






Article

Descriptive Genomic Analysis and Sequence Genotyping of the Two Papaya Species (*Vasconcellea pubescens* and *Vasconcellea chilensis*) Using GBS Tools

Basilio Carrasco ^{1,*}, Bárbara Arévalo ¹, Ricardo Perez-Diaz ¹, Yohaily Rodríguez-Alvarez ², Marlene Gebauer ², Jonathan E. Maldonado ^{3,4} , Rolando García-González ⁵, Borys Chong-Pérez ⁵ , José Pico-Mendoza ⁶, Lee A. Meisel ⁷ , Ray Ming ⁸  and Herman Silva ^{3,*} 

¹ Centro de Estudios en Alimentos Procesados (CEAP), Talca 3480094, Chile

² Departamento de Ciencias Vegetales, Facultad de Agronomía e Ingeniería Forestal, Pontificia Universidad Católica de Chile, Santiago 7820436, Chile

³ Laboratorio de Genómica Funcional y Bioinformática, Facultad de Ciencias Agronómicas, Universidad de Chile, Santiago 8820808, Chile

⁴ Laboratorio de Multiómica Vegetal y Bioinformática, Departamento de Biología, Facultad de Química y Biología, Universidad de Santiago de Chile, Santiago 9160000, Chile

⁵ Sociedad de Investigación y Servicios, BioTECNOS Ltda., San Javier 3660000, Chile

⁶ Facultad de Ingeniería Agronómica, Universidad Técnica de Manabí, Portoviejo 130105, Ecuador

⁷ Laboratorio de Genética Molecular Vegetal, Instituto de Nutrición y Tecnología de los Alimentos, Universidad de Chile, Santiago 7830490, Chile

⁸ Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

* Correspondence: bcarrasco@ceap (B.C.); hesilva@uchile (H.S.); Tel.: +56-979971795 (B.C.); +56-229785726 (H.S.)



Citation: Carrasco, B.; Arévalo, B.; Perez-Diaz, R.; Rodríguez-Alvarez, Y.; Gebauer, M.; Maldonado, J.E.; García-González, R.; Chong-Pérez, B.; Pico-Mendoza, J.; Meisel, L.A.; et al. Descriptive Genomic Analysis and Sequence Genotyping of the Two Papaya Species (*Vasconcellea pubescens* and *Vasconcellea chilensis*) Using GBS Tools. *Plants* **2022**, *11*, 2151. <https://doi.org/10.3390/plants11162151>

Academic Editor: Tika Adhikari

Received: 11 July 2022

Accepted: 3 August 2022

Published: 18 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: A genotyping by sequencing (GBS) approach was used to analyze the organization of genetic diversity in *V. pubescens* and *V. chilensis*. GBS identified 4675 and 4451 SNPs/INDELs in two papaya species. The cultivated orchards of *V. pubescens* exhibited scarce genetic diversity and low but significant genetic differentiation. The neutrality test yielded a negative and significant result, suggesting that *V. pubescens* suffered a selective sweep or a rapid expansion after a bottleneck during domestication. In contrast, *V. chilensis* exhibited a high level of genetic diversity. The genetic differentiation among the populations was slight, but it was possible to distinguish the two genetic groups. The neutrality test indicated no evidence that natural selection and genetic drift affect the natural population of *V. chilensis*. Using the *Carica papaya* genome as a reference, we identified critical SNPs/INDELs associated with putative genes. Most of the identified genes are related to stress responses (salt and nematode) and vegetative and reproductive development. These results will be helpful for future breeding and conservation programs of the Caricaceae family.

Keywords: *Vasconcellea*; SNPs/INDELs; GBS

1. Introduction

The family Caricaceae belongs to Brassicales, including papaya, mustard, and *Arabidopsis thaliana*, among other species [1]. Caricaceae includes five American genera, *Carica*, *Vasconcellea*, *Horovitzia*, *Jarilla*, and *Jacaratia*, and one African genus, *Cylicomorpha*. It has a total of 35 species, most of them described as diploids ($2n = 2x = 18$), with variable reproductive systems including monoecy, dioecy, and polygamy [2–5].

Carica papaya L. is the most successfully cultivated family and is considered to be the fourth most important tropical fruit crop in the world [6]. However, the *Vasconcellea* genus, also known as highland papayas or mountain papayas [7], is also relevant for biodiversity and genetics. This is a valuable genetic resource for *Carica papaya* genetic improvement because some species have shown resistance to PRSV-P (Papaya ringspot virus-P), cold tolerance, and a higher sugar content than tropical papayas [8].

Vasconcellea pubescens (ex *V. cundinamarcensis*) is a perennial, herbaceous, and trioecy cross-pollinated species with male, female, and hermaphrodite plants [9]. Its life cycle (Figure 1A) begins after emerging from a brief fall-winter break between June and August, when the average temperature reaches >20 °C. In springtime, the growth of new leaves, flowers, and fruits starts, and the fruits of the previous season also begin to mature. This process is maintained continuously between September and May of each year. Female and hermaphroditic plants can reach 4 m in height (Figure 2B) and present white flowers (Figure 2C,D); their fruits are big (180–200 g), spherical to cylindrical, and their color varies from green in immature fruit to yellow in ripe fruit (Figure 2C–E); they contain around 200 seeds.

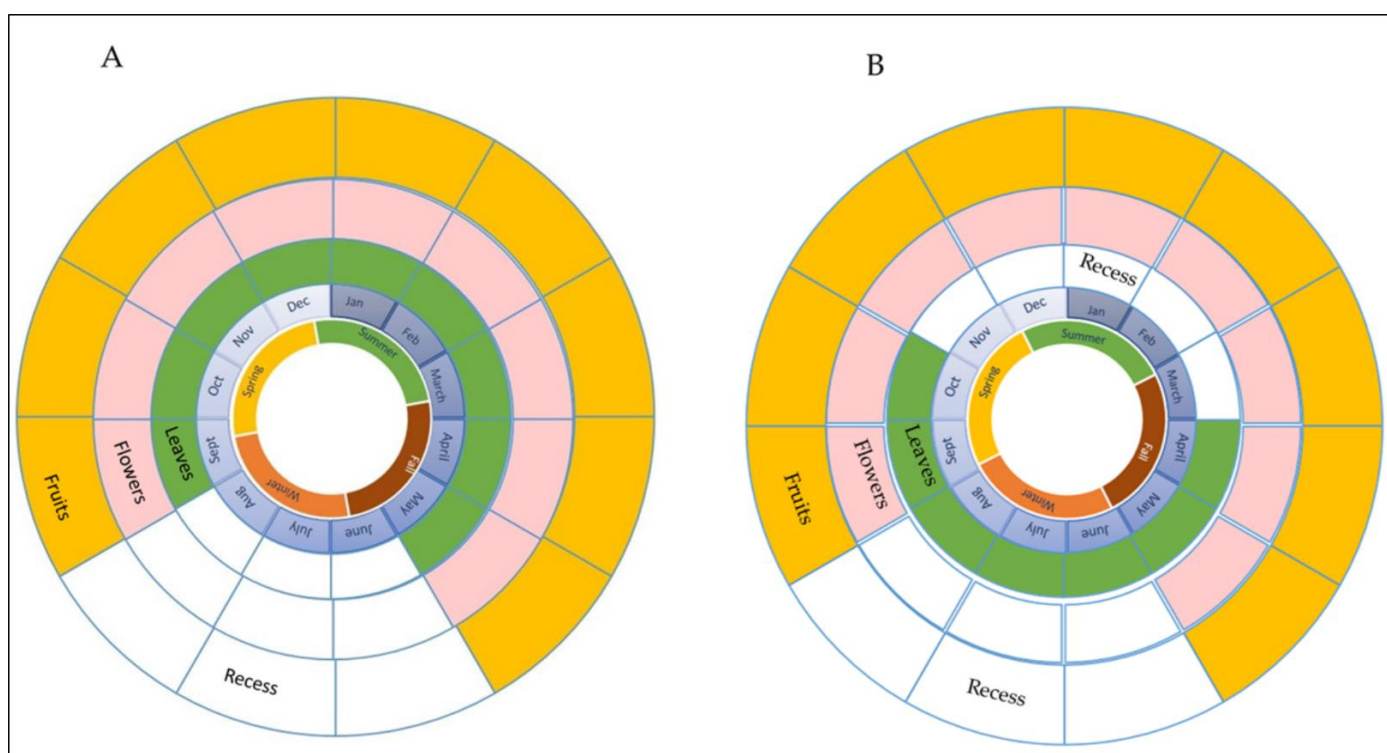


Figure 1. Life cycle of *V. pubescens* (A) and *V. chilensis* (B). Green, pink, and yellow indicate the period of presence and the development of new leaves, flowers, and fruits. White indicates the growth arrest of leaves, flowers, and fruits in *V. pubescens*, while, for *V. chilensis*, it indicates the fall of leaves, flowers, and fruits.

V. pubescens is cultivated in several South American countries [6,10]. In Chile, it is grown from $29^{\circ}54' \text{ S}$ – $71^{\circ}15' \text{ W}$ (La Serena, Chile) to $36^{\circ}08' \text{ S}$ – $72^{\circ}47' \text{ W}$ (Cofquecura, Chile), always under moderate temperatures (between 10 °C and 25 °C) because it cannot tolerate frost or temperatures over 30 °C. For years, its fruits have been used to produce canned fruit, honey, juice, and candies to satisfy the local market, and, recently, these have been exported to the United States and Europe. The origin of *V. pubescens* is in the Andes Mountain between Colombia and Peru [11]. The time of its introduction to northern Chile is still not precise, but the first Spanish chronicles indicate that native people had already cultivated them before Spanish conquerors arrived in this region around 1535 [12].

Another *Vasconcellea* species in Chile is *V. chilensis* (Planch. ex A. DC.) Solms (“Palo Gordo”, native papaya, or austral papaya), a small, endemic, deciduous, and succulent shrub with female and hermaphrodite plants (andro-gynodioecy).

It grows in semi-arid environmental conditions. Its life cycle starts with the development of its leaves at the beginning of April (autumn), when temperatures reach an average of 10 °C, and it loses them in November when the temperature rises above 20 °C. In springtime (September), the first flowers appear, followed by the fruits, which continue

until May of the following year. From November to May, the plant without leaves only maintains flowers and fruits (Figure 1B); the photosynthesis activity is carried out through its herbaceous stems.

