Ecology, Genetic Diversity, and Phylogeographic Structure of Andes Virus in Humans and Rodents in Chile[∇]

Rafael A. Medina,¹ #‡ Fernando Torres-Perez,^{1,2,3} # Hector Galeno,⁴ Maritza Navarrete,⁵ Pablo A. Vial,⁶
R. Eduardo Palma,^{3,7} Marcela Ferres,⁸ Joseph A. Cook,² and Brian Hjelle^{1,2,9*}

Center for Infectious Diseases and Immunity, Department of Pathology, University of New Mexico Health Sciences Center, Albuquerque, New Mexico 87131¹; Department of Biology, University of New Mexico, Albuquerque, New Mexico 87131²; Center for Advanced Studies in Ecology and Biodiversity, Pontificia Universidad Católica de Chile, Santiago, Chile³; Public Health Institute of Chile, Santiago, Chile⁴; Clínica Hospital of Valdivia, School of Medicine, Universidad Austral de Chile, Valdivia, Chile⁵; Institute of Biomedical Sciences, Clínica Alemana School of Medicine, Universidad del Desarrollo, Santiago, Chile⁶; Department of Ecology, Pontificia Universidad Católica de Chile, Santiago, Chile⁷; Department of Pediatrics and Virology Laboratory, Pontificia Universidad Católica de Chile, Santiago, Chile⁸; and Department of Molecular Genetics and Microbiology, University of New Mexico, Health Sciences Center, Albuquerque, New Mexico 87131⁹

Received 20 May 2008/Accepted 20 December 2008

Andes virus (ANDV) is the predominant etiologic agent of hantavirus cardiopulmonary syndrome (HCPS) in southern South America. In Chile, serologically confirmed human hantavirus infections have occurred throughout a wide latitudinal distribution extending from the regions of Valparaíso (32 to 33°S) to Aysén (46°S) in southern Patagonia. In this study, we found seropositive rodents further north in the Coquimbo region (30°S) in Chile. Rodent seroprevalence was 1.4%, with *Oligoryzomys longicaudatus* displaying the highest seroprevalence (5.9%), followed by *Abrothrix longipilis* (1.9%) and other species exhibiting ≤0.6% seropositivity. We sequenced partial ANDV small (S) segment RNA from 6 HCPS patients and 32 rodents of four different species collected throughout the known range of hantavirus infection in Chile. Phylogenetic analyses showed two major ANDV South (ANDV Sout) clades, congruent with two major Chilean ecoregions, Mediterranean (Chilean matorral [shrubland]) and Valdivian temperate forest. Human and rodent samples grouped according to geographic location. Phylogenetic comparative analyses of portions of S and medium segments (encoding glycoproteins Gn and Gc) from a subset of rodent specimens exhibited similar topologies, corroborating two major ANDV Sout clades in Chile and suggesting that yet unknown factors influence viral gene flow and persistence throughout the two Chilean ecoregions. Genetic algorithms for recombination detection identified recombination events within the S segment. Molecular demographic analyses showed that the virus is undergoing purifying selection and demonstrated a recent exponential growth in the effective number of ANDV Sout infections in Chile that correlates with the increased number of human cases reported. Although we determined virus sequences from four rodent species, our results confirmed *O. longicaudatus* as the primary ANDV Sout reservoir in Chile. While evidence of geographic differentiation exists, a single cosmopolitan lineage of ANDV Sout remains the sole etiologic agent for HCPS in Chile.

Hantaviruses are segmented negative-strand RNA viruses of the family *Bunyaviridae*; these viruses contain small (S), medium (M), and large (L) segments, encoding nucleocapsid (N), glycoproteins Gn and Gc, and RNA-dependent RNA polymerase, respectively (28). Although at least 24 named members of the genus *Hantavirus* have been associated with hantavirus cardiopulmonary syndrome (HCPS) or hemorrhagic fever with renal syndrome, it is likely that some of the recently named viruses are not distinct from hantaviruses described previously. Of the 16 or more named hantaviruses that have been linked to HCPS in the Western Hemisphere, Sin Nombre virus (SNV) and Andes virus (ANDV) represent the most important etio-

logic agents in North and South America, respectively (54). These viruses, like all pathogenic hantaviruses, are carried by wild rodent hosts and are transmitted to humans via accidental exposure to infected rodent excreta and secretions (9, 42).

Diverse hantavirus genotypes have been reported to occur throughout a vast geographical range in southern South America (8, 35); these include five different lineages of ANDV that have been proposed and named according to their geographic origin in Argentina and Chile (43). These viral genotypes include the following: ANDV Central Plata (ANDV Cent Plata) found on both sides of the Rio de la Plata (in Uruguay and Argentina) and central Buenos Aires province in Argentina; ANDV Central Buenos Aires (ANDV Cent Bs.As.) found in numerous localities of Buenos Aires province; ANDV Central Lechiguanas (ANDV Cent Lec) detected in the Entre Rios province in central Argentina; ANDV North (ANDV Nort) identified in the northern Argentine provinces of Jujuy, Salta, and Oran; and ANDV South (ANDV Sout), which has been identified throughout Chile and in the southern Andean region of Argentina (35, 43). Overall, ANDV may differ from other hantaviruses in that person-to-person transmission has been

* Corresponding author. Mailing address: Center for Infectious Diseases and Immunity, Department of Pathology, Health Sciences Center, MSC08 4640, University of New Mexico, Albuquerque, NM 87131. Phone: (505) 272-0624. Fax: (505) 272-4401. E-mail: bhjelle@salud.unm.edu.

R.A.M. and F.T.-P. contributed equally to this study.

‡ Present address: Department of Microbiology, Mount Sinai School of Medicine, New York, NY 10029.

[∇] Published ahead of print on 30 December 2008.

well documented in at least nine cases in Argentina and Chile, including in nosocomial settings (18, 38, 44).

In Chile, HCPS cases have occurred across a wide geographic range, from the Valparaíso region (32° to 33°S) in the north to the Aysén region (46°S) in southern Patagonia. After a few sporadic cases were recognized in 1995 and 1996 in the southern regions, HCPS struck Chile in an outbreak of 25 cases in the southern regions of Aysén and Los Lagos near Coyhaique in the summer of 1997 to 1998 (62). By the summer of 2000 to 2001, cases began to appear in the highly populated metropolitan region near Santiago, Chile. Between 2000 and 2007, a total of 439 HCPS cases were reported across nine regions in Chile (<http://epi.minsal.cl/>). The increased vigilance and greater physician recognition of the symptoms of HCPS likely played some role in the increase in recorded cases. Human cases have been identified and diagnosed every year since the disease was first reported in Chile.

ANDV was identified as the etiologic agent of the first recognized outbreak of hantavirus disease in southern Chile and is regarded as the major, if not the sole, etiologic agent of hantavirus cardiopulmonary syndrome in Chile (62). The predominant reservoir species for ANDV is *Oligoryzomys longicaudatus* (the long-tailed pigmy rice rat or colilargo) (35). Several other rodent species have been found to be seropositive in natural settings in Chile (42, 62); however, whether these rodents play a role in the natural transmission cycle or in transmission to humans remains unclear.

Little information is available about the genetic diversity and specific geographic distribution of hantaviruses in Chile. The number of rodent species involved in the enzootic and the degree of genetic differentiation of hantaviruses among humans and rodents in different regions of the country has not been fully assessed (63). In this study, we sought to determine whether hantaviruses other than ANDV and host reservoir species other than *O. longicaudatus* might be contributing to the caseload of HCPS in Chile. We found that ANDV in Chile consists largely of a single geographically structured lineage found throughout the range of the host. This lineage corresponds to the ANDV Sout lineage (43); however, within this lineage we identified two distinct clades that correspond to ecogeographic regions. Phylogenetic topologies showed no evidence of reassortment between the S and M segments; nonetheless, we identified recombination events within the S segment. The ANDV Sout population is under purifying selection in nature. A diverse group of rodents can be infected with ANDV Sout, but *O. longicaudatus* exhibits the highest seroprevalence. Due to the commensal relationship between this species and humans in agrarian communities, it is the primary host of epidemiologic significance to human infections.

MATERIALS AND METHODS

Human samples. All patients examined were officially enrolled in a study that was approved by the human institutional review boards of the University of New Mexico and Pontificia Universidad Católica de Chile. Patients with symptoms of HCPS were considered to have acute ANDV infection on the basis of the following serological criteria: the presence of immunoglobulin M antibodies against the viral N antigen, with or without specific immunoglobulin G antibodies. We identified patients through a national passive surveillance/diagnostic referral system administered through the Institute of Public Health in Santiago, Chile (18). We randomly selected patient blood or tissue samples for further study according to availability and geographic diversity.

