APPENDIX S1.- List of individuals by catalogue number, specie, locality, initials, geographic coordinates and GeneBank accession number.

Catalogue Number	Species	Locality	Initials	Latitude (S)	Longitude (W)	Genebank accesion number
NK 106069	*Phyllotis darwini*	Cerro el Roble, Til-Til	CR 69	32°59´ 24´´	71°1´12 ´´	JN226723
NK 106137	*Phyllotis darwini*	Cerro el Roble, Til-Til	CR 37	32°59´ 24´´	71°1´12 ´´	JN226724
NK 106141	*Phyllotis darwini*	Cerro el Roble, Til-Til	CR 41	32°59´ 24´´	71°1´12 ´´	JN226725
NK 106146	*Phyllotis darwini*	Cerro el Roble, Til-Til	CR 46	32°59´ 24´´	71°1´12 ´´	JN226726
NK 105341	*Phyllotis darwini*	Cerro Santa Inés, Los Vilos	CSI 41	32°9´0 ´´	71°28´4 8´´	JN226718
NK 105348	*Phyllotis darwini*	Cerro Santa Inés, Los Vilos	CSI 48	32°9´0 ´´	71°28´4 8´´	JN226719
NK 105349	*Phyllotis darwini*	Cerro Santa Inés, Los Vilos	CSI 49	32°9´0 ´´	71°28´4 8´´	JN226720
NK 95618	*Phyllotis darwini*	Chillepín	CH 18	31°52’ 48´´	70°47´2 4´´	JN226690

NK 95620	*Phyllotis darwini*	Chillepín	CH 20	31°52’ 48´´	70°47´2 4´´	JN226691
NK 95624	*Phyllotis darwini*	Chillepín	CH 24	31°52’ 48´´	70°47´2 4´´	JN226692
NK 120051	*Phyllotis darwini*	Concepción, Tomé Agua Tendida	Ate 51	36°38´ 24´´	72°47´2 4´´	JN226727
NK 120068	*Phyllotis darwini*	Concepción, Tomé Agua Tendida	Ate 68	36°38´ 24´´	72°47´2 4´´	JN226728
ER3	*Phyllotis magister*	Desembocadura Río Loa	Ph m3	21°25´ 33´´	70°33´3 6´´	JN226732
ER4	*Phyllotis magister*	Desembocadura Río Loa	Ph m4	21°25´ 33´´	70°33´3 6´´	JN226733
ER5	*Phyllotis magister*	Desembocadura Río Loa	Ph m5	21°25´ 33´´	70°33´3 6´´	JN226734
ER7	*Phyllotis magister*	Desembocadura Río Loa	Ph m7	21°25´ 33´´	70°33´3 6´´	JN226735
SSUC-Ma-00451	*Phyllotis darwini*	El Guanaco, Quirihue, Ñuble	QUI 51	36°24´ 0´´	72°37´3 0´´	JN226683
SSUC-Ma-00452	*Phyllotis darwini*	El Guanaco, Quirihue, Ñuble	QUI 52	36°24´ 0´´	72°37´3 0´´	JN226684
SSUC-Ma-00453	*Phyllotis darwini*	El Guanaco, Quirihue, Ñuble	QUI 53	36°24´ 0´´	72°37´3 0´´	JN226685
SSUC-	*Phyllotis*	El Guanaco,	QUI	36°24´	72°37´3	JN226686

