



Genotyping by Sequencing and Genome–Environment Associations in Wild Common Bean Predict Widespread Divergent Adaptation to Drought

Andrés J. Cortés^{1*†} and Matthew W. Blair²

¹ Department of Biological and Environmental Sciences, University of Gothenburg, Gothenburg, Sweden, ² Department of Agricultural and Environmental Science, Tennessee State University, Nashville, TN, United States

OPEN ACCESS

Edited by:

Nicholas Provart,
University of Toronto, Canada

Reviewed by:

Frédéric Marsolais,
Agriculture and Agri-Food Canada,
Canada

Umesh K. Reddy,
West Virginia State University,
United States

*Correspondence:

Andrés J. Cortés
acortes@corpoica.org.co

† Present address:

Andrés J. Cortés,
Centro de Investigación La Selva,
Corporación Colombiana de
Investigación Agropecuaria, Vereda
Llano Grande, Colombia

Specialty section:

This article was submitted to
Plant Abiotic Stress,
a section of the journal
Frontiers in Plant Science

Received: 06 November 2017

Accepted: 23 January 2018

Published: 21 February 2018

Citation:

Cortés AJ and Blair MW (2018)
Genotyping by Sequencing
and Genome–Environment
Associations in Wild Common Bean
Predict Widespread Divergent
Adaptation to Drought.
Front. Plant Sci. 9:128.
doi: 10.3389/fpls.2018.00128

Drought will reduce global crop production by >10% in 2050 substantially worsening global malnutrition. Breeding for resistance to drought will require accessing crop genetic diversity found in the wild accessions from the driest high stress ecosystems. Genome–environment associations (GEA) in crop wild relatives reveal natural adaptation, and therefore can be used to identify adaptive variation. We explored this approach in the food crop *Phaseolus vulgaris* L., characterizing 86 geo-referenced wild accessions using genotyping by sequencing (GBS) to discover single nucleotide polymorphisms (SNPs). The wild beans represented Mesoamerica, Guatemala, Colombia, Ecuador/Northern Peru and Andean groupings. We found high polymorphism with a total of 22,845 SNPs across the 86 accessions that confirmed genetic relationships for the groups. As a second objective, we quantified allelic associations with a bioclimatic-based drought index using 10 different statistical models that accounted for population structure. Based on the optimum model, 115 SNPs in 90 regions, widespread in all 11 common bean chromosomes, were associated with the bioclimatic-based drought index. A gene coding for an ankyrin repeat-containing protein and a phototropic-responsive NPH3 gene were identified as potential candidates. Genomic windows of 1 Mb containing associated SNPs had more positive Tajima's D scores than windows without associated markers. This indicates that adaptation to drought, as estimated by bioclimatic variables, has been under natural divergent selection, suggesting that drought tolerance may be favorable under dry conditions but harmful in humid conditions. Our work exemplifies that genomic signatures of adaptation are useful for germplasm characterization, potentially enhancing future marker-assisted selection and crop improvement.

Keywords: drought tolerance, adaptation, genomic signatures of selection, divergent selection, SNP markers, Tajima's D

INTRODUCTION

Understanding the genomic signatures associated with environmental variation is of great interest as it provides insights into how organisms adapt to their environment (Stinchcombe and Hoekstra, 2008; Hoffmann and Sgro, 2011; Savolainen et al., 2013). Recent genomic studies in wild plant populations have demonstrated that genome–environment associations (GEA)

can be used to identify adaptive loci and predict phenotypic variation. Generally, the studies have associated single nucleotide polymorphism (SNP) alleles and parameters from the accessions' environment of origin, to determine the potential for abiotic stress adaptation. For instance, Turner et al. (2010) predicted genetic adaptive variation to serpentine soils in *Arabidopsis lyrata*, Hancock et al. (2011) identified climate-adaptive genetic loci among a set of geographically diverse *Arabidopsis thaliana*, Fischer et al. (2013) predicted adaptive variation to topo-climatic factors in *Arabidopsis halleri*, Pluess et al. (2016) predicted genetic local adaptation to climate at a regional scale in *Fagus sylvatica*, and Yeaman et al. (2016) detected convergent local adaptation in two distantly related species of conifers.

The GEA approach has been also explored in some crop accessions as a prospection strategy for discovering wild germplasm or landraces for breeding as an alternative to traditional phenotyping. For example, Lasky et al. (2012) and Des Marais et al. (2014) looked at natural variation in *Arabidopsis* for drought resistance and water use efficiency, Yoder et al. (2014) was able to capture adaptive variation to thermal tolerance, drought tolerance, and resistance to pathogens in *Medicago truncatula*, Lasky et al. (2015) predicted genotype \times environment interactions to drought stress and aluminum toxicity in *Sorghum bicolor*, and Berthouly-Salazar et al. (2016) uncovered genomic regions involved in adaptation to abiotic and biotic stress on two climate gradients in *Cenchrus americanus*. Our group has focused on exploring the marker \times environment association approach with wild relatives of the food crop common bean (*Phaseolus vulgaris* L.) and we have shown candidate genes such as DREB, ASR, and ERECTA to have haplotypes associated with drought tolerance (Cortés et al., 2012a,b; Blair et al., 2016). In this study, we couple genome–environment association with estimates of genome-wide diversity in wild accessions by using a whole-genome marker method with thousands of SNPs combined with climatic indices.

Among the most comprehensive marker systems for common bean SNP detection is the method called genotyping by sequencing (GBS). This technique has the flexibility of being adaptable to wild relatives with no previous sequence information needed although a reference genome for the species is useful (Elshire et al., 2011). To date, the GBS method has mostly been conducted for cultivars of common bean but rarely for wild accessions, which was one of the purposes of our research. A critical step for GBS assays is the preparation of restriction enzyme digested reduced representation libraries of genomic DNA that is barcoded for evaluation on a high throughput sequencer. As an initial example of the method for genetic mapping in common bean, Hart and Griffiths (2015) compared *Pst*I as a methylation sensitive enzyme to the non-methylation sensitive *Ape*K1 proposed for most small-genome species by Elshire et al. (2011) and evaluated a population of 84 lines and 12 parental checks. Bhakta et al. (2015) used the *Pst*I GBS method to perform high resolution mapping on a population of 188 RILs from the cross of Jamapa \times Calima, cultivated beans from opposite gene pools and discovered nearly 2,000 usable SNP loci for genetic mapping. In parental surveys, Zou et al. (2013) used *Hae*III digestion and library construction in 36 Canadian

small seeded breeding lines to discover nearly 25,000 physically tagged SNPs. Subsequently, Ariani et al. (2016) used another new enzyme, the four base pair cutter *Cvi*AII, to evaluate a broader range of 18 common beans (one ancestral wild, four Andean domesticates, four wild Mesoamericans and nine cultivated Mesoamerican) finding 3,200–21,000 SNPs/genotype. Schroder et al. (2016), in contrast, used a combination of *Mse*I and *Taq*I enzymes along with size selections to compare the feasibility of double digestion and small fragment size generation with four bp cutters, to the *Ape*K1 method for 25 Mesoamerican beans, finding up to 12.5 times more usable SNPs with an 8X coverage. More recently, Ariani et al. (2017) used the same protocol than Ariani et al. (2016) to reveal the spatial and temporal scales of range expansion in wild common bean. In all cases size selection, sequencing depth and a reference genome from Schmutz et al. (2014) were critical for calling SNPs and avoiding missing data or false positives due to poor sequence coverage. DARTseq, a modification of the GBS method, has been used on 188 cultivated Brazilian genotypes for diversity assessment finding its value in population structure analysis (Valdisser et al., 2017). With these results in the literature, we were confident that the GBS method would be practical for generating many SNPs in wild accessions of common bean and could be used for both diversity evaluation and association analysis.

Wild bean are thought to have diverged from its sister species in the tropical Andes (Rendón-Anaya et al., 2017) to later diversified in South and Central America from an original range in Central America, after which domestication in the southern and northern ends of each region gave origin to Andean and Mesoamerican domesticates, respectively (Gepts and Debouck, 1991; Kwak and Gepts, 2009; Schmutz et al., 2014). Studies of wild beans, show that despite the bimodality in domestication, they are loosely structured across the full range of environments from Northern Mexico to Northern Argentina, with identifiable sub-populations of wild types centered in Argentina/Southern Bolivia, Ecuador/Northern Peru, Colombia, Guatemala, Highland Mexico, and Lowland Mexico (Blair et al., 2012; Cortés, 2013; Cortés et al., 2013). The southernmost sub-population is typical of the Andean gene pool while the two northern most sub-populations parallel the races found in Mesoamerican common bean with the centrally located sub-populations being intermediate. Among the wild beans many accessions survive and reproduce well in drought-affected regions of the New World, but have not been used for breeding the important seed types of the region.

Further wild germplasm characterization is important as common bean 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). This situation may become worst as increased drought due to climate change will reduce global crop production in >10% by 2050 with a potential to substantially worsen global malnutrition (Tai et al., 2014). Meanwhile wild beans are adapted as herbaceous species to dry forests and semi-arid regions of the Americas. Therefore,

increasing drought tolerance in common bean commercial varieties is highly desirable and one potential source of resistance is from the wild accessions of the species. Given this, our milestones for this study were to: (1) implement GBS technology in wild accessions of common bean, (2) characterize geo-referenced climate information about the geographic origins of the wild accessions, and (3) find marker \times environment associations at the whole genome level using the SNP loci discovered in the germplasm set. Specifically, we quantified SNP allelic associations with a bioclimatic-based drought index in order to identify adaptive variation suitable to breed drought tolerant varieties.