TABLE 1. Primers used for nested and heminested RT-PCR and sequencing

Primer ^a	Segment	Sequence (5' to 3')
S35+	Small	AAGCTGGAATGAGCACCTCCAAGA
S184+	Small	CGGGCAGCTGTGTCTACATTGGA
S495+	Small	CAACAAAGGGACCAGGATAAGGT
S626– ^b	Small	GGCGTGATTTCTTCAGCCTTCAT
S1185–	Small	GATTATGATCTTCTGGTCCAGTTGG
S1283– ^b	Small	TCGATCAGAGATTGTGCTAGCTG
Gn1+	Medium	TAGTAGTAGACTCCGCAAGAAGAA G
Gn540–	Medium	AGCCATACAACCTCTTATTGACAT
Gn559– ^b	Medium	TCCTGCT(G/T)(G/T)TAAACACACTAG CCAT
Gc2488+	Medium	GGAACAGAACAAACATGCAAGC
Gc2517+ ^b	Medium	TGCAAATGATTGTGTAGTAACACCA
Gc3143–	Medium	GTATTAGAGCCCCTAGCACAGGTT
Gc3171– ^b	Medium	TGACCTCCCTTACCAACCACTT

^a Primer coordinates incorporated into the primer name are according to the ANDV S- and M-segment genomic sequence.

^b Primers used for the RT step.

Rodent samples. We collected rodents in live traps according to standard protocols, as previously described (63). We followed established safety guidelines for rodent capture and processing (39). Trapping grids were set up at sites adjacent to case households and in locations chosen to provide geographic diversity. Animals in the field were anesthetized using ketamine, blood was drawn from the retro-orbital sinus, and the rodents were then sacrificed via cervical dislocation. We necropsied rodents on-site and collected and preserved hearts, kidneys, spleens, livers, and lungs in liquid nitrogen. A subset of rodent tissues was obtained through a national survey overseen by the Chilean Institute of Public Health.

Rodent antibody screening. Rodent antibody against ANDV was detected in blood samples using a strip immunoblot assay as previously described (63, 66). We used positive serology to identify individuals for viral sequence determination.

RT-PCR. Heminested reverse transcription-PCR (RT-PCR) was conducted essentially as described previously (21). Typically, 50 to 100 mg of heart or lung tissue or 100 µl of serum was used to extract RNA using an RNeasy minikit or a QIAmp viral RNA minikit (Qiagen Inc., Valencia, CA), respectively. We optimized S- and M-segment primers (Table 1) according to ANDV sequences ascertained through GenBank. We used the SuperScript II (Invitrogen, Carlsbad, CA) enzyme and 2.5 µM of an antisense primer to reverse transcribe ~1 µg of total RNA or 5 µl of RNA derived from serum. For the S segment, the antisense primer (coordinate S 626) was used to carry out the RT step, followed by an outer PCR using 5 µl of cDNA and 100 µM of the sense (coordinate S 35) and antisense S626 primers in a final volume of 50 µl. We used either a *Taq* gold polymerase (Applied Biosystems, Foster City, CA) or the *Taq* DNA polymerase (Roche, Indianapolis, IN), and reaction mixtures were run in a PE Biosystems model 9700 cyclor under the following conditions: 5 min at 94°C; 35 cycles, with 1 cycle consisting of 10 s at 94°C, 10 s at 48°C, and 60 s at 72°C; followed by a final extension step (5 min at 72°C) and a soak step at 4°C. Five microliters of each outer product was then subjected to 94°C for 5 min and 35 cycles (with 1 cycle consisting of 10 s at 94°C, 30 s at 50°C, and 60 s at 72°C), with a final extension step (5 min at 72°C) and a soak step at 4°C using 100 µM of the inner heminested primers S184+ and S626– in a final volume of 50 µl. The products were agarose gel purified and directly sequenced in both directions. We increased the length of S-segment sequences (to 943 nucleotides [nt]) in a subset of *O. longicaudatus* samples (pruned data set, 14 individuals). These samples were selected as representative samples to cover the geographic range of ANDV-seropositive individuals in Chile (Table 2). We amplified and sequenced two portions of the M segment, referred to as Gn (490 nt; coordinates 26 to 516) and Gc (577 nt; coordinates 2543 to 3119). The Gn gene was amplified by a heminested RT-PCR, using primer Gn559– for the RT step as described above and primers Gn1+ and Gn559– for the outer reaction and primers Gn1+ and Gn540– for the inner reaction. PCR conditions were similar except that the annealing temperature and time in the outer reaction were set at 48°C and 30 s, respectively. We used a nested RT-PCR to amplify the M-Gc segment using primer Gc3171– for the RT step, and primers Gc2488+ and Gc3171– for the outer PCR. Primers Gc2517+ and Gc3143– were used for the inner reaction. The conditions for the outer

TABLE 2. Detailed information on the samples used in this study including hantavirus species, source of the viral RNA (human or rodent), location, map number, allele letters in Fig. 1 and 2, and GenBank accession numbers

Hantavirus species	Source (human, rodent species)	Code or museum no.	Site	Administrative region in Chile	Country	Map no.	S-segment allele letters in Fig. 1	Virus sequence abbreviation in Table 4	GenBank accession number for:		
									S segment	Glycoprotein Gn	Glycoprotein Gc
Samples from this study											
Andes	<i>Oligoryzomys longicaudatus</i>	NK 96865	Fray Jorge	Coquimbo	Chile	1	MA	FJ865	EU241689	EU241637	EU241651
Andes	<i>Oligoryzomys longicaudatus</i>	NK 96887	Fray Jorge	Coquimbo	Chile	1	MB	FJ887	EU241690	EU241638	EU241652
Andes	<i>Oligoryzomys longicaudatus</i>	NK 108859	Los Vilos	Coquimbo	Chile	2	MC		EU241665		
Andes	<i>Oligoryzomys longicaudatus</i>	NK 104619	San Carlos de Apoquindo	Metropolitana	Chile	3	MD	SCA619	EU241691	EU241639	EU241653
Andes	<i>Oligoryzomys longicaudatus</i>	NK 105110	San Carlos de Apoquindo	Metropolitana	Chile	3	MD	SCA110	EU241667		
Andes	<i>Oligoryzomys longicaudatus</i>	NK 96325	San Carlos de Apoquindo	Metropolitana	Chile	3	MD	SCA325	EU241666		
Andes	Human	AND1134S	Lo Herrera	Metropolitana	Chile	4	MG		EU241669		
Andes	Human	AND1133S	Peñaflor	Metropolitana	Chile	5	MF		EU241668		
Andes	<i>Rattus rattus</i>	Rr55	Melipilla	Metropolitana	Chile	6	MH		EU241670		
Andes	<i>Oligoryzomys longicaudatus</i>	OL60S	Quebrada de la Plata	Metropolitana	Chile	7	MI		EU241671		
Andes	<i>Oligoryzomys longicaudatus</i>	OL269S	Picarquin	O'Higgins	Chile	8	ML		EU241674		
Andes	<i>Oligoryzomys longicaudatus</i>	NK 105929	San Fernando	O'Higgins	Chile	9	MK		EU241673		
Andes	Human	HNancagua	Nancagua	O'Higgins	Chile	10	MJ		EU241672		
Andes	<i>Oligoryzomys longicaudatus</i>	NK 120092	Curicó	Maule	Chile	11	MN	CUR092	EU241692	EU241640	
Andes	<i>Oligoryzomys longicaudatus</i>	NK 108501	Duao	Maule	Chile	12	MO		EU241675		
Andes	<i>Oligoryzomys longicaudatus</i>	NK 108505	Duao	Maule	Chile	12	MO		EU241676		
Andes	<i>Oligoryzomys longicaudatus</i>	NK 108843	Los Ruites	Maule	Chile	13	MM	LR843	EU241693	EU241641	EU241655
Andes	Human	HPWS	Chillán	Concepción	Chile	14	TB		EU241678		
Andes	Human	HChillanS	Chillán	Concepción	Chile	14	TA		EU241677		
Andes	<i>Oligoryzomys longicaudatus</i>	NK 120076	Tomé	Concepción	Chile	15	TC		EU241680		
Andes	Human	H792	Concepción	Concepción	Chile	16	TA	TUC979	EU241679	EU241642	EU241656
Andes	<i>Oligoryzomys longicaudatus</i>	NK 105979	Tucapel	Concepción	Chile	17	TD		EU241694		
Andes	<i>Oligoryzomys longicaudatus</i>	OL25S	Quilleco	Concepción	Chile	18	TE		EU241681		
Andes	<i>Oligoryzomys longicaudatus</i>	OL26S	Quilleco	Concepción	Chile	18	TF		EU241682		
Andes	<i>Oligoryzomys longicaudatus</i>	NK 96703	Temuco	Temuco	Chile	19	TG	TEM703	EU241697	EU241643	EU241657
Andes	<i>Oligoryzomys longicaudatus</i>	NK 105362	Parque Nacional Huerquehue	Temuco	Chile	20	TI	PNH362	EU241695	EU241645	EU241659
Andes	<i>Oligoryzomys longicaudatus</i>	NK 95083	Parque Nacional Villarrica	Temuco	Chile	21	TJ	PNV083	EU241696	EU241644	EU241658
Andes	<i>Oligoryzomys longicaudatus</i>	NK 95613	Carahue	Temuco	Chile	22	TH		EU241683		
Andes	<i>Oligoryzomys longicaudatus</i>	NK 104535	Panguipulli	Los Ríos	Chile	23	TK	PAN535	EU241698	EU241646	EU241660
Andes	<i>Oligoryzomys longicaudatus</i>	NK 104544	Panguipulli	Los Ríos	Chile	23	TI	PAN544	EU241699	EU241647	EU241661
Andes	<i>Abrothrix longipilis</i>	AL16	Neltume	Los Ríos	Chile	24	TL		EU241684		
Andes	<i>Oligoryzomys longicaudatus</i>	NK 96555	Reserva Nacional Río Simpson	Aysén	Chile	25	TQ	RNR5555	EU241700	EU241648	EU241662
Andes	<i>Oligoryzomys longicaudatus</i>	NK 96566	Reserva Nacional Río Simpson	Aysén	Chile	25	TR	RNR5566	EU241701	EU241649	EU241663
Andes	<i>Oligoryzomys longicaudatus</i>	OL178S	Coyhaique	Aysén	Chile	26	TS		EU241686		
Andes	<i>Abrothrix longipilis</i>	AL218	Coyhaique	Aysén	Chile	26	TS		EU241687		
Andes	<i>Loxodontomys micropus</i>	LM859	Coyhaique	Aysén	Chile	26	TT		EU241688		
Andes	<i>Oligoryzomys longicaudatus</i>	NK 120378	Chile Chico	Aysén	Chile	27	TO	CHC378	EU241702	EU241650	EU241664
Andes	<i>Oligoryzomys longicaudatus</i>	NK 120387	Chile Chico	Aysén	Chile	27	TP	CHC387	EU241685

