Theor Appl Genet (2012) 125:1069–1085
DOI 10.1007/s00122-012-1896-5

ORIGINAL PAPER

Nucleotide diversity patterns at the drought-related *DREB2* encoding genes in wild and cultivated common bean (*Phaseolus vulgaris* L.)

Andrés J. Cortés · Dominique This · Carolina Chavarro · Santiago Madriñán · Matthew W. Blair

Received: 1 June 2011 / Accepted: 11 May 2012 / Published online: 8 July 2012

Abstract Common beans are an important food legume faced with a series of abiotic stresses the most severe of which is drought. The crop is interesting as a model for the analysis of gene phylogenies due to its domestication process, race structure, and origins in a group of wild common beans found along the South American Andes and the region of Mesoamerica. Meanwhile, the DREB2 transcription factors have been implicated in controlling non-ABA dependent responses to drought stress. With this in mind our objective was to study in depth the genetic diversity for two DREB2 genes as possible candidates for association with drought tolerance through a gene phylogenetic analysis. In this genetic diversity assessment, we analyzed nucleotide diversity at the two candidate genes *Dreb2A* and *Dreb2B*, in partial core collections of 104 wild and 297 cultivated common beans with a total of 401 common bean genotypes from world-wide germplasm analyzed. Our wild population sample covered a range of semi-mesic to very dry habitats, while our cultivated samples presented a wide spectrum of low to high drought tolerance. Both genes showed very different patterns of nucleotide variation. *Dreb2B* exhibited very low nucleotide diversity relative to neutral reference loci previously surveyed in these populations. This suggests that strong purifying selection has been acting on this gene. In contrast, *Dreb2A* exhibited higher levels of nucleotide diversity, which is indicative of adaptive selection and population expansion. These patterns were more distinct in wild compared to cultivated common beans. These approximations suggested the importance of *Dreb2* genes in the context of drought tolerance, and constitute the first steps towards an association study between genetic polymorphism of this gene family and variation in drought tolerance traits. We discuss the utility of allele mining in the DREB gene family for the discovery of new drought tolerance traits from wild common bean.

Communicated by H. Becker.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-012-1896-5) contains supplementary material, which is available to authorized users.

A. J. Cortés
Evolutionary Biology Centre, Uppsala University, 751 05 Uppsala, Sweden

D. This · C. Chavarro · M. W. Blair
ADOC Project, Generation Challenge Program, Montpellier, France

D. This
CIRAD, UMR Développement et Amélioration des Plantes, TA-A 96/03, 34398 Montpellier, France

S. Madriñán
Laboratorio de Botánica y Sistemática, Universidad de los Andes, Apartado Aéreo, 4976 Bogotá, Colombia

M. W. Blair (✉)
Department of Plant Breeding and Genetics, Cornell University, 242 Emerson Hall, Ithaca, NY 14853, USA
e-mail: mwb1@cornell.edu

Introduction

Common bean (*Phaseolus vulgaris* L.) is a key source of nutrients and dietary protein for over 500 million people in Latin America and Africa and more than 4.5 out of 23 million hectares are grown in zones where drought is severe, such as in northeastern Brazil, coastal Peru, the central and northern highlands of Mexico, and in Eastern and Southern Africa (Broughton et al. 2003). Therefore, increasing drought tolerance in common bean commercial

varieties and landraces is highly desirable. A considerable reservoir for this task may be available in the wild and cultivated collections of common bean, as can be suggested by their high genetic diversity and phenotypic variability (Gepts et al. 2008).

Drought tolerance is a quantitative trait with genetic, epigenetic, and environmental components modulated by a set of characterized and un-characterized transcription factors. Some of these are regulators of the ABA response mechanism, while others are involved in ABA-independent pathways (Bartels and Sunkar 2005; Taiz and Zeiger 2006). In particular, drought-responsive element binding (DREB) protein encoding *Dreb* genes are plant-specific, stress-regulated transcription factors which belong to the AP2/EREBP family and which are in the ABA-independent pathway for drought stress response (Nayak et al. 2009). These transcription factors interact with DRE elements or promoters found near a large number of genes involved in adaptation to drought (Kizis et al. 2001; Nayak et al. 2009; Riechmann and Meyerowitz 1998). Expression analysis has demonstrated their explicit role in conferring increased drought, cold, and salt tolerance in the cereals, barley, rice, and wheat, as well as other plants such as chrysanthemum, *Brassica*, *Arabidopsis*, and *Aloe vera* (Chen et al. 2008; Morran et al. 2011; Wang et al. 2008; Wang and He 2007; Yang et al. 2009; Zhao et al. 2007; Zhuang et al. 2010; Kim et al. 2012). Moreover, a study by Nayak et al. (2009) showed that the extent of nucleotide diversity in *Dreb2A* genes in rice, barley, sorghum, common bean, and chick-pea provided some evidence of adaptive variation.

Common bean is a good model to study *Dreb* genes because of its rich evolutionary history and multiple domestication process. In particular, wild bean are thought to have diversified in South and Central America from an original range in Ecuador and northern Peru, after which domestication in the southern and northern ends of each region gave origin to Andean and Mesoamerican domesticates, respectively (Gepts 1998; Gepts and Debouck 1991; Gepts et al. 1986). Additional structure within each of these gene pools has been described: with the races Nueva Granada, Peru, and Chile identifiable within the Andean genepool (Benchimol et al. 2007; Blair et al. 2007, 2009; Kwak and Gepts 2009; Paredes et al. 2010), and the races Mesoamerica, Durango–Jalisco, and Guatemala observable within the Mesoamerican genepool (Blair et al. 2009; Diaz and Blair 2006). Both genepools followed somewhat parallel pathways of dissemination through the world, generating new secondary centers of diversity in Africa and Asia (Asfaw et al. 2009; Zhang et al. 2008).

The cultivated genepool structure contrasts with the population structure obtained for wild common bean in which four main clusters are seen: the Colombo–Mesoamerican, the Mexican, the Andean, and the Peruvian–Ecuadorian (Kwak and Gepts 2009; Payró et al. 2005; Rossi et al. 2009). Introgression between gene pools and between cultivated and wild genotypes has been a historical, long-term, and re-iterative process (Blair et al. 2006a, b, 2007; Papa and Gepts 2003). While many cultivated accessions of the core landrace collection described by Tohme et al. (1995) have been evaluated for drought tolerance, smaller core collection of wild beans (Tohme 1996) has not been evaluated for drought tolerance given the mostly climbing bean nature of wild common beans (Blair et al. 2012). However, many accessions of wild beans are from dryland areas or seasonally wet and dry forest margins.

In terms of agro-ecology of the cultivated races, the group Durango–Jalisco is the only one of the groups of races that has high drought tolerance, with part of this group distributed in semi-arid areas of Mexico (Diaz and Blair 2006; Payró et al. 2005); race Chile has adaptation to relative drier areas as well, but is only found in the southern Andes (Blair et al. 2007; Becerra et al. 2010). Races Mesoamerica and Guatemala, or Nueva Granada and Peru occupy low to mid altitude or highland regions, respectively, of Latin America and generally have higher water requirements for production. Although cultivars from the Durango–Jalisco complex have important levels of drought tolerance, one may expect to find higher levels in certain wild germplasm (Singh 2005).

Nucleotide diversity is a powerful tool for studying reference collections of cultivated and wild genotypes that allow population genetic tests based on the departure from the neutral equilibrium models to identify the diverse selective modes that shaped the evolution of specific genes chosen for analysis (Wakeley 2008). The particular role of speciation, duplication, lineage sorting, sub-functionalization, and ecological constrains on gene evolution can be inferred. Furthermore, population structure and SNP diversity are the principal activities which must be realized in any phylogenetic study of a gene family (Rafalski 2010) and help to determine evolutionary associations between population subgroups and gene diversity. The objectives of this research, therefore, were to (1) evaluate the allele diversity of two DREB-encoding genes (*2A* and *2B*) in wild and cultivated common beans, (2) determine the extent of haplotype diversity; (3) evaluate the correlation of alleles with genepool origins, and (4) measure association with geographic origin and drought tolerance. We also evaluated whether the genetic diversity at these candidate genes was dissimilar between wild and cultivated beans, and if their patterns of nucleotide variation were determined by adaptation to hydrological environments or by evolutionary inertia. Our ultimate goal was to define whether DREB genes may play a role in common bean drought adaptation and whether specific DREB may be useful for marker-assisted breeding of the crop.

Materials and methods

Plant material

A total of 401 accessions (104 wild and 297 cultivated) from the FAO germplasm collection were used in this study (Supplemental Table 1). Reference collections were selected to be representative samples of the genepools and races and to be a subset of the core collection for cultivated and wild beans (Tohme et al. 1995). Control genotypes included the Andean genotypes Calima/G4494 and Chaucha Chuga/G19833, as well as the Mesoamerican genotypes ICA Pijao/G5773 and Dorado/DOR364 (with common name and germplasm entry or advanced line name listed in each case). Seed samples for wild accessions were provided by the Genetic Resource Unit (http://isa.ciat.cgiar.org/urg/main.do).

