



**HAL**  
open science

# Genetic variation of loci potentially under selection confounds species-genetic diversity correlations in a fragmented habitat

Angeline Bertin, Nicolas Gouin, Alex Baumel, Ernesto Gianoli, Juan Serratos, Rodomiro Osorio, Stephanie Manel

## ► To cite this version:

Angeline Bertin, Nicolas Gouin, Alex Baumel, Ernesto Gianoli, Juan Serratos, et al.. Genetic variation of loci potentially under selection confounds species-genetic diversity correlations in a fragmented habitat. *Molecular Ecology*, 2017, 26 (2), pp.431 - 443. 10.1111/mec.13923 . hal-01605033

**HAL Id: hal-01605033**

**<https://hal.science/hal-01605033>**

Submitted on 25 Apr 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Genetic variation of loci potentially under selection confounds species-genetic diversity correlations in a fragmented habitat

ANGELINE BERTIN<sup>1</sup>, NICOLAS GOUIN<sup>1,2,3</sup>, ALEX BAUMEL<sup>4</sup>, ERNESTO GIANOLI<sup>1,5</sup>, JUAN SERRATOSA<sup>6</sup>, RODOMIRO OSORIO<sup>1</sup> and STEPHANIE MANEL<sup>7</sup>

<sup>1</sup>Departamento de Biología, Universidad de La Serena, La Serena, Chile

<sup>2</sup>Centro de Estudios Avanzados en Zonas Áridas, La Serena, Chile

<sup>3</sup>Instituto de Investigación Multidisciplinar en Ciencia y Tecnología, Universidad de La Serena, La Serena, Chile

<sup>4</sup>Institut Méditerranéen de Biodiversité et d'Ecologie marine et continentale (IMBE) Aix Marseille Univ, Univ Avignon, CNRS, IRD, Marseille, France.

<sup>5</sup>Departamento de Botánica, Universidad de Concepción, Concepción, Chile

<sup>6</sup>Facultad de Ciencias del Mar, Doctorado en Biología y Ecología Aplicada, Universidad Católica del Norte

<sup>7</sup>EPHE, PSL Research University, CNRS, UM, SupAgro, IRD, INRA, UMR 5175 CEFV, F-34293 Montpellier, FRANCE

*Keywords:* community genetics, species-genetic diversity correlations, outlier loci, habitat connectivity, high Andean wetlands

*Corresponding Author:*

Angéline Bertin

Departamento de Biología, Facultad de Ciencias, Universidad de La Serena, Av. Raúl Bitrán 1305, Casilla 554, La Serena, Chile.

Fax : (56) (51) 2 204383, electronic mail address: abertin@userena.cl;

*Running Title:* Influence of outlier loci on SGDCs

## 1 **Abstract**

2 Positive species-genetic diversity correlations (SGDCs) are often thought to result from the  
3 parallel influence of neutral processes on genetic and species diversity. Yet, confounding effects  
4 of non-neutral mechanisms have not been explored. Here, we investigate the impact of non-  
5 neutral genetic diversity on SGDCs in high Andean wetlands. We compare correlations between  
6 plant species diversity (SD) and genetic diversity (GD) calculated with and without loci  
7 potentially under selection (outlier loci). The study system includes 2188 specimens from five  
8 species (three common aquatic macroinvertebrate and two dominant plant species) that were  
9 genotyped for 396 amplified fragment length polymorphism loci. We also appraise the  
10 importance of neutral processes on SGDCs by investigating the influence of habitat  
11 fragmentation features. Significant positive SGDCs were detected for all five species (mean  
12  $SGDC = 0.52 \pm 0.05$ ). While only a few outlier loci were detected in each species, they resulted  
13 in significant decreases in GD and in SGDCs. This supports the hypothesis that neutral processes  
14 drive species-genetic diversity relationships in high Andean wetlands. Unexpectedly, the effects  
15 on genetic diversity of the habitat fragmentation characteristics in this study increased with the  
16 presence of outlier loci in two species. Overall, our results reveal pitfalls in using habitat features  
17 to infer processes driving SGDCs and show that a few loci potentially under selection are enough  
18 to cause a significant downward bias in SGDC. Investigating confounding effects of outlier loci  
19 thus represents a useful approach to evidence the contribution of neutral processes on species-  
20 genetic diversity relationships.

## 21 **Introduction**

22 Investigating spatial or temporal patterns of covariation between species diversity within a  
23 community (SD) and genetic diversity within a species (GD) is one of the main objectives  
24 of community genetics (Antonovics 1992). It is a key step towards a conceptual synthesis at  
25 the interface between ecology and evolutionary biology (Vellend & Geber 2005). It is also  
26 a core issue in conservation biology as it provides valuable information to understand and  
27 predict different facets of biodiversity (Hu *et al.* 2009; Lavergne *et al.* 2010; Mouquet *et al.*  
28 2012).

29         Studies addressing species genetic diversity correlations (SGDCs) in the wild have  
30 shown that the relationship between GD and SD can take different forms depending on the  
31 mechanisms at work (Kahilainen *et al.* 2014), with reports of positive (He *et al.* 2008;  
32 Papadopoulou *et al.* 2011; Lamy *et al.* 2013), negative (Silvertown *et al.* 2009; Xu *et al.*  
33 2016) and no correlation (Taberlet *et al.* 2012) between the two components. Concordant  
34 spatial patterns between species and genetic diversity can emerge from various processes.  
35 They may result from direct causal relationships between the two biodiversity components  
36 (Vellend and Geber 2005). For instance, GD can influence SD when genetic variation of  
37 habitat-forming species, such as forest trees, structures dependent communities (i.e. micro-  
38 organisms, fungi, arthropods and vertebrates, Whitham *et al.* 2006). Alternatively, SD in  
39 the community can impose variable selective pressures on populations, and therefore alter  
40 the levels of population GD (Strauss *et al.* 2005; Vellend and Geber 2005; Johnson and  
41 Stinchcombe 2007; Lankau and Strauss 2007). SD and GD can become indirectly  
42 associated as well, by responding similarly to common mechanisms (Vellend and Geber  
43 2005). Indeed, both ecological and genetic models posit similar effects of drift and

44 migration on species and genetic diversity (e.g. depletion effects of drift and diversity  
45 increases from immigration, Hu *et al.* 2009). Thus, concordant spatial patterns between  
46 species and genetic diversity can be common under conditions of demographic stochasticity  
47 and dispersal limitation (Rosenzweig 1995; Frankham 1996; Frankham 1997). To date,  
48 little research has been conducted into the drivers of negative SGDCs (Xu *et al.* 2016).  
49 According to a modeling study, negative SGDCs can occur under particular conditions of  
50 high genetic mutation rates, depending on the relative importance of mutation, immigration  
51 and local competition (Laroche *et al.* 2015). They may also occur in situations where  
52 increases in SD, promoted by environmental heterogeneity, are associated with a reduction  
53 in niche breadth and/or population size (Taberlet *et al.* 2012; Vellend *et al.* 2014; Xu *et al.*  
54 2016).

55         While both neutral and selective mechanisms can generate positive covariation  
56 between genetic and species diversity (Vellend and Geber 2005), evidence suggests that  
57 neutral processes play a dominant role of in this respect (Odat *et al.* 2004; Vellend 2004;  
58 Struebig *et al.* 2011; Papadopoulou *et al.* 2011; Lamy *et al.* 2013; Vellend *et al.* 2014).  
59 Positive SGDCs are common in discrete habitat units like oceanic islands, lakes, or forest  
60 fragments, and less frequent in arbitrarily defined spatial units such as grid cells or  
61 vegetation plots (Vellend and Geber 2005; Whitlock 2014; Vellend *et al.* 2014). In  
62 fragmented landscapes, migration and drift are likely major contributors to within-patch  
63 diversity. Thus, positive SGDCs among islands and habitat patches have been postulated to  
64 arise from the parallel influence of neutral processes on both species and genetic diversity  
65 (Vellend and Geber 2005; Whitlock 2014; Vellend *et al.* 2014). This is supported by a  
66 number of empirical studies, which have demonstrated that local habitat characteristics  
67 influencing rates of stochastic immigration and extinction contribute markedly to species-

68 genetic diversity relationships in fragmented habitats (Vellend 2005; He *et al.* 2008;  
69 Papadopoulou *et al.* 2011; Lamy *et al.* 2013). These habitat features include habitat size,  
70 connectivity and perturbation. Indeed, both habitat carrying capacity and heterogeneity are  
71 expected to increase with habitat size. Larger patches can thus support larger populations,  
72 which are both more genetically diverse and more resilient to genetic drift (Frankham  
73 1997). In addition, by offering more diverse ecological niches, larger habitats are also  
74 expected to hold more diverse communities, and are consequently less susceptible to  
75 random species loss through ecological drift (MacArthur and Wilson 1967, Hubbell 2001).  
76 Because habitat connectivity influences rates of migration, it is expected to result in  
77 increases in both genetic and species diversity, by facilitating the introduction of new  
78 alleles in a population or new species in the community (Hu *et al.* 2009). Regarding habitat  
79 perturbation, it can cause parallel effects on species and genetic diversity by provoking  
80 extinction of species and alleles irrespectively of their identities (Vellend 2003).

