine to purine or pyrimidine to pyrimidine) and about 1/3 are transversions (purine/pyrimidine) (Wang et al. 1998; Brookes 1999). Our study showed a similar ratio of transitions to transversions among the 15 genotypes. However, there was a relative greater proportion of transitions over transversions within the American soybean mapping parents. This result contrasts with Zhu et al. (2003), who reported essentially equal numbers of transitions and transversions. A summary of single and multiple nucleotide substitutions as



**Figure 1.** Distribution of minor allele frequencies at 97 SNP loci discovered in 110 gene fragments analyzed in 15 soybean genotypes.

well as single and multiple base indels is presented in Table 3. Three examples of triallelic SNPs were discovered, and a number of dinucleotide and trinucleotide SNPs were observed. One-, two-, and six-base indels were also detected (Table 3). A relatively small number of indels were observed, and these were all in noncoding DNA (Table 2).

Because the set of six diverse North American genotypes, Archer, PI 209332, Peking, Minsoy, Noir 1, and Evans, was identified by Zhu et al. (2003) as the minimum needed to maximize the discovery of sequence variation in the United States, each of the Korean lines was added to this set to determine if a significant increase in SNP discovery would result (Table 4). The number of SNPs discovered by adding either Pureunkong or Hwaecomputkong to the set of six U.S. mapping parents is presented in Table 4. Adding Pureunkong resulted in an increase in total SNPs discovered from 78 to 84 (86.6% of the total SNPs). The addition of Hwaecomputkong increased the number of SNPs discovered from 78 to 92 (94.8% of the total SNPs), suggesting that the addition of Hwaecomputkong as a seventh genotype would clearly increase the efficiency of SNP discovery in soybean.

There were five gene fragments in which Hwaecomputkong-specific SNPs were discovered. The length of the amplified regions (5' UTR, exon, intron, or 3' UTR), the sequenced length of each gene fragment, the position of the mutation and the allele present in Hwaecomputkong as well as in the six U.S. mapping parents is presented in Table 5. The allele present in two genotypes, PI 209332 and Minsoy, could not be determined in GenBank accession number AF327903 (functional candidate resistance protein, KR1) despite several sequencing attempts. Of the 32 gene fragments in which SNPs were discovered in this subset of 7 genotypes, 14 had a single SNP and multiple SNP loci were detected in the remainder (data not shown). Interestingly, AJ003246 (putatively encoding 2-hydroxydihydrodiazinone reductase) contained 12 SNPs and 1 indel in these seven mapping parents (Table 5). Eight SNP loci were discovered in the exon in these seven genotypes (data not shown). In the coding region, five synonymous changes were identified as well as three nonsynonymous changes each of which had a single base substitution in the first position. Boldface

**Table 3.** Single nucleotide substitution and indel frequencies for SNPs identified in 110 soybean genes

	Korea	USA	Korea + USA
<b>Transitions</b>			
C/T	19	23	29
G/A	20	22	26
<b>Transversions</b>			
C/G	2	2	4
A/T	4	3	7
C/A	5	4	4
T/G	8	9	11
Total	58	65	81
<b>Unusual substitutions</b>			
A/C/T	0	1	1
A/C/G	0	1	1
A/G/-	0	1	1
AA/GT	1	1	1
AC/GT	1	0	1
CC/TG	1	0	1
CT/TC	0	1	1
ATA/T-	1	0	0
AAA/ATA/T-	0	1	1
Total	4	6	8
<b>Single nucleotide indel</b>			
A/-	1	2	2
T/-	0	2	2
G/-	0	1	1
<b>Dinucleotide indel</b>			
TC/-	1	0	1
TG/-	1	1	1
<b>Hexanucleotide indel</b>			
TTATCC/-	1	1	1
Total	4	7	8

characters in Table 5 indicate the additional SNPs discovered by the inclusion of Hwaecomputkong as the seventh genotype. Among 13 SNP loci unique to Hwaecomputkong, 9 were discovered in just one GenBank accession (AJ003246).

