Vega-Polo, Pamela, Cobo, Maria Mercedes, Argudo, Andrea, Gutierrez, Bernardo, Rowntree, Jennifer and Torres, Maria de Lourdes (2020) Characterizing the genetic diversity of the Andean blueberry (Vaccinium floribundum Kunth.) across the Ecuadorian Highlands. PLOS ONE, 15 (12). e0243420-e0243420.

DOI: https://doi.org/10.1371/journal.pone.0243420
Publisher: Public Library of Science (PLoS)
Version: Published Version
Downloaded from: https://e-space.mmu.ac.uk/626975/
Usage rights: Creative Commons: Attribution 4.0
Additional Information: This is an Open Access article published in PLOS ONE.

Enquiries:
If you have questions about this document, contact rsl@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party’s rights have been compromised through this document please see our Take Down policy (available from https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines)
Abstract

The Ecuadorian páramo, a high altitude tundra-like ecosystem, is a unique source of various ecosystem services and distinct biodiversity. Anthropogenic activities are associated with its fragmentation, which alters ecological factors and directly threatens resident species. Vaccinium floribundum Kunth., commonly known as Andean blueberry or mortiño, is a wild shrub endemic to the Andean region and highly valued in Ecuador for its berries, which are widely used in food preparations and hold an important cultural value. Since it is a wild species, mortiño could be vulnerable to environmental changes, resulting in a reduction of the size and distribution of its populations. To evaluate the extent of these effects on the mortiño populations, we assessed the genetic diversity and population structure of the species along the Ecuadorian highlands. We designed and developed a set of 30 species-specific SSR (simple sequence repeats) markers and used 16 of these to characterize 100 mortiño individuals from 27 collection sites. Our results revealed a high degree of genetic diversity ($H_E = 0.73$) for the Ecuadorian mortiño, and a population structure analyses suggested the existence of distinct genetic clusters present in the northern, central and southern highlands. A fourth, clearly differentiated cluster was also found and included individuals from locations at higher elevations. We suggest that the population structure of the species could be explained by an isolation-by-distance model and can be associated with the geological history of the Andean region. Our results suggest that elevation could also be a key factor in the differentiation of mortiño populations. This study provides an extensive overview of the species across its distribution range in Ecuador, contributing to a better understanding of its conservation status. These results can assist in the development of conservation programs for this valuable biological and cultural resource and for the páramo ecosystem as a whole.
**Introduction**

*Vaccinium floribundum* Kunth., commonly known as *mortiño* or Andean blueberry, is a woody perennial shrub from the Ericaceae. It is endemic to the Andean region in South America ranging from Venezuela to Bolivia and can be found between 1600 to 4500 meters above sea level (masl) [1]. *V. floribundum* grows in high-altitude ecosystems such as cool montane forests and páramos (tundra-like ecosystems) at temperatures ranging from 7 to 18˚C and displays adaptations to withstand frost and freezing conditions [1–3].

In Ecuador, *mortiño* is valued primarily for its black-purple fruit widely used for the preparation of traditional drinks, ice creams, wines and preserves [1,4]. The fruits have high concentrations of bioactive compounds such as polyphenols, antioxidants, anthocyanins, and flavonoids with potential beneficial effects on human health [4–6]. Phytochemical studies have reported anti-inflammatory, tumor suppressing and blood sugar regulation properties [5–12]. For instance, it has been shown that the proanthocyanidin-enriched fraction (PAC) from *V. floribundum* efficiently reduces lipid accumulation in vitro through the inhibition of adipogenesis [10].

*Mortiño* berries are hand-picked directly from the páramos and sold in local markets. To our knowledge, *V. floribundum* has not yet been domesticated and established as a crop. Despite various attempts, its propagation is problematic due to low rates of seed germination and a lack of knowledge regarding the factors required for its growth [13,14].

*V. floribundum* plays an important environmental and ecological role, being one of the first species that recovers after bouts of deforestation and man-made fires in the páramo ecosystems [15]. This is partly due to its high regenerative capacity driven by propagation from roots and other woody structures [16]. However, the páramo is considered a fragile ecosystem overall due to low levels of primary productivity and slow natural succession, which makes any recovery after human intervention relatively slow [17,18]. Different anthropogenic activities have affected the páramo ecosystem during the past century, and its fragmentation in the Andean region has directly impacted the diversity of its species [19–23]. Studies have shown that habitat fragmentation processes are associated with the generation of patches that prevent seed and pollen dispersal between populations, resulting in the isolation and reduction of the populations’ size [24–29]. The effects of habitat fragmentation and other anthropogenic factors on the conservation status of *V. floribundum* in the Ecuadorian páramo are currently unknown.