81       Neutral explanations for positive SGDCs have also come from the wide use of  
82 presumed neutral molecular markers to estimate genetic diversity in SGDC investigations  
83 (Vellend and Geber 2005; Vellend *et al.* 2014). Genetic variation of molecular markers  
84 with supposedly no effects on fitness, such as microsatellite or amplified fragment length  
85 polymorphism (AFLP) loci, is considered to mirror neutral processes, including mutation,  
86 genetic drift and gene flow (Holderegger *et al.* 2006; Kirk and Freeland 2011). SGDCs  
87 estimated with these markers have therefore been assumed to emerge from concordant  
88 changes in species and genetic diversity induced by neutral processes (Vellend and Geber  
89 2005; Lamy *et al.* 2013; Vellend *et al.* 2014). However, no attention has been paid to  
90 confounding effects of non-neutral mechanisms, which may lead to erroneous conclusions  
91 regarding the mechanisms causing SD-GD covariation (Fig. 1). Indeed, selection generates

92 changes in allele frequencies at target genes, but also in adjacent, linked regions (genetic  
93 hitchhiking; Barton 2000). Such linked DNA regions, even non-coding DNA, can thus bear  
94 signatures of selection (confounding effect 1 in Fig. 1; Barton 2000; Oleksyk *et al.* 2010).  
95 When used to infer neutral processes, such loci may bias population parameter estimates  
96 (Luikart *et al.* 2003). Their expected effects on genetic diversity depend on the type of  
97 selection operating, as each selection mode leads to specific patterns of allele frequency  
98 changes (Oleksyk *et al.* 2010). For instance, divergent directional selection decreases  
99 genetic variation within populations, while balancing selection maintains variation within  
100 populations. A few loci potentially under divergent selection are enough to cause a  
101 significant downward bias in the assessment of GD estimated from markers assumed to be  
102 neutral (García-Verdugo *et al.* 2015). The presence of such markers can thus be misleading  
103 for SGDC studies. When SGDCs are driven by neutral mechanisms, loci under selection  
104 are likely to decrease the correlation between species and genetic diversity, since only  
105 neutral GD is expected to covary with SD in such circumstances. In contrast, no such  
106 effects are expected when SGDCs arise from non-neutral processes.

107         Separating the neutral and non-neutral components of GD not only opens new  
108 perspectives for the study of SGDCs, but can also be critical to interpreting apparent effects  
109 of local habitat features in a neutral framework. Several studies have investigated how  
110 disturbance regime, habitat size and/or connectivity relate to SGDCs in order to infer  
111 underlying mechanisms (Vellend and Geber 2005; Papadopoulou *et al.* 2011; Lamy *et al.*  
112 2013; Vellend *et al.* 2014). These studies postulate that these local features become  
113 associated with neutral components of SD/GD due to their modulating role on rates of  
114 migration and stochastic loss of species/alleles. However, these relationships can be  
115 mistakenly attributed to neutral processes (confounding effect 2 in Fig. 1) if disturbance

116 regime, habitat size and/or connectivity and selective environmental factors are spatially  
117 correlated, as they would lead to indirect associations between non-neutral components of  
118 GD and the local habitat features. These confounding effects have been documented for  
119 species richness-island area associations (Ricklefs and Lovette 1999), and are expected at  
120 the genetic level since populations at the margins of a species distribution range are not  
121 only more fragmented and isolated than central populations, but also subject to different  
122 ecological constraints (central-marginal hypothesis, Eckert *et al.* 2008). In such situations,  
123 factoring out the effects of disturbance regime, habitat size and/or connectivity bears the  
124 risk of overestimating the contribution of neutral processes on SD and GD, as it is likely to  
125 control for some influence of selective factors as well.

126         In this study, we aimed to disentangle the confounding effects of non-neutral  
127 genetic diversity on species-genetic diversity relationships in a highly fragmented  
128 ecosystem, the high Andean wetlands of Chile's Norte Chico. These wetlands, formed by  
129 groundwater upwelling, resemble oases in an arid matrix (Fig. 2, Squeo *et al.* 2006). They  
130 occupy a latitudinal gradient of both aridity (increasing northwards) and wetland density  
131 (Fig. 2). This suggests that wetlands differing in connectivity also experience different  
132 selective pressures. Given the high degree of fragmentation of these ecosystems, we  
133 hypothesized that neutral processes should play a critical role in species-genetic diversity  
134 relationships. Thus, we expected loci influenced by selection to decrease SGDCs. We also  
135 postulated that such loci can complicate the evaluation of the importance of neutral  
136 processes on SGDC by factoring out the effects of habitat connectivity on SD and GD.  
137 Indeed, due to the spatial covariation between habitat fragmentation and the latitudinal  
138 environment gradient, GD of loci under environmental gradient pressure is also expected to  
139 correlate with fragmentation characteristics. To test these conjectures, we investigated



140 SGDCs between species diversity of high Andean wetland vegetation and GDs of five  
141 species (two plants and three macroinvertebrates) with contrasting life-histories. We used a  
142 genome scan approach to screen for AFLP markers potentially under selection (i.e. outlier  
143 loci; Storz *et al.* 2004) and analyzed SGDCs and the effects of habitat characteristics (size,  
144 connectivity and stability) on SGDCs with and without these loci. Our results showed that,  
145 as expected, confounding effects of outlier loci on GD decrease SGDCs. This trend was  
146 particularly evident with loci potentially under divergent selection. We also found that, in  
147 some species, GD at outlier loci correlate with habitat features supposedly linked to neutral  
148 processes, revealing pitfalls in using such characteristics to infer processes driving SGDCs.  
149

## 150 **Material and methods**

### 151 **Study system**

152 We analyzed species and genetic diversity for 21 high Andean wetlands located along a ca.  
153 600 km latitudinal range (26°S–32°S) in north-central Chile (Fig. 2, and see Bertin *et al.*  
154 2015). The region, known as Norte Chico, is a biodiversity hotspot, and is characterized by  
155 remarkably high levels of endemism (Arroyo *et al.* 2004). It includes five hydrologically  
156 unconnected river basins. The climate of the region varies from hyperarid in the north to  
157 mediterranean in the south, with mean annual precipitation ranging between 35 mm and  
158 200 mm for the northernmost and southernmost valleys, respectively.

159

### 160 **Data collection**

#### 161 ***Community and population sampling***

162 Sampling of plant and benthic macroinvertebrate specimens was carried out between March  
163 and April 2011. For plant diversity assessment, the length of each wetland was divided into  
164 five sectors and a 30 x 30 cm quadrat was randomly placed within each sector. Plant  
165 species were separated and identified in the laboratory.

166 For genetic diversity evaluation, we selected five common species: two dominant  
167 plants, *Carex gayana* (Cyperaceae) and *Patosia clandestina* (Juncaceae), and three  
168 abundant aquatic macroinvertebrates, *Andesiops peruvianus* (Insecta, Ephemeroptera),  
169 *Austrelmis sp.* (Insecta, Coleoptera) and *Hyallela fossamancinii* (Crustacea, Amphipoda).  
170 Samples were collected by hand or using a small fishing net in the case of aquatic  
171 organisms. The entire wetland surface area was sampled in each case, with individual  
172 wetlands ranging in size from 3.7 to 38 ha. Leaf samples were conserved individually in  
173 silica gel until DNA extraction, whereas macroinvertebrates were stored in 95% alcohol.  
174 The plant *C. gayana* was the only species found to occur within all 21 wetlands sampled.  
175 Details regarding sampling sites and size for each species are given in Table S1.

176

### 177 **Genetic data**

178 For all five species, DNA extraction was performed using a CTAB protocol. AFLP  
179 amplifications and genotyping were carried out using standard procedures and following  
180 the protocol of Meudt and Clarke (2007), available at [http://clarkeresearch.org/aflp\\_2012-](http://clarkeresearch.org/aflp_2012-01-26/aflp.html)  
181 [01-26/aflp.html](http://clarkeresearch.org/aflp_2012-01-26/aflp.html), with very few modifications. In each species, four combinations of  
182 specific primers were selected based on their reliability and number of fragments amplified.  
183 AFLP fragments were separated by capillary electrophoresis and genotyping was  
184 performed with the software GeneMarker v2.4 (Softgenetics). Following Bonin *et al.*  
185 (2007), we retained only AFLP loci with genotyping error rate < 10%, with the exception

186 of the crustacean *H. fossamancinii*, for which we included loci with up to 12% genotyping  
187 error in order to maintain a comparable number of loci for all species. Loci with redundant  
188 information were filtered using the software AFLPOP (Duchesne and Bernatchez 2002).  
189 Detailed procedures and AFLP primer pairs are described in Supplementary Information.  
190 Only polymorphic markers (those with band frequencies between 5% and 95%) were kept  
191 in the analysis (Bonin *et al.* 2007).