DNA extraction and PCR amplification

Total DNA was extracted with the method of (Afanador et al. 1993). In the PCR reactions, a touchdown profile was used for *Dreb2A*, with a hot start of 95 °C for 5 min, a denaturing temperature of 72 °C for 45 s, an initial annealing temperature of 60 °C for 45 s, and an extension temperature of 72 °C for 1.5 min. Amplification conditions for *Dreb2B* used thermocycling conditions of 95 °C hot start for 5 min, followed by 35 cycles of 95 °C denaturation for 45 s, 48 °C annealing for 45 s, and 72 °C extension for 1.5 min. Annealing temperature dropped 1 °C per cycle for seven cycles, followed by 28 cycles at 53 °C. An extension period of 5 min at 72 °C was used as post-thermocycling for both genes. The PCR reactions were carried out in a 25-µl final volume containing 65 ng of genomic, 1× PCR buffer (10 mM of Tris–HCl pH 8.8, 50 mM of KCl, 0.1 % of TritonX-100); different amounts of each of the forward and reverse primers, namely 0.3 µM for *Dreb2B* and 0.4 µM for *Dreb2A* as shown in Table 1 along with additional $MgCl_2$; 2.5 mM (for *Dreb2A*) or 3 mM (for *Dreb2B*) of and total dNTPs; 0.4 mM (for *Dreb2A*) or 0.5 mM (for *Dreb2B*) so as to ensure good PCR amplification strength. Finally, 1.5 U of Taq Polymerase (Fermentas) was used for the amplification of both genes.

Product clean-up and sequencing

All PCR products were electrophoresed through 1.5 % agarose-Tris-Borate-EDTA gels containing SYBR-Green and they were purified using Exo-Sap clean-up reactions. These PCR products were used as templates for subsequent Sanger sequencing reactions, using BigDye Terminator v3.1 Cycle Sequencing Kit. The samples were run on an ABI 3730 automated sequencers at the Cornell University (Ithaca, New York) Biotechnology Resource Center for the wild collection and at the GeneScope (Paris, France) sequencing facility for the cultivated accessions. Four control genotypes were sequenced in both facilities. Base-pair calls, quality score assignment, and construction of contigs were carried out using Sequencher4.7 (GeneCodesCorp. Ann Arbor).

Gene characterization, domain detection, and protein alignment

Coding regions, UTRs, reading frames, and conserved domains were determined through blastx of the new sequences against the non-redundant (nr) protein database with a gap penalty of 11, an extension penalty of 1, and a BLOSUM62 matrix available from http://blast.ncbi.nlm.nih.gov. The AP2 domain region was confirmed using the Pfam website (http://pfam.sanger.ac.uk/). Nucleotide and protein alignments, as well as Neighbor-Joining trees, were constructed using orthologous and paralogous genes to verify conserved regions. Nucleotide alignments were carried out with MUSCLE algorithm (Edgar 2004) on Geneious 4.0 software (Biomatters Ltd.).

Polymorphism and neutrality at Dreb2 genes

Levels of genetic diversity within domesticated and wild common bean were quantified with measures of nucleotide diversity based on the number of segregating sites (θ_W) (Watterson 1975) and based on the average number of nucleotide differences per site between sequences (π) (Nei 1987) using the software program DnaSP 5.10 (Rozas et al. 2003). The number of haplotypes and the haplotype diversity (Hd) were calculated with the same software based on the haplotype reconstruction that was carried out with PHASE (Stephens and Donnelly 2003). Meanwhile, tests for selection were performed to estimate whether the DREB genes followed the Wright–Fisher model of neutral evolution in each subpopulation. Tajima's D (Tajima 1989) tests were carried out with DnaSP using 5,000 coalescent simulations (Wakeley 2008) for each division of wild versus cultivated beans, Andean versus Mesoamerican genepools, and for the races or other sub-populations found in each group. Moreover, we also followed a non-model-based test for selection in which the nucleotide diversity of Nei was contrasted against an estimated, nearly-neutral, distribution (evolutionary background) using gene base-SNP markers (Cortés et al. 2011). In addition, linkage disequilibrium analysis was conducted with Tassel 2.1 software (Bradbury et al. 2007).

Haplotype analysis

On the other hand, median joining haplotype networks were built using Network 4.5.1 (Bandelt et al. 1999) and

Table 1 Primers used for PCR amplifications of candidate genes in *P. vulgaris* in the cultivated and wild collections and annotation of sequences obtained after the diversity analysis

Gene name	*Dreb2A*	*Dreb2B*
EST sequence	CV535836	BQ481823
Primer name forward	PvDreb2a_CV53_ADOC_F	PvDreb2b_BQ48_ADOC_F
Sequence primer forward (5′–3′)	CTAATTCTGCATCTCCCTCAGGTC	TCTCCTTCAGCTATGAGTCC
Primer name reverse	PvDreb2a_CV53_ADOC_R	PvDreb2b_BQ48_ADOC_R
Sequence primer reverse (5′–3′)	CAGCTCAGCAGCAGCGTCTACT	AGAGGGGAGAGGCTTGTAG
Ta (°C)	60–53 (touchdown profile)	48
Source	Nayak et al. (2009)	CIAT / LCGF
Blastn (nucleotide db)	cDNA, clone: GMFL01-11-B11 (*G. max*, AK285532.1, 5E−121)	cDNA, clone: GMFL01-14-P08 (*G. max*, AK244839.1, 5E−121)
Blastn (EST database)	NOD_221_E03 Nodule EST library (*P. vulgaris*, CV535836.1, 1E−168)	PV_GEa0130c02r Leaf EST library (*P. vulgaris*, BQ481823.1, 1E−147)
Blastx	DREB2 (*G. max*, 7E−57, AAQ57226.1)	AP2/ERF transcription factor (*P. trichocarpa*, XP_002310574.1, 1E−31)

Neighbor Joining trees were constructed using 1,000 bootstrap replicates for node support carried out with the software program Mega4 (Tamura et al. 2007). Network trees accounted for population subdivision (*K* values) as accessed by Blair et al. (2009). Furthermore, it considered drought tolerance estimated previously in cultivated and wild accessions using field trials and ecological analysis (Pérez et al. 2008), respectively. The field trials were carried out at the International Center for Tropical Agriculture (CIAT) in Palmira, Valle de Cauca, Colombia. The experiment design consisted of 10 × 10 lattice with three repetitions each and two environments (drought and irrigated) evaluated at 2009 following the same methodology reported by Blair et al. (2010). The traits evaluated were days to flowering (DF), days to maturity (DM), pods per plant (PP), seed per pod (SP), seed per plant (SPL), empty pod % (EP), pod length average (PLA), weight 100 seeds (P100), and yield. A drought susceptibility index was calculated according to Rosales et al. (2000). An analogous index was calculated for wild accessions based on potential evotranspiration and mean monthly precipitation, according to Thornthwaite (1955).

Results

Overall characteristics, polymorphism, and neutrality of *Dreb2* Genes

The AP2/EREBP family members analyzed were found to be highly distinct with the *Dreb2B* gene being smaller than *Dreb2A* gene (Fig. 1). Both genes were intron-less as is characteristic of DREB genes. This structure was confirmed with the alignment between common bean ESTs (CV535836 and TC2798 for *Dreb2A* and BQ481823 as well as CA910244 for *Dreb2B*) and genomic sequences from the GSS collection of the NCBI database. The AP2/EREBP domain was highly conserved (identity of 96 %) between the two *Dreb2* proteins (Supplemental Fig. 1). Similarity between both proteins was 73 % overall. In general, *Dreb2* genes presented high polymorphism, although this was higher for the *Dreb2A* gene (34 SNPs) than for the *Dreb2B* gene (22 SNPs). Polymorphic sites were more common in non-conserved regions than in conserved domains (Table 2). However, polymorphism in the AP2/ERF conserved domain of *Dreb2B* was higher than the polymorphism in that same conserved domain of *Dreb2A*. Furthermore, recombination parameters per gene and minimum number of recombination events using the four-gamete test were slightly higher for conserved than for non-conserved regions.

Finally, molecular variation was significantly structured and related with the geographic origin for *Dreb2A*, but not for *Dreb2B* (G_{ST}, F_{ST} and S_{nn} values were bigger for *Dreb2A* than for *Dreb2B*). Furthermore, germplasm sub-groups were partially recognizable only for *Dreb2A*. Overall diversity was higher in the wild collection than in the cultivars (in terms of number of SNPs and polymorphic information content) (Table 3). Transitions were more widespread than transversions in *Dreb2B*, while this relationship was less pronounced for *Dreb2A*. Moreover, some of the *Dreb2A* polymorphic sites distinguished the four cultivars used as controls between Mesoamericans and Andeans, while *Dreb2B* did not distinguished between them.

