

Tennessee State University

## Digital Scholarship @ Tennessee State University

---

Agricultural and Environmental Sciences  
Faculty Research

Department of Agricultural and Environmental  
Sciences

---

12-19-2018

# Does the Genomic Landscape of Species Divergence in Phaseolus Beans Coerce Parallel Signatures of Adaptation and Domestication?

Andrés J. Cortés  
*Universidad Nacional de Colombia*

Paola Skeen  
*Universidad Nacional de Colombia*

Matthew W. Blair  
*Tennessee State University*

María I. Chacón-Sánchez  
*Universidad Nacional de Colombia*

Follow this and additional works at: <https://digitalscholarship.tnstate.edu/agricultural-and-environmental-sciences-faculty>



Part of the [Plant Sciences Commons](#)

---

### Recommended Citation

Cortés Andrés J., Skeen Paola, Blair Matthew W., Chacón-Sánchez María I., Does the Genomic Landscape of Species Divergence in Phaseolus Beans Coerce Parallel Signatures of Adaptation and Domestication?, *Frontiers in Plant Science*, VOLUME 9, 2018, PAGES 1816, ISSN 1664-462X, <https://www.frontiersin.org/article/10.3389/fpls.2018.01816>

This Article is brought to you for free and open access by the Department of Agricultural and Environmental Sciences at Digital Scholarship @ Tennessee State University. It has been accepted for inclusion in Agricultural and Environmental Sciences Faculty Research by an authorized administrator of Digital Scholarship @ Tennessee State University. For more information, please contact [XGE@Tnstate.edu](mailto:XGE@Tnstate.edu).



# Does the Genomic Landscape of Species Divergence in *Phaseolus* Beans Coerce Parallel Signatures of Adaptation and Domestication?

Andrés J. Cortés<sup>1,2\*</sup>, Paola Skeen<sup>3,4</sup>, Matthew W. Blair<sup>5</sup> and María I. Chacón-Sánchez<sup>3</sup>

<sup>1</sup> Corporación Colombiana de Investigación Agropecuaria (Agrosavia) – Centro de Investigación La Selva, Rionegro, Colombia, <sup>2</sup> Universidad Nacional de Colombia – Sede Medellín, Facultad de Ciencias Agrarias – Departamento de Ciencias Forestales, Medellín, Colombia, <sup>3</sup> Universidad Nacional de Colombia – Bogotá, Facultad de Ciencias Agrarias – Departamento de Agronomía, Bogotá, Colombia, <sup>4</sup> Department of Plant Sciences, University of California, Davis, Davis, CA, United States, <sup>5</sup> Department of Agricultural and Environmental Science, Tennessee State University, Nashville, TN, United States

## OPEN ACCESS

### Edited by:

Octavio Salgueiro Paulo,  
Universidade de Lisboa, Portugal

### Reviewed by:

Margarida Matos,  
Universidade de Lisboa, Portugal  
Anti Vasemägi,  
University of Turku, Finland

### \*Correspondence:

Andrés J. Cortés  
acortes@agrosavia.co

### Specialty section:

This article was submitted to  
Evolutionary and Population Genetics,  
a section of the journal  
Frontiers in Plant Science

**Received:** 02 April 2018

**Accepted:** 22 November 2018

**Published:** 19 December 2018

### Citation:

Cortés AJ, Skeen P, Blair MW and  
Chacón-Sánchez MI (2018) Does  
the Genomic Landscape of Species  
Divergence in *Phaseolus* Beans  
Coerce Parallel Signatures  
of Adaptation and Domestication?  
*Front. Plant Sci.* 9:1816.  
doi: 10.3389/fpls.2018.01816

Exploring the genomic architecture of species and populations divergence aids understanding how lineages evolve and adapt, and ultimately can show the repeatability of evolutionary processes. Yet, the genomic signatures associated with divergence are still relatively unexplored, leading to a knowledge gap on whether species divergence ultimately differs in its genetic architecture from divergence at other spatial scales (i.e., populations, ecotypes). Our goal in this research was to determine whether genomic islands of speciation are more prone to harbor within-species differentiation due to genomic features, suppressed recombination, smaller effective population size or increased drift, across repeated hierarchically nested levels of divergence. We used two species of *Phaseolus* beans with strong genepool and population sub-structure produced by multiple independent domestications each especially in Andean and Mesoamerican / Middle American geographies. We genotyped 22,531 GBS-derived SNP markers in 209 individuals of wild and cultivated *Phaseolus vulgaris* and *Phaseolus lunatus*. We identified six regions for species-associated divergence. Out of these divergence peaks, 21% were recovered in the four within-species between-genepool comparisons and in the five within-genepool wild-cultivated comparisons (some of the latter did retrieve genuine signatures of the well described multiple domestication syndromes). However, genomic regions with overall high relative differentiation (measured by  $F_{ST}$ ) coincided with regions of low SNP density and regions of elevated delta divergence between-genepools ( $\Delta_{Div}$ ), independent of the scale of divergence. The divergence in chromosome Pv10 further coincided with a between-species pericentric inversion. These convergences suggest that shared variants are being recurrently fixed at replicated regions of the genome, and in a similar manner across different hierarchically nested levels of divergence, likely as result of genomic features that make certain regions more prone to accumulate islands of speciation and within-species divergence. In summary, neighboring signatures of speciation, adaptation and domestication in *Phaseolus* beans are influenced by ubiquitous genomic constrains, which may continue to fortuitously shape genomic differentiation at various others scales of divergence.

**Keywords:** genomic islands of speciation, genomic signatures of selection, domestication syndrome, convergent evolution, gene flow, genomics constrains, GBS-derived SNP markers

## INTRODUCTION

Exploring the genomic architecture of species' and populations' divergence helps understanding how lineages evolve and adapt, and ultimately how repeatable evolutionary processes are (Gould, 1990; Lobkovsky and Koonin, 2012; Tenaillon et al., 2012). Recent sequencing studies have demonstrated that the genomic consequences of divergence are mixed (Ellegren and Wolf, 2017; Ravinet et al., 2017) and even misleading (Pennisi, 2014). Therefore, a knowledge gap on whether species divergence differs in its genetic architecture from divergence at other spatial scales (i.e., populations, ecotypes) still remains. Addressing this knowledge breach is a much-needed step to fully comprehend the relative contribution of variable gene flow and drift, divergent natural and human selection, and genome-wide heterogeneous recombination and effective population size in shaping the genomic landscape of divergence across and within species.

Genomic signatures associated with species, genepools and ecotypes' divergence can result from reduced gene flow due to interspecific cross-incompatibility, but also random genetic drift and selection (Nei, 2010). The origin of outlier 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). One approach that can help to distinguish these underlying causes of divergence is carrying out a replicated sampling of contrasting populations (Roesti et al., 2014; Lotterhos and Whitlock, 2015). If genetic drift rather than selection is responsible for the divergence, it is unlikely that signals of differentiation reappear consistently across replicates (Lotterhos and Whitlock, 2015). On the other hand, if selection acted on the same genetic variants at the replicated contrasting pairs, genomic regions with comparatively high divergence between individuals from contrasting populations should be identical at each of the replicated populations and those are likely to contain genes involved in adaptive divergence (Nielsen, 2005). Parallel selection on shared genetic variation, like in the well-described *Atp1a1*, *EDA* and *Pitx1* genes in sticklebacks (Jones et al., 2012a,b), should therefore lead to low divergence within populations and across replicates, in the exact genomic regions where equivalent variants are selected at each contrasting population (Roesti et al., 2014). In other words, divergent regions would co-localize with regions of reduced divergence in within-population comparisons if those regions differentiated as a result of parallel divergence from shared variation rather than due to novel variation evolving at each site (Roesti et al., 2014; Irwin et al., 2016). Discerning among gene flow, genetic drift and selection as the cause of parallel genomic divergence is possible as long as there is some degree of replication considered in the sampling of contrasting populations.

The genomic landscape of divergence can also be influenced by differences in ancestral variation and recombination in the genome (Strasburg et al., 2011; Wolf and Ellegren, 2016). Lineage sorting may be enhanced relative to background levels by a reduction in the effective population size ( $N_e$ ) due to processes other than gene flow, like low recombination (Jones et al., 2012b; Zhou et al., 2014; Wolf and Ellegren, 2016), creating genomic islands that eventually expand by genetic hitchhiking

(Ma et al., 2018). Since differentiation is further speeded up in low-recombining regions because of linked selection (Ellegren and Galtier, 2016), the imprint caused by genomic features on the differentiation landscape should be ubiquitous across different levels of divergence. Comparing relative ( $F_{ST}$ ) differentiation (Nachman and Payseur, 2012; Cruickshank and Hahn, 2014; Irwin et al., 2016) across hierarchical nested scales of divergence (Pereira et al., 2016), and coupling this with estimates of the recombination rate (Maynard Smith and Haigh, 1974; Nielsen, 2005; Storz, 2005), can therefore allow for further inferences on the processes giving rise to parallel divergence patterns, such as genomic constrains (Hermisson and Pennings, 2005; Barrett and Schluter, 2008; Pritchard et al., 2010). Therefore, besides a replicated sampling of contrasting populations, a hierarchical nested sampling across various scales of divergence is advisable in order to examine whether genomic islands of divergence may display differentiation due to suppressed recombination, smaller effective population size and increased drift.

*Phaseolus* beans, with their striking genepool structure and multiple domestications, constitute an excellent model system (Broughton et al., 2003; Bitocchi et al., 2017) to test these approaches and to explore to what extent genomic features, besides reduced gene flow and divergent selection, may lead to genomic divergence between (i.e., speciation islands) and within species (i.e., during the natural colonization of new habitats as well as part of the domestication syndromes, which is the suite of phenotypic and genetic changes arising during domestication that distinguish crops from their wild ancestors). Common and Lima beans are the only bean species with multiple domestications among the five domesticated species of *Phaseolus* (Bitocchi et al., 2017). Wild common bean (*P. vulgaris* L.) diverged from its sister species in the tropical Andes (Rendon-Anaya et al., 2017) and colonized South and Central America, originating what nowadays are known as the Andean and Mesoamerican genepools. Independent domestication events for each genepool gave rise to the Andean and Mesoamerican cultivars, although the exact location is still under debate, with a strong role of wild to cultivated introgression likely in the development of cultivar races (Gepts and Debouck, 1991; Kwak and Gepts, 2009; Schmutz et al., 2014). Similarly, wild Lima bean (*P. lunatus* L.) diverged from other *Phaseolus* species in the Andes, after which natural spread also led to a strong genepool structure, with the development of two Andean and two Mesoamerican genepools. Furthermore, independent domestications happened at least in two of these genepools, creating a very diverse range of phenotypes/genotypes for wild and cultivated lima beans that is still under exploration and only recently been analyzed (Chacon-Sanchez and Martinez-Castillo, 2017). To date, genotyping by Sequencing – GBS (Elshire et al., 2011) is among the preferable marker systems for *de novo* common bean SNP detection given its versatility through a wide range of applications (see Cortés and Blair, 2018a, and references therein, and compare with Cortés et al., 2011; Galeano et al., 2012; Kelleher et al., 2012; Blair et al., 2013). This technique has recently been implemented in Lima bean as well (Chacon-Sanchez and Martinez-Castillo, 2017). Therefore, common and Lima beans not only offer an exceptional arrangement of replicated

hierarchical-nested scales of divergence to test a wide spectrum of debatable evolutionary hypotheses, but also possess the necessary genetic resources to accurately quantify genome-wide patterns of differentiation.

In this study we took advantage of the recurrent phylogeographic splits and nested domestication events of common and Lima beans to examine whether genomic islands of speciation in *Phaseolus* species are more prone to harbor within-species divergence due to reduced recombination and increased drift. To accomplish this goal, we asked (1) are between-species  $F_{ST}$  outliers recovered in within-species comparisons, (2) is there any parallelism in the within-species divergence  $F_{ST}$  profiles, and (3) does divergence across repeated and hierarchically nested scales of divergence correlate with intrinsic genomic features (i.e., low-recombining regions such as centromeres and chromosomal rearrangements)? We predicted that if there were some parallelisms in the genetic adaptations to the Mesoamerican and Andean environments or in the genetic consequences of the domestication syndromes, then there would be matching signals of differentiation in the within-species between-genepools divergence  $F_{ST}$  profiles, or in the within-genepool wild-cultivated divergence  $F_{ST}$  profiles, respectively. These patterns of repeatability would not be observed if between-genepools and wild-cultivated divergence outliers were due to genetic drift (Lotterhos and Whitlock, 2015), if selection pressures were different (Ravinet et al., 2017), or if equivalent selective forces did not act on the same shared variation (Roesti et al., 2014; Ravinet et al., 2016). Yet, genomic constrains, rather than true signals of convergent adaptation and domestication, could still be the reason for these parallelisms. If genomic features were indeed constraining divergence, then genomic islands of differentiation would coincide with low-diversity and low-recombining regions, as well as with chromosomal rearrangements, regardless the nature and the scale of divergence.