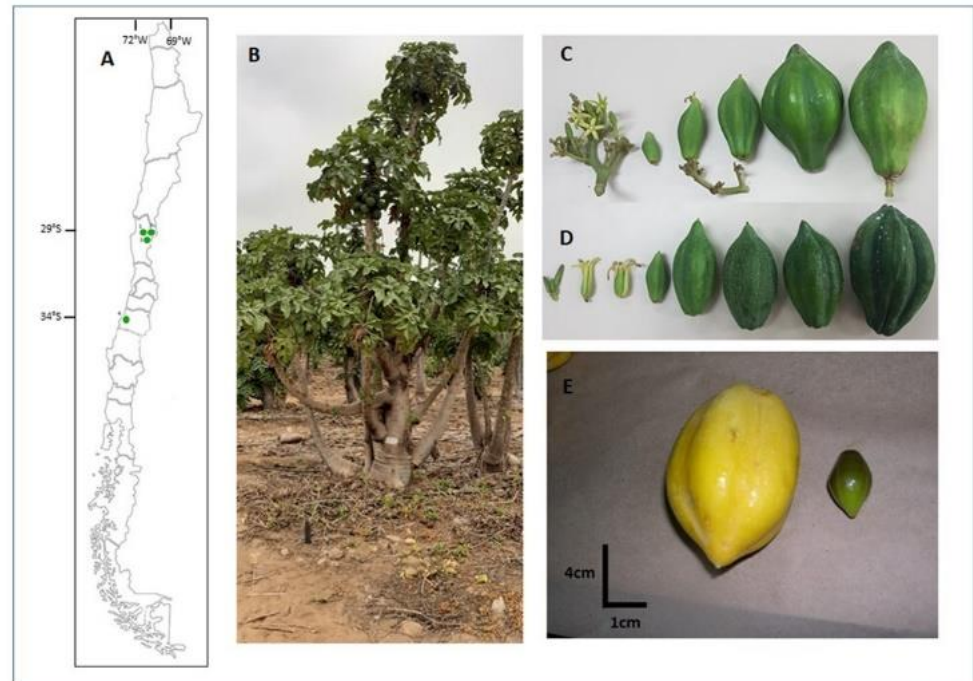


Figure 2. (A) Geographic location (29° to 34° South Latitude) of the four orchards of *V. pubescens* sampled in this study: Guallarauco Ltda, La Serena, Chile (1), Sociedad Agrícola HC, La Serena, Chile (2), AgroSaturno SA, La Serena, Chile (3), and Frutos de Lipimavida Ltda, Lipimavida, Chile (4). (B) A 4 m-tall plant of *V. pubescens* grows in an orchard in La Serena, Chile. (C) Hermaphrodite and (D) female flowers and immature fruits. (E) Yellow (*V. pubescens*, 180 g) and green (*V. chilensis*, 4 g) ripe fruits.

Female and hermaphrodite plants present small reddish flowers (Figure 3C). Its fruits are small and of various shapes; their color varies from green to brown (Figure 3D), although it is also possible to find some yellow ones; every fruit contains 6 to 7 seeds. Its current fragmented natural populations are between 29°25' S–71°17' W and 30°42' S–71°22' W (Figure 3A); thus, it has been classified as a vulnerable species [13]. Due to its capacity to resist abiotic and biotic stresses, it is a relevant resource for genetic improvements in *Carica* and *Vasconcellea* species [14,15].

Despite their importance, there have been few genetic studies on *Vasconcellea* species. The additional analysis of these two species will generate valuable information for future conservation and breeding strategies for the *Vasconcellea* and *Carica* genus [6]. Most genetic studies have been carried out on *Carica papaya* [16–18]. Co-dominant markers such as simple sequence repeats (SSRs) (i.e., microsatellites) [19–22] have been identified and used to analyze genetic diversity [16] and develop genetic linkage maps [19,23–26]. The SSRs isolated in *Carica papaya* have shown a low transferability to other *Caricaceae* [27]; thus, these molecular markers have not been used routinely to evaluate *Vasconcellea* species. Inter-simple sequence repeats (ISSRs) have been used to analyze the genetic structure of cultivated and natural populations of *V. pubescens* and *V. chilensis*, respectively [28,29]. However, the genetic information obtained from dominant markers (for example, ISSR) has limited use in genomic applications because the homology of co-migrating amplification products will need to be clarified, and the frequency of heterozygotes is unknown because they are indistinguishable from homozygotes [16,30,31].

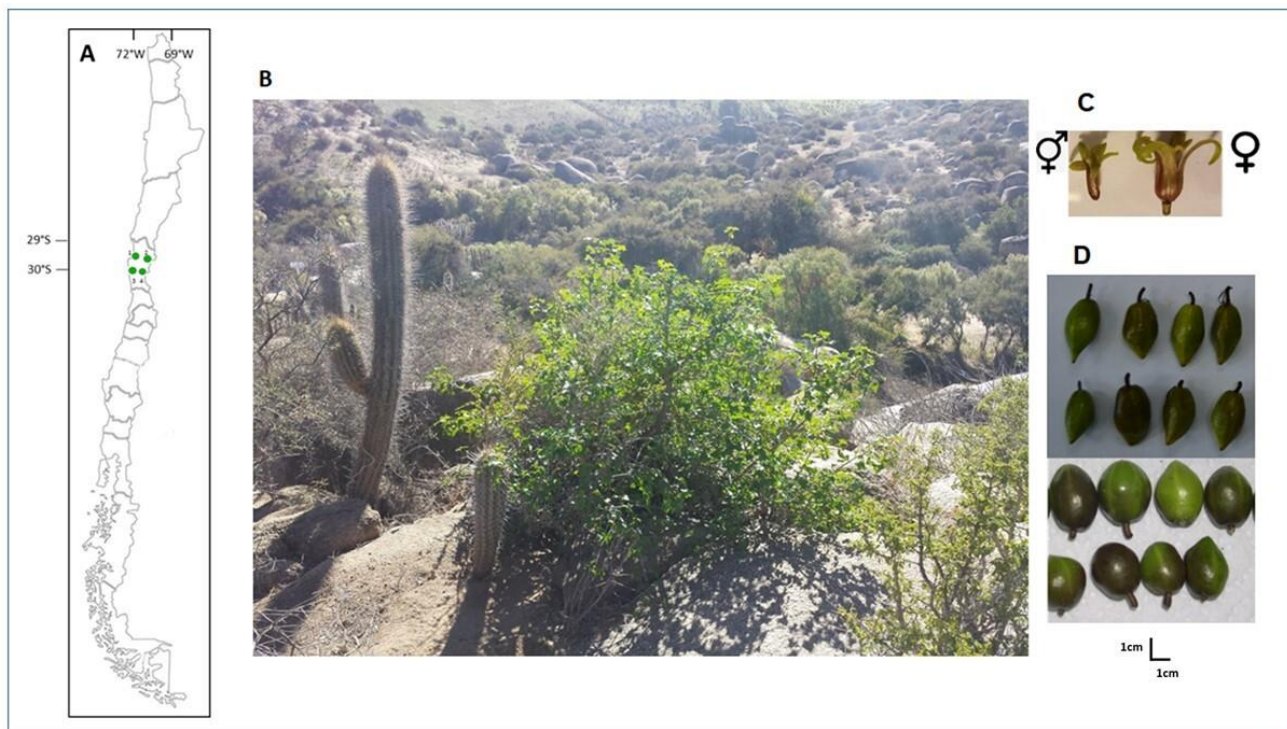


Figure 3. (A) Geographic location (29° to 30° South Latitude) of the four natural populations sampled in this study: Chungungo (1), Huachalalume (2), National Park Fray Jorge (3), and Valle del Encanto (4). (B) Shrub of *V. chilensis* grows in its natural semi-desert habitat; leaves can be seen from April to October. (C) Female and hermaphrodite flowers with reddish coloration. (D) Piriform and oval fruits (3–5 g).

Next-generation sequencing approaches (NGSs) have enabled the release of whole-genome reference sequences for many plant species [32,33]. On the other hand, the draft genome of *Carica papaya* L. is available as a model for tropical fruit trees [1,34]. In this regard, the genome size of *C. papaya* is 372 Mbp, where 24,746 genes have been identified, with an average gene length of 2373 bp per gene. Additionally, it has been estimated that repetitive sequences accounted for about 52% of the papaya genome, where transposable elements are the most abundant (51.9%), followed by microsatellite sequences (0.19%) [1,34].

The use of next-generation sequencing technology has revealed hundreds of thousands of single nucleotide polymorphisms (SNPs), which are the most abundant type of polymorphism found in eukaryotic genomes [35,36]. Unlike microsatellites, the power of SNPs comes not from the number of alleles but from a large number of loci assessed simultaneously and their genome-wide distribution.

A genotyping method based on partial genome sequencing has been developed, which uses restriction enzymes to digest the genome, reducing its complexity; after the DNA digestion, barcoded adapters are linked to the DNA fragments, which are then sequenced by high-throughput methods, obtaining hundreds of thousands of SNPs simultaneously [37,38]. Those methods have been called reduced-representation sequencing, one of which is genotyping by sequencing (GBS).

The GBS approach is well suited for genetic analysis because it provides useful information for developing conservation and breeding strategies. This information can help to identify priority populations to be conserved, select proper parental lines, and identify advanced breeding lines from the progenies. Additionally, the published reference genome of *C. papaya* has enabled the analysis of the physical location of the SNPs and INDELs on the genome, showing the relationships of collinearity and synteny between related species [39,40]. GBS has been used to study several cultivated plant species [41],

such as barley [37], wheat [42], rice [43], soybean [44], potato [45], apricot [46], peach [47], Chrysanthemum [48], and Japanese plum [49].

The use of mass sequencing tools and the *Carica papaya* genome sequence as a reference can be a valuable alternative to identify molecular markers such as SNPs and INDELs in *Vasconcellea* species. To date, there has been no report describing the SNP/INDEL diversity in any species of the family Caricaceae. Thus, a comparative genomics analysis among related *Vasconcellea* species would facilitate the identification of biologically relevant polymorphisms to study the evolutionary processes affecting their natural and cultivated populations and enable the development of molecular markers for breeding and conservation programs [50].

This study analyzed the genetic diversity of *V. chilensis* and *V. pubescens* using SNPs/INDELs and identified some critical SNPs/INDELs associated with putative genes for future conservation and breeding programs for the Caricaceae family.

2. Results

2.1. Genetic Diversity and Structure

The *C. papaya* genome (372 Mb) was used as a reference to identify SNPs/INDELs. In this regard, the GBS had 5 Mb and 4.7 Mb coverage for *V. pubescens* and *V. chilensis*, respectively. The sequencing depth was 14X for *V. pubescens* and 27X for *V. chilensis*.

From 92 samples of *V. pubescens*, 232,682,246 reads were obtained by GBS. Of them, 89% contained barcode and restriction sites. Finally, 14,213,301 reads were accepted in TOPM. After quality filtering, 4675 SNPs/INDELs were used for genetic analysis (Supplementary Materials File S1).

Low levels of polymorphic loci ($p = 62\%$ [se] = 1.12%) and heterozygosity (0.04) were found in cultivated *V. pubescens* (Table 1). The average nucleotide diversity per site (π) was 0.0032. Heterozygosity and nucleotide diversity did not exhibit significant differences between populations. Tajima's D test showed a negative value ($D = -2.281$) and was significantly different from zero according to the beta distribution at 95% confidence [51]. The F_{IS} value was significant ($F_{IS} = 0.50$ standard error [se] = 0.006), indicating deviation of the Hardy–Weinberg equilibrium and heterozygous deficits; additionally, the average selfing rate (s) was high ($s = 0.67$), suggesting a high level of inbreeding.

