


# Ancient DNA Reveals That the Variability of the DRD4 -521 C/T SNP Associated with Novelty Seeking Behavior is Influenced by Selection in Western South American Populations

Austin de Rubira<sup>1</sup> · Lea Georges<sup>1,2</sup> ·  
Lars Fehren-Schmitz<sup>1,2</sup> 

Received: 17 July 2015 / Revised: 20 August 2015 / Accepted: 24 August 2015 /  
Published online: 1 September 2015  
© Springer International Publishing 2015

**Abstract** Certain Polymorphisms in the DRD4 gene have been associated with behavioral traits such as novelty seeking (NS), impulsivity, and ADHD. Research has suggested that these personality traits may affect individual fitness differently depending on social structure, subsistence patterns, and mobility. To test this hypothesis this study examines the diachronic change in the frequency of the C allele of the -521 C/T SNP in the DRD4 promoter region, thought to contribute to the NS phenotype, analyzing ancient DNA from pre-Columbian South America. We used a SBE PCR based approach to analyze the DRD4 SNP in 125 prehistoric samples dating from ~8000 BP to the end of the pre-Columbian period and deriving from a diversity of socioeconomic environments. We performed forward simulations of genetic drift with selection to test if difference in allele frequencies over time can solely be explained by random evolutionary processes or if selection has to be considered. Further, we performed statistical tests to explore possible association of observed differences with socioeconomic or environmental factors. DRD4 genotypes for 79 individuals were obtained. We find signals of slight selection acting on the DRD4-521 T allele. While the postulated association between the -521 C/T DRD4 SNP with modes of subsistence was not confirmed the results suggest a possible impact of socioeconomic complexity

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s40750-015-0033-5) contains supplementary material, which is available to authorized users.

---

✉ Lars Fehren-Schmitz  
lfahrens@ucsc.edu

<sup>1</sup> Department of Anthropology, University of California, Santa Cruz, 1156 High Street, Santa Cruz Ca 95064, USA

<sup>2</sup> Historical Anthropology and Human Ecology, University Goettingen, D-37073 Goettingen, Germany

and political organization on the variability of the SNP in in pre-Columbian western South American populations.

**Keywords** Ancient DNA · DRD4 · Human behavior · South America · Novelty seeking · Socioeconomic complexity

## Introduction

The dopamine receptor D4 (DRD4) gene has been of particular interest in recent studies both because it is highly variable in humans, and it seems to affect complex behaviors. Polymorphisms in this gene have been associated with behavioral traits such as novelty seeking and impulsivity (Ebstein et al. 1996; Okuyama et al. 2000), the desire for migration (Chen et al. 1999; Matthews and Butler 2011), sensation seeking (C. J. Thomson et al. 2013), as well as modern disorders such as ADHD (Bellgrove et al. 2005) and schizophrenia (Lai et al. 2010).

Certain polymorphisms in the DRD4 gene such as the 7 repeat allele in a 48 bp VNTR in exon 3 and the C allele of the -521 C/T Single Nucleotide Polymorphism (SNP) in the promoter region (rs1800955) have been discussed to cause a reduced sensitivity of the rewards system (Hattori et al. 2009; Lichter et al. 1993; Munafò et al. 2008; Okuyama et al. 2000) that can be associated with a tendency for novelty seeking behavior. Novelty seeking (NS) is a personality trait is defined by a tendency towards exploratory behavior and intense excitement in response to novel stimuli, impulsive decision making, and quick loss of temper and avoidance of frustration. Research has suggested that in resource deprived or constantly changing environments, this can be a beneficial trait, allowing a person to find novel solutions to survive. On the other hand, in populations with abundant resources and highly structured social environments, people with this trait would exhibit restless behavior, and the same trait would become detrimental to success (Jensen et al. 1997; Williams and Taylor 2006).

A recent study (Eisenberg et al. 2008) showed evidence of this pattern in a modern population. In this study on the Ariaal people of Africa, the 7 repeat allele VNTR form was found to correlate to better nutrition in nomads, but worse nutrition in agriculturalists. Another study comparing indigenous hunter-gatherers and agriculturalists in South America found a positive correlation with foraging subsistence and high frequencies of the 7 repeat allele suggesting that the subsistence mode has selective impact on the polymorphism (Tovo-Rodrigues et al. 2010). Both studies have been conducted on modern populations, thus it has to be considered that observed frequencies might have been biased by past admixture with other populations or even that the modes of subsistence as well as the grades of social complexity in these populations changed over time, meaning that e.g. populations classified as hunter-gatherers had a history of being horticulturalists, pastoralists, or agriculturalist and more recently switched to a foraging lifestyle. Here we suggest another research strategy that allows us to test possible selective influences of socioeconomic changes on DRD4 polymorphisms over time: the analysis of ancient DNA from prehistoric South American populations dating to different time periods and associated with different modes of subsistence and organization through the archaeological record. If NS has been an evolutionary beneficial adaptation in the human past, and there is a genetic predisposition for this trait, we

should find high frequencies of alleles associated with NS in prehistoric hunter and gatherers. We should also observe a frequency decrease of alleles determining this personality trait with increasing social complexity, structure and the transition to producing economies in human groups, if NS is really maladaptive in these social environments, due to selective evolutionary processes.

Most studies on the DRD4 gene have largely focused on the 7 repeat allele of the 48 base-pair VNTR in exon 3. Since we aim to address our hypothesis using ancient DNA, issues of preservation would prevent us from observing a polymorphism involving so many base pairs. For this reason, this study looks at the  $-521$  C/T SNP (rs1800955) instead. In addition, this SNP is also potentially a more reliable marker for examining novelty seeking. Meta-analysis has shown that the C allele of this SNP could account for up to 2 % of phenotypic variance of the novelty seeking personality trait, while no association with this trait was found for the VNTR (Munafò et al. 2008; Schinka et al. 2002). The promoter region SNP has been reported to be in *linkage disequilibrium* with the exon III polymorphism (Ekelund et al. 2001), and is associated with variation in expression of the D4 receptor, with the T allele associated with a reduction in transcription levels of up to 40 % compared with the C allele (Ronai et al. 2001).