## MATERIALS AND METHODS

### Plant Material

A total of 209 accessions from *Phaseolus vulgaris* and *Phaseolus lunatus* were used in this study (Supplementary Table S1). All the genotypes were provided by the Genetic Resources Unit at the International Center for Tropical Agriculture (CIAT) and are preserved under the treaty for genetic resources from the Food and Agriculture Organization (FAO). These accessions are representative samples of the core collections for wild (Tohme et al., 1996) and cultivated common (Díaz and Blair, 2006; Blair et al., 2007) bean; as well as a diverse sampling of Lima beans (Motta-Aldana et al., 2010; Serrano-Serrano et al., 2010; Martínez-Castillo et al., 2014), for which there is not yet a core collection. All genepools and races/subpopulations for the wild and cultivated accessions of the two species have already been uncovered by the marker studies of Blair et al. (2012) and Chacon-Sanchez and Martínez-Castillo (2017). A total of 79 materials from *P. vulgaris* comprised 52 Mesoamerican (22 wild and 30 cultivated) and 27 Andean (9 wild and 18 cultivated) accessions; whereas a total of 130 materials from *P. lunatus*

comprised 18 Andean I (8 wild and 10 cultivated), 9 Andean II (all wild), 59 Mesoamerican I (16 wild and 43 cultivated) and 44 Mesoamerican II (33 wild and 11 cultivated) accessions. This sampling spanned all genepools and wild and cultivated populations within species across a replicated hierarchical-nested framework of divergence (Supplementary Figure S1).

### 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). Three 96-plex genotyping-by-sequencing (GBS) assays were made according to Elshire et al. (2011) for the 209 accessions, with one accession per assay chosen at random as duplicate. Library preparation with *ApeKI* digestions and paired-end sequencing were performed at the Biotechnology Resource Center of the Institute for Genomic Diversity (Cornell University, United States) and at the Australian Genome Research Facility (Melbourne, VIC, Australia). Genotyping and SNP calling were analyzed with the software NGSEP (Duitama et al., 2014). Sequence tags were aligned to the *P. vulgaris* v.2.1 reference genome (Schmutz et al., 2014), which is well-annotated and in high synteny with the *P. lunatus* genome (Bonifácio et al., 2012), using the BWA method (Li and Durbin, 2007). Trait Analysis by association, Evolution and Linkage (TASSEL) software (Glaubitz et al., 2014b) was used for filtering.

### Overall Patterns of Population Structure

We explored per-species genepool and subpopulation structure in the 209 accessions of common and Lima beans using principal coordinates analysis (PCoA) implemented in TASSEL (Glaubitz et al., 2014a). Customized scatter plots were drawn using R v.3.3.1 (R Core Team) and were colored according to the domestication status (wild or cultivated) and the genepool/race/subpopulation identity, following Blair et al. (2009, 2012) and Chacon-Sanchez and Martínez-Castillo (2017). Based on this exploration, Andean wild accessions of common bean, from the Ecuador north-Peru subpopulation, and Andean II wild accessions of Lima bean were included as controls for population structure but excluded from the oncoming analyses because they aggregated into independent genetic clusters with low representation, as had already been reported by Blair et al. (2012) and Chacon-Sanchez and Martínez-Castillo (2017).

### Patterns of Genomic Divergence

Patterns of genomic divergence were explored at three different hierarchically nested levels, based on the overall trends of population structure revealed in the previous section. At the top level, species-associated divergence (*P. lunatus* vs. *P. vulgaris*) was used as a proxy to identify islands of speciation, following Nosil and Feder (2011). At the intermediate level, within-species

between-genepools divergence was considered as a consequence of microevolution, that is population divergence and potentially different adaptations to the Mesoamerican and the Andean regions. This level contained four comparisons, one within *P. vulgaris* (Andean vs. Mesoamerican gene pools) and three within *P. lunatus* (Andean I vs. Mesoamerican I, Andean I vs. Mesoamerican II and Mesoamerican I vs. Mesoamerican II gene pools). At the bottom level, within-genepool wild-cultivated comparisons were indicative of the domestication syndrome. This level contained a total of five comparisons, two within *P. vulgaris* (wild vs. cultivated within the Andean and Mesoamerican gene pools) and three within *P. lunatus* (wild vs. cultivated within the Andean I, Mesoamerican I and Mesoamerican II gene pools). This set up led to a total of nine different comparisons of divergence, four of which are replicated within-species between-genepools, and five of which are replicated within-genepool wild-cultivated comparisons.

A sliding window analysis (window size =  $1 \times 10^7$  bps, step size = 500 kb) was used for contrasting between individuals at each of the ten comparisons. Window and step sizes had already been optimized in similar GBS-based genomic scans carried out in *Phaseolus* (Cortés and Blair, 2018a). We calculated relative differentiation computed as the fixation index –  $F_{ST}$  (Weir and Cockerham, 1984) and delta divergence –  $\Delta_{Div}$  (Roesti et al., 2014).  $F_{ST}$  values were averaged across replicated comparisons within each level but were also kept independent for comparative purposes. Confidence intervals around all per-comparison genome-wide average  $F_{ST}$  estimates were computed using the *R quantile* function at an  $\alpha$  value of 0.05.  $F_{ST}$  outliers were also identified at an  $\alpha$  of 0.05 based on the overall  $F_{ST}$  distribution for each comparison. Calculations were done using 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 R v.3.3.1 (R Core Team).

In order to make sense of the landscape of genomic divergence, we conducted the following comparisons:

- (1) We analyzed whether the  $F_{ST}$  outliers between species coincided with high  $F_{ST}$  values at within-species comparisons. This pattern is expected if genomic islands of speciation are repeatedly more prone to harbor within-species divergence as a result of limited recombination (Wolf and Ellegren, 2016).
- (2) We further assessed whether the within-species between-genepools divergence  $F_{ST}$  profiles were similar among the four available comparisons (Andean vs. Mesoamerican gene pools of *P. vulgaris*, and Andean I vs. Mesoamerican I, Andean I vs. Mesoamerican II and Mesoamerican I vs. Mesoamerican II gene pools of *P. lunatus*). This trend is expected if the same variants were selected as the result of similar selective pressures at the Mesoamerican and the Andean regions but not if divergence outliers were due to population divergence, that is genetic drift (Lotterhos and Whitlock, 2015).
- (3) We also assessed whether the within-genepool wild-cultivated divergence  $F_{ST}$  profiles were similar among the five available comparisons (wild vs. cultivated

*P. vulgaris* within the Andean and Mesoamerican gene pools, and wild vs. cultivated *P. lunatus* within the Andean I, Mesoamerican I and Mesoamerican II gene pools). This coincidence is expected if the same variants were selected as the result of parallel domestication syndromes but not if divergence was due to dissimilar domestication pressures or genetic drift (Lotterhos and Whitlock, 2015).

- (4) Finally, we explored if regions of high within-species  $F_{ST}$  co-localized with regions of low  $F_{ST}$  in within-population comparisons.  $\Delta_{Div}$  was used to analyze the difference between these two  $F_{ST}$  values in each window. Peaks in the  $\Delta_{Div}$  statistic point to genomic regions that diverged as a result of parallel divergence from shared variation rather than due to novel variation evolving at each site (Roesti et al., 2014). As a byproduct of the  $\Delta_{Div}$  calculation, the  $F_{ST}$  had to be computed in a total of eleven within-population comparisons. At the intermediate level of comparisons,  $F_{ST}$  was computed in three between-species within-genepool contrasts. These comparisons were (1) between the Andean gene pool of *P. vulgaris* and the Andean I gene pool of *P. lunatus*, (2) between the Mesoamerican gene pool of *P. vulgaris* and the Mesoamerican I gene pool of *P. lunatus*, and (3) between the Mesoamerican gene pool of *P. vulgaris* and the Mesoamerican II gene pool of *P. lunatus*. At the most nested level of comparisons,  $F_{ST}$  values were computed in a total of eight different within-population contrasts. For common bean there were two between-genepool wild-wild cultivated-cultivated comparisons, as follows: (1) between wild Andean and wild Mesoamerican accessions and (2) between cultivated Andean and cultivated Mesoamerican accessions. For Lima bean,  $F_{ST}$  values were calculated in three between-genepool wild-wild comparisons, as follows: (1) between wild Andean I and wild Mesoamerican I, (2) between wild Andean I and wild Mesoamerican II, and (3) between wild Mesoamerican I and wild Mesoamerican II. Also for Lima bean,  $F_{ST}$  values were estimated in three between-genepool cultivated-cultivated comparisons, as follows: (1) between cultivated Andean I and cultivated Mesoamerican I, (2) between cultivated Andean I and cultivated Mesoamerican II, and (3) between cultivated Mesoamerican I and cultivated Mesoamerican II. We refrained from calculating further single summary statistics across or within hierarchy replicates because of potential particularities in the processes behind each comparison that required individual inspection.

## Genome-Wide Patterns of Nucleotide Diversity

In order to describe patterns of diversity across the genome, we implemented a sliding window approach: window size =  $1 \times 10^6$  bps, step size = 200 kb, already optimized for the exploration of genome-wide diversity by Cortés and Blair (2018a). We computed per-window averages of SNP density, nucleotide

diversity as measured by  $\pi$  (Nei, 1987) 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). Since the distribution of the genomic islands of speciation and the within-species divergence peaks could have been constrained by genomic features, the 1 Mb flanking region of each  $F_{ST}$ -based outlier window midpoint was highlighted on the same plot depicting diversity statistics across the genome. Different colors were used to distinguish among comparisons between species (*P. lunatus* vs. *P. vulgaris*), between genepools (average of four within-species between-genepools comparisons), and between domestication statuses for *P. vulgaris* (average of two within-genepool wild-cultivated comparisons) and *P. lunatus* (average of three within-genepool wild-cultivated comparisons). Means and standard errors of all three summary statistics were computed and plotted in R v.3.3.1 (R Core Team) across all types of comparisons.

## RESULTS

### GBS Results

The raw Illumina DNA sequence data (745,927,060 high-quality barcoded reads) were processed through the GBS analysis pipeline as implemented in NGSEP (Duitama et al., 2014). The GBS analysis generated 2,000,294 unique sequence clusters (tags, Elshire et al., 2011; Glaubitz et al., 2014b). Of the total number of tags, 58.9% aligned uniquely to the *P. vulgaris* reference genome (Schmutz et al., 2014), 12.8% had multiple matches and 28.2% were unaligned. A total of 592,171 putative biallelic SNP markers were identified in the aligned tags after filtering for minimum read depth of 5X gene coverage and minimum quality score of 40. Of these, 87% with more than 20% missing data, a default threshold used for GBS studies (Glaubitz et al., 2014b), and further 59% with minimum allele frequencies (MAF) below 0.05 were excluded from the dataset. The high number of unaligned sequences and filtered markers is usual in GBS analyses (Glaubitz et al., 2014b) and in this particular case can be attributed to the naturally high levels of sequence divergence between species and genepools (Cortés and Blair, 2018a). Despite this, this GBS study yielded 22,531 SNP markers of high quality.

### Within-Species Population Structure Matched Genepool Identity and Domestication Status

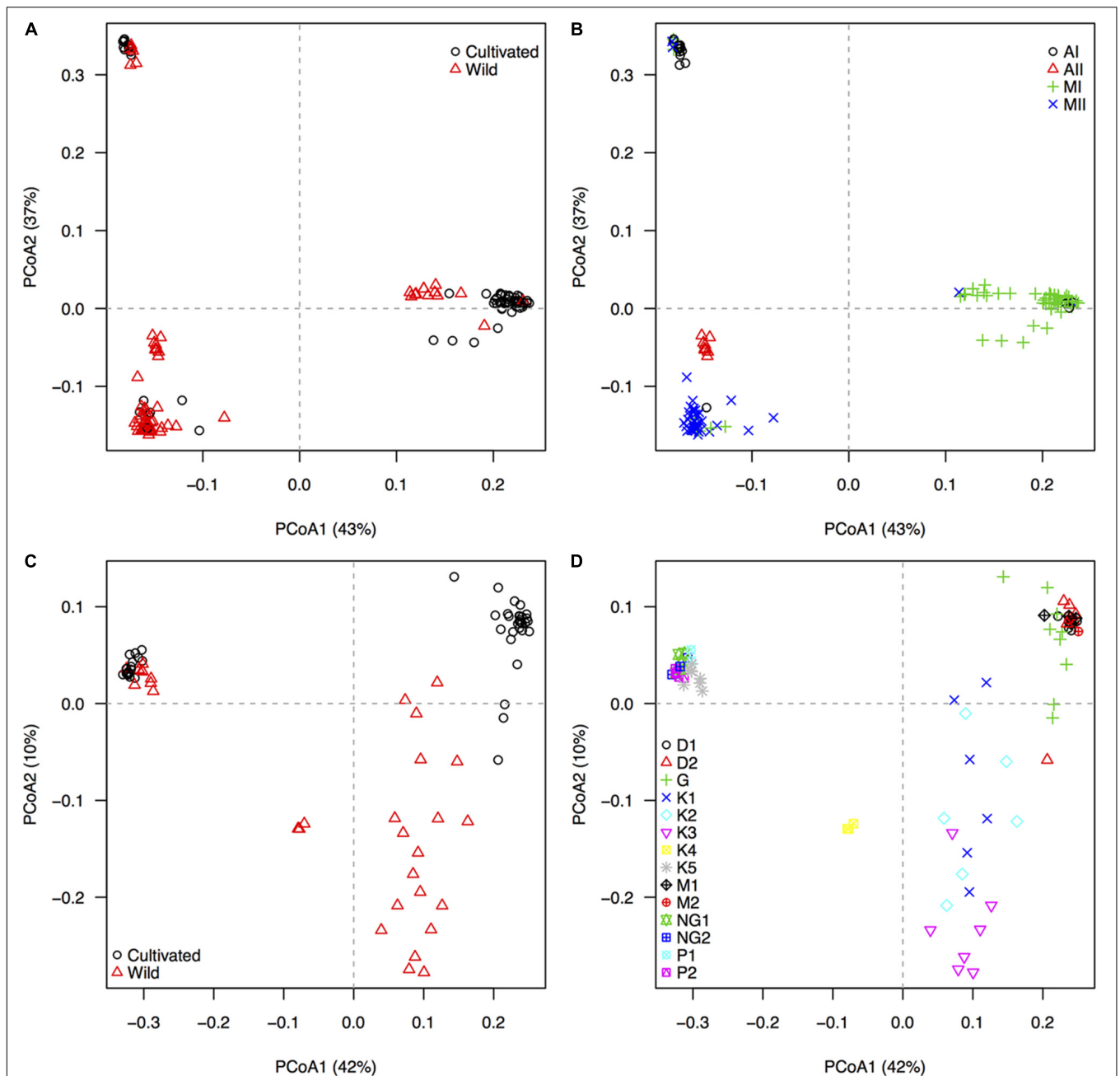
The 22,531 GBS-derived SNP markers recovered, through a principal components analysis with 209 accessions, the distinctive genepool structure of *P. lunatus* and *P. vulgaris* (Figure 1). The wild-cultivated split, although noticeable, was less marked in the three main genepools of *P. lunatus* (Figure 1A) and in the Andean genepool of *P. vulgaris* (Figure 1C). Within *P. lunatus*, the Mesoamerican I genepool was as separated from the Mesoamerican II genepool as any of the Mesoamerican genepools was from the Andean I genepool (Figure 1B). The Andean II genepool was intermediate to the other three