Ma-00457	*darwini*	Quirihue, Ñuble	57	0´´	0´´	
SSUC-	*Phyllotis*	El Guanaco,	QUI	36°24´	72°37´3	JN226687
Ma-00458	*darwini*	Quirihue, Ñuble	58	0´´	0´´	
SSUC-	*Phyllotis*	El Guanaco,	QUI	36°24´	72°37´3	JN226688
Ma-00459	*darwini*	Quirihue, Ñuble	59	0´´	0´´	
SSUC-	*Phyllotis*	El Guanaco,	QUI	36°24´	72°37´3	JN226689
Ma-00460	*darwini*	Quirihue, Ñuble	60	0´´	0´´	
NK 96868	*Phyllotis*	Fray Jorge	FJ6	30°40´	71°37´4	JN226713
	darwini		8	12´´	8´´	
NK 96869	*Phyllotis*	Fray Jorge	FJ6	30°40´	71°37´4	JN226714
	darwini		9	12´´	8´´	
NK 96871	*Phyllotis*	Fray Jorge	FJ7	30°40´	71°37´4	JN226715
	darwini		1	12´´	8´´	
NK 96874	*Phyllotis*	Fray Jorge	FJ7	30°40´	71°37´4	JN226716
	darwini		4	12´´	8´´	
NK	*Phyllotis*	Fray Jorge	FJ1	30°40´	71°37´4	JN226722
105515	*darwini*		5	12´´	8´´	
SSUC-	*Phyllotis*	Las Tacas	LT2	30°0´0	71°22´2	JN226672
Ma-00426	*darwini*		6	´´	6´´	
SSUC-	*Phyllotis*	Las Tacas	LT2	30°0´0	71°22´2	JN226673
Ma-00427	*darwini*		7	´´	6´´	
SSUC-	*Phyllotis*	Las Tacas	LT2	30°0´0	71°22´2	JN226674
Ma-00428	*darwini*		8	´´	6´´	

SSUC-Ma-00429	*Phyllotis darwini*	Las Tacas	LT29	30°0´0´´	71°22´26´´	JN226675
SSUC-Ma-00431	*Phyllotis darwini*	Las Tacas	LT31	30°0´0´´	71°22´26´´	JN226676
SSUC-Ma-00433	*Phyllotis darwini*	Las Tacas	LT33	30°0´0´´	71°22´26´´	JN226677
NK 160004	*Phyllotis darwini*	Llanos del Challe	LCH04	28°4´48´´	71°8´24´´	JN226731
SSUC-Ma-00442	*Phyllotis darwini*	Los Molles	LMo42	30°14´24´´	71°30´0´´	JN226681
NK 96846	*Phyllotis darwini*	Observatorio La Silla	OS46	29°15´0´´	70°43´48´´	JN226708
NK 96847	*Phyllotis darwini*	Observatorio La Silla	OS47	29°15´0´´	70°43´48´´	JN226709
NK 96852	*Phyllotis darwini*	Observatorio La Silla	OS52	29°15´0´´	70°43´48´´	JN226710
NK 96855	*Phyllotis darwini*	Observatorio La Silla	OS55	29°15´0´´	70°43´48´´	JN226711
NK 96864	*Phyllotis darwini*	Observatorio La Silla	OS64	29°15´0´´	70°43´48´´	JN226712
NK 142983	*Phyllotis darwini*	Pan de Azucar	Paz83	26°8´24´´	70°39´36´´	JN226729
NK	*Phyllotis*	Pan de Azucar	Paz	26°8´2	70°39´3	JN226730

142984	*darwini*		84	4´´	6´´	
NK 96575	*Phyllotis*	Pelambres	PE7	31°49´	70°34´4	JN226700
	darwini		5	12´´	8´´	
NK 96586	*Phyllotis*	Pelambres	PE8	31°49´	70°34´4	JN226701
	darwini		6	12´´	8´´	
NK 96591	*Phyllotis*	Pelambres	PE9	31°49´	70°34´4	JN226702
	darwini		1	12´´	8´´	
NK 96595	*Phyllotis*	Pelambres	PE9	31°49´	70°34´4	JN226703
	darwini		5	12´´	8´´	
NK 96596	*Phyllotis*	Pelambres	PE9	31°49´	70°34´4	JN226704
	darwini		6	12´´	8´´	
NK	*Phyllotis*	Llanos del Challe	LC	28°4´4	71°8´24	JN226721
105381	*darwini*		H81	8´´	´´	
SSUC-	*Phyllotis*	Punta de Choros	PCh	29°15´	71°27´3	JN226678
Ma-00435	*darwini*		35	0´´	6´´	
SSUC-	*Phyllotis*	Punta de Choros	PCh	29°15´	71°27´3	JN226679
Ma-00437	*darwini*		37	0´´	6´´	
SSUC-	*Phyllotis*	Punta de Choros	PCh	29°15´	71°27´3	JN226682
Ma-00445	*darwini*		45	0´´	6´´	
SSUC-	*Phyllotis*	Quebrada de la	PCh	29°15´	71°27´3	JN226680
Ma-00438	*darwini*	Plata	38	0´´	6´´	
NK 96738	*Phyllotis*	Quebrada del Tigre	QT3	32°33´	71°25´4	JN226705
	darwini		8	36´´	8´´	