Deviations from Wright–Fisher neutrality were significant in the cultivated germplasm collection for *Dreb2A*

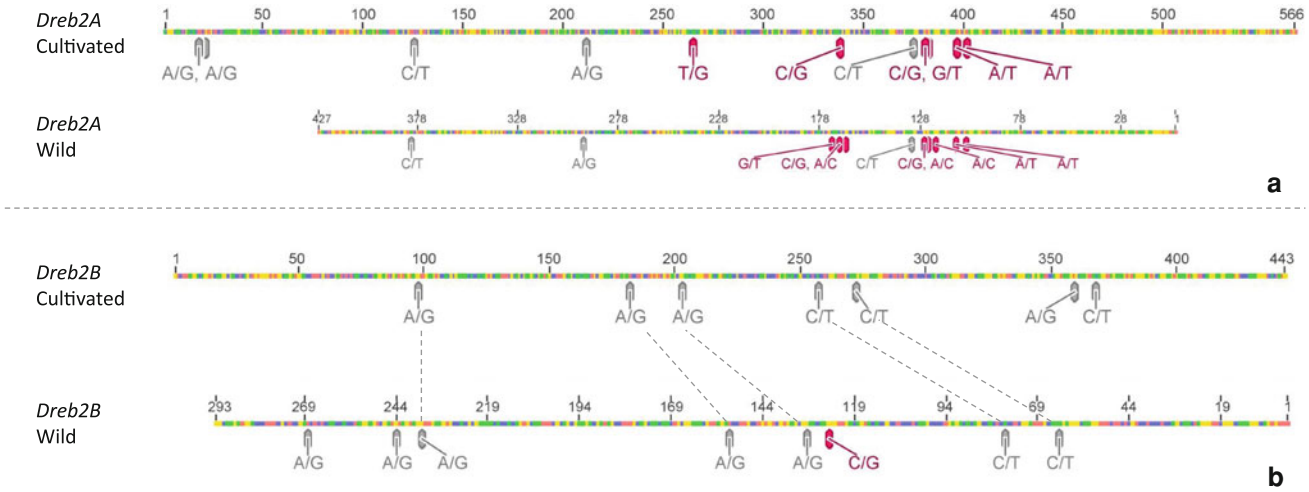

Fig. 1 Regions considered for the diversity analysis of *Dreb2A* (**a**) and *Dreb2B* (**b**) genes in the wild and cultivated collections. *Light gray markers* shown as *ovals* below the sequence are transitions and *dark gray markers* are transversions

Table 2 Population structure and recombination statistics for *Dreb2A* and *Dreb2B* genes in the wild genotypes

Gene	Region	G_{st}	F_{st}	S_{nn}	p Value S_{nn}	R	R_m
Dreb2A	Total	0.360	0.535	0.612	***	16.2	3
	AP2 domain	0.015	0.042	0.265	NS	NA	0
	Non_AP2	0.360	0.571	0.612	***	14.7	3
Dreb2B	Total	0.116	0.218	0.301	*	6.1	2
	AP2 domain	0.134	0.189	0.303	*	16.8	2
	Non_AP2	0.117	0.248	0.309	**	40.1	1

G_{ST}, F_{ST} as described in text

S_{nn} Genetic differentiation (Hudson 2000); R recombination parameter per gene = 4Nr (Hudson et al. 1987); R_m minimum number of recombination events (using the four-gamete test)

NS not significant, *NA* not-applicable

* $p < 0.05$, ** $p < 0.01$

when the analysis was carried out globally without considering population structure (Table 4); however, values varied upon consideration of structure. For example, Tajima's D was significant positive, even after considering genepool structure (Supplemental Table 2). Ramos-Onsins and Rozas' R^2 values, which tests for population growth, was not significant except in some isolated populations for *Dreb2B*. A background distribution for the nucleotide diversity based on the number of differences between any pair of individuals (π) was previously generated by SNP analysis of cultivated common beans (Cortés et al. 2011), and the π values for *Dreb2A* and *Dreb2B* were compared against this.

The background distribution was not neutral and tended to inflate π values. *Dreb2B* had a slightly tendency toward low values. On the other hand, *Dreb2A* presented extreme high values in relation with the evolutionary background, accounting for a p value less than 0.001 (Fig. 2). These values were correlated with the fact that the mismatch distributions presented a bimodal pattern for *Dreb2A,* but not for *Dreb2B* (supplemental Fig. 2). Moreover, the incongruence between the observed and the expected mismatch distributions was slightly reduced after considering population structure.

Linkage disequilibrium was found to be somewhat extensive along the analyzed genes (Fig. 3), and it decayed slightly as a function of physical distance, even when data were corrected for population structure. These results indicated little concerted evolution between genes, physical proximity between markers, and extended population structure. Finally, non-synonymous mutations were less frequent than synonymous mutations.

Haplotype analysis of DREB genes

Globally, *Dreb2A* presented more haplotypes than *Dreb2B*. Moreover, haplotypes with low frequency were more common at *Dreb2A* than at *Dreb2B* (Table 5). At least one

Table 3 Genetic diversity values for SNPs across the 401 accessions of wild (104 accessions) and cultivated (297 accessions) common bean collections

Gene	Group	SNP	Position consensus	SNP type	SNP frequency (%)	Majority allele	Minority allele	Homozygous majority allele	Homozygous minority allele	Heterozygous	Majority allele frequency	Minority allele frequency	PIC
Dreb2A	Total	1	53	[T/C]	48.50	C	T	151	142	8	0.51	0.49	**0.50**
		2	139	[A/G]	42.70	G	A	171	127	3	0.57	0.43	**0.49**
		3	192	[T/G]	0.70	T	G	299	2	0	0.99	0.01	0.01
		4	262	[A/C]	1.00	C	A	297	2	2	0.99	0.01	0.02
		5	266	[C/G]	17.10	C	G	248	50	3	0.83	0.17	**0.28**
		6	269	[T/G]	0.30	G	T	300	1	0	1.00	0.00	0.01
		7	302	[T/C]	16.80	T	C	250	50	1	0.83	0.17	**0.28**
		8	308	[C/G]	46.20	C	G	159	136	6	0.54	0.46	**0.50**
		9	310	[T/G]	37.40	G	T	185	109	7	0.63	0.37	**0.47**
		10	314	[T/G]	0.30	G	T	300	1	0	1.00	0.00	0.01
		11	324	[A/T]	35.20	A	T	193	104	4	0.65	0.35	**0.46**
		12	329	[A/T]	35.20	A	T	190	101	10	0.65	0.35	**0.46**
	Cultivated	1	14	[A/G]	41.80	A	G	117	84	1	0.58	0.42	**0.49**
		2	17	[A/G]	41.80	A	G	117	84	1	0.58	0.42	**0.49**
		3	122	[T/C]	42.60	C	T	113	83	6	0.57	0.43	**0.49**
		4	208	[A/G]	43.60	A	G	113	87	2	0.56	0.44	**0.49**
		5	261	[T/G]	1.00	T	G	200	2	0	0.99	0.01	0.02
		6	335	[C/G]	1.70	C	G	198	3	1	0.98	0.02	0.03
		7	371	[T/C]	1.70	T	C	198	3	1	0.98	0.02	0.03
		8	377	[C/G]	43.60	G	C	113	87	2	0.56	0.44	**0.49**
		9	379	[T/G]	41.80	G	T	117	84	1	0.58	0.42	**0.49**
		10	393	[A/T]	41.10	A	T	117	81	4	0.59	0.41	**0.48**
		11	398	[A/T]	40.30	A	T	117	78	7	0.60	0.40	**0.48**
	Wild	1	57	[A/T]	26.00	T	A	74	25	3	0.74	0.26	**0.38**
		2	62	[A/T]	24.50	T	A	77	25	0	0.75	0.25	**0.37**
		3	72	[A/C]	1.00	C	A	101	1	0	0.99	0.01	0.02
		4	76	[A/C]	29.40	C	A	69	27	6	0.71	0.29	**0.42**
		5	78	[C/G]	25.50	G	C	74	24	4	0.75	0.25	**0.38**
		6	84	[A/G]	46.10	A	G	55	47	0	0.54	0.46	**0.50**
		7	117	[A/C]	1.00	C	A	101	1	0	0.99	0.01	0.02
		8	120	[C/G]	47.10	G	C	53	47	2	0.53	0.47	**0.50**
		9	124	[T/G]	2.90	G	T	98	2	2	0.97	0.03	0.06
		10	247	[T/C]	15.20	C	T	86	15	1	0.85	0.15	**0.26**
		11	333	[A/G]	39.20	A	G	61	39	2	0.61	0.39	**0.48**