Table 1. Summary of population genetic statistics for *V. pubescens* and *V. chilensis*.

Species	N	L	Ho	π (10^{-3})	D
<i>V. pubescens</i>	92	4675	0.04	3.2	−2.281
Guallarauco Ltd.a	24		0.05	3.5	
Saturno SA	23		0.04	3.2	
Sociedad Agrícola HC	24		0.04	3.0	
Frutos de Lipimávida	21		0.05	3.1	
<i>V. chilensis</i>	94	4451	0.12	15.4	−1.065
Chungungo	26		0.12	17.0	
Fray Jorge	27		0.13	17.8	
Huachalalume	24		0.10	15.2	
Valle del Encanto	17		0.12	11.4	

N, number of individuals; L, number of loci (SNPs/INDELs); Ho, observed heterozygosity; π , mean pairwise difference between two sequences; D, Tajima's D; $= \pi - q/\sqrt{\text{var}(\pi - q)}$, where $q = \text{expected } \pi$.

According to the analysis of molecular variance (AMOVA), most of the SNPs/INDELs were distributed among the studied individuals (97%), and the genetic differentiation was low among orchards ($\phi_{PT} = 0.03$; $p < 0.001$).

In *V. chilensis*, the number of samples by population was 26, 27, 24, and 17 from “Chungungo”, National Park “Fray Jorge”, “Huachalalume”, and Historical Park “Valle del Encanto”, respectively. It is relevant to highlight that this sample size was small because *V. chilensis* is scarce through its remnant natural populations. In total, 310,093,649 reads were obtained, and, finally, 257,142,135 reads were accepted using TOPM. After filtering, 4451 SNPs/INDELs were obtained for population genetic analysis (Supplementary Materials File S1). In *V. chilensis*, a low level of polymorphic loci ($p = 63.8\%$ [se] = 3.16%) was estimated, but a moderate level of heterozygosity (0.12) was observed (Table 1). Additionally, the inbreeding coefficient (F_{IS}) and selfing rate were 0.18 and 0.31, respectively. The nucleotide diversity per site (π) was 0.0154, and it was similar between populations. Tajima’s D neutrality test showed a negative value ($D = -1.065$) and was not significantly different from zero according to the beta distribution at 95% confidence.

The SNPs/INDELs were mainly distributed among individuals (98%); only a 2.6% genetic differentiation was observed between the populations ($\phi_{PT} = 0.026$; $p < 0.001$).

2.2. Comparative Genomics

Genomic analysis allowed us to identify 730 common polymorphic SNPs/INDELs between *V. pubescens* and *V. chilensis*. Those SNPs/INDELs were subjected to AMOVA to determine their relevance for the interspecific and intraspecific genetic differentiation (ϕ_{PT}). SNPs/INDELs were selected when they reached ϕ_{PT} values equal or superior to 0.15 after 1000 permutations ($p < 0.001$). The AMOVA identified 58 SNPs/INDELs (Supplementary Materials File S1), which explained 53% ($\phi_{PT} = 0.53$ $p < 0.001$) of the genetic differences between *V. pubescens* and *V. chilensis*. The Bayesian analysis and Evanno methodology for those 58 SNPs/INDELs indicated that the best estimation of delta K was reached when the K value was equal to 2 (delta K = 430.049 for $k = 2$), clearly separating both *Vasconcellea* species (Figure 4A).

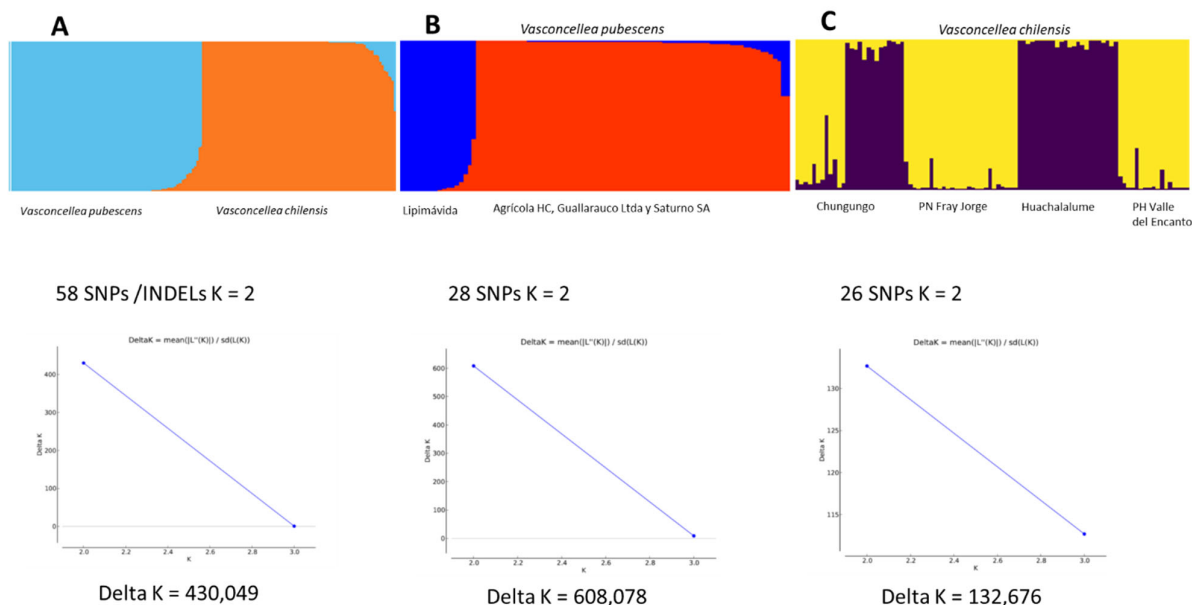


Figure 4. Assignment rate (Q values) using informative SNPs/INDELs. (A) *V. pubescens* vs. *V. chilensis* using 58 SNPs/INDELs. (B) Cultivated orchards of *V. pubescens* using 28 SNPs/INDELs. (C) Natural populations of *V. chilensis* using 26 SNPs/INDELs.

The *Carica papaya* genome (ASGPBv0.4) was used as a reference to locate the 58 SNPs/INDELs over the exon, intron, and intergenic regions. Most of the SNPs/INDELs

(72%) were positioned over exons, 17% were over introns, and only 11% were located in intergenic regions (Supplementary Materials File S1). The heterozygosity (H_o) for those 58 SNPs/INDELs was 26% for *V. pubescens* and 52% for *V. chilensis*, showing a negative correlation ($r = -0.316$; $p = 0.0156$) between both species. Out of the 58 SNPs/INDELs, 52 were positioned over 30 genes according to the *Carica* genome. Some of those SNPs/INDELs reported extreme genetic differentiation values ($\phi_{PT} > 0.5$; p) between *V. pubescens* and *V. chilensis*: tubulin alpha-2 chain ($\phi_{PT} = 0.91$); phosphatidylinositol monophosphate 5 kinase ($\phi_{PT} = 0.91$); nucleoporin 156 ($\phi_{PT} = 0.89$); cyclic nucleotide-regulated ion channel family protein ($\phi_{PT} = 0.88$); remorin family protein ($\phi_{PT} = 0.84$); autoinhibited Ca^{2+} -ATPase 1 and 2 ($\phi_{PT} = 0.77$); signal transduction histidine kinase; hybrid-type, ethylene sensor ($\phi_{PT} = 0.73$); S-adenosyl-L-methionine-dependent methyltransferase superfamily protein ($\phi_{PT} = 0.70$); nucleoporin 155 ($\phi_{PT} = 0.67$); DNA repair (Rad51) family protein ($\phi_{PT} = 0.62$); sec7 domain-containing protein ($\phi_{PT} = 0.61$); Deoxyxylulose-5-phosphate synthase ($\phi_{PT} = 0.57$); and disease resistance-responsive (dirigent-like protein) family protein ($\phi_{PT} = 0.52$).

When the orchards of *V. pubescens* were compared, 33.4% ($\phi_{PT} = 0.334$, $p < 0.001$) of the genetic differentiation was determined by 28 SNPs/INDELs (Supplementary Materials File S1), and the Bayesian analysis identified two clusters ($K = 2$; $\Delta K = 608.078$; Figure 4B). Out of the 28 SNPs/INDELs analyzed, 27 were located in intergenic regions, and only 1 SNP was found over an exon (PAC: 16403913.CDS.1) of the putative gene that codifies for the protein CK203. This SNP showed a high level of genetic differentiation ($\phi_{PT} = 0.615$). For *V. chilensis*, 26 SNPs/INDELs (Supplementary material S1) explained 42% ($\phi_{PT} = 0.42$, $p < 0.001$) of the genetic differentiation among four natural populations. The Bayesian clustering distinguished two groups ($K = 2$, $\Delta K = 132.676$; Figure 4C). The first group includes the southern populations “Valle del Encanto” and “Fray Jorge”, and the second cluster includes the northern populations “Chungungo” and “Huachalalume”. Out of the 26 SNPs/INDELs, 24 were located in intergenic regions. Furthermore, only 2 SNPs/INDELs were found over coding regions—one over an intron (PAC: 16403969) and the other one over an exon (PAC:16403933 CDS.1). These two loci showed a high level of genetic differentiation ($\phi_{PT} = 0.306$ and $\phi_{PT} = 0.44$, respectively) and were located in the annotated gene RING/U-box superfamily protein.

3. Discussion

Currently, no public information is available for Caricaceae species regarding population genetic diversity using polymorphisms identified by GBS. Some research has been carried out using dominant molecular markers such as RAPD [52–54], AFLP [55], ISSR [28,29], resistance gene analogs [21], and co-dominant markers such as SSR [15,18,56].

At the DNA level, primary molecular variations are single nucleotide polymorphisms (SNPs), one to several hundred base-pair insertions or deletions (INDELs), and variations in the number of tandem repeats (i.e., simple sequence repeat: SSR). Although the SSRs are considered a relevant DNA variation for genetic analysis [57], the SNPs and INDELs have gained strong acceptance because new next-generation technologies have made them more accessible to researchers. The main drawback of SNPs/INDELs is related to the possibility that they are bi-allelic, making them less variable than other genomic variants such as SSR. However, they are ubiquitous, high in frequency, and widely distributed throughout the genome. Sequencing technologies and bioinformatics tools have enabled the discovery of thousands of SNPs and INDELs for population structure studies of wild and cultivated plant species and serve as a valuable genetic tool to identify candidate polymorphisms related to phenotypic variation [58,59]. For instance, the power of SNPs/INDELs comes from the large number of loci that can be assessed simultaneously instead of the number of alleles per locus [60]. SNPs identified in papaya species started to be used only in the last ten years. Nucleotide variabilities for ITS, atpB-rbcL, and trnL-F spacers in *Jacaratia Mexicana* have been used to identify cultivars in papaya [58,61]. More recently, SNPs/INDELs have been used to study sexual determinations in species of the Caricaceae family. In this regard,

14,528 nucleotide substitutions and 965 short indels associated with sexual identity have been identified [62]. Additionally, a low nucleotide diversity has been found in the X-linked region of papaya [63]; SNPs in candidate genes associated with sex determination have been described elsewhere [64].

