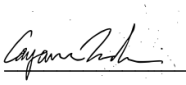
Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter know, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.



Name

04/10/23

Today's Date

Detecting Human Adaptations in Populations of the Andean Highlands

by

Cayanne Ruth Korder

John Lindo
Adviser

Anthropology

John Lindo

Adviser

Dr. Marcela Benítez

Committee Member

Dr. Arri Eisen

Committee Member

2023

Detecting Human Adaptations in Populations of the Andean Highlands

by

Cayanne Ruth Korder

John Lindo

Adviser

An abstract of
A thesis submitted to the Faculty of Emory College of Arts and Sciences
Of Emory University in partial fulfillment
Of the requirements of the degree of
Bachelor of Science with Honors

Anthropology and Human Biology

2023

Abstract

Detecting Human Adaptations in Populations of the Andean Highlands
By Cayanne Ruth Korder

Few whole genome population genetics studies have been conducted in the highlands of Ecuador, limiting what is known about human settlement and consequent adaptations to high-altitude regions of the Andes Mountains. Human adaptations arise over time that are specific to populations. If an isolated population is exposed to an environment over a significant period, these adaptations may become incorporated into a population's DNA. To test this, whole genomes of 15 modern individuals living in northern and southern Ecuador are analyzed using selection scans Ohana and iHS. The programs detect recent positive selection acting on the two populations of Ecuadorians, identifying genetic loci contributing to phenotypes, functional consequences, involved in human adaptions to high-altitude regions. It is hypothesized that of the 15 individuals living in high-altitude environments, there are positions on the chromosomes under positive selection that have functional consequences on the individuals' phenotypes. To validate the putative positive selection of the intronic, exonic, and intergenic loci amongst the populations, genome browsers are used to investigate the phenotypes related to the associated genes of the loci. Amongst the most strongly selected genes were *ANO7, PIK3CB, GALNT13, OCA2,* and *HAND1*. Functional consequences of these genes include the promotion of angiogenesis and erythropoiesis. *OCA2* was one of the most strongly selected genes; it assists in increased melanin production. The findings identify specific genes contributing to adaptations still prevalent and utilized in the DNA of Indigenous individuals living in the Andes Mountains today.

Detecting Human Adaptations in Populations of the Andean Highlands

By

Cayanne Ruth Korder

John Lindo

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Anthropology and Human Biology

2023

Acknowledgments

This work is not possible without the 15 individuals that donated their DNA to the project—thank you to the participants and their ancestors. And thank you to Dr. John Lindo, Yemko Pryor, Sophie Joseph, Mari Young, Lizzy Wagman, Richard Lunniss, and JJ Ortiz for the academic support throughout my research. Thank you Dr. Ari Eisen and Dr. Marcela E. Benítez for serving on my committee. Thank you to Kristin Phillips, Heather Carpenter, Robert A. Paul, the anthropology honors community, the Emory College Woodruff Scholars community, and the Halle Institute for Global Research for funding, guidance, and encouragement. Thank you to my mentors and chosen family Ben Whitcomb, Jess Whitcomb, Amy Rudolph, and Fred Rudolph—I am forever grateful for your kindness. Thank you to Lucille and Brian Huang for keeping me fed and sane. Lastly, I want to thank my family, including my eleven siblings Areia, Samantha, Cross, Courtney, Caleb, Emilee, Trysten, John, Zachariah, Asher, and Emerson. Thank you for inspiring me to do my best; everything I do is for you.

Table of Contents

Chapter 1: An Introduction to Ancient DNA

On an otherwise routine lecture day, I discovered my love for biological anthropology, effectively altering the trajectory of my undergraduate experience.

Dear Dr. Lindo,

Thank you for your lecture today. I am curious to learn more about biological anthropology. Can you point me in the direction of more literature in the field?

Sending this email catapulted me into the world of research. Within a few months, I declared my major in Anthropology & Human Biology. I joined Dr. John Lindo's ancient DNA lab as a sophomore and soon learned the technical procedures required to sequence ancient DNA including sterilization, extraction, purification, and building libraries. Beyond the technical training, I quickly realized the true impact of our work—John Lindo's lab explores the history of human movement and adaptation by applying methods of population genetics. We decode the human history written into our DNA, elucidating untold stories of human evolution, migration, relatedness, and indigeneity.