NK 96754	*Phyllotis darwini*	Quebrada del Tigre	QT54	32°33´36´´	71°25´48´´	JN226706
NK 95968	*Phyllotis darwini*	Rinconada Huechún	RHU68	33°0´36´´	70°48´36´´	JN226693
NK 95973	*Phyllotis darwini*	Rinconada Huechún	RHU73	33°0´36´´	70°48´36´´	JN226694
NK 96312	*Phyllotis darwini*	San Carlos Apoquindo	SA12	33°24´0´´	70°28´48´´	JN226695
NK 96319	*Phyllotis darwini*	San Carlos Apoquindo	SA19	33°24´0´´	70°28´48´´	JN226696
NK 96323	*Phyllotis darwini*	San Carlos Apoquindo	SA23	33°24´0´´	70°28´48´´	JN226697
NK 96349	*Phyllotis darwini*	San Carlos Apoquindo	SA49	33°24´0´´	70°28´48´´	JN226698
NK 96359	*Phyllotis darwini*	San Carlos Apoquindo	SA59	33°24´0´´	70°28´48´´	JN226699
NK 96817	*Phyllotis darwini*	San Carlos Apoquindo	SA17	33°24´0´´	70°28´48´´	JN226707
NK 105329	*Phyllotis darwini*	San Carlos Apoquindo	SA29	33°24´0´´	70°28´48´´	JN226717
SSUC-Ma-00416	*Phyllotis darwini*	Tranque de relaves Barahona	TrBa16	34°4´51´´	70°31´12´´	JN226664
SSUC-	*Phyllotis*	Tranque de relaves	TrB	34°4´5	70°31´1	JN226665

Ma-00417	*darwini*	Barahona	a17	1´´	2´´	
SSUC-	*Phyllotis*	Tranque de relaves	TrB	34°4´5	70°31´1	JN226666
Ma-00420	*darwini*	Barahona	a20	1´´	2´´	
SSUC-	*Phyllotis*	Tranque de relaves	TrB	34°4´5	70°31´1	JN226667
Ma-00421	*darwini*	Barahona	a21	1´´	2´´	
SSUC-	*Phyllotis*	Tranque de relaves	TrB	34°4´5	70°31´1	JN226668
Ma-00422	*darwini*	Barahona	a22	1´´	2´´	
SSUC-	*Phyllotis*	Tranque de relaves	TrB	34°4´5	70°31´1	JN226669
Ma-00423	*darwini*	Barahona	a23	1´´	2´´	
SSUC-	*Phyllotis*	Tranque de relaves	TrB	34°4´5	70°31´1	JN226670
Ma-00424	*darwini*	Barahona	a24	1´´	2´´	
SSUC-	*Phyllotis*	Tranque de relaves	TrB	34°4´5	70°31´1	JN226671
Ma-00425	*darwini*	Barahona	a25	1´´	2´´	

APPENDIX S2.- Fully detailed methodology.

Specimens and localities Sixty eight specimens of *Phyllotis darwini* were analyzed representing 18 localities across central Chile (Fig. 1). The southern distribution of *Phyllotis* was poorly represented because we were unable to capture individuals between 34° S and 36° S. The same is true for the northernmost portion of the range since we did not obtain samples between "Parque Nacional Pan de Azúcar" (26° S) and "Parque Nacional Llanos de Challe" (28°S). Voucher specimens were deposited in the Colección de Flora y Fauna Profesor Patricio Sánchez Reyes (SSUC), Departamento de Ecología, Pontificia Universidad Católica de Chile, Santiago, Chile, and the Museum of Southwestern Biology (MSB), University of New Mexico, Albuquerque, New Mexico. The list of specimens, localities and abbreviations is given in Appendix S1. Tissues and data associated to each specimen are cross-referenced and stored in the collection under a field catalog number: NK is the field catalog used by the SSUC and the MSB; ER is the field catalog of Dr. Enrique Rodriguez-Serrano. We followed the ASM guidelines during the collection and care of the animals used in this work (Sikes & Gannon, 2011)