The goal of this study is to determine whether or not the frequency of C alleles in the SNP  $-521$  C/T decreased over time in native South American populations due to selection for the T allele, and if possible to determine which societal changes most likely had the largest impact on allele frequencies.

## Materials and Methods

The samples for this study derive from 17 archaeological sites mostly from the Central Andes with the exception of the site Arroyo Seco 2 situated in the Argentinian Pampas (Fig. 1). All samples analyzed in this study have been part of previous studies investigating mitochondrial (mt) and nuclear genetic loci or are currently analyzed in genome wide sequencing projects (L Fehren-Schmitz et al. 2010; Lars Fehren-Schmitz et al. 2014, 2015). The sites studied were chosen from the collections of the UCSC Human Paleogenomics Lab (UCSC\_HPL) so that they represent a diversity of archaeological periods, associated subsistence strategies, and grades of socioeconomic complexity. The samples from 125 individuals selected from these sites represent the ones with the best DNA preservation observed in previous investigations (endpoint PCR, qPCR, shotgun sequencing), and therefore were most promising to also contain autosomal DNA. For comparison analysis we collected genotype data from the 1000 Genome Project Lima (PEL) population (Altshuler et al. 2012).

All populations analyzed besides their chronological division have been assigned to basic categories with regards to their subsistence (hunting/foraging; horticulture, pastoralism, agriculture, and mixed economies), social organization following the simplified system of Service (Service, E 1962); bands, tribes, chiefdoms, states), mobility (nomadic, semi-sedentary, sedentary), and altitude (lowlands/coast, highlands above 2500 m.a.s.l.). For a list of all populations/samples studied and their assigned categories refer to Table 1 and supplementary Table 1. Assignments to categories were based on the available archaeological literature for the sites (e.g. Grieder 1988; Reindel 2009; Pucciarelli et al. 2010). It has to be acknowledged that these categorizations cannot



**Fig. 1** Map showing the archaeological sites from which samples derived for this study

reflect the individual complexity of the archaeological groups included in this study and that transitions between assigned categories can be more or less fluent, but as a model for the questions addressed in this study the chosen systematics are the only ones feasible.

DNA was extracted from the teeth samples employing the protocol described in Fehren-Schmitz et al. (2014, 2015). At least two independent extracts were made for each sample to allow an authentication of analysis results by comparison. All analyses were carried out in laboratories entirely dedicated to ancient DNA analysis at UCSC (UCSC\_HPL) and the Department of Anthropology Goettingen (GoA); hence no modern DNA-based studies have been performed in the laboratory. All analyses were carried out according to the strict precautions and contamination prevention strategies necessary for ancient DNA analysis described in detail in Fehren-Schmitz et al. (2010, 2015).

To determine the genotypes of the  $-521$  C/T (rs1800955) Single Nucleotide Polymorphism in the dopamine D4 receptor gene (DRD4) we developed a PCR amplifying 87 bp spanning the sequence containing the SNP of interest coupled with a Single Base Extension (SBE) assay to directly type the SNP. Primers for the main PCR and the SBE PCR were designed using the Primer Select software (Lasergene 8.0 package, DNASTar). The minisequencing/SBE primers were designed one base contiguous to the polymorphic site of interest in reverse orientation. An additional, polymeric-A tail was added to the 5' end in order to ensure an effective separation of product lengths

**Table 1** List of sites tested showing archaeological dates and categorizations with regards to social organization, subsistence, and altitude

Site/region	Period	Date	n	Altitude	Time category	Mobility	Organization	Subsistence mode
Arroyo Seco	Archaic, Early	7500–6500 BP	2	Low	Early	nomadic	Band	Hunting/Foraging
Lauricocha	Archaic, Early	8000–7000 BP	2	High	Early	nomadic	Band	Hunting/Foraging
Pemil Alto	Archaic, Middle	5500–5000 BP	5	Low	Early	semi-sedentary	Band/Tribe	Hunting/Foraging/ Horticulture
La Galgada	Archaic, Late	4500 BP	6	High	Early	semi-sedentary	Tribe	Hunting/Foraging/ Pastoralism
Lauricocha	Archaic, Late	5000 BP	1	High	Early	semi-sedentary	Tribe	Hunting/Foraging/ Pastoralism
Lauricocha	Formative Period, Initial	3800 BP	1	High	Intermediate	sedentary	Tribe	Hunting/Foraging/ Pastoralism
Pampas Gramalote	Formative Period, Initial	3800 BP	1	Low	Intermediate	sedentary	Tribe	Agriculture/Foraging
Pisagua	Formative Period, Initial	3800–2800 BP	1	Low	Intermediate	sedentary	Tribe	Agriculture/Foraging
Palpa	Formative Period, Early Horizon	2400–2200 BP	4	Low	Intermediate	sedentary	Tribe	Agriculture/Foraging
Palpa	Early Intermediate Period	1700–1500 BP	13	Low	Intermediate	sedentary	Chiefdom	Agriculture
Loma Salvatierra	Early Intermediate Period	1800–1500 BP	1	Low	Intermediate	sedentary	Chiefdom	Agriculture
Laramate	Middle Horizon	1350–1100 BP	18	High	Late	sedentary	State	Agriculture
Laramate	Late Intermediate Period	1000–800 BP	20	High	Late	sedentary	State	Agriculture
San Sebastian	Late Horizon	650–500 BP	4	High	Late	sedentary	State	Agriculture

during electrophoresis (Nelson et al. 2007). Primer sequences can be found in Supplementary Table 2. The main PCR amplification was carried out in a total volume of 12.5 ul containing 6.25 ul Qiagen Multiplex PCR Master Mix (Qiagen, Hilden, Germany), 0.2 uM of each primer, 2–5 ul of DNA template, and DNase free RT-PCR grade H<sub>2</sub>O (Ambion1, Austin, TX) to complete the final reaction volume. PCR amplification took place in a C1000 Touch Thermal Cycler (Biorad, Hercules, CA) under the following conditions: initialization at 95 °C for 5 min; 40 cycles at 94 °C for 1 min, 65 °C for 1 min, and 72 °C for 1 min; final elongation at 60 °C for 30 min. PCR success and PCR product quantity were checked by gel electrophoresis on 2.5 % agarose gels.