Through the work of Dr. John Lindo's Ancient DNA lab of Emory's Anthropology department, my project aims to recover a part of Ecuador's ancient human history. I analyze the DNA of 15 modern individuals living in high-altitude regions of the Andes Mountains to expand on evidence of human adaptations to high-altitude environments. I hypothesize that amongst the 15 individuals living in high-altitude environments, there are positions on the chromosomes under positive selection that have functional consequences on the individuals' phenotypes. The following introductory chapter provides context about human history research, effectively

"setting the stage" for my analyses. These topics include public interest in genetics research, the significant role of ancient genetics research in understanding human history, the scientific community's violation of the civil rights of Indigenous constituents, and a discussion on the ongoing lack of representation of Indigenous populations in genetics studies.

1.1 Public Interest in Genetics Research

In the wake of pandemic-induced boredom, I gave in and took the 23andMe genetics test to understand my own ethnic background more fully. In exchange for my genetic privacy, I received an interactive world map, lit up in countries where my DNA could be traced back to. Shockingly, the map highlighted all North and South America. While a quarter of my ancestry is likely associated with various regions of Mexico, the app could not rule out my connection to the entire western hemisphere. This ambiguity exists because direct-to-consumer (DTC) genetic company databases lack nuanced data sets for people from North, Central, and South America.

I am not alone in my curiosity to explore familial ancestry through genetics. The top three DTC genomic testing companies established in the U.S.—23andMe, Ancestry, and Family Tree DNA—collectively serve over 27 million people worldwide.[1] Decreasing prices of DTC genetic services make it increasingly accessible for the public to participate in genetic testing—services that cost as little as 59 dollars.[2] These companies provide consumers with coveted information that Dr. Henry Greely of the Stanford School of Medicine calls "ethnicity and family information," considered to be the main attraction of genetic testing.

These tests allow people to connect with their ethnic background and participating family members through personalized information about their ethnicity and biogeographical ancestry.

---

[1] Henry T. Greely, "The Future of DTC Genomics and the Law," *Journal of Law, Medicine & Ethics* 48, no. 1 (ed 2020): 151–52, https://doi.org/10.1177/1073110520917003.
[2] Greely, "The Future of DTC Genomics and the Law."

Ethnic information, in this case, is the estimated percentage breakdown of ethnic groups that a participant's ancestors are most genetically like. DTC genetics companies estimate biogeographical ancestry by comparing common genetic markers to populations that share similar allelic frequencies.[3] Family information is calculated similarly to ethnic information, but the genetic markers are compared to other participants in a DTC company database.[4] If a person chooses to participate, they can reach out and connect with other willing participants that share a percentage of genetic overlap. DTC companies like 23andMe present these features, along with medical and trait-based genetic reports to their consumers.[5] The consumer presumably derives personal meaning and value, as I did, from these tests; they provide an opportunity to look to the past and discover evidence to inform a deeper understanding of oneself.

1.2 Advances and Origins of the Study of Ancient DNA

Ancient DNA research is a relatively new field, but its legitimacy and relevance are nonetheless gaining worldwide attraction. The 2022 Nobel Prize in Physiology or Medicine, for example, was awarded to Swedish scientist Svante Pääbo for his reconstruction of a Neanderthal's whole genome from a 40,000-year-old finger bone of an individual in a Siberian cave.[6] The sequences of this individual allowed the team to study how Denisovan ancient DNA (aDNA) influenced human adaptation to high-altitudes.[7] His team also revealed modern whole genomes of individuals living in Europe and Asia that contained 1-4% of Neanderthal DNA.[8]

---

[3] Greely, 152.
[4] Greely, 152.
[5] Greely, 152.
[6] Benjamin Mueller, "Nobel Prize Awarded to Scientist Who Sequenced Neanderthal Genome," *The New York Times*, October 3, 2022, sec. Health, https://www.nytimes.com/2022/10/03/health/nobel-prize-medicine-physiology-winner.html.
[7] Richard G. Klein, "Profile of Svante Pääbo: 2022 Nobel Laureate in Physiology or Medicine," *Proceedings of the National Academy of Sciences* 120, no. 1 (January 3, 2023): e2217025119, https://doi.org/10.1073/pnas.2217025119.
[8] Klein.

Overlap in human and Neanderthal DNA was especially prevalent as humans migrated out of Africa and into Europe beginning 1000,000 years ago.[9] This level of genetic overlap came as a surprise for many; scientists once thought that *Homo sapiens* rarely interbred with Neanderthals. Pääbo's contributions influenced new discoveries across the field, and, importantly, he has made the world an attentive audience to human origin, migration, and evolution. Fifty years ago, it was not possible to sequence aDNA, and now the field of ancient genetics can study human evolution on a scale of 50,000 plus years.

