

Gören, Erkan

**Working Paper**

## The Biogeographic Origins of Novelty-Seeking Traits

Oldenburg Discussion Papers in Economics, No. V-366-14

**Provided in Cooperation with:**

University of Oldenburg, Department of Economics

*Suggested Citation:* Gören, Erkan (2014) : The Biogeographic Origins of Novelty-Seeking Traits, Oldenburg Discussion Papers in Economics, No. V-366-14, University of Oldenburg, Department of Economics, Oldenburg

This Version is available at:

<https://hdl.handle.net/10419/105054>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



# Oldenburg Discussion Papers in Economics

The Biogeographic Origins of Novelty-Seeking Traits

Erkan Gören

V – 366 – 14

May 2014

**Department of Economics**  
University of Oldenburg, D-26111 Oldenburg

# THE BIOGEOGRAPHIC ORIGINS OF NOVELTY-SEEKING TRAITS\*

ERKAN GÖREN †

Department of Economics, Carl von Ossietzky University Oldenburg

THIS VERSION: MAY 2014

## Abstract

This paper empirically investigates the biogeographic determinants of the human DRD4 exon III locus, a particular gene variant associated with the human personality trait of novelty-seeking behavior. Providing a novel compilation of worldwide DRD4 exon III allele frequencies in a large sample of indigenous populations around the world, this study employs population-specific biogeographic characteristics using high-resolution geospatial data. The estimates suggest that migratory distance from East Africa naturally selects for specific novelty-seeking traits, even controlling for a broad range of biogeographic determinants. Notably, land suitability for pastoral nomadism is significantly related to DRD4 exon III diversity. This result provides further credence to the general observation that novelty-seeking traits are quite common in nomadic populations, explaining why some societies failed to settle and to develop centralized states.

**Keywords:** Novelty-Seeking Behavior, Entrepreneurial Traits, Biogeography,  
Out of Africa Hypothesis, Gene-Culture Co-Evolution, Natural Selection

**JEL Classification Numbers:** N50, O10, Z10

---

\*I would like to thank Jürgen Bitzer and the seminar participants at the Carl von Ossietzky University Oldenburg 2013 for useful comments and suggestions. All remaining errors are my own.

†Carl von Ossietzky University Oldenburg, Department of Economics, Campus Haarentor, Building A5, 26111 Oldenburg, Germany, Tel.: +49-441-798-4292, e-mail: erkan.goeren@uni-oldenburg.de.

## 1 Introduction

The role of geography in development has attracted considerable attention in the recent economic literature. There is broad consensus about the role of geographical factors such as latitude and climate in economic development, but the concrete mechanisms by which geography affects development are still inadequately understood and are the subject of some controversy. Direct effects of geography on development include reduced productivity, while its indirect effects are transmitted through the distribution of specific human traits, institutions, cultural attitudes, and biogeographic endowments. A vast literature has investigated the relationship between economic development and geography (Bloom and Sachs, 1998; Gallup et al., 1999), temperature and climate (Masters and McMillan, 2001; Sachs, 2001), the health environment (Sachs, 2001; Gallup and Sachs, 2001; Sachs and Malaney, 2002), agricultural productivity (Gallup and Sachs, 2000), natural resources (Sachs and Warner, 2001) and biogeographic endowments (Diamond, 1997; Olsson and Hibbs, 2005). An intriguing study in Diamond's *Guns, Germs and Steel* argues that the dominance of Western countries over areas in the new world, is due to the favorable biogeographic endowments of the former. The large size of Eurasia, its East-West orientation, and the number of plants and animals suitable for domestication resulted in an earlier adoption of agricultural practices (Neolithic transition) and the diffusion of agricultural innovations across space. The accompanying explosion of populations, innovations, and ideas on the one hand and early exposure to diseases accompanied with animal domestication on the other explains why Europeans were able to conquer some areas of the new world so effectively.

The main argument of the present study is that the differences that have emerged between populations since the Neolithic transition may account for the differences in development between countries. For example, the transition from hunter-gatherer economy to agricultural production occurred at different times in different regions of the world, affecting the amount of time it took populations to adapt to environmental changes. In line with previous findings on some of the deep-rooted factors that have influenced national development paths, this work empirically analyzes the biogeographic origins of a specific human dopamine gene related to novelty-seeking or exploratory behavior and its interaction with the cultural value of pastoral nomadism. Theoretical models of gene-culture co-evolution would predict that the high frequency of novelty-seeking traits in populations is substantially correlated with the production mode of pastoral nomadism, therefore providing an explanation of why some populations have failed to settle and to develop centralized states, resulting in economic backwardness. In contrast to agricultural practices, pastoral nomadism is technologically adapted to the natural environment to utilize

marginal resources effectively in dry-land areas that are otherwise unsuitable for sedentary agriculture (Johnson, 1969).<sup>1</sup> This mode of production enables people to handle environmental risk through a wide range of mechanisms including food diversification and physical storage, regional mobility, and exchange (Koocheki and Gliessman, 2005). From a historical perspective, pastoral nomadism emerged at approximately the same time as animal domestication and sedentary agriculture. It is viewed as a parallel production mode to agricultural intensification using improved irrigation systems that is based exclusively on animal domestication (McCorkle, 1992; Koocheki and Gliessman, 2005). However, while some nomads gradually developed sedentary communities with complex social institutions, others preserved their traditional way of life (Khazanov, 2003). The discovery of overseas trade routes and the great technological innovations since the industrial revolution, for example, have sharply decreased the military superiority and trade importance of nomadic lifestyles, resulting in the relative economic backwardness of such societies.<sup>2</sup> This would provide evidence of the complex gene-culture co-evolution process between the human trait of novelty-seeking and the effective intergenerational transmission of nomadic lifestyles.

From an evolutionary perspective, novelty-seeking behavior could be seen as means to expand a society's knowledge frontier. In a study on evolved altruism, Williams and Taylor (2006) assert that in societies with non-explorative behavior, people do not explore new food variants adequately and therefore suffer from malnutrition. On the other hand, explorative behavior by identical individuals is inherently inefficient because it results in redundant information. Therefore, only in societies made up primarily of non-explorative individuals can a small number of explorative individuals push the knowledge frontier of the overall society further, with the high costs being borne by a small group of individuals. More recently, Galor and Michalopoulos (2012) developed an evolutionary growth model, arguing that novelty-seeking traits, which they interpret as entrepreneurial traits, played a significant role in the process of economic development. They elaborate that in early stages of development, novelty-seeking, growth-promoting traits created an evolutionary advantage, and that a large number of such individuals in a society accelerated technological progress. But in advanced stages of economic development, the evolutionary advantage of novelty-seeking traits reversed and risk-averse individuals gained an

---

<sup>1</sup>See also Salzman (2002) for some general observations on nomadic lifestyles based on research on nomadic populations residing in Iran.

<sup>2</sup>Khazanov (2003) provides an excellent discussion of the great social and military achievements of nomadic empires that were adopted by sedentary societies and that became precursors to state formation, the division of labor, and industrial production.

evolutionary advantage, thus decreasing technological progress.

The importance of novelty-seeking traits in society and the fact that their distribution is strongly influenced by biogeographic factors is widely acknowledged by population geneticists. However, there exists no study, to the best of the author's knowledge, that investigates the influence of biogeographic factors on the distribution of novelty-seeking traits within the population in detail. The only existing studies on this topic, Chen et al. (1999) and Matthews and Butler (2011), examine the influence of migratory distance from East Africa but fail to control for biogeographic factors as a possible source of human adaptiveness in resource-depleted, time-critical, and frequently changing environments. They do not consider important biogeographic indicators such as land suitability for agriculture, the fraction of land allocated to pasture, and terrain ruggedness. Furthermore, these studies are based on limited data sets that include only 39 and 18 populations, respectively.

This is the starting point for the present paper. By creating a new data set that covers the distribution of worldwide allele frequencies of the DRD4 exon III gene, this study reaches a broader coverage of populations across the world. Furthermore, this unique data set permits the in-depth analysis of additional biogeographic factors. The biogeographic indicators are constructed specifically for the location covered by each population using high-resolution geospatial data. Given the DRD4 exon III allele frequencies, it is possible to construct population-specific expected heterozygosity measures for this particular gene variant. The baseline results confirm the high predictive power of migratory distance from East Africa in previous studies. The estimates suggest that DRD4 exon III expected heterozygosity increases with migratory distance from East Africa. Although the adaptive nature of specific novelty-seeking traits in migratory societies appears important in areas unsuitable for sedentary practices, a thorough empirical investigation of this important issue is still lacking. In contrast to previous work, this study highlights the importance of DRD4 exon III polymorphism in diverse environments and among genetically distinct populations. The estimates reveal that migratory distance from East Africa maintains its predictive power, even when controlling for a broad set of biogeographic factors and among different specifications. It is worth mentioning that land suitability for pastoral nomadism is significantly correlated with DRD4 exon III expected heterozygosity.<sup>3</sup> The same picture holds when focusing on the fraction of long allele types as the key dependent variable,

---

<sup>3</sup>This hypothesis is reinforced by a recent study among nomadic and recently settled Ariaal men of northern Kenya. Specifically, Eisenberg et al. (2008) have shown that novelty-seeking behavior was associated with better nutritional indices among the nomadic group, but worse in the recently settled group.

measuring the extent of novelty-seeking traits in society. In addition, variation in terrain ruggedness is negatively related to DRD4 exon III diversity. This observation provides evidence that terrain variability acts as a natural barrier to migratory behavior. Furthermore, land suitability for agriculture is significantly related to the 2-repeat allele. This particular allele may be adapted in populations practicing agriculture for a long period. Indeed, this particular allele is widespread in Asia, which experienced the Neolithic transition much earlier than other parts of the world.

Linking DRD4 exon III to biogeography has important implications for as yet unaddressed empirical research questions. For example, biogeographic factors can serve as qualified instrumental variables in the analysis of the indirect impacts of geography on economic development through the human trait associated with novelty-seeking behavior. Furthermore, from an econometric perspective, biogeographic factors are purely exogenous when investigating the impact of populations on economic development, and are therefore able to at least partially overcome potential endogeneity bias arising in data on post-1500 migration flows. This can provide a more thorough understanding of the effects of biogeography on economic outcomes, and when considering the distribution of specific human traits, it can help to disentangle the direct and indirect effects of biogeography.

The remainder of the paper is organized as follows. Section 2 provides a discussion of the hypothesized association between novelty seeking and a particular dopamine gene that received considerable attention from population geneticists. Section 3 presents the data descriptions on DRD4 exon III allele frequencies across regions and their relationship to migratory distance from East Africa. Section 4 reports the data sources and methods on the construction of population-specific biogeographic indicators. Section 5 discusses the empirical results on the biogeographic origins of observed DRD4 exon III allele frequencies across populations using high-resolution geospatial data on land quality and terrain ruggedness. Finally, Section 6 concludes by summarizing the main results and providing areas for future research.

## 2 Population Studies of DRD4 Exon III and Novelty Seeking

Because the use of genetic data in economics is still relatively rare, a few basic definitions are presented first to aid in understanding the subsequent parts of the paper. Genetic information is present in every cell in the form of *chromosomes*. Chromosomes are made up of long threads of relatively small nucleotides. These long threads are the carriers of genetic information and are called *DNA*. The DNA thread is extremely long for each specific chromosome. However, one can

recognize shorter segments of a DNA thread, called *genes*. The genes are responsible for specific activities in the cells. Each gene can take many different forms, which are called the *alleles* of the gene. For example, the specific gene that is responsible for eye color may contain many different alleles (like blue, green, and brown eye color). When population geneticists detect two or more alleles of a specific gene within a population (e.g., the gene for eye color), then this gene is called *polymorphic*.<sup>4</sup>

The human DRD4 gene is located at the distal tip of the short arm of chromosome eleven and contains of four encoded regions, called exons. The most extensive polymorphism is found in the third exon. This polymorphism occurs in a 48-base pair (bp) variable number of tandem repeats (VNTR), ranging from 2 to eleven repeats, with 2, 4, and 7 repeats being the most common alleles across populations (Van Tol et al., 1991, 1992; Lichter et al., 1993). Dopamine is an important neurotransmitter and affects the control of locomotion, endocrine function, reward, cognition, and emotion (Oak et al., 2000). The view that human personality has a heritable component has been debated by Plomin et al. (1994). Most of the evidence pertaining to the influence of genetics on complex human behaviors results from twin studies on personality, vocational interests, scholastic achievement, and cognitive abilities (Bouchard, 1994). Since then, molecular genetics has tried to identify specific genes or gene variants that influence human behavior. Progress in the classification of personality was first made by Cloninger (1987). The author categorized four different personality traits: novelty seeking, harm avoidance, reward dependence, and persistence. Especially the first personality trait of novelty seeking is a natural candidate to be biologically influenced by dopamine genes (Paterson et al., 1999). Novelty-seeking individuals are characterized as impulsive, exploratory, fickle, excitable, quick-tempered, and extravagant, whereas those with low novelty-seeking scores tend to be reflective, rigid, loyal, stoic, slow-tempered, and frugal (Cloninger et al., 1991; Ebstein et al., 1996). Two influential studies on the genetic association between a specific DRD4 exon III allele (namely, the 7-repeat) and novelty-seeking behavior by Benjamin et al. (1996) and Ebstein et al. (1996) triggered a wide range of empirical investigations linking human behavior to specific genes. Genetic studies have been carried out since then to find a possible association between DRD4 exon III and novelty seeking, neuropsychiatric diseases, substance abuse, Attention Deficit Hyperactivity Disorder (ADHD), and mood disorders, with partly inconclusive results.<sup>5</sup>

---

<sup>4</sup>Cavalli-Sforza et al. (1994) provide a thorough technical discussion of genetic material and the worldwide distribution of human genes.