To the best of our knowledge, our study is the first report on the organization of genomic diversity in *V. chilensis* and *V. pubescens* using SNPs/INDELs obtained by GBS. In this regard, it was possible to identify 4451 to 4675 high-quality SNPs/INDELs with a depth of coverage of 14X (5 Mb) and 27X (4.7 Mb). We found low levels of genetic variation in these two *Vasconcellea* species, which contrasts with the estimation of genetic diversity reported for plant species with similar biological characteristics such as autogamy, allogamy, and seed dispersal by gravity or ingestion [65].

The cultivated *V. pubescens* showed a high inbreeding level ($F_{is} = 0.50$), and the selfing rate ($s = 0.67$) suggested assortative mating. The scarce gene diversity for this species had previously been reported using 114 inter-simple sequence repeat loci (ISSR [28]). The heterozygosity was inferior to that reported for other allogamous fruit species such as peach ($H_o = 0.28$ [66]) and grape ($H_o = 0.39$ [67]).

The nucleotide diversity ($\pi = 3.2 \times 10^{-3}$) was similar to those reported for apple ($\pi = 3.7 \times 10^{-3}$ [68]), peach ($\pi = 2.1 \times 10^{-3}$ [66]), grape ($\pi = 5.1 \times 10^{-3}$ [67]), Chrysanthemum ($\pi = 5.5 \times 10^{-3}$ [48]), Loblolly pine ($\pi = 3.9 \times 10^{-3}$ [67]), and *Thymus vulgaris* ($\pi = 3.1 \times 10^{-3}$ [68]).

However, the nucleotide diversity in *V. pubescens* was superior to that found for *Carica papaya* (π autosome = 1.7×10^{-3} and π sexual chromosome = 0.38×10^{-3} [63]).

The domestication process (selection and seed propagation) may have played an essential role in reducing the genetic diversity for *V. pubescens* in Chile [63], and it could explain the high level of genetic similarity between individuals in each orchard. There are no records of introducing new germplasms from its origin (high mountains of Colombia and Ecuador), which could have reduced the availability of new genetic variants.

The highland papaya (*V. pubescens*) has been cultivated in northern Chile since the 16th century [12]. Interestingly, the seedlings used to establish new orchards are produced from selected plants from the same orchard; this process is repeated every 6 to 8 years. If this propagation system had been applied for a hundred years, it might have affected the level and organization of its genetic diversity. Additionally, crossing between relatives into the orchards could be responsible for the high homozygosity currently observed. Moreover, *V. pubescens* is self-compatible and displays both cross-pollination and self-pollination simultaneously [6,69].

The genetic differentiation observed between northern and southern orchards ($\phi_{PT} = 0.03$, $p < 0.001$) can be explained because seedlings established in the southern orchards represent a small portion of the genetic diversity of the northern orchards, and the selection practices applied by the growers may have increased the genetic differentiation between them. These results are coincident with those previously reported [28]. Tajima's D statistic, which is used to detect divergence from neutrality, showed significant negative values ($D = -2.281$), suggesting that the pattern of genetic diversity observed in *V. pubescens* has been affected by a selective sweep and population expansion after a bottleneck during its domestication [51].

In contrast, the natural populations of *V. chilensis* showed higher heterozygosity ($H_o = 0.120$) and nucleotide diversity ($p = 15.4 \times 10^{-3}$) than cultivated *V. pubescens* but also a lower inbreeding coefficient ($F_{is} = 0.18$) and selfing rate ($s = 0.3$).

The genetic differentiation among populations was low but significant ($\phi_{PT} = 0.026$; $p < 0.001$). The Bayesian analysis identified two genetic groups ($K = 2$), including individuals from the four analyzed populations. The Mantel test ($r(AB) = 0.32$, $p = 0.05$) indicated that there is a slight relationship between the geographic location of the populations and the pattern of genetic diversity. The Tajima D neutrality test ($D = -1.065$) was negative but not significant, suggesting that there is no evidence that natural selection and genetic drift can affect genetic diversity.

V. chilensis is a gynodioecious species that favors allogamy and heterozygosity. However, high levels of homozygosity and a high selfing rate were determined. The fragmentation and the small population size can partly explain these results. The highly fragmented populations are separated from each other, 45 to 85 km apart. Our results showed a tendency for isolation by distance between southern (“Valle del Encanto” Historical Park and “Fray Jorge” National Park) and northern populations (“Chungungo” and “Huachalalume”). Therefore, the pattern of genetic diversity in *V. chilensis* may reflect the outcome of its natural history; thus, the current fragmentation of its populations may reflect recent events that have not yet affected its genetic structure.

From a conservation point of view, the populations located in the “Valle del Encanto” and “Fray Jorge” are protected because the former are in a sector declared a National Historic Monument, and the latter are within a National Park. However, the “Huachalalume” and “Chungungo” populations are not protected and thus suffer from intense anthropogenic pressure.

Comparative genomics between species and populations allowed for the tagging of some genes annotated in the *C. papaya* genome. Several genes were found to be associated with abiotic and biotic stress responses and organogenesis. At the interspecific level, 58 SNPs/INDEL explained 53% of the genetic differentiation between *V. pubescens* and *V. chilensis*, of which 90% were tagged over exons and introns of 30 genes, most of them involved in vegetative and reproductive processes associated with adaptations to different stress conditions (Supplementary Materials File S1).

The SNPs/INDEL with the highest significant value of genetic differentiation were tagged over genes such as phosphatidylinositol monophosphate 5 kinase gene, which codes for an enzyme localized in the apical plasma membrane and adjacent cytosolic region of pollen tubes. It has been implicated in vesicle trafficking and cytoskeletal rearrangements during the carbohydrate metabolic process, cellular amino acid metabolic process, and pollen tip growth [70]. Other genes tagged by SNPs/INDELS that have been described in detail are Nucleoporin 155, 156 codes for proteins located in the nucleolus, plasma membrane, chloroplast, and nuclear pore complex, where they are implicated in nucleocytoplasmic transporter activity. The amino acid sequences of the Nup155 orthologous are highly conserved in the evolution of eukaryotes [71]. The S-adenosyl-L-methionine-dependent methyltransferase superfamily protein gene codes for a protein with methyltransferase activity. This gene is located in the Golgi apparatus, playing an important role in shoot development, the homogalacturonan biosynthetic process, the response to cytokinin stimuli, and root development [72]. DNA repair (Rad51) family protein genes code for a protein involved in the homologous recombination and repair of DNA. Its expression is restricted to pollen mother cells in anthers and megaspore mother cells in ovules [73,74]. The Sec7-domain-containing protein gene codes for a protein involved in the specification of apical-basal pattern formation. This protein is essential for cell division, expansion, and adhesion, and it is located in the endoplasmic reticulum and Golgi apparatus. It is related to endosomal trafficking, which plays an important role in regulating plant growth and development in optimal and stress conditions [75]. Tubulin alpha-2 chain is part of a gene family that codes for a structural protein constituent of the cytoskeleton (tubulin complex, cytosol, cell wall, plasmatic membrane). Tubulins are dimeric proteins involved in controlling fundamental processes such as cell division, the polarity of growth, cell wall deposition, intracellular trafficking, communications, and the response to salt stress [76,77]. Cyclic nucleotide-regulated ion channel family protein genes are involved in the defense response and calcium ion import; they enable the diffusion of calcium ions through a transmembrane aqueous pore or channel [78]. They have also been identified as putative heat sensors, which might play a key role in regulating the heat stress response [79] and modifying proteins located in the plasma membrane, where they participate in plant-microbe interaction signals [80] and virus trafficking through plasmodesmata and the plasma membrane [81]. They are also involved in the process of dehiscence and aging tissues, where mature branched plasmodesmata are predominant [82]. Autoinhibited

Ca²⁺-ATPase 1 and 2 genes are self-association proteins located in the plasma membrane and involved in calcium-transport, ATPase activity, and calmodulin-binding. They have also been associated with the response to nematode wounding-related signals in leaves and roots [83].

Signal transduction histidine kinase, a hybrid-type gene, codes for a protein located in membrane components, where it can sense ethylene and affect the metabolism of ABA and other plant hormones such as auxin, cytokinins, and gibberellic acid. It has been implicated in flavonol synthesis, seed germination, and salt tolerance [84].

The deoxyxylulose-5-phosphate synthase gene codes for a protein located in chloroplasts involved in the chlorophyll biosynthetic process, isopentenyl diphosphate biosynthetic process, mevalonate-independent pathway, and response to the light stimulus and mycorrhiza interaction [85,86]. The disease-resistance-responsive (dirigent-like protein) family protein gene codes for a protein located in the endomembrane system. It is believed to be involved in the lignan biosynthetic process and defense response [87]. Their evolutionary history could explain the genetic differences observed between *V. pubescens* and *V. chilensis*.

Other SNPs/INDELs with a high value of genetic differentiation were located over other additional genes, of which thirteen have been annotated, and their putative function has been described. Ten of these genes have been involved in plant development, and three have been associated with stress conditions (more details in Supplementary Materials File S1).

Most of the genetic differentiation among populations of *V. pubescens* was determined by SNPs/INDELs located in intergenic regions. Only one was found over an exon (PAC:16403913.CDS.1), which showed high genetic differentiation. This exon was located in the putative gene CK203 (42% homology), which responds to stress conditions mediated by abscisic acid [88].

Similarly, for *V. chilensis*, only two variants were found over coding regions (PAC:16403969 and PAC:16403933 CDS.1); both showed a high level of genetic differentiation and were located in the gene RING/U-box superfamily protein, which encodes a novel E3 ubiquitin ligase that acts as a central negative regulator of floral organ size [89].

V. pubescens and *V. chilensis* have certainly been subjected to different events during their natural history. Domestication, selection, genetic drift, and inbreeding have affected the genetic structure of both species in different ways. Most of the genetic differentiation among populations has been observed in intergenic regions, probably associated with the control of gene expression, although it is not possible to rule out random events.

In this study, we used GBS to understand the genomic diversity and population structure of the cultivated *V. pubescens* and wild *V. chilensis* to identify genetic polymorphisms (SNPs/INDELs) associated with interpopulation and interspecific differences between these two papaya species. Further research could focus on the isolation of these genes reported here; moreover, screening of their expression patterns will provide valuable information to carry out future conservation and genomics-assisted breeding plans.