With the expansion of aDNA technology and perspectives, biological anthropologists are motivated to uncover more prehistoric evidence to answer questions of human evolution, adaptation, phylogeny, and migration.[10] [11] [12] [13] Recent advances in ancient genomics allow scientists to recover ancient DNA on a whole genome scale with greater resolution than we have ever seen, even with the most degraded genetic material, as demonstrated by Dr. Pääbo's work. Fortunately, the field of aDNA research is growing, especially within the last thirty years with the development of next-generation sequencing (NGS) and polymerase chain reaction (PCR) technologies.[14]

The problem with ancient DNA research is that it has mainly focused on human

---

[9] Klein.
[10] Frédéric Delsuc et al., "Ancient Mitogenomes Reveal the Evolutionary History and Biogeography of Sloths - ScienceDirect," accessed April 10, 2023, https://www-sciencedirect-com.proxy.library.emory.edu/science/article/pii/S096098221930613X?via%3Dihub.
[11] Dan Chang et al., "The Evolutionary and Phylogeographic History of Woolly Mammoths: A Comprehensive Mitogenomic Analysis," *Scientific Reports* 7 (March 22, 2017): 44585, https://doi.org/10.1038/srep44585.
[12] Eleftheria Palkopoulou et al., "A Comprehensive Genomic History of Extinct and Living Elephants," *Proceedings of the National Academy of Sciences* 115, no. 11 (March 13, 2018): E2566–74, https://doi.org/10.1073/pnas.1720554115.
[13] Martin Sikora et al., "The Population History of Northeastern Siberia since the Pleistocene," *Nature* 570, no. 7760 (June 2019): 182–88, https://doi.org/10.1038/s41586-019-1279-z.
[14] Wenhao Xu et al., "An Efficient Pipeline for Ancient DNA Mapping and Recovery of Endogenous Ancient DNA from Whole-Genome Sequencing Data," *Ecology and Evolution* 11, no. 1 (2021): 391, https://doi.org/10.1002/ece3.7056.

population movements and evolution in Europe.[15] However, the use of ancient DNA to help characterize the nature of adaptive events—pathogen resistance, change in diet, and UV radiation—has yet to be fully realized. Because DNA is passed on from generation to generation, there are sections of our DNA that can link us to early civilizations and cultures. This is perhaps most pertinent to regions of the world such as Central and South America where many connections to ancient histories have been wiped away by European colonization.

1.3 Genetic Research of Indigenous Communities

With the development of aDNA research and DTC genetic testing, informed consent and data privacy have quickly become a source of concern—especially in the case of marginalized Indigenous peoples.[16] Limited genetic flow and cultural exchange, induced by environmental isolation, cultivated incredibly diverse languages, customs, cultures, and adaptations to one's environment. As is the case in Dr. Lindo's ancient DNA lab, genetic researchers are particularly interested in how genomes of isolated populations are unique in their variations compared to other populations. Research involving indigenous peoples raises legal, social, political, and spiritual issues—intersections that Euro-American-driven research often fail to address.[17] [18]Scientists conducting research with indigenous peoples' DNA have often failed to receive proper consent from indigenous communities. Two such cases include the misuse of DNA of

[15] Stephanie Marciniak and George H. Perry, "Harnessing Ancient Genomes to Study the History of Human Adaptation," *Nature Reviews Genetics* 18, no. 11 (November 2017): 661–64, https://doi.org/10.1038/nrg.2017.65.
[16] Greely, "The Future of DTC Genomics and the Law."
[17] Lorrieann Santos, "Genetic Research in Native Communities," *Progress in Community Health Partnerships : Research, Education, and Action* 2 (December 1, 2008): 321–27, https://doi.org/10.1353/cpr.0.0046.
[18] Nanibaa' A. Garrison, "Genetic Ancestry Testing with Tribes: Ethics, Identity & Health Implications," American Academy of Arts & Sciences, March 22, 2018, https://www.amacad.org/publication/genetic-ancestry-testing-tribes-ethics-identity-health-implications.

Havasupai and the Nuu-Chah-Nulth peoples.[19] [20] [21] In both cases, these issues came to light because the communities spoke out against the mistreatment of their DNA, but the researchers misuse of Havasupai and the Nuu-Chah-Nulth people's consent should not have happened. Genetic researchers have a responsibility to abide by proper conduct to protect research participants. Failing to do so is extremely harmful to the participants of the study and further perpetuates the violation of indigenous peoples' human and civil rights.