<sup>5</sup>See also the contributions in Oak et al. (2000) and Savitz and Ramesar (2004) for a general overview of DRD4 genetic association studies and Munafò et al. (2008) for a large-scale meta-study of the human DRD4



Genetic data on the worldwide distribution of DRD4 exon III allele frequencies among different populations have been compiled from a number of published molecular genetic studies, drawing mainly from studies reporting allele frequencies from healthy (non-psychiatric) individuals. The most reliable data on DRD4 exon III allele frequencies consist of 1,327 individuals from 36 various populations compiled by Chang et al. (1996). The populations analyzed in this study are historically native to their current geographical location and have not undergone recent genetic mixing with other populations. This fact is particularly important because genetic mixing of populations since the post-Columbian era has introduced a serious causality problem between DRD4 exon III allele frequencies on the one hand and historical migration routes and biogeographic variables on the other.<sup>6</sup> Another important data source for DRD4 exon III allele frequencies is the online Allele Frequency Database (ALFRED) supported by the U.S. National Science Foundation.<sup>7</sup> Furthermore, these data sources were additionally extended with DRD4 exon III allele frequencies from compiled molecular genetic studies in order to obtain a greater coverage of populations for the empirical analysis.<sup>8</sup>

### 3 Migratory Distance and DRD4 Exon III Polymorphism

Recent empirical and theoretical studies suggest that DRD4 exon III polymorphism underwent a rare mutational event of positive selection at some point after the exodus from East Africa. Specifically, Chen et al. (1999) showed the explanatory power of migratory distance from Africa for the 7-repeat allele in DRD4 exon III for various populations around the world. Furthermore, Wang et al. (2004) proposed a model of selection acting on the DRD4 exon III gene with known

---

gene's association with novelty-seeking behavior. Paterson et al. (1999) provide an excellent critical analysis of the putative genetic association of DRD4 with observed human behaviors.

<sup>6</sup>It should be noted that there exists another reliable data source for human alleles: the Human Genome Diversity Cell Line Panel from the Human Genome Diversity Project-Centre d'Etude du Polymorphisme Humain (HGDP-CEPH) consists of 53 indigenous populations around the globe. Nevertheless, the present study cannot employ this additional source of genetic data because allele frequencies for the DRD4 exon III gene were not sequenced for this population sample. For a more detailed discussion of this database, the interested reader is referred to Cann et al. (2002).

<sup>7</sup>ALFRED provides allele frequencies and information for a wide range of DNA polymorphism and populations. The main objective of ALFRED is the storage of genetic data from the scientific literature. By this means, it acts as a central database for future anthropological and evolutionary research projects. See also Rajeevan et al. (2005) for an introduction to ALFRED.

<sup>8</sup>Additional details regarding the sampled populations and data sources can be found in the separate Appendix to this paper, which is available from the author upon request.

biochemical and physiological differences between receptor variants. Their argumentation is that different blunted responses of DRD4 exon III alleles to dopamine levels may account for migratory behavior out of Africa, resulting in different human personality types prevailing under different environmental conditions.<sup>9</sup> It has been suggested by Jensen et al. (1997) and Wang et al. (2004) that resource-depleted, time-critical, or rapidly changing environments might select for individuals with "response-ready" adaptations, whereas resource-rich, time-optimal, or little-changing environments might select against such adaptations. Perhaps individuals with a strong inclination toward "response-ready" adaptations, such as novelty-seeking behavior, drove the wave of migration out of Africa (Ding et al., 2002).

This study utilizes the complete genetic information present in DRD4 exon III allele frequencies to capture the fact that some alleles may be associated with others. Furthermore, as noted above, the 2- and 7-repeat alleles are genetically and functionally related. The focus on a single allele may bias the results when analyzing the hypothesized association of novelty-seeking behavior with biogeographic variables. To avoid this pitfall, the present study makes use of a commonly used measure in the genetics literature: an index called *expected heterozygosity* that population geneticists utilize to measure the degree of diversity of a single gene within a particular population. Given the DRD4 exon III allele frequencies, this measure of genetic diversity is defined as follows:

$$H_k^{DRD4} = 1 - \sum_{a \in S_{DRD4}} p_{ak}^2, \quad (1)$$

where  $H_k^{DRD4}$  is the expected heterozygosity of population  $k$  with respect to the DRD4 exon III locus,  $S_{DRD4}$  refers to the set of observed DRD4 exon III alleles in population  $k$ , and  $p_{ak}$  is the frequency for the  $a$ th DRD4 exon III allele for this particular population. This measure corresponds to the probability that two randomly selected individuals from the *same* population will have different DRD4 exon III allele variants.

Empirical support for the relationship between geographic distance and variation in DRD4 exon III allele frequencies was first uncovered by Chen et al. (1999) and further confirmed by

---

<sup>9</sup>In their model, the 4-repeat variant is seen as functioning well evolutionarily, whereas the 2- and 7-repeat variants are suboptimal yet confer a behavioral advantage in some environments. Furthermore, the last two variants were shown to be genetically and functionally related, each exhibiting blunted responses to dopamine signals. Furthermore, the authors suggest that the inhibitory neurons utilizing the 2- and 7-repeat variants requires a higher dopamine level than the 4-repeat to allow for "normal" functioning. The authors take these increased dopamine levels as evidence of novelty-seeking behavior. The 2-repeat variant is hypothesized to produce intermediate human behaviors relative to individuals exhibiting the 4- and 7-repeat variants.

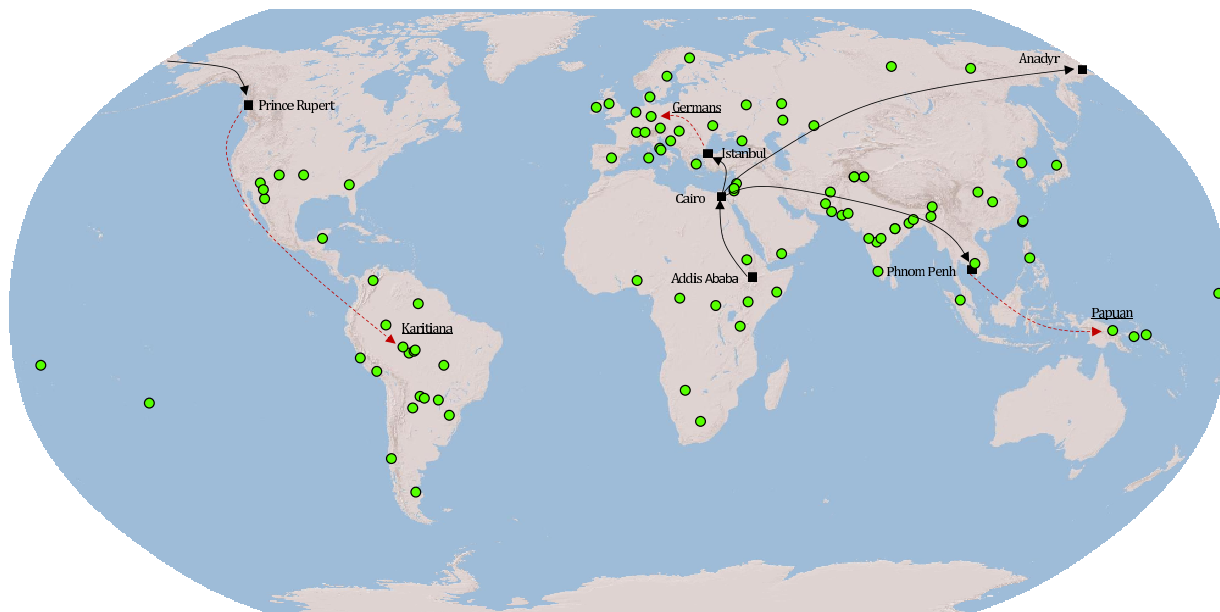
Matthews and Butler (2011).<sup>10</sup> The main hypothesis was that long-distance migration selects for particular gene variants because they reward exploratory behavior. This might have been an evolutionary advantage to populations that constantly moved through unfamiliar and frequently changing environments. To derive the migratory distance from East Africa, the great circle or geodesic calculation approach outlined in Ramachandran et al. (2005) was employed for the sampled indigenous populations considered in this study. Most of the information on the historical geographical location of the populations was provided by ALFRED. For populations where ranges of latitude and longitude coordinates are available, the mean of the latitudes and the mean of the longitudes were used. Specifically, Ramachandran et al. (2005) calculated great circle distances using Addis Ababa as starting point in East Africa, further restricting the pre-historic human migration routes to five obligatory, land-restricted waypoints, namely Cairo (Egypt), Istanbul (Turkey), Phnom Penh (Cambodia), Anadyr (Russia), and Prince Rupert (Canada). For example, the settlement of South America is then computed as the sum of the geodesic distance from Addis Ababa via Cairo, Anadyr, and Prince Rupert, plus the distance from Prince Rupert to the location of the group. Figure 1 illustrates the worldwide distribution of indigenous populations used in the empirical analysis along with the intermediate waypoints out of East Africa. Each green colored circle corresponds to a particular population.<sup>11</sup> Figure 2 provides a first insight into the relationship between DRD4 exon III expected heterozygosity and migratory distance from East Africa.<sup>12</sup> Greater distance from the cradle of humankind in East Africa is accompanied by higher DRD4 exon III expected heterozygosity measures. The linear regression depicted in Figure 2 provides a significant positive impact of migratory distance from East Africa on observed DRD4 exon III expected heterozygosity. For every 1,000 km distance from Addis Ababa, the expected heterozygosity increases by about 0.0069 (0.0018), with standard errors, clustered across regions, shown in parentheses. The  $R^2$  for this linear re-

---

<sup>10</sup>In a more general setting pertaining to the relationship between genetic and geographic distance, Ramachandran et al. (2005) have shown that measures of pairwise genetic distances and the expected heterozygosity of the 53 HGDP-CEPH sample is highly correlated with their measure of migratory distance from East Africa. They took this observation as evidence of a serial founder effect starting from Addis Ababa, Ethiopia (9N, 39E), through five obligatory (land-restricted) waypoints: Cairo, Egypt (30N, 31E); Istanbul, Turkey (41N, 28E); Phnom Penh, Cambodia (11N, 104E); Anadyr, Russia (64N, 177E); and Prince Rupert, Canada (54N, 130W). Following their study, the current analysis argues for the natural selection of specific DRD4 exon III gene variants since the exodus out of Africa.

<sup>11</sup>Detailed geographical information on the sampled populations is provided in Tables 7 and 8 in the Appendix.

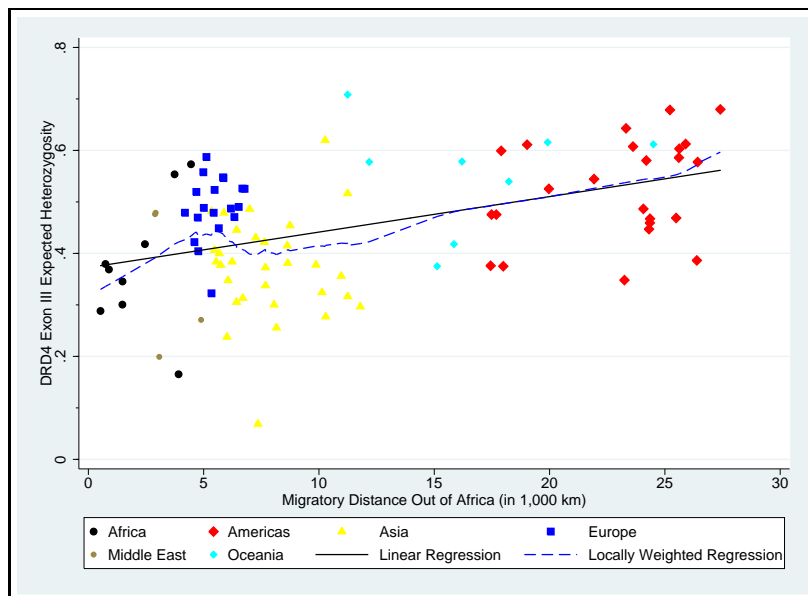
<sup>12</sup>Each point in Figure 2 corresponds to a particular population. Multiple observations for populations from compiled molecular genetic studies were weighted with their respective sample sizes.