FIG. 1. Phylogenetic relationship and diversity of ANDV S-segment alleles in Chile. (A) Map of the sampled localities (numbers from Table 2) of the Andes virus in Chile and in the East Andes in Argentina). Black circles represent samples sequenced in this study, and black squares are samples obtained from GenBank. (B) Maximum likelihood tree obtained from aligned sequences (397 nt) of the S segment. Numbers above branches indicate pseudoreplicate bootstrap values obtained by maximum likelihood analysis (the first value [outside parentheses]) and maximum parsimony analysis (value in parentheses). Numbers below branches indicate posterior probabilities for Bayesian inference. The bar indicates 0.1 nucleotide substitution/site. Taxa shown in italic type are representatives of previously reported ANDV lineages as described in Table 2 and in Results. For further details, refer to Materials and Methods.

TreePuzzle software (55) using the Dayhoff plus amino acid frequencies (F) model as calculated by ProtTest (1). We performed a phylogenetic analysis with concatenated sequences of the pruned data set ($n = 21$) that included sequences of the S, M-Gn, and M-Gc (1,668-nt) segments. Given that potential recombination was detected in the S segment (see Results), we used S-segment partition 2 (600 nt) and all Gn and Gc sequences for the concatenated analysis. Missing data were added for Curicó, Chile, in the Gc sequence. To estimate the PP of the phylogenetic trees, we used the Bayesian Markov chain Monte Carlo sampling procedure using BayesPhylogenies (45), which allows the combination of multi-

ple genes and models into one data set without a priori partitions. We used a general likelihood-based “mixture model” based on the GTR of gene sequence evolution to estimate the likelihood of each tree. We determined the likelihood of the trees by first using a simple GTR matrix (nQ) and then using $nQ + \Gamma$ mixture models where n varied between one and six independent rate matrices (Q s) and Γ represents a gamma-distributed rate heterogeneity model. A 50% majority rule consensus tree was constructed.

Out-group choice. For phylogenetic analyses, we included multiple prototypic viruses containing sequences overlapping with those of the S-segment por-

TABLE 3. ANDV seroprevalence in wild-caught small mammals in Chile

Species or parameter	No. of small mammals captured ^a	No. of seropositive animals/total no. of animals (%)			Overall seroprevalence (%)
		Females	Males	Undetermined sex	
<i>Oligoryzomys longicaudatus</i>	1,259	12/627 (1.91)	63/632 (9.97)		5.9
<i>Abrothrix longipilis</i>	1,269	5/544 (0.92)	19/725 (2.62)		1.9
<i>Abrothrix olivaceus</i>	1,930	1/786 (0.1)	1/1144 (0.13)		0.1
<i>Phyllotis darwini</i>	566	1/257 (0.39)	0/309 (0)		0.2
<i>Rattus rattus</i>	178	0/82 (0)	1/96 (1.04)		0.6
Other ^b					
Sex determined	1,967	0/985	0/982		0.0
Sex not determined	61			0/61	0.0
Total no. of small mammals captured	7,230	19/3281 (0.58) ^c	84/3888 (2.16) ^c		1.4 ^d

^a Small mammals trapped in Chile from 2000 to 2006.

^b Other small mammal species captured.

^c These values (number of seropositive animals and percentages) do not include animals with undetermined sex.

^d The overall percent seroprevalence calculation includes animals with undetermined sex.

tion sequenced in this study (Table 2). Trees were rooted using the SNV-S sequence. For the pruned data set, we selected out-groups that had previously been reported to be closely related to the in-group (ANDV) that had available sequences of S and M segments in GenBank (Table 2).

Molecular diversity and demographic analyses. We evaluated the molecular variability of the grouped S-segment sequences (397 nt) in the two primary clades (Fig. 1) using DnaSp 4.0 (52). We assessed population equilibrium of the two clades by performing Tajima's *D* test (60), Fu's *F_s* test (19), and *R₂* test (49), testing the significance of the statistics from 10,000 simulated samples. We also examined the population dynamics for the ANDV S segment using the Bayesian skyline plot (BSP) (16) as implemented in the BEAST program (15). This method is a nonparametric coalescence analysis and uses standard Markov chain Monte Carlo sampling procedures to estimate past population dynamics from the posterior distribution of the effective population size (*N_et*) (14), with no a priori assumption of a specific demographic model. *N_et* can be thought as the number of ANDV infections contributing to new infections. Analyses were performed comparing uniform rates across branches (strict clock) and uncorrelated relaxed clock assumptions and using the model of sequence evolution as derived from the Modeltest program. Analyses were performed using a fixed mean substitution rate of 2.84×10^{-3} recently reported for the S segment of hantaviruses (50), and the results were plotted together with the number of confirmed human cases as reported yearly by the Chilean Ministry of Health (<http://epi.minsal.cl/>).

We calculated the mean number of nonsynonymous (*dN*) to synonymous (*dS*) nucleotide substitutions within each clade by using the Nei-Gojobori method implemented in the PAML 3.15 program (65). Because only a few sites may be under selection and to compensate for any lack of statistical power of this approach for detecting selection (32), we also included three complementary approaches (single-likelihood ancestor, random effects likelihood, and fixed effects likelihood), which estimate the rates of nonsynonymous and synonymous changes at each site in a sequence for detecting sites under selection (32). In addition, the partitioning approach for robust inference of selection (53), a maximum likelihood method that takes into account recombination and synonymous rate variation, was used to infer whether a proportion of sites in the alignment evolve with a *dN/dS* ratio of >1 . Analyses were performed with various levels of significance, and the methods are implemented in the DataMonkey web-based gateway (30), which executes the molecular evolution analysis platform HyPhy (31).