The first case involves the Native American community of the Havasupai, located in Arizona. Havasupai joined an Arizona State University research project in the early 1990s in hopes of exploring the community's high rates of type-2 diabetes.[22] The researchers requested fingerprints, handprints, and blood samples from the tribal members. Within the decade, the researchers never communicated the results of the diabetes study to Havasupai. A freezer malfunction in the laboratory left the tribal community under the impression that the study concluded with no further results. However, the samples damaged during the freezer malfunction were salvaged and sent to other researchers to conduct further studies.[23]

Many of the researchers' violations arose when they shared the Havasupai samples to explore research that did not align with the goals of the original diabetes studies. The Havasupai were unaware of the duration or scope of the continued studies, and the researchers failed to share any results with the Havasupai. Due to the alleged lack of informed consent, breach of fiduciary duty, misrepresentation of the tribal nation, and conversion, negligence, and violation

---

[19] Robyn L. Sterling, "Genetic Research among the Havasupai: A Cautionary Tale," *AMA Journal of Ethics* 13, no. 2 (February 1, 2011): 113–17, https://doi.org/10.1001/virtualmentor.2011.13.2.hlaw1-1102.
[20] "Genetic Researcher Uses Nuu-Chah-Nulth Blood for Unapproved Studies in Genetic Anthropology | Ha-Shilth-Sa Newspaper," July 22, 2013, https://hashilthsa.com/news/2013-07-22/genetic-researcher-uses-nuu-chah-nulth-blood-unapproved-studies-genetic.
[21] C W Schmidt, "Indi-Gene-Ous Conflict.," *Environmental Health Perspectives* 109, no. 5 (May 2001): A216–19.
[22] Sterling, "Genetic Research among the Havasupai."
[23] Sterling.

of civil rights, these violations led to lawsuits filed on behalf of the Havasupai.[24] The district court upheld the claims of negligence and civil rights violations. The Havasupai plaintiffs decided to move the case from the federal to the state level with the remaining claims. When the case was settled, the Havasupai were awarded with concessions and a monetary award. This case serves as one example of how scientists studying Native American DNA have misused and mistreated Indigenous peoples' consent in the field of genetics. The researchers at ASU needed to communicate with and updated the Havasupai on the status and results of all the studies and, in failing to deliver on these accounts, ASU violated the rights of the Havasupai.

The second case involves the Nuu-Chah-Nulth First Nation, located in British Columbia, Canada. Researchers conducted a study that explored rheumatoid arthritis, but the data from the study was then shared with other investigators. The researchers de-identified the samples and conducted various research inquiries unrelated to rheumatoid arthritis, which, at the time, was common practice. However, the Nuu-Chah-Nulth, after discovering that their samples were shared, requested that their samples be returned. The community believed the researchers disregarded the consent given for the original study.[25]

In each scenario, researchers conducted population genetic studies using de-identified DNA samples of the populations. The results suggested that humans first populated North, Central, and South America by migrating from Asia through the Bering Strait. This evolutionary history does not align with longstanding oral histories of ancestral origins, further fueling tension between traditional knowledge and scientific research. The studies did not result from the direct consent of the Nuu-Chah-Nulth and Havasupai, thus the tribes could not contribute their

[24] Sterling.
[25] "Genetic Researcher Uses Nuu-Chah-Nulth Blood for Unapproved Studies in Genetic Anthropology | Ha-Shilth-Sa Newspaper."

concerns about the results that were released to the public. For example, inbreeding coefficients were included in the results of the studies, which suggested that inbreeding was a common practice amongst the tribes. As a result of the controversy generated by these research studies, many tribes instituted their own independent review boards to inspect and actively engage with genetic research involving the tribe.

Although review boards assist in protecting tribes against extractive research conducted in Euro-American tradition, it is essential that researchers are following community-based participatory research (CBPR). Established in the early 1990s, Dr. DeCambra and Dr. Enos developed the guidelines to encourage reciprocal dialogue and feedback between scientists and Native communities through the research process.[26] This approach has gained traction amongst Native American review boards and researchers alike, but there is still a lack of guidance.[27] The principles of CBPR encourage research that will:

"(1) address health of the community within its broader cultural, social, economic, and political context, (2) involve community at all levels, from priority setting and planning to interpretation and dissemination of findings, (3) identify community needs and concerns that need to be addressed, (4) build on the strengths and resources within the community, (5) promote co-learning and knowledge transfer, (6) provide tangible benefits to the community, (7) do no harm." [28]

If these guidelines were honored while working with the communities of the Nuu-Chah-Nulth and Havasupai, the violations of these tribal communities could have been prevented. I

[26] D. S. Matsunaga et al., "Participatory Research in a Native Hawaiian Community. The Wai'anae Cancer Research Project," *Cancer* 78, no. 7 Suppl (October 1, 1996): 1582–86.
[27] Soroya Julian McFarlane et al., "Community-Based Participatory Research (CBPR) to Enhance Participation of Racial/Ethnic Minorities in Clinical Trials: A 10-Year Systematic Review," *Health Communication* 37, no. 9 (August 2022): 1075–92, https://doi.org/10.1080/10410236.2021.1943978.
[28] Santos, "Genetic Research in Native Communities," 1.

highlight the CBPR approach because it directly applies to the work of Dr. Lindo's lab; we process the DNA of ancient and modern indigenous individuals throughout Central and South America. We uphold CBPR practices and prioritize working directly with the ancestors of the ancient individuals we study.