**Figure 1:** WORLDWIDE DISTRIBUTION OF SAMPLED POPULATIONS AND MIGRATORY PATHS FROM ADDIS ABABA

gression is about 0.21. This means that the geographic distance from East Africa alone predicts about 21% in the variation of DRD4 exon III expected heterozygosity. The remainder of the non-explained variation is therefore due to random genetic drift, population-specific histories, and environmental conditions.<sup>13</sup> Figure 3 provides a graphical illustration of DRD4 exon III allele frequencies across regions. The vertical axis shows the allele frequency for each region, whereas the horizontal axis lists the variable number of tandem repeats. Descriptive statistics are shown in Table 1. Consistent with previous findings in Chang et al. (1996), a clear geographic pattern of DRD4 exon III allele frequencies is apparent. The 4-repeat allele is observed in almost all populations with frequencies ranging from 0.149 to 0.964, while the 7-repeat allele was the second most observed allele, predominant in American populations but nearly absent in Asian populations. The 2-repeat allele was common in Asian and Oceanic populations but relatively uncommon in Africa. DRD4 exon III expected heterozygosity measures are shown in the last column of Table 1. Furthermore, the regional pattern provides additional information on DRD4 exon III diversity across populations within regions. While the standard deviation of DRD4 exon III expected heterozygosity is similar across populations in the Americas, Asia, and Oceania, higher ranges on this measure can be found in African and Middle Eastern populations. European populations show the lowest range in DRD4 exon III expected heterozygosity due to their high genetic resemblance. The remaining repeat-number alleles show no clear geographic

<sup>13</sup>Especially the last point will undergo a thorough empirical investigation in later parts of the paper.



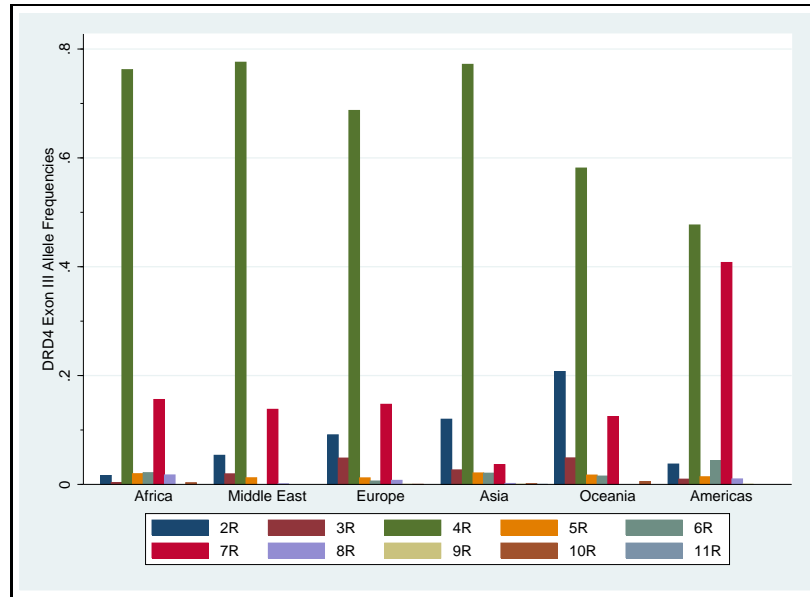
**Figure 2:** RELATIONSHIP BETWEEN DRD4 HETEROZYGOSITY AND MIGRATORY DISTANCE FROM EAST AFRICA

pattern. Overall, the above discussion shows that higher values in DRD4 exon III expected heterozygosity have to be interpreted accordingly. Wang et al. (2004) noted that the 4-repeat allele is the ancestral allele, where especially the 2- and 7-repeat variants have emerged through positive selection. For example, in American populations, a clear movement of the 4-repeat allele toward other allele variants, especially the 7-repeat, is obvious. This shift to other allele variants is in turn responsible for the high polymorphic nature of the DRD4 exon III gene across populations.

In addition to the use of migratory distance out of East Africa, this study also further explores the biogeographic origins of the DRD4 exon III gene in depth. An intriguing strand of research along the lines of Livingstone (1958), Wiesenfeld (1967), and Durham (1982) has emphasized the importance of biogeographic factors on the natural selection of specific human traits and their co-evolution with cultural values.<sup>14</sup> More recently, Cook (2011, 2013) reported evidence of a predominant role of biogeographic endowments in the spread and adaptation of human genes responsible for the ability to digest milk and immunity to infectious crowd diseases.<sup>15</sup>

<sup>14</sup>For example, as proposed by Livingstone (1958), and further analyzed by Wiesenfeld (1967), there exists a high association between the frequency of a specific sickle cell trait in Sub-Saharan African populations and the population's dependence on agriculture. The reason is that land cultivation in humid tropical areas promotes the spread of mosquitoes bearing malaria falciparum. Those individuals who are carriers of an intermediate sickle cell trait developed a high resistance to malaria over time.

<sup>15</sup>Infectious crowd diseases are defined as those capable of persisting only in large human populations. Wolfe et al. (2007) suggest that a number of temperate diseases (diphtheria, influenza A, measles, mumps, pertussis, rotavirus, smallpox, and tuberculosis) probably spread to humans from domesticable animals. Therefore, the rise



**Figure 3:** DRD4 EXON III ALLELE FREQUENCIES ACROSS REGIONS

In a related context, Michalopoulos (2012) analyzed the determinants of ethnolinguistic diversity within countries, arguing that differences in land endowments gave rise to location-specific human capital, leading to genetic drift and the formation of localized ethnic groups. The following section therefore analyzes the association of DRD4 exon III with biogeographic indicators, utilizing high-resolution geospatial data.

## 4 Data Sources on Biogeographic Indicators

This section discusses the data sources used in the construction of various biogeographic indicators. The Center for Sustainability and the Global Environment (SAGE) at the University of Wisconsin provides data on land suitability for agriculture, patterns of global land use, population density, and elevation at a very fine resolution of 0.5 decimal degrees latitude  $\times$  longitude. The global data on land suitability for agriculture were estimated by Ramankutty et al. (2002) to investigate the sensitivity of cultivable land areas to possible climate change. The estimated index of land suitability for agriculture represents the probability that a particular grid cell will be cultivated.<sup>16</sup> This index was obtained by combining existing relationships between croplands,

---

of agriculture and the frequent contact of farmers with domesticable animals played an important role in the evolution of animal pathogens into human pathogens. It is hypothesized that the early exposure of populations to these diseases fostered the natural selection of disease-related genes, leading in turn to higher life expectancy in the contemporary era.

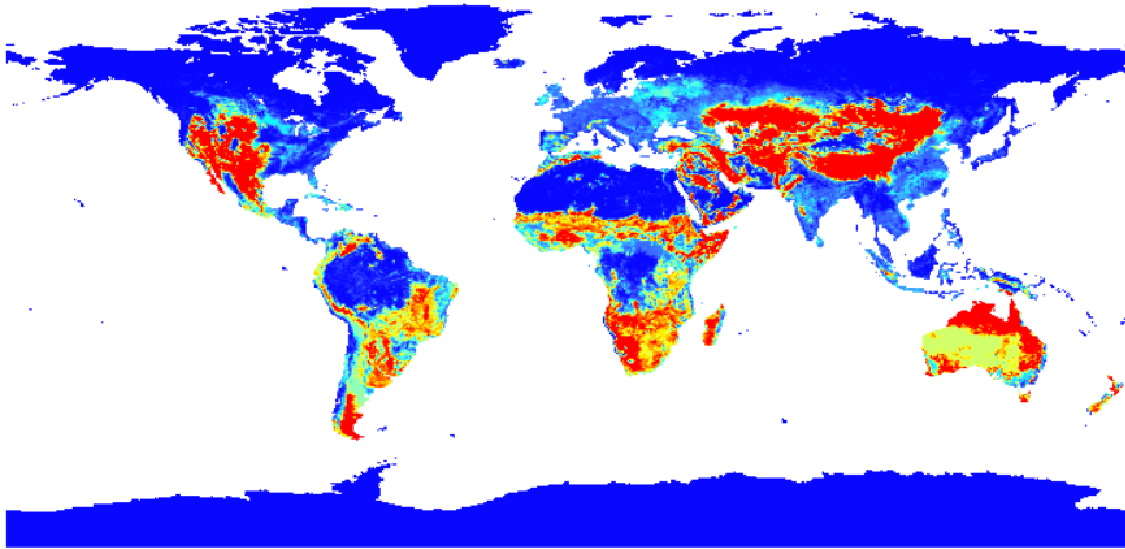
<sup>16</sup>This index does not control for agricultural productivity of a particular grid cell, but instead for whether the characteristics of the land area make it suitable for cultivation.

**Table 1:** DRD4 Exon III Allele Frequencies and Expected Heterozygosity Across Regions.

	2R	3R	4R	5R	6R	7R	8R	9R	10R	11R	$H^{DRD4}$
<b>Africa</b>											
Mean	0.016	0.004	0.762	0.020	0.022	0.156	0.017	0.000	0.003	0.000	0.377
S.D.	0.021	0.008	0.099	0.024	0.031	0.073	0.036	0.000	0.006	0.000	0.128
Min.	0.000	0.000	0.607	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.165
Max.	0.050	0.023	0.909	0.063	0.091	0.268	0.112	0.000	0.016	0.000	0.573
<i>Total Number of Alleles: 1,092</i>											
<i>Number of Populations: 9</i>											
<b>Americas</b>											
Mean	0.037	0.010	0.477	0.014	0.044	0.408	0.010	0.000	0.000	0.000	0.526
S.D.	0.064	0.021	0.181	0.020	0.054	0.197	0.036	0.001	0.000	0.000	0.099
Min.	0.000	0.000	0.149	0.000	0.000	0.114	0.000	0.000	0.000	0.000	0.348
Max.	0.250	0.083	0.773	0.074	0.179	0.781	0.179	0.007	0.000	0.000	0.680
<i>Total Number of Alleles: 2,024</i>											
<i>Number of Populations: 24</i>											
<b>Asia</b>											
Mean	0.120	0.027	0.772	0.021	0.021	0.036	0.002	0.000	0.001	0.000	0.373
S.D.	0.067	0.031	0.078	0.019	0.019	0.041	0.004	0.002	0.007	0.000	0.100
Min.	0.023	0.000	0.520	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.069
Max.	0.320	0.115	0.964	0.095	0.060	0.125	0.015	0.009	0.040	0.001	0.620
<i>Total Number of Alleles: 14,696</i>											
<i>Number of Populations: 31</i>											
<b>Europe</b>											
Mean	0.091	0.048	0.687	0.012	0.006	0.147	0.007	0.000	0.000	0.000	0.489
S.D.	0.027	0.022	0.050	0.010	0.012	0.049	0.008	0.001	0.000	0.000	0.061
Min.	0.048	0.015	0.602	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.322
Max.	0.154	0.100	0.808	0.034	0.055	0.198	0.030	0.002	0.001	0.000	0.587
<i>Total Number of Alleles: 19,971</i>											
<i>Number of Populations: 19</i>											
<b>Middle East</b>											
Mean	0.054	0.019	0.776	0.012	0.000	0.138	0.001	0.000	0.000	0.000	0.356
S.D.	0.008	0.016	0.112	0.017	0.000	0.113	0.002	0.000	0.000	0.000	0.143
Min.	0.046	0.000	0.667	0.000	0.000	0.038	0.000	0.000	0.000	0.000	0.199
Max.	0.064	0.038	0.892	0.037	0.000	0.269	0.004	0.000	0.000	0.000	0.479
<i>Total Number of Alleles: 808</i>											
<i>Number of Populations: 4</i>											
<b>Oceania</b>											
Mean	0.207	0.049	0.581	0.017	0.015	0.125	0.000	0.000	0.005	0.000	0.553
S.D.	0.161	0.039	0.123	0.018	0.019	0.105	0.000	0.000	0.015	0.000	0.109
Min.	0.000	0.000	0.417	0.000	0.000	0.026	0.000	0.000	0.000	0.000	0.375
Max.	0.409	0.099	0.752	0.043	0.043	0.304	0.000	0.000	0.042	0.000	0.708
<i>Total Number of Alleles: 1,002</i>											
<i>Number of Populations: 8</i>											

**Notes:** The number of observed DRD4 exon III alleles was derived, in part, from its corresponding allele frequencies. Multiple observations for populations from compiled molecular genetic studies were weighted with their respective sample sizes. See the Appendix for additional details.

climate indices, and soil characteristics. The same data source further provides information on the fraction of land that is allocated to pasture for each grid cell. Because one main hypothesis is that novelty-seeking traits are substantially correlated with the cultural value of pastoral nomadism, it would be interesting to find out whether land area suitable for pasture is significantly related to DRD4 exon III expected heterozygosity. Furthermore, this study also considers the



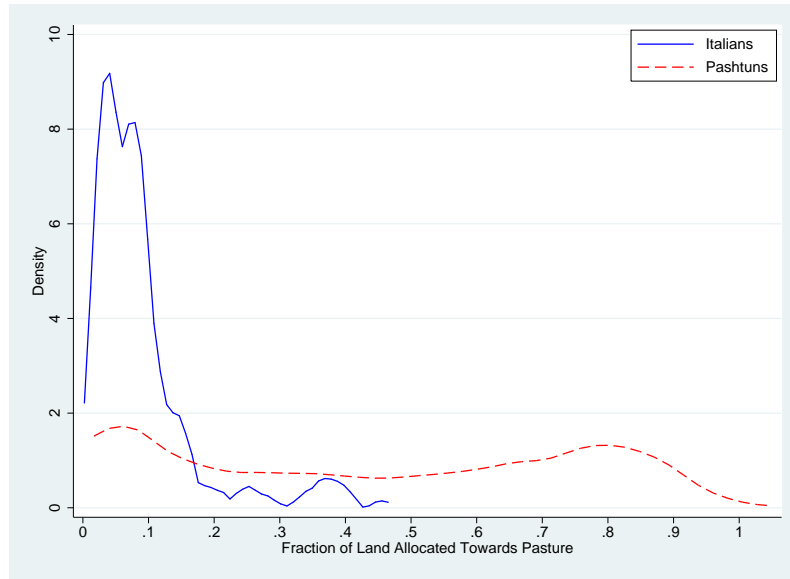
**Figure 4:** GLOBAL DISTRIBUTION OF LAND SUITABILITY FOR PASTURE

role of elevation in DRD4 exon III outcomes to capture the notion of human adaptiveness to prevailing biogeographic conditions. This would provide further credence for the role of terrain ruggedness in the natural selection of specific DRD4 exon III alleles.