genepools, appearing slightly closer to the Mesoamerican II genepool. Within *P. vulgaris*, the Mesoamerican genepool was more scattered than the Andean genepool, mainly due to the wild Mesoamerican subpopulations that were clearly split from the cultivars. The Ecuador northern-Peru wild subpopulation was intermediate, appearing closer to the Mesoamerican wild subpopulations (Figure 1D). Race Guatemala was closer to the Mesoamerican wild subpopulations than any other Mesoamerican race.

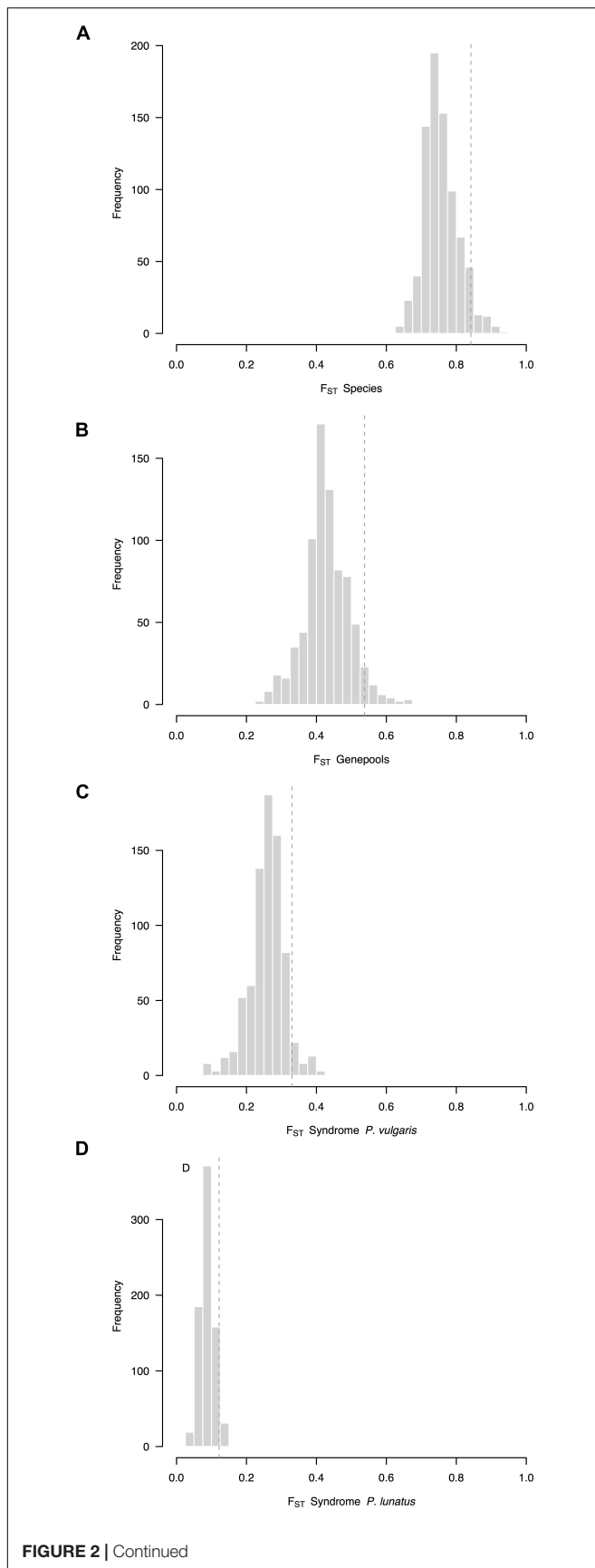
Overall, the average  $F_{ST}$  between species ( $F_{ST} = 0.76$ , CI 95%: 0.72–0.87, Figure 2A) was higher than the average  $F_{ST}$  between genepools ( $F_{ST} = 0.43$ , CI 95%: 0.39–0.56, Figure 2B). Both were higher than the average  $F_{ST}$  between wild and cultivated accessions of common bean ( $F_{ST} = 0.26$ , CI 95%: 0.24–0.36, Figure 2C), which in turn was higher than the average  $F_{ST}$  between wild and cultivated accessions of Lima bean ( $F_{ST} = 0.09$ , CI 95%: 0.07–0.14, Figure 2D). Within-species between-genepools comparisons were asymmetric. Although the average  $F_{ST}$  values between the Andean and Mesoamerican genepools of *P. lunatus* were indistinguishable (for the Andean I vs. Mesoamerica I comparison,  $F_{ST} = 0.61$ , CI 95%: 0.53–0.88, Supplementary Figure S1C; for the Andean I vs. Mesoamerica II comparison,  $F_{ST} = 0.54$ , CI 95%: 0.48–0.83, Supplementary Figure S1D), both of them were higher than the average  $F_{ST}$  values between the Mesoamerican genepools of *P. lunatus* ( $F_{ST} = 0.35$ , CI 95%: 0.31–0.52, Supplementary Figure S1E) and between the Andean and Mesoamerican genepools of *P. vulgaris* ( $F_{ST} = 0.22$ , CI 95%: 0.18–0.42, Supplementary Figure S1B). Within-genepool wild-cultivated comparisons were also asymmetric. The average  $F_{ST}$  between wild and cultivated accessions of Andean common bean ( $F_{ST} = 0.37$ , CI 95%: 0.35–0.44, Supplementary Figure S1F) exceed the average  $F_{ST}$  between wild and cultivated accessions of Mesoamerican common bean ( $F_{ST} = 0.15$ , CI 95%: 0.10–0.30, Supplementary Figure S1G). Both comparisons surpassed the average  $F_{ST}$  values between wild and cultivated accessions within genepools of Lima bean, which were equivalent among them (within the Andean I genepool,  $F_{ST} = 0.05$ , CI 95%: 0.03–0.13, Supplementary Figure S1H; within the Mesoamerican I genepool,  $F_{ST} = 0.12$ , CI 95%: 0.10–0.20, Supplementary Figure S1I; within the Mesoamerican II genepool,  $F_{ST} = 0.09$ , CI 95%: 0.08–0.17, Supplementary Figure S1J).

### The Between-Species and Within-Species Genomic Divergence Profiles Overlapped Scatteredly

The genomic landscape of species divergence revealed six outlier regions in chromosomes Pv3, Pv7, Pv8, Pv10, and Pv11 (Figure 3A). Chromosome Pv10 exhibited two outlier regions split by a 'high valley.' On the other hand, the average genomic landscape of genepool divergence revealed four outlier regions in chromosomes Pv1, Pv5, Pv10, and Pv11 (Figure 3B). The region in chromosome Pv10 overlapped with a between-species peak, whereas the region in chromosome Pv11 fell beside a between-species outlier region. Finally, the average genomic landscape of domestication revealed six outlier regions in common bean (Figure 3C) and seven outlier regions



**FIGURE 1 |** Population structure, as revealed by an unsupervised machine learning strategy implemented through a principal coordinates analyses (PCoAs), based on 22,531 GBS-derived SNP markers genotyped in 209 individuals of wild and cultivated Lima bean (*Phaseolus lunatus*) and common bean (*P. vulgaris*). First and second row of diagrams respectively display accessions of (A,B) *P. lunatus* and (C,D) *P. vulgaris*, colored by (A,C) domestication status and (B,D) genepool/race identity. Black circles and red triangles in the left panels (A,C) mark cultivated and wild accessions, respectively. For Lima bean (B), four genepool populations, according to Chacon-Sanchez and Martinez-Castillo (2017), are indicated by different symbols. Andean genepool genotypes of *P. lunatus* are divided into subgroups AI and All, whereas Mesoamerican genepool genotypes of *P. lunatus* are divided into subgroups MI and MII. For common bean (D), nine within-genepool cultivated subpopulations, according to Blair et al. (2009), and five within-genepool wild subpopulations, according to Blair et al. (2012), are indicated by different symbols. Andean genepool genotypes of *P. vulgaris* (cloud on the upper left quadrant) are divided into cultivated subgroups NG1 and NG2 (race Nueva Granada), cultivated subgroups P1 and P2 (race Peru) and wild subgroup K5 (Andean). Mesoamerican genepool genotypes of *P. vulgaris* are divided into cultivated subgroups D1 and D2 (Durango–Jalisco complex), cultivated subgroup G (race Guatemala), cultivated subgroups M1 and M2 (race Mesoamerica), and wild subgroups K1 (Mesoamerican), K2 (Guatemalan), K3 (Colombian), and K4 (Ecuador and northern-Peru). The percentage of explained variation by each axis is shown within parenthesis in the label of the correspond axis.



**FIGURE 2 |** Frequency distributions of relative differentiation across hierarchically nested levels of divergence in *Phaseolus*. Relative differentiation is calculated as the fixation index ( $F_{ST}$ ). Average  $F_{ST}$  values per sliding window (window size =  $1 \times 10^7$  bps, step size = 500 kb) were computed **(A)** between species (*P. lunatus* vs. *P. vulgaris*), **(B)** between gene pools (average of four within-species between-gene pools comparisons), and between domestication statuses for **(C)** *P. vulgaris* (average of two within-gene pool wild-cultivated comparisons, **Supplementary Figures S1F,G**) and **(D)** *P. lunatus* (average of three within-gene pool wild-cultivated comparisons, **Supplementary Figures S1H–J**). Single within-species between-gene pools comparisons were: Andean vs. Mesoamerican gene pools of *P. vulgaris* (**Supplementary Figure S1B**), and Andean I vs. Mesoamerican I (**Supplementary Figure S1C**), Andean I vs. Mesoamerican II (**Supplementary Figure S1D**) and Mesoamerican I vs. II (**Supplementary Figure S1E**) gene pools of *P. lunatus*. Dashed lines are thresholds for the identification of outliers ( $\alpha = 0.05$ ).