DNA extraction and sequencing We used frozen liver for DNA extraction using the Wizard Genomic DNA Purification Kit (PROMEGA, Madison, Wisconsin). The DNA extraction in *Phyllotis magister* (used as outgroup) was performed from ethanol preserved ear tissue. We amplified via the Polymerase Chain Reaction (PCR) the Hypervariable domain II (HV2) of the mitochondrial DNA (mtDNA) control region in 72 individuals. We used HV2 instead of the traditionally Hypervariable domain I (HV1) because the substitution rate of the former was enough to solve the evolutionary relationships among *Phyllotis darwini´s* haplotypes. Although the HV1 has a significantly higher variability

than HV2 (measure as nucleotide diversity), the former was about only 10% more variable than HV2 (data not shown). Primers used for PCR were 282 and 283 (Bacigalupe *et al.*, 2004), and the thermal profile was: initial denaturation at 95°C for 7 min, followed by 30 cycles of 94°C (30 s), 59°C (16 s), and 72 °C (1 min 15 s). A final extension followed at 72 °C for 4 min. PCR products were purified with PCR Preps (QIAGEN). Cycle sequencing (Murray 1989) was performed using primer 283 labeled with the Big Dye terminator kit (Perkin Elmer, Norwalk, Connecticut). Sequencing reactions were analyzed on an Applied Biosystems Prism 310 (Foster City, California) automated sequencer. We sequenced a total of 412 base pairs of the mtDNA control region and those sequences have been deposited in Gene Bank (Gene Bank accession numbers JN226664 - JN226735). Sequences were aligned using BIOEDIT (Hall, 1999) and by eye. In addition, the complete sequence of *Auliscomys pictus* mtDNA control region (Gene Bank accession number AF296272) was used as a reference for alignment. Finally, saturation of the molecular marker was evaluated using the Xia test (Xia *et al.*, 2003a) implemented in DAMBE (Xia & Xie, 2001). The assumption of neutrality was tested calculating Tajima's D index (Tajima, 1989) implemented in the DnaSP 4.1 software (Rozas *et al.*, 2003), as well as the nucleotide and haplotype diversity indexes.

Haplotype network and intraspecific phylogeny Haplotype network and demographic analyses were performed over an haplotype file, built in the DnaSP 4.1 software (Rozas *et al.*, 2003). For phylogenetic analyses four specimens of *Phyllotis magister*, the sister species of *P. darwini* (Steppan, 1998), was used as outgroup. The Markov Chain Monte Carlo (MCMC) method within a Bayesian framework (hereafter BMCMC) was used to estimate the posterior probability of phylogenetic trees. The MCMC

procedure ensures that trees are sampled in proportion to their probabilities of occurrence under the model of gene-sequence evolution. Approximately 10,000,000 phylogenetic trees were generated using the BMCMC procedure, sampling every 1000^{th} trees to ensure that successive samples were independent. The first 50 trees of the sample were removed to avoid including trees sampled before convergence of the Markov Chain. The pattern of molecular evolution from the control region in mammals is very complex. In rodents, strong rate heterogeneity among sites has been detected, as well as a variable length and number of tandem repeated elements, even between subspecies. Moreover, the HV2 domain may feature heterogeneous patterns of molecular evolution because it posses three conserved blocks, and it is functionally important given the presence of the replication origin of the H (Heavy) strand (Larizza *et al.*, 2002). Because of this, we used a general likelihood-based mixture model (MM; Pagel & Meade, 2004), based on the general time-reversible (GTR) model of gene-sequence evolution to estimate the likelihood of each tree. This model accommodates cases in which different sites in the alignment evolved in qualitatively distinct ways, but does not require prior knowledge of these patterns or partitioning of the data. These analyses were conducted using the software Bayes Phylogenies, available at the website http://www.evolution.rdg.ac.uk/SoftwareMain.html. In order to find the best mixture model of gene-sequence evolution, we obtained the likelihood of the trees by first using a GTR matrix plus the gamma distributed rate heterogeneity model (1GTR + G) and then continuing to add up to five GTR + G matrices were determined. For the posterior analyses, only the combination of matrices with the fewest number of parameters that significantly increased the likelihood was used. Posterior probabilities for topologies were assessed as the proportion of trees sampled after burn-in, in which that particular topology was observed.