Figure 4 depicts the worldwide distribution of land dedicated to pasture as a fraction of an area's overall land use, where warmer colors (red) indicate a larger fraction of land allocated to pasture and cooler colors (blue) indicate a smaller fraction. Using this geospatial raw data, the variables utilized in the empirical analysis were constructed by intersecting the global grid cells with data on the location covered by each population. Data on the location inhabited by each population is obtained from ALFRED. This database provides information on the approximate location of the population's homeland in the form of a geographic rectangle.

The spatial distribution of land quality varies considerably across populations. The following Figure 5 plots the distribution in the fraction of land that is allocated to pasture among Italians and among the Pashtun people, who inhabit parts of Afghanistan and Pakistan. The variation of land allocated to pasture is very limited in the Italian region, with an average fraction of 0.09 and a standard deviation of approximately 0.08. In contrast, the Pashtun people inhabit a geographical region that is much more suitable to the practice of pastoral nomadism: the average fraction of land allocated to pasture is 0.43, with an accompanying standard deviation of 0.32. It is not surprising that one of the major production modes of the Pashtun people is pastoral nomadism, given the variation in the quality of land allocated to pasture. Notably, the high quality of land suitable for agriculture in the Italian region provides evidence of the emergence of sedentary practices.





**Figure 5:** KERNEL DENSITY OF LAND SUITABILITY FOR PASTURE FOR ITALIANS AND PASHTUNS

The average number of gridded cells on biogeographic characteristics across all populations is about 823, ranging from 2 cells for Samaritans to 24,000 cells for Russians, with a standard deviation of about 2,647. The median cell size is about 90. In summary, there exist enough regional observations for each population to ensure a certain degree of variability in the construction of population-specific biogeographic indicators.

## 5 Empirical Results

This section investigates the biogeographic determinants of DRD4 exon III expected heterozygosity on the population level. It starts with migratory distance from East Africa as the only explanatory variable and then includes various measures of land characteristics in the regression. This procedure permits the predictive power of migratory distance from East Africa to be subjected to population-specific biogeographic controls. As stated elsewhere, linking DRD4 exon III expected heterozygosity to biogeographic characteristics and migratory distance from East Africa has a clear advantage for future empirical research questions. For example, biogeographic factors serve as qualified controls when analyzing the impact of the human trait associated with novelty-seeking behavior on historical or contemporary economic development, either in terms of population density or current levels of per capita income.<sup>17</sup>

<sup>17</sup>See also the theoretical model discussed in Galor and Michalopoulos (2012) for the role of novelty-seeking individuals in the process of economic development.

Specifically, the current analysis employs the following logistic regression model:

$$H_k^{DRD4} = \mathbf{F} \left( \alpha + region_k + \beta abs(latitude) + \mathbf{\Gamma}'_k \beta_{land} + \mathbf{\Lambda}'_k \beta_{elevation} + \mathbf{\Xi}'_k \beta_{pasture} \right) + \varepsilon_k, \quad (2)$$

where  $H_k^{DRD4}$  is observed DRD4 exon III expected heterozygosity for population  $k$  and  $\mathbf{F}(\cdot)$  is the cumulative distribution function of the logistic model. The variable *region* refers to continent fixed effects (e.g., regional dummies for Africa, the Americas, Asia, and Europe), *abs(latitude)* is the absolute value of a population's latitude in decimal degrees,  $\mathbf{\Gamma}$  refers to a vector of agricultural conditions (e.g., land suitability for agriculture, dispersion of land suitability, and land suitability Gini),  $\mathbf{\Lambda}$  is a vector of terrain ruggedness characteristics (e.g., elevation, and variation of elevation). Finally,  $\mathbf{\Xi}$  is a vector indicating land suitability for pasture (e.g., the fraction of land allocated to pasture, variation in land suitable for pasture, and dispersion of pasture land) and  $\varepsilon$  is a population-specific error term. Descriptive statistics and pairwise correlations for the variables employed in the empirical analysis are presented in Tables 5 and 6 in the Appendix. Given that migratory distance from East Africa has been found to be an important predictor of observed DRD4 exon III expected heterozygosity, the first specification in Table 2 considers it as the only regressor. As expected, migratory distance from East Africa is positively associated with observed DRD4 exon III expected heterozygosity. The associated coefficient is precisely estimated and statistically significant at the 5% level. Evaluating its marginal impact at mean values would state that for every 1,000 km distance from Addis Ababa, the expected heterozygosity increases by about 0.006 points. Column (2) maintains the robustness of the previous result to the inclusion of absolute latitude. As expected, higher absolute latitude regions, as a rough measure of harsh climate conditions, is positively associated with DRD4 exon III expected heterozygosity, although its impact is not precisely estimated. In the third column of Table 2, the robustness of migratory distance is confirmed for the inclusion of other biogeographic characteristics. In light of the possibility that the impact of biogeographic characteristics could be non-monotonic, the specifications examined also control for squared biogeographic controls. Interestingly, once controlling for migratory distance from East Africa, there remains virtually no association of DRD4 exon III expected heterozygosity with measures of land suitability for agriculture. Since geographic variability gave rise to random genetic drift, column (4) includes different measures for elevation. Once again, the robustness of migratory distance is verified. In column (5), measures of land suitability for pasture are included. The coefficients associated with pasture and its squared value enter in a hump-shaped fashion and are statistically significant at the 10% and 5% level, respectively. This result provides evidence that DRD4 exon III expected heterozygosity is the highest in intermediate land areas suitable for pastoral

**Table 2:** Biogeographic Determinants of DRD4 Exon III Expected Heterozygosity.

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: DRD4 Exon III Expected Heterozygosity Measure					
Migratory Distance	0.026** (0.008)	0.029*** (0.004)	0.028*** (0.006)	0.029*** (0.005)	0.032*** (0.003)	0.033 (0.018)
Absolute Latitude		0.005 (0.004)	0.006 (0.005)	0.004 (0.004)	0.004 (0.003)	0.004 (0.002)
Land Suitability			-0.041 (1.018)	0.019 (1.124)	-0.912 (0.828)	0.164 (1.322)
Land Suitability squared			-0.111 (1.035)	-0.014 (1.119)	0.655 (0.831)	-0.194 (1.171)
Dispersion of Land Suitability			0.147 (0.296)	0.251 (0.330)	0.438 (0.223)	0.015 (0.323)
Land Suitability Gini			-0.441 (0.450)	-0.046 (0.402)	-0.237 (0.256)	0.063 (0.361)
Elevation				-0.180 (0.147)	-0.145 (0.192)	0.072 (0.313)
Elevation squared				0.052 (0.036)	0.035 (0.055)	-0.008 (0.082)
Std. Dev. of Elevation				-0.217** (0.065)	-0.191* (0.085)	-0.249** (0.082)
Pasture					2.992* (1.167)	3.089*** (0.691)
Pasture squared					-4.400** (1.689)	-4.756*** (0.997)
Std. Dev. of Pasture					-0.135 (1.239)	-0.111 (1.325)
Dispersion of Pasture					-0.453 (0.583)	-0.151 (0.496)
Constant	-0.508** (0.189)	-0.685*** (0.124)	-0.624* (0.280)	-0.628* (0.278)	-0.553 (0.309)	-0.367 (0.538)
Number of Populations	92	92	92	92	92	92
$R^2$	0.94	0.95	0.95	0.95	0.95	0.96
Log Likelihood	73.95	75.43	76.21	78.74	81.61	90.53
Continent Effects	No	No	No	No	No	Yes

*Notes:* The dependent variable is observed DRD4 exon III expected heterozygosity on the population level.

*Independent variables:* *Migratory Distance* refers to the migratory distance (in 1,000 km) from East Africa starting from Addis Ababa, Ethiopia (9N, 39E) to the location of the population through five obligatory (land-restricted) waypoints: namely Cairo, Egypt (30N, 31E); Istanbul, Turkey (41N, 28E); Phnom Penh, Cambodia (11N, 104E); Anadyr, Russia (64N, 177E); and Prince Rupert, Canada (54N, 130W). *Absolute Latitude* is the absolute value of a population's latitude in decimal degrees. *Land Suitability* is a geospatial indicator, ranging from 0 to 1, of land suitability for agriculture across 0.5 decimal degree grid cells in the location inhabited by each population. *Dispersion of Land Suitability* is the difference between the maximum and minimum index of land suitability across 0.5 decimal degree grid cells in the location inhabited by each population. *Land Suitability Gini* is the Gini coefficient of the land suitability index for agriculture across 0.5 decimal degree grid cells in the location inhabited by each population. *Elevation* is the mean elevation (in km) above the sea level across 0.5 decimal degree grid cells in the location inhabited by each population. *Std. Dev. of Elevation* is the standard deviation of elevation across 0.5 decimal degree grid cells in the location inhabited by each population. *Pasture* is the mean value of the fraction of land that is allocated to pasture across 0.5 decimal degree grid cells in the location inhabited by each population. *Std. Dev. of Pasture* is the standard deviation of pasture across 0.5 decimal degree grid cells in the location inhabited by each population. *Dispersion of Pasture* is the difference between the maximum and minimum index of pasture across 0.5 decimal degree grid cells in the location inhabited by each population.

Standard errors, clustered with respect to regions, are reported in parentheses.

\*: Significant at the 10% level. \*\*: Significant at the 5% level. \*\*\*: Significant at the 1% level.

nomadism. One explanation could be that areas unsuitable for pasture (e.g., agricultural land areas) naturally select against novelty-seeking traits, whereas areas highly suitable for pasture naturally favor novelty-seeking traits due to the inherited evolutionary advantage they provide in resource-depleted areas. In both cases, the measures for DRD4 exon III expected heterozygosity would be low. Hence, only in intermediate areas suitable for pasture, where sedentary practices can also develop, do novelty-seeking and non-novelty-seeking traits coexist, leading in turn to higher DRD4 exon III expected heterozygosity values. This result provides evidence that novelty-seeking traits are quite common in areas suitable for pastoral nomadism. Finally, the last column controls for continent fixed effects. The coefficients associated with pasture and its squared value remain consistent and rather stable. However, the coefficient associated with migratory distance out of East Africa now becomes statistically insignificant. This is mainly due to the fact that this variable is highly correlated with the regional dummies, since the migratory distance out of East Africa is the lowest for African countries in comparison to the other regions. The empirical relationship between DRD4 exon III expected heterozygosity and biogeographic factors may be confounded by the effects of the underlying population structure. To control for this possibility, the following Table 3 truncates the regression analysis by population structure, regional coverage, and the number of sampled alleles. The specification in column (1) restricts the analysis to the 28 populations, as defined by the Human Genome Diversity Cell Line Panel. This definition was used to control for possible influences of population structure on DRD4 exon III alleles. This results in the exclusion of 64 populations. The coefficient associated with migratory distance from Africa is positive, but statistically insignificant, albeit weakly at the 10% level ( $t$ -statistic = 1.88). Given the limited sample of populations and the increased degrees of freedom in this specification, the falling significance level appears unsurprising.<sup>18</sup> Column (2) excludes HGDP populations from the analysis. The coefficient assessing the impact of migratory distance from Africa is positive and statistically significant at the 1% level. Interestingly, variation in terrain ruggedness is negatively associated with DRD4 exon III diversity. This result suggests that terrain variability acts as a barrier against the natural selection of novelty-seeking traits because it dampens geographic mobility of populations. Column (3) excludes Oceanic populations from the main analysis, because this region is biogeographically distinct from Eurasia, the Americas, and Africa, and thus may bias the results. This procedure results in the exclusion of 5 populations. The estimated coefficients are rather stable compared to the base

---

<sup>18</sup>The HGDP-CEPH project originally consisted of 53 populations, yet only 28 populations were found for DRD4 exon III alleles.