Table 3 continued

Gene	Group	SNP	Position consensus	SNP type	SNP frequency (%)	Majority allele	Minority allele	Homozygous majority allele	Homozygous minority allele	Heterozygous	Majority allele frequency	Minority allele frequency	PIC
Dreb2B	Total	1	8	[A/G]	0.40	G	A	262	1	0	1.00	0.00	0.01
		2	15	[A/G]	5.30	G	A	249	14	0	0.95	0.05	*0.10*
		3	99	[A/G]	6.10	A	G	247	16	0	0.94	0.06	*0.11*
		4	120	[A/G]	10.30	G	A	236	27	0	0.90	0.10	*0.18*
		5	126	[C/G]	0.40	G	C	262	1	0	1.00	0.00	0.01
		6	174	[T/C]	8.40	T	C	241	22	0	0.92	0.08	*0.15*
		7	189	[T/C]	8.40	T	C	241	22	0	0.92	0.08	*0.15*
	Cultivated	1	8	[A/G]	0.40	G	A	262	1	0	1.00	0.00	0.01
		2	15	[A/G]	5.30	G	A	249	14	0	0.95	0.05	*0.10*
		3	99	[A/G]	6.10	A	G	247	16	0	0.94	0.06	*0.11*
		4	120	[A/G]	10.30	G	A	236	27	0	0.90	0.10	*0.18*
		5	126	[C/G]	0.40	G	C	262	1	0	1.00	0.00	0.01
		6	174	[T/C]	8.40	T	C	241	22	0	0.92	0.08	*0.15*
		7	189	[T/C]	8.40	T	C	241	22	0	0.92	0.08	*0.15*
	Wild	1	93	[A/G]	27.60	A	G	42	16	0	0.72	0.28	**0.40**
		2	108	[A/G]	27.60	A	G	42	16	0	0.72	0.28	**0.40**
		3	156	[C/G]	1.70	C	G	57	1	0	0.98	0.02	0.03
		4	162	[T/C]	36.20	C	T	37	21	0	0.64	0.36	**0.46**
		5	183	[T/C]	24.10	T	C	44	14	0	0.76	0.24	**0.37**
		6	267	[T/C]	17.20	C	T	48	10	0	0.83	0.17	**0.29**
		7	274	[T/C]	1.70	C	T	57	1	0	0.98	0.02	0.03
		8	298	[T/C]	22.40	T	C	45	13	0	0.78	0.22	**0.35**

In bold: PIC values higher than 0.2, in italic: PIC values between 0.1 and 0.1

Table 4 Diversity analysis for *Dreb2A* and *Dreb2B* in the wild and cultivated collections

Gene	Group	No. SNPs	No. alleles	No. polymorphic sites	Watterson's theta	Pi	D Tajima	*p* Value	D Tajima (Andean)	*p* Value (Andean)	D Tajima (Mesoamerican)	*p* Value (Mesoamerican)	No. haplotypes	Heterozygosity
Dreb2A	Total	12	602	12	0.0045	0.0091	2.2283	*	2.0537	NS	2.0764	NS	22	0.697
	Cultivated	11	404	11	0.0036	0.0085	3.0722	**	0.1097	NS	0.6118	***	7	0.529
	Wild	11	204	11	0.0048	0.0088	1.9451	NS	0.3857	NS	0.8521	NS	19	0.823
Dreb2B	Total	7	526	7	0.0039	0.0028	−0.5590	NS	−1.2281	NS	−0.0827	NS	9	0.225
	Cultivated	7	414	7	0.0032	0.0011	−1.2848	NS	−1.7097	NS	−1.3711	NS	4	0.133
	Wild	8	116	8	0.0046	0.0071	1.3374	NS	−0.0306	NS	1.6246	NS	9	0.648

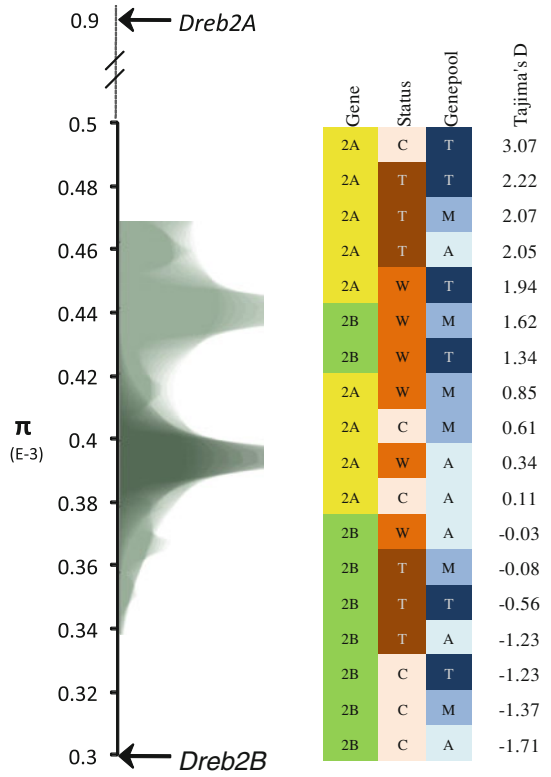

Fig. 2 Comparison of *Dreb2* genes to general nucleotide diversity in *Phaseolus vulgaris* as was computed previously with SNP markers (Cortés et al. 2011). Population structure and adaptive selection are associated with a high π values, but bottlenecks and selective sweeps are associated with low π value. *T* Total, *C* Cultivated, *W* Wild, *A* Andean, *M* Mesoamerican

hypothetical haplotype was required to analyze variation at each gene. Some of the patterns of population structure suggested in the previous section were revealed by total haplotype frequency for each gene, as well as for each population. In particular, haplotypes with the highest frequency were shared by accessions from Mesoamerican, Guatemala, Colombia, and Andean populations for both genes. A pair of equally high-frequency haplotypes shared by more than two populations was found for *Dreb2A* (Fig. 4). Interestingly, Ecuador-North Peru population did not have any of these haplotypes. A Mesoamerican versus Andean based genepool division was clear for *Dreb2A* than for *Dreb2B* (Fig. 5). Congruence between diversity control haplotypes and wild population haplotypes was straightforward for *Dreb2A*.

In summary, the combination of estimated habitat drought stress, drought severity index, population structure, and candidate gene haplotypes in a qualitative analysis over network trees revealed four main categories of relationships: First, some haplotypes included sequences of accessions from the same wild population but with very different estimated levels of drought stress associated with

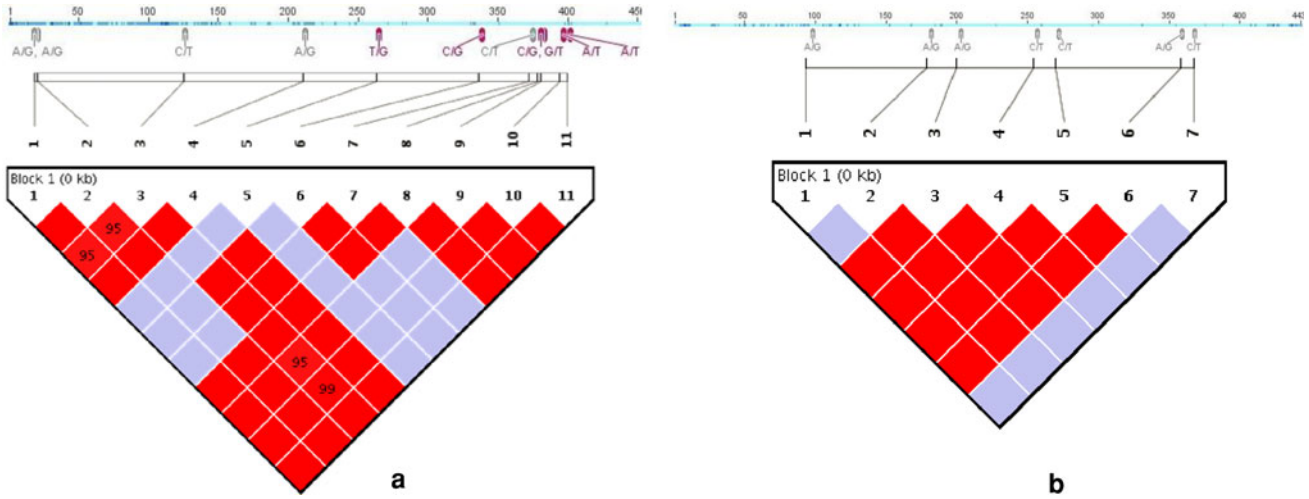

Fig. 3 Linkage disequilibrium patterns within *Dreb2A* (**a**) and *Dreb2B* (**b**) genes for wild common bean. *Dark-colored squares* indicate significant linkage disequilibrium (LD) between the markers (p value < 0.05). *Light gray marks* shown as *ovals* below the sequence are transitions and *dark gray marks* are transversions

their habitat. Second, some haplotypes included accessions from dissimilar populations that presented a similar estimated habitat drought stress. Third, some accessions had distinct haplotypes but were categorized in a similar rank of estimated drought stress, irrespective of their populations. Fourth, some accessions of the same population had dissimilar haplotypes and disparate drought stress ranks. All these categories were confirmed over the neighbor-joining trees, as well.

Discussion

Wild common bean is a reservoir of genetic variation at Dreb2

Common bean DREB-encoding genes are an example of a large gene family with a simple structure and a proven role in drought tolerance. This paper is the first attempt to characterize their diversity in both wild and cultivated common beans and builds on the characterization of the gene family across legumes and cereals by Nayak et al. (2009). Furthermore, this research integrates different lines of evidence from coalescent theory and diversity analysis to suggest the possible role of these genes in terms of drought tolerance. Traditionally, wild relatives of cultivated plants were not subjected to bottlenecks or selective sweeps. However, these processes do occur during the domestication, when various traits are selected for improvement by various stages of agriculture. Hence, we can presume that wild genepools have not been genetically eroded and that they conserve much of the original variation present in a species especially in a widespread plant such as wild *P. vulgaris* (Singh 2001).

Wild relatives are often considered to be better adapted to some detrimental environmental conditions from their original habitat than are their cultivated versions. Moreover, wild accession often presents higher levels of exogamy than cultivars. Such adaptations were lost in the transition toward the field and gave way to a superior allocation of resources for yield. Consequently, wild relatives are expected to have higher genetic diversity and phenotypic variability than the cultivated individuals, especially for genes not related to the traits that were subjected to diversifying selection or genes that are part of the domestication syndrome. This trend has been demonstrated in rice as well (Li et al. 2011; Philippe et al. 2010; Zhao et al. 2010a, b).