To remove excess primers and dNTPs from the previous reaction the 7.5 ul PCR product of the basis PCR were first incubated together with 2.5 units (U) rAPid Alkaline Phosphatase (1 U/ul, Roche, Mannheim, Germany), and 1 U Exo I (20 U/ul, New England Biolabs) in a thermocycler for 1 hour at 37 °C, and following 15 min at 75 °C for heat inactivation. Afterwards 1 ul of the purified PCR product was used for the SBE reaction in a final reaction volume of 5ul consisting of the template, 2.5 ul SNaPshot Ready Reaction Mix (Applied Biosystems, Carlsbad, CA, USA), 0.1 uM SB primer and 20uM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (Merck, Mannheim, Germany). The ammonium sulfate was added for the suppression of nonspecific peaks in the primer extension reaction (Doi et al. 2004). Subsequently, the reaction mixture was treated with 2.5 units of rAPid Alkaline Phosphatase (Roche) at 37 °C for 1 h, and 75 °C for 15 min. to remove the unincorporated fluorescent ddNTPs remaining in the mixture. The Primer extension reaction products were separated and detected by capillary electrophoresis using POP4 polymer on an ABI PRISM 310 Genetic Analyzer (Applied Biosystems). The data was analyzed using GeneScan Analysis Software Version 3.1.2 (Applied Biosystems), and DRD4 -521C/T genotypes were determined by confirming the base substitution at the SNP site. Supporting Information Table 1 shows the base substitutions for the SNP for each sample successfully tested.

To determine the mitochondrial haplogroups of individuals not tested before we employed a multiplex SBE assay typing 26 SNPs in the mitochondrial genome in parallel determining common Native American haplogroups. Protocols and conditions for the AmericaPlex26 were followed as described in Coutinho et al. (Coutinho et al. 2014). Analysis of mtDNA was performed to ensure the Native American ancestry of the individuals.

Each PCR was conducted at least two times from each of the two independent DNA extracts. Additionally 70 samples were also tested in a second ancient DNA lab at the University of Goettingen (Goa) employing the same protocols as described above. If results were inconsistent and there was a risk of false allele determination due to allelic dropout, experiments were repeated more often to find a consensus or discarded. Negative extraction controls and negative PCR controls were employed in this study.

Allele frequencies for the DRD4 SNP were obtained directly by gene counting. The Hardy–Weinberg equilibrium was assessed by the exact test of Guo and Thompson (Guo and Thompson 1992), using the ARLEQUIN v.3.5 software.

To evaluate if changes in the DRD4 rs1800955 T allele frequency observed in the time transect can be explained by genetic drift, or if natural selection needs to be considered we used a forward simulation approach based on Wilde et al. (Wilde et al. 2014) that also allows the estimation of the selection coefficient (*s*). To reflect uncertainty in the ancient allele frequencies due to low sample size we first drew the allele

frequency estimate for the oldest population (Early/Middle Archaic), from a random Beta ( $n_p + 1, n_q + 1$ ) distribution, where  $n_p$  and  $n_q$  were equal to the number of the respective ancestral and derived allele, in each forward simulation. We then obtained 10,000 MCMC samples by binomial sampling across generations for each combination of priors. We used a forward simulation for drift with selection using equation 3.5 from Maynard Smith (Maynard Smith 1998). The number of generation forward simulated was determined by defining the number of generations between the oldest population (Early/Middle Archaic) and the youngest population (Late Horizon/Modern Lima) tested (318 generations, assuming 25 years generation time). We tested 110 different combinations of priors. The variable priors tested were effective population size ( $N_e$ ) at the time of the ancient sample ( $N_e = 100\text{--}100,000$ ) and the selection coefficient ( $s = 0.00\text{--}0.02$ ). The upper limits of both priors tested were defined by the empirical observations made during the simulations. As there is no evidence for a dominant or recessive mode of action for the SNP, an codominant model of inheritance was applied (Thomson et al. 2014). Subsequently, the simulated distribution of allele frequencies at generation 0 were compared with those observed using the equation  $1 - 2 \times |0.5 - P|$ , where  $P$  is the proportion of simulated modern allele frequencies that are greater than that observed, yielding a two-tailed empirical  $P$  value for the observed allele frequency changes for all prior combinations tested (Voight et al. 2005; Wilde et al. 2014).

In an attempt to go beyond the detection of a potential selective impact on the variability of DRD4 -521 we also tested the association between the allele and genotype frequency of the SNP and populations in different ways of living following the categorizations described in supplementary Table 1. Allele distributions were compared using  $X^2$  Tests or Fisher's exact test depending on the sample size of the combinations tested. All tests were two-tailed, and significance level was set at 0.05. The analyses were performed with the SPSS software, version 16 (SPSS, USA). Population genetic structure was investigated using analysis of molecular variance (AMOVA; Excoffier et al. 1992). Subsistence, socioeconomic complexity, mobility, and geography grouping approaches for subdivision were considered using the ARLEQUIN v. 3.5 program (Excoffier and Lischer, 2010).

## Results

A total of 79 ancient samples out of the 125 samples tested were successfully genotyped for the DRD4 -521 C/T alleles. Only those samples were integrated into the downstream analyses that reproducibly yielded the same genotypes in at least four independent PCRs from two independent extracts to exclude wrong genotype calls due to allelic dropout. The successfully analyzed samples distribute over all archaeological sites and periods that have been tested with the later sites from the Andean highlands showing significantly better DNA preservation. The individual genotyping results are found in Supplementary Table 1. All individuals tested with the 26 mitochondrial SNP AmericaPlex assay belong to one of the four Native American mt-haplogroups A, B, C, and D (cf. Supplementary Table 1).