**Table 3:** Biogeographic Determinants of DRD4 Exon III Expected Heterozygosity (Assessing Population Structure).

	(1)	(2)	(3)	(4)	(5)	(6)
	HGDP Sample	Non-HGDP Sample	Without Oceania	Without Americas	Sampled Alleles $\geq 50$	Sampled Alleles $\geq 50$
Dependent Variable: DRD4 Exon III Expected Heterozygosity Measure						
Migratory Distance	0.024 (0.013)	0.034*** (0.005)	0.033*** (0.004)	0.007 (0.024)	0.029*** (0.006)	0.063** (0.023)
Absolute Latitude	-0.005 (0.013)	0.007 (0.006)	0.007* (0.003)	0.001 (0.004)	0.003 (0.003)	0.001 (0.002)
Land Suitability	-2.045 (3.351)	-0.580 (0.612)	-0.182 (0.537)	-0.083 (0.724)	1.300* (0.633)	3.069** (1.100)
Land Suitability squared	1.814 (2.673)	0.550 (0.647)	0.127 (0.605)	-0.355 (0.679)	-0.842 (0.546)	-2.123* (0.875)
Dispersion of Land Suitability	0.613 (0.920)	0.254 (0.305)	0.227 (0.147)	0.336 (0.268)	-0.160 (0.183)	-0.828* (0.323)
Land Suitability Gini	0.009 (1.437)	0.102 (0.191)	-0.048 (0.225)	-0.002 (0.510)	0.694* (0.327)	1.298** (0.381)
Elevation	-0.173 (0.316)	-0.077 (0.165)	-0.056 (0.322)	-0.227 (0.155)	-0.096 (0.245)	-0.043 (0.241)
Elevation squared	0.028 (0.060)	0.041 (0.030)	0.017 (0.090)	0.050 (0.043)	0.010 (0.062)	0.016 (0.060)
Std. Dev. of Elevation	-0.248 (0.304)	-0.185*** (0.038)	-0.248** (0.067)	-0.199 (0.107)	-0.254** (0.076)	-0.303** (0.086)
Pasture	4.702 (8.780)	2.518 (1.797)	3.296* (1.267)	1.345 (2.562)	2.874 (1.566)	2.306* (1.100)
Pasture squared	-7.513 (11.426)	-4.075 (2.684)	-4.893* (1.791)	-3.134 (3.102)	-4.477* (2.219)	-3.618* (1.610)
Std. Dev. of Pasture	-2.106 (5.047)	0.244 (0.773)	-0.184 (0.848)	1.604 (1.717)	-1.734 (1.346)	-1.422 (1.042)
Dispersion of Pasture	-0.231 (1.203)	-0.358 (0.542)	-0.266 (0.623)	-0.945* (0.372)	0.240 (0.651)	0.709 (0.456)
Constant	-0.211 (0.639)	-0.798* (0.377)	-0.909*** (0.133)	-0.114 (0.398)	-1.058*** (0.181)	-1.682** (0.623)
Number of Populations	28	64	87	68	74	74
$R^2$	0.95	0.96	0.96	0.95	0.97	0.98
Log Likelihood	27.03	57.46	81.07	62.28	78.27	90.77
Continent Effects	No	No	No	No	No	Yes

*Notes:* The dependent variable is observed DRD4 exon III expected heterozygosity on the population level.

*Independent variables:* *Migratory Distance* refers to the migratory distance (in 1,000 km) out of East Africa starting from Addis Ababa, Ethiopia (9N, 39E) to the location of the population through five obligatory (land-restricted) waypoints: namely Cairo, Egypt (30N, 31E); Istanbul, Turkey (41N, 28E); Phnom Penh, Cambodia (11N, 104E); Anadyr, Russia (64N, 177E); and Prince Rupert, Canada (54N, 130W). *Absolute Latitude* is the absolute value of a population's latitude in decimal degrees. *Land Suitability* is a geospatial indicator, ranging from 0 to 1, of land suitability for agriculture across 0.5 decimal degree grid cells in the location inhabited by each population. *Dispersion of Land Suitability* is the difference between the maximum and minimum index of land suitability across 0.5 decimal degree grid cells in the location inhabited by each population. *Land Suitability Gini* is the Gini coefficient of the land suitability index for agriculture across 0.5 decimal degree grid cells in the location inhabited by each population. *Elevation* is the mean elevation (in km) above the sea level across 0.5 decimal degree grid cells in the location inhabited by each population. *Std. Dev. of Elevation* is the standard deviation of elevation across 0.5 decimal degree grid cells in the location inhabited by each population. *Pasture* is the mean value of the fraction of land that is allocated to pasture across 0.5 decimal degree grid cells in the location inhabited by each population. *Std. Dev. of Pasture* is the standard deviation of pasture across 0.5 decimal degree grid cells in the location inhabited by each population. *Dispersion of Pasture* is the difference between the maximum and minimum index of pasture across 0.5 decimal degree grid cells in the location inhabited by each population. Standard errors, clustered with respect to regions, are reported between parenthesis.

\*: Significant at the 10% level. \*\*: Significant at the 5% level. \*\*\*: Significant at the 1% level.

specification in Table 2, column (5). However, the coefficient associated with absolute latitude, as a rough measure of harsh climatic conditions, now becomes statistically significant at the 10% level. The analysis in column (4) assesses the statistical significance of migratory distance and

biogeographic factors to the exclusion of populations in the Americas. The coefficient associated with migratory distance from East Africa remains positive but loses its statistical significance. One reason may be that migratory distance from Africa naturally selects for DRD4 exon III alleles only in very distant settlements. Hence, excluding populations from the American continent results in a considerable decrease in the variation of DRD4 exon III expected heterozygosity. Column (5) checks the robustness of the main results in restricting the analysis to populations with at least 50 sampled alleles (this criterion corresponds to a minimum number of 25 sampled individuals for each population) in the DRD4 exon III gene. This results in the exclusion of 18 populations due to small sample sizes. Interestingly, a large number of biogeographic indicators now become statistically significant. While land suitability for agriculture acts in a hump-shaped fashion, although the squared term is not statistically significant at the 10% level, greater inequality in land quality increases DRD4 exon III expected heterozygosity. Once controlling for continent fixed effects, as shown in column (6), the empirical results become even clearer, with a strong increase in the significance levels. The statistical significance of the hump-shaped effect for the fraction of land allocated to pasture is once again maintained. In general, greater variability in land quality for agriculture, as an indication of resource-depleted environments, naturally selects for specific alleles, thus increasing DRD4 exon III diversity.

The following Table 4 shows the biogeographic determinants of specific DRD4 exon III alleles that have received considerable attention in molecular genetic studies. In addition, this part of the analysis assesses the biogeographic impact on different definitions used to categorize novelty-seeking alleles. Column (1) utilizes the frequency of the 2-repeat allele as the dependent variable. Notably, this specific allele, which has a similar function to that of the 7-repeat, shows no association with migratory distance from East Africa. In contrast, it is highly correlated with the land suitability measures for agriculture. This result provides evidence that the 2-repeat allele may be adapted by populations practicing agriculture for a long period of time. This hypothesis seems plausible for Asian populations residing in China and India, which experienced the Neolithic transition much earlier than in other parts of the world and which have high 2-repeat allele frequencies. Column (2) presents the results for the biogeographic determinants of the 4-repeat allele. Consistent with the findings in Wang et al. (2004), the 4-repeat allele seems to be the ancestral allele in human populations because its frequency is highest in populations residing near the cradle of humankind. Column (3) establishes the general finding that since the exodus out of East Africa, a clear shift has been observed from the 4-repeat allele towards the 7-repeat allele. Notably, the only selective force acting on the 7-repeat allele is migratory distance

**Table 4:** Biogeographic Determinants of Specific DRD4 Exon III Alleles (Assessing Different Key Dependent Variables).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Dependent Variable						
	2-Repeat	4-Repeat	7-Repeat	2- and 7-Repeat	2- and 7-Repeat	Long Alleles $\geq 6$ -Repeat	Long Alleles $\geq 6$ -Repeat
Migratory Distance	-0.032 (0.030)	-0.058*** (0.009)	0.094** (0.029)	0.061*** (0.011)	0.043 (0.031)	0.088** (0.023)	0.002 (0.020)
Absolute Latitude	-0.020 (0.012)	0.003 (0.003)	-0.020 (0.019)	-0.006 (0.004)	-0.017** (0.005)	-0.025 (0.018)	-0.030*** (0.004)
Land Suitability	12.320** (4.161)	-1.294 (1.158)	0.648 (2.112)	1.868 (1.508)	1.905 (1.051)	0.751 (1.634)	-0.539 (0.796)
Land Suitability squared	-8.778** (2.969)	1.770 (1.225)	-1.362 (1.991)	-2.209 (1.367)	-2.335** (0.802)	-1.469 (1.491)	-0.892 (0.574)
Dispersion of Land Suitability	-0.706** (0.268)	0.182 (0.499)	-0.806 (0.681)	-0.382 (0.555)	-0.504 (0.313)	-0.559 (0.705)	-0.503 (0.318)
Land Suitability Gini	4.824** (1.874)	0.097 (0.392)	-0.301 (0.450)	0.001 (0.570)	-0.054 (0.726)	-0.155 (0.294)	-1.150 (1.206)
Elevation	-0.605 (0.348)	0.252 (0.362)	-0.002 (0.310)	-0.267 (0.388)	-0.313 (0.486)	0.119 (0.332)	-0.621 (0.435)
Elevation squared	0.140* (0.064)	-0.074 (0.103)	0.038 (0.089)	0.063 (0.112)	0.068 (0.132)	0.014 (0.093)	0.201 (0.123)
Std. Dev. of Elevation	0.392 (0.280)	0.341* (0.168)	-0.332 (0.502)	-0.211 (0.246)	-0.028 (0.256)	-0.573 (0.386)	0.035 (0.455)
Pasture	-10.193** (3.416)	-1.469 (1.236)	0.460 (2.747)	0.283 (1.607)	0.116 (1.353)	2.020 (2.249)	3.304*** (0.365)
Pasture squared	11.467** (3.704)	2.186 (2.013)	0.216 (4.524)	-0.055 (2.937)	0.542 (2.756)	-2.116 (3.509)	-3.813** (1.104)
Std. Dev. of Pasture	4.514 (2.357)	1.606 (3.423)	-7.862 (5.027)	-2.342 (4.925)	-3.088 (4.403)	-6.956* (3.210)	-9.116** (3.369)
Dispersion of Pasture	-1.511** (0.520)	-0.104 (1.194)	2.007 (1.298)	0.123 (1.450)	0.394 (1.061)	1.876 (1.155)	2.324** (0.784)
Constant	-4.592** (1.366)	0.981** (0.330)	-1.714** (0.547)	-1.264** (0.414)	-1.061** (0.282)	-1.570** (0.390)	-0.652 (0.359)
Number of Populations	92	92	92	92	92	92	92
$R^2$	0.69	0.98	0.85	0.90	0.92	0.88	0.93
Log Likelihood	126.7	80.54	81.05	85.22	94.63	81.03	107.9
Continent Effects	No	No	No	No	Yes	No	Yes

*Notes:* The dependent variables are various measures of observed DRD4 Exon III allele frequencies on the population level.

*Independent variables:* *Migratory Distance* refers to the migratory distance (in 1,000 km) out of East Africa starting from Addis Ababa, Ethiopia (9N, 39E) to the location of the population through five obligatory (land-restricted) waypoints: namely Cairo, Egypt (30N, 31E); Istanbul, Turkey (41N, 28E); Phnom Penh, Cambodia (11N, 104E); Anadyr, Russia (64N, 177E); and Prince Rupert, Canada (54N, 130W). *Absolute Latitude* is the absolute value of a population's latitude in decimal degrees. *Land Suitability* is a geospatial indicator, ranging from 0 to 1, of land suitability for agriculture across 0.5 decimal degree grid cells in the location inhabited by each population. *Dispersion of Land Suitability* is the difference between the maximum and minimum index of land suitability across 0.5 decimal degree grid cells in the location inhabited by each population. *Land Suitability Gini* is the Gini coefficient of the land suitability index for agriculture across 0.5 decimal degree grid cells in the location inhabited by each population. *Elevation* is the mean elevation (in km) above the sea level across 0.5 decimal degree grid cells in the location inhabited by each population. *Std. Dev. of Elevation* is the standard deviation of elevation across 0.5 decimal degree grid cells in the location inhabited by each population. *Pasture* is the mean value of the fraction of land that is allocated to pasture across 0.5 decimal degree grid cells in the location inhabited by each population. *Std. Dev. of Pasture* is the standard deviation of pasture across 0.5 decimal degree grid cells in the location inhabited by each population. *Dispersion of Pasture* is the difference between the maximum and minimum index of pasture across 0.5 decimal degree grid cells in the location inhabited by each population.

Standard errors, clustered with respect to regions, are reported between parenthesis.

\*: Significant at the 10% level. \*\*: Significant at the 5% level. \*\*\*: Significant at the 1% level.

from Africa. The same picture emerges from the analysis shown in column (4), when grouping together the 2- and 7-repeat alleles because of the similar blunted responses to dopamine signals. Once controlling for continent fixed effects, the coefficient associated with migratory distance from Africa becomes statistically insignificant, as shown in column (5). However, as stated elsewhere, this result is due to the fact that regional variables are highly correlated with

migratory distance from Africa, rendering the latter factor statistically insignificant. Column (6) employs the fraction of long alleles ( $\geq 6$ -repeat) as the key dependent variable, measuring the frequency of novelty-seeking traits within a population. The coefficient associated with migratory distance from Africa retains its statistical significance. Column (7) includes continent fixed effects in the analysis. Although the coefficient for migratory distance from Africa loses its statistical significance on the same grounds discussed above, the fraction of land allocated to pasture naturally selects for long allele types in a non-linear, hump-shaped fashion.

In summary, the empirical findings reveal the biogeographic origins of DRD4 exon III expected heterozygosity, a particular gene variant associated with novelty-seeking behavior. The estimates suggest that the human exodus out of East Africa in prehistoric times naturally selected for different DRD4 exon III alleles, leading to higher observed DRD4 exon III expected heterozygosity in populations that settled at a greater distance from East Africa. Furthermore, the results show that DRD4 exon III expected heterozygosity is quite common in populations inhabiting areas suitable for pastoral nomadism.