The genetic diversity of a number of species of the genus *Vaccinium* has been extensively studied, (e.g. *V. corymbosum* and *V. macrocarpon*) with the use of RAPD (random amplified polymorphic DNA), ISSR (inter simple sequence repeats), and SSR (simple sequence repeats) molecular markers, with a focus on plant breeding and horticultural crop programs [30–35]. However, genetic diversity studies of *Vaccinium* wild species, adapted to live in unique ecosystems like the páramo, are scarce. In a previous study, we used 11 heterologous (non-species-specific) SSR markers developed for *V. corymbosum* and found a moderate degree of *V. floribundum* genetic diversity in three provinces in northern Ecuador. However, this approach presented some limitations, mainly regarding the use of heterologous markers and the restricted sampling range [36]. Therefore, the aim of this study was to characterize the genetic diversity and population structure of *V. floribundum* with the use of species-specific SSR markers across the complete distribution range of the species in Ecuador. For this purpose, we designed and developed 30 species-specific SSR markers and used a subset of these (n = 16) for the genetic characterization of 100 *mortiño* individuals from different localities across all ten provinces in the Ecuadorian Highlands. The data presented here contribute to our understanding of the genetic makeup of this species in Ecuador.
Materials and methods

Sampling and DNA extraction

From June 2017 to August 2018, a total of 100 samples from individual plants were obtained from 27 collection sites (CS) distributed across ten provinces in the Ecuadorian Highlands: Carchi, Imbabura, Pichincha, Cotopaxi, Bolivar, Tungurahua, Chimborazo, Canar, Azuay, and Loja. Three to seven samples were collected at each site (S1 Table). Samples were collected between 2881 and 4160 masl and all individuals were georeferenced (geographical location and altitude) using a Garmin ETRex 10 Outdoor Handheld GPS Navigation Unit (Garmin International Inc.). For the purposes of this study, we considered the Ecuadorian Highlands latitudinal range (~657 km measured from north to south) and assigned each collection site to one of three regions: northern, central and southern (each region has a vertical extension of approximately 219 km). The northern region comprises individuals from CS1 (La Cofradia in Carchi) to CS11 (Sigchos in Cotopaxi), the central region includes individuals from CS12 (Quilotoa in Cotopaxi) to CS20 (Surimpalti in Cañar), and the southern region comprises individuals from CS21 (San Miguel in Cañar) to CS27 (Podocarpus in Loja) (Fig 1; S1 Table).

Five to seven young leaves were collected from each individual and transported at 4˚C to the Plant Biotechnology Laboratory at Universidad San Francisco de Quito, where they were stored at -20˚C. Total genomic DNA was extracted from frozen leaves using the CTAB method [37]. DNA was quantified by spectrophotometry using a Nanodrop 2000 (Thermo Scientific). Genetic data for specimens were obtained under the Genetic Resources Permit Number: MAE-DNB-CM-2016-0046 granted to Universidad San Francisco de Quito by Ministerio del Ambiente Ecuador, in accordance with the Ecuadorian law.

Development of SSR markers for V. floribundum

Microsatellite markers were developed from a single DNA extraction of V. floribundum using the Galaxy-based bioinformatics pipeline reported by Griffiths et al (2016) [38] at the University of Manchester genomics facility. The Illumina MiSeq platform was used with the shotgun 2 x 250 paired-end sequencing methodology (Nextera DNA Preparation Kit, Illumina, USA) and the sample used 0.33 of a flow cell. Post sequencing, data was first assessed for quality using FastQC [39], then filtered and trimmed using Trimmomatic [40]. Microsatellites were identified using Pal_finder v.0.02 [41], configured to search for sequences with a minimum of 8 repeated units, varying from two to six nucleotide motifs. The primers were designed using Primer3 [42,43]. Imperfect and interrupted repeats, loci where the primer sequence occurred in >1 reads and loci without primers were removed using the function Pal_filter [38]. Paired end reads were then assembled using Pal_filter and PANDASeq [38]. This pipeline produced a table of simple loci plus primers, ordered from the largest to the smallest repeat motif. Primer design was optimized for use with Platinum Taq DNA polymerase (Invitrogen, USA) with an optimal melting temperature (T_m) of 62˚C (Min 59˚C; Max 65˚C) and a maximum difference among primer pairs of 3˚C.

A total of 2 x 1,680,936 raw sequence reads was produced, with none flagged as poor quality. Sequence length ranged from 50–300 bp with a reported %GC content of 39. A total of 358 primer pairs were suggested, all amplifying SSRs with simple motifs. The majority of SSRs (333) were 2 bp motifs, but there were 22 with 3 bp motifs, one with 4 bp motif, one with 5 bp motif, and one with 6 bp motif.

From the initial 358 microsatellites markers, the first 30 were selected as candidates for our study; 25 with three or more base pairs motifs (3–6 bp), and five with 2 bp motifs. The melting temperature (Tm) of the primer sets ranged from 58˚C to 63˚C, with lengths between 18 and
25 bp (S2 Table). The primers were synthetized, including a universal tail (Tail A) in the locus-specific forward primers for genotyping. The universal tail A is complementary to fluorescently labelled universal primers (VIC, 6-FAM, NED or PET), which allows them to be
incorporated during the amplification, obtaining a fluorescently labelled amplicon ready to be genotyped [44].

Sample preparation and genotyping
Total genomic DNA (20 ng) was amplified in a 30μl reaction containing 1X PCR buffer, 2mM MgCl₂, 0.15μM modified locus-specific forward primer, 0.5μM reverse primer, 0.5μM universal Tail A with a fluorophore (VIC, 6-FAM, NED or PET), 0.2mM dNTPs, and 1U Platinum Taq Polymerase. The Tail A primers label the amplified products at the 5’ end through a three-primer system as reported by Blacket et al (2012) [44], as the modified locus-specific forward primer is exhausted in early cycles and the fluorophore (VIC, 6-FAM, NED or PET) gets incorporated into the PCR fragments in the subsequent cycles.