Nucleotide sequence accession numbers. The sequences from all the ANDV Sout S and M segments described in this study have been deposited in GenBank and have been assigned accession numbers EU241637 to EU241653 and EU241655 to EU241702 (for details, refer to Table 2). Other previously published sequences used in the study were obtained from GenBank (Table 2).

RESULTS

Rodent seroprevalence and study design. During the years from 2000 to 2006, 7,230 rodents and other small mammals

were captured throughout Chile as part of an NIAID-funded International Collaboration in Infectious Diseases project. Tissue samples from 3,304 wild rodents were frozen. The sampling scheme involved much of the range occupied by *Oligoryzomys longicaudatus* from Carrizal Bajo in the northern desert (28°14'S) to as far south as Torres del Paine National Park in southern Patagonia (51°S) in Chile. Of the total number of specimens analyzed, 103 rodents were seropositive for ANDV by using a strip immunoblot assay, with an overall seroprevalence of 1.4% (Table 3). *O. longicaudatus* displayed the highest seroprevalence (5.9%), followed by *Abrothrix longipilis* (1.9%). In both species, males had significantly higher seroprevalence rates than females (9.9% versus 1.91%, respectively, for *O. longicaudatus* and 2.6% versus 0.92%, respectively, for *A. longipilis*). Other rodent species showed much lower seroprevalence ($\leq 0.6\%$), with only one or two animals testing positive for ANDV (Table 3). Of all the seropositive animals, tissue samples were available for 56 rodents from which viral RNA isolation was attempted. We were able to amplify and obtain viral sequences from 23 individuals (Table 2). Fifteen additional specimens were also amplified and sequenced from human and rodent samples obtained through the Public Health Institute of Chile (Table 2). Sequences were obtained from four different rodent species: *O. longicaudatus*, *Rattus rattus*, *Abrothrix longipilis*, and *Loxodontomys micropus*.

ANDV in Chile consists of a single cosmopolitan lineage. To assess phylogenetic relationships of ANDV in Chile, we used 38 viral sequences from human and rodent samples comprising the entire geographic distribution of HCPS cases and throughout the natural habitat of the colilargo in Chile (Fig. 1A). We found no evidence of recombination within the S-segment (397-nt) data set. ML, MP, and Bayesian inferences recovered all viral alleles in a monophyletic group, highly supported by Bayesian inference (PP = 1.00), but weakly supported by maximum likelihood bootstrap (MLB = 52) and maximum parsimony bootstrap (MPB = 57) values (Fig. 1B). All Chilean samples and the western Argentinean sample (El Bolsón, allele TN) located near the Andes mountains, were included in the ANDV Sout clade (in-group). Hantavirus samples from central, northern, and eastern Argentina were placed as a basal

FIG. 2. Unrooted network of ANDV Sout S segment using 397 nt with alleles depicted according to geographical distribution. The size of the circles represents the number of individuals per allele, and the designations represent alleles (Table 2). White and gray symbols represent alleles from the Chilean Mediterranean and temperate forest ecoregions, respectively. The origins of the samples are indicated as follows: squares, humans; circles, *O. longicaudatus*; rhombus, *A. longipilis*; pentagon, *L. micropus*; triangle, *R. rattus*. The circle and rhombus indicate that an allele was shared between *O. longicaudatus* and *A. longipilis*.

group outside the ANDV Sout clade, with Lechiguanas (ANDV Cent Lec), Hu39694 (ANDV Cent Bs.As.), and Oran (ANDV Nort) samples grouped in a supported clade (PP = 1.00; Fig. 1B). Within the in-group, two groups largely corresponded to two major ecogeographical regions in Chile, the Chilean matorral (shrubland; Mediterranean clade) and the Valdivian temperate forest (Temperate Forest clade). The Temperate Forest clade was well supported (PP = 0.98), while there was weak support for the Mediterranean clade (Fig. 1B). Within the Temperate Forest clade, samples were further subdivided into southern and northern Temperate Forest clades. Human samples and those of the four rodent species were placed in trees according to their ecogeographical origin. *Abrothrix longipilis* alleles TL and TS and *L. micropus* (allele TT)

were placed in the Temperate Forest clade, while *R. rattus* (allele MH) was placed in the Mediterranean clade. Human samples were placed in either the Mediterranean (alleles MF, MG, and MJ) or Temperate Forest (alleles TA and TB) clades.

The unrooted network resulting from median-joining analysis (Fig. 2) is concordant with the tree resulting from maximum likelihood analysis (Fig. 1B) showing two major groups (separated by 25 mutations), with alleles associated by their ecogeographical distribution in either Mediterranean or temperate forest. Within the temperate forest group, samples from the south and north were further separated by 13 mutations (Fig. 2). Together, these results suggest that at least two clades occur throughout the geographical distribution of ANDV Sout in Chile.

Phylogeographic relationships of ANDV in Chile. To begin to test for the presence of reassortant viruses and to better establish the phylogeographic relationships of ANDV in Chile, we compared the phylogenies of the M-Gn, M-Gc, and S segments obtained from a subset of 14 colilargos (Fig. 3). For the S segment, we used 943-nt sequences to perform sequence identity analysis. Not surprisingly, Chilean ANDV Sout S-segment sequences showed a high level of similarity that ranged from 96.8 to 100% at the amino acid level and 84.9 to 100% at the nucleotide level (Table 4). Identity ranged from 91.4 to 100% (amino acids) and 83 to 99.7% (nucleotides) for Gn and from 97.3 to 100% (amino acids) and 85.5 to 100% (nucleotides) for Gc (data not shown). Analyses of the S segment for potential recombination events showed evidence of discordant phylogenetic signals and detected a putative recombination breakpoint between nucleotide positions 334 and 342. Hence, S-segment phylogenetic inferences were performed independently with bipartitioned sequences using nucleotides 1 to 333 (partition 1) and nucleotides 343 to 942 (partition 2) of our sequence alignment. To determine presumptive recombinant sequences, we conducted analyses that excluded one sequence in every run. The recombination signal disappeared when one of three sequences were excluded: Tucapel, Los Angeles, or Panguipulli (sample 104535) (all of these were components of the Temperate Forest clade). By contrast, no evidence of recombination was evident for the Gn and Gc sequences.

When rooted with Laguna Negra virus, all phylogenetic analyses recovered ANDV Sout as monophyletic and with high support (Fig. 3). Within the ANDV Sout clade, the topology obtained using S-segment partition 1 recovered Mediterranean samples basal to the monophyletic Temperate Forest clade (Fig. 3A); S-segment partition 2 yielded two reciprocally monophyletic clades with high support, corresponding to the Mediterranean clade (MLB = 88, MPB = 98, and PP = 1.99) and the Temperate Forest clade observed previously (MLB = 54, MPB = 98, and PP = 0.88; Fig. 3B). When using the amino acid sequence of the N protein, the clades (Mediterranean and Temperate Forest) were less well supported in phylogenetic inferences (data not shown). Phylogenies using the Gn (Fig. 3C) and Gc (Fig. 3D) nucleotide sequences showed the same pattern by recovering two main groups within the ANDV Sout clade. The Mediterranean clade was supported using the Gn (MLB = 100, MPB = 99, and PP = 1.00) and Gc (MLB = 88, MPB = 95, and PP = 1.00) sequences. The Temperate Forest clade was supported by the Gc (MLB = 65, MPB = 65, and PP = 0.96) and Gn sequences using maximum parsimony (MPB = 99) and Bayesian inference (PP = 0.99).

Bayesian inference of the concatenated sequences of Gn, Gc, and S portions (1,668 nt) recovered both the Temperate Forest and Mediterranean clades with high posterior probability values (PP = 1.00; Fig. 4). Within the Mediterranean clade, two supported groups were recovered, representing northern (PP = 1.0) and southern (PP = 0.96) Mediterranean subgroups. Samples from the northern temperate forest ecoregion are derived, while the Chile Chico sample is basal (but poorly supported) to all samples of this clade (Fig. 4). Overall, these data confirm that at least two primary clades of ANDV Sout exist throughout the geographical range of this virus in Chile.