## 6 Conclusion

Based on recent empirical evidence about the role of migratory distance from East Africa and the human personality trait associated with novelty-seeking behavior, this study explores the biogeographic determinants of the human DRD4 exon III gene in depth. This particular gene variant is highly polymorphic across populations and exhibits strong signs of natural selection since the exodus out of East Africa. Providing a novel compilation of DRD4 exon III allele frequencies from various molecular genetic studies in a large sample of indigenous populations across the globe, this study utilizes population-specific biogeographic characteristics (e.g., land suitability for agriculture, fraction of land allocated to pasture, and elevation) using high-resolution geospatial data. The results suggest that migratory distance out of East Africa retains its strong explanatory power for novelty-seeking alleles at the DRD4 exon III locus, even controlling for a broad range of biogeographic determinants. This is the first study demonstrating that the quality of land endowments, to capture the notion of resource-depleted environments, is significantly correlated with variation in DRD4 exon III allele frequencies across populations. These observations lend support to the notion that novelty-seeking traits are adaptive in areas unsuitable for sedentary lifestyles, exposing human genes to strong selective pressure and enhancing the reproductive success of individuals exhibiting these traits in otherwise harsh and frequently changing environments.



A particularly promising research question would be the role of individuals bearing novelty-seeking traits in the process of economic development in a society. Recent empirical investigations confirm the indirect role of biogeography through the distribution of specific human traits, highlighting the importance of populations (in terms of culture, ideas, and beliefs), instead of a direct impact of locations on a country's economic development today. The main argument in this literature is that differences in biogeographic endowments gave rise to different production modes (sedentary versus nomadic lifestyles)—that is, cultures—across populations, and that the resulting differences in cultural values can be measured partly in human genes through a complex gene-culture co-evolution process. Because novelty-seeking traits are quite common in nomadic populations, it would be interesting to analyze whether the high prevalence of novelty seeking individuals within a society may have acted as an obstacle to permanent settlement because of the effective intergenerational transmission of nomadic lifestyles to future generations, resulting in a later agricultural transition and low economic development today.

## References

- Benjamin, J., Li, L., Patterson, C., Greenberg, B. D., Murphy, D. L., and Hamer, D. H. (1996). Population and Familial Association Between the D4 Dopamine Receptor Gene and Measures of Novelty Seeking. *Nature*, 12:81–84.
- Bhaduri, N., Das, M., Das, A. B., and Mukhopadhyay, K. (2007). Dopamine Receptor D4 Exon 3 Variable Number of Tandem Repeat Polymorphism: Distribution in Eastern Indian Population. *Indian Journal of Human Genetics*, 13:54–58.
- Bloom, D. E. and Sachs, J. D. (1998). Geography, Demography, and Economic Growth in Africa. *Brookings Papers on Economic Activity*, 1998:207–295.
- Bouchard, T. J. (1994). Genes, Environment and Personality. *Science*, 264:1700–1701.
- Cann, H. M., de Toma, C., Cazes, L., Legrand, M. F., Bodmer, V. M. L. P. J., Bodmer, W. F., Bonne-Tamir, B., Cambon-Thomsen, A., Chen, Z., Chu, J., Carcassi, C., Contu, L., Du, R., Excoffier, L., Ferrara, G. B., Friedlaender, J., Groot, H., Gurwitz, D., Jenkins, T., Herrera, R. J., Kidd, X. H. J., Kidd, K. K., Langaney, A., Lin, A. A., Mehdi, S. Q., Parham, P., Piazza, A., Pistillo, M. P., Qian, Y., Shu, Q., Xu, J., Zhu, S., Weber, J. L., Greely, H. T., Feldman, M. W., Thomas, G., Dausset, J., and Cavalli-Sforza, L. L. (2002). A Human Genome Diversity Cell Line Panel. *Science*, 296:261–262.
- Cavalli-Sforza, L. L., Menozzi, P., and Piazza, A. (1994). *The History and Geography of Human Genes*. Princeton University Press, Princeton, NJ.
- Chang, F.-M., Kidd, J. R., Livak, K. J., Pakstis, A. J., and Kidd, K. K. (1996). The World-Wide Distribution of Allele Frequencies at the Human Dopamine D4 Receptor Locus. *Human Genetics*, 98:91–101.
- Chen, C., Burton, M., Greenberger, E., and 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.
- Cloninger, C. R. (1987). A Systematic Method for Clinical Description and Classification of Personality Variants. A Proposal. *Archives of General Psychiatry*, 44:573–588.
- Cloninger, C. R., Przybeck, T. R., and Svrakic, D. M. (1991). The Tridimensional Personality Questionnaire: U.S. Normative Data. *Psychological Reports*, 69:1047–1057.

- Cook, C. J. (2011). Genetic Determinants of Health Differentials: The Role of Infectious Disease in Natural Selection Since the Neolithic Revolution. *mimeo*.
- Cook, C. J. (2013). The Role of Lactase Persistence in Precolonial Development. *mimeo*.
- Diamond, J. (1997). *Guns, Germs, and Steel: The Fates of Human Societies*. Norton, New York and London.
- Ding, Y.-C., Chi, H.-C., Grady, D. L., Morishima, A., Kidd, J. R., Kidd, K. K., Flodman, P., Spence, M. A., Schuck, S., Swanson, J. M., Zhang, Y.-P., and Moyzis, R. K. (2002). Evidence of Positive Selection Acting at the Human Dopamine Receptor D4 Gene Locus. *Proceedings of the National Academy of Sciences*, 99:309–314.
- Durham, W. H. (1982). Interactions of Genetic and Cultural Evolution: Models and Examples. *Human Ecology*, 10:289–323.
- Ebstein, R. P., Novick, O., Umansky, R., Priel, B., Osher, Y., Blaine, D., Bennett, E. R., Nemanov, L., Katz, M., and Belmaker, R. H. (1996). Dopamine D4 Receptor (D4DR) Exon III Polymorphism Associated with the Human Personality Trait of Novelty Seeking. *Nature*, 12:78–80.
- Eisenberg, D. T. A., Campbell, B., Gray, P. B., and Sorenson, M. D. (2008). Dopamine Receptor Genetic Polymorphisms and Body Composition in Undernourished Pastoralists: An Exploration of Nutrition Indices Among Nomadic and Recently Settled Ariaal Men of Northern Kenya. *BMC Evolutionary Biology*, 8.
- Gallup, J. L. and Sachs, J. D. (2000). Agriculture, Climate, and Technology: Why Are the Tropics Falling Behind? *American Journal of Agricultural Economics*, 82:731–737.
- Gallup, J. L. and Sachs, J. D. (2001). The Economic Burden of Malaria. *American Journal of Tropical Medicine and Hygiene*, 64:85–96.
- Gallup, J. L., Sachs, J. D., and Mellinger, A. D. (1999). Geography and Economic Development. *International Regional Science Review*, 2:179–232.
- Galor, O. and Michalopoulos, S. (2012). Evolution and the Growth Process: Natural Selection of Entrepreneurial Traits. *Journal of Economic Theory*, 147:759–780.
- Jensen, P., Mrazek, D., Knapp, P. K., Steinberg, L., Pfeffer, C., Schowalter, J., and 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*, 36:1672–1679.

- Johnson, D. L. (1969). The Nature of Nomadism: A Comparative Study of Pastoral Migrations in Southwestern Asia and Northern Africa. *Department of Geography, University of Chicago, Research Paper No. 118*.
- Khazanov, A. M. (2003). Nomads of the Eurasian Steppes in Historical Retrospective. In Kradin, N. N., Bondarenko, D. M., and Barfield, T. J., editors, *Nomadic Pathways in Social Evolution*, volume 5 of 8, chapter 2, pages 25–49. Center for Civilizational and Regional Studies of the Russian Academy of Sciences, Moscow.
- Koocheki, A. and Gliessman, S. R. (2005). Pastoral Nomadism, a Sustainable System for Grazing Land Management in Arid Areas. *Journal of Sustainable Agriculture*, 25:113–131.
- Lichter, J. B., Barr, C. L., Kennedy, J. L., Tol, H. H. V., Kidd, K. K., and Livak, K. J. (1993). A Hypervariable Segment in the Human Dopamine Receptor D<sub>4</sub> (DRD4) Gene. *Human Molecular Genetics*, 2:767–773.
- Livingstone, F. B. (1958). Anthropological Implications of Sickle Cell Gene Distribution in West Africa. *American Anthropologist*, 60:533–562.
- Masters, W. A. and McMillan, M. S. (2001). Climate and Scale in Economic Growth. *Journal of Economic Growth*, 6:167–186.
- Matthews, L. J. and 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*, 145:382–389.
- McCorkle, C. (1992). Plants, animals and people. In *Agropastoral Systems Research. Special Studies in Social, Political and Economic Development*, pages 3–18. Westview Press, Boulder, CO.
- Michalopoulos, S. (2012). The Origins of Ethnolinguistic Diversity. *American Economic Review*, 102:1508–1539.
- Munafò, M. R., Yalcin, B., Willis-Owen, S. A., and Flint, J. (2008). Association of the Dopamine D4 Receptor (DRD4) Gene and Approach-Related Personality Traits: Meta-Analysis and New Data. *Biological Psychiatry*, 63:197–206.
- Naka, I., Ohashi, J., Kimura, R., Furusawa, T., Yamauchi, T., Nakazawa, M., Natsuhara, K., Ataka, Y., Nishida, N., Ishida, T., Inaoka, T., Matsumura, Y., and Ohtsuka, R. (2012). DRD4 VNTR Polymorphism in Oceanic Populations. *Anthropological Science*, 120:151–155.

- Oak, J. N., Oldenhof, J., and Tol, H. H. V. (2000). The Dopamine D<sub>4</sub> Receptor: One Decade of Research. *European Journal of Pharmacology*, 405:303–327.
- Olsson, O. and Hibbs, D. A. (2005). Biogeography and Long-Run Economic Development. *European Economic Review*, 49:909–938.
- Paterson, A. D., Sunohara, G. A., and Kennedy, J. L. (1999). Dopamine D4 Receptor Gene: Novelty or Nonsense? *Neuropsychopharmacology*, 21:3–16.
- Plomin, R., Owen, M. J., and McGuffin, P. (1994). The Genetic Basis of Complex Human Behaviors. *Science*, 264:1733–1739.
- Rajeevan, H., Cheung, K.-H., Gadagkar, R., Stein, S., Soundararajan, U., Kidd, J. R., Pakstis, A. J., Miller, P. L., and Kidd, K. K. (2005). ALFRED: An Allele Frequency Database for Microevolutionary Studies. *Evolutionary Bioinformatics Online*, 1:1–10.
- Ramachandran, S., Deshpande, O., Roseman, C. C., Rosenberg, N. A., Feldman, M. W., and Cavalli-Sforza, L. L. (2005). Support from the Relationship of Genetic and Geographic Distance in Human Populations for a Serial Founder Effect Originating in Africa. *Proceedings of the National Academy of Sciences*, 44:15942–15947.
- Ramankutty, N., Foley, J. A., Norman, J., and McSweeney, K. (2002). The Global Distribution of Cultivable Lands: Current Patterns and Sensitivity to Possible Climate change. *Global Ecology and Biogeography*, 11:377–392.
- Sachs, J. D. (2001). Tropical Underdevelopment. *National Bureau of Economic Research Working Paper 8119*.
- Sachs, J. D. and Malaney, P. (2002). The Economic and Social Burden of Malaria. *Nature*, 415:680–685.
- Sachs, J. D. and Warner, A. M. (2001). The Curse of Natural Resources. *European Economic Review*, 45:827–838.
- Salzman, P. C. (2002). Pastoral Nomads: Some General Observations Based on Research in Iran. *Journal of Anthropological Research*, 58:245–264.
- Savitz, J. B. and Ramesar, R. S. (2004). Genetic Variants Implicated in Personality: A Review of the More Promising Candidates. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, 131B:20–32.

- Tishkoff, S. A., Reed, F. A., Friedlaender, F. R., Ehret, C., Ranciaro, A., Froment, A., Hirbo, J. B., Awomoyi, A. A., Bodo, J.-M., Doumbo, O., Ibrahim, M., Juma, A. T., Kotze, M. J., Lema, G., Moore, J. H., Mortensen, H., Nyambo, T. B., Omar, S. A., Powell, K., Pretorius, G. S., Smith, M. W., Thera, M. A., Wambebe, C., Weber, J. L., and Williams, S. M. (2009). The Genetic Structure and History of Africans and African Americans. *Science*, 324:1035–1044.
- Tovo-Rodrigues, L., Callegari-Jacques, S. M., Petzl-Erler, M. L., Tsuneto, L., Salzano, F. M., and Hutz, M. H. (2010). Dopamine Receptor D4 Allele Distribution in Amerindians: A Reflection of Past Behavior Differences? *American Journal of Physical Anthropology*, 143:458–464.
- Van Tol, H. H. M., Bunzow, J. R., Guan, H.-C., Sunahara, R. K., Seeman, P., Niznik, H. B., and Civelli, O. (1991). Cloning of the Gene for a Human Dopamine D<sub>4</sub> Receptor with High Affinity for the Antipsychotic Clozapine. *Nature*, 350:610–614.
- Van Tol, H. H. M., Wu, C. M., Guan, H.-C., Ohara, K., Bunzow, J. R., Civelli, O., Kennedy, J., Seeman, P., Niznik, H. B., and Jovanovic, V. (1992). Multiple Dopamine D4 Receptor Variants in the Human Population. *Nature*, 358:149–152.
- Wang, E., Ding, Y.-C., Flodman, P., Kidd, J. R., Kidd, K. K., Grady, D. L., Ryder, O. A., Spence, M. A., Swanson, J. M., and Moyzis, R. K. (2004). The Genetic Architecture of Selection at the Human Dopamine Receptor D4 (DRD4) Gene Locus. *American Journal of Human Genetics*, 74:931–944.
- Wiesenfeld, S. L. (1967). Sickle-Cell Trait in Human Biological and Cultural Evolution. *Science*, 157:1134–1140.
- Williams, J. and Taylor, E. (2006). The Evolution of Hyperactivity, Impulsivity and Cognitive Diversity. *Journal of the Royal Society Interface*, 3:399–413.
- Wolfe, N. D., Dunavan, C. P., and Diamond, J. (2007). Origins of Major Human Infectious Diseases. *Nature*, 447:279–283.