Amplification consisted of 15 min at 94˚C, followed by 40 cycles of 30 sec at 94˚C, 90 sec at the standardized annealing temperature (58–63˚C) (S2 Table), 60 sec a 72˚C, and a final elongation of 5 min at 72˚C. Successful PCR amplification was identified in 1.5% agarose gel electrophoresis, where one or two bands per locus were present. Labelled amplified products were genotyped by Macrogen (Seoul, Korea) on an ABI 3100 Genetic Analyzer (Applied Biosystems) automatic capillary sequencer, using 500 LIZ as a size standard.

Genetic analyses
SSR marker performance, genetic diversity and population genetics. GeneMarker software (SoftGenetics LLC) was used to identify individual alleles and allele size. Polymorphic information content values for each SSR locus were obtained with the polysat R-based statistical package [45]. Expected heterozygosity (Hₑ), observed heterozygosity (Hₒ) and the number of alleles (Nₐ) per locus and per collection site were estimated using the adegenet [46] and hierfstat [47] R packages. Private alleles were identified using the poppr R package [48], while null allele frequencies for all the SSR loci were determined with the FreeNA software through the EM algorithm [49]. Mean allelic richness (Aᵣ), standardized to the minimum sample size (n = 3) through rarefaction, was estimated using the diversity R package [50]. Fixation indices (Fₛₜ) were calculated using the adegenet R package [46] and described as estimates of inbreeding.

Population structure and genetic differentiation. A principal coordinate analysis (PCoA) was performed with the ade4 package [51] and the first three components plotted using the ggplot package [52]. In addition, a Bayesian analysis with an admixture model was performed using the program STRUCTURE 2.3.4 [53]. The potential number of genetic clusters (K) was evaluated between 1 and 10, with 10 independent repetitions for each K value. A 100,000 step burn-in period was used, followed by 1,000,000 Markov Chain Monte Carlo (MCMC) steps. The optimum value of K was evaluated through de Evanno method [54] using the web-based program STRUCTURE HARVESTER [55]. The independent replicates for the optimum value of K were aligned with the program CLUMPP [56] and the final STRUCTURE graph was plotted with DISTRUCT [57].

The R package poppr was used to perform an analysis of molecular variance (AMOVA) to evaluate genetic differentiation among and within the identified genetic clusters [48]. Pairwise Fₛₜ genetic distances between these clusters were estimated with the hierfstat package [47] using the Weir & Cockerham equation [58]. The same genetic distances, corrected for null alleles, were calculated with the FreeNA software [49]. Expected heterozygosity (Hₑ), observed heterozygosity (Hₒ), number of alleles (Nₐ), number of private alleles (Nₚₐ), and the mean allelic richness (Aᵣ) standardized (n = 14) through rarefaction, were also calculated for each of the clusters. The differences in the Hₑ between clusters were calculated through Monte-Carlo
tests using adegenet R package [46]. Bonferroni corrections for multiple paired comparisons were applied. Furthermore, directional relative migration was calculated and plotted using the diveRsity R-based package [50], based on Nei genetic distances, to better describe the patterns of gene flow between the identified genetic clusters.

Relationship between elevation, geographic distances and genetic variability of V. floribundum. Given the geographic coordinates and the elevation (masl) of each sampled individual, a Mantel test was performed (10,000 permutations) using the ape R statistical package [59], to evaluate the relationship between the genetic (Nei’s) and geographic distances among the sampled individuals. Furthermore, to determine whether elevation was associated with the genetic diversity of V. floribundum, stats [60] and ggpubr R [61] statistical packages were used to perform a Pearson correlation test between the expected heterozygosity ($H_E$) of individuals found at any given collection site and the mean elevation of that site.

Results
Validation of a set of species-specific SSR markers developed for V. floribundum

Of the 30 SSR markers tested, four were excluded due to the absence of PCR products, and an additional 10 presented different artifacts like stutter peaks and excessive baseline noise. The remaining 16 species-specific SSR markers showed reliable PCR amplification and highly fluorescent peaks (>1000 relative fluorescent units) after genotyping, and were selected for further analysis. These markers proved to be polymorphic and highly informative with an average polymorphic information content (PIC) value of 0.69. Mo025 (PIC = 0.90) and Mo020 (PIC = 0.88) were the most polymorphic markers, while Mo011 was the least polymorphic (PIC = 0.26) (Table 1). A total of 179 alleles were identified ($N_A$), with an average number of

Table 1. Genetic diversity indices of the 100 V. floribundum individuals analyzed with 16 SSR markers.