The hypothesis of dissimilar grades of variation between wild and cultivated plants has been reinforced for common bean in the present research using candidate gene sequences instead of genomic variation as was used previously for this species. The evaluation of haplotype variability at an important stress-related gene such as DREB allows comparisons of adaptive variation, and not just neutral polymorphism as with random molecular markers characterized previously (Blair et al. 2009; Papa et al. 2007; Kwak and Gepts 2009). The analysis of presumable adaptive variation closes the gap between diversity analysis and functional genomics. Furthermore, the selective and demographic hypotheses, which aim to explain the lack of neutrality, are distinguished easily when adaptive and neutral variation is compared in a common framework. As we will discuss in the following sections, our survey at DREB-encoding

Table 5 Haplotype information for SNPs across the 401 genotypes of common bean wild and cultivated collections

Dreb2A					*Dreb2B*				
Group	Haplotype name	Haplotype sequence	Size	Frequency	Group	Haplotype name	Haplotype sequence	Size	Frequency (%)
Total	1	CGTCCGTCTGAA	4	0.70	Total	1	GAAAGCC	26	4.90
	2	CATCCGTGGGAA	251	**41.70**		2	GGGAGCC	8	1.50
	3	CGTCCGTCTGAT	1	0.20		3	GGGAGTT	16	3.00
	4	CATCCGTGTGAA	1	0.20		4	GGGACCC	2	0.40
	5	CATCCGTCGGAA	1	0.20		5	GGAGGCC	4	0.80
	6	CGTCGGCCTGAA	1	0.20		6	AGGAGTT	2	0.40
	7	CAGCCGTGGGAA	4	0.70		7	GGAGGTT	462	**87.80**
	8	CGTAGGCCTGAA	1	0.20		8	GAAGGTT	2	0.40
	9	TGTCGGTCGGAA	1	0.20		9	GGGGGCC	4	0.80
	10	TGTAGGCCTGAA	1	0.20	Cultivated	1	GAGTTAT	17	4.10
	11	TGTACGTCTTTT	2	0.30		2	GGACCGC	4	1.00
	12	TGTCCGTGGGAA	21	3.50		3	GAGTTAC	385	**93.00**
	13	CGTCCGTGGGAA	1	0.20		4	AAACCGC	8	1.90
	14	TGTCCGTCTGTA	3	0.50	Wild	1	AACCTCCT	64	**55.20**
	15	CGTCCGTCTGTA	1	0.20		2	AACTCCCC	16	**13.80**
	16	TGTCCGTCTGTT	204	**33.90**		3	GGCCTCCT	4	3.40
	17	CGTCGTCCGGAA	2	0.30		4	GGCTCCCC	4	3.40
	18	TGTCCGTCTGAT	3	0.50		5	GGCCCCCC	4	3.40
	19	TGTCGGCCGGAA	55	9.10		6	AACCTTCT	2	1.70
	20	TGTACGTCTGTT	2	0.30		7	GGCTTTCT	18	**15.50**
	21	CGTCGGTCTGAA	1	0.20		8	AACTCCTT	2	1.70
	22	CGTCGGCCGGAA	41	6.80		9	GGGTCCCC	2	1.70
Cultivated	1	AACATCTGGAA	224	**55.40**					
	2	AACAGCTGGAA	4	1.00					
	3	GGCGTCTCTAA	3	0.70					
	4	GGTGTCTCTTA	2	0.50					
	5	GGCGTCTCTTA	1	0.20					
	6	GGTGTCTCTTT	163	**40.30**					
	7	AATGTGCCGAA	7	1.70					

Table 5 continued

Dreb2A					*Dreb2B*				
Group	Haplotype name	Haplotype sequence	Size	Frequency	Group	Haplotype name	Haplotype sequence	Size	Frequency (%)
Wild	1	TTCCGGACGCG	2	1.00					
	2	TTCCGGCCGCA	48	**23.50**					
	3	TTCCCACGGCA	22	**10.80**					
	4	TTCAGGCCTCG	1	0.50					
	5	ATCAGACGGCA	3	1.50					
	6	TTCAGACCGCG	1	0.50					
	7	TACAGACGGCA	1	0.50					
	8	AACAGACGGCA	45	**22.10**					
	9	TTCCCACGGTG	29	**14.20**					
	10	TTCAGGCCGCG	1	0.50					
	11	AAAAGACGTCA	2	1.00					
	12	TTCAGGCCTCA	1	0.50					
	13	ATCAGACGGCG	1	0.50					
	14	TTCACACGGTG	1	0.50					
	15	TTCAGACGGCG	1	0.50					
	16	AACAGACGTCA	2	1.00					
	17	TTCCGACGGTG	1	0.50					
	18	TTCCGACCGCG	1	0.50					
	19	TTCCGGCCGCG	41	**20.10**					

In bold: values higher than 10 %

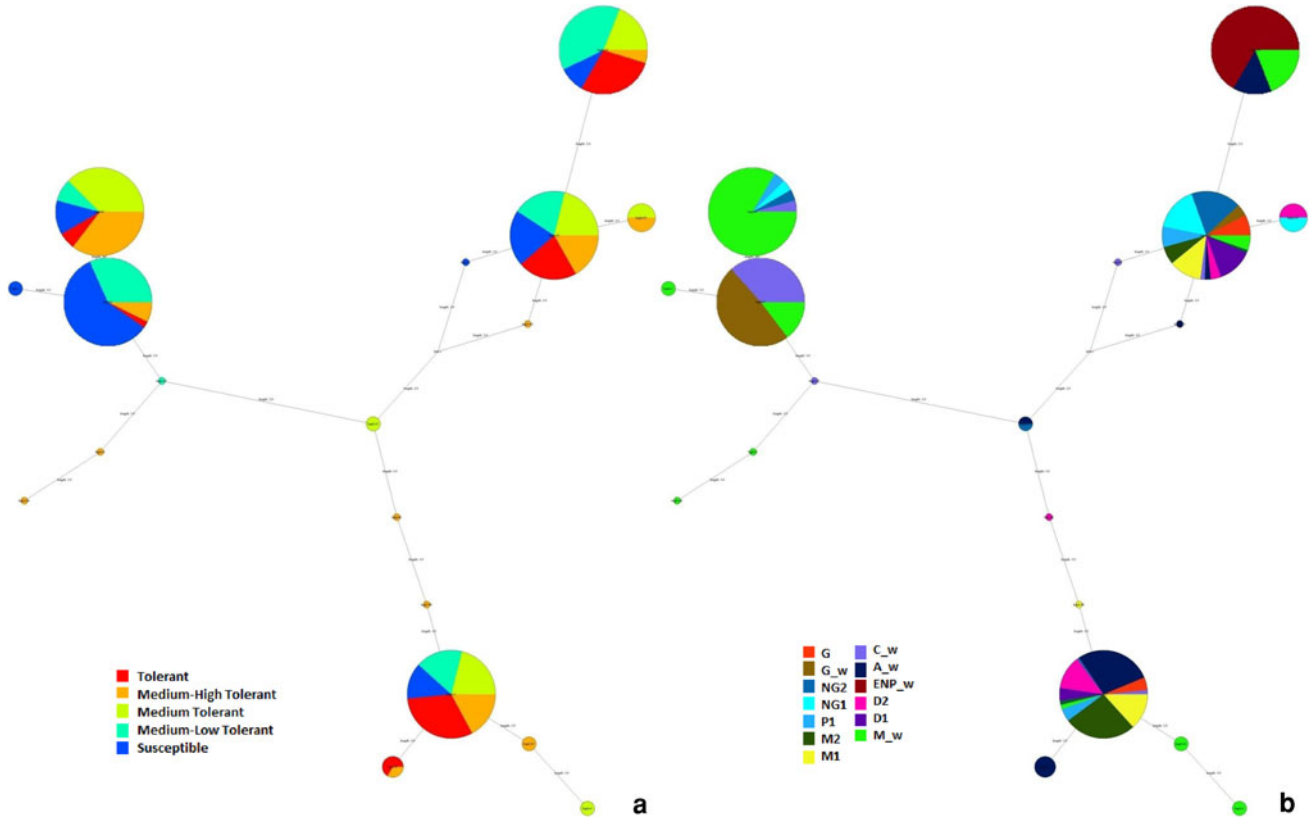

Fig. 4 Haplotype networks for *Dreb2A*. Each node represents a haplotype, its size being proportional to its frequency. An intervening segment between *two circles* corresponds to the substitutions between those haplotypes. *Small red circles* are hypothetical haplotypes. Subfigure **a** contains the drought tolerance states. Subfigure **b** shows the following groups: cultivated Mesoamerican sub-groups (M1 and M2), cultivated Durango sub-groups (D1 and D2), cultivated Guatemala race (G), cultivated Nueva Granada sub-groups (NG1 and NG2) and Peru race (P1), wild Mesoamerican (M_w), wild Guatemala (G_w), wild Colombian (C_w), wild Ecuador-North Peru (ENP_w) and wild Andean (A_w) according to previous studies (Broughton et al. 2003; Kwak and Gepts 2009) (color figure online)

genes allowed us to conclude dissimilar diversity between wild and cultivated common beans, different evolutionary imprints in each gene, and correlations between each candidate gene and drought tolerance even when accounting for the evolutionary inertia and the confounding effects of population structure.