The successfully sequenced samples were grouped into categories by altitude, mobility, subsistence pattern, chronology, and social organization for analysis.

Table 2 shows the genotype and allele frequencies for all the categories in the different group comparisons.

We observed a slight increase of the DRD4 -541 T allele frequency from the oldest population associated with a foraging lifestyle to the youngest pre-Columbian populations (cf. Table 2). It is worth mentioning that the DRD4 SNP genotype- and allele-frequencies of the late pre-Columbian populations resemble that of the modern Peruvian population from Lima. Chi square tests show that the allele frequencies between the oldest and youngest group are slightly non-significant ( $df = 1$ ,  $\chi^2 = 2.24$ ,  $p = 0.088$ ) while stepwise comparison of the four chronological groups (Early, Intermediate, Late, Modern) was tested slightly significant ( $df = 3$ ,  $\chi^2 = 7.82$ ,  $p = 0.045$ ). Since the accuracy of these standard tests probably suffers from the small sample size we further tested if the observed allele frequency changes can be explained by neutral genetic drift alone using a binomial forward simulation approach. While neutrality (selection coefficient  $s = 0$ ) for the DRD4 SNP was not completely rejected assuming very low population sizes ( $N_e = 100$ – $1000$ ) scenarios considering light selection found higher support using a codominant model (Fig. 2, Supplementary Table 3). The

**Table 2** Genotypic and allelic frequencies of the DRD4 -521 C/T polymorphism observed in the ancient South American samples grouped by Chronology, Altitude, Mobility, Subsistence, and political organization (ref. Supplementary Table 1 for individual genotypes)

Comparison	Categories	n	Genotypes			Alleles	
			CC	CT	TT	C	T
Chronology	Early-/Middle Archaic	9	0.33	0.33	0.33	0.500	0.500
	Late Archaic	7	0.29	0.57	0.14	0.571	0.429
	Formative Period	6	0.17	0.67	0.17	0.500	0.500
	Early Intermediate	14	0.29	0.50	0.21	0.536	0.464
	Middle Horizon	18	0.17	0.39	0.44	0.361	0.639
	Late Intermediate	20	0.10	0.50	0.40	0.350	0.650
	Late Horizon	4	0.25	0.25	0.50	0.375	0.625
	Modern Lima	85	0.14	0.45	0.41	0.365	0.635
Broader chronological groups	Early	17	0.31	0.44	0.25	0.531	0.469
	Intermediate	23	0.29	0.52	0.19	0.548	0.452
	Late	30	0.14	0.43	0.43	0.357	0.643
	Modern	62	0.14	0.45	0.41	0.365	0.635
Altitude	Low	27	0.33	0.41	0.26	0.537	0.463
	High	52	0.15	0.48	0.37	0.394	0.606
Mobility	Nomadic	4	0.00	0.50	0.50	0.250	0.750
	Semi-Sedentary	12	0.42	0.42	0.17	0.625	0.375
	Sedentary	63	0.19	0.46	0.35	0.421	0.579
Mode of subsistence	Mixed-Foraging	17	0.29	0.47	0.24	0.529	0.471
	Agriculture	62	0.19	0.45	0.35	0.419	0.581
Political organization	Non-State	37	0.30	0.49	0.22	0.541	0.459
	State	42	0.14	0.43	0.43	0.357	0.643



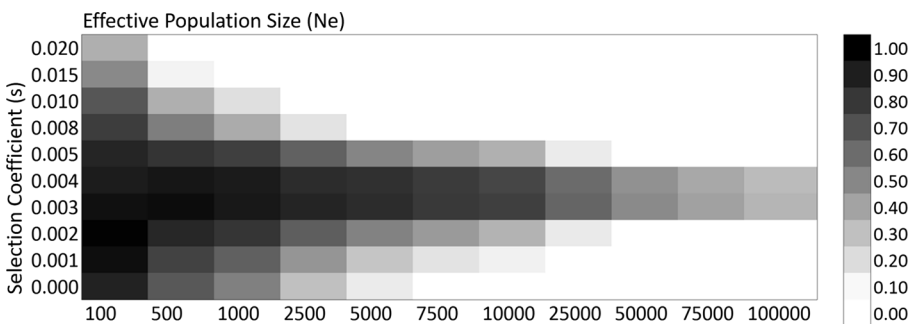
selection coefficient that best explained the observed frequency changes of the DRD -541 T allele between the Early/Middle Archaic populations and the Late Horizon/Modern Population was  $s = 0.003$  (Fig. 2).

To further explore if the observed selection acting on the DRD4 -541 T allele can be explained due to the association with reduced susceptibility to NS and the proposed negative association of the trait with increasing social complexity and agricultural lifestyle we compared the allele and genotype frequencies between the social and geographical categories.

For altitude, those populations below 2500 m.a.s.l. were classed as low altitude, those above were high altitude. Higher frequencies of the C allele were observed at low altitude. A chi squared significance test showed that the difference in allele frequency between high and low altitude was not significant ( $df = 1$ ,  $\chi^2 = 2.39$ ,  $p = 0.1221$ ). Allele frequency comparison of the populations grouped by subsistence mode (mixed foraging/horticulture vs. agriculture) revealed no significant difference ( $df = 1$ ,  $\chi^2 = 0.9$ ,  $p = 0.3428$ ) as did the grouping by mobility category ( $p = 0.09243$ ). Social organization was differentiated as state and non-state (band, tribe, or chiefdom) individuals. Using a chi squared test, the difference in allele frequencies was found to be statistically significant ( $df = 1$ ,  $\chi^2 = 4.65$ ,  $p = 0.0311$ ) while genotype difference were found to be slightly non-significant ( $df = 2$ ,  $\chi^2 = 5.02$ ,  $p = 0.0813$ ). The additional tests for genetic structure using AMOVA (Table 3) however found no significant population structure ( $P = 0.115$ ; Table 3) even though the comparison exhibited the highest level of molecular variation between the suggested divisions (5.18 %). None of the other divisions tested by AMOVA exhibited significant amounts of among group genetic variation (Table 3). Also, none of the groups showed significant deviation from Hardy-Weinberg equilibrium.