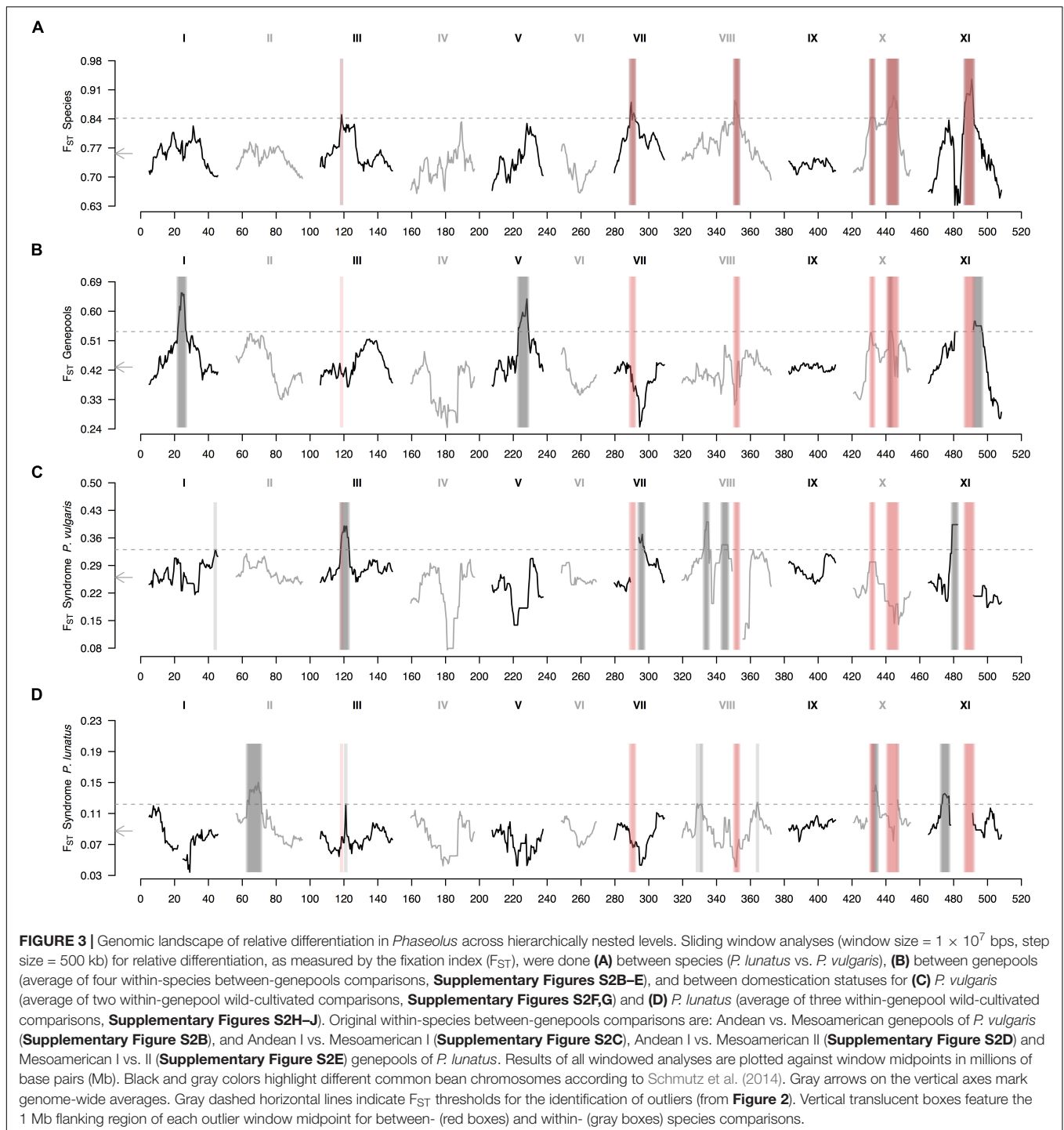
in Lima bean (**Figure 3D**), of which only one was shared (in chromosome Pv3). Of these eleven different regions, only three, in chromosomes Pv3 and Pv10, overlapped to outlier regions in the between-species comparisons.

When exploring in more detail the genomic landscape of relative differentiation in *Phaseolus* beans at different hierarchically nested levels, parallel divergence among between-species and within-species comparisons remained limited (**Supplementary Figure S2**). The only between-species divergence peak that reappeared consistently in individual between-gene pools comparisons was the one in chromosome Pv11 (**Supplementary Figures S2B–D**). This same region was outlier in the wild vs. cultivated  $F_{ST}$  profile of Mesoamerican I Lima bean (**Supplementary Figure S2I**). Similarly, the only between-species divergence peak that reappeared somehow steadily in individual wild-cultivated comparisons was the one in chromosome Pv3 (**Supplementary Figures S2F,G**), which was recovered in the  $F_{ST}$  profiles of both common bean domestications (**Supplementary Figures S2F,G**), and was adjacent to an outlier region in the  $F_{ST}$  profile of the Andean I Lima bean domestication (**Supplementary Figure S2H**). Also recurrent across profiles, with less consistency though, was the between-species divergence region in chromosome Pv10 that coincided with  $F_{ST}$  peaks in all between-gene pools comparisons of *P. lunatus* (**Supplementary Figures S2C–E**), except for the Andean I vs. Mesoamerican II contrast (**Supplementary Figure S2D**), as well as in the domestication profiles of Andean common bean (**Supplementary Figure S2F**), and Andean I (**Supplementary Figure S2H**) and Mesoamerica II Lima bean (**Supplementary Figure S2J**).

### Within-Species Divergence Revealed Some Signatures of Parallelism and Shared Variation

Within-species between-gene pools divergence  $F_{ST}$  profiles (**Supplementary Figures S2B–E**) were partially similar among the four available comparisons (Andean vs. Mesoamerican gene pools of *P. vulgaris*, and Andean I vs. Mesoamerican I, Andean I vs. Mesoamerican II and Mesoamerican I vs. Mesoamerican II gene pools of *P. lunatus*). Peaks in chromosomes





Pv1 and Pv11 reappeared in all four comparisons, although the peak in chromosome Pv11 was slightly shifted leftward in the comparison between the Mesoamerican gene pools of *P. lunatus* (**Supplementary Figure S2E**), so that it did not overlap the between-species peak that the other comparisons did. This same comparison lacked a peak in chromosome Pv5 that was consistently observed in the other three comparisons (**Supplementary Figures S2B–D**), and exhibited an outlier peak

in chromosome Pv10 that was also detected in the Andean I vs. Mesoamerican I comparison within Lima bean (**Supplementary Figure S2C**), overlying a between-species peak. Single peaks in chromosomes Pv2 and Pv8, and in chromosomes Pv4 and Pv6, respectively appeared only once in that comparison (Mesoamerican gene pools of *P. lunatus*, **Supplementary Figure S2E**) and in the between-gene pools comparison within common bean (**Supplementary Figure S2B**).

Within-genepool wild-cultivated divergence  $F_{ST}$  profiles (**Supplementary Figures S2F–J**) were also moderately recurrent across all five available comparisons (wild vs. cultivated *P. vulgaris* within the Andean and Mesoamerican genepools, and wild vs. cultivated *P. lunatus* within the Andean I, Mesoamerican I and Mesoamerican II genepools). A couple of peaks re-appeared in several profiles of the domestication syndrome and overlapped differentiated regions between species. The first one, a divergence peak in chromosome Pv3, was repeatedly retrieved in the genomic profiles of the common bean domestications (**Supplementary Figures S2F,G**) and the Lima bean Andean I domestication (**Supplementary Figure S2H**). Likewise, an outlier region at the beginning of chromosome Pv10 was recurrently recovered in the genomic profiles of the Andean domestications of both species (**Supplementary Figures S2F,H**) and the Lima bean Mesoamerican II domestication (**Supplementary Figure S2J**). Among the few consistent signatures of domestication, a divergence peak in chromosome Pv11 re-appeared in the genomic profiles of all Mesoamerican syndromes (**Supplementary Figures S2G,I,J**). Toward the left of this region, another shared outlier  $F_{ST}$  peak in chromosome Pv11 was detected in the Andean I (**Supplementary Figure S2H**) and the Mesoamerican I (**Supplementary Figure S2I**) domestication profiles of Lima bean. These two peaks were separated by a ‘high valley’ and fell beside the outlier peak detected for all within-species between-genepools comparisons (**Supplementary Figures S2B–E**), which coincided with a between-species peak (**Supplementary Figure S2A**) only recovered by the profile of the Mesoamerican I Lima bean domestication (**Supplementary Figure S2I**). An outlier region also captured in all  $F_{ST}$  profiles of the Mesoamerican domestications (**Supplementary Figures S2G,I**), except for the one of the Mesoamerican II Lima bean (**Supplementary Figure S2J**), was located in chromosome Pv8. At the beginning of the same chromosome, three outlier peaks almost overlapped between both common bean domestications (**Supplementary Figures S2F,G**) and the Lima bean Mesoamerican II domestication (**Supplementary Figure S2J**). All outlier windows in chromosome Pv8 flanked a between-species divergence peak (**Supplementary Figure S2A**). Also tangentially, a peak in chromosome Pv2 observed in the Mesoamerican common bean domestication  $F_{ST}$  profile (**Supplementary Figure S2G**) re-appeared in the profiles of all the Lima bean domestications (**Supplementary Figures S2H–J**). Single outlier peaks in chromosomes Pv4 and Pv7 were respectively detected for the Andean (**Supplementary Figure S2F**) and the Mesoamerican (**Supplementary Figure S2G**) domestications of common bean, and in chromosomes Pv1 and Pv4 for the Andean domestication of Lima bean (**Supplementary Figure S2H**).

Finally, the within-species  $\Delta_{Div}$  statistic (**Figure 4** and **Supplementary Figure S3**), indicative of parallel divergence from shared variation, revealed that regions of high  $F_{ST}$  generally co-localized with regions of low  $F_{ST}$  in within-population comparisons (**Supplementary Figure S4**). This pattern mostly held for the average landscape of genepool divergence, for which all within-species between-genepools differentiated regions overlapped peaks in the  $\Delta_{Div}$  profile (**Figure 4B**); as well as

for the average landscape of domestication in common bean, for which five out of six wild-cultivated differentiated regions overlaid peaks in the  $\Delta_{Div}$  profile (**Figure 4C**, the exception was the region in the tail of chromosome Pv8). However, the average landscape of domestication in Lima bean did not have any wild-cultivated differentiated region intersecting a peak in the  $\Delta_{Div}$  profile (**Figure 4D**), meaning that in this case domestication likely did not involve as much shared variation than for common bean. The outlier region in chromosome Pv3 that was shared by the two average wild-cultivated  $F_{ST}$  profiles intersected a  $\Delta_{Div}$  peak for the domestication of common bean (**Figure 4C**) but not for the domestication of Lima bean (**Figure 4D**).

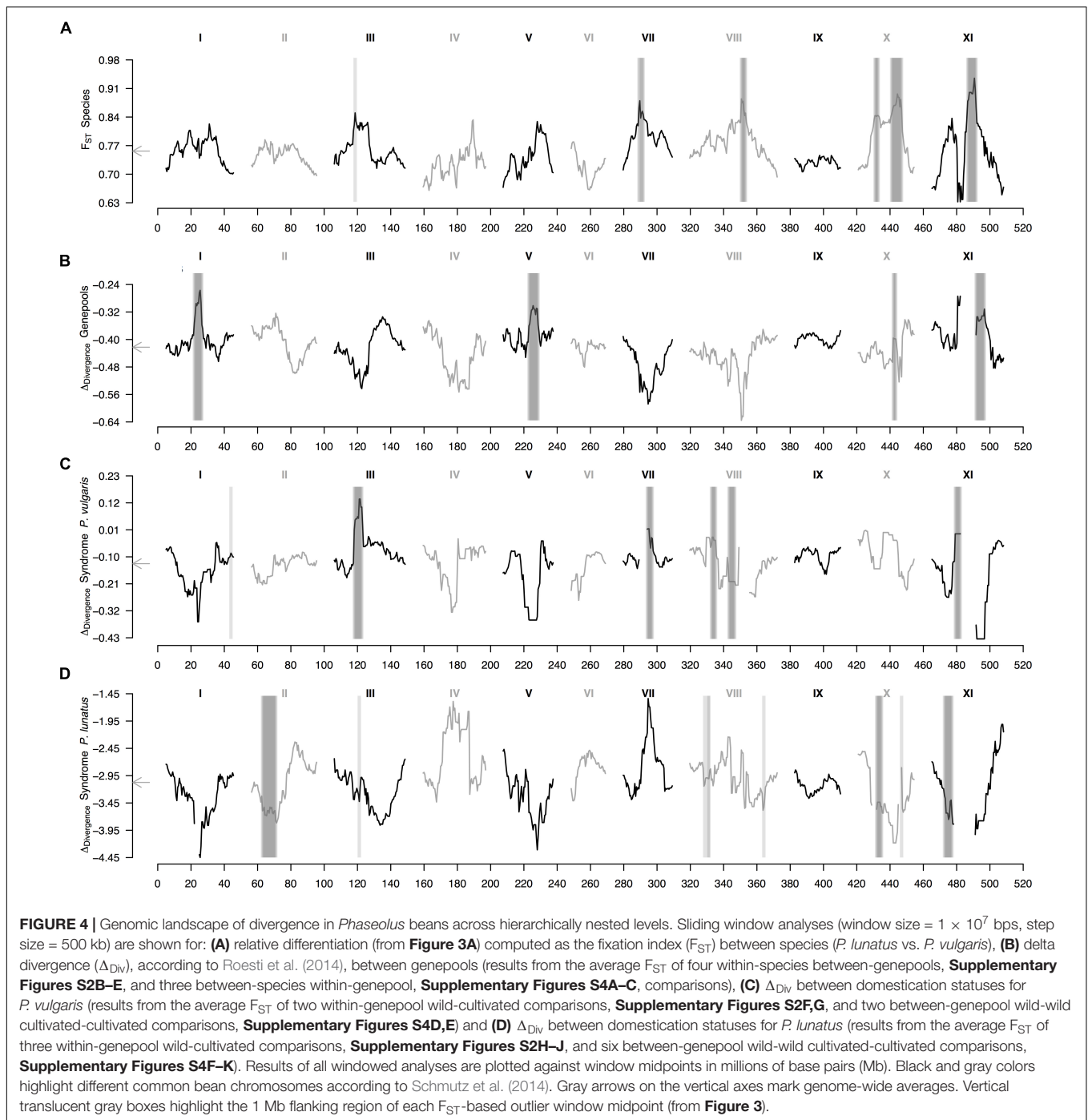
## Genomic Islands Correlated With Intrinsic Genomic Features

A sliding window analysis (window size =  $1 \times 10^6$  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 24 SNPs per million base pairs (95% CI, 7–73, **Figure 5A**). Average nucleotide diversity as measured by  $\pi$  was 0.40 per million base pairs (95% CI, 0.38–0.47, **Supplementary Figure 5B**). Average Tajima’s  $D$  was 1.36 per million base pairs (95% CI, 1.24–1.79, **Figure 5C**). High values of  $\pi$  and Tajima’s  $D$  mirrored strong hierarchically nested population structure.