Molecular diversity and demographic analyses of ANDV S segment in Chile. Estimations of allele and gene diversity and

results of Tajima's D test, Fu's F_S test, and R_2 tests are summarized in Table 5. Analyses were performed by grouping S-segment samples into "Mediterranean" and "temperate forest" groups. Overall, Fu's F_S test, Tajima's D test, and R_2 test were nonsignificant for both groups, as well as for the complete data set, indicating that the hypothesis of stable demographic history cannot be rejected in favor of a recent expansion. Nucleotide diversity was similar among both groups, with 0.056 (± 0.003) for the Mediterranean clade and 0.048 (± 0.004) for the Temperate Forest clades (Table 5). BSP analysis (Fig. 5) revealed, however, that the effective number of infections through time increased exponentially from 1999, reaching a peak during the years 2001 to 2002, after which the growth rate decreased. When the total number of human HCPS cases per year are superimposed on the BSP (data from the website <http://epi.minsal.cl/>), we observed a proportional relationship among the number of HCPS cases reported and the number of effective ANDV infections through time, with a ~ 1 -year delay observed in the latter (Fig. 5).

Within ANDV Sout, the dN/dS ratios for the S-segment sequences using PAML were $\ll 1.000$ in the two clades, with a mean dN/dS ratio of 0.003 in the Mediterranean and temperate forest sequences (Table 5). These results were congruent with three complementary approaches (single-likelihood ancestor counting, random effects likelihood, and fixed effects likelihood) and the partitioning approach for robust inference of selection method (Table 5), which showed no evidence for positive selection, indicating that most substitutions were silent, and likely reflect strong purifying selection acting over the S-segment gene.

DISCUSSION

Since the discovery of Sin Nombre hantavirus in 1993, studies in the New World have mainly focused on screening other potential hantavirus hosts, describing new strains, and establishing the phylogenetic relationships of newly described strains to other known strains. Finer-scale studies of geographic variation within strains (phylogeography) are valuable for elucidating population genetic processes in one or more related viral strains (10, 11, 24, 57) and inferring evolutionary dynamics of associated hosts (6). Herein, we demonstrate that ANDV is widespread in Chile and thus far appears to be the sole etiologic agent for human hantavirus disease in Chile. ANDV S-segment sequences from humans and four rodent species (*O. longicaudatus*, *R. rattus*, *A. longipilis*, and *L. micropus*) in Chile that exhibited antibodies against ANDV N antigen correspond to the prototypical Argentinean and Chilean virus first described in 1996, ANDV Sout (36, 37). We believe that this virus should be regarded a cosmopolitan strain because of its unusually widespread distribution and moderate genetic divergence.

Previous studies have reported *O. longicaudatus* as the main reservoir of ANDV Sout in Argentina and Chile based on small-scale serology and sequence analyses (42, 62, 63). The drastically higher seroprevalence observed in *O. longicaudatus* (5.9%) compared to other rodent species ($\leq 1.9\%$) and our greater ability to detect viral RNA from *O. longicaudatus* further confirm this rodent species as the main reservoir of ANDV Sout in Chile. Because PCR primers used to detect

viral RNA are not optimized to amplify unexpected “spillover” viruses (23), cross-infection by distinct hantavirus strains might be more common than generally appreciated. In the case of ANDV Sout and *O. longicaudatus*, it is likely that spillover infection occurs in other host species (e.g., sympatric *A. longipilis* and *L. micropus*) (41, 58). Relatively high seroprevalence observed in *A. longipilis* (1.9%) and the presence of the TS allele in *O. longicaudatus* and *A. longipilis* captured during the same trapping event and locality (Aysén region in Chile) are consistent with spillover (Table 2). Nevertheless, our data and previous studies (42) indicate that, although a number of rodent species can be infected by ANDV Sout, accidental infections do not contribute significantly to the replication cycle or to the maintenance of this virus in nature. Naturally infected rodents other than *O. longicaudatus* (e.g., *A. longipilis*) are unable to transmit virus under experimental settings (42). Thus, *O. longicaudatus* appears to be the sole epidemiologically relevant host of ANDV Sout in Chile.

In Chile, hantavirus-seropositive rodents were obtained from Coquimbo region (30°56'S) to as far south as Aysén region (46°46'S), in contrasting landscapes with vast differences in temperature, precipitation, humidity, and vegetation. Our data suggest a degree of geographic structure of ANDV Sout in Chile, as defined by two well-supported clades when using both S- and M-segment sequences (Fig. 1 to 4). Those clades are congruent with two major ecoregions in Chile. The Chilean matorral or Mediterranean region, located from the Coquimbo region (30°S) south to the Concepción region (36° to 38°S) in Chile, is mainly characterized by a highly heterogeneous vegetation mosaic, with dry xerophytic thorn scrub, evergreen sclerophyllous trees in mountain foothills, and forest of winter deciduous trees (3, 4, 40). The Valdivian temperate forest region in Chile is characterized predominantly by evergreen broadleaf forest mixed with needle-leaved trees, and ranges from 35° to 48° (Aysén region) south latitude (3, 40). Of particular interest is the area between latitudes 36° and 40°S (Bío Bío and Araucanía regions in Chile) that has been reported as an ecoregion interface, where abiotic conditions (e.g., climate, relative humidity, temperature, rainfall, physical and chemical soil composition) have determined differences in the species composition in a number of organisms (4, 12, 64). The ANDV Sout phylogeographic break occurs around 36°S with viral sequences from geographically proximate localities, such as Los Riles (map number 13) and Chillán (map number 14) genetically divergent and placed in different clades (Table 2 and Fig. 1).

Implications of host phylogeographic breaks are still poorly understood but have been related to potentially different pathogenicity (13). Human hantavirus infections in Chile have occurred across a wide geographic range. Sequence differences observed in our analyses did not indicate the presence of

strains or independent lineages other than ANDV, with samples from all four rodent species and humans exhibiting antibodies against hantaviruses when infected with ANDV Sout (Tables 2 and 3 and Fig. 1). Thus, the phylogenetic break observed in ANDV Sout suggests that geographical and historical factors (51) have influenced viral gene flow across these two ecoregions.

Phylogenies obtained using portions of the S segment and two regions of the M segment (Gn and Gc) were highly congruent. All topologies demonstrated similar phylogeographic patterns, with samples placed according to their latitudinal distribution (Fig. 3). Although not fully supported in all the topologies obtained, this pattern was further refined with the detection of two geographically restricted subclades within the Temperate Forest clade (Fig. 1B, 2, and 3B). More extensive sampling and longer viral sequences are necessary to determine finer geographical structure of ANDV Sout. Our multigene analyses revealed that using S-segment partition 2 (600 nt) reconstructed the phylogeographic relationships within ANDV Sout (Fig. 3B) and that using M-segment sequences and concatenating portions from genes with different variability rates increased statistical support for some nodes (Fig. 4). Given that patients with HCPS often have multiple potential sites of exposure, our study also demonstrates the utility of sequencing ANDV Sout to identify the geographic origin of human cases for surveillance purposes (22).

The two clades recovered within ANDV Sout are not observed in phylogenetic reconstructions of its major rodent host, *O. longicaudatus* (46). ANDV Sout did not show any sign of demographic expansion using neutrality tests (Table 5), contrasting with the results obtained by coalescence analysis-based tests performed using the Bayesian approach. BSP showed a sudden growth in the effective number of ANDV infections during 2000 to 2002, which coincides with the onset of human HCPS cases (Fig. 5). These results show the utility of BSP in examining demographic histories of RNA viruses (7, 16, 25, 48) and show that algorithms based on neutrality tests might be too conservative to analyze these viruses. Given that small data sets are prone to disagreement among methods used to detect sites under selection (32), congruence is interpreted as a robust result. Thus, nucleotide substitutions in the S segment of ANDV in Chile were found to be silent, where most amino acid replacements are likely deleterious, a result congruent with the functional importance of this protein. This further supports the idea that ANDV Sout is only moderately diversified in Chile and contrasts with higher levels of diversity observed among hantaviruses found in other regions in South America (8, 10, 27, 43).

There was no evidence of reassortment of the S and M segments, further indicating that a single virus strain circulates in Chile. Of interest, however, was the detection of recombining

FIG. 3. Phylogenetic comparison of the S and M segments of ANDV Sout in Chile. Maximum likelihood phylogenetic trees from aligned sequences of the pruned data set are shown. (A) S-segment phylogenetic tree partition 1 (1 to 333 nt), (B) S-segment phylogenetic tree partition 2 (600 nt), (C) M-Gn phylogenetic analysis (490 nt), and (D) M-Gc phylogenetic tree obtained using 577 nt. The numbers next to the branches are the pseudoreplicate bootstrap values obtained by maximum likelihood (first value) and maximum parsimony (second value) analyses and posterior probabilities obtained with Bayesian inference (third value) (—, no value). Names in italic type are previously reported ANDV lineages (Table 2). Recombining sequences are underlined in panels A and B. The bars indicate 0.05 or 0.1 nucleotide substitution/site.