## Discussion

This study is the first test the hypothesis that polymorphisms associated with the novelty seeking personality trait are under selection due to changes in subsistence modes or social complexity from a diachronic perspective using ancient DNA analysis. Using a forward simulation approach testing genetic drift plus selection and



**Fig. 2** Heatmap illustrating the two-tailed empirical  $P$  values for the similarity between the observed allele frequencies in the Late Horizon/Modern Central Andean population and the distribution of simulated frequencies, given the priors ( $x = Ne$ ;  $y = s$ ) modeled between the oldest and youngest populations

considering bias due to low sample size we find that the DRD –521 T allele has been under light selection ( $s = 0.003$ ) in western South American populations considering a codominant allelic condition of the SNP. To explore the possible impact of other allelic conditions on our results we also tested dominant or recessive models for some prior combinations and found that neutrality is rejected for both while the selection coefficient estimated differs slightly.

However, to infer natural selection based on temporal differences in allele frequency it is necessary to assume population continuity. Previous ancient DNA studies, including the samples analyzed, found clear evidence of population continuity in western South America, or the Central Andes, in the monitored time frame based on mitochondrial, autosomal, y-chromosomal markers (Kemp et al. 2009; Carnese et al. 2010; Fehren-Schmitz et al. 2010; Fehren-Schmitz et al. 2011; Fehren-Schmitz et al. 2014; Fehren-Schmitz et al. 2015). Studies of uniparental inherited and genome-wide genetic markers link these ancient populations to the modern population of the Central Andes (Bisso-Machado et al. 2012; Lewis et al. 2007; Reich et al. 2012; Wang et al. 2007). These studies also suggest that geographic structure was established very early in South America followed by very limited gene flow into the Central Andes until more recent admixture events with other continental populations occurred after the European contact (Reich et al. 2012). To exclude a possible bias due to the inclusion of a geographically distant population into our population sets we excluded the Arroyo

**Table 3** AMOVA for the partitions of variance considering the six types of classification

Comparison	n Groups	Percentage of variation			FSC ( <i>p</i> -value)	FST ( <i>p</i> -value)	FCT ( <i>p</i> -value)
		Among groups	Among populations within-groups	Within- populations			
Chronology	7	–29.03	32.71	96.32	0.254 (0.116 + –0.003)	0.037 (0.059 + –0.002)	–0.290 (0.999 + –0.000)
Broader chrono groups	3	2.69	2.06	95.24	0.021 (0.117 + –0.003)	0.048 (0.302 + –0.005)	0.0267 (0.178 + –0.003)
Altitude	2	1.86	3.00	95.14	0.031 (0.117 + –0.003)	0.049 (0.197 + –0.003)	0.019 (0.192 + –0.004)
Mobility	2	–0.60	4.30	96.30	0.043 (0.118 + –0.003)	0.037 (0.121 + –0.003)	–0.006 (0.346 + –0.005)
Subsistence	2	1.80	3.13	95.07	0.032 (0.114 + –0.003)	0.049 (0.181 + –0.004)	0.018 (0.189 + –0.004)
Organization	2	5.18	0.86	93.96	0.009 (0.387 + –0.004)	0.060 (0.114 + –0.003)	0.052 (0.115 + –0.003)

Seco samples from the Early/Middle Archaic population and re-run the simulations confirming that selection acted on the DRD4 -521 T allele.

The 1000 Genome Project and the ALFRED database demonstrate relatively low variability for the DRD4 SNP in the global population. The populations from the Americas exhibit the highest mean frequency of the T allele (0.63) referring to the 100 Genome Project data (Altshuler et al. 2012). On the individual population level ALFRED reveals a high amount of diversity between native American groups with T allele frequencies ranging from ~0.62–0.95 (Osier et al. 2002). The Karitiana, an indigenous Amazonian group practicing a foraging lifestyle mixed with horticulture and little contact to agricultural groups, exhibit one the highest DRD4 -521 T allele frequencies. In the first instance this might contradict the outcome of our study as we would expect higher frequencies of the Novelty Seeking associated G allele in the Karitiana. However, Native South American populations show a high amount of genetic diversity not shared between geographic regions (Wang et al. 2007). The previously mentioned processes of early geographic structuring followed by isolation and drift and observed high inter-population differentiation between the Central Andes and other regions of South America are suitable to explain this observation, especially when considering that the effective population size of the initial population wave dispersing into South America was very low (Hey 2005; Reich et al. 2012; Tamm et al. 2007; Wang et al. 2007). Additionally, it has to be considered that the population size of the Karitiana is very low and gene-flow is very limited increasing the impact of genetic drift on allele frequencies (Zigtkiewicz et al. 1997). The post-European population size decline had massive impact on Amazonian groups suggesting a bottleneck effect that might have influenced the patterns of genetic diversity we observe in contemporary indigenous populations even though they had very limited contact to the outside world in the post-Colombian era (Livi-Bacci 2006; O’Fallon and Fehren-Schmitz 2011). All these observations support our hypothesis that the observed relatively low DRD4 -521 T frequency in the early Central Andean populations might be explained by founder effects, and that subsequent selection lead to an increase in the alleles frequency. This would also mean that selection for this marker was limited to specific social and environmental factors found in the research area.

The data we collected on altitude gave interesting insight to other potential sources of selective pressure. One study found an association between the T allele of the -521 C/T SNP and increased susceptibility to pre-eclampsia, an illness that would likely be fatal more often at high altitudes where considerable cardiac stress already exists (Korobochka et al. 2006). Although the difference was found not to be significant, the much higher proportion of the T allele at higher altitudes (cf. Table 2) suggests that the SNP is not associated with differential fitness at high altitudes.