<table>
<thead>
<tr>
<th>Locus</th>
<th>T (°C)</th>
<th>Motif</th>
<th>Expected size (bp)</th>
<th>$N_A$</th>
<th>$H_O$</th>
<th>$H_E$</th>
<th>PIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mo001</td>
<td>58</td>
<td>(TTCCTG)48</td>
<td>300–400</td>
<td>10</td>
<td>0.61</td>
<td>0.82</td>
<td>0.79</td>
</tr>
<tr>
<td>Mo002</td>
<td>58</td>
<td>(TTTGG)50</td>
<td>100–200</td>
<td>6</td>
<td>0.39</td>
<td>0.58</td>
<td>0.49</td>
</tr>
<tr>
<td>Mo004</td>
<td>58</td>
<td>(TCC)33</td>
<td>300–400</td>
<td>9</td>
<td>0.45</td>
<td>0.79</td>
<td>0.77</td>
</tr>
<tr>
<td>Mo005</td>
<td>58</td>
<td>(ACC)33</td>
<td>200–300</td>
<td>12</td>
<td>0.46</td>
<td>0.87</td>
<td>0.85</td>
</tr>
<tr>
<td>Mo007</td>
<td>63</td>
<td>(TGC)24</td>
<td>200–300</td>
<td>9</td>
<td>0.35</td>
<td>0.58</td>
<td>0.50</td>
</tr>
<tr>
<td>Mo008</td>
<td>63</td>
<td>(ACC)27</td>
<td>300–400</td>
<td>11</td>
<td>0.41</td>
<td>0.77</td>
<td>0.74</td>
</tr>
<tr>
<td>Mo009</td>
<td>63</td>
<td>(AGT)24</td>
<td>400–500</td>
<td>7</td>
<td>0.25</td>
<td>0.59</td>
<td>0.55</td>
</tr>
<tr>
<td>Mo010</td>
<td>63</td>
<td>(TTG)30</td>
<td>250–350</td>
<td>11</td>
<td>0.32</td>
<td>0.66</td>
<td>0.62</td>
</tr>
<tr>
<td>Mo011</td>
<td>58</td>
<td>(TTC)24</td>
<td>300–400</td>
<td>6</td>
<td>0.22</td>
<td>0.26</td>
<td>0.26</td>
</tr>
<tr>
<td>Mo015</td>
<td>60</td>
<td>(TTG)27</td>
<td>200–300</td>
<td>8</td>
<td>0.24</td>
<td>0.63</td>
<td>0.57</td>
</tr>
<tr>
<td>Mo016</td>
<td>60</td>
<td>(TTG)33</td>
<td>350–450</td>
<td>13</td>
<td>0.45</td>
<td>0.79</td>
<td>0.76</td>
</tr>
<tr>
<td>Mo018</td>
<td>60</td>
<td>(TGC)27</td>
<td>200–300</td>
<td>9</td>
<td>0.42</td>
<td>0.72</td>
<td>0.69</td>
</tr>
<tr>
<td>Mo020</td>
<td>60</td>
<td>(ATT)33</td>
<td>300–400</td>
<td>20</td>
<td>0.59</td>
<td>0.89</td>
<td>0.88</td>
</tr>
<tr>
<td>Mo021</td>
<td>60</td>
<td>(TTTGG)33</td>
<td>300–400</td>
<td>16</td>
<td>0.49</td>
<td>0.88</td>
<td>0.87</td>
</tr>
<tr>
<td>Mo024</td>
<td>60</td>
<td>(ATC)24</td>
<td>300–400</td>
<td>12</td>
<td>0.56</td>
<td>0.88</td>
<td>0.87</td>
</tr>
<tr>
<td>Mo025</td>
<td>60</td>
<td>(TTTG)34</td>
<td>250–350</td>
<td>20</td>
<td>0.46</td>
<td>0.91</td>
<td>0.90</td>
</tr>
</tbody>
</table>

$N_A$ (number of alleles); $H_O$ (observed heterozygosity); $H_E$ (expected heterozygosity); and PIC (polymorphic information content). Optimum annealing temperature (in °C) for each set of primers and the expected size of the amplified band (or the allele found) are also presented. Primers sequences and SSR motifs are presented in the S2 Table.

https://doi.org/10.1371/journal.pone.0243420.t001
11.2 alleles per locus (ranging from 6 alleles for Mo002 to 20 alleles for Mo020 and Mo025) (Table 1). Null alleles presented low to mid frequencies across all the SSR loci, ranging from 0.034 for Mo002 to 0.225 for Mo025 (S3 Table). Expected heterozygosity ($H_E$) estimates ranged from 0.26 for Mo001 to 0.91 for Mo025 (Table 1).

**Genetic diversity of *V. floribundum* in the Ecuadorian Highlands**

After genotyping 100 individuals from all the collection sites, an average number of 41.7 unique alleles per collection site were identified. CS18 (Quimiac) ($N_A = 65; H_E = 0.6$) and CS14 (Carihuairazo) ($N_A = 60; H_E = 0.61$), in the central region, presented the highest number of alleles and expected heterozygosity estimates. In contrast, CS24 (Cruces) ($N_A = 28; H_E = 0.28$) and CS23 (Cajas) ($N_A = 26; H_E = 0.24$) in the southern region, and CS12 (Quilotoa) ($N_A = 26; H_E = 0.21$) in the central region presented the lowest genetic diversity indicators (S4 Table). Similar results were found when analyzing the mean allelic richness ($A_R$) per collection site, an estimate that corrects for the sampling imbalance between collection sites (S4 Table). The global expected heterozygosity ($H_E = 0.73$) (Table 1) revealed a high degree of genetic diversity for *V. floribundum* in the Ecuadorian Highlands.