The goal of the ancient DNA lab, on a larger scale, is to assist in filling gaps in the known genetic makeup of Central and South America. As the database of whole genomes grows, genetic mapping techniques will make the data available to the public through the development of a website. As of right now, the lack of data collected on the genetic nuances between populations in Central and South America makes it challenging to accurately map the genetic makeup of indigenous populations throughout Central and South America.

1.4 Lack of Indigenous Representation in Genetic Databases

Genetic ancestry databases represent a geographical region with the limited samples they have in a particular region. And, statistically speaking, as genetic ancestry databases grow, they become increasingly accurate. If a population of indigenous peoples is underrepresented in a database, which they often are, there is limited information a genetic testing company can provide to that population. Individuals from a region that has been under sampled cannot effectively represent the entire genetic diversity of the region, as the individual can lack common genetic variants or misrepresent a rare genetic variant as more common than it is. This phenomenon may be referred to as a sampling error. As with any research study, a larger sampling size decreases the chance of sampling errors—a statistical error wherein an analyst selects a sample that does not accurately represent the population. Because Indigenous peoples are more commonly underrepresented within genetic testing, they are more susceptible to this form of error.

Unsurprisingly, European ancestry is well documented throughout genetic databases, resulting in more accurate DNA reports for people of European descent that indigenous peoples. Indigenous representation in genetic testing databases decreased from 0.06 percent to 0.05 percent from 2009 to 2016, limiting DTC companies' ability to generalize with statistical accuracy about Indigenous populations. I focus on this issue because our laboratory strives to provide people from Central and South America with accurate representation of their ancestry in genetic databases. With these objectives in mind, I am hopeful that our lab contributes substantial genetic work to inform our constituents and their communities. In the next chapter, I provide context surrounding the seven Kichwa individuals participating in our study through the exploration of national identity.

Chapter 2: A National Identity

Ecuadorians currently living in Northern regions of Ecuador and the southernmost Loja provence donated saliva samples to our project; half of these individuals identify as Kichwa. The genetic makeup of these individuals is analyzed to make sense of selection pressures that may have existed in high-altitude regions of the Andes. This work is particularly significant because Kichwa represents the largest group of indigenous people in Ecuador. Of the almost three million people living in Ecuador today, 20% of the country's population are Kichwa.[29] The majority of Ecuadorian Kichwa live in the Andean highlands.[30] More specficially, many Kichwa live in the Carachi province in the north and Loja province in the south.[31] Despite ongoing challenges and

[29] Gregory Knapp, "The Changing Kichwa Language Map in Ecuador," 2020, 1731–42, https://doi.org/10.1007/978-3-319-73400-2_51-1.
[30] Fabricio González Andrade and Dora Sánchez, "Genetic Profile of the Kichwas (Quichuas) from Ecuador by Analysis of STR Loci," *Human Biology* 76, no. 5 (2004): 723–30, https://doi.org/10.1353/hub.2005.0007.
[31] González Andrade and Sánchez.

struggles, the Kichwa people remain to be a vibrant and important part of the cultural landscape of South America.

Although the Kichwa language connects Indigenous populations' nationality across Ecuador, there are a variety of identities, traditions, and social cultural characteristics within Kichwa populations.[32] This generalized term represents several culturally distinct groups, including the Karanki, Natabuela, Otavalo, Kayambi, Kitu-kara, Panzaleo, Chibuleo, Kichwa del Tungurahua, Salasaka, Puruhá, Kanòari, and Saraguro.[33] The collective Kichwa nationality recognizes the existence of peoples prior to the formation of the Ecuadorian state. During the colonial period, the Kichwa people were subjected to forced labor and exploitation by the Spanish; many Kichwa communities were forced to work in mines and on large-scale agricultural estates, and their cultural traditions and language were suppressed.[34] In the 20th century, there was a resurgence of Kichwa culture and identity throughout Ecuador.[35] Kichwa leaders and activists worked to promote indigenous rights and recognition, and the Kichwa language was recognized as an official language of the country in 1998.[36] Today, the Kichwa people throughout Ecuador continue to maintain their cultural traditions and connection to the land through music, dance, and other forms of artistic expression.[37]

