GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Assessment of genetic diversity in lentils (*Lens culinaris* Medik.) based on SNPs

R. Basheer-Salimia[1], B. Camilli[2], S. Scacchi[3], E. Noli[3] and M. Awad[1]

[1]Department of Plant Production and Protection, Faculty of Agriculture, Hebron University, Hebron, Palestine
[2]Overseas Onlus, Spilamberto, Modena, Italy
[3]Department of Agricultural Sciences, Alma Mater Studiorum University of Bologna, Bologna, Italy

Corresponding author: R. Basheer-Salimia
E-mail: rezqbasheer@hebron.edu

Genet. Mol. Res. 14 (2): 5870-5878 (2015)
Received June 18, 2014
Accepted December 2, 2014
Published June 1, 2015
DOI http://dx.doi.org/10.4238/2015.June.1.4

ABSTRACT. This study is the first attempt to establish an SNP database for the purpose of estimating the genetic diversity and relatedness of Palestinian lentil genotypes. A total of 14 lentil accessions (11 local, two supplied by ICARDA, and one introduced from Italy) were investigated. By sequencing two genes, lectin and lipid transfer protein 5 (*LTP5*), four SNPs were detected (three in the first and one in latter gene) with average frequencies of one SNP every 228 and 578 bp, respectively. In addition, in *LTP5* two single-nucleotide indels were observed in the non-coding part of the gene. Four haplotypes were identified in the lectin gene, three in *LTP5*. One lectin haplotype coincided with that present in GenBank belonging to two cultivated varieties, two were rather similar to this, whereas the last one turned out closer to the sequence of one wild lentil accession, indicating the existence of diversity in the Palestinian germplasm. These results, enhancing the available knowledge of lentil

genetic resources in Palestine, may contribute to their conservation and utilization in breeding projects.

Key words: Genetic fingerprinting; Lentils (*Lens culinaris* Medik.); SNP markers

INTRODUCTION

Lentil (*Lens culinaris* Medik.) is one of the oldest domesticated, annual, cool-season pulse crops. Based on archaeological findings, the first domestication of lentils occurred in Syria and Turkey approximately 8500 BC (Cubero 1981; Zohary and Hopf, 2000). Lentil is a diploid (2n = 2x = 14) and autogamous species with two main subspecies: ssp *culinaris*, the lentil crop, and ssp *orientalis*, its wild progenitor (Galasso et al., 2004). Economically, due to their high-quality protein and fiber content, lentil seeds are an important food source for humans, especially in developing countries. Additionally, lentil straw is a valued animal feed (Fratini et al., 2007). In Palestine, lentil is an essential crop in the local agricultural system, and it is grown mainly under rain-fed conditions with a total production of about 170 tons obtained from 260 ha (FAO, 2013; http://faostat.fao.org). On a worldwide scale, Palestinian production is rather low; also yield is below average and varying widely from year to year as well as from location to location.

The long history of lentil cultivation and the wide range of Palestinian agro-climatic conditions produce a complex picture of lentil genotypes that are either misidentified or called by different names in different areas. Furthermore, the severe decline of its cultivation increases the risk of losing useful genetic resources for any future breeding programs and variety development (Ladizinsky 1993; Piergiovanni 2000). Therefore, it is critical to discriminate among genotypes both for the conservation of lentil genetic resources as well as for crop improvement purposes.

During the last decades, different molecular markers, including RAPD and ISSRs (Sonnate and Pignone, 2001), AFLP (Sharma et al., 1996), SSR (Reddy et al., 2009) have been used to estimate genetic variation in lentil. Single nucleotide polymorphisms (SNPs) are the most recent markers used to detect genetic variation in many organisms (Mammadov et al., 2012) and are typically preferred because of the following reasons: 1) high stability of inheritance; 2) location in coding regions, resulting in amino acid changes in the encoded polypeptides; and 3) high abundance therefore high discriminatory power. Thus, currently they are widely used for the certification of cultivars and lines (Shirasawa et al., 2006).

Considering the importance of this crop, and the need to enhance the availability and utilization of its genetic resources, our study was aimed at establishing a SNP database for Palestinian lentil accessions. These data will allow a preliminary evaluation of the genetic diversity and the assessment of the relationships among accessions for conservation purposes.