4. Materials and Methods

4.1. Plant Material

Young leaves were collected from 92 and 94 adult plants of *V. pubescens* and *V. chilensis*, respectively. The samples were frozen in liquid nitrogen and stored at -80°C until DNA extraction.

Three orchards of *V. pubescens* were collected in the Coquimbo Region (Figure 2A). The orchards were Guallarauco Ltd.a. ($32^{\circ}23'31''\text{ S}-71^{\circ}22'29''\text{ W}$, 210 masl), Sociedad Agrícola HC ($29^{\circ}5'639''\text{ S}-71^{\circ}09'12''\text{ W}$, 180 masl), and AgroSaturno SA ($29^{\circ}59'25''\text{ S}-71^{\circ}01'38''\text{ W}$, 264 masl). Additionally, one population was collected from Frutos de Lipimavida Ltd.a., located in the Maule Region ($34^{\circ}84'42''\text{ S}-72^{\circ}14'08''\text{ W}$, 0 masl).

Currently, only four natural populations of *V. chilensis* represent the most significant numbers of individuals for this species. These populations are located in the Coquimbo Region of Chile (Figure 3A) and were sampled to carry out this study. The

populations sampled were Chungungo (29°25'19" S–71°17'17" W, 20 m above sea level [masl]), Huachalalume (29°59'52" S–71°09'06" W, 485 masl), Fray Jorge National Park (30°38'06" S–71°34'16" W, 600 masl), and Valle del Encanto National Historical Monument (30°42'12" S–71°22'22" W, 200 masl).

4.2. DNA Extraction and Genotyping by Sequencing (GBS)

The DNA of young leaves from 92 individuals of *V. pubescens* and 94 individuals of *V. chilensis* was isolated using a modified CTAB approach [29]. A concentration of at least 100 ng/μL of double-stranded DNA was corroborated using a Qubit 3.0 Fluorimeter (Thermo Fisher Scientific, Waltham, Massachusetts, USA).

GBS protocols were carried out in the Institute of Biotechnology, Cornell University, Ithaca, NY, USA (<http://www.biotech.cornell.edu/brc/genomics-facility>, accessed on 17 June 2022) [37]. Libraries were developed using the restriction enzyme PstI (CTGCAG) and two different adapters. Sequencing was performed using Illumina HiSeq2000 (Illumina Inc., San Diego, CA, USA). The raw files (FASTQ) were aligned to the papaya genome version ASGPBv0.4 [34] using the Burrows–Wheeler alignment tool, version 0.7.8-r441 (<http://bio-bwa.sourceforge.net/>, accessed on 22 April 2021, [90]). A ‘master’ TagCounts file was produced, which was aligned to the papaya genome v1.0, and a Tags on Physical Map (TOPM) file was built, containing the best genomic position of each tag. The barcode information in the original FASTQ files is used to tally the number of times each tag in the master tag list is observed in each sample (‘taxon’), and these counts are stored in a TagsByTaxon (TBT file) file. The information recorded in the TOPM and TBT was used to call SNPs and INDEL. The information obtained for the genotypes was stored as a filtered VCF file for further analysis with Tassel 5.2.52 (<https://www.maizegenetics.net/tassel/docs/TasselPipelineGBS.pdf>, accessed on 22 April 2021, [38,91]).

4.3. Population Genetic Analysis

SNPs/INDELS were identified following the nomenclature defined in the papaya genome v1.0 [34]. Heterozygosity and Fis were estimated using Tassel software v5.2.52, Genealex v6.5 (<https://biology-assets.anu.edu.au/GenALEX/Download.html>, accessed on 15 May 2022, [92]) and Popgene v1.31 (<https://sites.ualberta.ca/~fyeh/popgene.html>, accessed on 3 May 2021, [93]). The selfing rate (s) was calculated as $s = 2 F_{IS} / (1 + F_{IS}) \times 100$ [93].

The nucleotide diversity (π) values were estimated as the mean pairwise difference between two sequences for all possible sequence combinations [45]. The expectation of π (identified as q in this text) was explored using Tassel v5.2.52 software [91], DnaSP (<http://www.ub.edu/dnasp/>, accessed on 3 May 2021, [94]), and Mega X (<https://www.megasoftware.net/>, accessed on 3 May 2021, [95]). The confidence limit of D was obtained by assuming the beta distribution [96].

Tajima’s D neutrality test (D) was used to explore the effect of selection and genetic drift on nucleotide diversity [51,97]. The value of D [$\pi - q / \sqrt{\text{var}(\pi - q)}$] equals zero under a neutral model. Negative values indicate an excess of rare variants, recent population expansion, and purifying selection. In contrast, positive values suggest a lack of rare alleles, balancing selection, and sudden population contraction [97].

The population structure was inferred by a Bayesian model clustering algorithm using Structure v.2.3.3 (<https://web.stanford.edu/group/pritchardlab/structure.html>, accessed on 3 May 2021, [98]). Five independent runs of the algorithm were used, assuming values of K from 1 to 10. The run parameters were performed with 100,000 Markov chain Monte Carlo (MCMC) repetitions and a burn-in period of 10,000, assuming an admixture population. The optimal K value was estimated using ΔK methodology [99]. The proportion of ancestry in a given cluster was calculated as an assignment rate (Q), where 0.8 was accepted as a correct assignment to a single cluster. The ΔK s obtained by Structure Harvester (<https://taylor0.biology.ucla.edu/structureHarvester/>, accessed on 3 May 2021, [100]) and Pophelper (<http://www.royfrancis.com/pophelper/articles/index.html>, accessed on 3 May 2021, [101]) were used to estimate the rate of change in the log probability of data between successive K

values and the variance of log probabilities. Each Q value plot was developed in Structure Plot v2.0 software (<http://omicsspeaks.com/strplot2/>, accessed on 3 May 2021, [102]).

The SNPs/INDELs that enable differentiation between *V. pubescens* and *V. chilensis* and among *V. pubescens* and *V. chilensis* populations were estimated using AMOVA ϕ PT. The ϕ PT estimates the proportion of the variance among populations relative to the total variance [103]. The criteria used to select those SNPs/INDELs were ϕ PT values ≥ 0.15 (values of ϕ PT ≥ 0.15 indicate a high level of differentiation), with a significance level inferior to 0.05 (p) and missing data per locus inferior to 0.10. The ϕ PT and its significance level for each locus were estimated with 1000 permutations through Genealex v6.5.

According to the *Carica papaya* genome, the selected SNPs/INDELs were used to identify its physical positions and putative genes. The location of each SNPs/INDEL within the exon and intron was determined using a Perl script (www.perl.org (accessed on 10 September 2021), determining the gene structure limits from the *Carica papaya* genome GFF annotation, version ASGPBv0.4 [34]).

5. Conclusions

GBS technology enabled the identification of 4675 and 4451 SNPs/INDELs in two papaya species: *V. pubescens* and *V. chilensis*, respectively. The cultivated *V. pubescens* had a reduced genetic diversity at the genomic level and small differentiation between the cultivated populations, although it was possible to distinguish slight differences in the population from southern Chile (Lipimavida, Chile). The effect of a population bottleneck during domestication in Chile could explain those results. Therefore, it is necessary to introduce new germplasms from their centers of origin and initiate a breeding program to develop varieties that combine the adaptation of local genotypes cultivated for hundreds of years in Chile with genotypes from Colombia and Ecuador. On the other hand, the fragmented and isolated natural populations of *V. chilensis* had a moderate genetic diversity and low differentiation between populations, although it was possible to appreciate a particular trend associated with their geographic distribution. Its adaptations to the semi-desert conditions of northern Chile (less than 70 mm of annual rainfall) and temperatures that fluctuate between 9 °C and 20 °C make this species a unique and valuable genetic resource. Finally, it is urgent to take measures to conserve its natural populations because they are suffering based on its lower number of plants due to human activities.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/plants11162151/s1>, File S1 with SNPs and genes characterization information. Sheet 1 Total SNPs of *Vasconcellea pubescens*; Sheet 2 Total SNPs of *Vasconcellea chilensis*; Sheet 3 Nucleotide codes; Sheet 4 Information about genetic differences between *V. pubescens* and *V. chilensis* and genes; Sheet 5 Information about genetic differences between orchards of *V. pubescens*; Sheet 6 Information about genetic differences between natural population of *V. chilensis*.