**Table 4.** Total SNPs and percentage of total SNPs discovered in 15 soybean genotypes by the analysis of selected subsets of genotypes

Genotypes included subset	SNPs no.	Percent of SNPs
Archer, PI 209332	29	29.9
Archer, PI 209332, Peking	50	51.5
Archer, PI 209332, Peking, Minsoy	62	63.9
Archer, PI 209332, Peking, Minsoy, Noir 1,	76	78.5
Archer, PI 209332, Peking, Minsoy, Noir 1, Evans,	78	80.4
Archer, PI 209332, Peking, Minsoy, Noir 1, Evans, Pureunkong	84	86.6
Archer, PI 209332, Peking, Minsoy, Noir 1, Evans, Hwaecomputkong	92	94.8
All 15 genotypes	97	100.0

**Table 5.** Characterizations of the additional detected SNPs and indels after inclusion of Hwaeomputkong as the seventh genotype

GenBank accession number	Sequenced region	Sequenced length (bp)	Mutated position (bp)	Context	Amino acid replacement	Archer	PI 209332	Peking	Minsoy	Noir I	Evans	Hwaeomputkong
AB040543	exon, intron	480	172	GTGGG(C/T)AATGC		C	C	C	T	T	T	C
			187	CTTAG(A/T)GTCTG		T	T	T	T	T	T	<b>A<sup>a</sup></b>
			240	ATTGT(C/T)GATTT		C	C	C	T	T	T	C
			270	GAAGC(A/G)TATTA		G	G	G	A	A	A	A
AF327903	exon, intron	663	186	AGAGG(A/G)GACCA	Gly → Gly	G	N <sup>b</sup>	A	A	A	A	G
			386	ACTGG(A/T)AGTTT	Gly → Gly	T	N	T	T	T	T	<b>A</b>
			590	ATTTT(A/T)TGTTT		A	N	T	N	T	T	T
			631	TATCC(TTATCC/-)GCACT		—	N	—	N	—	TTATCC	—
			663	CAAGA(C/T)TGCAA		T	N	C	N	C	C	T
AJ003246	exon, intron	530	282	CCTTT(CC/TG)AAGTT		CC	CC	CC	CC	CC	CC	<b>TG</b>
			293	TGTCC(A/C)TGTGT		A	A	A	A	C	A	C
			302	GTGTG(C/T)CCTTC		C	C	C	C	C	C	<b>T</b>
			312	CAAAA(T/C)GATTA		T	T	T	T	T	T	<b>C</b>
			318	GATTA(C/T)GGATG		C	C	C	C	C	C	<b>T</b>
			324	GGATG(A/G)TTACG		A	A	A	A	A	A	<b>G</b>
			326	ATGAT(T/C)ACGAT		T	T	T	T	C	T	C
			359	GCGAC(G/T)TAGTC		G	G	G	G	T	G	G
			362	ACGTA(GT/AC)CTTGA		GT	GT	GT	GT	GT	GT	<b>AC</b>
			367	GTCTT(G/A)AATGA		G	G	G	G	G	G	<b>A</b>
			385	AGGAA(T/C)AACTT		T	T	T	T	C	T	T
			417	CAAAT(T/C)GATTC		T	T	T	T	T	T	<b>C</b>
			465	TTGTG(A/G/-)TCACA		A	A	A	—	G	A	G
J02746	5' UTR	555	192	TAATT(T/A)CCCAA		T	A	A	A	T	A	T
			306	GACCG(C/A)GTAGA		C	A	A	A	C	A	A
			444	TATGG(T/C)TATATT		T	T	T	T	T	T	<b>C</b>
			455	TCCTA(C/G)ACAGC		C	C	C	C	C	C	<b>G</b>
			516	TATCC(T/C)GAAAT		T	C	C	C	T	C	C
L01430	exon, 3' UTR	517	392	TTGTG(G/A)TATAG		A	A	A	A	A	A	<b>G</b>

<sup>a</sup> Boldface character indicates SNP was identified by addition of Hwaeomputkong.<sup>b</sup> N = not determined.