---

[32] Felipe Esteban Acosta Muñoz, “Shunguhuan Yuyai: The Battle for Kichwa Language and Culture Revitalization in Ecuador As,” n.d.
[33] Carlos Espinosa, “Los indígenas Karanki, Natabuela, Otavalo y Kayambis danzaron en Ibarra,” El Comercio, November 30, 1AD, https://www.elcomercio.com/actualidad/ecuador/indigenas-karanki-natabuela-otavalo-kayambis.html.
[34] Nicholas Limerick, “Speaking for a State: Standardized Kichwa Greetings and Conundrums of Commensuration in Intercultural Ecuador,” 186–87, accessed April 10, 2023, https://doi.org/10.1086/708164.
[35] Muñoz, “Shunguhuan Yuyai: The Battle for Kichwa Language and Culture Revitalization in Ecuador As,” 18–25.
[36] Nicholas Limerick, “Speaking for a State.”
[37] Felipe Esteban Acosta Muñoz, “Shunguhuan Yuyai: The Battle for Kichwa Language and Culture Revitalization in Ecuador As,” 73–75.

The spelling and meaning of the word "Kichwa," varies in different contexts, especially when referring to the variety of languages spoken in different regions of the Americas. Each linguistic variety of Quechua is distinct amongst geographical regions, influenced by numerous factors. Firstly, the Quechua language originates from the Inca Empire.[38] Due to the expansion of the Inca Empire through the 1400's, it is the most widely spoken Indigenous language in the Americas—approximately 8 million speakers throughout the Andes.[39] [40] Quechua, referred to as *Runa Simi* by those who speak the language, is predominant amongst Incan descendants living in Bolivia, Peru, and Argentina.[41] In Ecuador, the dialect is Quichua—notice, it is spelled with an 'i' rather than an 'e'.[42] This spelling derives from an admixture of Spanish and Indigenous people.[43] Again, speakers refer to the language as *Runa Simi*.[44] And while 'Quichua' is a phrase used widely to refer to the Ecuadorian variety of languages, the alternative spelling variation is Kichwa, an ancestral name referring to the population that speaks Quichua.[45] Thus, the language of Kichwa is a product of a mixture of aboriginal language, pre-Inca language, and Quichua from Incas.[46] When I use the term 'Kichwa' throughout the remainder of the paper, please note the extensive cultural, linguistic, and regional variation associated with this general categorization.

Along with the Kichwa individuals included in the analysis, eight Oñacapac and Gañil individuals with over 95% indigenous ancestry donated DNA to the project. To give context, the Oñacapac and Gañil communities are in the Loja Province, the southernmost region of Ecuador.

---

[38] "Quechua: The Surviving Language of the Inca Empire | GVI," accessed April 10, 2023, https://www.gviusa.com/blog/quechua-the-surviving-language-of-the-inca-empire/.
[39] "Quechua: The Surviving Language of the Inca Empire | GVI."
[40] "Quechua," accessed April 10, 2023, https://plc.sas.upenn.edu/quechua.
[41] "Quechua."
[42] C.P. Ninacuri, "The influence of Kichwa in Ecuadorian Andean Spanish: the case of the morpheme -ka," *Boletin de Filologia* 57, no. 1 (2022): 209–31, https://doi.org/10.4067/S0718-93032022000100209.
[43] "Quechua."
[44] "Quechua."
[45] Ninacuri, "La influencia del kichwa en el castellano andino ecuatoriano ambateño: el caso del morfema -ka."
[46] Ninacuri.

The Oñacapac community is in the southernmost part of the Loja Province, near the town of Vilcabamba—known for its agriculture, including coffee, avocados, and bananas.[47] The Oñacapac have a strong tradition of weaving and producing textiles that are sold both locally and internationally. The community of Gañil is in the northern part of the Loja Province, near the town of Saraguro. Individuals living in Gañil are known for their skill in working with silver, creating a variety of silver jewelry/ The community is also known for its traditional clothing, which features brightly colored fabrics and intricate embroidery. Each of these communities primarily speaks Quichua.

The Gañil people are part of the broader group of indigenous people known as the Saraguros, which has a long history of resisting Spanish colonization and preserving its cultural traditions. Before the arrival of the Spanish in the 16th century, the Saraguro people were a highly organized society with a sophisticated system of governance, agriculture, and trade. Both the Oñacapac and Gañil communities have faced challenges in recent years, including the loss of traditional lands and the encroachment of modern development. However, these communities continue to work to preserve their cultural heritage and way of life.