Author Contributions: B.C. and H.S. conceived the research plans; B.C. and H.S.; performed most of the experiments; J.E.M. performed some of the bioinformatic analyses; B.C., H.S. and R.M. analyzed the data; B.C., M.G., Y.R.-A., J.E.M., R.G.-G., B.C.-P., R.P.-D., J.P.-M., L.A.M., B.A. and H.S. wrote the article. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Fondecyt Regular # 1150919 and Programa de Fortalecimiento de Centros Regionales R19A10001. It was sponsored by Agencia Nacional de Investigación y Desarrollo de Chile (ANID), Chile.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: B.C. thanks ANID, Fondecyt Regular # 1150919, and the Programa de Fortalecimiento de Centros Regionales # R19A10001.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Moore, P.H.; Ming, R. Papaya Genome: A Model for Tropical Fruit Trees and Beyond. *Trop. Plant Biol.* **2008**, *1*, 179–180. [[CrossRef](#)]
2. Badillo, V.M. *Monografía de la Familia Caricaceae*; Asociación de Profesores: 1971. Badillo, V. M. Monografía de la Familia Caricaceae; Universidad Central de Venezuela: Maracay, Venezuela, 1971.
3. Badillo, V.M. *Caricaceae: Segundo Esquema*; Revista de la Facultad de Agronomía (UCV), Alcance; Universidad Central de Venezuela, Facultad de Agronomía: Caracas, Venezuela, 1993; Volume 43, p. 111.
4. Badillo, V.M.; Carica, L. *Vasconcella* St. Hil. (Caricaceae): Con La Rehabilitación de Este Último. *Ernstia* **2000**, *10*, 74–79.
5. Badillo, V.M. Nota Correctiva *Vasconcella* St. Hil. y No *Vasconcella* (Caricaceae). *Ernstia* **2001**, *11*, 75–76.
6. Scheldeman, X.; Willems, L.; Coppens d’Eeckenbrugge, G.; Romeijn-Peters, E.; Restrepo, M.T.; Romero Motoche, J.; Jiménez, D.; Lobo, M.; Medina, C.I.; Reyes, C.; et al. Distribution, Diversity and Environmental Adaptation of Highland Papayas (*Vasconcella* spp.) in Tropical and Subtropical America. In *Plant Conservation and Biodiversity; Topics in Biodiversity and Conservation*; Hawksworth, D.L., Bull, A.T., Eds.; Springer: Dordrecht, The Netherlands, 2007; pp. 293–310. ISBN 978-1-4020-6444-9.
7. National Research Council. *Lost Crops of the Incas: Little-Known Plants of the Andes with Promise for Worldwide Cultivation*; The National Academies Press: Washington, DC, USA, 1999; ISBN 978-0-309-04264-2.
8. Fabiane, R.C.; Pereira, T.N.S.; Hodnett, G.L.; Pereira, M.G.; Stelly, D.M. Fluorescent in Situ Hybridization of 18S and 5S rDNA in Papaya (*Carica papaya* L.) and Wild Relatives. *Caryologia* **2008**, *61*, 411–416. [[CrossRef](#)]
9. Salvatierra-González, M.A.; Jana-Ayala, C. Floral Expression and Pollen Germination Ability in Productive Mountain Papaya (*Vasconcella pubescens* A.DC.) Orchards. *Chil. J. Agric. Res.* **2016**, *76*, 136–142. [[CrossRef](#)]
10. Chong-Pérez, B.; Carrasco, B.; Silva, H.; Herrera, F.; Quiroz, K.; Garcia-Gonzales, R. Regeneration of Highland Papaya (*Vasconcella pubescens*) from Anther Culture. *Appl. Plant Sci.* **2018**, *6*, e01182. [[CrossRef](#)] [[PubMed](#)]
11. Correa, Q.J.E.; Yesid Bernal, H. *Especies Vegetales Promisorias de Los Países Del Convenio Andrés Bello*; Secretaría Ejecutiva del Convenio Andrés Bello: Bogotá, Colombia, 1989.
12. Latham, R.E. *La Agricultura Precolombina en Chile y los Países Vecinos*; Ediciones de la Universidad de Chile: Santiago, Chile, 1936.
13. Squeo, F.; Arancio, G.; Gutierrez, J. *Libro Rojo de La Flora Nativa y de Los Sitios Prioritarios Para Su Conservación: Región de Coquimbo*; Ediciones de la Universidad de La Serena: La Serena, Chile, 2001.
14. Jordan, Z.M. In Vitro Morphogenic Responses of *Vasconcella chilensis* Planch. Ex, A. DC (Caricaceae). *Agron. Colomb.* **2011**, *29*, 481–485.
15. Ocampo, J.; Dambier, D.; Ollitrault, P.; D’eeckenbrugge, G.; Brottier, P.; Froelicher, Y.; Risterucci, A.-M. Microsatellite Markers in *Carica papaya* L.: Isolation, Characterization and Transferability to *Vasconcella* Species. *Mol. Ecol. Notes* **2006**, *6*, 212–217. [[CrossRef](#)]
16. Chávez-Pesqueira, M.; Núñez-Farfán, J. Genetic Diversity and Structure of Wild Populations of *Carica papaya* in Northern Mesoamerica Inferred by Nuclear Microsatellites and Chloroplast Markers. *Ann. Bot.* **2016**, *118*, 1293–1306. [[CrossRef](#)]
17. Chávez-Pesqueira, M.; Núñez-Farfán, J. Domestication and Genetics of Papaya: A Review. *Front. Ecol. Evol.* **2017**, *5*, 155–163. [[CrossRef](#)]
18. Chávez-Pesqueira, M.; Suárez-Montes, P.; Castillo, G.; Núñez-Farfán, J. Habitat Fragmentation Threatens Wild Populations of *Carica papaya* (Caricaceae) in a Lowland Rainforest. *Am. J. Bot.* **2014**, *101*, 1092–1101. [[CrossRef](#)] [[PubMed](#)]
19. Chen, C.; Yu, Q.; Hou, S.; Li, Y.; Eustice, M.; Skelton, R.L.; Veatch, O.; Herdes, R.E.; Diebold, L.; Saw, J.; et al. Construction of a Sequence-Tagged High-Density Genetic Map of Papaya for Comparative Structural and Evolutionary Genomics in Brassicales. *Genetics* **2007**, *177*, 2481–2491. [[CrossRef](#)] [[PubMed](#)]
20. Dai, F.; Miao, Q.; Zhou, B.; Yang, L.; Liu, Z.-L. Protective Effects of Flavonols and Their Glycosides against Free Radical-Induced Oxidative Hemolysis of Red Blood Cells. *Life Sci.* **2006**, *78*, 2488–2493. [[CrossRef](#)]
21. Sengupta, S.; Das, B.; Acharyya, P.; Prasad, M.; Ghose, T.K. Genetic Diversity Analysis in a Set of Caricaceae Accessions Using Resistance Gene Analogues. *BMC Genet.* **2014**, *15*, 137. [[CrossRef](#)]
22. Vidal, N.M.; Grazziotin, A.L.; Ramos, H.C.C.; Pereira, M.G.; Venancio, T.M. Development of a Gene-Centered SSR Atlas as a Resource for Papaya (*Carica papaya*) Marker-Assisted Selection and Population Genetic Studies. *PLoS ONE* **2014**, *9*, e112654. [[CrossRef](#)] [[PubMed](#)]
23. Blas, A.L.; Yu, Q.; Veatch, O.J.; Paull, R.E.; Moore, P.H.; Ming, R. Genetic Mapping of Quantitative Trait Loci Controlling Fruit Size and Shape in Papaya. *Mol. Breed.* **2012**, *29*, 457–466. [[CrossRef](#)]
24. Na, J.-K.; Wang, J.; Murray, J.E.; Gschwend, A.R.; Zhang, W.; Yu, Q.; Navajas-Pérez, R.; Feltus, F.A.; Chen, C.; Kubat, Z.; et al. Construction of Physical Maps for the Sex-Specific Regions of Papaya Sex Chromosomes. *BMC Genom.* **2012**, *13*, 176. [[CrossRef](#)]
25. Wai, C.M.; Moore, P.H.; Paull, R.E.; Ming, R.; Yu, Q. An Integrated Cytogenetic and Physical Map Reveals Unevenly Distributed Recombination Spots along the Papaya Sex Chromosomes. *Chromosom. Res.* **2012**, *20*, 753–767. [[CrossRef](#)]
26. Yu, Q.; Tong, E.; Skelton, R.L.; Bowers, J.E.; Jones, M.R.; Murray, J.E.; Hou, S.; Guan, P.; Acob, R.A.; Luo, M.-C.; et al. A Physical Map of the Papaya Genome with Integrated Genetic Map and Genome Sequence. *BMC Genom.* **2009**, *10*, 371. [[CrossRef](#)]
27. Kyndt, T.; Droogenbroeck, B.V.; Haegeman, A.; Roldán-Ruiz, I.; Gheysen, G. Cross-Species Microsatellite Amplification in *Vasconcella* and Related Genera and Their Use in Germplasm Classification. *Genome* **2006**, *49*, 786–798. [[CrossRef](#)]
28. Carrasco, B.; Avila, P.; Perez-Diaz, J.; Muñoz, P.; Garcia-Gonzales, R.; Lavandero, B.; Zurita-Silva, A.; Retamales, J.; Caligari, P. Genetic Structure of Highland Papayas (*Vasconcella pubescens* (Lenn, et C. Koch) Badillo) Cultivated along a Geographic Gradient in Chile as Revealed by Inter Simple Sequence Repeats (ISSR). *Genet. Resour. Crop Evol.* **2008**, *56*, 331–337. [[CrossRef](#)]