In contrast to the study by Tovo-Rodrigues et al. (2010)) we also find no significant genetic differences between hunter-gatherers and agriculturalists. Considering the exploration-exploitation trade-off we would assume the comparison to be the most obvious to show significant differences when increased explorative behavior due to NS should be beneficial in settings with less resource security and susceptibility to environmental changes (Humphries et al. 2012; Yamaguchi et al. 2015). It has to be stressed that this study did not analyze the same genetic marker but looked at the 48 bp VNTR in the Exon III region of the DRD4 gene. Thus, a direct comparison of the results is not possible. The obviously low sample size of our study also might have

biased the results of the comparison due to lack of sensitivity. Associations between climate and economy have been found for the Val158Met polymorphism in the COMT gene recently suggesting that future studies trying to test this hypothesis should expand the number of genetic markers studied (Piffer 2013).

We find a slightly significant lower frequency of the C allele that is thought to contribute to the novelty seeking phenotype in samples deriving from state level societies (Wari, Inca, Modern) of the Central Andes than in all earlier populations, but no significant genetic substructure could be found. The assumed contribution of the tested SNP to the NS phenotype was determined to be about 3 % (Munafò et al. 2008). A recent study also found evidence for differential susceptibility in individuals exhibiting specific DRD4 genotypes towards behavioral phenotypes highlighting the variability in individuals' sensitivity, or responsiveness, to environmental influences, whether positive or negative (Sweitzer et al. 2013). These points and the fact that selection of course acts on phenotypes and not genotypes suggest that the selective impact detectable for the single SNP we analyzed and the chances to find significant differences between our rather general classifications is very low. Nevertheless, we found evidence for selection acting on the DRD4 -521 T allele and the comparisons we made suggest that differences in social structure and political organization might be the driving factors. This of course would make it necessary to assume that NS or risk-taking behavior is actually beneficial in human groups exhibiting lower grades of hierarchical structure and specialization. Studies of human and non-human primates have reported correlations with Dopamine receptor availability (RA) and social affiliation/status. They demonstrate that increased RA is found in individuals of higher social dominance while lower RA is observed for individuals of lower social status or desire for social dominance (Cervenka et al. 2010; Martinez et al. 2010; Yamaguchi et al. 2015). This to some extent correlates with the hypothesis that exploratory behavior and intense excitement in response to novel stimuli, impulsive decision making, and quick loss of temper and avoidance of frustration associated with the NS phenotype are less beneficial in highly structured social environments with specialized social/economic roles when the individual does not belong to a socially privileged group or has political power (Jensen et al. 1997; Williams and Taylor 2006). In the contrary, individuals living in an egalitarian society, as it is often attributed to foragers (Service, E 1962), should exhibit neither positive nor negative effects based on dopamine RA in association to social status, while increased exploratory behavior might constitute a beneficial trait to secure resources as mentioned before. This discussion is rather suggestive as our study has the power to detect selection but cannot directly relate that observation to a specific phenotype. Other factors like the function of gender in the societal context would have to be considered too. Taken together, our results neither verify nor reject the general hypothesis that it is the NS phenotype that is under selection due to specific socioeconomic factors and modes of subsistence in Central Andean populations. Other environmental and genetic factors that we could not test due to the limited sample size and geographical diversity in our study could be causal. In order to really understand the possible impact of socioeconomic and subsistence change on

genes contributing to complex behavioral phenotypes, genome wide studies of prehistoric and historic human populations involving a larger and more diverse sample would be more sufficient.

**Acknowledgments** This research was supported by a Faculty Research Grant awarded by the Committee on Research from the University of California, Santa Cruz. The authors greatly thank the Peruvian Institute for Cultural Heritage (Instituto Nacional de Cultura del Perú) for granting us the exportation permission for the sample material (Resolución Directoral Nacional No. 1346) and Gustavo Politis for granting us access to the materials from Arroyo Seco 2

## References

- Altshuler, D. M., Durbin, R. M., Abecasis, G. R., Bentley, D. R., Chakravarti, A., Clark, A. G., et al. (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature*, *491*, 56–65.
- Bellgrove, M. A., Hawi, Z., Lowe, N., Kirley, A., Robertson, I. H., & Gill, M. (2005). DRD4 gene variants and sustained attention in attention deficit hyperactivity disorder (ADHD): Effects of associated alleles at the VNTR and –521 SNP. *American Journal of Medical Genetics - Neuropsychiatric Genetics*, *136 B*(1), 81–86.
- Bisso-Machado, R., Bortolini, M. C., & Salzano, F. M. (2012). Uniparental genetic markers in south amerindians. *Genetics and Molecular Biology*, *35*(2), 365–387.
- Carnese, F. R., Mendisco, F., Keyser, C., Dejean, C. B., Dugoujon, J. M., Bravi, C. M., et al. (2010). Paleogenetical study of pre-columbian samples from pampa grande (Salta, Argentina). *American Journal Physical Anthropology*, *141*, 452–462.
- Cervenka, S., Gustavsson, J. P., Halldin, C., & Farde, L. (2010). Association between striatal and extrastriatal dopamine D2-receptor binding and social desirability. *NeuroImage*, *50*(1), 323–328.
- Chen, C., Burton, M., Greenberger, E., & Dmitrieva, J. (1999). Population migration and the variation of dopamine D4 receptor (DRD4) allele frequencies around the globe. *Evolution and Human Behavior*, *20*, 309–324.
- Coutinho, A., Valverde, G., Fehren-Schmitz, L., Cooper, A., Barreto Romero, M. I., Flores Espinoza, I., et al. (2014). AmericaPlex26: a snapshot multiplex system for genotyping the main human mitochondrial founder lineages of the americas. *PLoS One*, *9*(3): e93292.
- Doi, Y., Yamamoto, Y., Inagaki, S., Shigeta, Y., Miyaishi, S., & Ishizu, H. (2004). A new method for ABO genotyping using a multiplex single-base primer extension reaction and its application to forensic casework samples. *Leg. Med. (Tokyo)*, *6*(1344–6223 (Print)), 213–223.
- Ebstein, R. P., Novick, O., Umansky, R., Priel, B., Osher, Y., Blaine, D., et al. (1996). Dopamine D4 receptor (D4DR) exon III polymorphism associated with the human personality trait of novelty seeking. *Nature Genetics*, *12*(1), 78–80.
- Eisenberg, D. T. A., Campbell, B., Gray, P. B., & Sorenson, M. D. (2008). Dopamine receptor genetic polymorphisms and body composition in undernourished pastoralists: an exploration of nutrition indices among nomadic and recently settled ariial men of northern Kenya. *BMC Evolutionary Biology*, *8*, 173.
- Ekelund, J., Suhonen, J., Järvelin, M. R., Peltonen, L., & Lichtermann, D. (2001). No association of the –521 C/T polymorphism in the promoter of DRD4 with novelty seeking. *Molecular Psychiatry*, *6*(6), 618–619.
- Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, *131*(2), 479–491.
- Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour*, *10*, 564–567.
- Fehren-Schmitz, L., Reindel, M., Tomasto-Cagigao, E., Hummel, S., & Herrmann, B. (2010). Pre-columbian population dynamics in coastal southern Peru: a diachronic investigation of mtDNA patterns in the palpa region by ancient DNA analysis. *American Journal of Physical Anthropology*, *141*(2), 208–221.
- Fehren-Schmitz, L., Warnberg, O., Reindel, M., Seidenberg, V., Tomasto-Cagigao, E., Isla-Cuadrado, J., et al. (2011). Diachronic investigations of mitochondrial and y-chromosomal genetic markers in pre-columbian andean highlanders from south Peru. *Ann.Hum.Genet.*, *75*, 266–283.