TABLE 4. S-segment nucleotide (top of diagonal) and amino acid (bottom of diagonal) identity matrix within ANDV Sout and among South American lineages^a

Virus sequence abbreviation	% Identity of viruses																			LN
	FJ865	FJ887	SCA619	CUR092	LR843	TUC979	PNH362	PNV083	TEM703	PAN535	PAN544	RNRS555	RNRS566	CHC378	CHI-7913	AND123	AH-1	LEC ^b	Hu39694 ^b	ORN ^b
FJ865		97.6	93.9	91.7	91.2	86.8	86.3	86	86.3	86.3	86.3	85.7	86.4	86.6	86	84.9	86	80	80.4	79.7
FJ887	97.7		92.5	93.4	92.9	88.2	87.7	87.5	87.7	87.7	87.7	87.4	87.8	87.8	87.4	86.3	87.4	81.3	81.8	81.3
SCA619	99.3	98		91.9	92	87	87	87.1	86.3	87.2	87	86.4	87.2	87.7	87.4	86.9	87.6	80.2	80.7	80.8
CUR092	97.4	99.6	97.7		97.5	88.8	89	89.2	88.2	89.2	89	88.5	88.9	89.3	88.3	88	88.9	82.8	82.1	81.7
LR843	97.7	100	98	99.6		89.2	89	89.2	88.4	89.2	89	88.6	88.9	89.8	88.9	88.5	89.5	82.3	82.6	82
TUC979	97.1	99.3	97.4	99	99.3		97.2	96.7	97.8	97.2	97.2	97.4	97.5	92.5	92.7	92.1	93.8	82.4	82	81.6
PNH362	97.1	99.3	97.4	99	99.3	100		99.1	98	98.7	100	97.4	97.3	92.3	92.5	92.1	93.4	82.1	81.8	81.6
PNV083	97.1	99.3	97.4	99	99.3	100	100		97.6	97.6	99.1	96.7	97	92.4	92.8	92.2	93.7	82.2	82	81.6
TEM703	97.1	99.3	97.4	99	99.3	100	100	100		97.6	98	97.6	97.3	92.1	92.7	91.9	93.4	82	81.7	81.7
PAN535	97.1	99.3	97.4	99	99.3	100	100	100	100		99.3	97	97.3	92.5	92.7	92.3	93.6	82.1	82	81.4
PAN544	97.1	99.3	97.4	99	99.3	100	100	100	100	100		97.4	97.3	92.3	92.5	92.1	93.4	82.1	81.8	81.6
RNRS555	96.8	99	97.1	98.7	99	99.6	99.6	99.6	99.6	99.6	99.6		97.1	92.1	92.3	91.5	92.9	81.5	81.3	81.3
RNRS566	97.1	99.3	97.4	99	99.3	100	100	100	100	100	100	99.6		93.3	93.5	92.8	94.3	82.5	81.8	81.6
CHC378	97.1	99.3	97.4	99	99.3	100	100	100	100	100	100	99.6	100		98	94.6	97.2	82.4	80.9	81.6
CHI-7913	97.1	99.3	97.4	99	99.3	100	100	100	100	100	100	99.6	100	100		94.9	97.4	82.2	81.2	81.1
AND123	97.1	99.3	97.4	99	99.3	100	100	100	100	100	100	99.6	100	100	100		95.3	82.8	81.5	81.6
AH-1	97.1	99.3	97.4	99	99.3	100	100	100	100	100	100	99.6	100	100	100	100		82.5	80.7	81.7
LEC ^b	93.6	95.8	93.9	96.1	95.8	95.8	95.8	95.8	95.8	95.8	95.8	95.5	95.8	95.8	95.8	95.8	95.8		85.4	85.3
Hu39694 ^b	94.2	96.4	94.5	96.8	96.4	96.4	96.4	96.4	96.4	96.4	96.4	96.1	96.4	96.4	96.4	96.4	96.4	99.3		85.9
ORN ^b	93.9	96.1	94.2	96.4	96.1	96.1	96.1	96.1	96.1	96.1	96.1	95.8	96.1	96.1	96.1	96.1	96.1	98	98.7	
LN	87.2	89.4	87.5	89.1	89.4	89.4	89.4	89.4	89.4	89.4	89.4	89.1	89.4	89.4	89.4	89.4	89.4	89.1	89.8	89.8

^a The virus sequence abbreviations shown in Table 2 are used.
^b Samples corresponding to previously reported ANDV lineages other than ANDV Sout: Hu39694 corresponds to ANDV Cent Bs.As., Lechiguanas corresponds to ANDV Cent Lec., and Oran corresponds to ANDV Nort. Sequences for the Andes Central Plata (ANDV Cent Plata) lineage were not available.

FIG. 4. Phylogenetic relationship of concatenated S- and M-segment sequences of ANDV Sout in Chile. Majority rule (50%) consensus tree for the $2Q+I$ model of concatenated M- and S-segment sequences (1,668 nt) using Bayesian inference. The consensus tree is based on 4,500 samples from a converged Markov chain. The numbers above branches are the posterior probability values. Sequence names shown in italic type indicate previously reported ANDV lineages (Table 2). The bar indicates 0.1 nucleotide substitution/site.

nation resulting in divergent topologies obtained using different portions of the S segment (Fig. 3A and B). Tucapel, Panguipulli (sample 104535), and Los Angeles (samples that grouped in the Temperate Forest clade) were identified as the

FIG. 5. Number of ANDV infections estimated from genetic data. The Bayesian skyline plot is derived from S-segment (397-nt) sequences. The thick solid line is the median effective number of ANDV infections (left-hand y axis) and represents the product of effective population size and generation time (in years). The shaded area indicates the 95% highest posterior density (HPD) region interval. Bars represent the total number of human HCPS cases per year (1997 to 2004; right-hand y axis) as reported by the Chilean Ministry of Health.

taxa that reported the recombination events. This is consistent with the idea of coinfection of rodent hosts resulting in recombination events within the S segments of circulating hantaviruses (29, 56). However, although recombination might be a common occurrence (detected in 3 out of 17 sequences analyzed), it is not convincingly contributing to the genetic diversity of ANDV Sout, because most changes between the sequences of the S segments and the Gn and Gc portions of the two clades were found to be synonymous (data not shown). The latter is corroborated by the high percentage of amino acid identity observed within ANDV Sout sequences (Table 4) and the tree topology obtained with the amino acid sequences of N protein (data not shown).

Demonstration of a widespread and well-established hantavirus enzootic in Chile indicates that ANDV Sout is likely to remain a public health concern with significant outbreak risk. The phylogenetic split observed in our study demonstrates that unknown factors contribute to the widespread characteristics

TABLE 5. Descriptive statistics of genetic variation and parameter estimates of ANDV S-segment sequences in Chile

Clade or parameter	N^a	N_h^b	S^c	Hd (mean $\pm$ SD) ^d	π (mean $\pm$ SD) ^e	Results of neutrality test ^f			dN/dS by: ^g		
						Tajima's D test	Fu's F_S test	R_2 test	PAML	FEL ^h	PARRIS (mean $\pm$ SD) ⁱ
Mediterranean	17	14	65	0.971 $\pm$ 0.032	0.056 $\pm$ 0.003	0.165 (NS)	-0.732 (NS)	0.146 (NS)	0.003	0.000	0.003 $\pm$ 0.000
Temperate Forest	26	22	76	0.988 $\pm$ 0.014	0.048 $\pm$ 0.004	-0.304 (NS)	-4.189 (NS)	0.116 (NS)	0.003	0.000	0.002 $\pm$ 0.000
Total	43	36	101	0.991 $\pm$ 0.007	0.077 $\pm$ 0.003	0.397 (NS)	-7.053 (NS)	0.136 (NS)	0.003	0.000	0.004 $\pm$ 0.000

^a N , number of individuals (small mammals or humans).

^b N_h , number of alleles.

^c S , number of segregating sites.

^d Hd, allele diversity.

^e π , nucleotide diversity.

^f NS, not significant ($P > 0.05$ for Tajima's D test and R_2 test and $P > 0.02$ for Fu's F_S test).

^g dN/dS is the ratio between nonsynonymous (dN) and synonymous (dS) nucleotide substitution rates by clade. The average value is shown for the dN/dS values found by using the PAML program and by using the FEL method.