MATERIALS AND METHODS

Plant Material and Compilation of the Bioclimatic-Based Drought Index

A total of 86 DNAs from *Phaseolus vulgaris* were used in this study, extracted from 86 geo-referenced wild common bean accessions (**Supplementary Table S1**). All the genotypes were provided by the Genetic Resources Unit at the International Center for Tropical Agriculture and are preserved under the treaty for genetic resources from the Food and Agriculture Organization hereafter abbreviated as the FAO collection. These accessions were selected to be a representative of a core collection for wild beans (Tohme et al., 1996) and all five subpopulations and gene pools were uncovered by single-sequence repeat (SSR) markers by Blair et al. (2012). Geographic information provided for each accession by the Genetic Resource Unit¹ was also used as described in Cortés et al. (2013) to estimate drought stress at each collection site based on the precipitation regimes coupled with potential evapotranspiration (PET) models (Thorntwaite and Mather, 1957; Hamon, 1961) accounting for the effects of historical temperature and radiation (Hijmans et al., 2005).

More concretely, monthly mean air temperature and monthly precipitation averaged for the years 1970–2000 were downloaded from WorldClim² for each accession coordinate. The Thorntwaite method was then used to calculate PET (Cortés et al., 2013 for the detailed equations) considering the effects of both temperature and radiation (Thorntwaite and Mather, 1957). PET and the precipitation data were then combined in a monthly drought index that was averaged and normalized across all 12 months (Cortés et al., 2013 for the detailed equations). This normalized annual mean drought index (**Supplementary Figure S1**, Shapiro–Wilk normality test P -value = 0.106), fed by monthly environmental data averaged for the years 1970–2000, is hereafter referred as bioclimatic-based drought index. Two properties of this index must be noted. First, it assumes that plant distribution is in equilibrium with niche requirements and ecological forces, so that the errant presence of poorly adapted genotypes can be discarded (Forester et al., 2016). Second, this index is stable across years as it is based on climate data averaged across three decades. The ecological optimum, the

stability and the normality of this index make it ideal for GEA analyses, in contrast with crude environmental measures.

Sample Collection, DNA Extraction, and Genotyping-by-Sequencing

Leaf tissue weighing approximately 20 mg was harvested at 40 days after plant germination and was immediately dried in Silica Gel (Sigma–Aldrich, Germany). Genomic DNA was extracted using the QIAGEN DNeasy Plant Mini Kit (QIAGEN, Germany), following the manufacturer's instructions, and quantified using a Qubit® dsDNA HS Fluorometer (Life Technologies, Stockholm, Sweden). One 96-plex GBS assay was made according to Elshire et al. (2011) for the 86 accessions, with library preparation performed with *ApeKI* digestions at the Biotechnology Resource Center of the Institute for Genomic Diversity (Cornell University, United States). Genotyping and SNP calling were done with TASSEL software v. 3.0 (Glauibitz et al., 2014) based on the reference genome v. 2.0 for *P. vulgaris* deposited at the Phytozome website³. Sequence tags were aligned to the *P. vulgaris* reference genome (Schmutz et al., 2014), using the BWA method (Li and Durbin, 2007). A customized physical map was calculated and drawn using R v.3.3.1 (R Core Team) to place each new GBS locus.

Identification of Loci Associated with the Bioclimatic-Based Drought Index

In order to account for possible demographic effects we examined subpopulation structure in the 86 geo-referenced wild common bean accessions using principal coordinates analysis (PCoA) implemented in the software Trait Analysis by aSSociation, Evolution and Linkage, Tassel v.5 (Bradbury et al., 2007). The same dataset and software were used to perform association analyses between the SNP markers and the bioclimatic-based drought index from Cortés et al. (2013). A total of ten generalized (GLM) and mixed (MLM) linear models were compared. Within each model family, five models were built as follows: (1) model with the gene pool identity and the first two PCoA axes scores as covariates; (2) models with the within-gene pool subpopulation identity, according to Blair et al. (2012), and the first two PCoA axes scores as covariates; (3) model with the first two PCoA axes scores as covariates; (4) model with the within-gene pool subpopulation identity, according to Blair et al. (2012), as covariate; and (5) model with the gene pool identity as covariate. Despite that GLMs usually exhibit a higher rate of false positives (Rosenberg et al., 2010), they were still considered in the present study for comparative purposes. All five MLMs used a centered IBS kinship matrix as a random effect to control for genomic background implementing the EMMA and P3D algorithms to reduce computing time (Zhang et al., 2010). QQ-plots of the P -values were inspected to assess whether excessive numbers of false positives were generated and choose in this way the optimum model. Significant associations were determined using a strict Bonferroni correction of P -values at $\alpha = 0.001$, leading to a significance threshold of 4.4×10^{-8} (0.001 divided by the

¹genebank.ciat.cgiar.org

²<http://www.worldclim.org>

³<https://phytozome.jgi.doe.gov/pz/portal.html>

number of markers, 22,845) or $-\log_{10}(4.4 \times 10^{-8}) = 7.36$. The construction of customized PCoA and Manhattan diagrams was carried out with the software R v.3.3.1 (R Core Team).

Potential candidate genes were identified within the 1,000 bp sections flanking each SNP marker that was associated with the bioclimatic-based drought index by using the common bean reference genome (Schmutz et al., 2014) and the PhytoMine and BioMart tools in Phytozome v.12³.

Sliding Window Analysis

We used a sliding window approach (window size = 1×10^6 bps, step size = 200 kb) to describe patterns of variation and overall divergence across the genome. We computed per-window averages of SNP density, nucleotide diversity as measured by π (Nei, 1987), Watterson's theta (θ) estimator (Watterson, 1975), and Tajima's D (Tajima, 1989) using the software Tassel v.5 (Bradbury et al., 2007) and customized R scripts. Results of all windowed analyses were plotted against window midpoints in millions of base pairs (Mb) also using the software R v.3.3.1 (R Core Team).

Bootstrap-based means and 95% confidence intervals around the mean were calculated for these four summary statistics (SNP density, π , θ and Tajima's D) when computed in sliding windows that contained or did not contain at least one marker that was associated with the bioclimatic-based drought index. Each summary statistic of windows containing and not containing associated SNPs was randomly resampled with replacement (bootstrapping) across windows within grouping factor (associated vs. no associated), and the overall mean was stored for each grouping factor. This step was iterated 1,000 times using customized R scripts. Bootstrapping was performed independently for each summary statistic in order to eliminate correlations among these.

RESULTS

GBS Results

The raw Illumina DNA sequence data (180,540,321 high-quality barcoded reads per lane) were processed through the GBS analysis pipeline as implemented in TASSEL-GBS v3.0. The GBS analysis generated 1,625,330 unique sequence clusters (tags, Elshire et al., 2011; Glaubitz et al., 2014). Of the total number of tags, 71.2% aligned uniquely to the *P. vulgaris* reference genome, 10.1% had multiple matches and 18.7% were unaligned. A total of 197,314 putative SNP markers were identified in the aligned tags. Of these, 47% with more than 20% missing data, a default threshold used for GBS studies (Glaubitz et al., 2014), and further 18% with minimum allele frequencies (MAF) below 0.05 were excluded from the dataset. The high number of missing data/unaligned sequences and low MAF value SNPs can be explained by the naturally high levels of sequence diversity in wild accessions (Glaubitz et al., 2014). Still the GBS project yielded 22,845 SNP markers of high quality with an average read depth of 13.6X gene coverage. The locations of these GBS-derived SNP loci for the wild accessions were identified by sequence similarity and drawn to scale based on their physical

map distribution in the common bean genome (Figure 1). The centromeres according to Schmutz et al. (2014) were marked with the extent of the centromeric repeats in comparison with the GBS derived markers.

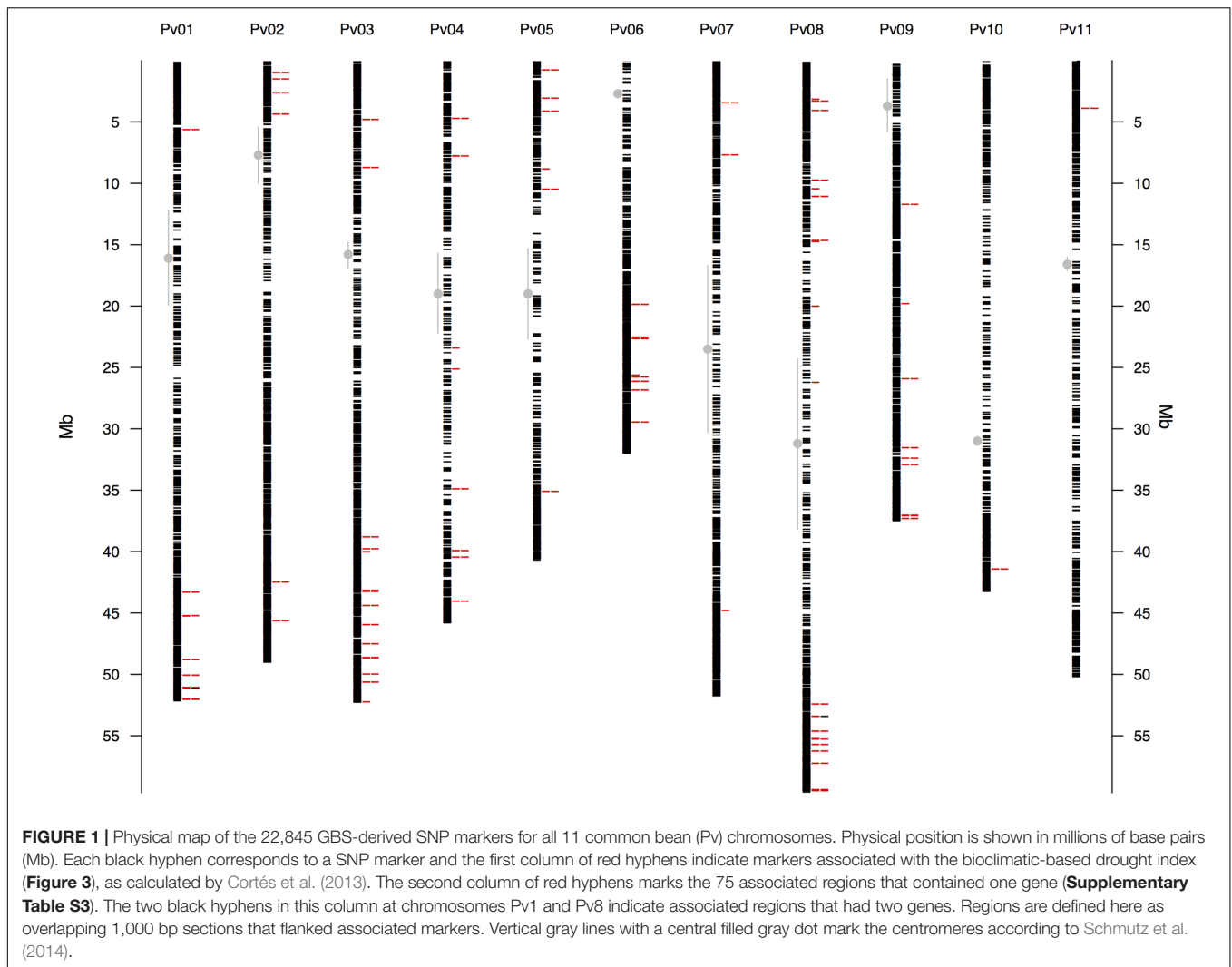
There Were 115 Associated SNPs Widespread in 90 Regions in All 11 Chromosomes