The evaluation of drought physiology traits in wild genotypes would have been impractical due to long growth cycle (Beebe et al. 2008). Hence, the ecological analysis predicted drought tolerance successfully where other sources of information were not available. However, caution must be taken because while cultivated common bean sacrifices the rusticity of wild beans to achieve a maximum efficiency, wild common bean is thought to delay reproductive development to face drought tolerance, especially in equatorial regions where precipitation pattern is bimodal. Moreover, plant genetic responses to biotic and abiotic stresses tend to be highly correlated with precipitation pattern and fungal incidence (Agarwal et al. 2006; Chai and Zhang 1999). Consequently, heat, salinity, and edaphic stresses can be highly associated with the ecological pattern as well (Ramírez-Villegas et al. 2010).

In our first findings, the genetic variation found at the level of cultivated races was considerable, but lower than that in wild common beans. In addition, some differences existed between the adaptation of wild and cultivated individuals to arid regimens. Several of the wild beans might be valuable for plant breeding, whilst others would not be useful given adaptation and photoperiod requirements of equatorial versus sub-tropical zones (Kelly 2000). Therefore, we propose that both cultivated and wild (primary) gene pools are taken into account to exploit variation for drought tolerance while considering the need for a good harvest index as part of various yield components. An approach to breeding with wild beans is offered by the advanced backcross technique pioneered in common beans by Blair et al. (2006b).

DREB encoding paralogs have experienced different evolutionary patterns

Selective process, such as purifying selection and local adaptation, imprint in different manners on different parts of genomes, causing the departure of genetic variation from

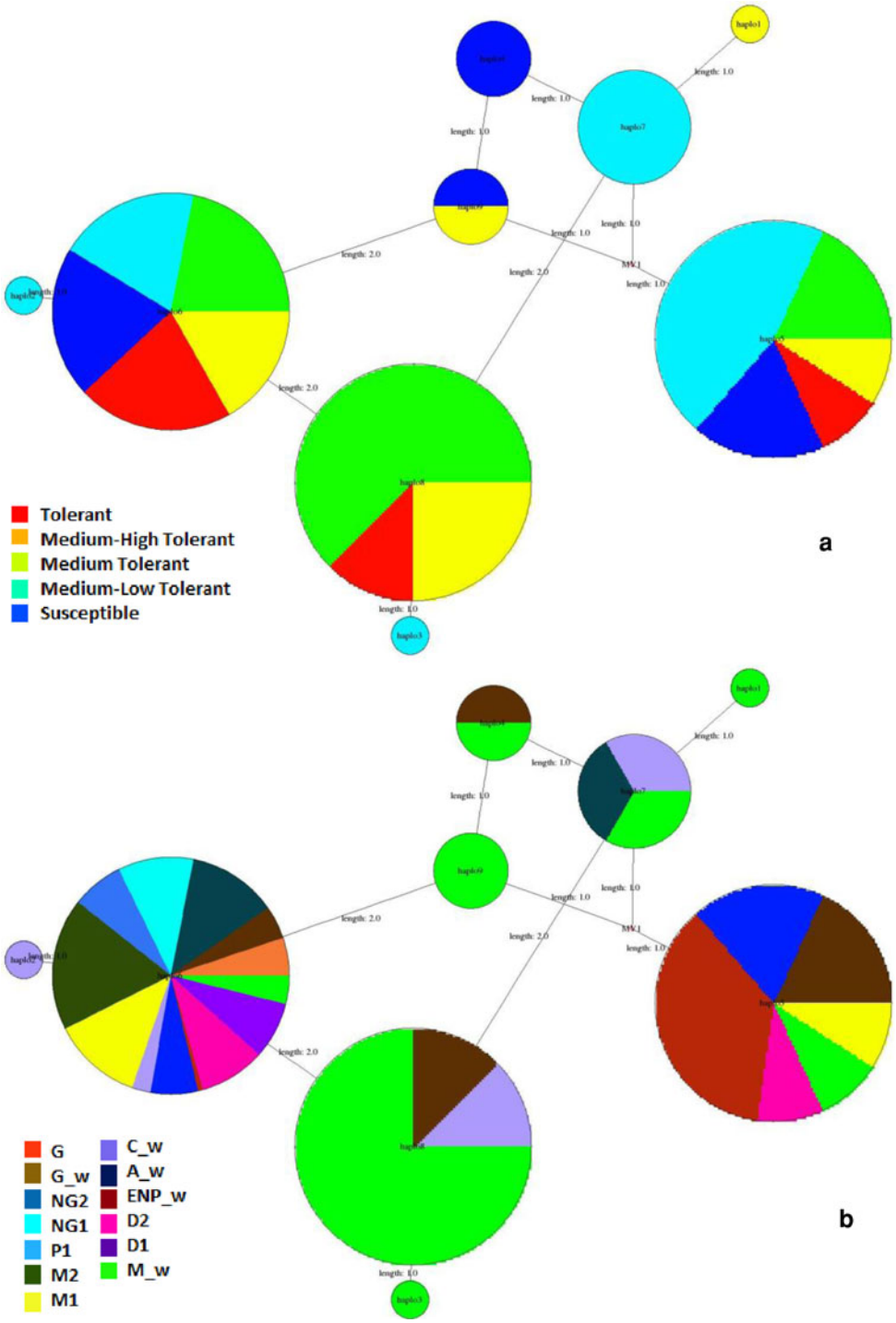

Fig. 5 Haplotype networks for *Dreb2B*. For details see legend of Fig. 4

the neutral expectations (Zhao et al. 2010a, b). Purifying selection is associated with low values of nucleotide diversity (π) and Tajima's D because only low-frequency polymorphisms can avoid being eliminated by widespread directional selection. However, recent population bottlenecks tend to achieve the same reduction in nucleotide variation.

On the other hand, local adaptation tends to homogenize haplotypes within the same niche, fix polymorphisms in different populations and eliminate low-frequency polymorphism. Consequently, few haplotypes with high frequency are generated, corresponding to high values of nucleotide diversity and Tajima's D. Nevertheless, independent domestication events, extensive population structure, and population expansions after bottlenecks can produce these same patterns (Wakeley 2008).

In the case of common bean, the two independent domestication processes generated extensive population

structure and genome-wide increases in the global nucleotide diversity (Gepts et al. 1986; Kwak and Gepts 2009). That is why we have observed a significant positive, bimodal distribution of π values. This pattern was lost when the two genepools were considered independently. It remains to be determined whether a population expansion after the two independent genepool bottlenecks could be relevant in explaining the observed pattern.

We did observe significant population expansion for the Mesoamerican wild population using the Ramos-Onsins and Rozas' R^2 statistic (Librado and Rozas 2009) and predict that an extensive survey of SNP markers would reinforce these conclusions. Particularly, a well-saturated genome-wide mismatch distribution will allow us to confirm the extent of population expansions after the domestication of the ancestral Andean and Mesoamerican wild genepools and the generation of the modern Andean wild genepool.

For wild and cultivated common bean, the global neutrality test against the Wright–Fisher neutral model was rejected for *Dreb2A* (with a bias toward high nucleotide diversity), but not for *Dreb2B*. However, when the neutrality test was made against the evolutionary background, neutrality was rejected for *Dreb2B* (with a bias toward low nucleotide diversity) and for *Dreb2A* as well (with a bias toward high nucleotide diversity). This result was likely to be a consequence of the fact that the Wright–Fisher neutral model did not account for the population structure nor the evolutionary processes of the species. Therefore, the evolutionary background is the ideal framework to make straightforward comparisons between candidate genes and genome-wide variation, and this is perfectly equivalent to applying the neutrality test to each one of the populations of wild versus cultivated accessions in each genepool as well. Both sources of evidence suggested that *Dreb2B* was subjected to selective sweeps at least in Andean and Mesoamerican wild common bean.

Selective sweeps have been a common finding in other genes and species. For example, ABA-related transcription factors in wild tomatoes (Xia et al. 2010) along with domestication genes in maize (Camus et al. 2008; Tiana et al. 2009) and rice (Caicedo et al. 2007; Li et al. 2011) have been associated with local or genomic selective sweeps. Indeed, the implications of these findings are important; because once we know how directional selective processes look like, we can identify this pattern elsewhere near other genes within the same species.

It is necessary to emphasize that population structure and climatic variability are partially correlated because both follow a latitudinal pattern (Chacón et al. 2005, 2007). This was found to be particularly true for *Dreb2A* because its variation overlapped with the evolutionary background and populations structure. Although population structure explained this gene's significantly high nucleotide diversity, it did not explain the association between *Dreb2A* and the estimated drought tolerance at the haplotype level for wild common bean.