192

### 193 **Habitat parameters**

194 Habitat attributes associated with connectivity, size and stability were assessed for each of  
195 the 21 sampled wetlands. Attributes were derived from maps of wetland distribution  
196 generated based on spectral analysis of Landsat 8 OLI satellite imagery  
197 (<http://glovis.usgs.gov/>, see Fig. 2) and elevation data. Normalized Difference Vegetation  
198 Index, a measure of actively photosynthesizing vegetation (NDVI, Rouse *et al.* 1973), was  
199 first calculated using Landsat data (January 2011) for a training dataset of 87 high Andean  
200 wetlands mapped on-site in the Norte Chico region by Chile's Ministry of Environment and  
201 Agricultural and Livestock Service. Based on the mean NDVI value of all pixels  
202 comprising the 87 field-mapped wetlands, and applying the minimum observed elevation  
203 value as a cut-off, we defined wetlands here as all vegetated pixels with NDVI values > 0.2,  
204 located above 1800 m a.s.l. To exclude sparse vegetation from the mapping, only vegetated  
205 areas  $\geq 0.36$  ha were considered.

206 We calculated various local structural connectivity metrics for each wetland using  
207 Graphab 1.0 (Foltete *et al.* 2012), including node degree (Dg), clustering coefficient (CC),  
208 closeness centrality (CCe), eccentricity (Ec), betweenness centrality (BC) and wetland  
209 density. These metrics were quantified by considering all the wetlands mapped within a 20

210 km radius from the focal site. In addition, we estimated the surface of each focal wetland  
211 from the GIS-based map (Fig. 2). As a proxy for wetland stability, we used the temporal  
212 range of the mean NDVI of each site calculated over 20 years (from 1991 to 2002 and from  
213 2004 to 2011). Before evaluating the impact of the habitat metrics on diversity, we removed  
214 strongly correlated predictors by excluding one variable when a Pearson's correlation  
215 between a pair of habitat metrics was  $> 0.7$ . The final set of predictors included wetland  
216 size, Dg, CCe, Ec, BC and the temporal range of NDVI wetland values.

217

## 218 **Statistical analysis**

### 219 **Identification of outlier loci**

220 We used the Bayesian likelihood  $F_{ST}$ -based method implemented in the software BayeScan  
221 (Foll and Gaggiotti 2008) to identify outlier loci. This approach estimates for each locus the  
222 posterior probabilities of two alternative models: a neutral one and a model including  
223 selection. We performed the analyses using the default parameters and considered a  
224 conservative value of 10 for the prior odds (10:1 odds in favor of the neutral model). Loci  
225 with  $F_{ST}$  values substantially higher than the observed norm are considered potentially  
226 under diversifying selection, while those with substantially lower  $F_{ST}$  values are considered  
227 potentially under balancing or purifying selection (Foll and Gaggiotti 2008). These outlier  
228 loci were identified with the R function `plot_bayescan` as described in the user's manual,  
229 using a false discovery rate (FDR, the q-value threshold) of 5%. Accordingly, we examined  
230 the effects of outlier loci by creating three marker datasets and considering their  $F_{ST}$  values:  
231 a dataset free of outliers (DS1), one with all the genotyped loci, thus including both the

232 non-outlier and all the outlier loci (i.e. with high and low  $F_{ST}$ , DS2) and one dataset  
233 including the non-outlier loci and the outlier loci with high  $F_{ST}$  only (DS3).

234

235 **Species richness, genetic diversity and the influence of outlier loci on genetic diversity**  
236 **estimates**

237 We used the statistical software R (R Core Team 2015) to calculate diversity indices and  
238 perform SGDC analyses. We estimated species diversity within each wetland as species  
239 richness of all the collected quadrats, and within-population genetic diversity as Nei's gene  
240 diversity (Nei 1987). We calculated the bootstrapped 95% confidence interval of species  
241 richness and Nei's indices by considering 10,000 bootstrap replications using the R  
242 package rich (Rossi 2011) and R functions of AFLPdat (Ehrich 2006), respectively. Nei's  
243 gene indices were computed for the three marker datasets separately (DS1, DS2 and DS3).  
244 We tested the effects of outlier loci on GD by carrying out linear mixed models (LMM) of  
245 Nei's gene diversity (dependent variable) in relation to the presence of outlier loci (fixed  
246 effect), and including species and site as crossed random factors. We performed two such  
247 analyses, testing the effects of all the outlier loci (with datasets DS1 and DS2), and of those  
248 with high  $F_{ST}$  only (with datasets DS1 and DS3). These analyses were also performed for  
249 each species separately.

250

251 **Effects of habitat characteristics on species and genetic diversity**

252 We used partial least-square (PLS) regressions to analyze the influence of habitat  
253 characteristics on species and genetic diversity. PLS regressions are recommended in  
254 analyses involving a large number of potentially correlated predictor variables and few  
255 observations (Mevik and Wehrens 2007), as in the case of the present study. Such analyses

256 first produce a set of uncorrelated linear combinations of the original predictors, the PLS  
257 components, calculated so as to maximize the covariance between the predictor and  
258 response variables (Mevik and Wehrens 2007). Then, the PLS components are used to  
259 derive the linear regression of the dependent variable (i.e. species or genetic diversity in  
260 this case). To establish the optimal number of PLS components, we used a leave-one-out  
261 cross-validation strategy (Mevik and Cederkvist 2004). Significance of the PLS model was  
262 assessed by testing the significance of the Pearson correlation between the observed and  
263 predicted values. To determine which habitat characteristics contributed most to the model,  
264 we examined the loading weights and performed approximate t-tests based on jackknife  
265 variance estimates of regression coefficients, calculated based on the optimal number of  
266 PLS components.

267

### 268 **Species and genetic diversity correlations**

269 For each marker dataset, we computed Pearson correlations between plant species richness  
270 and Nei's gene diversity index. One-tail tests with Bonferroni-Holm corrections were  
271 applied to correct for multiple comparisons. In order to assess the overall trend in SGDCs,  
272 we calculated Student's t-tests to determine whether the average correlation calculated from  
273 the five species was greater than zero. To investigate whether the relationships between  
274 species and genetic diversity were mediated by habitat isolation, stability, and size, we  
275 performed partial Pearson correlations after controlling for significant effects of habitat  
276 characteristics on species and genetic diversity. Significance of the residual correlations  
277 was tested as described above for the raw SGDC coefficients. LMMs were performed to  
278 test for differences in SGDCs between DS1 and DS2 (dataset as fixed effect), using species  
279 as a random factor.

## 280 **Results**

### 281 **Identification of the outlier loci**

282 Depending on the species, between 66 and 90 polymorphic AFLP loci were retained in our  
283 study (Table S2). The BayeScan analysis identified between one and ten outlier loci per  
284 species, representing between 1.2% and 15.1% of the AFLP loci, respectively, with an  
285 average of 6.8% over all species (Table S2). Most of these outlier loci had high  $F_{ST}$  values,  
286 but in *P. clandestina* and *Austrelmis sp.* two loci showed low  $F_{ST}$  values (Table S2).

287

### 288 **Species richness and genetic diversity**

289 Plant species richness ranged from 6 to 21 species per wetland (Table S3) and decreased  
290 with latitude (Fig. S1). The two plant species showed low genetic diversity overall (range  
291 of average Nei's gene diversity indices: 0.11–0.13; Table S4), varying moderately to  
292 substantially between wetlands (more than six-fold in *C. gayana*; Table S4). Genetic  
293 diversity estimates of macroinvertebrates were higher (range of average Nei's gene  
294 diversity indices: 0.11–0.28; Table S4) but less variable (Table S4). The presence of outlier  
295 loci significantly influenced Nei's gene diversity. These effects were detected in four of the  
296 five species (Fig. 3), and as an overall trend across species (LMM, dataset effect:  $LR =$   
297  $12.52$ ,  $df = 1$ ,  $P < 0.001$  for DS1 vs DS2;  $LR = 24.77$ ,  $df = 1$ ,  $P < 0.001$  for DS1 vs DS3).  
298 The only species for which we failed to detect outlier loci effects on genetic diversity was  
299 *C. gayana*, but only one outlier was identified for this species. Overall, GD estimates  
300 decreased with the presence of outlier loci (Fig. 3, Fig. S1), and this trend was more  
301 pronounced when only outliers with high  $F_{ST}$  were included in the genetic datasets than  
302 when both outliers with high and low  $F_{ST}$  were included.