To assess whether the hierarchical relationships between haplotypes (inferred from BMCMC) were consistent with its reticulate associations, and to explicitly assess the geographical pattern and frequency associated with each haplotype, a network of haplotypes was calculated using the median joining algorithm (Bandelt *et al.,* 1999) implemented in NETWORK 4.5 (Rohl & Mihn, 1997).

Population genetic analyses To evaluate the genetic structure within *P. darwini* we identified the populations within the species using the GENELAND software (Guilliot *et al.*, 2005). This approach is a Bayesian cluster analysis that uses individual geo-referenced genetic data to detect the number and geographic position of populations (Guilliot *et al.*, 2008). The algorithm identifies genetic discontinuities while estimates both the number and locations of populations without any a priori knowledge on the populational units and limits. Once the number and limits of populations are established, the population membership probability is calculated from the posterior probability distribution of the MCMC. First, one independent run was performed by 10,000,000 of generations, sampling every 1000 generations of the Markov chain and treating the number of populations as unknown. Then, we choose the better of five independent runs, each of 10,000,000 of generations and sampling every 1,000 but now treating the number of populations as a fix parameter estimated from the first independent run. From the posterior distribution, we draw a map of probability isoclines of population membership, one for each population or cluster inferred by the model.

Once the geographic location of cluster units and phylogenetic relationships was known, we assigned the haplotypes to each of the two major phylogenetic groups according to its geographic location and performed a Mantel test (Mantel, 1967). Across populations

of each clade, we evaluated correlations between the genetic and the geographic distance to test for isolation by distance (Rousset, 1997). The Mantel test was performed in the PopTools software (Hood, 2010); first we computed two matrices for populations of each of the two major clades, one of genetic differentiation index (pairwise Fst), and another of pairwise geographic distances between localities. The frequency distribution of correlation coefficients expected by chance was approximated through randomization of both genetic and geographic distances matrices between the haplotypes, with 10.000 replicates for each matrix. The significance of the correlation between genetic and geographic distances was assessed as the cumulative probability of the correlation coefficients from the random distribution that exceeded the value of the observed correlation coefficient between genetic and geographic distances.

To achieve insights about the demographic history of *Phyllotis darwini* that could explain the genealogical patterns, we evaluated the sudden expansion model in the distribution of pairwise genetic differences (Rogers & Harpending, 1992; Schneider & Excoffier, 1999). This analysis was performed from the haplotypes dataset using an infinite-sites model that took into account multiple substitutions and allow mutation rates to vary through DNA sequence. To compute this Mismatch distribution and test its goodness-of-fit to the sudden expansion model (Schneider & Excoffier, 1999), we used the software ARLEQUIN 3.1 (Schneider *et al.*, 2000). The least-squares deviation method was used as a test of goodness-of-fit (Schneider & Excoffier, 1999).

Clock calibration The age of intraspecific divergence events was estimated in a relaxed molecular clock approach implemented in the software BEAST v.1.7.4 (Drummond *et al.*, 2006). The node ages for the main phylogroups recovered inside *Phyllotis darwini* were co-

estimated from a subsample of the intraspecific phylogeny, but rooted with a D-loop sequence from *Auliscomys pictus* (Gene Bank accession number AF296272). To make a gross estimation, we used the 3.0 -5.1 Myr basal split in *Phyllotis* suggested by (Steppan *et al.*, 2007) and 10% molecular divergence rate estimated for D-loop in rodents (Brown, 1986). The analysis implemented a GTR + G + I model with rate variation (four gamma categories) and a Yule branching rate prior. Rate variation across branches was assumed to be uncorrelated and lognormally distributed (Drummond *et al.*, 2006). The MCMC chain was run for 10 000 000 generations (burn-in 10 000 generations), with parameters sampled every 1000 steps.