The haplotype analysis was carried using a mixed model which accounted for population structure, so a false-positive association that are especially common in haplotype–phenotype correlations can be rejected (Sahana et al. 2010). Local adaptation explains these discoveries perfectly because it imprints the genetic regions with global high nucleotide diversity and is congruent with ecological (estimated drought tolerance) and genetic (population structure) characteristics as well. In conclusion, for *Dreb2A*, neither population structure nor population expansions after bottlenecks explain per se and *in plene* both results. Similar patterns have been found at other candidate genes for drought tolerance in tomatoes (Frankel et al. 2006, 2003; Xia et al. 2010) and *Arabidopsis* (Kim et al. 2012). Hence, we report here the first case of paralogous genes with the AP2/EREBP domain that present quite different selection and evolutionary signatures.

As a further result of our study, we found signatures of purifying selection and local adaptation in the wild and cultivated collections. However, an analysis limited to one of the genepools was unsuccessful to achieve our findings. This is a consequence of the enormous genetic and phenotypic variation stored in both genepools and in both the wild and cultivated sets of accessions. In addition, there are differential selection patterns between wild and cultivated beans, which reinforce the previous point. While common bean is selected in modern breeding programs as a bush bean, mechanized crop, wild common bean is a viney annual plant that germinates among small trees and shrubs in forest clearings or in disturbed environments with the onset of seasonal rains. The growth cycle of the wild common bean is from 8 to 10 months in length. In tropical environments with bimodal rainfall, a mid-season dry period occurs that can last 2 to 4 weeks near the sub-tropics, to as long as 3 months on the equator. In response to this mid-cycle drought the wild *P. vulgaris* enters a survival mode of slow growth and reduced physiological activity, until rainfall resumes and flowering occurs. However, cultivated beans are not subjected frequently to these environmental pressures. Moreover, it has been reported that wild common bean occupy many geographical regions at temperate zones of high-altitude regions with extensive drought stress. Those regions include the Andean arid areas of Peru, Bolivia and Chile, and the infertile highlands of northwest Mexico.

In short, the adaptive importance of DREB-encoding genes was inferred because we detected a distortion in genetic patterns in relation with the evolutionary background. Sub-functionalization is a tentative explanation of

this discrepancy, so each paralogous gene may confer more or less tolerance to drought stress, and hence is imprinted in a different manner by natural selection. In this study, association analysis using estimated drought tolerance has reinforced the importance of *Dreb2* genes in the context of this abiotic stress tolerance. Consequently, we have characterized valuable genes for plant breeding in common bean and also we have proven that signatures of selection are good predictors of markers and genes associated with a desired trait.

The practical results of this research is the potential to practice allele selection on the *Dreb* genes in a predictive manner and then to use marker-assisted selection based on SNPs within the genes themselves to move the new alleles from wild or unadapted landraces into modern cultivars. We predict that in general modern cultivars, bred on the fertile and irrigated soils of experiment stations of national and international breeding programs, do not have the ideal alleles for drought tolerance at key genes in the pathway to drought tolerance. Meanwhile, wild beans and landraces from arid regions were shown here to share common alleles at *Dreb2* genes which suggest that they are the ideal alleles to focus on for transferring to improved genetic backgrounds. Several breeding methods can be contemplated for breeding with unadapted germplasm, most involving advanced, inbred or recurrent backcrossing or recurrent selection. Among these options, recurrent backcrossing along with marker-assisted selection can be used to create isolines with the *Dreb2* alleles in an improved background. Once incorporated into a bush bean background, the novel *Dreb2* alleles can be combined with QTL loci discovered within cultivated beans (Blair et al. 2012) and the epistatic interactions between the loci can be evaluated.

Acknowledgments The authors wish to thank A. Soler and L. Díaz for DNA extractions and the personnel of the CIAT Genetic Resource Unit and CIAT bean program for seed characterization and multiplication. They also acknowledge the Cornell Biotechnology Center for support in sequencing. Sequencing by D. Brunel and team is gratefully acknowledged, too. This research was supported by the Generation Challenge Program and had technical support from the ICRISAT genomics center.

References

Afanador LK, Hadley SD, Kelly JD (1993) Adoption of a mini-prep DNA extraction method for RAPD marker analysis in common bean. Bean Improv Cooperative 35:10–11

Agarwal PK, Agarwal P, Reddy MK, Sopory SK (2006) Role of DREB transcription factors in abiotic and biotic stress tolerance in plants. Plant Cell Rep 25:1263–1274

Asfaw A, Blair MW, Almekinders C (2009) Genetic diversity and population structure of common bean (*Phaseolus vulgaris* L.) landraces from the East African Highlands. Theor Appl Genet 120:1–12

Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Bio Evol 16:37–48

Bartels D, Sunkar R (2005) Drought and salt tolerance in plants. Critical Rev Plant Sci 24:23–58

Becerra V, Paredes M, Rojo C, Díaz LM, Blair MW (2010) Microsatellite marker characterization of Chilean common bean (*Phaseolus vulgaris* L.) germplasm. Crop Sci 50:1–10

Beebe S, Rao IM, Cajiao C, Grajales M (2008) Selection for drought resistance in common bean also improves yield in phosphorus limited and favorable environments. Crop Sci 48:582–592

Benchimol LL, de Campos T, Carbonell SAM, Colombo CA, Chioratto AF, Formighieri EF, Gouvea LRL, de Souza AP (2007) Structure of genetic diversity among common bean (*Phaseolus vulgaris* L.) varieties of Mesoamerican and Andean origins using new developed microsatellite markers. Genet Res Crop Evol 54:1747–1762

Blair MW, Giraldo MC, Buendia HF, Tovar E, Duque MC, Beebe SE (2006a) Microsatellite marker diversity in common bean (*Phaseolus vulgaris* L.). Theor Appl Genet 113:100–109

Blair MW, Iriarte G, Beebe S (2006b) QTL analysis of yield traits in an advanced backcross population derived from a cultivated Andean x wild common bean (*Phaseolus vulgaris* L.) cross. Theor Appl Genet 112:1149–1163

Blair MW, Diaz JM, Hidalgo R, Diaz LM, Duque MC (2007) Microsatellite characterization of Andean races of common bean (*Phaseolus vulgaris* L.). Theor Appl Genet 116:29–43

Blair M, Diaz LM, Buendia HF, Duque MC (2009) Genetic diversity, seed size associations and population structure of a core collection of common beans (*Phaseolus vulgaris* L.). Theor Appl Genet 119:955–972

Blair MW, Medina JI, Astudillo C, Rengifo J, Beebe SE, Machado G, Graham R (2010) QTL for seed iron and zinc concentration and content in a Mesoamerican common bean (*Phaseolus vulgaris* L.) population. Theor Appl Genet 121:1059–1070

Blair MW, Galeano CH, Tovar E, Muñoz-Torres MC, Velasco A, Beebe SE, Rao IM (2012) Development of a Mesoamerican intra-genepool genetic map for QTL detection in a drought tolerant x susceptible common bean (*Phaseolus vulgaris* L.) cross. Mol Breed 29:71–88

Bradbury PJ, Zhang Z, roon DE, Casstevens RM, Ramdoss Y, Buckler ES (2007) TASSELL Software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635

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

Caicedo AL, Williamson SH, Hernandez RD, Boyko A, Fledel-Alon A, York TL, Polato NR, Olsen KM, Nielsen R, McCouch SR, Bustamante CD, Purugganan MD (2007) Genome-wide patterns of nucleotide polymorphism in domesticated rice. PLoS Genet 3:1745–1756

Camus L, Chevin LM, Cordet CT, Charcosset A, Manicacci D, Tenaillon MI (2008) Patterns of molecular evolution associated with two selective sweeps in the Tb1-Dwarf8 region in maize. Genetics 180:1107–1121

Chacón MI, Pickersgill B, Debouck DG (2005) Domestication patterns in common bean (*Phaseolus vulgaris* L.) and the origin of the Mesoamerican and Andean cultivated races. Theor Appl Genet 110:432–444

Chacón MI, Pickersgill B, Debouck DG, Arias JS (2007) Phylogeographic analysis of the chloroplast DNA variation in wild common bean (*Phaseolus vulgaris* L.) in the Americas. Plant Syst Evol 266:175–195

Chai TY, Zhang YX (1999) Gene expression analysis of a proline-rich protein from bean under biotic and abiotic stress acta botanica sínica 41:111–113

Chen JQ, Meng XP, Zhang Y, Xia M, Wang XP (2008) Over-expression of OsDREB genes lead to enhanced drought tolerance in rice. Biotechnol Lett 30:2191–2198

Cortés AJ, Chavarro MC, Blair MW (2011) SNP marker diversity in common bean (*Phaseolus vulgaris* L.). Theor Appl Genet doi. doi:10.1007/s00122-011-1630-8

Diaz LM, Blair MW (2006) Race structure within the Mesoamerican gene pool of common bean (*Phaseolus vulgaris* L.) as determined by microsatellite markers. Theor Appl Genet 114:143–154

Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113

Frankel N, Hasson E, Iusem ND, Rossi M (2003) Adaptive evolution of the eater stress—induced gene Asr2 in *Lycopersicon* species dwelling in arid habitats. Mol Bio Evol 20:1955–1962

Frankel N, Carrari F, Hasson E, Iusem ND (2006) Evolutionary history of the *Asr* gene family. Gene 378:74–83