Chapter 3: Populating High Altitude Regions of the Andes

The archeological evidence supports long standing sedentary communities in the Andean highlands, a critical environmental factor for inducing human adaptive processes amongst the individuals living in the Loja province and northern Ecuador. To give context, the Andes Mountains are characterized by high altitudes, with some regions reaching over 13,100 feet

[47] Brian S. Bauer, Javier Fonseca Santa Cruz, and Miriam Aráoz Silva, *Vilcabamba and the Archaeology of Inca Resistance*, 2015, 3–6, https://escholarship.org/uc/item/3xk2n538.

above sea level.[48] The Andean Mountains run through the Loja Province, which is in Southern Ecuador with approximately 4,166 square miles of terrain.[49] The province begins in the Andes Mountains, is flanked by Ecuadorian provinces, and borders Peru on the province's western-most border. Villages throughout the region vary in elevation; the capital city Loja, located at the base of the Cuxibamba valley, has an elevation of approximately 7,300 feet.[50] Although it is not entirely clear when people were first permanently settled in the Andes, current archeological evidence suggests that "intermittent exploration" of the Andes may have led to logistical or seasonal occupation, followed by permanent settlements.[51]

The timeline of permanent settlement in highlands regions of the Andes is opaque because its archeological context is limited. This is due, in part, to the humid conditions of the Andes, which are not conducive to preserving organic materials such as wood and coal.[52] Well-preserved archeological sites suggest permanent living in the Andes as early as 12,000 years ago or 10,000 BCE.[53] For example, archaeological evidence at the Cuncaicha rock shelter, located in the highlands of Southern Peru, supports this timeline.[54] Even as populations were faced with the challenges of low-oxygen levels and cold temperatures, evidence suggests that people began residing in high-altitude environments of the Andes within 2,000 years of the first entry into the Americas.[55]

---

[48] "Andes Mountains | Definition, Map, Plate Boundary, & Location | Britannica," accessed April 10, 2023, https://www.britannica.com/place/Andes-Mountains.
[49] "Loja | Ecuador | Britannica," accessed April 10, 2023, https://www.britannica.com/place/Loja-Ecuador.
[50] "Loja | Ecuador | Britannica."
[51] Kurt Rademaker et al., "Cuncaicha Rockshelter, a Key Site for Understanding Colonization of the High Andes: Reply to Capriles et Al.," *Current Anthropology* 57, no. 1 (February 2016): 101–3, https://doi.org/10.1086/684826.
[52] Rademaker et al., 102.
[53] Kurt Rademaker et al., "Paleoindian Settlement of the High-Altitude Peruvian Andes," *Science (New York, N.Y.)* 346, no. 6208 (October 24, 2014): 466–69, https://doi.org/10.1126/science.1258260.
[54] Rademaker et al., "Cuncaicha Rockshelter, a Key Site for Understanding Colonization of the High Andes," 102.
[55] Rademaker et al., "Paleoindian Settlement of the High-Altitude Peruvian Andes."

One of the most well-studied archaeological sites specific to the Loja province is the archaeological complex of Vilcabamba, a mountainous town near the Oñacapac community.[56] The site includes impressive stone structures and a network of terraced fields and irrigation canals that were used for agriculture. These agricultural systems allowed for cultivation crops such as maize, quinoa, and potatoes. The terraced fields are an impressive engineering feat, designed to maximize sunlight exposure and water retention on each level of the terrace.[57]

Another important site in the Loja province is the archaeological complex of Puerta de la Tierra, which dates to the Late Formative Period (500 BCE-500 CE). [58] This site includes a large earthen mound believed to be used for ceremonial purposes, as well as a network of canals and irrigation systems used for agriculture. Other archeological sites in the Loja province include the cave complex of the Toquilla Bamba, which includes evidence of human habitation dating back over 7,000 years. Additionally, the archaeological site of La Chimba includes a network of terraced fields and stone structures that were used for agriculture and religious ceremonies.[59] Overall, archaeological history of the Loja Province provides insights into the types of diverse cultures and societies that inhabited this region over the millennia. Humans have inhabited the Andean highlands for approximately 12,000 years, allowing adaptation to arise in the human genome over thousands of years.[60]