### 303 **Species-genetic diversity correlations and the influence of outlier loci**

304 Our analyses revealed positive correlations between species richness and genetic diversity  
305 (Table S5). The average SGDCs of the five species was high (ranging from  $0.51 \pm 0.07$  for  
306 DS3 to  $0.54 \pm 0.05$  for DS1, Table S5) and significantly positive overall (*t*-tests:  $t = 17.4$ –  
307  $25.5$ ,  $df = 4$ ,  $P < 0.001$  in all cases). All SGDCs were positive and significant with respect  
308 to each species ( $P < 0.05$  in all cases after correcting for multiple comparisons). They  
309 ranged from 0.47 to 0.60 for DS1 (Table S5). The presence of outlier loci significantly  
310 influenced the SGDCs (Fig. 4A, LMM, dataset effect:  $LR = 6.05$ ,  $df = 2$ ,  $P = 0.05$ ), causing  
311 a decrease in SGDCs overall (Fig 4A, Table S5). This trend was particularly marked in *P.*  
312 *clandestina* (Fig. 4A), for which the SGDC estimate dropped by 13.5% when outlier loci  
313 with high  $F_{ST}$  were included (Fig. 4A).

314

### 315 **Influence of habitat fragmentation on species richness, genetic diversity and SGDCs**

316 The PLS regressions explained 49% of the variation in wetland plant richness (Table 1).  
317 Significant influence of wetland size and connectivity on genetic diversity was found for  
318 the plants and the mayfly species (*A. peruvianus*), but no such effects were detected for the  
319 long-lived aquatic species (*H. fossamancinii* and *Austrelmis sp*). Wetland isolation was an  
320 important determinant of plant species richness and genetic diversity. We found a negative  
321 influence of the average distance to neighboring patches (i.e. closeness centrality, Table 1)  
322 and positive effects of the number of neighboring patches (i.e. node degree, Table 1) on  
323 genetic and species diversity. A positive effect of the distance to the farthest wetland (i.e.  
324 eccentricity) on species richness and genetic diversity was observed for the plants and the  
325 mayfly species (Table 1).



326 The effects of habitat fragmentation on GD were not lowered by the presence of  
327 outlier loci (i.e. DS2 and DS3, Table 1). In fact, the opposite trend was observed in *P.*  
328 *clandestina* and *A. peruvianus*, with habitat connectivity explaining a greater proportion of  
329 the variance in GD for DS2/DS3 than for DS1 (Table 1).

330 Factoring out the effects of habitat fragmentation on species and genetic diversity  
331 significantly influenced SGDC estimates (LMM, fragmentation habitat control effect:  $LR =$   
332  $8.42$ ,  $df = 1$ ,  $P < 0.01$ ), decreasing average partial SDGCs relative to average original  
333 values (Fig. 4B, Table S5). The decreases in partial SGDCs were more consistent and  
334 pronounced with DS1 than with DS2 (Fig. 4B). In *A. peruvianus*, controlling for habitat  
335 fragmentation had strikingly contrasting effects depending on whether or not the outlier loci  
336 were included in the analysis (Fig. 4B). As in other species, the partial SGDC decreased  
337 with DS1, but increased considerably with DS2, reaching a maximum of 0.7 (Table S5). To  
338 confirm that this discrepancy was specifically linked to the three outlier loci detected, we  
339 applied a bootstrap procedure to simulate the null distribution of the difference in partial  
340 SGDCs between DS1 and DS2 when randomly eliminating three non-outlier loci from  
341 DS2. None of the 1,000 bootstrap replications resulted in SGDC differences greater than or  
342 equal to the one observed between DS1 and DS2.

343

## 344 **Discussion**

### 345 **Confounding effects of outlier loci in SGDC studies**

346

347 There is increased awareness that neutral molecular markers are not always free of  
348 selection pressures; disregarding this fact can bias population genetic inferences (Luikart et  
349 al. 2003, Landguth & Balkenhol 2012). To avoid this problem, Luikart et al. (2003)

350 advocated excluding  $F_{ST}$  outlier loci from genetic analyses when the focus is on neutral  
351 processes. In this study, we screened for AFLP outlier loci and investigated their impact on  
352 SGDC estimates. We also examined how these loci influence the apparent contribution of  
353 habitat characteristics linked to neutral processes (migration rates and rates of stochastic  
354 loss of alleles/species) on SGDCs. Our results show that outlier loci led to a marked  
355 decrease in GD, which in turn downplayed the strength of SGDCs.

356         Outlier loci have been suggested to be involved in adaptive processes (Hancock *et*  
357 *al.* 2011; Manel *et al.* 2012). Yet, alternative mechanisms, which may or may not be  
358 indirectly associated with selective forces, have also been proposed to explain outlier loci,  
359 such as correlations in co-ancestry between subpopulations in hierarchical structures, gene  
360 surfing, background selection, and even endogenous genetic barriers (Bierne *et al.* 2011,  
361 2013). Regardless of the actual mechanisms involved in the atypical of the actual  
362 mechanisms involved in the atypical differentiation patterns of the outlier loci, those with  
363 markedly high  $F_{ST}$  demonstrated the most pronounced effects in terms of reductions in GD.  
364 This result concurs with previous empirical observations based on microsatellite data  
365 (Garcia-Verdugo *et al.* 2015), and thus suggests that GD might often be underestimated  
366 when such loci are included in population genetic analyses.

367         We found a significant downward effect of outlier loci on SGDCs, indicating that  
368 SD correlates more strongly with GD of non-outlier loci than with GD of outlier loci. This  
369 pattern is consistent with the expectation that SD is more closely related to neutral GD than  
370 to non-neutral GD, due to the parallel influence of neutral processes on both diversity  
371 levels. Investigating the effects of filtering out outlier loci from SGDC studies is therefore a  
372 potentially useful approach to evidence the contribution of neutral processes. In addition, it  
373 may contribute to revealing SGDCs, and explain some of the variation in species-genetic

374 diversity relationships observed in the field. Indeed, the effects of outlier loci are likely to  
375 depend on the strength of neutral mechanisms on both diversity components. In high  
376 Andean wetlands, and highly fragmented ecosystems in general, neutral evolutionary  
377 processes are likely key determinants of biodiversity. As a result, they are expected to  
378 generate strong species-genetic diversity covariations (Vellend and Geber 2005; Whitlock  
379 2014; Vellend *et al.* 2014), as were found here, with these SGDCs ranking among the  
380 highest values reported to date (Vellend and Geber 2005; Whitlock 2014; Vellend *et al.*  
381 2014). In such conditions, non-neutral processes, and thus outlier loci, should only have  
382 minor confounding effects. The influence of outlier loci on SGDCs is likely to become  
383 increasingly problematic as the relative importance of neutral processes in shaping diversity  
384 patterns decreases. Extending the present framework to less fragmented ecosystems may  
385 potentially uncover further SGDC patterns, thereby advancing current understanding of  
386 species-genetic diversity relationships.

387         The presence of outlier loci also complicated the evaluation of the contribution to  
388 SGDCs of habitat characteristics supposedly linked to neutral processes. Without such loci,  
389 the partial SGDCs, calculated after filtering out the effects of habitat fragmentation on SD  
390 and GD, decreased compared to the corresponding non-partial estimates. This indicates  
391 concordant responses of SD and GD at non-outlier loci to these habitat characteristics.  
392 Outlier loci somehow blurred these trends however, as reduction of partial SGDCs was less  
393 consistent and less pronounced among species with outlier loci present. In the mayfly *A.*  
394 *peruvianus*, the SGDC with outlier loci actually increased when habitat effects were  
395 factored out (Fig. 4B). This indicates that contrasting effects of habitat characteristics on  
396 SD and GD at outlier loci were partly masking the species-genetic diversity relationship in  
397 this species. It further shows that GD at outlier loci and neutral GD can correlate differently

398 with habitat features, thus confounding the apparent importance of neutral mechanisms on  
399 GD. In fact, for both *P. clandestina* and *A. peruvianus*, the habitat features showed stronger  
400 correlations with GD at outlier than at non-outlier loci (Table 1). Habitat stability, size, and  
401 connectivity, being intrinsic determinants of population demography, are expected to  
402 influence all neutral loci equally (Luikart *et al.* 2003). Their stronger association with  
403 outlier loci in two species thus suggests that they can also be indirectly associated with  
404 selective processes and local adaptation. In our case, high altitude wetland densities decline  
405 in parallel with latitude, and both wetland size and connectivity show spatial structure along  
406 this gradient (Fig. S2, Supplementary Material). In such circumstances, the standard  
407 approach to disentangling the influence of neutral processes on SGDCs – i.e., factoring out  
408 the effects of habitat features– may be ineffective, since it can simultaneously omit the  
409 effects of selective environmental factors. Our results thus illustrate the difficulties in  
410 inferring evolutionary processes from partial SGDCs when habitat density is paired with  
411 environmental gradient, a situation likely to be common in ecosystems subject to climatic  
412 and/or topographic influence. To overcome pitfalls in interpreting the contribution of  
413 disturbance regime, habitat size and/or connectivity to SGDCs, it would be advisable to  
414 analyze the spatial distribution of these variables and evaluate their correlation with other  
415 environmental factors as potential actors of selection.