Gepts P (1998) Origin and evolution of common bean : past events and recent trends. HortScience 33:1119–1135

Gepts P, Debouck DG (1991) Origin, domestication, and evolution of the common bean. In: van Schoonhaven A, Voysest O (eds) Common beans: research for Crop Improvement Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia

Gepts P, Osborn TC, Rashka K, Bliss FA (1986) Phaseolin-protein variability in wild forms and landraces of the common bean (*Phaseolus vulgaris*): evidence for multiple centers of domestication. Econ Bot 40:451–468

Gepts P, Aragão F, Barros E, Blair M, Brondani R, Broughton W, Galasso I, Hernandez G, Kami J, Lariguet P, McClean P, Melotto M, Miklas P, Pauls P, Pedrosa-Harand A, Porch T, Sánchez F, Sparvoli F, Yu K (2008) Genomics of *Phaseolus* beans, a major source of dietary protein and micronutrients in the tropics. In: Moore P, Ming R (eds) Genomics of tropical crop plants. Springer, New York

Hudson RR (2000) A new statistic for detecting genetic differentiation. Genetics 155:2011–2014

Hudson RR, Kreitman M, Aguade M (1987) A test of neutral molecularevolution based on nucleotide data. Genetics 116:153–159

Kelly JD (2000) Remaking bean plant architecture for efficient production. Adv Agron 71:109–143

Kim J-S, Mizoi J, Yoshida T, Fujita Y, Nakajima J, Ohori T, Todaka D, Nakashima K, Hirayama T, Shinozaki K, Yamaguchi-Shinozaki K (2012) An ABRE promoter sequence is involved in osmotic stress-responsive expression of the DREB2A gene, which encodes a transcription factor regulating drought-inducible genes in Arabidopsis. Plant Cell Physiol 52:2136–2146

Kizis D, Lumbreras V, Pagès M (2001) Role of AP2/EREBP transcription factors in gene regulation during abiotic stress. FEBS Lett 498:187–189

Kwak M, Gepts P (2009) Structure of genetic diversity in the two major gene pools of common bean (*Phaseolus vulgaris* L., Fabaceae). Theor Appl Genet 118:979–992

Li ZM, Zheng XM, Ge S (2011) Genetic diversity and domestication history of African rice (*Oryza glaberrima*) as inferred from multiple gene sequences. Theor Appl Genet 123:21–31

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

Morran S, Eini O, Pyvovarenko T, Parent B, Singh R, Ismagul A, Eliby S, Shirley N, Landridge P, Lopato S (2011) Improvement of stress tolerance of wheat and barley by modulation of expression of DREB/CBF factors. Plant Biotechnol J 9:230–249

Nayak SN, Balaji J, Upadhyaya HD, Hash T, Kishor K, Chattopadhyay D, Rodriguez LM, Blair MW, Baume M, McNally K, This D, Hoisington DA, Varshney RK (2009) Isolation and sequence analysis of DREB2A homologues in three cereal and two legume species. Plant Sci 177:460–467

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

Papa R, Gepts P (2003) Asymmetry of gene flow and differential geographical structure of molecular diversity in wild and domesticated common bean (*Phaseolus vulgaris* L.) from Mesoamerica. Theor Appl Genet 106:239–250

Papa R, Bellucci E, Rossi M, Leonardi S, Rau D, Gepts P, Nanni L, Attene G (2007) Tagging the signatures of domestication in common bean (*Phaseolus vulgaris*) by means of pooled DNA samples. Ann Bot 100:1039–1051

Paredes M, Becerra V, Tay J, Blair MW, Bascur G (2010) Selection of a representative core collection from the Chilean common bean germplasm. Chil J Agric Res 70:3–15

Payró E, Gepts P, Colunga P, Zizumbo D (2005) Spatial distribution of genetic diversity in wild populations of *Phaseolus vulgaris* L. from Guanajuato and Michoacán, México. Genet Resour Crop Evol 52:589–599

Pérez JC, Monserrate F, Beebe S, Blair M (2008) Field evaluation of a common bean reference collection for drought tolerance. In: CIAT (ed) Annual report outcome line SBA-1 product 2: beans that are more productive in smallholder systems of poor farmers. CIAT, Palmira, Colombia

Philippe R, Courtois B, McNally KL, Mournet P, El-Malki R, Le Paslier MC, Fabre D, Billot C, Brunel D, Glaszmann JC, This D (2010) Structure, allelic diversity and selection of Asr genes, candidate for drought tolerance, in *Oryza sativa* L. and wild relatives. Theor Appl Genet 121:769–787

Rafalski JA (2010) Association genetics in crop improvement. Curr Opin Plant Biol 13:174–180

Ramírez-Villegas J, Khoury C, Jarvis A, Debouck DG, Guarino L (2010) A gap analysis methodology for collecting crop genepools: a case study with *Phaseolus* beans. PLoS ONE 5:e13497

Riechmann JL, Meyerowitz EM (1998) The AP2/EREBP family of plant transcription factors. Biol Chem 379:633–646

Rosales R, Ramírez P, Acosta JA, Castillo F, Kelly JD (2000) Grain yield and drought tolerance of common bean under field conditions. Agrociencia 34:153–165

Rossi M, Bitocchi E, Bellucci E, Nanni L, Rau D, Attene G, Papa R (2009) Linkage disequilibrium and population structure in wild and domesticated populations of *Phaseolus vulgaris* L. Evol Appl 2:504–522

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19:2496–2497

Sahana G, Guldbrandtsen B, Janss L, Lund MS (2010) Comparison of association mapping methods in a complex pedigreed population. Genet Epidemiol 34:455–462

Singh SP (2001) Broadening the genetic base of common bean cultivars: a review. Crop Sci 41:1659–1675

Singh SP (2005) Common bean (*Phaseolus vulgaris* L.). In: Singh RJ, Jauhar PP (eds) Genetic resources, chromosome engineering, and crop improvement, grain legumes. CRC Press, London

Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Amer J Hum Genet 73:1162–1169

Taiz L, Zeiger E (2006) Plant physiology, 4th edn. Sinauer Associates, Sunderland, MA

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

Thornthwaite, CW, Mather JR (1955) The water balance. Climatology 8(1):1–104

Tiana F, Stevensa NM, Buckler ES (2009) Tracking footprints of maize domestication and evidence for a massive selective sweep on chromosome 10. Proc Nat Acad Sci USA 106:9979–9986

Tohme J, Jones P, Beebe S, Iwanaga M (1995) The combined use of agroecological and characterization data to establish the CIAT *Phaseolus vulgaris* core collection. In: Hodgkin T, Brown AH, van Hintum JL, Morales EA (eds) Core collections of plant genetic resources. John Wiley & Sons, New York

Tohme J, Gonzalez O, Beebe S, Duque MC (1996) AFLP analysis of gene pools of a wild bean core collection Crop Sci 36:1375–1384

Wakeley J (2008) Coalescent theory: an introduction. Harvard University, Cambridge

Wang Y-M, He C-F (2007) Isolation and characterization of a cold-induced DREB gene from *Aloe vera* L. Plant Mol Biol Rep 25:121–132

Wang Q, Guan Y, Wu Y, Chen H, Chen F, Chu C (2008) Overexpression of a rice OsDREB1F gene increases salt, drought, and low temperature tolerance in both *Arabidopsis* and rice. Plant Mol Biol 67:589–602

Watterson GA (1975) Number of segregating sites in genetic models without recombination. Theor Popul Biol 7:256–276

Xia H, Camus-Kulandaivelu LT, Stephan W, Tellier AL, Zhang Z (2010) Nucleotide diversity patterns of local adaptation at drought-related candidate genes in wild tomatoes. Mol Ecol (Online version)

Yang Y, Wu J, Zhu K, Liu L, Chen F, Yu D (2009) Identification and characterization of two chrysanthemum (*Dendronthema x moriforlium*) DREB genes, belonging to the AP2/EREBP family. Mol Biol Rep 36:71–81

Zhang XY, Blair MW, Wang SM (2008) Genetic diversity of chinese common bean (*Phaseolus vulgaris* L.) landraces assessed with simple sequence repeat markers. Theor Appl Genet 117:629–640

Zhao J, Ren W, Zhi D, Wang L, Xia G (2007) Arabidopsis DREB1A / CBF3 bestowed transgenic tall fescue increased tolerance to drought stress. Plant Cell Rep 26:1521–1528

Zhao KY, Wright M, Kimball J, Eizenga G, McClung A, Kovach M, Tyagi W, Ali ML, Tung CW, Reynolds A, Bustamante CD, McCouch SR (2010a) Genomic diversity and introgression in *O. sativa* reveal the impact of domestication and breeding on the rice genome. Plos One 5(5):e10780

Zhao L, Hu Y, Chong K, Wang T (2010b) ARAG1, an ABA-responsive DREB gene, plays a role in seed germination and drought tolerance of rice. Ann Bot 105:401–409

Zhuang J, Xiong AS, Peng RH, Gao F, Zhu B, Zhang JA, Fu XY, Jin XF, Chen JM, Zhang Z, Qiao YS, Yao QH (2010) Analysis of *Brassica rapa* ESTs: gene discovery and expression patterns of AP2/ERF family genes. Mol Biol Rep 37:2485–2492