# A Descriptive Statistics

**Table 5:** Summary Statistics for DRD4 Exon III and Biogeographic Determinants.

	Obs.	Mean	S.D.	Min.	Max.
$H^{DRD4}$	92	0.4452	0.1204	0.0689	0.7083
2-Repeat	92	0.0808	0.0751	0.0000	0.3590
4-Repeat	92	0.6675	0.1673	0.1493	0.9643
7-Repeat	92	0.1796	0.1825	0.0000	0.7813
2- and 7-Repeat	92	0.2603	0.1643	0.0000	0.8358
Long Alleles ( $\geq 6$ -Repeat)	92	0.2102	0.1936	0.0000	0.7813
Migratory Distance	92	10.8834	7.8858	0.5264	27.4029
Absolute Latitude	92	28.8792	16.9585	1.5000	67.5000
Land Suitability	92	0.4817	0.2670	0.0068	0.9585
Land Suitability squared	92	0.3026	0.2594	0.0000	0.9187
Dispersion of Land Suitability	92	0.6478	0.3299	0.0143	1.0000
Land Suitability Gini	92	0.2760	0.2103	0.0127	0.7696
Elevation	92	0.5760	0.6802	-0.2898	3.7392
Elevation squared	92	0.7893	2.1019	0.0001	13.9818
Std. Dev. of Elevation	92	0.4819	0.4532	0.0170	2.0337
Pasture	92	0.2312	0.1981	0.0007	0.6257
Pasture squared	92	0.0923	0.1196	0.0000	0.3915
Std. Dev. of Pasture	92	0.1287	0.1017	0.0008	0.3851
Dispersion of Pasture	92	0.4869	0.3330	0.0022	1.0000

**Table 6:** Pairwise Correlations for DRD4 Exon III and Biogeographic Determinants.\*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
(1) $H^{DRD4}$	1.0000																		
(2) 2-Repeat	0.1510	1.0000																	
(3) 4-Repeat	-0.7093	0.0378	1.0000																
(4) 7-Repeat	0.4433	-0.4359	-0.8764	1.0000															
(5) 2- and 7-Repeat	0.5615	-0.0271	-0.9561	0.9114	1.0000														
(6) Long Alleles ( $\geq 6$ -Repeat)	0.4967	-0.4431	-0.8934	0.9774	0.8831	1.0000													
(7) Migratory Distance	0.4258	-0.1503	-0.7090	0.6902	0.6980	0.7150	1.0000												
(8) Absolute Latitude	0.0351	0.1680	0.2671	-0.3419	-0.3029	-0.3870	-0.2805	1.0000											
(9) Land Suitability	0.0530	0.2259	0.0531	-0.1218	-0.0320	-0.1227	-0.0444	0.0897	1.0000										
(10) Land Suitability squared	0.0215	0.2052	0.1012	-0.1509	-0.0738	-0.1604	-0.0886	0.1347	0.9641	1.0000									
(11) Dispersion of Land Suitability	-0.0276	0.0422	0.2705	-0.3050	-0.3194	-0.2696	-0.2069	0.2139	0.0506	-0.1014	1.0000								
(12) Land Suitability Gini	-0.1399	-0.1197	0.2952	-0.2539	-0.3367	-0.2416	-0.2429	0.2879	-0.6900	-0.6667	0.4371	1.0000							
(13) Elevation	-0.1141	-0.1677	0.1399	-0.0785	-0.1638	-0.0269	-0.0067	-0.0683	-0.1821	-0.2114	0.2910	0.3997	1.0000						
(14) Elevation squared	-0.0239	-0.1117	0.0307	-0.0053	-0.0569	0.0362	0.0618	-0.0709	-0.1176	-0.1452	0.2237	0.2805	0.9061	1.0000					
(15) Std. Dev. of Elevation	-0.1972	0.0754	0.2956	-0.2954	-0.2937	-0.2746	-0.0493	-0.0544	-0.0941	-0.1380	0.4700	0.3831	0.4766	0.3571	1.0000				
(16) Pasture	-0.0162	-0.2875	0.0259	0.0642	-0.0601	0.1176	0.1165	-0.1536	-0.2370	-0.2457	0.2012	0.3756	0.4117	0.3130	0.1805	1.0000			
(17) Pasture squared	-0.0042	-0.2820	-0.0357	0.1261	0.0112	0.1736	0.1744	-0.1429	-0.2585	-0.2528	0.1060	0.3271	0.3607	0.2858	0.0788	0.9696	1.0000		
(18) Std. Dev. of Pasture	-0.1800	-0.1931	0.2539	-0.1864	-0.2953	-0.1448	-0.0880	-0.1360	-0.3771	-0.3934	0.3794	0.5619	0.3945	0.2429	0.5388	0.6569	0.5260	1.0000	
(19) Dispersion of Pasture	-0.1828	-0.2319	0.2605	-0.1863	-0.3129	-0.1342	-0.0800	-0.0849	-0.3382	-0.3657	0.4879	0.5651	0.3700	0.1932	0.5341	0.6190	0.4959	0.9311	1.0000

\* Number of Populations: 92

## B Data Description and Sources

**DRD4 Exon III Expected Heterozygosity Measure.** The expected heterozygosity measure from observed DRD4 exon III allele frequencies across populations. This measure corresponds to the probability that two randomly selected individuals from the same population will differ with respect to the DRD4 exon III gene. Additional details regarding the sampled populations and data sources can be found in the separate Appendix to this paper, available from the author upon request.

**Migratory Distance from East Africa.** The migratory distance (in 1,000 km) out of East Africa from Addis Ababa, Ethiopia (9N, 39E) to the location of the population through five obligatory (land-restricted) waypoints: namely Cairo, Egypt (30N, 31E); Istanbul, Turkey (41N, 28E); Phnom Penh, Cambodia (11N, 104E); Anadyr, Russia (64N, 177E); and Prince Rupert, Canada (54N, 130W). Specifically, the computations are based on the great circle or geodesic calculation approach outlined in Ramachandran et al. (2005). Most of the information on the historical geographical location of the populations was provided by the Allele Frequency Database (ALFRED). For populations where ranges of latitude and longitude coordinates are available, the mean of the latitudes and the mean of the longitudes were used (see the separate Appendix for additional details regarding the sampled populations, available from the author upon request).

**Absolute Latitude.** The absolute value of a population’s approximate latitude in decimal degrees. The data source for the population’s representative latitude is from ALFRED.

**Land Suitability for Agriculture.** A geospatial indicator, ranging from 0 to 1, of land suitability for agriculture in the location inhabited by each population. The raw data are provided in high resolution gridded form of 0.5 decimal degrees latitude  $\times$  longitude by Ramankutty et al. (2002). This index represents the probability that a particular grid cell will be cultivated. The measures of land suitability for agriculture are based on indicators of climatic suitability,  $S_{clim}$ , (e.g., growing degree days (GDD) and a moisture index ( $\alpha$ ), capturing the availability of water to plants) and soil suitability for cultivation,  $S_{soil}$ , (e.g., soil carbon density ( $C_{soil}$ ) and an indicator ( $pH_{soil}$ ) capturing the extent of acidic or alkaline soil characteristics). The land suitability index for agriculture is then constructed as  $S = S_{clim} \times S_{soil}$ , where the functional forms for  $S_{clim}$  and  $S_{soil}$  are derived from probability density functions of actual cropland area,



$A_{crop}$ , versus each component in  $GDD$ ,  $\alpha$ ,  $C_{soil}$ , and  $pH_{soil}$ , respectively. Ramankutty et al. (2002) have chosen empirically fitted sigmoidal functions given by  $S_{clim} = f_1(GDD) \times f_2(\alpha)$  and  $S_{soil} = g_1(C_{soil}) \times g_2(pH_{soil})$ , respectively. Specifically,  $f_1(GDD) = \frac{1}{[1+e^{a(b-GDD)}]}$  and  $f_2(\alpha) = \frac{1}{[1+e^{c(d-\alpha)}]}$ , with  $a = 0.0052$ ,  $b = 1334$ ,  $c = 14.705$ , and  $d = 0.3295$ . The empirically fitted functions for  $g_1(C_{soil})$  and  $g_2(pH_{soil})$  are given by  $g_1(C_{soil}) = \frac{a}{[1+e^{b(c-C_{soil})}]} \times \frac{a}{[1+e^{d(e-C_{soil})}]}$ , with  $a = 3.9157$ ,  $b = 1.3766$ ,  $c = 3.468$ ,  $d = -0.0791$ ,  $e = -27.33$ , and

$$g_2(pH_{soil}) = \begin{cases} -2.085 + 0.475pH_{soil}, & \text{if } pH_{soil} \leq 6.5, \\ 1.0, & \text{if } 6.5 < pH_{soil} \leq 8 \\ 1.0 - 2.0pH_{soil}, & \text{if } pH_{soil} \geq 8 \end{cases}$$

The raw data are available online at the Center for Sustainability and the Global Environment (SAGE), <http://www.sage.wisc.edu>, at the University of Wisconsin. The values for the population level analysis are constructed by intersecting the global grid cells with data on the location inhabited by each population, which are obtained from ALFRED. This database provides information on the approximate location of each population's homeland in the form of a geographic rectangle.

**Dispersion of Land Suitability for Agriculture.** The difference between the maximum and minimum index of land suitability for agriculture across 0.5 decimal degree grid cells in the location inhabited by each population. For additional details, see the discussion about the construction of the land suitability index for agriculture.

**Land Suitability Gini.** This measure corresponds to the Gini coefficient of the land suitability index for agriculture across 0.5 decimal degree grid cells in the location inhabited by each population. For additional details, see the discussion about the construction of the land suitability index for agriculture.

**Elevation.** The mean elevation in the location inhabited by each population in km above the sea level. This measure is constructed using geospatial data at a 0.5 decimal degree resolution level, which in turn is based on more disaggregated data at a 0.083333328 decimal degree resolution level from the National Oceanic and Atmospheric Administration (NOAA) and U.S. National Geophysical Data Center, TerrainBase, release 1.0 (CD-ROM), Boulder, Colo. The raw data set can be downloaded online from <http://www.sage.wisc.edu/atlas/data.php?incdataset=Topography>.

**Standard Deviation of Elevation.** The standard deviation of elevation across 0.5 decimal degree grid cells in the location inhabited by each population. For additional details, see the discussion about the construction of the mean elevation index.

**Land Suitability for Pasture.** The mean value of the fraction of land that is allocated to pasture across 0.5 decimal degree grid cells in the location inhabited by each population. The raw data are available at <http://www.sage.wisc.edu/iamdata/>.

**Standard Deviation of Pasture.** The standard deviation of the fraction of land allocated to pasture across 0.5 decimal degree grid cells in the location inhabited by each population. For additional details, see the discussion about the construction of the land suitability for pasture measure.

**Dispersion of Pasture.** The difference between the maximum and minimum index of the fraction of land allocated to pasture across 0.5 decimal degree grid cells in the location inhabited by each population. For additional details, see the discussion about the construction of the land suitability for pasture measure.