^h FEL, fixed effects likelihood method.

ⁱ PARRIS, partitioning approach for robust inference of selection methods.

of ANDV Sout, as well as its maintenance and persistence in two very different ecoregions in Chile. Nevertheless, whether ecological, historical, or other factors contribute to this phenomenon remains to be investigated. Both clades of this virus appear among patients with HCPS, suggesting that they have similar pathogenic potential to humans. Due to the highly pathogenic characteristic of this virus and its unique ability to transmit from person to person (18, 34, 38), it is imperative that we understand the underlying historical and ecological factors of ANDV in the hope of designing and implementing effective preventive and surveillance strategies.

ACKNOWLEDGMENTS

We thank Eliecer Villagra and Jorge Fernandez for contributing tissue samples and sequences and Paula Godoy for excellent technical assistance. We are grateful to Gregory Ebel and Eric Bortz for helpful comments and critically reviewing the manuscript and to the anonymous reviewers for valuable suggestions. We also thank the Servicio Agrícola y Ganadero (SAG) and Corporación Nacional Forestal (CONAF) for trapping permits and the Ministry of Health in Chile and local health services that provided us with information on sampling sites.

This study was supported by United States Public Health Service grants, UO1 AI56618, UO1 AI54779, R56 AI 034448, U19 AI45452, UO1 AI054779, and FONDECYT 1070331. R.A.M. and F.T.-P. were supported by the Fogarty Actions for Building Capacity award 5D43 TW01133.

REFERENCES

- Abascal, F., R. Zardoya, and D. Posada. 2005. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21**:2104–2105.
- Alfaro, M. E., S. Zoller, and F. Lutzoni. 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. *Mol. Biol. Evol.* **20**:255–266.
- Amigo, J., and C. Ramirez. 1998. A bioclimatic classification of Chile: woodland communities in the temperate zone. *Plant Ecol.* **136**:9–26.
- Armesto, J. J., M. T. K. Arroyo, L. F. Hinojosa, T. T. Veblen, K. R. Young, and A. R. Orme. 2007. The Mediterranean environment of central Chile, p. 184–199. *In* T. T. Veblen, A. R. Orme, and K. R. Young (ed.), *The physical geography of South America*. Oxford University Press, Oxford, United Kingdom.
- Bandelt, H. J., P. Forster, and A. Rohl. 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **16**:37–48.
- Biek, R., A. J. Drummond, and M. Poss. 2006. A virus reveals population structure and recent demographic history of its carnivore host. *Science* **311**:538–541.
- Biek, R., J. C. Henderson, L. A. Waller, C. E. Rupprecht, and L. A. Real. 2007. A high-resolution genetic signature of demographic and spatial expansion in epizootic rabies virus. *Proc. Natl. Acad. Sci. USA* **104**:7993–7998.
- Bohman, M. C., S. P. Morzunov, J. Meissner, M. B. Taylor, K. Ishibashi, J. Rowe, S. Levis, D. Enria, and S. C. St. Jeor. 2002. Analysis of hantavirus genetic diversity in Argentina: S segment-derived phylogeny. *J. Virol.* **76**:3765–3773.
- Botten, J., K. Mirowsky, C. Y. Ye, K. Gottlieb, M. Saavedra, L. Ponce, and B. Hjelle. 2002. Shedding and intracage transmission of Sin Nombre hantavirus in the deer mouse (*Peromyscus maniculatus*) model. *J. Virol.* **76**:7587–7594.
- Chu, Y. K., B. Milligan, R. D. Owen, D. G. Goodin, and C. B. Jonsson. 2006. Phylogenetic and geographical relationships of hantavirus strains in eastern and western Paraguay. *Am. J. Trop. Med. Hyg.* **75**:1127–1134.
- Davis, P. L., E. C. Holmes, F. Larrous, W. H. van der Poel, K. Tjornehoj, W. J. Alonso, and H. Bourhy. 2005. Phylogeography, population dynamics, and molecular evolution of European bat lyssaviruses. *J. Virol.* **79**:10487–10497.
- di Castri, F., and E. Hajek. 1976. *Bioclimatología de Chile*. Ediciones de la Pontificia Universidad Católica de Chile, Santiago, Chile.
- Dragoo, J. W., J. A. Lackey, K. E. Moore, E. P. Lessa, J. A. Cook, and T. L. Yates. 2006. Phylogeography of the deer mouse (*Peromyscus maniculatus*) provides a predictive framework for research on hantaviruses. *J. Gen. Virol.* **87**:1997–2003.
- Drummond, A. J., G. K. Nicholls, A. G. Rodrigo, and W. Solomon. 2002. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics* **161**:1307–1320.
- Drummond, A. J., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**:214.
- Drummond, A. J., A. Rambaut, B. Shapiro, and O. G. Pybus. 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* **22**:1185–1192.
- Felsenstein, J. 1985. Confidence limits on phylogenies—an approach using the bootstrap. *Evolution* **39**:783–791.
- Ferres, M., P. Vial, C. Marco, L. Yanez, P. Godoy, C. Castillo, B. Hjelle, I. Delgado, S. J. Lee, and G. J. Mertz. 2007. Prospective evaluation of household contacts of persons with hantavirus cardiopulmonary syndrome in Chile. *J. Infect. Dis.* **195**:1563–1571.
- Fu, Y. X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**:915–925.
- Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**:95–98.
- Hjelle, B., F. Chavez-Giles, N. Torrez-Martinez, T. Yamada, J. Sarisky, M. Ascher, and S. Jensen. 1994. Dominant glycoprotein epitope of Four Corners hantavirus is conserved across a wide geographical area. *J. Gen. Virol.* **75**:2881–2888.
- Hjelle, B., N. Torrez-Martinez, F. T. Koster, M. Jay, M. S. Ascher, T. Brown, P. Reynolds, P. Ettestad, R. E. Voorhees, J. Sarisky, R. E. Enscoe, L. Sands, D. G. Mosley, C. Kioski, R. T. Bryan, and C. M. Sewell. 1996. Epidemiologic linkage of rodent and human hantavirus genomic sequences in case investigations of hantavirus pulmonary syndrome. *J. Infect. Dis.* **173**:781–786.
- Hjelle, B., and T. Yates. 2001. Modeling hantavirus maintenance and transmission in rodent communities. *Curr. Top. Microbiol. Immunol.* **256**:77–90.
- Holmes, E. C. 2004. The phylogeography of human viruses. *Mol. Ecol.* **13**:745–756.
- Hon, C. C., T. Y. Lam, A. Drummond, A. Rambaut, Y. F. Lee, C. W. Yip, F. Zeng, P. Y. Lam, P. T. Ng, and F. C. Leung. 2006. Phylogenetic analysis reveals a correlation between the expansion of very virulent infectious bursal disease virus and reassortment of its genome segment B. *J. Virol.* **80**:8503–8509.
- Huelsenbeck, J. P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**:754–755.
- Johnson, A. M., L. T. M. de Souza, I. B. Ferreira, L. E. Pereira, T. G. Ksiazek, P. E. Rollin, C. J. Peters, and S. T. Nichol. 1999. Genetic investigation of novel hantaviruses causing fatal HPS in Brazil. *J. Med. Virol.* **59**:527–535.
- Jonsson, C. B., and C. S. Schmaljohn. 2001. Replication of hantaviruses. *Curr. Top. Microbiol. Immunol.* **256**:15–32.
- Klempa, B., H. A. Schmidt, R. Ulrich, S. Kaluz, M. Labuda, H. Meisel, B. Hjelle, and D. H. Kruger. 2003. Genetic interaction between distinct Dobrava hantavirus subtypes in *Apodemus agrarius* and *A. flavicollis* in nature. *J. Virol.* **77**:804–809.
- Kosakovsky Pond, S. L., and S. D. Frost. 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* **21**:2531–2533.
- Kosakovsky Pond, S. L., S. D. Frost, and S. V. Muse. 2005. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* **21**:676–679.
- Kosakovsky Pond, S. L., and S. D. Frost. 2005. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol. Biol. Evol.* **22**:1208–1222.
- Kosakovsky Pond, S. L., D. Posada, M. B. Gravenor, C. H. Woelk, and S. D. Frost. 2006. Automated phylogenetic detection of recombination using a genetic algorithm. *Mol. Biol. Evol.* **23**:1891–1901.
- Lazaro, M. E., G. E. Cantoni, L. M. Calanni, A. J. Resa, E. R. Herrero, M. A. Iacono, D. A. Enria, and S. M. G. Cappa. 2007. Clusters of hantavirus infection, southern Argentina. *Emerg. Infect. Dis.* **13**:104–110.
- Levis, S., S. P. Morzunov, J. E. Rowe, D. Enria, N. Pini, G. Calderon, M. Sabatini, and S. C. St. Jeor. 1998. Genetic diversity and epidemiology of hantaviruses in Argentina. *J. Infect. Dis.* **177**:529–538.
- Lopez, N., P. Padula, C. Rossi, M. E. Lazaro, and M. T. Franze-Fernandez. 1996. Genetic identification of a new hantavirus causing severe pulmonary syndrome in Argentina. *Virology* **220**:223–226.
- Lopez, N., P. Padula, C. Rossi, S. Miguel, A. Edelstein, E. Ramirez, and M. T. Franze-Fernandez. 1997. Genetic characterization and phylogeny of Andes virus and variants from Argentina and Chile. *Virus Res.* **50**:77–84.
- Martinez, V. P., C. Bellomo, J. San Juan, D. Pinna, R. Forlenza, M. Elder, and P. J. Padula. 2005. Person-to-person transmission of Andes virus. *Emerg. Infect. Dis.* **11**:1848–1853.
- Mills, J. N., T. L. Yates, J. E. Childs, R. R. Parmenter, T. G. Ksiazek, P. E. Rollin, and C. J. Peters. 1995. Guidelines for working with rodents potentially infected with hantavirus. *J. Mammal.* **76**:716–722.
- Olson, D. M., E. Dinerstein, E. D. Wikramanayake, N. D. Burgess, G. V. N. Powell, E. C. Underwood, J. A. D'Amico, I. Itoua, H. E. Strand, J. C. Morrison, C. J. Loucks, T. F. Allnutt, T. H. Ricketts, Y. Kura, J. F. Lamoreux, W. W. Wettengel, P. Hedao, and K. R. Kassem. 2001. Terrestrial ecoregions of the world: a new map of life on Earth. *Bioscience* **51**:933–938.
- Ortiz, J. C., W. Venegas, J. A. Sandoval, P. Chandia, and F. Torres-Pérez.