The GBS-derived SNP markers recovered the well-described Andean and Mesoamerican genepool structure and five within-genepool subpopulations, as depicted by the principal components analysis (Figure 2, compare to Blair et al., 2012). QQ-plots from the association analyses between the 22,845 SNP markers and the bioclimatic-based drought index from Cortés et al. (2013) indicated that GLM analyses likely had excessive rates of false positives (Supplementary Figure S2) whereas MLM models controlling for population structure and using a kinship matrix more effectively reduced the false positive rate (Supplementary Figure S3). The MLM model with the first two PCoA axes scores used as covariates, was the best at controlling for false positives because it exhibited the smoothest transition toward significant outliers.

This last model yielded a total of 115 SNP markers associated with the bioclimatic-based drought index at the Bonferroni-corrected significance threshold of 7.36 $-\log_{10}$ (*P*-value) (Figure 3). This group of SNPs had an average R^2 of $51.3\% \pm 0.4$, fairly stable across markers (Supplementary Table S2). The 115 SNPs were clustered in 90 different regions, defined here as overlapping 1,000 bp sections that flanked associated markers. Associated SNPs and regions were widespread in all 11 common bean chromosomes (Table 1).

Chromosomes Pv3 and Pv8 had the highest number of associated SNPs with 21 and 32 SNPs clustered in 16 and 21 different regions, respectively. Chromosomes Pv1, Pv2, Pv4, Pv5, Pv6, and Pv9 contained an intermediate number of associated SNPs with 11, 6, 11, 7, 12, and 9 SNPs clustered in 11, 6, 8, 6, 8, and 9 different regions, respectively. Chromosomes Pv7, Pv10 and Pv11 had the fewest number of associated SNPs with 3, 2, and 1 SNPs clustered in 3, 1, and 1 different regions, respectively. Chromosome Pv8 had more regions with at least two associated SNPs than any other chromosome, and these regions had more associated SNPs than in any other chromosome for a total of five regions with an average number of 3.2 associated SNPs in each region. The single region that contained the most associated SNPs was also situated on chromosome Pv8 with 6 SNPs and an average R^2 of $51.1\% \pm 0.3$. After this chromosome, Pv3 was also notable for having four regions with an average number of 2.5 associated SNPs per region.

A total of 75 regions, comprising 99 SNP markers associated with the bioclimatic-based drought index, contained at least one gene, for a total of 77 genes identified as potential candidates for drought tolerance from the wild accession analysis (Supplementary Table S3). Most genes were from chromosomes Pv1, Pv3, and Pv8 with 11, 14, and 16 genes. Only two regions, at chromosomes Pv1 and Pv8 and containing a total of seven different SNPs, spanned two or more genes. The one in Pv8 was



the region with more associated SNPs. One of the two genes in this region encoded an ankyrin repeat-containing protein, which was associated with osmotic regulation via the assembly of cation channels in the membranes (Voronin and Kiseleva, 2008). Among other identified potential candidate genes there was a phototropic-responsive NPH3 gene (Pedmale and Liscum, 2007) in Pv3. The associated SNPs flanking these genes had an average R^2 of $49\% \pm 0.2$.

Associated Genomic Windows Were Enriched for SNP Density and Positive Tajima's D Scores

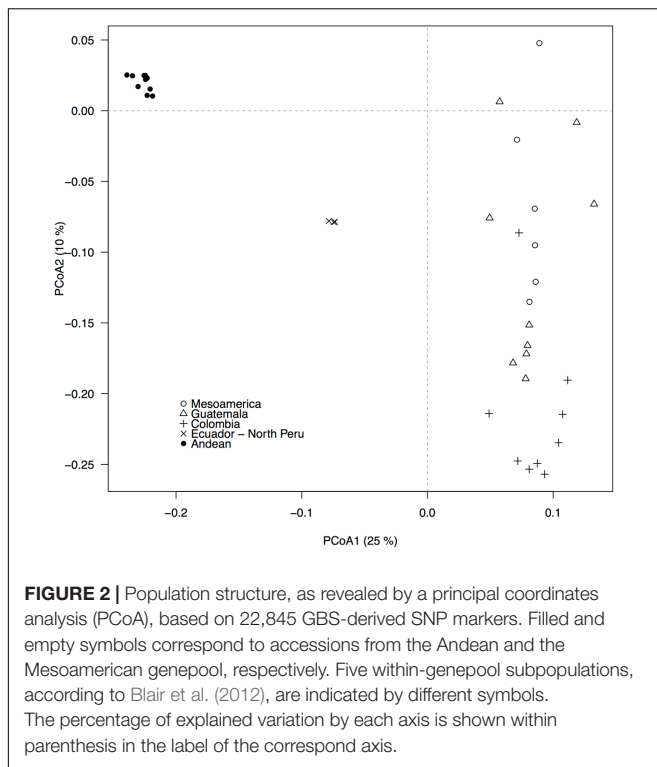
A sliding window analysis (window size = 1×10^6 bp, step size = 200 kb) was used to explore the patterns of genome-wide diversity. Marker density decayed drastically toward the centromeres. Average marker density was 44 SNPs per million base pairs (95% CI, 4–143, Figure 4A). Average nucleotide diversity as measured by π was 0.3 per million base pairs (95% CI, 0.2–0.4, Figure 4B). Average Watterson's theta (θ) was 0.20 per million base pairs (95% CI, 0.19–0.21, Figure 4C). Average

Tajima's D was 0.68 per million base pairs (95% CI, 0.05–1.22, Figure 4D).

These statistics were compared between 1 Mb sliding windows that contained (associated windows) or did not contain at least one marker that was associated with the bioclimatic-based drought index (non-associated windows). Genomic windows containing at least one associated SNP had overall higher SNP density (79 ± 6 vs. 39 ± 2 , Figure 5A), lower scores for Watterson's theta (θ) scores (0.2016 ± 0.0001 vs. 0.2026 ± 0.0001 , Figure 5C) and more positive Tajima's D scores (0.71 ± 0.02 vs. 0.678 ± 0.009 , Figure 5D) than windows without associated markers. Nucleotide diversity, as measured by π , was slightly elevated in associated windows when compared with no associated windows (0.322 ± 0.006 vs. 0.317 ± 0.003 , Figure 5B).

DISCUSSION

This research indicates that adaptation to drought in common bean, as estimated by bioclimatic variables, is widespread across all 11 common bean chromosomes and has been under natural



divergent selection (inflated Tajima's D values). Both of these results imply that the genetic sources for drought tolerance are abundant in wild common bean and that natural drought tolerance may be favorable under dry conditions but harmful in more humid conditions. The awareness about this trade-off, upon which natural divergent selection has been acting for

thousands of years, may aid the breeding of new drought tolerant varieties specifically adapted to unique micro-environments and local regions rather than varieties, eventually obsolete, originally intended for a wider range of environments. Below we first discuss the evidence that support the main conclusion of widespread divergent adaptation to drought. Later, we follow up by discussing the implications of this finding.

Wild Common Bean Exhibit Widespread Divergent Adaptation to Drought

It is well known that selective processes, such as purifying selection and local adaptation (divergent selection), differentially imprint regions within the same genome, causing a heterogeneous departure of genetic variation from the neutral expectations and from the background trend (Ellegren and Galtier, 2016). Specifically, habitat-mediated purifying selection is associated with localized low values of nucleotide diversity (π) (Nei, 1987) and Tajima's D (Tajima, 1989) and high scores of the Watterson's theta (θ) estimator (Watterson, 1975) because only low-frequency polymorphisms can avoid being eliminated by widespread directional selection. Although recent population bottlenecks tend to achieve the same reduction in nucleotide variation, this pattern is expected at a more genome-wide level. Similarly, local adaptation tends to homogenize haplotypes within the same niche and fix polymorphisms in different populations. Consequently, few haplotypes with high frequency are retained, corresponding to high values of nucleotide diversity (π) and Tajima's D, and low scores of the Watterson's theta (θ) estimator (Wakeley, 2008). While independent domestication events, extensive population structure, and population expansions after bottlenecks can produce the same patterns (Nei, 2010), these demographic processes are expected to be observed at a more genome-wide level.