MATERIAL AND METHODS

Plant materials

A total of fourteen lentil accessions were investigated in this study (Table 1). Eleven were collected from farms in different locations of the West-Bank and the Gaza Strip, considering the different agro-climatic zones of Palestine, whereas the remaining accessions were introduced (two supplied by ICARDA and one from Italy).

Table 1. List of lentil accessions analyzed in this study.

Code	Source
LCOH1	Hebron
LCOH2	Hebron
LCOJ1	Jenin
LCOJ2	Jenin
LGZ1	Gaza Strip
LH1	Hebron
L-I	Italy
LIC2	ICARDA
LIC4	ICARDA
LJ1	Jenin
LJ2	Jenin
LR1	Ramallah
LR2	Ramallah
LRA3	Ramallah

Genomic DNA extraction

For each accession, 20 seeds were germinated. After 20 days, leaf tissue samples were taken from the seedlings and ground into a fine powder in liquid nitrogen. Genomic DNA was extracted using the DNeasy Plant Mini Kit (Qiagen Inc.) following manufacturer protocols. The DNA quality and concentration were checked using 1.0% agarose gel electrophoresis and absorbance readings (Jenway 6305, UV-Visible Spectrophotometer, UK), respectively.

Primer design and polymerase chain reaction (PCR)

PCRs were performed using the locus-specific oligonucleotides designed on the published DNA sequences of lentil lectin (NCBI accession No. AJ318216) and lipid transfer protein 5 precursor (NCBI accession No. AY793556.1). PCR primer pairs were designed using the Primer-Blast software of NCBI (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and synthesized by Invitrogen (http://www.lifetechnologies.com/ps/en/home/brands/invitrogen.html). For the lectin gene primers were: forward 5'AGGTGAACTCAACAGAAACCACT3', reverse 5'TGAACTTCATGTGCTGCAA3'. For the lipid transfer protein 5 precursor gene primers were: forward 5'TGGCAAGAAGCATGAAGTTA3', reverse 5'CCAAAACAGTCTAGTACA ACAAGAGG3'.

PCR amplifications were performed in a total volume of 100 µL containing 1.0 X Taq buffer (Hylabs), 1.0 µM of each primer, 200 µM dNTPs (Fermentas), 2.0 mM $MgCl_2$ (Hylabs), 1.0 U *Taq* DNA polymerase (Hylabs), and 30 ng genomic DNA. HPLC water was added to the PCR mix to achieve the appropriate final volume. DNA was then amplified on a Peltier Thermal Cycler-200 (MJ Research. Inc, Watertown, MA, USA) with an initial denaturation of the DNA template for 10 min at 95°C, followed by 35 cycles of denaturation (95°C for 45 s), annealing (57.5°C for 45 s), and extension (72°C for 60 s), with a final extension step of 72°C for 10 min. In order to check the quality of amplification before the sequencing step, 10 µL PCR product of each accession together with 2 µL loading buffer were separated on a 2% (v/v) agarose gel. After electrophoresis, the gel was stained with ethidium bromide, and bands were visualized with UV light. **(Links ok)**

Sequencing and SNP discovery

PCR products were submitted for automated sequencing (Hy Laboratory Ltd., Rehovot, Israel). SNP discovery was conducted using Chromas Lite 2.1 (Technelysium Pty Ltd., Australia) and MEGA 4 software (Tamura et al., 2007) to detect and characterize the polymorphisms among the sequencing traces obtained.

Genetic diversity and data analyses

Statistical analyses of the polymorphisms were performed using the DnaSP software version 3.51 (Librado and Rozas 2009). Nucleotide diversity (π, i.e. the average number of nucleotide differences per site; Nei, 1987), the number of segregating sites (Watterson, 1975), and their sampling variances were calculated. Using Tajima's D test, implemented in DnaSP, π and θ values were compared to test the neutrality of molecular polymorphisms. Relationships among DNA sequences were investigated by calculating the proportion of nucleotide sites at which two sequences being compared were different (p-distance) and performing cluster analysis according to UPGMA using the MEGA4 software. For the purposes of haplotype analysis and tree calculation indels were treated as equivalent to single nucleotide differences. Branch support of the similarity trees was computed using a bootstrap test of 500 replications.