- Fehren-Schmitz, L., Haak, W., Mächtle, B., Masch, F., Llamas, B., Tomasto Cagigao, E., et al. (2014). Climate change underlies global demographic, genetic, and cultural transitions in pre-columbian southern Peru. *Proceedings of the National Academy of Sciences of the United States of America*, 111(26), 9443–9448.
- Fehren-Schmitz, L., Llamas, B., Lindauer, S., Tomasto-Cagigao, E., Kuzminsky, S., Rohland, N., et al. (2015). A Re-appraisal of the early andean human remains from lauricocha in Peru. *PloS One*, 10(6), e0127141.
- Grieder, T. E. (1988). *La Galgada, Peru*. Austin:University of Texas Press.
- Guo, S. W., & Thompson, E. A. (1992). Performing the exact test of hardy-weinberg proportion for multiple alleles. *Biometrics*, 48(2), 361–372.
- Hattori, E., Nakajima, M., Yamada, K., Iwayama, Y., Toyota, T., Saitou, N., & Yoshikawa, T. (2009). Variable number of tandem repeat polymorphisms of DRD4: re-evaluation of selection hypothesis and analysis of association with schizophrenia. *European Journal of Human Genetics: EJHG*, 17(6), 793–801.
- Hey, J. (2005). On the number of New World founders: a population genetic portrait of the peopling of the Americas. *PLoS.Biol.*, 3, e193.
- Humphries, M. D., Khamassi, M., & Gurney, K. (2012). Dopaminergic control of the exploration-exploitation trade-off via the basal ganglia. *Frontiers in Neuroscience*, 6, 9.
- Jensen, P. S., Mrzacek, D., Knapp, P. K., Steinberg, L., Pfeffer, C., Schowalter, J., & Shapiro, T. (1997). Evolution and revolution in child psychiatry: ADHD as a disorder of adaptation. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36(12), 1672–1679 discussion 1679–1681.
- Kemp, B. M., Tung, T. A., & Summar, M. L. (2009). Genetic continuity after the collapse of the wari empire: mitochondrial DNA profiles from wari and post-wari populations in the ancient andes. *Am.J.Phys.Anthropol.*, 140(1), 80–91.
- Korobochka, R., Gritsenko, I., Gonen, R., Ebstein, R. P., & Ohel, G. (2006). Association between a functional dopamine D4 receptor promoter region polymorphism (–C521T) and pre-eclampsia: a family-based study. *Molecular Human Reproduction*, 12(2), 85–88.
- Lai, J. H., Zhu, Y. S., Huo, Z. H., Sun, R. F., Yu, B., Wang, Y. P., et al. (2010). Association study of polymorphisms in the promoter region of DRD4 with schizophrenia, depression, and heroin addiction. *Brain Research*, 1359, 227–232.
- Lewis, C. M., Lizarraga, B., Tito, R. Y., Medina, A., Martinez, R., Polo, S., et al. (2007). Mitochondrial DNA and the peopling of south america. *Human.Biology.*, 79(2), 159–178.
- Lichter, J. B., Barr, C. L., Kennedy, J. L., Van Tol, H. H., Kidd, K. K., & Livak, K. J. (1993). A hypervariable segment in the human dopamine receptor D4 (DRD4) gene. *Human Molecular Genetics*, 2(6), 767–773.
- Livi-Bacci, M. (2006). The depopulation of hispanic america after the conquest. *Population .Developmental Review.*, 32, 199–232.
- Martinez, D., Orlowska, D., Narendran, R., Slifstein, M., Liu, F., Kumar, D., et al. (2010). Dopamine type 2/3 receptor availability in the striatum and social status in human volunteers. *Biological Psychiatry*, 67(3), 275–278.
- Matthews, L. J., & Butler, P. M. (2011). Novelty-seeking DRD4 polymorphisms are associated with human migration distance out-of-Africa after controlling for neutral population gene structure. *American Journal of Physical Anthropology*, 145(3), 382–389.
- Maynard Smith, J. (1998). *Evolutionary genetics* (2nd ed., ). Oxford:Oxford University Press.
- Munafò, M. R., Yalcin, B., Willis-Owen, S. A., & Flint, J. (2008). Association of the dopamine D4 receptor (DRD4) gene and approach-related personality traits: meta-analysis and new data. *Biological Psychiatry*, 63(2), 197–206.
- Nelson, T. M., Just, R. S., Loreille, O., Schanfield, M. S., & Podini, D. (2007). Development of a multiplex single base extension assay for mitochondrial DNA haplogroup typing. *Croatica Medical Journal*, 48, 460–472.
- O’Fallon, B. D., & Fehren-Schmitz, L. (2011). Native Americans experienced a strong population bottleneck coincident with European contact. *Proceedings of the National Academy of Sciences of the United States of America*, 108: 20444–20448. doi:10.1073/pnas.1112563108
- Okuyama, Y., Ishiguro, H., Nankai, M., Shibuya, H., Watanabe, A., & Arinami, T. (2000). Identification of a polymorphism in the promoter region of DRD4 associated with the human novelty seeking personality trait. *Molecular Psychiatry*, 5(1), 64–69.
- Osier, M. V., Cheung, K.-H., Kidd, J. R., Pakstis, A. J., Miller, P. L., & Kidd, K. K. (2002). ALFRED: an allele frequency database for anthropology. *American Journal of Physical Anthropology*, 119(1), 77–83.
- Piffer, D. (2013). Correlation of the COMT Val158Met polymorphism with latitude and a hunter-gatherer lifestyle suggests culture–gene coevolution and selective pressure on cognition genes due to climate. *Anthropological Science*, 121(3), 161–171.