**Table 8:** DRD4 Exon III Allele Frequencies, Expected Heterozygosity, and Migratory Distance from East Africa.

Population	2R	3R	4R	5R	6R	7R	8R	9R	10R	11R	$H^{DRD4}$	$N^a$	Lat <sup>b</sup>	Long <sup>b</sup>	Mig. Dist. <sup>c</sup>	HGDP
<b>Europe</b>																
Greeks	0.111	0.031	0.700	0.015	0.013	0.130	0.000	0.000	0.000	0.000	0.479	614	38.30	24.00	4.187	0
Ukrainian	0.085	0.015	0.750	0.000	0.055	0.065	0.030	0.000	0.000	0.000	0.422	200	48.50	31.00	4.598	0
Hungarian	0.095	0.033	0.659	0.010	0.004	0.189	0.009	0.001	0.001	0.000	0.519	5128	47.00	19.50	4.683	0
Croatian	0.113	0.029	0.706	0.010	0.004	0.133	0.002	0.002	0.000	0.000	0.469	1260	44.38	16.50	4.743	0
Adygei	0.048	0.048	0.760	0.029	0.000	0.115	0.000	0.000	0.000	0.000	0.404	104	44.50	39.75	4.766	1
Roman Jews	0.074	0.074	0.630	0.019	0.000	0.185	0.019	0.000	0.000	0.000	0.558	54	42.00	13.00	4.984	0
Italians	0.104	0.031	0.691	0.011	0.007	0.151	0.005	0.000	0.000	0.000	0.488	2805	42.45	12.75	5.006	1
Austrian	0.126	0.055	0.602	0.031	0.004	0.173	0.008	0.000	0.000	0.000	0.587	254	47.75	13.25	5.120	0
Sardinian	0.154	0.038	0.808	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.322	26	40.00	9.00	5.339	1
Swiss	0.105	0.043	0.697	0.007	0.000	0.147	0.000	0.000	0.000	0.000	0.479	400	46.75	8.25	5.433	0
Germans	0.076	0.049	0.655	0.014	0.004	0.197	0.004	0.001	0.000	0.000	0.523	3263	51.00	10.50	5.474	0
French	0.103	0.024	0.722	0.008	0.008	0.135	0.000	0.000	0.000	0.000	0.449	126	46.75	5.25	5.656	1
Dutch	0.088	0.045	0.635	0.013	0.006	0.198	0.015	0.000	0.000	0.000	0.548	616	52.13	5.38	5.844	0
Danes	0.070	0.100	0.640	0.000	0.000	0.170	0.020	0.000	0.000	0.000	0.546	100	56.35	10.50	5.855	0
Swedes	0.051	0.059	0.690	0.009	0.006	0.176	0.009	0.000	0.000	0.000	0.487	783	62.13	17.63	6.181	0
Spain	0.092	0.033	0.705	0.015	0.003	0.151	0.001	0.000	0.000	0.000	0.471	1572	40.00	-2.88	6.331	0
English	0.108	0.053	0.689	0.005	0.000	0.141	0.003	0.000	0.000	0.000	0.490	1304	54.50	-4.00	6.519	0
Finn	0.072	0.078	0.664	0.034	0.000	0.146	0.006	0.000	0.000	0.000	0.526	972	67.50	27.50	6.678	0
Irish	0.056	0.082	0.654	0.000	0.005	0.192	0.010	0.000	0.000	0.000	0.525	390	53.50	-8.50	6.772	0
<b>Middle East</b>																
Israeli	0.046	0.025	0.694	0.037	0.000	0.194	0.004	0.000	0.000	0.000	0.476	520	31.38	35.00	2.891	0
Samaritans	0.064	0.000	0.667	0.000	0.000	0.269	0.000	0.000	0.000	0.000	0.479	78	32.00	35.00	2.920	0
Druze	0.054	0.015	0.892	0.000	0.000	0.038	0.000	0.000	0.000	0.000	0.199	130	33.25	36.00	3.074	1
Jews, Yemenite	0.050	0.038	0.850	0.013	0.000	0.050	0.000	0.000	0.000	0.000	0.271	80	15.00	48.00	4.888	0
<b>Oceania</b>																
Malay	0.292	0.042	0.417	0.000	0.042	0.167	0.000	0.000	0.042	0.000	0.708	24	3.00	100.00	11.242	0
Filipino	0.359	0.000	0.538	0.038	0.026	0.038	0.000	0.000	0.000	0.000	0.578	78	13.90	121.50	12.176	0
Papua New Guineas	0.000	0.000	0.750	0.000	0.000	0.250	0.000	0.000	0.000	0.000	0.375	40	-5.00	145.00	15.123	1
Melanesian	0.074	0.037	0.752	0.033	0.012	0.091	0.000	0.000	0.000	0.000	0.418	242	-6.43	151.52	15.852	0
Melanesian, Nasioi	0.000	0.087	0.565	0.043	0.000	0.304	0.000	0.000	0.000	0.000	0.578	46	-6.00	155.00	16.201	1
Micronesians	0.267	0.034	0.621	0.009	0.043	0.026	0.000	0.000	0.000	0.000	0.540	116	4.68	176.30	18.232	0
Samoans	0.409	0.091	0.455	0.000	0.000	0.045	0.000	0.000	0.000	0.000	0.616	22	-14.00	-172.00	19.917	0
Polynesians	0.258	0.099	0.553	0.014	0.000	0.076	0.000	0.000	0.000	0.000	0.612	434	-23.84	-142.13	24.500	0

<sup>a</sup> The total number of sampled DRD4 exon III alleles, e.g., the number of individuals  $\times$  2. The number of observed chromosomes for each population was derived, in part, from its corresponding allele frequencies.

<sup>b</sup> Most of the information on the historical geographical location of the populations was provided by the Allele Frequency Database (ALFRED). For populations where ranges of latitude and longitude coordinates are available, the mean of the latitudes and the mean of the longitudes were used. Latitude and longitude coordinates are displayed in decimal degrees, where negative values indicate south/west, respectively.

For the following populations, the geographic coordinates were obtained from the studies listed below:

Kaingang, Guarani (Tovo-Rodrigues et al., 2010); Muskoke, Quechua, Colombians (Guaibito), Roman Jews, Thoti, Malay (Chen et al., 1999); Jemez Pueblo, Papua New Guineas, Samoans, Sardinians (Chang et al., 1996); Manipuri (Meitei) (Bhaduri et al., 2007); Melanesian, Micronesians, Polynesians (Naka et al., 2012); Makrani (HGDP-CEPH; Cann et al. (2002)); Hadza (Supporting online material in Tishkoff et al. (2009)).

<sup>c</sup> The migratory distance (in 1,000 km) out of East Africa starting from Addis Ababa, Ethiopia (9N, 39E) to the location of the population through five obligatory (land-restricted) waypoints: namely Cairo, Egypt (30N, 31E); Istanbul, Turkey (41N, 28E); Phnom Penh, Cambodia (11N, 104E); Anadyr, Russia (64N, 177E); and Prince Rupert, Canada (54N, 130W). Specifically, the computations are based on the great circle or geodesic calculation approach introduced in Ramachandran et al. (2005).

### **Zuletzt erschienen /previous publications:**

- V-313-09 **Heinz Welsch**, Implications of Happiness Research for Environmental Economics
- V-314-09 **Heinz Welsch, Jan Kühling**, Determinants of Pro-Environmental Consumption: The Role of Reference Groups and Routine Behavior
- V-315-09 **Christoph Böhringer and Knut Einar Rosendahl**, Green Serves the Dirtiest: On the Interaction between Black and Green Quotas
- V-316-09 **Christoph Böhringer, Andreas Lange, and Thomas P. Rutherford**, Beggar-thy-neighbour versus global environmental concerns: an investigation of alternative motives for environmental tax differentiation
- V-317-09 **Udo Ebert**, Household willingness to pay and income pooling: A comment
- V-318-09 **Udo Ebert**, Equity-regarding poverty measures: differences in needs and the role of equivalence scales
- V-319-09 **Udo Ebert and Heinz Welsch**, Optimal response functions in global pollution problems can be upward-sloping: Accounting for adaptation
- V-320-10 **Edwin van der Werf**, Unilateral climate policy, asymmetric backstop adoption, and carbon leakage in a two-region Hotelling model
- V-321-10 **Jürgen Bitzer, Ingo Geishecker, and Philipp J.H. Schröder**, Returns to Open Source Software Engagement: An Empirical Test of the Signaling Hypothesis
- V-322-10 **Heinz Welsch, Jan Kühling**, Is Pro-Environmental Consumption Utility-Maximizing? Evidence from Subjective Well-Being Data
- V-323-10 **Heinz Welsch und Jan Kühling**, Nutzenmaxima, Routinen und Referenzpersonen beim nachhaltigen Konsum
- V-324-10 **Udo Ebert**, Inequality reducing taxation reconsidered
- V-325-10 **Udo Ebert**, The decomposition of inequality reconsidered: Weakly decomposable measures
- V-326-10 **Christoph Böhringer and Knut Einar Rosendahl**, Greening Electricity More Than Necessary: On the Excess Cost of Overlapping Regulation in EU Climate Policy
- V-327-10 **Udo Ebert and Patrick Moyes**, Talents, Preferences and Inequality of Well-Being
- V-328-10 **Klaus Eisenack**, The inefficiency of private adaptation to pollution in the presence of endogeneous market structure
- V-329-10 **Heinz Welsch**, Stabilität, Wachstum und Well-Being: Wer sind die Champions der Makroökonomie?
- V-330-11 **Heinz Welsch and Jan Kühling**, How Has the Crisis of 2008-2009 Affected Subjective Well-Being?
- V-331-11 **Udo Ebert**, The redistribution of income when needs differ
- V-332-11 **Udo Ebert and Heinz Welsch**, Adaptation and Mitigation in Global Pollution Problems: Economic Impacts of Productivity, Sensitivity, and Adaptive Capacity
- V-333-11 **Udo Ebert and Patrick Moyes**, Inequality of Well-Being and Isoelastic Equivalence Scales
- V-334-11 **Klaus Eisenack**, Adaptation financing as part of a global climate agreement: is the adaptation levy appropriate?
- V-335-11 **Christoph Böhringer and Andreas Keller**, Energy Security: An Impact Assessment of the EU Climate and Energy Package
- V-336-11 **Carsten Helm and Franz Wirl**, International Environmental Agreements: Incentive Contracts with Multilateral Externalities
- V-337-11 **Christoph Böhringer, Bouwe Dijkstra, and Knut Einar Rosendahl**, Sectoral and Regional Expansion of Emissions Trading
- V-338-11 **Christoph Böhringer and Victoria Alexeeva-Talebi**, Unilateral climate policy and competitiveness: The implications of differential emission pricing
- V-339-11 **Christoph Böhringer, Carolyn Fischer, and Knut Einar Rosendahl**, Cost-Effective Unilateral Climate Policy Design: Size Matters
- V-340-11 **Christoph Böhringer, Jared C. Carbone, Thomas F. Rutherford**, Embodied Carbon Tariffs
- V-341-11 **Carsten Helm and Stefan Pichler**, Climate Policy with Technology Transfers and Permit Trading
- V-342-11 **Heinz Welsch and Jan Kühling**, Comparative Economic Performance and Institutional Change in OECD Countries: Evidence from Subjective Well-Being Data

- V-343-11 **Heinz Welsch and Jan Kühling**, Anti-Inflation Policy Benefits the Poor: Evidence from Subjective Well-Being Data
- V-344-12 **Klaus Eisenack und Leonhard Kähler**, Unilateral emission reductions can lead to Pareto improvements when adaptation to damages is possible
- V-345-12 **Christoph Böhringer, Brita Bye, Taran Fæhn, Knut Einar Rosendahl**  
Alternative Designs for Tariffs on Embodied Carbon: A Global Cost-Effectiveness Analysis
- V-346-12 **Christoph Böhringer, Jared C. Carbone, Thomas F. Rutherford**, Efficiency and Equity Implications of Alternative Instruments to Reduce Carbon Leakage
- V-347-12 **Christoph Böhringer, Andreas Lange, Thomas F. Rutherford**, Optimal Emission Pricing in the Presence of International Spillovers: Decomposing Leakage and Terms-of-Trade Motives
- V-348-12 **Carsten Helm, Dominique Demougin**, Incentive Contracts and Efficient Unemployment Benefits in a Globalized World
- V-349-12 **Heinz Welsch**, Organic Food and Human Health: Instrumental Variables Evidence
- V-350-12 **Heinz Welsch, Jan Kühling**, Competitive Altruism and Endogenous Reference Group Selection in Private Provision of Environmental Public Goods
- V-351-13 **Jürgen Bitzer, Erkan Gören**, Measuring Capital Services by Energy Use: An Empirical Comparative Study
- V-352-13 **Erkan Gören**, Economic Effects of Domestic and Neighbouring Countries' Cultural Diversity
- V-353-13 **Erkan Gören**, How Ethnic Diversity affects Economic Development?
- V-354-13 **Christoph Böhringer, Thomas F. Rutherford, Marco Springmann**; Clean-Development Investments: An Incentive-Compatible CGE Modelling Framework
- V-355-13 **Christoph Böhringer, Knut Einar Rosendahl, Jan Schneider**, Unilateral Climate Policy: Can Opec Resolve the Leakage Problem?
- V-356-13 **Heinz Welsch, Jan Kühling**, Income Comparison, Income Formation, and Subjective Well-Being: New Evidence on Envy versus Signaling
- V-357-13 **Anna Pechan, Klaus Eisenack**, The impact of heat waves on electricity spot markets
- V-358-13 **Heinz Welsch, Katrin Rehdanz, Daiju Narita, Toshihiro Okubo**, Well-being effects of a major negative externality: The case of Fukushima
- V-359-13 **Heinz Welsch, Philipp Biermann**, Electricity Supply Preferences in Europe: Evidence from Subjective Well-Being Data
- V-360-14 **Christoph Böhringer, Jared C. Carbone, Thomas F. Rutherford**, The Strategic Value of Carbon Tariffs
- V-361-14 **Christoph Böhringer, André Müller**, Environmental Tax Reforms in Switzerland A Computable General Equilibrium Impact Analysis
- V-362-14 **Christoph Böhringer, Nicholas Rivers, Thomas Rutherford, Randall Wigle**, Sharing the burden for climate change mitigation in the Canadian federation
- V-363-14 **Christoph Böhringer, Alexander Cuntz, Diemtar Harhoff, Emmanuel A. Otoo**, The Impacts of Feed-in Tariffs on Innovation: Empirical Evidence from Germany
- V-364-14 **Christoph Böhringer, André Müller, Jan Schneider**, Carbon Tariffs Revisited
- V-365-14 **Anna Pechan**, Which Incentives Does Regulation Give to Adapt Network Infrastructure to Climate Change? - A German Case Study
- V-366-14 **Erkan Gören**, The Biogeographic Origins of Novelty-Seeking Traits