RESULTS AND DISCUSSION

The present study is the first report to estimate the genetic diversity and the relationships among lentil genotypes grown in Palestine using a SNP-based approach. Searching the NCBI database for *L. culinaris* ssp *culinaris* sequences provided single copies of lectin and lipid transfer protein 5 (*LTP5*) genes. For the latter, the complete coding sequence including one intron was retrieved, whereas for the first only part of the coding sequence was obtained. The use of the lectin gene in our study is related to its importance as one of the major storage proteins in legume seeds (Vitale and Bollini, 1995). It has been suggested that lectins play a role in seed maturation, cell-wall assembly, defense mechanisms, and rhizobial nodulation of legume roots. On the other hand, the *LTP5* precursor transfers phospholipids as well as galactolipids across plant membranes. Further, it may play a role in wax or cutin deposition in the cell walls of expanding epidermal cells and certain secretory tissues (http://www.uniprot.org/uniprot/A0AT31).

Amplifications of lectin and *LTP5* genes conducted with the designed primers produced amplified fragments of 685 and 615 bp, respectively. However, in order to avoid the disturbed parts of the sequences close to the annealing sites of primers the lengths of the fragments analyzed were reduced to 672 bp for lectin and 578 bp *LTP5*.

In total, for all examined lentil genotypes, four SNPs were detected, three in the lectin gene and one in *LTP5* (Tables 2 and 3), yielding a frequency of 1 SNP per 224 bp for the first gene and 1 SNP per 578 bp for the second one. The level of polymorphism found here for lectin (4.5 SNP/kb) was comparable to that found in other self-pollinating plant species such as rice (4.2) and barley (6.3), but on the other hand, the lower frequency of nucleotide changes detected in the *LTP5* sequence (1.7 SNP/kb) was similar to that reported in other legumes such as pea (1.9), *Medicago truncatula* (2.0), and soybean (2.1) (Leonforte et al., 2013).

Table 2. SNPs located in the lectin gene.

Accession	Haplotype	Allele and positions			Aminoacids[1] and positions		
		51	248	594	17	83	198
LJ1	Hap_1	A	T	G	K	M	L
LR2		A	T	G	K	M	L
LH1		A	T	G	K	M	L
LI		A	T	G	K	M	L
LGZ1		A	T	G	K	M	L
LCOH1	Hap_2	A	T	T	K	M	F
LR1		A	T	T	K	M	F
LJ2		A	T	T	K	M	F
LIC2		A	T	T	K	M	F
LCOJ2	Hap_3	C	T	T	N	M	F
LCOH2	Hap_4	C	C	G	N	T	L
LCOJ1		C	C	G	N	T	L
LRA3		C	C	G	N	T	L
LIC4		C	C	G	N	T	L
Change of AA class[2]					from B-P to N-P	from N-NP to N-P	No change

[1]F = phenylalanine; K = lysine; L = leucine; M = methionine; N = asparagine; T = threonine. [2]B-P: Basic-polar; N-P: neutral-polar; N-NP: neutral-non polar.

Table 3. DNA polymorphisms located in *LTP5*.

Accession	Haplotypes	SNP		INDELs	
		Alleles and positions	Aminoacids[1] and positions	Alleles and positions[2]	
		160	54	326	365
LJ1	Hap_1	C	R	T	T
LR2		C	R	T	T
LH1		C	R	T	T
LI		C	R	T	T
LGZ1		C	R	T	T
LCOJ1		C	R	T	T
LCOH2	Hap_2	C	R	-	-
LRA3		C	R	-	-
LIC4		C	R	-	-
LCOJ2		C	R	-	-
LCOH1	Hap_3	T	C	-	-
LR1		T	C	-	-
LJ2		T	C	-	-
LIC2		T	C	-	-
Change of AA class[3]			from B-P to N-SP		

Allele positions refer to the whole sequence. [1]R = arginine; C = cysteine. [2]Polymorphism detected in non-coding regions. [3]B-P: Basic-polar; N-SP: neutral- slightly polar.

Of the four SNPs detected, two were transversions and two were transitions. All of them were exonic, being localized inside protein coding regions, and were non-synonymous, resulting in amino acid substitutions.

Two single-nucleotide indels were also observed in the intronic sequence of *LTP5*, whereas none was detected in the coding region of both genes.

A detailed description of the nucleotide diversity of the two genes among the genotypes under study is presented in Table 4.

Table 4. Nucleotide diversity and neutrality test of the lectin and *LTP5* sequences analyzed in the 14 accessions.