416

#### 417 **Species-genetic diversity correlations in high altitude wetlands: patterns and processes**

418 We found strong spatial associations between species richness and AFLP genetic diversity  
419 estimates in high altitude wetlands. Consistent with simulation models suggesting stronger  
420 species-genetic diversity relationships when GD is measured in more common species  
421 (Vellend 2005), highest SGDC was obtained for *C. gayana*, the most abundant plant

422 species of high altitude wetlands of Chile's Norte Chico. Overall, our results add to  
423 previous meta-analysis evidence that SGDCs are widespread in discrete habitats  
424 functioning as islands (Vellend and Geber 2005; Whitlock 2014; Vellend *et al.* 2014).  
425 Furthermore, our results show that such correlations are not limited to component species  
426 of the community. For instance, we detected strong SGDCs between plant species richness  
427 and GD of the benthic macro-invertebrate taxa, despite the fact that the latter possess very  
428 different life-history traits, particularly regarding their dispersal modes.

429         SGDCs in patchy ecosystems are thought to result from parallel influences of  
430 neutral forces governing variation at the genetic and species diversity levels (Vellend and  
431 Geber 2005; Whitlock 2014; Vellend *et al.* 2014). Here, the contribution of such  
432 mechanisms is substantiated both by the lowering effect of outlier loci on SGDCs, as well  
433 as the apparent common responses of SD and GD at non-outlier loci to habitat features  
434 (size, connectivity and stability). The persistence of positive, and overall significant, partial  
435 SGDCs after removing all confounding effects (outlier loci and habitat features) suggests,  
436 however, that other causal mechanisms might also be involved. These can include both  
437 parallel responses to processes unrelated to habitat features, and direct relationships  
438 between SD and GD. While our results indicate that potentially selected loci tend to  
439 decrease SGDCs overall, they do not rule out the possibility that some selective  
440 mechanisms may contribute to species-genetic diversity relationships. For instance, a strong  
441 selective agent, e.g. a voracious herbivore or an aggressive exotic plant, acting with  
442 increasing intensity upon plant communities located along a spatial gradient might cause a  
443 gradual reduction in SD (due to differences in species' abilities to cope with increased  
444 predation/competition) together with a reduction of GD in the "surviving" species, where  
445 only some tolerant genotypes could remain; this would generate a positive correlation

446 between SD and GD. In our study system, plant SD may affect soil and water  
447 characteristics and, as such, may influence local GD by determining the abundance of plant  
448 and macroinvertebrate populations. Alternatively, genetic variation within foundation  
449 species such as *P. clandestina*, which hosts other plant species, can affect the composition  
450 and diversity of associated communities (Whitham *et al.* 2006). Nevertheless, this  
451 conjecture would imply a significant correlation between GD at non-outlier loci and GD of  
452 specific genes encoding ecologically relevant traits influencing plant assemblages.

453 Further investigation is required to fully unravel the mechanisms driving SGDCs in  
454 high Andean wetlands. Future studies will benefit from next-generation sequencing applied  
455 to genome scans, which will very soon allow for more reliable estimates of genes under  
456 selection (Manel *et al.* 2016), essential to distinguishing neutral from adaptive markers.  
457 Ideally, field studies could complement these efforts by identifying strong interactions  
458 between species in the community, putative selective pressures, and patterns of selection  
459 and adaptive evolution.

460

## 461 **Acknowledgements**

462 We thank F. Squeo, G. Arancio, D. Rodriguez, R. Inzulza Ayala, L. Cifuentes, E. Alvarez,  
463 C. Urqueta, R. Hereme and A. Troncoso for their help with field work and/or for their  
464 technical assistance, and Craig Weideman for revising the English of the manuscript. Our  
465 thanks go to P. Jarne and F. Laroche for insightful discussion during the preparation of this  
466 manuscript. We also thank four anonymous reviewers for their comments on previous  
467 versions of the manuscript. The authors are grateful to La Junta de Vigilancia del Elqui, la  
468 Comunidad Agrícola los Huascoaltinos, La Sociedad de Parceleros de Coiron, La Sociedad

469 de Parceleros de San Agustín, La Sociedad de Parceleros Hacienda Illapel, la Hacienda  
470 Tulahuén Oriente, la Hacienda El Maiten de Pedregal and la Hacienda El Bosque for giving  
471 us access to private wetlands. This research was funded by a grant from the Chilean  
472 National Research Agency FONDECYT #1110514 and the exchange research grant ECOS-  
473 CONICYT C12B02.

474

## 475 **References**

- 476 Antonovics J (1992) Toward community genetics. In: *Plant resistance to herbivores and*  
477 *pathogens: ecology, evolution, genetics* (eds. Fritz RS, Simms E), pp. 426-449. The  
478 University of Chicago Press, Chicago.
- 479 Arroyo MTK, Marquet P, Marticorena C *et al.* (2004) Chilean winter rainfall-Valdivian  
480 forests. In: *Hotspots revised: Earth's biologically wealthiest and most threatened*  
481 *ecosystems.* (eds. Mittermeier RA, Gil PR, Hoffmann M *et al.*), pp. 99-103. CEMEX,  
482 México D.F.
- 483 Barton NH (2000) Genetic hitchhiking. *Philosophical Transactions of the Royal Society of*  
484 *London Series B-Biological Sciences*, **355**, 1553-1562.
- 485 Bertin A, Alvarez E, Gouin N *et al.* (2015) Effects of wind-driven spatial structure and  
486 environmental heterogeneity on high-altitude wetland macroinvertebrate assemblages  
487 with contrasting dispersal modes. *Freshwater Biology*, **60**, 297-310.
- 488 Bierne N, Welch J, Loire E *et al.* (2011) The coupling hypothesis: why genome scans may  
489 fail to map local adaptation genes. *Molecular Ecology*, **20**, 2044-2072.
- 490 Bierne N, Roze D, Welch JJ (2013). Pervasive selection or is it ...? Why are Fst outliers  
491 sometimes so frequent? *Molecular Ecology*, **22**, 2061-2064.

492 Bonin A, Ehrich D, Manel S (2007) Statistical analysis of amplified fragment length  
493 polymorphism data: a toolbox for molecular ecologists and evolutionists  
494 6. *Molecular Ecology*, **16**, 3737-3758.

495 Duchesne P, Bernatchez L (2002) AFLPOP: a computer program for simulated and real  
496 population allocation, based on AFLP data. *Molecular Ecology Notes*, **2**, 380-383.

497 Eckert C, Samis K, Lougheed S (2008) Genetic variation across species' geographical  
498 ranges: the central-marginal hypothesis and beyond. *Molecular Ecology*, **17**, 1170-  
499 1188.

500 Ehrich D (2006) AFLPDAT: a collection of R functions for convenient handling of AFLP  
501 data. *Molecular Ecology Notes*, **6**, 603-604.

502 Foll M, Gaggiotti O (2008) A Genome-Scan Method to Identify Selected Loci Appropriate  
503 for Both Dominant and Codominant Markers: A Bayesian Perspective. *Genetics*, **180**,  
504 977-993.

505 Foltete JC, Clauzel C, Vuidel G (2012) A software tool dedicated to the modelling of  
506 landscape networks. *Environmental Modelling & Software*, **38**, 316-327.

507 Frankham R (1996) Relationship of genetic variation to population size in wildlife.  
508 *Conservation Biology*, **10**, 1500-1508.

509 Frankham R (1997) Do island populations have less genetic variation than mainland  
510 populations? *Heredity*, **78**, 311-327.

511 García-Verdugo C, Sajeve M, La Mantia T *et al.* (2015) Do island plant populations really  
512 have lower genetic variation than mainland populations? Effects of selection and  
513 distribution range on genetic diversity estimates. *Molecular Ecology*, **24**, 726-741.