Genomic differentiation across hierarchically nested scales of divergence always coincided with regions of relatively low SNP density regardless the nature of the comparison (average SNP density for all windows  $19 \pm 2$  vs. divergent windows between: species  $5 \pm 4$ , genepools  $5 \pm 4$ , and domestication syndromes in *P. vulgaris*  $8 \pm 4$  and in *P. lunatus*  $10 \pm 4$ , **Figure 6A**). However, the  $F_{ST}$ -based outlier regions spanned a wide range of  $\pi$  values (all windows  $0.405 \pm 0.005$  vs. divergent windows between: species  $0.45 \pm 0.01$ , genepools  $0.42 \pm 0.03$ , and domestication syndromes in *P. vulgaris*  $0.38 \pm 0.04$  and in *P. lunatus*  $0.42 \pm 0.02$ , **Figure 6B**) and Tajima’s  $D$  values (all windows  $1.38 \pm 0.01$  vs. divergent windows between: species  $1.65 \pm 0.03$ , genepools  $1.46 \pm 0.08$ , and domestication syndromes in *P. vulgaris*  $1.2 \pm 0.1$  and in *P. lunatus*  $1.45 \pm 0.06$ , **Figure 6C**). The latter were significantly inflated in divergent windows between species and reduced in divergent windows between wild and cultivated common bean.

## DISCUSSION

We have found that genomic islands of speciation are not necessarily more prone to harbor within-species divergence, yet subjacent genomic constrains could still be shaping parallel divergence at broader genomic scales. With that in mind, we first discuss some evidence of parallelisms in the genetic adaptations to the Mesoamerican and Andean environments in common and Lima beans resulting from shared variation. Later, we review cases of moderate repeatability in the genomic consequences of multiple domestication events. Finally, we

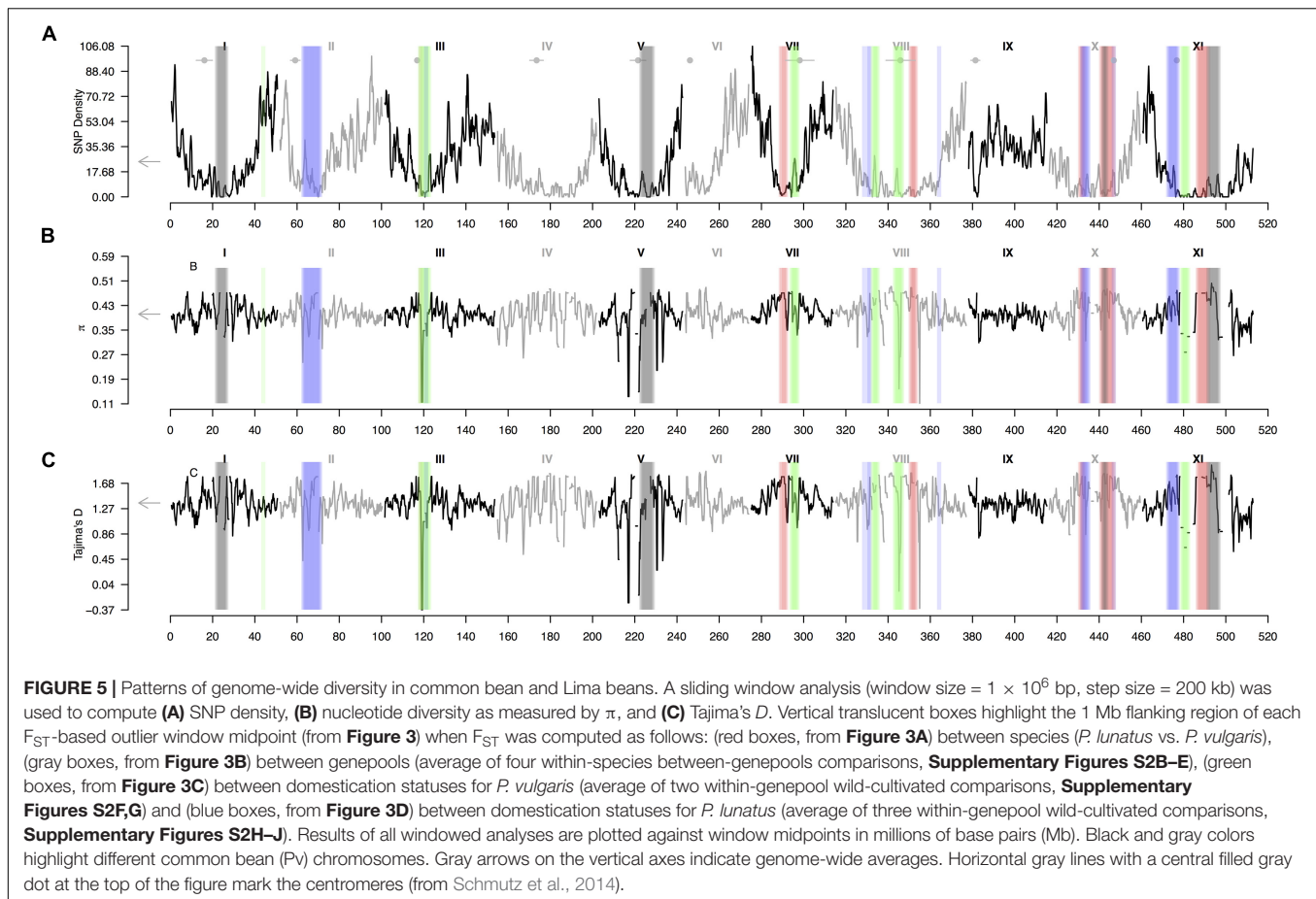


argue that despite those genetic variants may have been recruited in parallel by selective processes alone, genomic features and linked selection could have also enhanced convergent differentiation in low-recombining regions. Our study shows that differentiation across repeated and hierarchically nested levels of divergence always co-occurs with regions of low SNP density, and in some cases with chromosomal rearrangements. These concurring signatures may be a byproduct of genomic constrains inherent to low-recombining regions and intrinsic genomic features. We advise a more systematic use of repeated

and hierarchically nested samplings in order to improve our understanding of the underlying causes of the genomic landscape of divergence.

## Genomic Islands of Speciation Are Not More Prone to Harbor Within-Species Divergence

We predicted that if genomic islands of speciation in *Phaseolus* species were more prone to harbor within-species

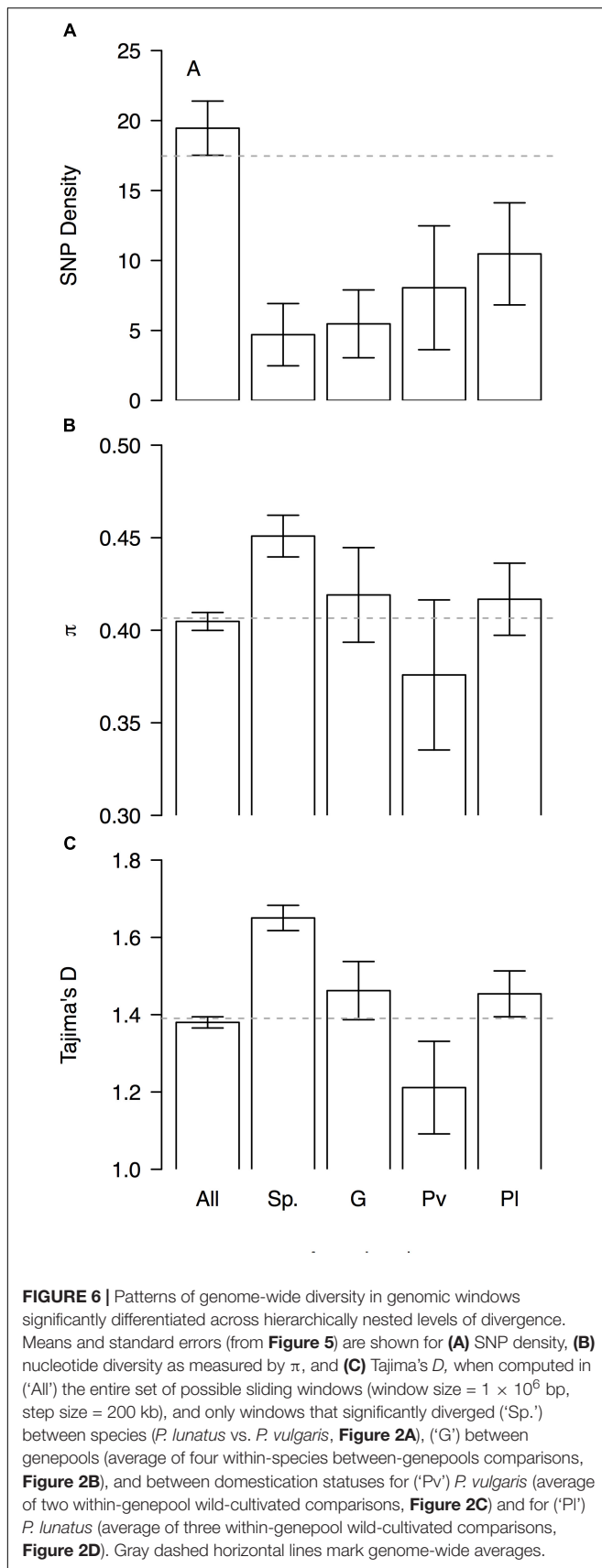


divergence due to reduced recombination and increased drift, then between-species  $F_{ST}$  outliers should be recovered in within-species comparisons. Of the six different speciation islands found in this study, only three, in chromosomes Pv3 and Pv10, are shared with any of the 11 regions found as outliers in the within-species comparisons, leading to an overlap of 21%. Parallel divergence among between-species and within-species comparisons remains limited when exploring in more detail the genomic landscape of relative differentiation in *Phaseolus* beans at different hierarchically nested levels. Recurrent across profiles is the between-species divergence region in chromosome Pv10, where a pericentric inversion between species is reported (Bonifácio et al., 2012) – a genomic feature that we do not discard could enhance divergence at other regions as well (i.e., Pv11). In addition, the only between-species divergence peak that reappears consistently in individual between-genepools comparisons is the one in chromosome Pv11, whereas the only between-species divergence peak that reappears somehow steadily in individual wild-cultivated comparisons is the one in chromosome Pv3. Therefore, we can conclude that between-species  $F_{ST}$  outliers in *Phaseolus* beans are not always recovered in within-species comparisons and that within-species divergence does not necessarily arise within the speciation islands. At first glance, this result could discard the phantom of spurious processes pushing up the

genomic islands of speciation (Pennisi, 2014; Ellegren and Wolf, 2017; Ravinet et al., 2017). However, as discussed in the last section of the discussion, subjacent genomic constraints may still shape parallel divergence at a broader scale across the genome.

### Within-Species Parallel Divergence Partially Results From Shared Variation

Our study provides evidence of some putative parallelisms in the genetic adaptations to the Mesoamerican and Andean environments in common and Lima beans. This finding assumes that there has been neglectful gene flow between cultivars and does not discard that genomic features can still play a role in shaping within-species genomic divergence. Of the eight outlier regions in the within-species between-genepools divergence  $F_{ST}$  profiles, three peaks (37%) in chromosomes Pv1, Pv5, and Pv11 reappear somehow consistently across all four comparisons. Besides this exploration, the landscape of genomic adaptation has remained largely unexplored in *Phaseolus* beans. Among the few other studies addressing this question, a panel of wild common bean sampled across the Andean and Mesoamerican ranges revealed that regardless the strength of the bottlenecks (Ariani et al., 2017), the signatures of divergent adaptation are widespread along the genome and



coincided with regions of elevated SNP density (Cortés and Blair, 2018a), frequent recombination and high gene content (Blair et al., 2018). However, these surveys have not explicitly addressed the colonization of the Andes by lineages coming from Central America and the corresponding change in selection pressures associated with different altitudes, latitudes and micro-environments. Topographically complex mountainous systems, such as the Andes, harbor an impressive heterogeneity of climates at a small scale (Cortés and Blair, 2018b; Cortés and Wheeler, 2018). The ridges and valleys constitute physical barriers that can limit dispersal and cause local variation in rainfall, resulting in genetic isolation and variation in habitats (Cortés et al., 2014; Sedlacek et al., 2014, 2015, 2016; Wheeler et al., 2014, 2015, 2016; Little et al., 2016). Both processes have likely speeded up the evolution of high species diversity in this region (Antonelli et al., 2009; Hoorn et al., 2010; Madriñán et al., 2013; Arnegard et al., 2014; Cortés et al., 2018). Yet, the relative effects of geographic isolation, environmental variation at a small scale, and their potential interactions across genepools remain poorly understood in wild beans. Characterizing in more detail the genomic consequences associated with the colonization of diverse habitats may disclose new cases of genetic parallelism in the adaptation of beans.