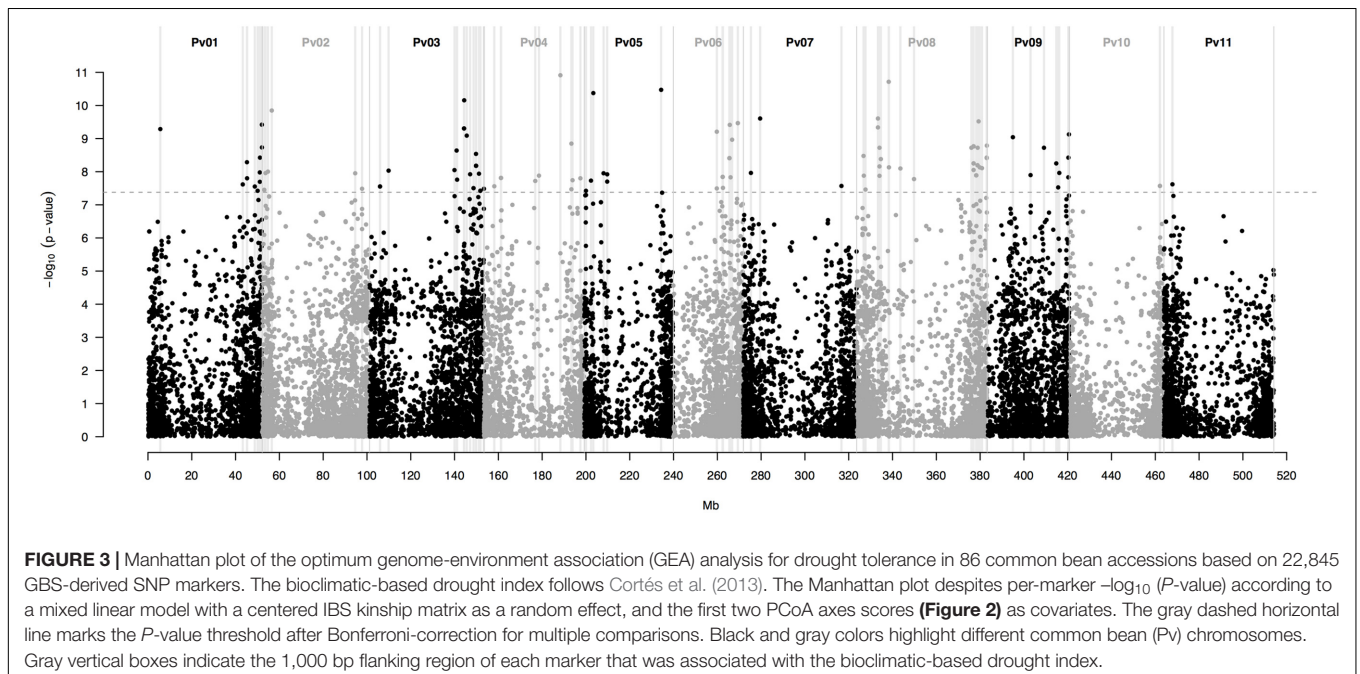
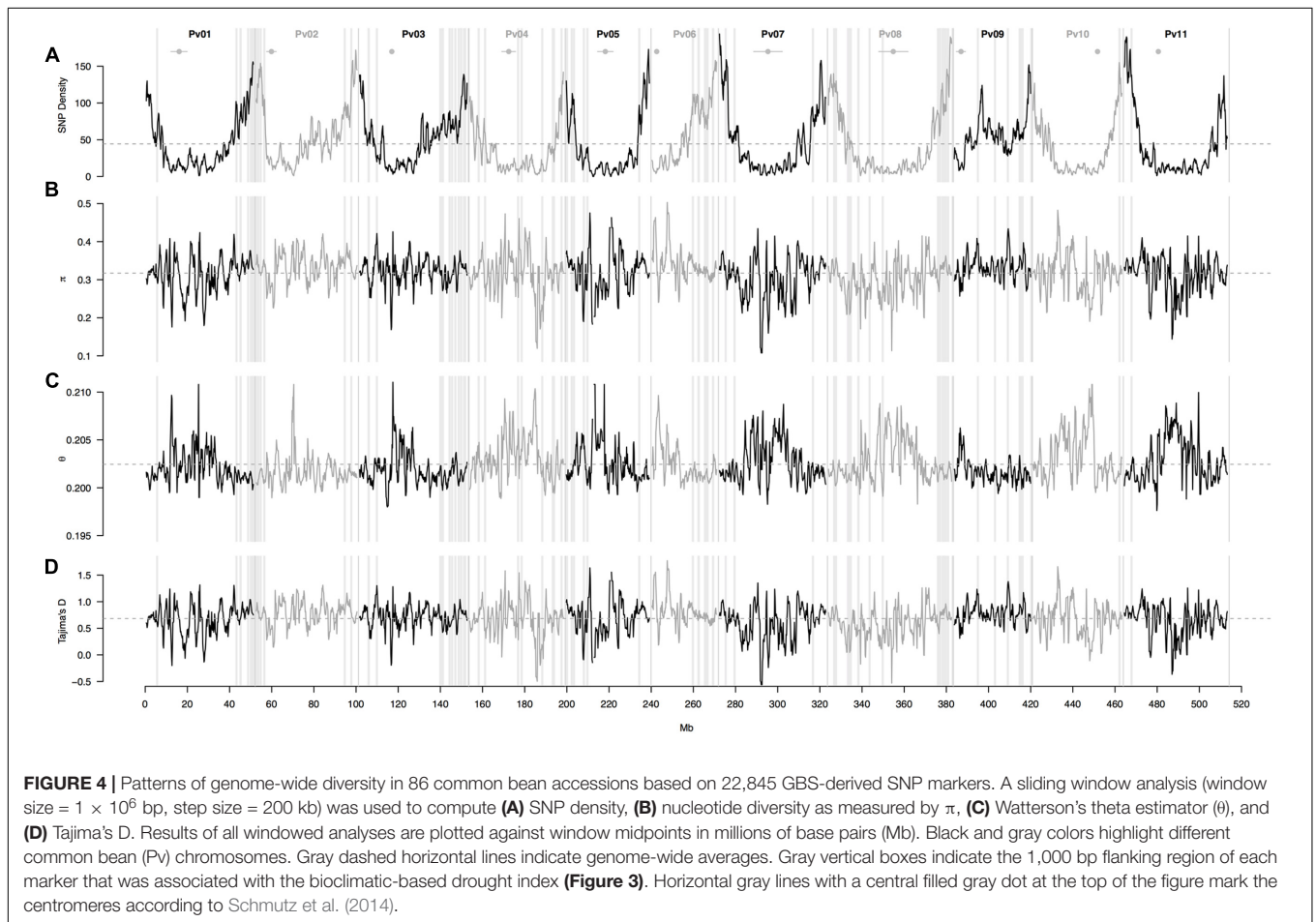


TABLE 1 | Per-chromosome (Pv) summary statistics for the 115 SNP markers associated with the bioclimatic-based drought index (Cortés et al., 2013) in 86 common bean accessions based on the optimum association analysis (Figure 3).

Pv	Number of associated SNPs	Average – \log_{10} (P-value)	Average R^2 (%)	Number of associated regions	Number of regions with more than one associated SNP	Average number of associated SNPs in regions with more than one associated SNP	Number of associated regions containing genes	Number of genes
Pv1	11	8 ± 2	51 ± 1	11	0	NA	10	11
Pv2	6	8 ± 3	50 ± 2	6	0	NA	6	6
Pv3	21	8 ± 2	51 ± 1	16	4	2.25	14	14
Pv4	11	8 ± 2	50 ± 1	8	2	2.50	6	6
Pv5	7	8 ± 3	52 ± 2	6	1	2.00	5	5
Pv6	12	8 ± 2	52 ± 1	8	4	2.00	7	7
Pv7	3	8 ± 5	50 ± 2	3	0	NA	2	2
Pv8	32	8 ± 1	52 ± 1	21	5	3.20	15	16
Pv9	9	8 ± 3	51 ± 1	9	0	NA	8	8
Pv10	2	7.6	47.8	1	1	2.00	1	1
Pv11	1	7.6	47.1	1	0	NA	1	1
Total	115	8 ± 1	51.3 ± 0.4	90	17	2.47	75	77