Distribution models We modeled the climatic niche of each intraspecific lineage to approximate the whole species' current distribution, and its distribution during the LGM under the assumptions that: (1) climate is an important factor driving the species' distribution; (2) the climatic niche of species remained conserved between the LGM and present time, and (3) overlapped lineage's distribution ranges will approach the whole species geographic range. The latter assumption was tested by overlapping distribution models of each intraspecific lineage in order to approach the full species distributional range, as the sum of ranges estimated for each lineage. The resultant distributional range was roughly contrasted with another model built for the whole species without considering phylogenetic structure.

The climatic niches were reconstructed using the methodology of ecological niche modeling, where environmental data are extracted from occurrence records and random points (represented by geographic coordinates). Habitat suitability was evaluated across the landscape using program specific algorithms (Elith *et al.*, 2006). The current models were

then projected on the climatic reconstructions of the LGM. For occurrence records, we used our unique sampling localities. In addition to full geographic distribution models for the species, we built climatic models for each major lineage recovered in the intraspecific phylogeny following the same approach. As a test of consistency we overlap the intraspecific lineage distribution models, in order to compare it to the full species distribution models.

The current climate was represented by bioclimatic variables from the WorldClim dataset v. 1.4 (http://www.world clim.org/; Hijmans *et al.*, 2005) that are derived from monthly temperature and precipitation data, and represent biologically meaningful aspects of local climate (Waltari *et al.*, 2007; Jezkova *et al.*, 2009).

For environmental layers representing the climatic conditions of the LGM, we used ocean–atmosphere simulations (Harrison, 2000) available through the Paleoclimatic Modeling Intercomparison Project (Braconnot *et al.*, 2007). These reconstructions of the LGM climate are based on simulated changes in concentration of greenhouse gases, ice sheet coverage, insulation and topography (caused by lowering sea levels). We used two models that have been previously downscaled for the purpose of ecological niche modeling (Waltari *et al.*, 2007): Community Climate System Model v. 3 (CCSM; Otto-Bliesner *et al.*, 2006) and the Model for Interdisciplinary Research on Climate v. 3.2 (MIROC; Hasumi & Emori, 2004). The original climatic variables used in these models have been downscaled to the spatial resolution of 2.5 min under the assumption that changes in climate are relatively stable over space (high spatial autocorrelation) and were converted to bioclimatic variables (Peterson & Nyári, 2007).

Climatic niche models were built in the software package MAXENT v. 3.2.1 (Phillips *et al.*, 2006), a program that calculates relative probabilities of the species'

presence in the defined geographic space, with high probabilities indicating suitable environmental conditions for the species (Phillips *et al.*, 2004). Trapping coordinates of each individual captured for DNA extraction were used as presence points. We used the default parameters in MAXENT (500 maximum iterations, convergence threshold of 0.00001, regularization multiplier of 1, and 10 000 background points) with the application of random seed and logistic probabilities for the output (Phillips & Dudik, 2008). We masked our models to four altitudinal categories resuming both, the abrupt altitudinal clines characteristic of central Chile, and some known altitudinal distribution limits for several vertebrate taxa in this area (Fuentes & Jaksic, 1979). This procedure was conducted because reducing the climatic variation being modeled to that which exists within a geographically realistic area improves model accuracy and reduces problems with extrapolation (Pearson *et al.*, 2002; Thuiller *et al.*, 2004; Randin *et al.*, 2006). We ran 10 replicates for each model, and an average model was presented using logistic probability classes of climatic niche suitability. The presence– absence map was determined using the 'maximum training sensitivity plus specificity logistic threshold' where the omission error of all occurrence records is set to zero (i.e., locations of all occurrence records are predicted as 'suitable'). Inside the suitability area determined by the threshold, we show the complete distribution of logistic probability values, in order to investigate if highest values were associated to the major features of the geographically complex scenario in central Chile (Andes, Coastal Cordillera, valley and coast). We used the receiver operating characteristic for its area under the curve (AUC) value to evaluate the model performance (Fielding & Bell, 1997; Raes & ter Steege, 2007). AUC values ranged from 0.5 for a random prediction to 1 for perfect prediction (Phillips *et al.*, 2004).