**Population structure and genetic differentiation**

A principal coordinate analysis (PCoA) was conducted to test genetic similarities between individuals and to define possible groupings. Two-dimensional plots were obtained (Fig 2) for the first three components explaining 42.83% of the total variance (PC1 = 21.25%; PC2 = 13.36%; PC3 = 8.22%). Overall, no clear differentiation between regions was observed; nonetheless, most of the individuals sampled in the northern region separated from the individuals sampled in the southern region through the variance explained by the second component (PC2) (Fig 2). Individuals from the central region were distributed along the PC2. Furthermore, a clearly differentiated group was identified (marked by a green ellipse, Fig 2) which included individuals from CS12 (Quilotoa) in the central region, and CS22, CS23, and CS24 (collection sites in Azuay) in the southern region. Individuals from this group were mainly separated from the others through the variance explained by the first component (PC1) (Fig 2).

![Fig 2. Principal coordinate analysis (PCoA) of *V. floribundum* using 16 SSR markers. The first three components represent 42.83% of the total variation. Individuals were assigned to their defined latitudinal region (northern, central and southern). One main group was identified (green ellipse) and was separated by the first component (PC1).](https://doi.org/10.1371/journal.pone.0243420.g002)
A Bayesian analysis was performed to better evaluate the genetic structure of *V. floribundum* in the Ecuadorian Highlands. The STRUCTURE results showed four possible genetic lineages (*K* = 4), based on the highest Δ*K* calculated through the Evanno method [54] (S5 Table). Individuals were assigned to four clusters based on their predominant inferred ancestry (Fig 3A). Cluster 1 contained most of the individuals from the northern region (in red), cluster 2 included the individuals from the central region plus CS9 (Lloa) and CS11 (Sigchos) from the northern region (in yellow), and cluster 3 comprised the individuals from the southern region plus CS19 (Cerro Abuga) and CS20 (Surimpalti) from the central region (in blue) (Fig 3A). Notably, cluster 4 was highly consistent with the group identified in the PCoA (Fig 2), comprising the individuals from CS12 (Quilotoa in Cotopaxi) in the central region, and CS22, CS23 and CS24 (collection sites in Azuay) in the southern region (in green) (Fig 3A). When the number of possible genetic lineages was reduced to *K* = 3 (Fig 3B) and *K* = 2 (Fig 3C), cluster 4 remain grouped separately, showing a clear differentiation of this cluster from the others.

An analysis of molecular variance (AMOVA) shows most of the variation (78.64%) occurs within the genetic clusters, and only 21.36% between these clusters (*p* = 0.001) (S6 Table). Corrected *F*<sub>ST</sub> genetic distances show that greater differentiation occurs between cluster 4 and clusters 1 (*F*<sub>ST</sub> = 0.171), 2 (*F*<sub>ST</sub> = 0.249) and 3 (*F*<sub>ST</sub> = 0.230). The lowest genetic distance was found between clusters 1 and 2 (*F*<sub>ST</sub> = 0.078), suggesting higher genetic similarities between these clusters (Table 2).
The F\textsubscript{ST} were estimated according to Weir and Cockerham (1984) and corrected for null alleles through FreeNA software. Correction for null alleles decreases the possibility of bias in the estimation of genetic differentiation values. Higher values indicate greater genetic differentiation between the analyzed clusters.

Finally, genetic diversity indices for each cluster were calculated (Table 3). The number of alleles (\(N\textsubscript{A}\)) ranged from 57 alleles for cluster 4 to 120 alleles for cluster 2. The private alleles for each genetic cluster ranged between 8 (14.03%) for cluster 4 and 25 for cluster 2 (20.83%). The highest expected heterozygosity (\(H\textsubscript{E}\)) estimate was observed for clusters 1 (\(H\textsubscript{E} = 0.67\)), 2 (\(H\textsubscript{E} = 0.65\)) and 3 (\(H\textsubscript{E} = 0.68\)); with cluster 4 showing a significantly lower genetic diversity (\(H\textsubscript{E} = 0.39, p = 0.001\)) (Table 3). No statistical differences in \(H\textsubscript{E}\) were found between cluster 1 and 2 (\(p = 0.259\)), 1 and 3 (\(p = 0.777\)), or 2 and 3 (\(p = 0.331\)). The lowest diversity indices were found in the four locations assigned to cluster 4 (CS12, CS22, CS23, and CS24) (S4 Table).

### Directional genetic differentiation and relative migration between V. floribundum genetic clusters

A directional relative migration network was inferred for the four genetic clusters, given their spatially structured distribution. These results revealed different levels of gene flow and possible migration rates between the four genetic clusters (Fig 4). As observed from genetic distances, clusters 1 and 2 presented the greatest genetic similarity (represented by the shorted distances between nodes). In contrast, cluster 4 presented the greatest differentiation in relation to the others, again in agreement with the results obtained by the genetic distances' analyses (Table 2). A relatively high rate of potential bidirectional asymmetric gene flow was found between clusters 1 and 2, being higher from cluster 2 to cluster 1 than in the opposite direction. In turn, both clusters present moderate rates of bidirectional asymmetric gene flow with cluster 3. Remarkably, none of the clusters shows evidence of gene flow with cluster 4, which in turn presents minimal unidirectional gene flow with the other clusters (Fig 4).