The genomic consequences of multiple domestication events are also moderately recurrent as revealed by our survey. From the twelve regions putatively differentiated as the result of the domestication syndrome, only 5 (42%) appear in more than one comparison but none appears in all. Two peaks in chromosome Pv3 and Pv10 are repeated across three different comparisons of all five profiles of the domestication syndromes. At least the region in chromosome Pv3 has been reported to be involved in the vernalization pathway (i.e., *Phvul.003G033400*) as part of the Mesoamerican domestication of common bean, as well as with seed growth (Schmutz et al., 2014). Two other divergence peaks in chromosome Pv8 and Pv11 are consistent across all three genomic profiles of the Mesoamerican domestication syndrome. The region in chromosome Pv8 is known for being related with the encoding of the nitrate reductase (i.e., *Phvul.008G168000*), a critical element for plant and seed growth, during the Mesoamerican domestication of common bean (Schmutz et al., 2014). Also as part of this domestication event, the region in chromosome Pv11 is associated with increased plant size through the ubiquitin ligase degradation pathway (i.e., *Phvul.011G213300*) that controls flower and stem size (Schmutz et al., 2014). More loosely, a peak at chromosome Pv2 in the Mesoamerican common bean domestication  $F_{ST}$  profile is recovered in the profiles of all three Lima bean domestications. This region has been linked with the domestication syndrome of Lima bean since it is involved in the regulation of seed germination (i.e., *Phvul.002G033500*) and leaf size (i.e., *Phvul.002G041800*) and is enriched by inflated linkage disequilibrium scores (Chacon-Sanchez and Martinez-Castillo, 2017). Also, a QTL (St) related with dehiscence has been reported in this region, but candidate genes are still unknown (Koinange et al., 1996). Although scattered, some of these few regions may reveal truth parallelisms in the domestication syndromes, whereas others may still be constrained by genomic features.

Also striking is the rarity of regions putatively involved in domestication and shared by several domestication events. This trend, mostly expected for quantitative traits with complex genetic architectures (Blair et al., 2006; Meyer et al., 2012) because of the complexity in the interaction of selection, drift, novel mutations and epistatic effects (Chevin and Hospital, 2008; Csillery et al., 2018), had already been noticed for common bean (Schmutz et al., 2014) – potentially applying for Lima bean as well (Chacon-Sanchez and Martinez-Castillo, 2017), and so does not necessarily speak for a prevalent role of drift. Yet, since divergence in the lack of repeatability is still a liable result of lineage sorting, caution must be undertaken while interpreting these signals. Single outlier regions for the common bean Andean–Mesoamerican split and the Lima bean Mesoamerican I–Mesoamerican II split are found in chromosomes Pv4 and Pv6, and in chromosomes Pv2 and Pv8, respectively. Similarly, single outlier peaks associated with the domestication of the Andean and Mesoamerican common bean and the Andean Lima bean are respectively found in chromosomes Pv4, Pv7, and Pv1 and Pv4. These eight singularities may result from different adaptive pressures across the Americas unique to each species, distinctive adaptation to the Mesoamerican micro-environments, dissimilar selection as part of each domestication event (Ravinet et al., 2017), equivalent selective forces acting on different genetic variants (Roesti et al., 2014; Ravinet et al., 2016), or genetic drift (Lotterhos and Whitlock, 2015). Discerning among these causes may require further genotyping in an extended panel specifically addressing each comparison. At least for the divergence peak at chromosome Pv7 in the wild-cultivated Mesoamerican common bean comparison, other drivers besides the domestication itself are an unlikely reason for divergence because a wide region in chromosome Pv7 region is known for being associated with increased seed weight (i.e., *Phvul.007G094299* – *Phvul.007G.99700*) during the Mesoamerican domestication of common bean (Schmutz et al., 2014), as well as with flowering regulation (i.e., *Phvul.007G096500* and *Phvul.007G065600*) as part of the domestication of Lima bean (Chacon-Sanchez and Martinez-Castillo, 2017) and both common bean genepools (Schmutz et al., 2014).

The origin of the variants exhibiting parallel divergence 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). The  $\Delta_{Div}$  statistic is indicative of parallel divergence from shared variation and reveals that regions of high  $F_{ST}$  for the within-species comparisons generally co-localized with regions of low  $F_{ST}$  in within-population comparisons. This pattern mostly holds for the average landscape of genepool divergence and for the average landscape of domestication in common bean and indicates a predominance of parallel divergence from shared variation rather than due to novel variation evolving at each site (Roesti et al., 2014). The average landscape of domestication in Lima bean, however, only has one out of seven wild-cultivated differentiated regions intersecting a peak in the  $\Delta_{Div}$  profile. This peak, in the same tail of chromosome Pv7 that has been associated with multiple domestication syndromes as discussed in the previous paragraph, may therefore also

be the result of divergence from shared variation. Despite that shared variation may have been recruited in parallel by selective processes alone, genomic features and linked selection could have also enhanced convergent differentiation (Wolf and Ellegren, 2016; Ellegren and Wolf, 2017), as discussed below.

## Genomic Features Constrain Divergence Across Scales

Our study shows that differentiation across repeated and hierarchically nested levels of divergence always co-occurs with regions of low SNP density. Increased lineage sorting, and consequently rapid differentiation, is a common phenomenon in low-recombining regions because of linked selection and a reduction in the effective population size (Jones et al., 2012b; Zhou et al., 2014; Wolf and Ellegren, 2016). Likewise, low-recombining regions also tend to exhibit a decline in diversity due to background selection and, to a lower extent, because of genetic hitchhiking (Ellegren and Galtier, 2016; Martin and Jiggins, 2017; Ma et al., 2018; Martin et al., 2018). In our study we have found evidence that regions with low SNP diversity are enriched for contiguous signatures of differentiation between bean species, between genepools and as part of the multiple domestication syndromes. These concurring signatures may be a byproduct of genomic constraints inherent to low-recombining regions.

One of the regions that repeatedly exhibit high differentiation across hierarchically nested levels of divergence in the presence of low SNP density was the centromeric section of chromosome Pv11. The wild-cultivated divergence peak in this chromosome is shared by three domestication syndromes and is located beside the outlier peak detected for all within-species between-genepools comparisons, which in turn coincides with a major between-species peak. In this wide section of chromosome Pv11 our analysis further revealed that convergent divergence is consistently correlated with very low SNP density, as expected because of combined effects of linked and background selection in low-recombining regions (Jones et al., 2012b; Zhou et al., 2014; Wolf and Ellegren, 2016; Ravinet et al., 2017). The observation that genomic constraints are biasing divergence across scales in this section of chromosome Pv11 is reinforced by the fact that previous genomic scans did not attribute to this region a consistent outstanding role during the domestication syndromes (Schmutz et al., 2014; Chacon-Sanchez and Martinez-Castillo, 2017) or in conferring adaptation to different environments and latitudes across the Americas (Cortés et al., 2012a,b, 2013; Blair et al., 2016; Cortés and Blair, 2018a). The only exception is the candidate gene influencing plant size (*Phvul.011G213300*) as part of the Mesoamerican domestication syndrome of common bean (Schmutz et al., 2014), but then this pattern has not been consistently reported for the other domestication events as to explain its steady repeatability across hierarchically nested levels of divergence in windows with low SNP density.

Other ‘hotspots’ for spurious divergence due to genomic constraints may be the regions with low SNP density in

chromosomes Pv8 and Pv10 that exhibit signatures of between-species divergence as well as repeated between genepools and within-genepool wild-cultivated divergence. The region in chromosomes Pv8 was previously reported to be highly divergent during the domestication of the Andean common bean, but then there were not candidate genes in this region associated with that domestication syndrome in particular (Schmutz et al., 2014), despite that the same region is known for being involved in plant and seed growth (i.e., *Phvul.008G168000*) during the Mesoamerican domestication of the same species. This paradox may then be a consequence of genomic constrains obscuring genuine anthropic selection and repeatedly forcing divergence in this region. Similarly, the wide divergent region in chromosomes Pv10, characterized by two outlier peaks split by a 'high valley,' actually matches a pericentric inversion between species (Bonifácio et al., 2012), exemplifying how genomic features inexorably condition differentiation across scales of divergence. This coincidence unlikely is an artifact in the genome assembly because both available versions of the reference genome are equally robust for mapping GBS reads and detecting the same overall  $F_{ST}$  estimates regardless reported inversions (**Supplementary Figure S5**).

The observation that low-recombining regions are enriched for differentiation across repeated and hierarchically nested levels of divergence in *Phaseolus* beans opposes the profiles of the genome-wide selection scans carried out in common bean. While low-recombining regions are more prone to exhibit signatures of divergence, regions toward the arms of the chromosomes with high SNP density more often harbor adaptive variation (Cortés and Blair, 2018a). This trend follows expectations because low-recombining regions are more liable to display divergence because of linked selection (Ellegren et al., 2012; Ellegren and Galtier, 2016; Ellegren and Wolf, 2017), whereas recombination hotspots usually exhibit higher SNP density and are enriched with functional genes (Rodgers-Melnick et al., 2015; Ellegren and Galtier, 2016) – an already well-described relationship for common bean (Bhakta et al., 2015; Blair et al., 2018). Also, adaptive divergent selection usually homogenizes haplotypes within the same niche and fixes polymorphisms in different populations, so that few haplotypes with high frequency remain. This selective process leads to high values of nucleotide diversity and Tajima's  $D$ , and low values of the Watterson's theta ( $\theta$ ) estimator (Wakeley, 2008), a tendency that was corroborated in wild common bean when looking for adaptive variants (Cortés and Blair, 2018a) but that was lacking in the present study while retrieving the genomic landscape of divergence between species, genepools and domestication statuses.

In short, our study provides comprehensive evidence that, despite some genuine parallelism accounting for the recruitment of shared variants at replicated comparisons in the between-genepools comparisons and as part of the domestication syndromes, the consequences of intrinsic genomic features are compelling across different hierarchically nested levels of divergence. Because certain regions are more prone to accumulate islands of divergence as the result of genomic constrains, we advocate that studies of genomic divergence should consider more systematically a dual-purpose

sampling, such as ours. First, using replicated populations under presumably similar selection pressures helps accounting for lineage sorting and characterizing the nature of the selected variants – i.e., novel vs. standing (Roesti et al., 2014). Second, a hierarchically nested sampling across various levels of divergence allows for further assessments on the processes, that like genomic constrains, may give rise to parallel divergence patterns (Hermisson and Pennings, 2005; Barrett and Schluter, 2008; Pritchard et al., 2010; Pereira et al., 2016). Finally, some of these examinations must be verified with genomic features and estimates of the recombination rate (Maynard Smith and Haigh, 1974; Nielsen, 2005; Storz, 2005). Ultimately, our work could be seen as the first exploration of this combined sampling tested in two *Phaseolus* bean species that exhibit strong genepool structure and multiple domestication events. We foresee that as the evidence of pervasive genomic constrains shaping genomic differentiation across species at countless scales of divergence accumulates, replicated samplings of contrasting populations in a hierarchically nested framework of divergence will become indispensable.

## PERSPECTIVES

The putatively adaptive divergence regions identified in this study must be validated. Similar selective pressures at the Mesoamerican and the Andean regions could account for some of the parallelisms in the between-genepools divergence profiles. Similarly, parallel domestication pressures acting on the same genetic variants could be the leading cause for some of the parallelisms observed in the wild-cultivated divergence profiles that were not biased by genomic constrains. A denser genotyping, e.g., by whole genome sequencing – WGS s. Lobaton et al. (2018), in a panel of contrasting genotypes specifically addressing key comparisons would allow narrowing these divergence peaks and identifying candidate pathways (Fuentes-Pardo and Ruzzante, 2017). Pursuing this research is crucial because understanding the genomic signatures of adaptation and domestication is useful for germplasm characterization and offers the potential to enhance breeding by exploiting naturally available genetic variants. On the other hand, among the five domesticated species in the *Phaseolus* genus, common and Lima beans are the only ones exhibiting range expansions toward South American and multiple domestications (Bitocchi et al., 2017). However, exploring the landscape of divergence in other domesticated *Phaseolus* species is equally insightful because of their overlapping distribution ranges, nested phylogenetic relationships and divergent adaptations. For instance, Year (*P. dumosus*) and Runner (*P. coccineus*) beans are Mesoamerican and well adapted to humid habitats, which makes them a potential source of resistance to biotic stresses. On the other hand, Tepary bean (*P. acutifolius*) is also Mesoamerican but is well known for growing in desert and semi-arid environments, which makes it a likely source of tolerance to abiotic stresses. These species also possess well-established genomic resources (Guerra-Garcia et al., 2017) that could speed up newer genome-wide comparisons. *Phaseolus* species that never underwent domestication are also abundant (ca. 70) and could

enrich our understanding of genomic divergence in this intricate genus. Considering all *Phaseolus* species will ultimately reinforce beans as a model for understanding speciation, adaptation and crop evolution (Broughton et al., 2003; Cortés, 2013; Bitocchi et al., 2017).