Locus	Region and lenght (bp)	Type of polymorphism	Polymorphic sites (N)	Haplotypes (N)	H_d	k	π	θ	Tajima's D test
Lectin	Coding 672	SNP	3	4	0.758	1.429	0.00213	0.00140	n.s.
		INDEL	0	1	0	0	–		
LTP5	Coding 314	SNP	1	2	0.440	0.440	0.00140	0.00100	n.s.
		INDEL	0	1	0	0	–		
	Non-coding 265	SNP	0	1	0	0	–		
		INDEL	2	2	0.527	1.055	0.00398		

H_d = Haplotype diversity, k = average number of nucleotide differences, π = nucleotide diversity, θ = number of segregating site, n.s.= not significant

The π value in the coding region of the two genes determined in the collection of lentil accessions here considered was higher for lectin (0.00213) than for *LTP5* (0.00140). In particular, the π value of the first gene was higher and the second very similar compared to that found (0.00134) by Alo et al. (2011) in cultivated lentils as the average of coding regions of 22 genes. Considering the coding regions of other leguminous species, the level of polymorphism here detected for both genes was higher than that reported as the average of 90 genes in soybean (0.00054; Zhu et al., 2003). Moreover, the polymorphism reported for the actin gene in faba bean (0.00250; Basheer-Salimia et al., 2014) was comparable to that found here for lectin gene, but higher than that revealed for *LTP5*. The low level of π in the second case was somehow unexpected since in sunflower the *LTP* gene had been previously reported as highly polymorphic (Giordani et al., 2011) and therefore it was chosen in this study as a candidate for cultivar discrimination.

On the other hand, regarding *LTP5* the two indels detected in the intron determined, as expected, a higher level of polymorphism compared to the coding region.

The results of the Tajima D test indicate that the polymorphisms detected in both genes are selectively neutral.

The trees based on the nucleotide polymophisms detected show four haplotypes for the lectin gene (Figure1) and three haplotypes for *LTP5* (Figure 2). Haplotype 4 of the lectin gene, differing from all other haplotypes at nucleotide position 248, is the most supported branch in the dendrogram. Similarly haplotype 1 of the *LTP5*, differing from the two other haplotypes for two single nucleotide insertions at positions 326 and 365, is the most supported branch in the dendrogram.

When considered together, the four SNPs and the two indels separate the 14 accession into five groups (Figure 3). It can be pointed out that only genotypes LCOJ1 and LCOJ2 showed a unique haplotype, whereas the remaining accessions fell into groups. Haplotype 2 of the lectin gene coincided with haplotype 3 of *LTP5*, resulting in a higher bootstrap value (87) in the dendrogram based on the two genes compared to the separate analysis (42 and 63, for lectin and *LTP5* genes, respectively). Also the reliability of the other branches was improved by the combined analysis, although they did not match exactly the previous groupings.

This classification, however, did not reflect the position of the collection sites of the accessions, nor there was information available regarding their relationships that could explain the similarities observed. The fact that accessions of different provenance share the same haplotype may actually reflect the common practice of seed exchange among farmers, also considering the relatively limited collection area.

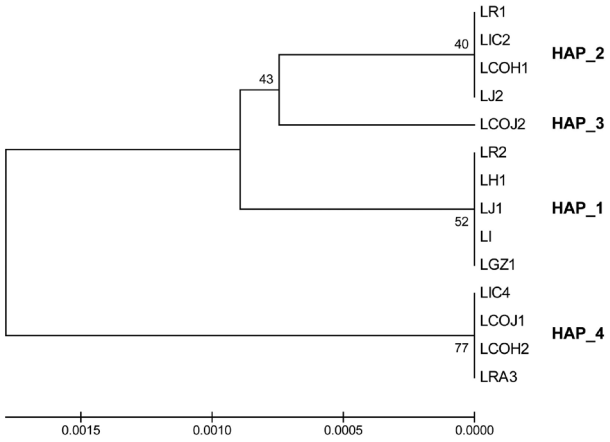

Figure 1. UPGMA tree of the 14 lentil accessions based on the SNP markers in the lectin gene.