### Table 2. Pairwise F\textsubscript{ST} values between the four genetic clusters identified for V. floribundum in the Ecuadorian Highlands.

<table>
<thead>
<tr>
<th></th>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 2</td>
<td>0.078</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>0.114</td>
<td>0.120</td>
<td>-</td>
</tr>
<tr>
<td>Cluster 4</td>
<td>0.171</td>
<td>0.249</td>
<td>0.230</td>
</tr>
</tbody>
</table>

https://doi.org/10.1371/journal.pone.0243420.t002

### Table 3. Genetic diversity indices of the four V. floribundum identified genetic clusters.

<table>
<thead>
<tr>
<th>Clusters</th>
<th>(N)</th>
<th>(N\textsubscript{A})</th>
<th>(N\textsubscript{PA})</th>
<th>(A\textsubscript{E})</th>
<th>(H\textsubscript{O})</th>
<th>(H\textsubscript{E})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 1</td>
<td>35</td>
<td>118</td>
<td>18</td>
<td>2.25</td>
<td>0.45</td>
<td>0.67</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>31</td>
<td>120</td>
<td>25</td>
<td>2.09</td>
<td>0.48</td>
<td>0.65</td>
</tr>
<tr>
<td>Cluster 3</td>
<td>20</td>
<td>103</td>
<td>15</td>
<td>2.03</td>
<td>0.51</td>
<td>0.68</td>
</tr>
<tr>
<td>Cluster 4</td>
<td>14</td>
<td>57</td>
<td>8</td>
<td>1.65</td>
<td>0.08</td>
<td>0.39</td>
</tr>
</tbody>
</table>

\(N\) (number of individuals); \(N\textsubscript{A}\) (number of alleles); \(N\textsubscript{PA}\) (number of private alleles); \(A\textsubscript{E}\) (mean allelic richness calculated using rarefaction); \(H\textsubscript{O}\) (observed heterozygosity); and \(H\textsubscript{E}\) (expected heterozygosity) are given for all 100 individuals genotyped at 16 SSR loci distributed in four genetic clusters.

https://doi.org/10.1371/journal.pone.0243420.t003
Relationship between elevation, geographic distances and genetic variability of *V. floribundum*

Mantel test results revealed a significant correlation between genetic and geographic distances \((r = 0.29; p = 0.0001)\) between individuals, suggesting that a possible isolation by distance dynamic might explain the observed geographic structure of the studied populations. Furthermore, we observed a significant negative correlation between elevation and genetic diversity (evaluated through expected heterozygosity estimates) among the different collection sites through the three regions \((r = -0.53; p = 0.0045)\) (Fig 5; S7 Table). The regression equation \((H_E = 0.968 - 0.00015 \times \text{Elevation})\) can be approximated as a mean genetic diversity decrease of 0.00015 per every 1-meter increment in elevation along our altitude sampling range.

**Discussion**

**Genetic diversity of *V. floribundum* in the Ecuadorian Highlands**

The mean number of alleles \((N_A = 11.2 \text{ alleles per locus})\) and the global expected heterozygosity \((H_E = 0.73)\) found in this study reveal a high degree of genetic diversity for *V. floribundum* in the Ecuadorian Highlands. In a previous study, where we assessed the genetic diversity of *V. floribundum* in the north of Ecuador, we reported on average 6.1 alleles per locus and an expected heterozygosity \((H_E)\) of 0.49, representing a moderate degree of genetic diversity [36]. Differences in both the number of alleles and the expected heterozygosity between these studies could be explained by the use of heterologous markers (designed for *V. corymbosum*) and the collection of individuals in a limited geographical area. The use of heterologous markers could generate the presence of null alleles that tend to bias the estimation of heterozygosity and inflate \(F_{ST}\) values [62–64]. Furthermore, the geographical range and distribution of populations have a significant impact on the amount of genetic variation of a species [65] and could...
explain the higher degree of genetic diversity found in this study. This can be particularly true for instances where geographically distant individuals are also genetically more distinct, highlighting the importance of sampling range as well as density of collection sites within a geographical region. A greater genetic variability has been previously reported for species distributed over a wide geographical range with respect to those found in a specific limited area [66–68].

The total number of alleles obtained for *V. floribundum* in this study is similar to those reported for wild *V. angustifolium* and *V. corymbosum* clones [35], and higher compared to those reported for *V. macrocarpon* cultivars [69]. Several studies report that intensive selection and inbreeding during domestication processes reduce the genetic diversity in many plant species and increase genetic drift [70,71], which could explain why wild *Vaccinium* species are more diverse than their domesticated counterparts. The global expected heterozygosity (He) in this study is similar to the He reported for *V. sieboldii* (H_E = 0.73) and *V. ciliatum* (H_E = 0.75), both wild endangered species [72]. A high level of genetic diversity in these species could indicate random mating among individuals, a common feature of wild plant species [71,72].