514 Hancock AM, Brachi B, Faure N *et al.* (2011) Adaptation to Climate Across the  
515 *Arabidopsis thaliana* Genome. *Science*, **334**, 83-86.



516 He T, Lamont BB, Krauss SL, Enright NJ, Miller BP (2008) Covariation between  
517 intraspecific genetic diversity and species diversity within a plant functional group.  
518 *Journal of Ecology*, **96**, 956-961.

519 Holderegger R, Kamm U, Gugerli F (2006) Adaptive vs. neutral genetic diversity:  
520 implications for landscape genetics. *Landscape Ecology*, **21**, 797-807.

521 Hu FS, Hampe A, Petit RJ (2009) Paleoecology meets genetics: deciphering past  
522 vegetational dynamics. *Frontiers in Ecology and the Environment*, **7**, 371-379.

523 Hubbell, S.P. (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*,  
524 Princeton University Press.

525 Johnson MTJ, Stinchcombe JR (2007) An emerging synthesis between community ecology  
526 and evolutionary biology. *Trends in Ecology & Evolution*, **22**, 250-257.

527 Kahilainen A, Puurtinen M, Kotiaho JS (2014) Conservation implications of species-  
528 genetic diversity correlations. *Global Ecology Conservation*, **2**, 315-323.

529 Kirk H, Freeland JR (2011) Applications and Implications of Neutral versus Non-neutral  
530 Markers in Molecular Ecology. *International Journal of Molecular Sciences*, **12**,  
531 3966-3988.

532 Lamy T, Jarne P, Laroche F *et al.* (2013) Variation in habitat connectivity generates  
533 positive correlations between species and genetic diversity in a metacommunity.  
534 *Molecular Ecology*, **22**, 4445-4456.

535 Lankau RA, Strauss SY (2007) Mutual feedbacks maintain both genetic and species  
536 diversity in a plant community. *Science*, **317**, 1561-1563.

537 Laroche F, Jarne P, Lamy T, David P, Massol F (2015) A Neutral Theory for Interpreting  
538 Correlations between Species and Genetic Diversity in Communities. *American*  
539 *Naturalist*, **185**, 59-69.

540 Lavergne S, Mouquet N, Thuiller W, Ronce O (2010) Biodiversity and Climate Change:  
541 Integrating Evolutionary and Ecological Responses of Species and Communities.  
542 *Annual Review of Ecology, Evolution, and Systematics, Vol 41*, **41**, 321-350.

543 Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of  
544 population genomics: From genotyping to genome typing  
545 13. *Nature Reviews Genetics*, **4**, 981-994.

546 MacArthur R.H., Wilson E.O. (1967) The theory of island biogeography. Princeton  
547 University Press, New Jersey.

548 Manel S, Perrier C, Pratlong M *et al.* (2016) Genomic resources and their influence on the  
549 detection of the signal of positive selection in genome scans. *Molecular Ecology*, **25**,  
550 170-184.

551 Manel S, Gugerli F, Thuiller W *et al.* (2012) Broad-scale adaptive genetic variation in  
552 alpine plants is driven by temperature and precipitation. *Molecular Ecology*, **21**,  
553 3729-3738.

554 Meudt HM, Clarke AC (2007) Almost forgotten or latest practice? AFLP applications,  
555 analyses and advances. *Trends in Plant Science*, **12**, 106-117.

556 Mevik BH, Cederkvist HR (2004) Mean squared error of prediction (MSEP) estimates for  
557 principal component regression (PCR) and partial least squares regression (PLSR).  
558 *Journal of Chemometrics*, **18**, 422-429.

559 Mevik BH, Wehrens R (2007) The pls package: Principal component and partial least  
560 squares regression in R. *Journal of Statistical Software*, **18**, 1-23.

561 Mouquet N, Devictor V, Meynard CN *et al.* (2012) Ecophylogenetics: advances and  
562 perspectives. *Biological Reviews*, **87**, 769-785.

563 Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York, N.  
564 Y.

565 Odat N, Jetschke G, Hellwig FH (2004) Genetic diversity of *Ranunculus acris* L.  
566 (Ranunculaceae) populations in relation to species diversity and habitat type in  
567 grassland communities. *Molecular Ecology*, **13**, 1251-1257.

568 Oleksyk TK, Smith MW, O'Brien SJ (2010) Genome-wide scans for footprints of natural  
569 selection. *Philosophical Transactions of the Royal Society B-Biological Sciences*,  
570 **365**, 185-205.

571 Papadopoulou A, Anastasiou I, Spagopoulou F *et al.* (2011) Testing the Species-Genetic  
572 Diversity Correlation in the Aegean Archipelago: Toward a Haplotype-Based  
573 Macroecology? *American Naturalist*, **178**, 241-255.

574 R Core Team. A language and environment for statistical computing. 2015. Vienna,  
575 Austria, R Foundation for Statistical Computing.

576 Ricklefs RE, Lovette IJ (1999) The roles of island area per se and habitat diversity in the  
577 species-area relationships of four Lesser Antillean faunal groups. *Journal of Animal*  
578 *Ecology*, **68**, 1142-1160.

579 Rosenzweig (1995) *Species diversity in space and time*, Cambridge University Press  
580 ednCambridge, UK.

581 Rossi J (2011) Rich: an R package to analyse species richness. *Diversity*, **3**, 112-120.

582 Rouse JW, Haas RH, Schell J, Deering D (1973) Monitoring vegetation systems in the  
583 Great Plains with ERTS. In: *Third ERTS Symposium* (ed. NASA SP-351 I), pp. 309-  
584 317.

585 Silvertown J, Biss PM, Freeland J (2009) Community genetics: resource addition has  
586 opposing effects on genetic and species diversity in a 150-year experiment. *Ecology*  
587 *Letters*, **12**, 165-170.

588 Squeo FA, Warner BG, Aravena R, Espinoza D (2006) Bofedales: high altitude peatlands  
589 of the central Andes. *Revista Chilena de Historia Natural*, **79**, 245-255.

590 Storz JF, Payseur BA, Nachman MW (2004) Genome scans of DNA variability in humans  
591 reveal evidence for selective sweeps outside of Africa. *Molecular Biology and*  
592 *Evolution*, **21**, 1800-1811.

593 Strauss SY, Sahli H, Conner JK (2005) Toward a more trait-centered approach to diffuse  
594 (co)evolution. *New Phytologist*, **165**, 81-89.

595 Struebig MJ, Kingston T, Petit EJ *et al.* (2011) Parallel declines in species and genetic  
596 diversity in tropical forest fragments. *Ecology Letters*, **14**, 582-590.

597 Taberlet P, Zimmermann NE, Englisch T *et al.* (2012) Genetic diversity in widespread  
598 species is not congruent with species richness in alpine plant communities. *Ecology*  
599 *Letters*, **15**, 1439-1448.

600 Vellend M (2004) Parallel effects of land-use history on species diversity and genetic  
601 diversity of forest herbs. *Ecology*, **85**, 3043-3055.

602 Vellend M (2005) Species diversity and genetic diversity: Parallel processes and correlated  
603 patterns. *American Naturalist*, **166**, 199-215.

604 Vellend M, Geber MA (2005) Connections between species diversity and genetic diversity.  
605 *Ecology Letters*, **8**, 767-781.

606 Vellend M, Lajoie G, Bourret A *et al.* (2014) Drawing ecological inferences from  
607 coincident patterns of population- and community-level biodiversity. *Molecular*  
608 *Ecology*, **23**, 2890-2901.

609 Whitham TG, Bailey JK, Schweitzer JA *et al.* (2006) A framework for community and  
610 ecosystem genetics: from genes to ecosystems. *Nature Reviews Genetics*, **7**, 510-523.

611 Whitlock R (2014) Relationships between adaptive and neutral genetic diversity and  
612 ecological structure and functioning: a meta-analysis. *Journal of Ecology*, **102**, 857-  
613 872.

614 Xu W, Liu L, He T *et al.* (2016) Soil properties drive a negative correlation between  
615 species diversity and genetic diversity in a tropical seasonal rainforest. *Scientific*  
616 *Reports*, **6**.

617

## 618 **Data accessibility**

619 Raw files of the AFLP genotypes (i.e. band presence and absence) for each of the five  
620 species and R scripts will be made available from the Dryad Digital Repository upon  
621 acceptance of the manuscript.

622

## 623 **Author contributions**

624 ABertin, NG, ABaumel, EG and SM conceived and designed the study. ABertin, NG, JS  
625 and RO performed the experiments. ABertin, NG and ABaumel analyzed the data. ABertin,  
626 NG, ABaumel and SM wrote the manuscript. All the authors contributed substantially to  
627 revisions.