These SNPs unique to Hwaecomputkong suggested that it is genetically distinct from the U.S. mapping parents at six SNP fragments because one small fragment had 8 of 13 SNPs. Phenotypic characteristics of Hwaecomputkong (including early maturity, short plant height, and large seed size) as well as its use as a vegetable soybean also suggest that this genotype is very unique among the nine Korean mapping parents.

In summary, we identified a total of 97 SNPs from 110 soybean genes in all 15 soybean genotypes. These data indicated the presence of abundant sequence diversity in the soybean cultivars assayed. The North American mapping parents showed greater genetic diversity than the Korean mapping parents. The addition of Hwaecomputkong to the six parents of the North American mapping populations increased the number of SNPs discovered from 78 to 92. Thus, these seven genotypes could be helpful for discovery of cSNPs and gene-targeted map constructions. If these and other SNPs were determined to be the causal mutations associated with phenotypic traits, the development of these functional markers could lead to the errorless marker assisted selection in soybean breeding.

## Acknowledgments

This work was supported by a grant (code number CG3121) from Crop Functional Genomic Center of the 21st Century Frontier Research Program funded by the Ministry of Science and Technology and Rural Development Administration of the Republic of Korea. We thank National Instrumentation Center for Environmental Management at Seoul National University in Korea.