Fig 5. Linear regression between *V. floribundum* genetic diversity and elevation. The plot shows the correlation between elevation (x-axis) and the genetic diversity (y-axis), assessed through the expected heterozygosity estimates per collection sites (confidence interval = 95%). Each data point represents one collection site where the individuals were sampled in this study (total of 27 collection sites).

https://doi.org/10.1371/journal.pone.0243420.g005
Population structure and genetic differentiation

The Bayesian approach in STRUCTURE showed that *V. floribundum* individuals were grouped into four genetic clusters (Fig 3). From these, only individuals from cluster 4 also grouped separately in the PCoA (Fig 2), and greatly differentiated from the other clusters based on genetic distances ($0.171 \leq F_{ST} \leq 0.249$). This cluster comprises individuals from CS12 (Quilotoa) in the central region, and CS22 (Toreadora), CS23 (Cajas) and CS24 (Cruces) in the southern region. Beyond their genetic similarities, these individuals share the uppermost elevations among the collection sites included in this study (3929–4160 masl) (S1 Table). This higher strata of the páramo altitudinal range has been described as a distinct ecosystem known as superpáramo [73,74], and its stark ecological differences compared to the lower altitude settings could account for some of the observed differences in the mortiño populations established here. First, the effects of elevation on the genetic diversity and differentiation on the species are highlighted by our results that show a correlation between the expected heterozygosity of individuals in specific collection sites and their elevation ($r = 0.29; p = 0.0001$), with (the high-altitude) cluster 4 presenting the lowest expected heterozygosity ($H_E = 0.27\pm0.06$) and exclusive allele number ($N_{PA} = 8$) (Table 3). Observations of two woody plant species growing in altitudinal clines (*Castanopsis eyeri* and *Daphniphyllum oldhamii*) suggest that factors like temperature, phenological characteristics, low population densities and small effective population sizes could play a role in the isolation and reduction of genetic diversity of populations at higher elevations [75].

Furthermore, the evident differentiation of this cluster of high-elevation populations can be explained by an isolation by elevation model where gene flow from lower to higher elevations is reduced [75,76]. Different mechanisms can drive this altitude-dependent isolation, such as reproductive barriers due to phenological shifts that results in differences in flowering time between higher and lower altitudes [74,76]. It is also possible that altitude clines affect pollinator behavior: pollinator visits are less frequent at higher altitudes [77,78], which could explain the results obtained in the migration network where the gene flow was minimal and unidirectional from cluster 4 to the other clusters (0.21–0.39) (Fig 4). As a consequence of these types of phenomena, habitats at higher elevations (like in the superpáramo) separated by lower elevation valleys can behave like islands where gene flow between different locations is decreased. This kind of “sky island” dynamics can further explain the larger divergence between the populations from cluster 4 and the rest of the mortiño genetic diversity in the Ecuadorian Andes.

Nonetheless, the questions of the similarities between geographically distant locations within cluster 4 (i.e. CS12 versus CS22, CS23 and CS24) and the genetic patterns observed in the remaining clusters persist. These patterns might be associated with the geological history and geography of the Andean region, important forces that drive the evolution and distribution of different species in the region [79,80]. In accordance with our previous study [36], the Mantel test results confirmed a significant correlation between genetic and geographic distances ($r = 0.29; p = 0.0001$), suggesting that an isolation-by-geographic distance model also explains part of the population structure of *V. floribundum* in Ecuador. Under this model, physical barriers could play an additional role in the reduced genetic flow between clusters [81,82], particularly between the first three genetic clusters which show moderate to high differentiation ($0.078 \leq F_{ST} \leq 0.120$) (Table 2). The moderate degree of differentiation ($F_{ST} = 0.078$) between clusters 1 and 2 could be explained by the emergence of mountainous barriers in the Andean Highlands during the Miocene, as we previously suggested [36] based on results which are consistent with the current study. However, an alternative hypothesis could be related to the spread of ice masses throughout the Andean region during the Pleistocene, a phenomenon that has been proposed as a factor that directly affected the diversity and
structure of different species [83,84]. The contraction of the páramos during this period could have caused V. floribundum populations to be separated by ice masses, reducing or preventing their gene flow and partially isolating both clusters [85]. During the glacial retreat (i.e. melting of the ice masses), individuals from clusters 1 and 2 could have come into contact again, a phenomenon known as secondary contact [84,86].

A third driving force of the observed population structure might include the role of adaptation of V. floribundum to different ecosystems, prompting both diversifying selection and convergent evolution. For instance, the physical characteristics of different soil types, determined by their distinct geological origins, partially overlap with the distribution of specific genetic clusters. For instance, individuals from clusters 1 and 2 grow over Andisols developed from volcanic projections characteristic to the northern Ecuadorian highlands, encompassing the northern region and most of the central region in this study (CS1-18). The presence of multiple volcanoes and eolian deposits of volcanic material in these areas have generated a distinct pyroclastic layer over the soils [83]. In contrast, individuals from cluster 3 are found on a wide monotonous plateau. These individuals grow on an old ferritic soil base, also known as metamorphic rock, which is characterized by the accumulation of clay [83]. The geological origin of the soils could then explain some of the differentiation between cluster 3 and both clusters 1 and 2 (FST1 = 0.114; FST2 = 0.120), as the unique characteristics between the volcanic and metamorphic rock could act as a force for diversifying selection. Similarly, extreme climatic conditions at higher elevations (characterized by low temperatures and limited nutrient availability [74,87]) can serve as effective selective pressures. Plants adapt to these conditions through strategies such as size reduction, clonal propagation that favors survival over long periods of climatic oscillation, and the fixation of specific alleles [88–90]. This kind of adaptive process could explain the genetic similarities of geographically distant populations within cluster 4 (Fig 2) even in the absence of gene flow. However, the interplay between isolation-by-distance and isolation-by-elevation scenarios is probably more nuanced, as studies have reported that under the latter, gene flow is higher between sites at similar elevations than across elevation clines [76,91].