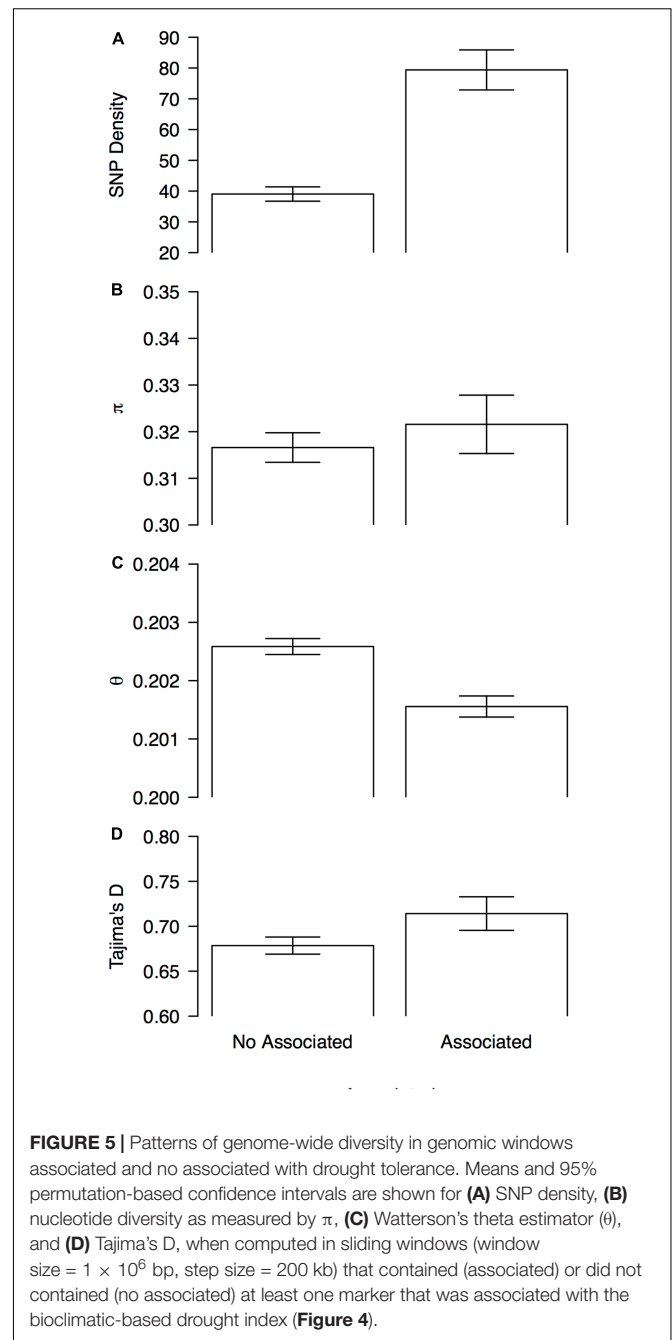
Given are results of a mixed linear model with a centered IBS kinship matrix as a random effect, and the first two PCoA axes scores (Figure 2) as covariates. Regions are defined here as overlapping 1,000 bp sections that flanked associated markers. Potential candidate genes were identified within the 1,000 bp section flanking each associated marker.



Since divergent selection tends to homogenize haplotypes within the same niche and fix polymorphisms in different populations, few haplotypes with high frequency are retained, matching high values of nucleotide diversity and Tajima's D, and low values of the Watterson's theta (θ) estimator (Wakeley, 2008). In this study we have found that the genomic regions associated with the bioclimatic-based drought index, widespread in all 11 common bean chromosomes, displayed these very same signatures: inflated Tajima's D values and lowered Watterson's theta (θ) estimator. In other words, few haplotypes with high frequency are being differentially fixed in populations coming from contrasting environments. This result speaks for divergent selection acting on the genetic variants that are associated with drought tolerance, as compared to the opposite signal of directional selection, in which few dominant haplotypes are expected.

Yet, genomic signatures associated with habitat heterogeneity can still result from causes other than adaptation and selection (Nei, 2010), for example random genetic drift and population structure, and are also influenced by differences in ancestral variation and recombination in the genome (Strasburg et al., 2011; Cortés, 2015; Cortés et al., 2015a; Ellegren and Galtier, 2016; Wolf and Ellegren, 2016). Moreover, the origin of habitat-associated variants from novel or standing genetic variation leads to distinctively different patterns of genomic divergence (Hermisson and Pennings, 2005; Barrett and Schluter, 2008; Pritchard et al., 2010). A way to distinguish these underlying causes of divergence is comparing summary statistics (i.e., Tajima's D) from different genomic sections, as we did here, because demographic processes usually leave genome-wide signatures while selection tends to imprint more localized regions (Wakeley, 2008). With this in mind, the signatures of divergent selection displayed in the genomic regions that in this study were associated with the bioclimatic-based drought index are unlikely due to confounding demographic processes such as independent domestication events, extensive population structure and population expansions after bottlenecks because the mixed linear model that we chose to identify the GEA accounted for population structure and demographic processes usually leave genome-wide signatures that should have imprinted the no associated windows as well. Therefore, the signatures displayed on the genetic variants that are associated with drought tolerance seem to reflect true divergent selection rather than confounding demographic processes like independent domestication events, extensive population structure, or population expansions after bottlenecks.

One potential caveat may concern the fact that we used GBS with the *ApeKI* restriction enzyme for the first time in wild accessions of common bean, possibly leading to problems like missing data due to sequence divergence compared to the cultivated reference genome. Nonetheless, wild-cultivated divergent regions with a high proportion of missing data seem to be rare because the U-shaped pattern of SNP density is just as expected when using a non-methylation sensitive enzyme. This decay in diversity proportional to the decay in the rate of recombination (toward the centromeres) was first described in *D. melanogaster* and has been confirmed in many organisms since



then (Ellegren and Galtier, 2016). The pattern was understood as an effect of genetic hitchhiking, but background selection may also be a contributing factor, perhaps even the dominating one.

It is also appropriate to clarify that the power to detect marker-trait associations in this study was not limited by the number of markers. Every three SNPs are redundant and replaceable given a criterion of linkage disequilibrium, or LD (Carlson et al., 2004), which is higher in selfing plants (Slatkin, 2008; Rossi et al., 2009), like common bean (Broughton et al., 2003; Rossi et al., 2009; Blair et al., 2018). At least in humans, a 0.8 R^2 threshold is used to identify maximally informative SNPs for association analyses by

reducing the redundancy due to LD (Carlson et al., 2004), and this could apply to other species as well (Slate et al., 2009). Here, average marker density was 44 SNPs per million base pairs, or average distance between any two SNP markers in the reference genome was 23 thousand base pairs. Since LD in wild common bean, measured as marker correlation R^2 , was reported to decay to 0.8 per every 81 thousand base pairs (Rossi et al., 2009; Blair et al., 2018), on average every 3, physically linked, GBS-derived SNP markers are in sufficient LD to be considered redundant by the association analysis.

Finally, despite that the modest sample size used in the association analysis may overlook the majority of associations with low effect sizes and large-effect genes that segregate at low frequency (Maher, 2008), genes with major effects that segregate at moderate frequencies are still identifiable. In fact, this may be the case of a gene coding for an ankyrin repeat-containing protein, associated with osmotic regulation (Voronin and Kiseleva, 2008), and a phototropic-responsive NPH3 gene (Pedmale and Liscum, 2007), found in regions in chromosomes Pv8 and Pv3 flanking associated SNPs with an average R^2 of 49%.

Wild Accessions May Be Useful to Breed Drought Tolerant Varieties in Common Bean

Previous research from Tiranti and Negri (2007) and Ariani et al. (2017) demonstrated that selective micro-environmental effects play a role in shaping genetic diversity and structure in common bean wild accessions. In this study, we have confirmed that ecological gradients related with drought stress are associated with divergent selection in wild common beans, after accounting for genepool and subpopulation structure. This divergent selective pressure might be a consequence of local level rainfall patterns. Specifically, in tropical environments near the equator with bimodal rainfall a mid-season dry period occurs that can last 2–4 weeks. In contrast in the sub-tropics, a dry period of three or more months can occur. In response to this mid-cycle drought of the sub-tropics, *P. vulgaris* enters a survival mode of slow growth and reduced physiological activity until rainfall resumes and flowering occurs (Beebe et al., 2008). Beans growing in wetter conditions on the other hand are less frequently subjected to these environmental pressures and have a fitness advantage to mature in a shorter length of time. Given these ecological differences, the reaction typically associated with drought tolerance although favorable under dry conditions seems detrimental under more humid conditions, which is consistent with the genomic signature of divergent selection observed in this study. This trade-off must be accounted for when breeding for drought tolerance, reinforcing the need of varieties locally adapted to unique micro-environments and narrow regions instead of varieties intended for a wider range of environments (Cortés and Blair, 2018a,b).

Furthermore, wild accessions of common bean occupy more geographical regions with extensive drought stress than cultivated accessions (Cortés et al., 2013). Those regions include the arid areas of Peru, Bolivia, and Argentina, and the valleys

of northwest Mexico. Hence, a broad habitat distribution for wild common bean has exposed these genotypes to both dry and wetter conditions, while cultivated common bean has a narrower distribution and is traditionally considered susceptible to drought. These differences in the ecologies of wild and cultivated common bean have been associated with higher genetic diversity in the former group when surveying candidate genes for drought tolerance such as the ASR (Cortés et al., 2012a), DREB (Cortés et al., 2012b), and ERECTA (Blair et al., 2016) gene families, once population structure (Blair et al., 2012) and the background distribution of genetic diversity have been accounted for.

Although not all associated markers may be causal, but rather linked with causal elements, in this study we were able to identify a wide genetic source for drought tolerance, including new potential candidates like a gene coding for an ankyrin repeat-containing protein (Voronin and Kiseleva, 2008) and a phototropic-responsive NPH3 gene (Pedmale and Liscum, 2007). We were also able to identify some further differences between the adaptations of wild accessions found in arid and more humid environments, which may be valuable for plant breeding. Therefore, we reinforce, as was envisioned by Acosta et al. (2007), that wild accessions and landraces of common bean be taken into account to exploit naturally available divergent variation for drought tolerance.

PERSPECTIVES

The potential candidate genes identified in this study should be validated as candidates for drought tolerance through re-sequencing and controlled drought stress treatments under green house and field conditions, ideally in a diverse panel of common bean accessions and closely related species, such as Tepary bean (*P. acutifolius*) that is known for growing in desert and semi-arid environments. Our group is currently running this initiative as part of a larger project spanning several other candidate genes, such as DREB, ASR, and ERECTA (Cortés et al., 2012a,b; Blair et al., 2016).