## References

- Brookes AJ, 1999. The essence of SNPs. *Gene* 234:177–186.
- Cho RJ, Mindrinos M, Richards DR, Sapolsky RJ, Anderson M, Drenkard E, Dewdney J, Reuber TL, Stammers M, Federspiel N, and others, 1999. Genome-wide mapping with biallelic markers in *Arabidopsis thaliana*. *Nat Genet* 23:203–207.
- Collins FS, Brooks LD, and Charkravarti A, 1998. A DNA polymorphism discovery resource for research on human genetic variation. *Genome Res* 8:1229–1231.
- Concibido VC, Lange DA, Denny RL, Orf JH, and Young ND, 1997. Genome mapping of soybean cyst nematode resistance genes in 'Peking', PI 90763 and PI 88788 using DNA markers. *Crop Sci* 37:258–264.
- Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N, VanToai TT, Lohnes DG, Chung J, and Specht JE, 1999. An integrated genetic linkage map of the soybean genome. *Crop Sci* 39:1464–1490.
- Gupta PK, Roy JK, and Prasad M, 2001. Single nucleotide polymorphisms: a new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. *Curr Sci* 80:524–535.
- Halushka MK, Fan JB, Bentley K, Hsie L, Shen N, Weder A, Cooper R, Lipshutz R, and Chakravarti A, 1999. Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. *Nat Genet* 22:239–247.
- Hoogendoorn B, Owen MJ, Oefner PJ, Williams N, Austin J, and O'Donovan MC, 1999. Genotyping single nucleotide polymorphisms by primer extension and high performance liquid chromatography. *Hum Genet* 104:89–93.
- Kanazin V, Talbert H, See D, DeCamp P, Nevo E, and Blake T, 2002. Discovery and assay of single-nucleotide polymorphisms in barley (*Hordeum vulgare*). *Plant Mol Biol* 48:529–537.
- Kim MY, Ha B-K, Jun T-H, Hwang E-Y, Van K, Kuk YI, and Lee S-H, 2004. Single nucleotide polymorphism discovery and linkage mapping of lipoxygenase-2 gene (*Lx2*) in soybean. *Euphytica* 135:169–177.
- Kota R, Rudd S, Facius A, Kolesov G, Thiel T, Zhang H, Stein N, Mayer K, and Graner A, 2003. Snipping polymorphisms from large EST collections in barley (*Hordeum vulgare* L.). *Mol Gen Genomics* 270:24–33.
- Kuppuswamy MN, Hoffman JW, Kasper CK, Spitzer SG, Groce SL, and Bajaj SP, 1991. Single nucleotide primer extension to detect genetic diseases: Experimental application to hemophilia B (factor IX) and cystic fibrosis genes. *Proc Natl Acad Sci USA* 88:1143–1147.
- Kwok P-Y and Gu Z, 1999. Single nucleotide polymorphism libraries: why and how are we building them?. *Mol Med Today* 5:538–543.
- Mansur LM, Orf JH, Chase K, Jarvik T, Cregan PB, and Lark KG, 1996. Genetic mapping of agronomic traits using recombinant inbred lines of soybean. *Crop Sci* 36:1327–1336.
- Marth GT, Korf I, Yandell MD, Yeh RT, Gu Z, Zakeri H, Stitzel NO, Hillier L, Kwok P-Y, and Gish WR, 1999. A general approach to single-nucleotide polymorphism discovery. *Nat Genet* 23:452–456.
- Mudge J, Cregan PB, Kenworthy JP, Kenworthy WJ, Orf JH, and Young ND, 1997. Two microsatellite markers that flank the major soybean cyst nematode resistance locus. *Crop Sci* 37:1611–1615.
- Nairz K, Stocker H, Schindelholz B, and Hafen E, 2002. High-resolution SNP mapping by denaturing HPLC. *Proc Natl Acad Sci USA* 99:10575–10580.
- Orf JH, Chase K, Jarvik T, Mansur LM, Cregan PB, Adler FR, and Lark KG, 1999. Genetics of soybean agronomic traits: I. Comparison of three related recombinant inbred populations. *Crop Sci* 39:1642–1651.
- Pastinen T, Raitio M, Lindroos K, Tainola P, Peltonen L, and Syvanen AC, 2000. A system for specific, high-throughput genotyping by allele-specific primer extension on microarrays. *Genet Res* 10:1031–1042.
- Picoult-Newberg L, Ideker TE, Pohl MG, Taylor SL, Donaldson MA, Nickerson DA, and Boyce-Jacino M, 1999. Mining SNPs from EST databases. *Genome Res* 9:167–174.
- Qiu BX, Arelli PR, and Slep DA, 1999. RFLP markers associated with soybean cyst nematode resistance and seed composition in a 'Peking' x 'Essex' population. *Theor Appl Genet* 98:356–364.
- Rafalski A, 2002. Applications of single nucleotide polymorphisms in crop genetics. *Curr Opin Plant Biol* 5:94–100.
- Ronaghi M, Uhlen M, and Nyren P, 1998. A sequencing method based on real-time pyrophosphate. *Science* 281:363–365.
- Tenaillon MI, Sawkins MC, Long AD, Gaut RL, and Doebley JF, 2001. Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. *mays* L.). *Proc Natl Acad Sci USA* 98:9161–9166.
- Van K, Hwang E-Y, Kim MY, Kim Y-H, Cho YI, Cregan PB, and Lee S-H, 2004. Discovery of single nucleotide polymorphisms in soybean using primers designed from ESTs. *Euphytica* 139:147–157.
- Wang DG, Fan JB, Siao CJ, Bero A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, and Spencer J, and others, 1998. A large-scale identification, mapping, and genotyping of single nucleotide polymorphisms in the human genome. *Science* 280:1077–1085.
- Wolford JK, Blunt D, Ballecer C, and Prochazka M, 2000. High-throughput SNP detection by using DNA pooling and denaturing high performance liquid chromatography (DHPLC). *Hum Genet* 107:483–487.

Wu DY and Wallace RB, 1989. The ligation amplification reaction (LAR)-amplification of specific DNA sequences using sequential rounds of template-dependent ligation. *Genomics* 4:460–569.

Ye F, Li M-S, Taylor JD, Nguyen Q, Colton HM, Casey WM, Wagner M, Weiner MP, and Chen J, 2001. Fluorescent microsphere-based readout technology for multiplexed human single nucleotide polymorphism analysis and bacterial identification. *Hum Mutat* 17: 305–316.

Zhu YL, Song QJ, Hyten DL, Van Tassell CP, Matukumalli LK, Grimm DR, Hyatt SM, Fickus EW, Young ND, and Cregan PB, 2003. Single-nucleotide polymorphisms in soybean. *Genetics* 163:1123–1134.

**Received January 4, 2005**

**Accepted March 31, 2005**

**Corresponding Editor:** J. Perry Gustafson