29. Carrasco, B.; García-González, R.; Díaz, C.; Ávila, P.; Cáceres, P.; Lobos, G.A.; Silva, H.; Caligari, P.D.S. Genetic and Morphological Characterization of the Endangered Austral Papaya *Vasconcellea chilensis* (Planch. Ex, A. DC.) Solms. *Genet. Resour. Crop. Evol.* **2014**, *61*, 1423–1432. [[CrossRef](#)]
30. Devey, D. *Ecological Genetics: Design, Analysis and Application*. Lowe, A., Harris S, Ashton, P. 2004. Oxford: Blackwell Publishing. £29.99 (Softback). 344 Pp. *Ann. Bot.* **2005**, *95*, 705. [[CrossRef](#)]
31. Nadeem, M.A.; Nawaz, M.A.; Shahid, M.Q.; Doğan, Y.; Comertpay, G.; Yıldız, M.; Hatipoğlu, R.; Ahmad, F.; Alsaleh, A.; Labhane, N.; et al. DNA Molecular Markers in Plant Breeding: Current Status and Recent Advancements in Genomic Selection and Genome Editing. *Biotechnol. Biotechnol. Equip.* **2018**, *32*, 261–285. [[CrossRef](#)]
32. Goodstein, D.M.; Shu, S.; Howson, R.; Neupane, R.; Hayes, R.D.; Fazo, J.; Mitros, T.; Dirks, W.; Hellsten, U.; Putnam, N.; et al. Phytozome: A Comparative Platform for Green Plant Genomics. *Nucleic Acids Res.* **2012**, *40*, D1178–D1186. [[CrossRef](#)] [[PubMed](#)]
33. Michael, T.P.; VanBuren, R. Progress, Challenges and the Future of Crop Genomes. *Curr. Opin. Plant Biol.* **2015**, *24*, 71–81. [[CrossRef](#)]
34. Ming, R.; Hou, S.; Feng, Y.; Yu, Q.; Dionne-Laporte, A.; Saw, J.H.; Senin, P.; Wang, W.; Ly, B.V.; Lewis, K.L.T.; et al. The Draft Genome of the Transgenic Tropical Fruit Tree Papaya (*Carica Papaya* Linnaeus). *Nature* **2008**, *452*, 991–996. [[CrossRef](#)]
35. Batley, J.; Edwards, D. SNP Applications in Plants. In *Association Mapping in Plants*; Oraguzie, N.C., Rikkerink, E.H.A., Gardiner, S.E., De Silva, H.N., Eds.; Springer: New York, NY, USA, 2007; pp. 95–102. ISBN 978-0-387-36011-9.
36. Sonah, H.; Bastien, M.; Iquira, E.; Tardivel, A.; Légaré, G.; Boyle, B.; Normandeau, É.; Laroche, J.; Larose, S.; Jean, M.; et al. An Improved Genotyping by Sequencing (GBS) Approach Offering Increased Versatility and Efficiency of SNP Discovery and Genotyping. *PLoS ONE* **2013**, *8*, e54603. [[CrossRef](#)]
37. Elshire, R.J.; Glaubitz, J.C.; Sun, Q.; Poland, J.A.; Kawamoto, K.; Buckler, E.S.; Mitchell, S.E. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS ONE* **2011**, *6*, e19379. [[CrossRef](#)] [[PubMed](#)]
38. Glaubitz, J.C.; Casstevens, T.M.; Lu, F.; Harriman, J.; Elshire, R.J.; Sun, Q.; Buckler, E.S. TASSEL-GBS: A High Capacity Genotyping by Sequencing Analysis Pipeline. *PLoS ONE* **2014**, *9*, e90346. [[CrossRef](#)]
39. Truong, H.T.; Ramos, A.M.; Yalcin, F.; de Ruiter, M.; van der Poel, H.J.A.; Huvenaars, K.H.J.; Hogers, R.C.J.; van Enckevort, L.J.G.; Janssen, A.; van Orsouw, N.J.; et al. Sequence-Based Genotyping for Marker Discovery and Co-Dominant Scoring in Germplasm and Populations. *PLoS ONE* **2012**, *7*, e37565. [[CrossRef](#)] [[PubMed](#)]
40. Voss-Fels, K.; Snowdon, R.J. Understanding and Utilizing Crop Genome Diversity via High-Resolution Genotyping. *Plant Biotechnol. J.* **2016**, *14*, 1086–1094. [[CrossRef](#)]
41. Poland, J.A.; Rife, T.W. Genotyping-by-Sequencing for Plant Breeding and Genetics. *Plant Genome* **2012**, *5*, 92–102. [[CrossRef](#)]
42. Romay, M.C.; Millard, M.J.; Glaubitz, J.C.; Peiffer, J.A.; Swarts, K.L.; Casstevens, T.M.; Elshire, R.J.; Acharya, C.B.; Mitchell, S.E.; Flint-Garcia, S.A.; et al. Comprehensive Genotyping of the USA National Maize Inbred Seed Bank. *Genome Biol.* **2013**, *14*, R55. [[CrossRef](#)] [[PubMed](#)]
43. Spindel, J.; Wright, M.; Chen, C.; Cobb, J.; Gage, J.; Harrington, S.; Lorieux, M.; Ahmadi, N.; McCouch, S. Bridging the Genotyping Gap: Using Genotyping by Sequencing (GBS) to Add High-Density SNP Markers and New Value to Traditional Bi-Parental Mapping and Breeding Populations. *Theor. Appl. Genet.* **2013**, *126*, 2699–2716. [[CrossRef](#)] [[PubMed](#)]
44. Jarquín, D.; Kocak, K.; Posadas, L.; Hyma, K.; Jedlicka, J.; Graef, G.; Lorenz, A. Genotyping by Sequencing for Genomic Prediction in a Soybean Breeding Population. *BMC Genom.* **2014**, *15*, 740. [[CrossRef](#)]
45. Uitdewilligen, J.G.A.M.L.; Wolters, A.-M.A.; D'hoop, B.B.; Borm, T.J.A.; Visser, R.G.F.; van Eck, H.J. A Next-Generation Sequencing Method for Genotyping-by-Sequencing of Highly Heterozygous Autotetraploid Potato. *PLoS ONE* **2013**, *8*, e62355. [[CrossRef](#)]
46. Gürcan, K.; Teber, S.; Ercisli, S.; Yilmaz, K.U. Genotyping by Sequencing (GBS) in Apricots and Genetic Diversity Assessment with GBS-Derived Single-Nucleotide Polymorphisms (SNPs). *Biochem. Genet.* **2016**, *54*, 854–885. [[CrossRef](#)]
47. Cao, K.; Li, Y.; Deng, C.H.; Gardiner, S.E.; Zhu, G.; Fang, W.; Chen, C.; Wang, X.; Wang, L. Comparative Population Genomics Identified Genomic Regions and Candidate Genes Associated with Fruit Domestication Traits in Peach. *Plant Biotechnol. J.* **2019**, *17*, 1954–1970. [[CrossRef](#)]
48. Chong, X.; Zhang, F.; Wu, Y.; Yang, X.; Zhao, N.; Wang, H.; Guan, Z.; Fang, W.; Chen, F. A SNP-Enabled Assessment of Genetic Diversity, Evolutionary Relationships and the Identification of Candidate Genes in Chrysanthemum. *Genome Biol. Evol.* **2016**, *8*, 3661–3671. [[CrossRef](#)]
49. Carrasco, B.; González, M.; Gebauer, M.; García-González, R.; Maldonado, J.; Silva, H. Construction of a Highly Saturated Linkage Map in Japanese Plum (*Prunus Salicina*, L.) Using GBS for SNP Marker Calling. *PLoS ONE* **2018**, *13*, e0208032. [[CrossRef](#)] [[PubMed](#)]
50. Koepke, T.; Schaeffer, S.; Harper, A.; Dicenta, F.; Edwards, M.; Henry, R.J.; Møller, B.L.; Meisel, L.; Oraguzie, N.; Silva, H.; et al. Comparative Genomics Analysis in Prunoideae to Identify Biologically Relevant Polymorphisms. *Plant Biotechnol. J.* **2013**, *11*, 883–893. [[CrossRef](#)] [[PubMed](#)]
51. Tajima, F. Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics* **1989**, *123*, 585–595. [[CrossRef](#)] [[PubMed](#)]
52. Jobin-Decor, M.P.; Graham, G.C.; Henry, R.J.; Drew, R.A. RAPD and Isozyme Analysis of Genetic Relationships between *Carica Papaya* and Wild Relatives. *Genet. Resour. Crop Evol.* **1997**, *44*, 471–477. [[CrossRef](#)]
53. Kanupriya, K.; Shobhana, M.; Vasugi, C.; Aswath, C.; Radhika, V.; Reddy, L.; Dinesh, M. Genetic Relationship among Papaya (*Carica Papaya*) and Wild Papaya (*Vasconcellea* Species) Using RAPD and ISSR Markers. *Indian J. Agric. Sci.* **2012**, *82*, 366–369.

54. Rodríguez, J.; Rodríguez, P.; González, M.E.; Martínez-Gómez, P. Molecular Characterization of Cuban Endemism *Carica Cubensis* Solms Using Random Amplified Polymorphic DNA (RAPD) Markers. *Agric. Sci.* **2010**, *1*, 95. [[CrossRef](#)]
55. Van Droogenbroeck, B.; Breyne, P.; Goetghebeur, P.; Romeijn-Peeters, E.; Kyndt, T.; Gheysen, G. AFLP Analysis of Genetic Relationships among Papaya and Its Wild Relatives (Caricaceae) from Ecuador. *Theor. Appl. Genet.* **2002**, *105*, 289–297. [[CrossRef](#)]
56. Brown, J.E.; Bauman, J.M.; Lawrie, J.F.; Rocha, O.J.; Moore, R.C. The Structure of Morphological and Genetic Diversity in Natural Populations of *Carica Papaya* (Caricaceae) in Costa Rica. *Biotropica* **2012**, *44*, 179–188. [[CrossRef](#)]
57. Singh, N.; Choudhury, D.R.; Singh, A.K.; Kumar, S.; Srinivasan, K.; Tyagi, R.K.; Singh, N.K.; Singh, R. Comparison of SSR and SNP Markers in Estimation of Genetic Diversity and Population Structure of Indian Rice Varieties. *PLoS ONE* **2013**, *8*, e84136. [[CrossRef](#)]
58. Arias, D.M.; Albarrán-Lara, A.L.; González-Rodríguez, A.; Peñaloza-Ramírez, J.; Dorado, O.; Leyva, E. Genetic Diversity and Structure of Wild Populations of the Tropical Dry Forest Tree *Jacaratia Mexicana* (Brassicales: Caricaceae) at a Local Scale in Mexico. *Rev. Biol. Trop.* **2012**, *60*, 1–10. [[CrossRef](#)]
59. Causse, M.; Desplat, N.; Pascual, L.; Le Paslier, M.-C.; Sauvage, C.; Bauchet, G.; Bérard, A.; Bounon, R.; Tchoumakov, M.; Brunel, D.; et al. Whole Genome Resequencing in Tomato Reveals Variation Associated with Introgression and Breeding Events. *BMC Genom.* **2013**, *14*, 791. [[CrossRef](#)] [[PubMed](#)]
60. Foster, J.T.; Allan, G.J.; Chan, A.P.; Rabinowicz, P.D.; Ravel, J.; Jackson, P.J.; Keim, P. Single Nucleotide Polymorphisms for Assessing Genetic Diversity in Castor Bean (*Ricinus Communis*). *BMC Plant Biol.* **2010**, *10*, 13. [[CrossRef](#)] [[PubMed](#)]
61. Abdullah, N.; Bakar, N.A.; Bakar, U.K.A.; Abidin, R.A.Z.; Yusof, M.F. Sequence Information on Single Nucleotide Polymorphism (SNP) through Genome Sequencing Analysis of *Carica Papaya* Variety Eksotika and Sekaki. *J. Trop. Agric. Food Sci.* **2016**, *44*, 219–228.
62. Ueno, H.; Urasaki, N.; Natsume, S.; Yoshida, K.; Tarora, K.; Shudo, A.; Terauchi, R.; Matsumura, H. Genome Sequence Comparison Reveals a Candidate Gene Involved in Male-Hermaphrodite Differentiation in Papaya (*Carica Papaya*) Trees. *Mol. Genet. Genom.* **2015**, *290*, 661–670. [[CrossRef](#)] [[PubMed](#)]
63. VanBuren, R.; Wai, C.M.; Zhang, J.; Han, J.; Arro, J.; Lin, Z.; Liao, Z.; Yu, Q.; Wang, M.-L.; Zee, F.; et al. Extremely Low Nucleotide Diversity in the X-Linked Region of Papaya Caused by a Strong Selective Sweep. *Genome Biol.* **2016**, *17*, 230. [[CrossRef](#)]
64. Lee, C.-Y.; Lin, H.-J.; Viswanath, K.K.; Lin, C.-P.; Chang, B.C.-H.; Chiu, P.-H.; Chiu, C.-T.; Wang, R.-H.; Chin, S.-W.; Chen, F.-C. The Development of Functional Mapping by Three Sex-Related Loci on the Third Whorl of Different Sex Types of *Carica Papaya* L. *PLoS ONE* **2018**, *13*, e0194605. [[CrossRef](#)]
65. Nybom, H. Comparison of Different Nuclear DNA Markers for Estimating Intraspecific Genetic Diversity in Plants. *Mol. Ecol.* **2004**, *13*, 1143–1155. [[CrossRef](#)]
66. Aranzana, M.J.; Illa, E.; Howad, W.; Arús, P. A First Insight into Peach [*Prunus Persica* (L.) Batsch] SNP Variability. *Tree Genet. Genomes* **2012**, *8*, 1359–1369. [[CrossRef](#)]
67. Emanuelli, F.; Lorenzi, S.; Grzeskowiak, L.; Catalano, V.; Stefanini, M.; Troggio, M.; Myles, S.; Martinez-Zapater, J.M.; Zyprian, E.; Moreira, F.M.; et al. Genetic Diversity and Population Structure Assessed by SSR and SNP Markers in a Large Germplasm Collection of Grape. *BMC Plant Biol.* **2013**, *13*, 39. [[CrossRef](#)]
68. Micheletti, D.; Troggio, M.; Zharkikh, A.; Costa, F.; Malnoy, M.; Velasco, R.; Salvi, S. Genetic Diversity of the Genus *Malus* and Implications for Linkage Mapping with SNPs. *Tree Genet. Genomes* **2011**, *7*, 857–868. [[CrossRef](#)]
69. Silva, J.a.T.; Rashid, Z.; Nhut, D.T.; Sivakumar, D.; Gera, A.; Souza, M.T.; Tennant, P. Papaya (*Carica Papaya* L.) Biology and Biotechnology. *Tree For. Sci. Biotechnol.* **2007**, *1*, 47–73.
70. Sousa, E.; Kost, B.; Malhó, R. Arabidopsis Phosphatidylinositol-4-Monophosphate 5-Kinase 4 Regulates Pollen Tube Growth and Polarity by Modulating Membrane Recycling. *Plant Cell* **2008**, *20*, 3050–3064. [[CrossRef](#)] [[PubMed](#)]
71. Zhang, X.; Yang, H.; Yu, J.; Chen, C.; Zhang, G.; Bao, J.; Du, Y.; Kibukawa, M.; Li, Z.; Wang, J.; et al. Genomic Organization, Transcript Variants and Comparative Analysis of the Human Nucleoporin 155 (NUP155) Gene. *Gene* **2002**, *288*, 9–18. [[CrossRef](#)]
72. Yuan, Y.; Qi, L.; Yu, J.; Wang, X.; Huang, L. Transcriptome-Wide Analysis of SAME Superfamily to Novelty Phosphoethanolamine N-Methyltransferase Copy in *Lonicera Japonica*. *Int. J. Mol. Sci.* **2014**, *16*, 521–534. [[CrossRef](#)] [[PubMed](#)]
73. Kobayashi, W.; Liu, E.; Ishii, H.; Matsunaga, S.; Schlögelhofer, P.; Kurumizaka, H. Homologous Pairing Activities of Arabidopsis *Thaliana* RAD51 and DMC1. *J. Biochem.* **2019**, *165*, 289–295. [[CrossRef](#)] [[PubMed](#)]
74. Pradillo, M.; López, E.; Linacero, R.; Romero, C.; Cuñado, N.; Sánchez-Morán, E.; Santos, J.L. Together Yes, but Not Coupled: New Insights into the Roles of RAD51 and DMC1 in Plant Meiotic Recombination. *Plant J.* **2012**, *69*, 921–933. [[CrossRef](#)] [[PubMed](#)]
75. Ashraf, M.A.; Rahman, A. Cold Stress Response in Arabidopsis *Thaliana* Is Mediated by GNOM ARF-GEF. *Plant J.* **2019**, *97*, 500–516. [[CrossRef](#)] [[PubMed](#)]
76. Kopczak, S.D.; Haas, N.A.; Hussey, P.J.; Silflow, C.D.; Snustad, D.P. The Small Genome of Arabidopsis Contains at Least Six Expressed Alpha-Tubulin Genes. *Plant Cell* **1992**, *4*, 539–547. [[CrossRef](#)]
77. Krupková, E.; Immerzeel, P.; Pauly, M.; Schmülling, T. The TUMOROUS SHOOT DEVELOPMENT2 Gene of Arabidopsis Encoding a Putative Methyltransferase Is Required for Cell Adhesion and Co-Ordinated Plant Development. *Plant J.* **2007**, *50*, 735–750. [[CrossRef](#)]
78. Sun, K.; Wolters, A.-M.A.; Loonen, A.E.H.M.; Huibers, R.P.; van der Vlugt, R.; Goverse, A.; Jacobsen, E.; Visser, R.G.F.; Bai, Y. Down-Regulation of Arabidopsis DND1 Orthologs in Potato and Tomato Leads to Broad-Spectrum Resistance to Late Blight and Powdery Mildew. *Transgenic Res.* **2016**, *25*, 123–138. [[CrossRef](#)]