The work presented here ultimately illustrates that genomic signatures of adaptation are useful for germplasm characterization, potentially enhancing future marker-assisted selection and crop improvement. We envision that GEA studies coupled with estimates of genome-wide diversity will become more common in the oncoming years. They both might allow assessing the naturally available genetic variability for adaptation to various types of stress that, like drought, are expected to worsen with climate change. For instance, frost (Wheeler et al., 2014, 2016), nutrient limitation (Sedlacek et al., 2014; Little et al., 2016), altered snowmelt (Cortés et al., 2015b; Cortés, 2016, 2017a), distorted biotic interactions (Wheeler et al., 2015), and altered growing (Sedlacek et al., 2015, 2016) and flowering (Cortés et al., 2014) seasons, can easily be modeled and associated with allelic variants using a genome-wide diversity background. GEA studies coupled with estimates of genome-wide diversity could also be expandable to heterogeneous collections that include a mix of commercial

accessions, landraces, wild populations and wild relatives, and that span a variety of ecologically diverse managed and natural ecosystems (Madriñán et al., 2013; Cortés and Wheeler, 2018) screened with a wide range of genotyping techniques (Cortés et al., 2011; Galeano et al., 2012; Kelleher et al., 2012; Blair et al., 2013). Genomic selection models, which are becoming increasingly popular (Desta and Ortiz, 2014), in conjunction with a wide-spectrum of stochastic models (Cortés, 2017b, 2018), could even incorporate at some point environmental variables and estimates of genome-wide diversity in order to improve the prediction of phenotypic variance and the estimation of the genotype \times environment interactions.

DATA ACCESSIBILITY

The pipeline configuration file and SNP dataset are archived at the Dryad Digital Repository under doi: 10.5061/dryad.j2c24.

AUTHOR CONTRIBUTIONS

AC designed the study with insights from MB. AC collected and analyzed the data. AC wrote the manuscript with contributions from MB.

FUNDING

This research was supported by the Lundell and Tullberg grants to AC as PI. The Geneco mobility fund from Lund University is acknowledged for making possible the synergistic kickoff meeting between AC and MB in the spring of 2015 at Nashville, TN, United States. Writing time by the first author was possible thanks to the grants CTS14:55 from the Carl Trygger Foundation to S. Berlin, and 4.1-2016-00418 from Vetenskapsrådet (VR) and BS2017-0036 from Kungliga Vetenskapsakademien (KVA) to AC, and by MB thanks to the Evans Allen fund from the United States Department of Agriculture. The publication fund of the Colombian Corporation for Agricultural Research is very much appreciated for supporting the publication of this work.

ACKNOWLEDGMENTS

We are grateful to Daniel G. Debouck and the Genetic Resources Unit at the International Center for Tropical Agriculture for donating the seeds and sharing the geographic coordinates of the plant material that was used in this study. Reassurance from J. M. Osorio toward this work is very much appreciated.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2018.00128/full#supplementary-material>

FIGURE S1 | Boxplot and histogram of the bioclimatic-based drought index (DI), according to Cortés et al. (2013), for the 86 common bean accessions used in this study.

FIGURE S2 | Genome–environment association (GEA) analyses for drought tolerance according to five different general linear models (GLMs) in 86 common bean accessions based on 22,845 GBS-derived SNP markers. Given are Manhattan plots (**A,C,E,G,I**) and QQ-plots (**B,D,F,H,J**) of $-\log_{10}$ (P -value) for the following models: (**A,B**) GLM with the genepool identity and the first two PCoA axes scores (**Figure 2**) as covariates; (**C,D**) GLM with the within-genepool subpopulation identity, according to Blair et al. (2012), and the first two PCoA axes scores (**Figure 2**) as covariates; (**E,F**) GLM with the first two PCoA axes scores (**Figure 2**) as covariates; (**G,H**) GLM with the within-genepool subpopulation identity, according to Blair et al. (2012), as covariate; and (**I,J**) GLM with the genepool identity as covariate. The bioclimatic-based drought index follows Cortés et al. (2013). The gray dashed horizontal line marks the P -value threshold after Bonferroni-correction for multiple comparisons. Black and gray colors highlight different common bean (Pv) chromosomes.

FIGURE S3 | Genome–environment association (GEA) analyses for drought tolerance according to five different mixed linear models (MLMs) in 86 common bean accessions based on 22,845 GBS-derived SNP markers. Given are Manhattan plots (**A,C,E,G,I**) and QQ-plots (**B,D,F,H,J**) of $-\log_{10}$ (P -value) for the following models: (**A,B**) MLM with the genepool identity and the first two PCoA axes scores (**Figure 2**) as covariates; (**C,D**) MLM with the within-genepool subpopulation identity, according to Blair et al. (2012), and the first two PCoA axes scores (**Figure 2**) as covariates; (**E,F**) MLM with the first two PCoA axes scores (**Figure 2**) as covariates (model shown in **Figure 3**); (**G,H**) MLM with the within-genepool subpopulation identity, according to Blair et al. (2012), as covariate; and (**I,J**) MLM with the genepool identity as covariate. All models used a centered IBS kinship matrix as a random effect. The bioclimatic-based drought index follows Cortés et al. (2013). The gray dashed horizontal line marks the P -value threshold after Bonferroni-correction for multiple comparisons. Black and gray colors highlight different common bean (Pv) chromosomes.

TABLE S1 | Identity of the 86 common bean accessions used in this study. The G identification number (from the Genetic Resources Unit at the International Center for Tropical Agriculture) and the country of origin are shown according to the following convention: ARG for Argentina, BOL for Bolivia, COL for Colombia, ECU for Ecuador, GTM for Guatemala, HND for Honduras, MEX for Mexico and PER for Peru. The bioclimatic-based drought index (DI) according to Cortés et al. (2013) is also shown (and in **Supplementary Figure S1**). It appears in red if the accession was found in a dry environment (high drought index, above zero) or in blue if it was found in a very wet habitat (low drought index, below zero).

TABLE S2 | Statistics from the optimum association analysis for the 115 SNP markers associated with the bioclimatic-based drought index in 86 common bean accessions based on 22,845 GBS-derived SNP markers. Given are results of a mixed linear model (depicted in **Figure 3**) with a centered IBS kinship matrix as a random effect, and the first two PCoA axes scores (**Figure 2**) as covariates. Among other statistics, the F -value, the $-\log_{10}$ (P -value) and the variation explained (R^2) by each marker, are shown. Marker names are given as chromosome and position on the common bean reference genome from Schmutz et al. (2014). Unique regions containing one or more SNPs are named with a consecutive number per chromosome under the column called “region.” Gray cells in this column mark regions with two or more associated SNPs. Regions are defined here as overlapping 1,000 bp sections that flanked associated markers.

TABLE S3 | List of 77 genes in the 75 associated regions that contained at least one gene. The position of each gene and its functional description are shown together with other identifiers and descriptors. The list of associated SNP markers surrounding each gene is also depicted, for a total of 99 SNPs. Genes co-occurring in the same region are marked with gray cell under the column “associated SNPs.” Regions are defined here as overlapping 1,000 bp sections that flanked associated markers.