628

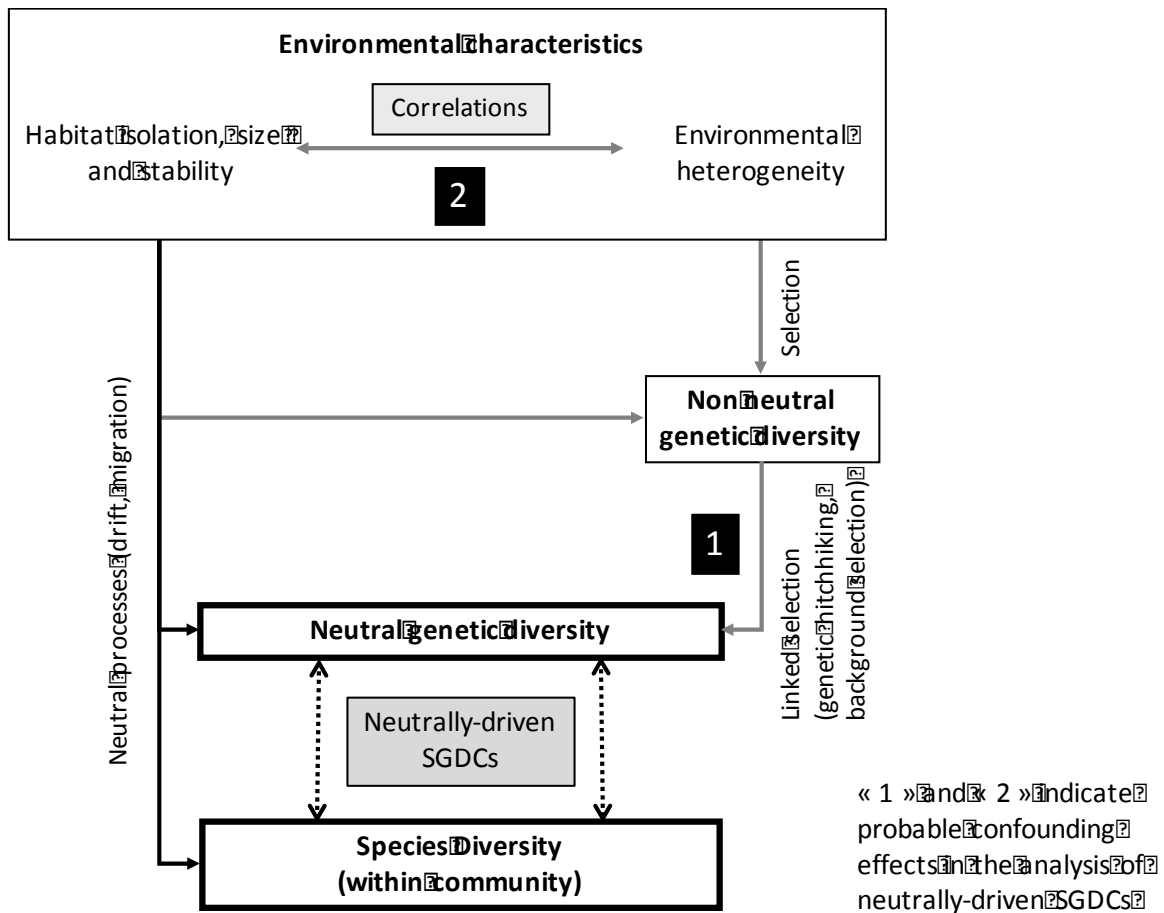
629 **Tables**

630 **Table 1.** Effects of habitat characteristics on genetic diversity and taxa diversity estimated using partial least-square regressions  
 631 (PLSR). The direction of the effect (positive/negative) of each habitat characteristic on the diversity parameters is reported in  
 632 parentheses. DS1 refers to the AFLP dataset excluding outlier loci, DS2 to the dataset including both non-outlier and outlier loci  
 633 (potentially under diversifying or balancing selection), and DS3 to the dataset including non-outlier and outlier loci with high  $F_{ST}$   
 634 (potentially under diversifying selection).

	AFLP dataset	Number of sites	Optimal number of PLSR axes	Most contributing variables	Explained variance ( $R^2$ )
<b><i>Species diversity</i></b>		21	3	Closeness centrality (<0), Eccentricity (>0), Node degree (>0)	0.49***
<b><i>Genetic Diversity</i></b>					
<i>C. gayana</i>	DS1	21	3	Node degree (>0), Closeness centrality (<0), Eccentricity (>0)	0.51***
	DS2 / DS3	21	3	Node degree (>0), Closeness centrality (<0), Eccentricity (>0)	0.50***
<i>P. clandestina</i>	DS1	16	1	Node degree (>0), Capacity (>0)	0.19 <sup>†</sup>

	DS2	16	1	Closeness centrality (<0), Eccentricity (>0), Node degree (>0), Capacity (>0)	0.35*
	DS3	16	2	Node degree (>0), Capacity (>0)	0.31 <sup>†</sup>
<i>H. fossamancinii</i>	DS1	14	0	-	-
	DS2 / DS3	14	0	-	-
<i>A. peruvianus</i>	DS1	14	0	-	-
	DS2 / DS3	14	1	Closeness centrality (<0)	0.43*
<i>Austrelmis sp</i>	DS1	15	0	-	-
	DS2	15	0	-	-
	DS3	15	0	-	-