- Pucciarelli, H. M., Perez, S. I., & Politis, G. G. (2010). Early holocene human remains from the argentinean pampas: additional evidence for distinctive cranial morphology of early south americans. *American Journal of Physical Anthropology*, 143(2), 298–305.
- Reich, D., Patterson, N., Campbell, D., Tandon, A., Mazieres, S., Ray, N., et al. (2012). Reconstructing native american population history. *Nature*, 488(7411), 370–374.
- Reindel, M. (2009). Life at the edge of the desert - archaeological reconstruction of the settlement history in the valleys of palpa, Peru. In M. Reindel, & G. Wagner (Eds.), *New Technologies for Archaeology: Multidisciplinary Investigations in Palpa and Nasca, Peru* (pp. 439–462). Berlin-Heidelberg-New York: Springer.
- Ronai, Z., Barta, C., Guttman, A., Lakatos, K., Gervai, J., Staub, M., & Sasvari-Szekely, M. (2001). Genotyping the -521C/T functional polymorphism in the promoter of region of dopamine D4 receptor (DRD4) gene. *Electrophoresis*, 22(6), 1102–1105.
- Schinka, J. A., Letsch, E. A., & Crawford, F. C. (2002). DRD4 and novelty seeking: results of meta-analyses. *American Journal of Medical Genetics Neuropsychiatric Genetics*, 114(6), 643–648.
- Service, E. (1962). *Primitive social organisation*. New York:Random House.
- Sweitzer, M. M., Halder, I., Flory, J. D., Craig, A. E., Gianaros, P. J., Ferrell, R. E., & Manuck, S. B. (2013). Polymorphic variation in the dopamine D4 receptor predicts delay discounting as a function of childhood socioeconomic status: evidence for differential susceptibility. *Social Cognitive and Affective Neuroscience*, 8(5), 499–508.
- Tamm, E., Kivisild, T., Reidla, M., Metspalu, M., Smith, D. G., Mulligan, C. J., et al. (2007). Beringian standstill and spread of Native American founders. *PLoS.ONE.*, 2(1932–6203 (Electronic) LA - eng PT - Journal Article), e829.
- Thomson, C. J., Hanna, C. W., Carlson, S. R., & Rupert, J. L. (2013). The -521 C/T variant in the dopamine-4-receptor gene (DRD4) is associated with skiing and snowboarding behavior. *Scandinavian Journal of Medicine & Science in Sports*, 23(2), e108–e113.
- Thomson, C. J., Rajala, A. K., Carlson, S. R., & Rupert, J. L. (2014). Variants in the dopamine-4-receptor gene promoter are not associated with sensation seeking in skiers. *PloS One*, 9(4), e93521.
- Tovo-Rodrigues, L., Callegari-Jacques, S. M., Petzl-Erler, M. L., Tsuneto, L., Salzano, F. M., & Hutz, M. H. (2010). Dopamine receptor D4 allele distribution in amerindians: a reflection of past behavior differences? *American Journal Physiology Anthropology*, 143, 458–464.
- Voight, B. F., Adams, A. M., Frisse, L. A., Qian, Y., Hudson, R. R., & Di Rienzo, A. (2005). Interrogating multiple aspects of variation in a full resequencing data set to infer human population size changes. *Proceedings of the National Academy of Sciences of the United States of America*, 102(51), 18508–18513.
- Wang, S., Lewis, C. M., Jakobsson, M., Ramachandran, S., Ray, N., Bedoya, G., et al. (2007). Genetic Variation and Population Structure in Native Americans. *PLoS.Genet*, 3(e185).
- Wilde, S., Timpson, A., Kirsanow, K., Kaiser, E., Kayser, M., Unterländer, M., et al. (2014). Direct evidence for positive selection of skin, hair, and eye pigmentation in europeans during the last 5,000 y. *Proceedings of the National Academy of Sciences of the United States of America*, 111(13), 4832–4837.
- Williams, J., & Taylor, E. (2006). The evolution of hyperactivity, impulsivity and cognitive diversity. *Journal of the Royal Society, Interface/the Royal Society*, 3(8), 399–413.
- Yamaguchi, Y., Lee, Y.-A., & Goto, Y. (2015). Dopamine in socioecological and evolutionary perspectives: implications for psychiatric disorders. *Frontiers in Neuroscience*, 9, 219.
- Zigtkiewicz, E., Yotova, V., Jarnik, M., Korab-Laskowska, M., Kidd, K. K., Modiano, D., et al. (1997). Nuclear DNA diversity in worldwide distributed human populations. *Genetic*, 205(1–2), 161–171.