Extensive samplings meant to validate putatively adaptive divergence regions and to better characterize the genomic landscape of divergence in all *Phaseolus* beans species could benefit from explicit comparisons of the profiles of relative ( $F_{ST}$ ) and absolute ( $D_{XY}$ ) between-population divergence.  $F_{ST}$  vs.  $D_{XY}$  comparisons can inform about population divergence in the presence of gene flow (co-occurrence of peaks in both profiles), recurrent selection across subpopulations (co-occurrence of  $F_{ST}$  peaks with shallow  $D_{XY}$  valleys) and selective sweeps predating the subpopulations' split (co-occurrence of  $F_{ST}$  peaks with deep  $D_{XY}$  valleys) (Nachman and Payseur, 2012; Cruickshank and Hahn, 2014; Irwin et al., 2016). Discerning among these processes will improve our understanding of the range expansions and the multiple domestications of *Phaseolus* beans.

In the long run, we are looking forward to seeing more coherent and systematic samplings of replicated contrasting populations across hierarchically nested levels of divergence in all kinds of species and biomes. Understanding the extent of repeatability and the causes of genomic divergence has always been challenging but the field is now moving forward toward a more cohesive framework. New ways in characterizing obscure genomic features promise aiding our understanding on how the genomic landscape of divergence is shaped.

## DATA ACCESSIBILITY

The sample key file, pipeline configuration file, and filtered dataset are archived at the Dryad Digital Repository under doi: 10.5061/dryad.54c5q66.

## AUTHOR CONTRIBUTIONS

AC and MC-S conceived this study with insights from MB. AC and MC-S carried out DNA extractions to produce GBS data. AC, MC-S, and PS analyzed genotyping by sequencing data. AC wrote the manuscript with contributions from all co-authors.

## FUNDING

This research was funded by a Colciencias (Colombia) grant awarded to MC-S under contract number FP44842-009-2015 and project code 1101-658- 42502, by the grant 3404 from Fundación para la Promoción de la Investigación y la Tecnología del Banco de la República de Colombia to MC-S and by the Lundell and Tullberg (Sweden) grants to AC. The Geneco mobility fund from Lund University is thanked for subsidizing the meeting between AC and MB in the spring of 2015 at Nashville (TN, United States). AC's writing time was sponsored by the grants 4J1-2016-00418 from Vetenskapsrådet (VR) and

BS2017-0036 from Kungliga Vetenskapsakademien (KVA). MB received support from the Evans Allen fund of the United States Department of Agriculture (Grant number TENX-1410 to TSU). The editorial fund from the Colombian Corporation for Agricultural Research is acknowledged for financing to AC.

## ACKNOWLEDGMENTS

We are grateful with D. G. Debouck and the Genetic Resources Unit at the International Center for Tropical Agriculture for providing the seeds of the accessions that were considered in this study. J. Duitama is thanked for bioinformatics assistance. Some of the ideas presented in this manuscript were refined thanks to comments from A. Caro, I. Cerón, C. Jiggins, D. Londoño, P. Reyes, C. Salazar, J. J. Wiens, and R. Yockteng during the VI Symposium of the Colombian Society for Evolution held in Cali (Colombia) on August 2017. Administrative support from J. Bergman, P. Brandén, A. Perrigo, and L. Sjöblom is appreciated. We also acknowledge the Genomic Diversity Facility of the Institute of Biotechnology at Cornell University for support in SNP genotyping.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2018.01816/full#supplementary-material>

**FIGURE S1** | Frequency distributions of relative differentiation across hierarchically nested levels of divergence in *Phaseolus* beans. Relative differentiation is calculated as the fixation index ( $F_{ST}$ ). Average  $F_{ST}$  values per sliding window (window size =  $1 \times 10^7$  bps, step size = 500 kb) were computed as follows: **(A)** between species (*P. lunatus* vs. *P. vulgaris*), **(B)** between Andean and Mesoamerican gene pools of *P. vulgaris*, **(C)** between Andean I and Mesoamerican I gene pools of *P. lunatus*, **(D)** between Andean I and Mesoamerican II gene pools of *P. lunatus*, **(E)** between Mesoamerican I and Mesoamerican II gene pools of *P. lunatus*, **(F)** between wild and cultivated accessions of Andean *P. vulgaris*, **(G)** between wild and cultivated accessions of Mesoamerican *P. vulgaris*, **(H)** between wild and cultivated accessions of Andean I *P. lunatus*, **(I)** between wild and cultivated accessions of Mesoamerican I *P. lunatus*, and **(J)** between wild and cultivated accessions of Mesoamerican II *P. lunatus*. Dashed lines indicate thresholds for the identification of outliers at an  $\alpha$  of 0.05.

**FIGURE S2** | Detail of the genomic landscape of relative differentiation in *Phaseolus* beans across hierarchically nested levels of divergence. Sliding window analyses (window size =  $1 \times 10^7$  bps, step size = 500 kb) for relative differentiation, as measured by the fixation index ( $F_{ST}$ ), were computed as follows: **(A)** between species (*P. lunatus* vs. *P. vulgaris*), **(B)** between Andean and Mesoamerican gene pools of *P. vulgaris*, **(C)** between Andean I and Mesoamerican I gene pools of *P. lunatus*, **(D)** between Andean I and Mesoamerican II gene pools of *P. lunatus*, **(E)** between Mesoamerican I and Mesoamerican II gene pools of *P. lunatus*, **(F)** between wild and cultivated accessions of Andean *P. vulgaris*, **(G)** between wild and cultivated accessions of Mesoamerican *P. vulgaris*, **(H)** between wild and cultivated accessions of Andean I *P. lunatus*, **(I)** between wild and cultivated accessions of Mesoamerican I *P. lunatus*, and **(J)** between wild and cultivated accessions of Mesoamerican II *P. lunatus*. Results of all windowed analyses are plotted against window midpoints in millions of base pairs (Mb). Black and gray colors highlight different common bean chromosomes according to Schmutz et al. (2014). Gray arrows on the vertical axes mark genome-wide averages. Gray dashed horizontal lines indicate  $F_{ST}$  thresholds for the identification of outliers at an  $\alpha$  of 0.05 (from **Supplementary Figure S1**). Vertical



translucent boxes feature the 1 Mb flanking region of each outlier window midpoint for between- (red boxes) and within- (gray boxes) species comparisons.

**FIGURE S3** | Detail of the genomic landscape of divergence in *Phaseolus* beans across hierarchical nested levels. Sliding window analyses (window size =  $1 \times 10^7$  bps, step size = 500 kb) are shown for: **(A)** relative differentiation (from **Figure 3A**) computed as the fixation index ( $F_{ST}$ ) between species (*P. lunatus* vs. *P. vulgaris*), **(B)** delta divergence ( $\Delta_{Div}$ ), according to Roesti et al. (2014), between gene pools (results from the average  $F_{ST}$  of four within-species between-gene pools, **Supplementary Figures S2B–E**, and three between-species within-gene pool, **Supplementary Figures S4A–C**, comparisons), **(C)**  $\Delta_{Div}$  between domestication statuses for *P. vulgaris* (results from the average  $F_{ST}$  of two within-gene pool wild-cultivated comparisons, **Supplementary Figures S2F,G**, and two between-gene pool wild-wild cultivated-cultivated comparisons, **Supplementary Figures S4D,E**) and **(D)**  $\Delta_{Div}$  between domestication statuses for *P. lunatus* (results from the average  $F_{ST}$  of three within-gene pool wild-cultivated comparisons, **Supplementary Figures S2H–J**, and six between-gene pool wild-wild cultivated-cultivated comparisons, **Supplementary Figures S4F–K**). Vertical translucent boxes highlight the 1 Mb flanking region of each  $F_{ST}$ -based outlier window midpoint (from **Figure 3**) when  $F_{ST}$  was computed as follows: (red boxes, from **Figure 3A**) between species (*P. lunatus* vs. *P. vulgaris*), (gray boxes, from **Figure 3B**) between gene pools (average of four within-species between-gene pools comparisons, **Supplementary Figures S2B–E**), (green boxes, from **Figure 3C**) between domestication statuses for *P. vulgaris* (average of two within-gene pool wild-cultivated comparisons, **Supplementary Figures S2F,G**) and (blue boxes, from **Figure 3D**) between domestication statuses for *P. lunatus* (average of three within-gene pool wild-cultivated comparisons, **Supplementary Figures S2H–J**). Results of all windowed analyses are plotted against window midpoints in millions of base pairs (Mb). Black and gray colors highlight different common bean chromosomes according to Schmutz et al. (2014). Gray arrows on the vertical axes mark genome-wide averages.

**FIGURE S4** | Genomic landscape of within-population relative differentiation in *Phaseolus* beans across hierarchically nested levels. Sliding window analyses (window size =  $1 \times 10^7$  bps, step size = 500 kb) for relative differentiation, as measured by the fixation index ( $F_{ST}$ ), were computed in three between-species within-gene pool comparisons as follows: **(A)** between the Andean gene pool of *P. vulgaris* and the Andean I gene pool of *P. lunatus*, **(B)** between the Mesoamerican gene pool of *P. vulgaris* and the Mesoamerican I gene pool of *P. lunatus*, and **(C)** between the Mesoamerican gene pool of *P. vulgaris* and the Mesoamerican II gene pool of *P. lunatus*.  $F_{ST}$  values were computed in two between-gene pool wild-wild cultivated-cultivated comparisons for common bean

as follows: **(D)** between wild Andean and wild Mesoamerican accessions of *P. vulgaris* and **(E)** between cultivated Andean and cultivated Mesoamerican accessions of *P. vulgaris*.  $F_{ST}$  values were calculated in three between-gene pool wild-wild comparisons for Lima bean as follows: **(F)** between wild Andean I and wild Mesoamerican I accessions of *P. lunatus*, **(G)** between wild Andean I and wild Mesoamerican II accessions of *P. lunatus*, and **(H)** between wild Mesoamerican I and wild Mesoamerican II accessions of *P. lunatus*. Finally,  $F_{ST}$  values were estimated in three between-gene pool cultivated-cultivated comparisons for Lima bean as follows: **(I)** between cultivated Andean I and cultivated Mesoamerican I accessions of *P. lunatus*, **(J)** between cultivated Andean I and cultivated Mesoamerican II accessions of *P. lunatus*, and **(K)** between cultivated Mesoamerican I and cultivated Mesoamerican II accessions of *P. lunatus*. Results of all windowed analyses are plotted against window midpoints in millions of base pairs (Mb). Black and gray colors highlight different common bean chromosomes based on Schmutz et al. (2014). Gray arrows on the vertical axes mark genome-wide averages. Dashed lines mark thresholds for the identification of outliers at  $\alpha$  of 0.05. Vertical red translucent boxes show the 1 Mb flanking region of each between-species  $F_{ST}$ -based outlier window midpoint (from **Figure 3A**).

**FIGURE S5** | Correlation of overall  $F_{ST}$  estimates when using the two available versions of the common bean assembly (Schmutz et al., 2014) as reference genome ( $R^2 = 0.948$ , 95% CI 0.946–0.950). Chromosomes **(A)** without and **(B)** with (Pv2, Pv9, and Pv10, the latter in red) reported inversions (Bonifácio et al., 2012) are shown.

**TABLE S1** | Identity of the 209 common bean and Lima bean accessions used in this study. The G identification number (from the Genetic Resources Unit at the International Center for Tropical Agriculture), the status and the gene pool/race identity are shown. For Lima bean, four gene pool populations, according to Chacon-Sanchez and Martinez-Castillo (2017), are indicated by different symbols. Andean gene pool genotypes of *P. lunatus* are divided into subgroups AI and AII, whereas Mesoamerican gene pool genotypes of *P. lunatus* are divided into subgroups MI and MII. For common bean, nine within-gene pool cultivated subpopulations, according to Blair et al. (2009), and five within-gene pool wild subpopulations, according to Blair et al. (2012), are indicated by different symbols. Andean gene pool genotypes of *P. vulgaris* are divided into cultivated subgroups NG1 and NG2 (race Nueva Granada), cultivated subgroups P1 and P2 (race Peru) and wild subgroup K5 (Andean). Mesoamerican gene pool genotypes of *P. vulgaris* are divided into cultivated subgroups D1 and D2 (Durango–Jalisco complex), cultivated subgroup G (race Guatemala), cultivated subgroups M1 and M2 (race Mesoamerica), and wild subgroups K1 (Mesoamerican), K2 (Guatemalan), K3 (Colombian) and K4 (Ecuador and north-Peru).