635 <sup>†</sup>:  $P < 0.1$ , \*:  $P < 0.05$  and \*\*\*:  $P < 0.001$

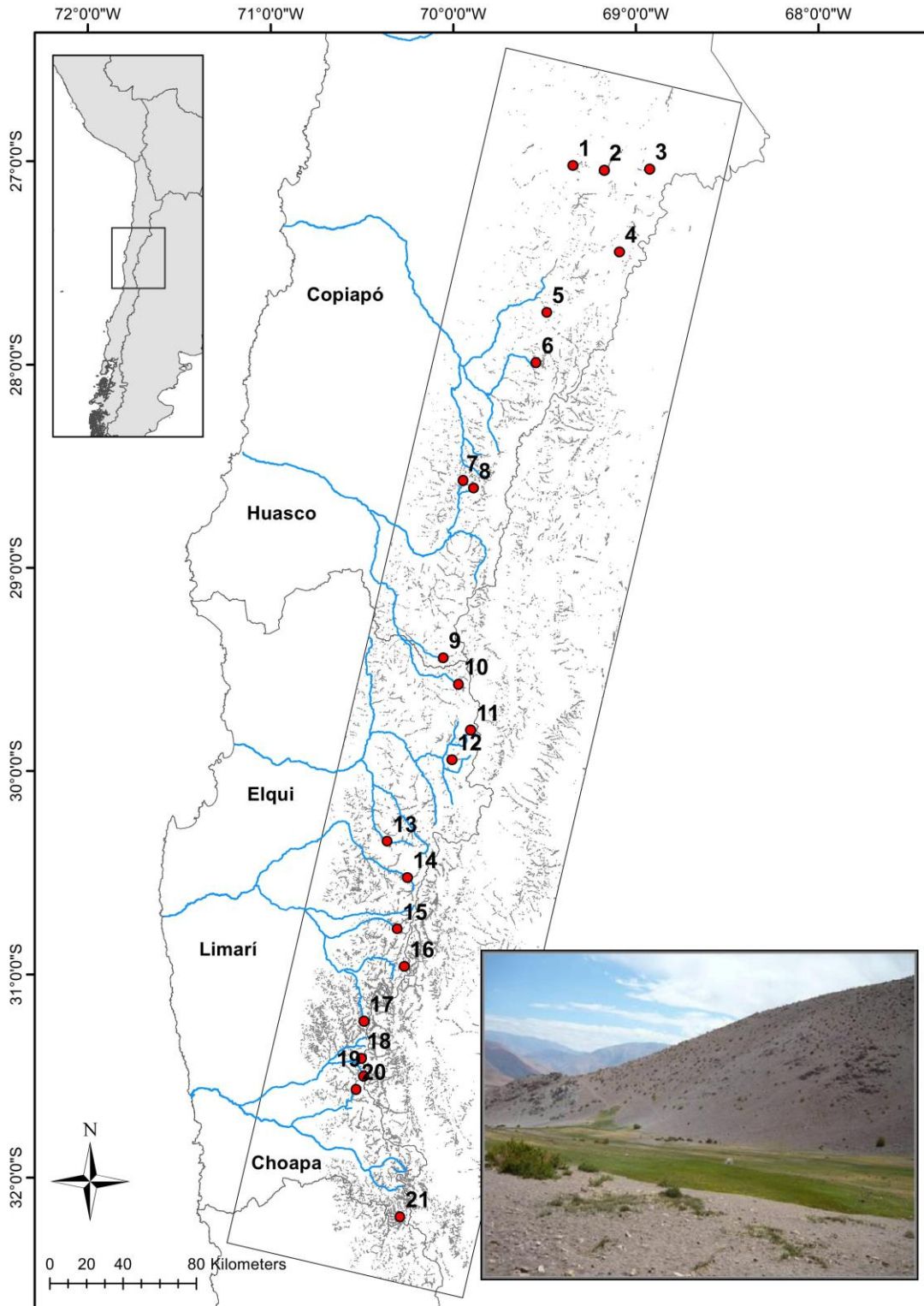


637

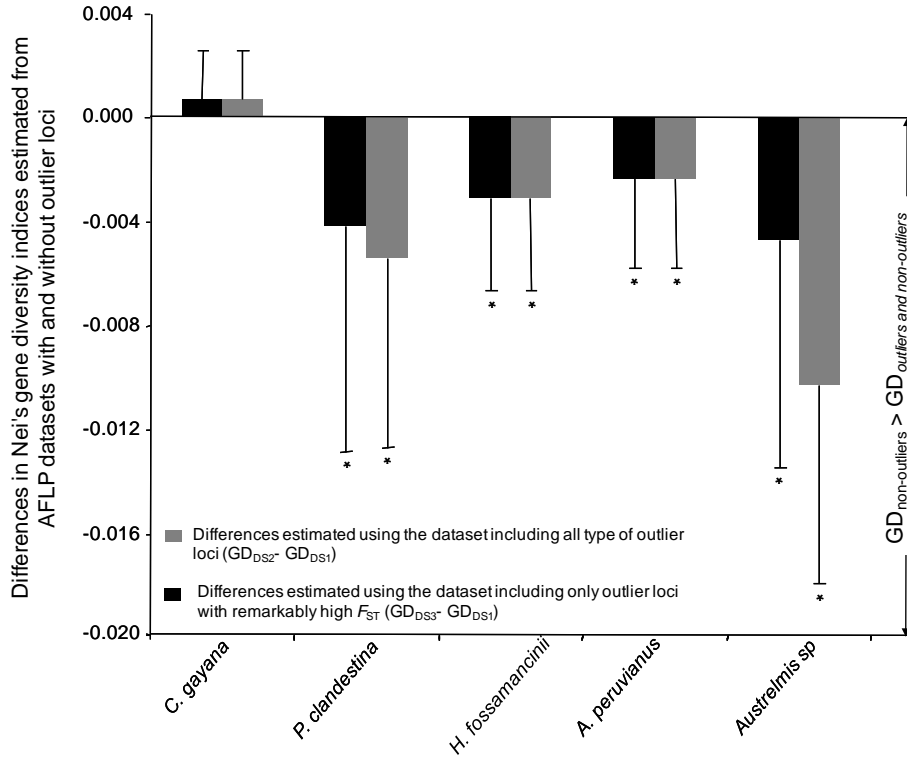
638

639 **Figure 1.** Relationship between species diversity within a community and neutral genetic  
 640 diversity within a species. Species-genetic diversity correlations (SGDCs) with neutral  
 641 genetic diversity are assumed to result from parallel responses to neutral processes (black  
 642 arrows), which relative importance is expected to be determined by habitat features such as  
 643 disturbance regime, habitat size and/or connectivity. Yet, non-neutral processes can have  
 644 confounding effects (grey arrows) in the investigation of neutrally-driven SGDCs. They can  
 645 confound the apparent level of SGDCs driven by neutral processes (confounding effect 1)  
 646 influencing neutral genetic diversity through linked selection. The contribution of habitat  
 647 characteristics regulating rates of stochastic immigration and extinction to SGDCs may also  
 648 be confounded by linked selection if adaptive genetic markers are selected by  
 649 environmental conditions covarying with those habitat characteristics (confounding effect  
 650 2).



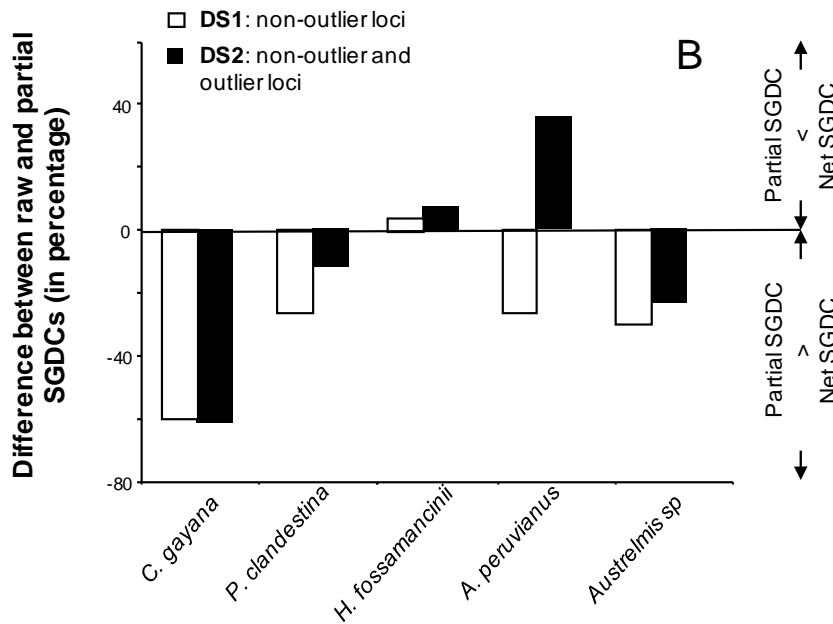
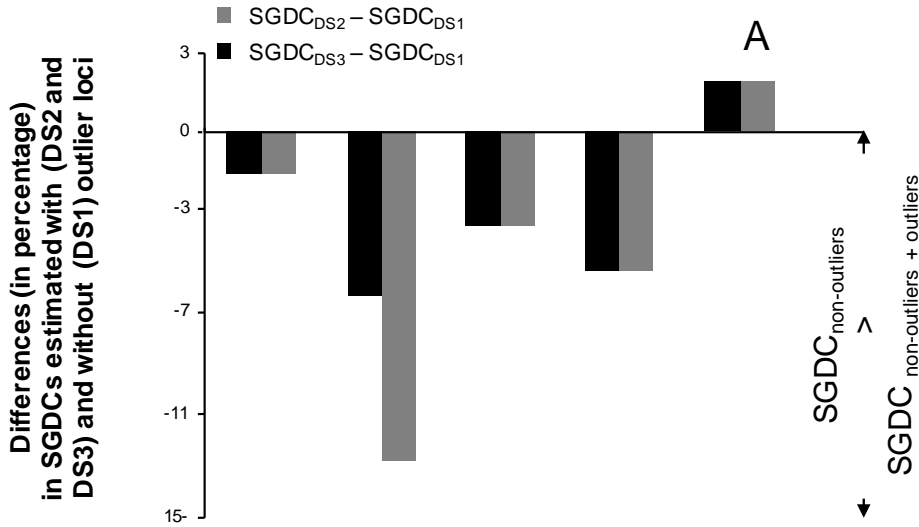


651 **Figure 2.** Picture of a high altitude wetland and geographical distribution of the sampling  
 652 sites along Chile's Norte Chico. The rectangle depicts the remote-sensing area used for  
 653 identification and mapping of high Andean wetlands (shown in grey).



654  
655  
656  
657  
658  
659

**Figure 3.** Average population difference + or – SD between Nei's gene diversity indices estimated from the AFLP datasets including outlier loci (DS2 and DS3) and the dataset free of outliers (DS1). Asterisks indicate significant differences ( $P < 0.05$ ) in average population genetic diversity estimated with and without outlier loci for each comparison in each species.



660  
 661 **Figure 4.** Differences between species genetic diversity correlations (SGDCs) estimated  
 662 under various scenarios. A) Differences between SGDCs estimated without (SGDC<sub>DS1</sub>) and  
 663 with outlier loci (SGDC<sub>DS2</sub> and SGDC<sub>DS3</sub>). SGDC<sub>DS2</sub> - SGDC<sub>DS1</sub> refers to the difference  
 664 between SGDCs calculated with the full AFLP dataset including both non-outlier and  
 665 outlier loci (DS2) and the dataset excluding all the outlier loci (DS1), and SGDC<sub>DS3</sub> -  
 666 SGDC<sub>DS1</sub> to the difference between SGDCs calculated with the dataset including non-  
 667 outlier and outlier loci with high  $F_{ST}$  only (DS3) and the dataset excluding all the outlier  
 668 loci (DS1). B) Difference between raw SGDCs and partial SGDCs, calculated after  
 669 accounting for habitat characteristics, for each focal species and for datasets without (DS1)  
 670 and with all the outlier loci (DS2).  
 671