REFERENCES

- Acosta, J. A., Kelly, J. D., and Gepts, P. (2007). Prebreeding in common bean and use of genetic diversity from wild germplasm. *Crop Sci.* 47, S44–S59. doi: 10.2135/cropsci2007.04.0008IPBS
- Ariani, A., Berny, J. C., and Gepts, P. (2016). Genome-wide identification of SNPs and copy number variation in common bean (*Phaseolus vulgaris* L.) Using Genotyping-by-Sequencing (Gbs). *Mol. Breed.* 36:87. doi: 10.1007/s11032-016-0512-9
- Ariani, A., Berny Mier, Y. T. J., and Gepts, P. (2017). Spatial and temporal scales of range expansion in wild *Phaseolus vulgaris*. *Mol. Biol. Evol.* 35, 119–131. doi: 10.1093/molbev/msx273
- Barrett, R. D., and Schluter, D. (2008). Adaptation from standing genetic variation. *Trends Ecol. Evol.* 23, 38–44. doi: 10.1016/j.tree.2007.09.008
- Beebe, S., Rao, I. M., Cajiao, C., and Grajales, M. (2008). Selection for drought resistance in common bean also improves yield in phosphorus limited and favorable environments. *Crop Sci.* 48, 582–592. doi: 10.2135/cropsci2007.07.0404
- Berthouly-Salazar, C., Thuillet, A. C., Rhone, B., Mariac, C., Ousseini, I. S., Couderc, M., et al. (2016). Genome scan reveals selection acting on genes linked to stress response in wild pearl millet. *Mol. Ecol.* 25, 5500–5512. doi: 10.1111/mec.13859
- Bhakta, M. S., Jones, V. A., and Vallejos, C. E. (2015). Punctuated distribution of recombination hotspots and demarcation of pericentromeric regions in *Phaseolus vulgaris* L. *PLoS One* 10:e0116822. doi: 10.1371/journal.pone.0116822
- Blair, M. W., Cortés, A. J., Farmer, A. R., Wei, H., Ambachew, D., Penmetsa, V., et al. (2018). Uneven recombination rate and linkage disequilibrium across a reference SNP map for common bean (*Phaseolus vulgaris* L.). *PLoS One*. doi: 10.1371/journal.pone.0189597
- Blair, M. W., Cortés, A. J., Penmetsa, R. V., Farmer, A., Carrasquilla-Garcia, N., and Cook, D. R. (2013). A high-throughput Snp marker system for parental polymorphism screening, and diversity analysis in common bean (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* 126, 535–548. doi: 10.1007/s00122-012-1999-z
- Blair, M. W., Cortés, A. J., and This, D. (2016). Identification of an *Erecta* gene and its drought adaptation associations with wild and cultivated common bean. *Plant Sci.* 242, 250–259. doi: 10.1016/j.plantsci.2015.08.004
- Blair, M. W., Soler, A., and Cortés, A. J. (2012). Diversification and population structure in common beans (*Phaseolus vulgaris* L.). *PLoS One* 7:e49488. doi: 10.1371/journal.pone.0049488
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). Tassel: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308
- Broughton, W. J., Hernandez, G., Blair, M., Beebe, S., Gepts, P., and Vanderleyden, J. (2003). Beans (*Phaseolus* Spp.) - model food legumes. *Plant Soil* 252, 55–128. doi: 10.1023/A:1024146710611
- Carlson, C. S., Eberle, M. A., Rieder, M. J., Yi, Q., Kruglyak, L., and Nickerson, D. A. (2004). Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am. J. Hum. Genet.* 74, 106–120. doi: 10.1086/381000
- Cortés, A. J. (2013). On the origin of the common bean (*Phaseolus vulgaris* L.). *Am. J. Plant Sci.* 4, 1998–2000. doi: 10.4236/ajps.2013.410248
- Cortés, A. J. (2015). *On the Big Challenges of a Small Shrub: Ecological Genetics of Salix Herbacea L.* Uppsala: Acta Universitatis Upsaliensis.
- Cortés, A. J. (2016). *Environmental Heterogeneity at a Fine Scale: Ecological and Genetic Implications in a Changing World.* Saarbrücken: LAP Lambert Academic Publishing.
- Cortés, A. J. (2017a). “Local scale genetic diversity and its role in coping with changing climate,” in *Genetic Diversity*, ed. L. Bitz (Rijeka: InTech), 140. doi: 10.5772/67166
- Cortés, A. J. (2017b). On how role versatility boosts an STI. *J. Theor. Biol.* 440, 66–69. doi: 10.1016/j.jtbi.2017.12.018
- Cortés, A. J. (2018). “Prevalence in MSM is enhanced by role versatility,” in *Big Data Analytics in HIV/AIDS Research*, ed. A. Mazari (Hershey: IGI Global), 250.
- Cortés, A. J., and Blair, M. W. (2018a). “Lessons from common bean on how wild relatives and landraces can make tropical crops more resistant to climate change,” in *Rediscovery of Landraces as a Resource for the Future*, ed. O. Grillo (Rijeka: InTech), 165.
- Cortés, A. J., and Blair, M. W. (2018b). “Naturally available genetic adaptation in Common bean and its response to climate change,” in *Climate Resilient Agriculture – Strategies and Perspectives*, eds C. Srinivasarao, A. K. Shanker, and C. Shanker (Rijeka: InTech), 133. doi: 10.5772/intechopen.72380
- Cortés, A. J., Chavarro, M. C., and Blair, M. W. (2011). Snp marker diversity in common bean (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* 123, 827–845. doi: 10.1007/s00122-011-1630-8
- Cortés, A. J., Chavarro, M. C., Madriñán, S., This, D., and Blair, M. W. (2012a). Molecular ecology and selection in the drought-related *Asr* gene polymorphisms in wild and cultivated common bean (*Phaseolus vulgaris* L.). *BMC Genet.* 13:58. doi: 10.1186/1471-2156-13-58
- Cortés, A. J., Liu, X., Sedlacek, J., Wheeler, J. A., Lexer, C., and Karrenberg, S. (2015a). “Maintenance of female-bias in a polygenic sex determination system is consistent with genomic conflict,” in *On the Big Challenges of a Small Shrub: Ecological Genetics of Salix herbacea L.*, ed. A. J. Cortés (Uppsala: Acta Universitatis Upsaliensis).
- Cortés, A. J., Monserrate, F., Ramírez-Villegas, J., Madriñán, S., and Blair, M. W. (2013). Drought tolerance in wild plant populations: the case of common beans (*Phaseolus vulgaris* L.). *PLoS One* 8:e62898. doi: 10.1371/journal.pone.0062898
- Cortés, A. J., This, D., Chavarro, C., Madriñán, S., and Blair, M. W. (2012b). Nucleotide diversity patterns at the drought-related *Dreb2* encoding genes in wild and cultivated common bean (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* 125, 1069–1085. doi: 10.1007/s00122-012-1896-5
- Cortés, A. J., Waeber, S., Lexer, C., Sedlacek, J., Wheeler, J. A., Van Kleunen, M., et al. (2014). Small-scale patterns in snowmelt timing affect gene flow and the distribution of genetic diversity in the alpine dwarf shrub *Salix herbacea*. *Heredity* 113, 233–239. doi: 10.1038/hdy.2014.19
- Cortés, A. J., and Wheeler, J. A. (2018). “The environmental heterogeneity of mountains at a fine scale in a changing world,” in *Mountains, Climate, and Biodiversity*, eds C. Hoorn, A. Perrigo, and A. Antonelli (Hoboken: Wiley), 584.
- Cortés, A. J., Wheeler, J. A., Sedlacek, J., Lexer, C., and Karrenberg, S. (2015b). “Genome-wide patterns of microhabitat-driven divergence in the alpine dwarf shrub *Salix herbacea* L.,” in *On the Big Challenges of a Small Shrub: Ecological Genetics of Salix herbacea L.*, ed. A. J. Cortés (Uppsala: Acta Universitatis Upsaliensis).
- Des Marais, D. L., Auchincloss, L. C., Sukamtoh, E., McKay, J. K., Logan, T., Richards, J. H., et al. (2014). Variation in *Mpk12* affects water use efficiency in *Arabidopsis* and reveals a pleiotropic link between guard cell size and ABA response. *Proc. Natl. Acad. Sci. U.S.A.* 111, 2836–2841. doi: 10.1073/pnas.1321429111
- Desti, Z. A., and Ortiz, R. (2014). Genomic selection: genome-wide prediction in plant improvement. *Trends Plant Sci.* 19, 592–601. doi: 10.1016/j.tplants.2014.05.006
- Ellegren, H., and Galtier, N. (2016). Determinants of genetic diversity. *Nat. Rev. Genet.* 17, 422–433. doi: 10.1038/nrg.2016.58
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6:e19379. doi: 10.1371/journal.pone.0019379
- Fischer, M. C., Rellstab, C., Tedder, A., Zoller, S., Gugerli, F., Shimizu, K. K., et al. (2013). Population genomic footprints of selection and associations with climate in natural populations of *Arabidopsis halleri* from the Alps. *Mol. Ecol.* 22, 5594–5607. doi: 10.1111/mec.12521
- Forester, B. R., Jones, M. R., Joost, S., Landguth, E. L., and Lasky, J. R. (2016). Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. *Mol. Ecol.* 25, 104–120. doi: 10.1111/mec.13476
- Galeano, C. H., Cortés, A. J., Fernandez, A. C., Soler, A., Franco-Herrera, N., Makunde, G., et al. (2012). Gene-based single nucleotide polymorphism markers for genetic and association mapping in common bean. *BMC Genet.* 13:48. doi: 10.1186/1471-2156-13-48
- Gepts, P., and Deboucq, D. (1991). “Origin, domestication and evolution of the common bean (*Phaseolus vulgaris* L.),” in *Common Beans: Research for Crop Improvement*, eds A. Van Shoonhoven and O. Voysest (Wallingford: Commonwealth Agricultural Bureau), 7–53.

- Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., et al. (2014). Tassel-Gbs: a high capacity genotyping by sequencing analysis pipeline. *PLoS One* 9:e90346. doi: 10.1371/journal.pone.0090346
- Hamon, W. R. (1961). Estimating potential evapotranspiration. *Proc. Am. Soc. Civil Eng.* 87, 107–120.
- Hancock, A. M., Brachi, B., Faure, N., Horton, M. W., Jarymowycz, L. B., Sperone, F. G., et al. (2011). Adaptation to climate across the *Arabidopsis thaliana* Genome. *Science* 334, 83–86. doi: 10.1126/science.1209244
- Hart, J. P., and Griffiths, P. D. (2015). Genotyping-by-sequencing enabled mapping and marker development for the Potyvirus resistance allele in common bean. *Plant Genome* 8. doi: 10.3835/plantgenome2014.09.0058
- Hermisson, J., and Pennings, P. S. (2005). Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* 169, 2335–2352. doi: 10.1534/genetics.104.036947
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* 25, 1965–1978. doi: 10.1002/joc.1276
- Hoffmann, A. A., and Sgro, C. M. (2011). Climate change and evolutionary adaptation. *Nature* 470, 479–485. doi: 10.1038/nature09670
- Kelleher, C. T., Wilkin, J., Zhuang, J., Cortés, A. J., Quintero, Á. L. P., Gallagher, T. F., et al. (2012). Snp discovery, gene diversity, and linkage disequilibrium in wild populations of *Populus tremuloides*. *Tree Genet. Genomes* 8, 821–829. doi: 10.1007/s11295-012-0467-x
- Kwak, M., and 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. doi: 10.1007/s00122-008-0955-4
- Lasky, J. R., Des Marais, D. L., McKay, J. K., Richards, J. H., Juenger, T. E., and Keitt, T. H. (2012). Characterizing genomic variation of *Arabidopsis thaliana*: the roles of geography and climate. *Mol. Ecol.* 21, 5512–5529. doi: 10.1111/j.1365-294X.2012.05709.x
- Lasky, J. R., Upadhyaya, H. D., Ramu, P., Deshpande, S., Hash, C. T., Bonnette, J., et al. (2015). Genome-environment associations in sorghum landraces predict adaptive traits. *Sci. Adv.* 1:e1400218. doi: 10.1126/sciadv.1400218
- Li, H., and Durbin, R. (2007). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324
- Little, C. J., Wheeler, J. A., Sedlacek, J., Cortés, A. J., and Rixen, C. (2016). Small-scale drivers: the importance of nutrient availability and snowmelt timing on performance of the alpine shrub *Salix herbacea*. *Oecologia* 180, 1015–1024. doi: 10.1007/s00442-015-3394-3
- Madrinán, S., Cortés, A. J., and Richardson, J. E. (2013). Páramo is the world's fastest evolving and coolest biodiversity hotspot. *Front. Genet.* 4:192. doi: 10.3389/fgene.2013.00192
- Maher, B. (2008). Missing Heritability. *Nature* 456, 18–21.
- Nei, M. (1987). *Molecular Evolutionary Genetics*. New York, NY: Columbia University Press.
- Nei, M. (2010). The neutral theory of molecular evolution in the genomic era. *Annu. Rev. Genomics Hum. Genet.* 11, 265–289. doi: 10.1146/annurev-genom-082908-150129
- Pedmale, U. V., and Liscum, E. (2007). Regulation of phototropic signaling in *Arabidopsis* via phosphorylation state changes in the phototropin 1-interacting protein Nph3. *J. Biol. Chem.* 282, 19992–20001. doi: 10.1074/jbc.M702551200
- Pluess, A. R., Frank, A., Heiri, C., Lalague, H., Vendramin, G. G., and Oddou-Muratorio, S. (2016). Genome-environment association study suggests local adaptation to climate at the regional scale in *Fagus sylvatica*. *New Phytol.* 210, 589–601. doi: 10.1111/nph.13809
- Pritchard, J. K., Pickrell, J. K., and Coop, G. (2010). The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Curr. Biol.* 20, R208–R215. doi: 10.1016/j.cub.2009.11.055
- Rendón-Anaya, M., Montero-Vargas, J. M., Saburido-aiLvarez, S., Vlasova, A., Capella-Gutierrez, S., Ordaz-Ortiz, J. J., et al. (2017). Genomic history of the origin and domestication of common bean unveils its closest Sister Species. *Genome Biol.* 18:60. doi: 10.1186/s13059-017-1190-6
- Rosenberg, N. A., Huang, L., Jewett, E. M., Szpiech, Z. A., Jankovic, I., and Boehnke, M. (2010). Genome-wide association studies in diverse populations. *Nat. Rev. Genet.* 11, 356–366. doi: 10.1038/nrg2760
- Rossi, M., Bitocchi, E., Bellucci, E., Nanni, L., Rau, D., Attene, G., et al. (2009). Linkage disequilibrium and population structure in wild and domesticated populations of *Phaseolus vulgaris* L. *Evol. Appl.* 2, 504–522. doi: 10.1111/j.1752-4571.2009.00082.x
- Savolainen, O., Lascoux, M., and Merilä, J. (2013). Ecological genomics of local adaptation. *Nat. Rev. Genet.* 14, 807–820. doi: 10.1038/nrg3522
- Schmutz, J., McClean, P. E., Mamidi, S., Wu, G. A., Cannon, S. B., Grimwood, J., et al. (2014). A reference genome for common bean and genome-wide analysis of dual domestications. *Nat. Genet.* 46, 707–713. doi: 10.1038/ng.3008
- Schroder, S., Mamidi, S., Lee, R., Mckain, M. R., McClean, P. E., and Osorno, J. M. (2016). Optimization of genotyping by sequencing (Gbs) data in common bean (*Phaseolus vulgaris* L.). *Mol. Breed.* 36:6. doi: 10.1007/s11032-015-0431-1
- Sedlacek, J., Bossdorf, O., Cortés, A. J., Wheeler, J. A., and Van-Kleunen, M. (2014). What role do plant-soil interactions play in the habitat suitability and potential range expansion of the alpine dwarf shrub *Salix herbacea*? *Basic Appl. Ecol.* 15, 305–315. doi: 10.1016/j.baec.2014.05.006
- Sedlacek, J., Cortés, A. J., Wheeler, J. A., Bossdorf, O., Hoch, G., Klapste, J., et al. (2016). Evolutionary potential in the alpine: trait heritabilities and performance variation of the dwarf willow *Salix herbacea* from different elevations and microhabitats. *Ecol. Evol.* 6, 3940–3952. doi: 10.1002/ece3.2171
- Sedlacek, J., Wheeler, J. A., Cortés, A. J., Bossdorf, O., Hoch, G., Lexer, C., et al. (2015). The response of the alpine dwarf shrub *Salix herbacea* to altered snowmelt timing: lessons from a multi-site transplant experiment. *PLoS One* 10:e0122395. doi: 10.1371/journal.pone.0122395
- Slate, J., Gratten, J., Beraldi, D., Stapley, J., Hale, M., and Pemberton, J. M. (2009). Gene mapping in the wild with SNPs: guidelines and future directions. *Genetica* 136, 97–107. doi: 10.1007/s10709-008-9317-z
- Slatkin, M. (2008). Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. *Nat. Rev. Genet.* 9, 477–485. doi: 10.1038/nrg2361
- Stinchcombe, J. R., and Hoekstra, H. E. (2008). Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity* 100, 158–170. doi: 10.1038/sj.hdy.6800937
- Strasburg, J. L., Sherman, N. A., Wright, K. M., Moyle, L. C., Willis, J. H., and Rieseberg, L. H. (2011). What can patterns of differentiation across plant genomes tell us about adaptation and speciation? *Philos. Trans. R. Soc. B Biol. Sci.* 367, 364–373. doi: 10.1098/rstb.2011.0199
- Tai, A. P. K., Martin, M. V., and Heald, C. L. (2014). Threat to future global food security from climate change and ozone air pollution. *Nat. Clim. Change* 4, 817–821. doi: 10.1038/nclimate2317
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595.
- Thornthwaite, C. W., and Mather, J. R. (1957). Instructions and tables for computing potential evapotranspiration and the water balance. *Climatology* 10, 185–311.
- Tiranti, B., and Negri, V. (2007). Selective microenvironmental effects play a role in shaping genetic diversity and structure in a *Phaseolus vulgaris* L. Landrace: implications for on-farm conservation. *Mol. Ecol.* 16, 4942–4955. doi: 10.1111/j.1365-294X.2007.03566.x
- Tohme, J., González, O., Beebe, S., and Duque, M. C. (1996). AFLP analysis of gene pools of a wild bean core collection. *Crop Sci.* 36, 1375–1384. doi: 10.2135/cropsci1996.0011183X003600050048x
- Turner, T. L., Bourne, E. C., Von Wettberg, E. J., Hu, T. T., and Nuzhdin, S. V. (2010). Population resequencing reveals local adaptation of *Arabidopsis lyrata* to serpentine soils. *Nat. Genet.* 42, 260–263. doi: 10.1038/ng.515
- Valdisser, P. A. M. R., Pereira, W. J., Almeida, J. E., Muñller, B. S. F., Coelho, G. R. C., and de Menezes, I. P. P. (2017). In-depth genome characterization of a Brazilian common bean core collection using dartseq high-density Snp genotyping. *BMC Genomics* 18:423. doi: 10.1186/s12864-017-3805-4
- Voronin, D. A., and Kiseleva, E. V. (2008). Functional role of proteins containing ankyrin repeats. *Cell Tissue Biol.* 2, 1–12. doi: 10.1134/S1990519X0801001X
- Wakeley, J. (2008). *Coalescent Theory: An Introduction*. Cambridge: Harvard University.
- Watterson, G. A. (1975). Number of segregating sites in genetic models without recombination. *Theor. Popul. Biol.* 7, 256–276. doi: 10.1016/0040-5809(75)90020-9
- Wheeler, J. A., Cortés, A. J., Sedlacek, J., Karrenberg, S., Van Kleunen, M., Wipf, S., et al. (2016). The snow and the willows: accelerated spring snowmelt

- reduces performance in the low-lying alpine shrub *Salix herbacea*. *J. Ecol.* 104, 1041–1050. doi: 10.1111/1365-2745.12579
- Wheeler, J. A., Hoch, G., Cortés, A. J., Sedlacek, J., Wipf, S., and Rixen, C. (2014). Increased spring freezing vulnerability for alpine shrubs under early snowmelt. *Oecologia* 175, 219–229. doi: 10.1007/s00442-013-2872-8
- Wheeler, J. A., Schnider, F., Sedlacek, J., Cortés, A. J., Wipf, S., Hoch, G., et al. (2015). With a little help from my friends: community facilitation increases performance in the dwarf shrub *Salix herbacea*. *Basic Appl. Ecol.* 16, 202–209. doi: 10.1016/j.baae.2015.02.004
- Wolf, J. B., and Ellegren, H. (2016). Making sense of genomic islands of differentiation in light of speciation. *Nat. Rev. Genet.* 18, 87–100. doi: 10.1038/nrg.2016.133
- Yeaman, S., Kathryn, A. H., Katie, E., Lotterhos, K. E., Suren, H., and Nadeau, S. (2016). Convergent local adaptation to climate in distantly related conifers. *Science* 353, 1431–1433. doi: 10.1126/science.aaf7812
- Yoder, J. B., Stanton-Geddes, J., Zhou, P., Briskine, R., Young, N. D., and Tiffin, P. (2014). Genomic signature of adaptation to climate in *Medicago truncatula*. *Genetics* 196, 1263–1275. doi: 10.1534/genetics.113.159319
- Zhang, Z., Ersoz, E., Lai, C. Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., et al. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* 42, 355–360. doi: 10.1038/ng.546
- Zou, X., Shi, C., Austin, R. S., Merico, D., Munholland, S., Marsolais, F., et al. (2013). Genome-wide single nucleotide polymorphism and insertion-deletion discovery through next-generation sequencing of reduced representation libraries in common bean. *Mol. Breed.* 33, 769–778. doi: 10.1007/s11032-013-9997-7

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Cortés and Blair. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.