## REFERENCES

- Antonelli, A., Nylander, J. A. A., Persson, C., and Sanmartin, I. (2009). Tracing the impact of the andean uplift on neotropical plant evolution. *Proc. Natl. Acad. Sci. U.S.A.* 106, 9749–9754. doi: 10.1073/pnas.0811421106
- 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
- Arnegard, M. E., Mcgee, M. D., Matthews, B., Marchinko, K. B., Conte, G. L., Kabir, S., et al. (2014). Genetics of ecological divergence during speciation. *Nature* 511, 307–311. doi: 10.1038/nature13301
- 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
- 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
- Bitocchi, E., Rau, D., Bellucci, E., Rodriguez, M., Murgia, M. L., Gioia, T., et al. (2017). Beans (*Phaseolus* Ssp.) as a model for understanding crop evolution. *Front. Plant Sci.* 8:722. doi: 10.3389/fpls.2017.00722
- Blair, M., Diaz, L. M., Buendia, H. F., and Duque, M. C. (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. doi: 10.1007/s00122-009-1064-8
- Blair, M. W., Cortes, A. J., Farmer, A. D., Huang, W., Ambachew, D., Penmetza, R. V., et al. (2018). Uneven recombination rate and linkage disequilibrium

- across a reference snp map for common bean (*Phaseolus vulgaris* L.). *PLoS One* 13:e0189597. doi: 10.1371/journal.pone.0189597
- Blair, M. W., Cortés, A. J., Penmetza, R. V., Farmer, A., Carrasquilla-García, 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., Diaz, J. M., Hidalgo, R., Diaz, L. M., and Duque, M. C. (2007). Microsatellite characterization of andean races of common bean (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* 116, 29–43. doi: 10.1007/s00122-007-0644-8
- Blair, M. W., Iriarte, G., and Beebe, S. (2006). 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. doi: 10.1007/s00122-006-0217-2
- 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
- Bonifácio, E. M., Fonsêca, A., Almeida, C., Dos Santos, K. G. B., and Pedrosa-Harand, A. (2012). Comparative cytogenetic mapping between the lima bean (*Phaseolus lunatus* L.) and the common bean (*P. vulgaris* L.). *Theor. Appl. Genet.* 124, 1513–1520. doi: 10.1007/s00122-012-1806-x
- 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
- Chacon-Sanchez, M. I., and Martinez-Castillo, J. (2017). Testing domestication scenarios of lima bean (*Phaseolus lunatus* L.) in mesoamerica: insights from genome-wide genetic markers. *Front. Plant Sci.* 8:1551. doi: 10.3389/fpls.2017.01551
- Chevin, L. M., and Hospital, F. (2008). Selective sweep at a quantitative trait locus in the presence of background genetic variation. *Genetics* 180, 1645–1660. doi: 10.1534/genetics.108.093351
- 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., and Blair, M. W. (2018a). 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
- Cortés, A. J., and Blair, M. W. (2018b). “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). doi: 10.5772/intechopen.71669
- 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 Genetics* 13:58. doi: 10.1186/1471-2156-13-58
- Cortés, A. J., This, D., Chavarro, M. 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., Garzón, L. N., Valencia, J. B., and Madriñán, S. (2018). On the causes of rapid diversification in the páramos: isolation by ecology and genomic divergence in *espeletia*. *Front. Plant Sci.* 9:1700. doi: 10.3389/fpls.2018.01700
- 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., 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 eds *Mountains, Climate, and Biodiversity*, eds C. Hoorn, A. Perrigo, and A. Antonelli (New York, NY: Wiley).
- Cruickshank, T. E., and Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol. Ecol.* 23, 3133–3157. doi: 10.1111/mec.12796
- Csillery, K., Rodriguez-Verdugo, A., Rellstab, C., and Guillaume, F. (2018). Detecting the genomic signal of polygenic adaptation and the role of epistasis in evolution. *Mol. Ecol.* 27, 606–612. doi: 10.1111/mec.14499
- Díaz, L. M., and Blair, M. W. (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. doi: 10.1007/s00122-006-0417-9
- Duitama, J., Quintero, J. C., Cruz, D. F., Quintero, C., Hubmann, G., Foulquie-Moreno, M. R., et al. (2014). An integrated framework for discovery and genotyping of genomic variants from high-throughput sequencing experiments. *Nucleic Acids Res.* 42:e44.
- Ellegren, H., and Galtier, N. (2016). Determinants of genetic diversity. *Nat. Rev. Genet.* 17, 422–433. doi: 10.1038/nrg.2016.58
- Ellegren, H., Smeds, L., Burri, R., Olason, P. I., Backström, N., Kawakami, T., et al. (2012). The genomic landscape of species divergence in *ficidula* flycatchers. *Nature* 491, 756–760. doi: 10.1038/nature11584
- Ellegren, H., and Wolf, J. B. W. (2017). Parallelism in genomic landscapes of differentiation, conserved genomic features and the role of linked selection. *J. Evol. Biol.* 30, 1516–1518. doi: 10.1111/jeb.13113
- 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
- Fuentes-Pardo, A. P., and Ruzzante, D. E. (2017). Whole-genome sequencing approaches for conservation biology: advantages, limitations and practical recommendations. *Mol. Ecol.* 26, 5369–5406. doi: 10.1186/1471-2156-13-48
- 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. doi: 10.1371/journal.pone.0090346
- Glaubitz, J. C., Casstevens, T. M., Harriman, J., Elshire, R. J., Sun, Q., and Buckler, E. S. (2014a). Tassel-Gbs: a high capacity genotyping by sequencing analysis pipeline. *PLoS One* 9:e90346. doi: 10.1371/journal.pone.0090346
- Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., et al. (2014b). Tassel-Gbs: a high capacity genotyping by sequencing analysis pipeline. *PLoS One* 9:e90346. doi: 10.1371/journal.pone.0090346
- Gould, S. (1990). *Wonderful Life: The Burgess Shale and the Nature of History*. New York, NY: W.W. Norton and Company.
- Guerra-García, A., Suarez-Atilano, M., Mastretta-Yanes, A., Delgado-Salinas, A., and Pinero, D. (2017). Domestication genomics of the open-pollinated scarlet runner bean (*Phaseolus coccineus* L.). *Front. Plant Sci.* 8:1891. doi: 10.3389/fpls.2017.01891
- 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
- Hoorn, C., Wesselingh, F. P., Steege, H. T., Bermudez, M. A., Mora, A., Sevink, J., et al. (2010). Amazonia through time: andean uplift, climate change, landscape evolution, and biodiversity. *Science* 330, 927–931. doi: 10.1126/science.1194585
- Irwin, D. E., Alcaide, M., Delmore, K. E., Irwin, J. H., and Owens, G. L. (2016). Recurrent selection explains parallel evolution of genomic regions of high relative but low absolute differentiation in a ring species. *Mol. Ecol.* 25, 4488–4507. doi: 10.1111/mec.13792
- Jones, F. C., Chan, Y. F., Schmutz, J., Grimwood, J., Brady, S. D., Southwick, A. M., et al. (2012a). A genome-wide Snp genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. *Curr. Biol.* 22, 83–90. doi: 10.1016/j.cub.2011.11.045
- Jones, F. C., Grabherr, M. G., Chan, Y. F., Russell, P., Mauceli, E., Johnson, J., et al. (2012b). The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484, 55–61. doi: 10.1038/nature10944
- 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
- Koinange, M. K., Singh, S. P., and Gepts, P. (1996). Genetic control of the domestication syndrome in common bean. *Crop Sci.* 36, 1037–1045. doi: 10.2135/cropsci1996.0011183X003600040037x
- 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
- 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
- Lobaton, J. D., Miller, T., Gil, J., Ariza, D., De La Hoz, J. F., Soler, A., et al. (2018). Resequencing of common bean identifies regions of inter-gene pool introgression and provides comprehensive resources for molecular breeding. *Plant Genome* 11. doi: 10.3835/plantgenome2017.08.0068
- Lobkovsky, A. E., and Koonin, E. V. (2012). Replaying the tape of life: quantification of the predictability of evolution. *Front. Genet.* 3:246. doi: 10.3389/fgene.2012.00246
- Lotterhos, K. E., and Whitlock, M. C. (2015). The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Mol. Ecol.* 24, 1031–1046. doi: 10.1111/mec.13100

- Ma, T., Wang, K., Hu, Q., Xi, Z., Wan, D., Wang, Q., et al. (2018). Ancient polymorphisms and divergence hitchhiking contribute to genomic islands of divergence within a poplar species complex. *Proc. Natl. Acad. Sci. U.S.A.* 115, E236–E243. doi: 10.1073/pnas.1713288114
- 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
- Martin, S. H., Davey, J., Salazar, C., and Jiggins, C. (2018). Recombination rate variation shapes barriers to introgression across butterfly genomes. *bioRxiv* [Preprint]. 10.1101/297531
- Martin, S. H., and Jiggins, C. D. (2017). Interpreting the genomic landscape of introgression. *Curr. Opin. Genet. Dev.* 47, 69–74. doi: 10.1016/j.cdev.2017.08.007
- Martínez-Castillo, J., Camacho-Peírez, L., Villanueva-Viramontes, S., Andueza-Noh, R. H., and Chacón-Sainchez, M. I. (2014). Genetic structure within the mesoamerican gene pool of wild phaseolus lunatus (fabaceae) from Mexico as revealed by microsatellite markers: implications for conservation and the domestication of the species. *Am. J. Bot.* 101, 851–864. doi: 10.3732/ajb.1300412
- Maynard Smith, J., and Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genet. Res.* 23, 23–35. doi: 10.1017/S0016672300014634
- Meyer, R. S., Duval, A. E., and Jensen, H. R. (2012). Patterns and processes in crop domestication: an historical review and quantitative analysis of 203 global food crops. *New Phytol.* 196, 29–48. doi: 10.1111/j.1469-8137.2012.04253.x
- Motta-Aldana, J. R., Serrano-Serrano, M. L., Hernáiz-Torres, J., Castillo-Villamizar, G., and Debouck, D. G. (2010). Multiple origins of lima bean landraces in the Americas: evidence from chloroplast and nuclear DNA polymorphisms. *Crop Sci.* 50, 1773–1787. doi: 10.2135/cropsci2009.12.0706
- Nachman, M. W., and Payseur, B. A. (2012). Recombination rate variation and speciation: theoretical predictions and empirical results from rabbits and mice. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 409–421. doi: 10.1098/rstb.2011.0249
- 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
- Nielsen, R. (2005). Molecular signatures of natural selection. *Annu. Rev. Genet.* 39, 197–218. doi: 10.1146/annurev.genet.39.073003.112420
- Nosil, P., and Feder, J. L. (2011). Genomic divergence during speciation: causes and consequences. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 332–342. doi: 10.1098/rstb.2011.0263
- Pennisi, E. (2014). Disputed islands. *Science* 345, 611–613. doi: 10.1126/science.345.6197.611
- Pereira, R. J., Barreto, F. S., Pierce, N. T., Carneiro, M., and Burton, R. S. (2016). Transcriptome-wide patterns of divergence during allopatric evolution. *Mol. Ecol.* 25, 1478–1493. doi: 10.1111/mec.13579
- 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
- Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlovic, M., et al. (2017). Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *J. Evol. Biol.* 30, 1450–1477. doi: 10.1111/jeb.13047
- Ravinet, M., Westram, A., Johannesson, K., Butlin, R., Andre, C., and Panova, M. (2016). Shared and nonshared genomic divergence in parallel ecotypes of littorina saxatilis at a local scale. *Mol. Ecol.* 25, 287–305. doi: 10.1111/mec.13332
- Rendon-Anaya, M., Montero-Vargas, J. M., Saburido-Alvarez, 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
- Rodgers-Melnick, E., Bradbury, P. J., Elshire, R. J., Glaubitz, J. C., Acharya, C. B., Mitchell, S. E., et al. (2015). Recombination in diverse maize is stable, predictable, and associated with genetic load. *Proc. Natl. Acad. Sci. U.S.A.* 112, 3823–3828. doi: 10.1073/pnas.1413864112
- Roesti, M., Gavrillets, S., Hendry, A. P., Salzburger, W., and Berner, D. (2014). The genomic signature of parallel adaptation from shared genetic variation. *Mol. Ecol.* 23, 3944–3956. doi: 10.1111/mec.12720
- 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
- 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.baee.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
- Serrano-Serrano, M. L., Hernandez-Torres, J., Castillo-Villamizar, G., Debouck, D. G., and Chacon Sanchez, M. I. (2010). Gene pools in wild lima bean (*Phaseolus lunatus* L.) from the Americas: evidences for an andean origin and past migrations. *Mol. Phylogenet. Evol.* 54, 76–87. doi: 10.1016/j.ympev.2009.08.028
- Storz, J. F. (2005). Using genome scans of DNA polymorphism to infer adaptive population divergence. *Mol. Ecol.* 14, 671–688. doi: 10.1111/j.1365-294X.2005.02437.x
- 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
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595.
- Tenaillon, O., Rodriguez-Verdugo, A., Gaut, R. L., McDonald, P., Bennett, A. F., Long, A. D., et al. (2012). The molecular diversity of adaptive convergence. *Science* 335, 457–461. doi: 10.1126/science.1212986
- 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
- Wakeley, J. (2008). *Coalescent Theory: An Introduction*. Cambridge: Harvard University.
- Weir, B. S., and Cockerham, C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358–1370.
- 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: 10.1111/mec.14264
- Wolf, J. B., and Ellegren, H. (2017). Making sense of genomic islands of differentiation in light of speciation. *Nat. Rev. Genet.* 18, 87–100. doi: 10.1038/nrg.2016.133
- Zhou, L., Bawa, R., and Holliday, J. A. (2014). Exome resequencing reveals signatures of demographic and adaptive processes across the genome and range of black cottonwood (*Populus trichocarpa*). *Mol. Ecol.* 23, 2486–2499. doi: 10.1111/mec.12752

**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.

The reviewer MM and handling Editor declared their shared affiliation.

Copyright © 2018 Cortés, Skeen, Blair and Chacón-Sánchez. 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(s) 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.