79. Katano, K.; Kataoka, R.; Fujii, M.; Suzuki, N. Differences between Seedlings and Flowers in Anti-ROS Based Heat Responses of Arabidopsis Plants Deficient in Cyclic Nucleotide Gated Channel 2. *Plant Physiol. Biochem.* **2018**, *123*, 288–296. [[CrossRef](#)] [[PubMed](#)]
80. Jarsch, I.K.; Ott, T. Perspectives on Remorin Proteins, Membrane Rafts, and Their Role during Plant-Microbe Interactions. *Mol. Plant Microbe Interact.* **2011**, *24*, 7–12. [[CrossRef](#)] [[PubMed](#)]
81. Raffaele, S.; Bayer, E.; Mongrand, S. Upregulation of the Plant Protein Remorin Correlates with Dehiscence and Cell Maturation. *Plant Signal. Behav.* **2009**, *4*, 915–919. [[CrossRef](#)] [[PubMed](#)]
82. Raffaele, S.; Bayer, E.; Lafarge, D.; Cluzet, S.; German Retana, S.; Boubekeur, T.; Leborgne-Castel, N.; Carde, J.-P.; Lherminier, J.; Noirot, E.; et al. Remorin, a Solanaceae Protein Resident in Membrane Rafts and Plasmodesmata, Impairs Potato Virus X Movement. *Plant Cell* **2009**, *21*, 1541–1555. [[CrossRef](#)]
83. Costa, A.; Luoni, L.; Marrano, C.A.; Hashimoto, K.; Köster, P.; Giacometti, S.; De Michelis, M.I.; Kudla, J.; Bonza, M.C. Ca²⁺-Dependent Phosphoregulation of the Plasma Membrane Ca²⁺-ATPase ACA8 Modulates Stimulus-Induced Calcium Signatures. *J. Exp. Bot.* **2017**, *68*, 3215–3230. [[CrossRef](#)] [[PubMed](#)]
84. Bakshi, A.; Wilson, R.L.; Lacey, R.F.; Kim, H.; Wuppalapati, S.K.; Binder, B.M. Identification of Regions in the Receiver Domain of the ETHYLENE RESPONSE1 Ethylene Receptor of Arabidopsis Important for Functional Divergence. *Plant Physiol.* **2015**, *169*, 219–232. [[CrossRef](#)]
85. Carretero-Paulet, L.; Cairó, A.; Talavera, D.; Saura, A.; Imperial, S.; Rodríguez-Concepción, M.; Campos, N.; Boronat, A. Functional and Evolutionary Analysis of DXL1, a Non-Essential Gene Encoding a 1-Deoxy-D-Xylulose 5-Phosphate Synthase like Protein in Arabidopsis Thaliana. *Gene* **2013**, *524*, 40–53. [[CrossRef](#)]
86. Floss, D.S.; Hause, B.; Lange, P.R.; Küster, H.; Strack, D.; Walter, M.H. Knock-down of the MEP Pathway Isogene 1-Deoxy-D-Xylulose 5-Phosphate Synthase 2 Inhibits Formation of Arbuscular Mycorrhiza-Induced Apocarotenoids, and Abolishes Normal Expression of Mycorrhiza-Specific Plant Marker Genes. *Plant J.* **2008**, *56*, 86–100. [[CrossRef](#)]
87. Kazenwadel, C.; Klebensberger, J.; Richter, S.; Pfannstiel, J.; Gerken, U.; Pickel, B.; Schaller, A.; Hauer, B. Optimized Expression of the Dirigent Protein AtDIR6 in Pichia Pastoris and Impact of Glycosylation on Protein Structure and Function. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 7215–7227. [[CrossRef](#)]
88. Riera, M.; Figueras, M.; Lopez, C.; Goday, A.; Pages, M. Protein Kinase CK2 Modulates Developmental Functions of the Abscisic Acid Responsive Protein Rab17 from Maize. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 9879–9884. [[CrossRef](#)]
89. Chen, D.; Molitor, A.M.; Xu, L.; Shen, W.-H. Arabidopsis PRC1 Core Component ATRING1 Regulates Stem Cell-Determining Carpel Development Mainly through Repression of Class I KNOX Genes. *BMC Biol.* **2016**, *14*, 112. [[CrossRef](#)] [[PubMed](#)]
90. Li, H.; Durbin, R. Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)] [[PubMed](#)]
91. Bradbury, P.J.; Zhang, Z.; Kroon, D.E.; Casstevens, T.M.; Ramdoss, Y.; Buckler, E.S. TASSEL: Software for Association Mapping of Complex Traits in Diverse Samples. *Bioinformatics* **2007**, *23*, 2633–2635. [[CrossRef](#)]
92. Peakall, R.; Smouse, P.E. GenAEx 6.5: Genetic Analysis in Excel. Population Genetic Software for Teaching and Research—An Update. *Bioinformatics* **2012**, *28*, 2537–2539. [[CrossRef](#)]
93. Yeh, F.; Yang, R.; Boyle, T. *Microsoft Windows-Based Freeware for Populations Genetic Analysis*; POPGENE Version 1.32, Undefined; University of Alberta: Edmonton, AB, Canada, 1999.
94. Rozas, J.; Ferrer-Mata, A.; Sánchez-DelBarrio, J.C.; Guirao-Rico, S.; Librado, P.; Ramos-Onsins, S.E.; Sánchez-Gracia, A. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Mol. Biol. Evol.* **2017**, *34*, 3299–3302. [[CrossRef](#)] [[PubMed](#)]
95. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)]
96. Hull-Sanders, H.M.; Eubanks, M.D.; Carr, D.E. Inbreeding Depression and Selfing Rate of Ipomoea Hederacea Var. Integriuscula (Convolvulaceae). *Am. J. Bot.* **2005**, *92*, 1871–1877. [[CrossRef](#)] [[PubMed](#)]
97. Nei, M.; Li, W.H. Mathematical Model for Studying Genetic Variation in Terms of Restriction Endonucleases. *Proc. Natl. Acad. Sci. USA* **1979**, *76*, 5269–5273. [[CrossRef](#)]
98. Pritchard, J.K.; Stephens, M.; Donnelly, P.J. Inference of population structure using multilocus genotype data. *Genetics* **2000**, *155*, 945–959. [[CrossRef](#)]
99. Evanno, G.; Regnaut, S.; Goudet, J. Detecting the Number of Clusters of Individuals Using the Software STRUCTURE: A Simulation Study. *Mol. Ecol.* **2005**, *14*, 2611–2620. [[CrossRef](#)]
100. Earl, D.A.; vonHoldt, B.M. STRUCTURE HARVESTER: A Website and Program for Visualizing STRUCTURE Output and Implementing the Evanno Method. *Conserv. Genet. Resour.* **2012**, *4*, 359–361. [[CrossRef](#)]
101. Francis, R. Pophelper: An R Package and Web App to Analyse and Visualize Population Structure. *Mol. Ecol. Resour.* **2016**, *17*, 27–32. [[CrossRef](#)] [[PubMed](#)]
102. Ramasamy, R.K.; Ramasamy, S.; Bindroo, B.B.; Naik, V.G. STRUCTURE PLOT: A Program for Drawing Elegant STRUCTURE Bar Plots in User Friendly Interface. *Springerplus* **2014**, *3*, 431. [[CrossRef](#)] [[PubMed](#)]
103. Excoffier, L.; Smouse, P.E.; Quattro, J.M. Analysis of Molecular Variance Inferred from Metric Distances among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. *Genetics* **1992**, *131*, 479–491. [[CrossRef](#)] [[PubMed](#)]