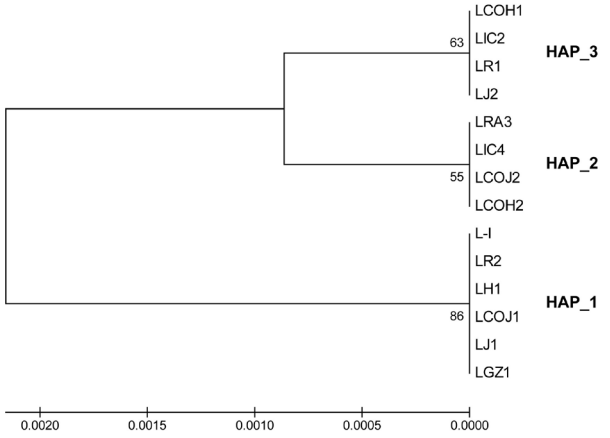

Figure 2. UPGMA tree of the 14 lentil accessions based on the SNP and indel markers in the *LTP5* gene.

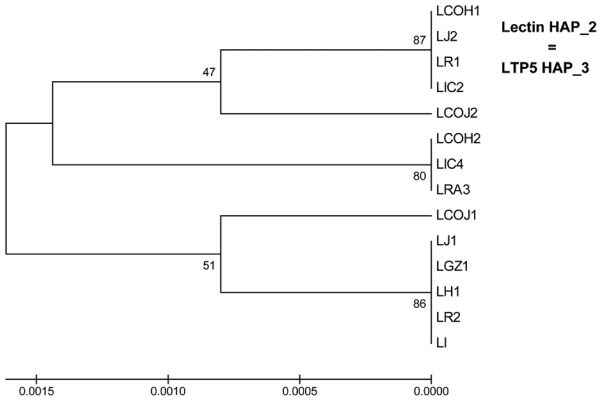

Figure 3. UPGMA tree based on the SNP and indel markers discovered in both lectin and *LTP5* genes.

Comparing the sequences of the lectin gene obtained in this study with those present in GenBank (Galasso et al., 2004), some Palestinian accessions shared the same haplotype (Hap_1) of the two *L. culinaris* ssp *culinaris* cultivars ('Eston' and 'Laird') described by those Authors, whereas the accessions belonging to Hap_4 turned out slightly more similar to the wild progenitor ssp *orientalis* (Figure 4).

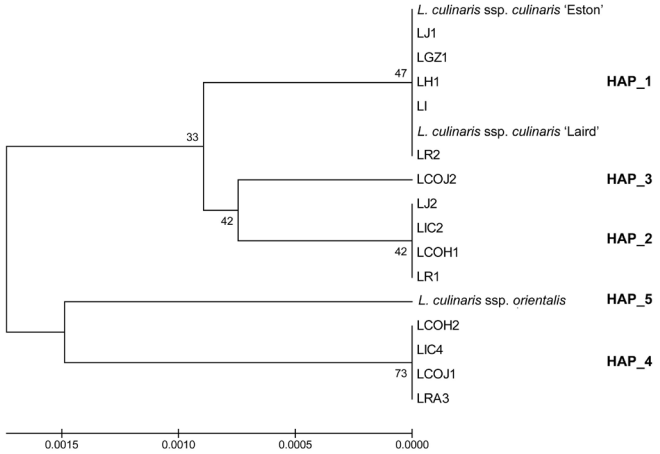

Figure 4. UPGMA tree of the 14 Palestinian lentil accessions together with two ssp *culinaris* and one spp *orientalis* accessions of different origin based on the SNP markers in the lectin gene.

The analysis of the lectin and *LTP5* genes did not allow distinguishing down to the level of single accessions (except in two cases), but just between several haplotypes; therefore more genes should be investigated to reach an unambiguous identification of the accessions. However, these results seem to indicate the existence of some diversity in Palestinian lentil germplasm and provide a preliminary tool for the description of the relationships between accessions.

CONCLUSIONS

In conclusion, the Palestinian lentil germplasm here considered shows some genetic variation for the lectin and *LTP5* precursor genes. This study provides preliminary information on the genetic diversity of this germplasm: as expected the analysis of the two genes did not allow to discriminate at the level of single accessions, hence additional genes need to be investigated to obtain a better picture of the genetic relationships among them and to achieve the goal of their unambiguous identification. Once an appropriate set of SNPs were identified across Palestinian lentil accessions, these polymorphisms could allow the development of cost-effective a robust genotyping assays for germplasm characterization and utilization.