2004. Hantavirus in rodents of the VIII Region of Chile. *Rev. Chil. Hist. Nat.* **77**:251–256.
42. Padula, P., R. Figueroa, M. Navarrete, E. Pizarro, R. Cadiz, C. Bellomo, C. Joffe, L. Zaror, E. Rodriguez, and R. Murua. 2004. Transmission study of Andes hantavirus infection in wild sigmodontine rodents. *J. Virol.* **78**:11972–11979.
43. Padula, P. J., S. B. Colavecchia, V. P. Martinez, M. O. G. Della Valle, A. Edelstein, S. D. L. Miguel, J. Russi, J. M. Riquelme, N. Colucci, M. Almiron, and R. D. Rabinovich. 2000. Genetic diversity, distribution, and serological features of hantavirus infection in five countries in South America. *J. Clin. Microbiol.* **38**:3029–3035.
44. Padula, P. J., A. Edelstein, S. D. L. Miguel, N. M. Lopez, C. M. Rossi, and R. D. Rabinovich. 1998. Hantavirus pulmonary syndrome outbreak in Argentina: molecular evidence for person-to-person transmission of Andes virus. *Virology* **241**:323–330.
45. Pagel, M., and A. Meade. 2004. A phylogenetic mixture model for detecting pattern-heterogeneity in gene sequence or character-state data. *Syst. Biol.* **53**:571–581.
46. Palma, R. E., E. Rivera-Milla, J. Salazar-Bravo, F. Torres-Pérez, U. F. J. Pardinas, P. A. Marquet, A. E. Spotorno, A. P. Meynard, and T. L. Yates. 2005. Phylogeography of *Oligoryzomys longicaudatus* (Rodentia: Sigmodontinae) in temperate South America. *J. Mammal.* **86**:191–200.
47. Posada, D., and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**:817–818.
48. Pybus, O. G., E. Barnes, R. Taggart, P. Lemey, P. V. Markov, B. Rasachak, B. Syhavong, R. Phetsouvanah, I. Sheridan, I. S. Humphreys, L. Lu, P. N. Newton, and P. Klennerman. 29 October 2008. Genetic history of hepatitis C virus in East Asia. *J. Virol.* doi:10.1128/JVI.01501-08.
49. Ramos-Onsins, S. E., and J. Rozas. 2002. Statistical properties of new neutrality tests against population growth. *Mol. Biol. Evol.* **19**:2092–2100.
50. Ramsden, C., F. L. Melo, L. M. Figueiredo, E. C. Holmes, and P. M. Zanotto. 2008. High rates of molecular evolution in hantaviruses. *Mol. Biol. Evol.* **25**:1488–1492.
51. Real, L. A., J. C. Henderson, R. Biek, J. Snaman, T. L. Jack, J. E. Childs, E. Stahl, L. Waller, R. Tinline, and S. Nadin-Davis. 2005. Unifying the spatial population dynamics and molecular evolution of epidemic rabies virus. *Proc. Natl. Acad. Sci. USA* **102**:12107–12111.
52. Rozas, J., J. C. Sanchez-DelBarrio, X. Messeguer, and R. Rozas. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**:2496–2497.
53. Scheffler, K., D. P. Martin, and C. Seoighe. 2006. Robust inference of positive selection from recombining coding sequences. *Bioinformatics* **22**:2493–2499.
54. Schmaljohn, C., and B. Hjelle. 1997. Hantaviruses: a global disease problem. *Emerg. Infect. Dis.* **3**:95–104.
55. Schmidt, H. A., K. Strimmer, M. Vingron, and A. von Haeseler. 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* **18**:502–504.
56. Sibold, C., H. Meisel, D. H. Kruger, M. Labuda, J. Lysy, O. Kozuch, M. Pejcoch, A. Vaheri, and A. Plyusnin. 1999. Recombination in Tula hantavirus evolution: analysis of genetic lineages from Slovakia. *J. Virol.* **73**:667–675.
57. Sironen, T., A. Vaheri, and A. Plyusnin. 2001. Molecular evolution of Puumala hantavirus. *J. Virol.* **75**:11803–11810.
58. Spotorno, A., R. E. Palma, and J. P. Valladares. 2000. Biología de roedores reservorios de hantavirus en Chile. *Rev. Chilena Infectol.* **17**:197–210.
59. Swofford, D. L. 2002. PAUP* beta version. Phylogenetic Analysis Using Parsimony (* and other methods). Sinauer Associated, Sunderland, MA.
60. Tajima, F. 1989. The effect of change in population size on DNA polymorphism. *Genetics* **123**:597–601.
61. Tamura, K., and M. Nei. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**:512–526.
62. Toro, J., J. D. Vega, A. S. Khan, J. N. Mills, P. Padula, W. Terry, Z. Yadon, R. Valderrama, B. A. Ellis, C. Pavletic, R. Cerda, S. Zaki, S. Wun-Ju, R. Meyer, M. Tapia, C. Mansilla, M. Baro, J. A. Vergara, M. Concha, G. Calderon, D. Enria, C. J. Peters, and T. G. Ksiazek. 1998. An outbreak of hantavirus pulmonary syndrome, Chile, 1997. *Emerg. Infect. Dis.* **4**:687–694.
63. Torres-Pérez, F., J. Navarrete-Droguett, R. Aldunate, T. L. Yates, G. J. Mertz, P. A. Vial, M. Ferres, P. A. Marquet, and R. E. Palma. 2004. Peridomestic small mammals associated with confirmed cases of human hantavirus disease on southcentral Chile. *Am. J. Trop. Med. Hyg.* **70**:305–309.
64. Villagran, C., L. F. Hinojosa, J. Llorente-Bousquets, and J. J. Morrone. 2005. Esquema biogeográfico de Chile, p. 551–577. Regionalización biogeográfica en Iberoamérica y tópicos afines. Ediciones de la Universidad Nacional Autónoma de México, Jiménez Editores, Mexico.
65. Yang, Z. H. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13**:555–556.
66. Yee, J., I. A. Wortman, R. A. Nofchissey, D. Goade, S. G. Bennett, J. P. Webb, W. Irwin, and B. Hjelle. 2003. Rapid and simple method for screening wild rodents for antibodies to Sin Nombre hantavirus. *J. Wildl. Dis.* **39**:271–277.