Future studies should address the occurrence of different adaptations in mortiño individuals living in specific habitats (such as the superpáramo) and explore the relationship between genetic diversity and the origin of the soils where the mortiño grows.

Conclusions

The genetic data presented in this study demonstrates that V. floribundum displays a high degree of genetic diversity in the Ecuadorian Highlands. Population structure analyses revealed that individuals grouped into four genetic clusters, distributed according to their geographic location (northern, central and southern regions). While genetic distances are associated with geographic distances between individuals, it is noteworthy that elevation also has an effect on the genetic diversity of the populations. Interestingly, the group identified in the PCoA (comprising individuals from CS12, CS22, CS23 and CS24) was consistent with the cluster 4, found through the bayesian analysis. Individuals from this cluster were collected at higher altitudes, presented lower genetic diversity and a higher degree of genetic differentiation from the remaining populations, as well as low gene flow with the populations assigned to the other clusters. It is evident that the structure of V. floribundum populations in Ecuador was driven by an interplay between isolation-by-distance and isolation-by-elevation dynamics, and the geography of the Andean region possibly plays a key role in driving these dynamics through a combination of different geographic features, historical climatic fluctuations and the establishment of scenarios favoring the action of adaptive processes.
Furthermore, the high values of genetic diversity that were found are encouraging for the conservation of this species. Since this study assessed the genetic diversity and population structure of wild *V. floribundum* in the ten provinces that encompass the Ecuadorian High-lands, the results presented here could serve as a basis for the development of conservation programs for this species and its habitat at the páramo, a relevant and unique ecosystem.

**Supporting information**

S1 Table. Information for the 27 *V. floribundum* Collection Sites (CS) from 3 defined regions in the Ecuadorian Highlands.

(PDF)

S2 Table. List of the 30 species-specific SSR markers designed for *V. floribundum*.

(PDF)

S3 Table. Estimate of null allele frequencies for each analyzed *V. floribundum* genetic cluster and SSR locus. Mean: the mean frequencies over the four genetic clusters.

(PDF)

S4 Table. Summary of the diversity parameters for each collection site where *V. floribundum* individuals were collected. $N_A$ (number of alleles); $N_{PA}$ (number of private alleles); $A_R$ (mean allelic richness calculated using rarefaction); $H_O$ (observed heterozygosity); $H_E$ (expected heterozygosity); and $F_{IS}$ (fixation index) are given for all 100 individuals genotyped at 16 SSR loci across 27 collection sites in three geographic regions.

(PDF)

S5 Table. Values for estimating the optimum K from the analysis in STRUCTURE with an admixture model.

(PDF)

S6 Table. Results of the analysis of molecular variance (AMOVA) performed for the *V. floribundum* genetic clusters that were identified in the population structure analysis ($p = 0.001$).

(PDF)

S7 Table. Output for general linear model for the effect of the elevation over the heterozygosit.

(PDF)

**Acknowledgments**

We would like to thank Diego Urquía for his contribution in data processing, as well as Dario Ramirez and Andrea Pinos for their assistance in the collection of the samples. We also thank all the members of the Plant Biotechnology Laboratory (Universidad San Francisco de Quito) who contributed to this research.

**Author Contributions**

**Conceptualization**: Maria de Lourdes Torres.

**Data curation**: Pamela Vega-Polo, Maria Mercedes Cobo.

**Formal analysis**: Pamela Vega-Polo, Maria Mercedes Cobo.

**Funding acquisition**: Maria de Lourdes Torres.
**Investigation:** Pamela Vega-Polo, Andrea Argudo, Maria de Lourdes Torres.

**Methodology:** Pamela Vega-Polo, Maria Mercedes Cobo, Andrea Argudo, Jennifer Rowntree, Maria de Lourdes Torres.

**Project administration:** Maria de Lourdes Torres.

**Resources:** Jennifer Rowntree, Maria de Lourdes Torres.

**Software:** Pamela Vega-Polo, Bernardo Gutierrez.

**Supervision:** Maria Mercedes Cobo, Maria de Lourdes Torres.

**Validation:** Bernardo Gutierrez.

**Visualization:** Pamela Vega-Polo, Maria Mercedes Cobo, Bernardo Gutierrez, Maria de Lourdes Torres.

**Writing – original draft:** Pamela Vega-Polo, Andrea Argudo, Jennifer Rowntree.

**Writing – review & editing:** Pamela Vega-Polo, Maria Mercedes Cobo, Bernardo Gutierrez, Maria de Lourdes Torres.

**References**


---

Characterizing the genetic diversity of Ecuadorian V. floribundum
Characterizing the genetic diversity of Ecuadorian *V. floribundum*