ACKNLOWEDGMENTS

Research made possible thanks to the support of the Italian Ministry of Foreign Affairs (MAE) and Regione Emilia Romagna in the frame of the "Project to reduce the poverty rate in the Palestinian rural areas through improving biodiversity and local organic farming

9582/OVERSEAS/TOC" that was promoted by Overseas onlus and ACS. The authors would also like to express their gratitude to the Agriculture Development Association-PARC and the technical staff of the project.

REFERENCES

Alo F, Furman BJ, Akhunov E, Dvorak J, et al. (2011). Leveraging genomic resources of model species for the assessment of diversity and phylogeny in wild and domesticated lentil. *J. Hered*. 102: 315-329.

Basheer-Salimia R, Camilli B, Scacchi S, Noli E, et al. (2014) Genetic Diversity among Palestinian Faba Bean (*Vicia faba* L.) Ecotypes based on single nucleotide polymorphisms. *Europ. J. Hort. Sci*. 79: 300-305.

Cubero JI (1981). Taxonomy, distribution and evolution of the lentil and its wild relatives. In: Genetic resources and their exploitation chickpeas, fava beans and lentils (Witcombe J and Erskine W, eds.). M. Nijhoff and W. Junk Publishers, Boston, 187-204.

Fratini R, DuránY, García P and Pérez de la Vega M (2007). Identification of quantitative trait loci (QTL) for plant structure, growth habit and yield in lentil. *Spanish J. Agric. Res.* 5: 348-356.

Galasso L, Lioi C, Lanave R, Bollini F et al. (2004). Identification and isolation of lectin nucleotide sequences and species relationships in the genus *Lens* (Miller). *Theor. Appl. Genet*. 108: 1098-1102.

Giordani T, Buti M, Natali L, Cattonaro F, et al. (2011) An analysis of sequence variability in eight genes putatively involved in drought response in sunflower (*Helianthus annuus* L.). *Theor. Appl. Genet.* 122: 1039-1049.

Ladizinsky G (1993). Wild lentils. *Crit. Rev. Plant Sci.* 12: 169-184.

Leonforte A, Sudheesh S, Cogan NOI, Salisbury PA et al. (2013). SNP marker discovery, linkage map construction and identification of QTLs for enhanced salinity tolerance in field pea (*Pisum sativum* L.). *BMC Plant Biology* 13: 161.

Librado P and Rozas J (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452.

Mammadov J, Aggarwal R, Buyyarapu R and Kumpatla S (2012) SNP markers and their impact on plant breeding. *Int. J. Plant Genomics* 12: 728398.

Nei M (1987). Molecular evolutionary genetics. Columbia University Press, New York.

Piergiovanni AR (2000). The evolution of lentil (*Lens culinaris* Medik.) cultivation in Italy and its effects on the survival of autochthonous populations. *Genet. Resour. Crop Evol.* 47: 305-314.

Reddy MRK, Rathour R, Kumar N, Katoch P et al. (2009). Cross-genera legume SSR markers for analysis of genetic diversity in *Lens* species. *Plant Breed.* 129: 514-518.

Sharma SK, Knox MR and Ellis THN (1996) AFLP analysis of the diversity and phylogeny of *Lens* and its comparison with RAPD analysis. *Theor. Appl. Genet.* 93: 751-758.

Shirasawa K, Shiokai S, Yamaguchi M, Kishitani S, et al. (2006). Dot-blot-SNP analysis for practical plant breeding and cultivar identification in rice. *Theor. Appl. Genet.* 113: 147-155.

Sonnante G and Pignone D (2001). Assessment of genetic variation in a collection of lentil using molecular tools. *Euphytica* 120: 301-307.

Tamura K, Dudley J, Nei M and Kumar S (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol*. 24: 1596-1599.

Vitale A and Bollini R (1995). Legume Storage Proteins. In: Seed Development and Germination (Kigel J and Galili G, eds.). Marcel Dekker, New York, 73-102.

Watterson GA (1975). On the number of segregating sites in genetical models without recombination. *Theor. Popul. Biol.* 7: 256-276.

Zhu YL, Song QL, Hyten DL and Van Tassell CP (2003). Single-nucleotide polymorphisms in soybean. *Genetics* 163: 1123-1134.

Zohary D and Hopf M (2000). The Domestication of Plants in the Old World. 3rd edn. Oxford University Press, Oxford.