---

[56] Matthew Hayes, "Into the Universe of the Hacienda: Lifestyle Migration, Individualism and Social Dislocation in Vilcabamba, Ecuador," *Journal of Latin American Geography* 14, no. 1 (2015): 79–100, https://doi.org/10.1353/lag.2015.0001.
[57] Hayes.
[58] John W. Hoopes and Maria Masucci, "Evaluating Pre-Columbian Contact between Ecuador and Costa Rica: A Ceramic Approach (Masucci and Hoopes 2022)," *Waves of Influence: Pacific Maritime Networks Connecting Mexico, Central America, and Northwestern South America*, January 1, 2022, https://www.academia.edu/79246890/Evaluating_Pre_Columbian_Contact_between_Ecuador_and_Costa_Rica_A_Ceramic_Approach_Masucci_and_Hoopes_2022_.
[59] Randall Haas et al., "Humans Permanently Occupied the Andean Highlands by at Least 7 Ka," *Royal Society Open Science* 4, no. 6 (June 2017): 170331, https://doi.org/10.1098/rsos.170331.
[60] Rademaker et al., "Paleoindian Settlement of the High-Altitude Peruvian Andes."

Human adaptation in highland regions like the Loja Province is, in part, induced by several strong selection pressures including harsh environments, the development of agriculture, consequential social inequality, and relatively high population densities.[61] I will focus most specifically on the effect of the environment. High altitude environments are considered 'harsh' because these environments contain less oxygen, making breathing more difficult. Thus, people living in high altitude environments for extended periods of time have been shown to be associated with a range of health consequences, including hypoxia, chronic mountain sickness (CMS), high-altitude pulmonary hypertension (HAPH), and high-altitude pulmonary edema (HAPE).[62]

If the health conditions are threatening to the biological success of a population, it is likely that natural selection plays a role in mitigating the effects of the serious health risks. One of the most common health conditions associated with high altitude environments is hypoxia, which occurs when there is a lack of oxygen in the body. When the human body cannot meet its needs with the given level of oxygen, this leads to symptoms such as shortness of breath, headache, dizziness, fatigue, and nausea.[63]

Hypoxia can put significant stress on the body, which eventually leads to long term conditions, such as chronic mountain sickness (CMS).[64] This condition, characterized by an increase in red blood cell production (polycythemia), can occur in people who have lived at high

[61] John Lindo et al., "The Genetic Prehistory of the Andean Highlands 7000 Years BP Though European Contact," *Science Advances* 4, no. 11 (November 8, 2018): eaau4921, https://doi.org/10.1126/sciadv.aau4921.
[62] Fabiola León-Velarde and Olga Mejía, "Gene Expression in Chronic High Altitude Diseases," *High Altitude Medicine & Biology* 9, no. 2 (2008): 130–31, https://doi.org/10.1089/ham.2007.1077.
[63] E. Garrido, J. Botella de Maglia, and O. Castillo, "Acute, Subacute and Chronic Mountain Sickness," *Revista Clínica Española (English Edition)* 221, no. 8 (October 1, 2021): 481–90, https://doi.org/10.1016/j.rceng.2019.12.009.
[64] León-Velarde and Mejía, "Gene Expression in Chronic High Altitude Diseases."

altitudes for an extended period.[65] CMS can cause the blood to become thick, leading to a range of symptoms such as headache, dizziness, fatigue, shortness of breath, and cyanosis (blue discoloration of the skin).[66] Over time, CMS can lead to additional health complications such as heart failure and even stroke.[67]

Consistent exposure to high altitude environments can also induce high-altitude pulmonary edema (HAPE), a condition that is characterized by fluid accumulating in the lungs, making it difficult to breathe.[68] The cause of HAPE is thought to be a combination of factors, such as increased pressure in the blood vessels in the lungs, inflammation, and changes in the permeability of the blood vessel walls. Like HAPH, symptoms of HAPE can include coughing, shortness of breath, chest tightness, and fatigue. If left untreated, HAPE is also life-threatening.[69]

High-altitude pulmonary hypertension (HAPH) is an additional health complication associated with living in high altitude regions—a condition wherein blood vessels of the lungs become narrowed resulting in increased blood pressure in the lungs.[70] HAPH is considered less common than HAPE, and the onset of symptoms is typically slower, taking weeks or even months to develop. HAPH usually occurs at higher altitudes above 3,000 meters, or 9,800 feet, and is more common in individuals who have been exposed to high altitude for extended periods, such as permanent high-altitude residents.[71] This condition causes the heart to chronically work harder to pump blood, leading to harmful symptoms such as shortness of breath, chest pain, and fatigue. If left untreated, HAPH can also lead to heart failure and other complications. Overall,

---

[65] León-Velarde and Mejía, 130.
[66] Garrido, Botella de Maglia, and Castillo, “Acute, Subacute and Chronic Mountain Sickness.”
[67] Garrido, Botella de Maglia, and Castillo.
[68] Swapnil J. Paralikar, “High Altitude Pulmonary Edema-Clinical Features, Pathophysiology, Prevention and Treatment,” *Indian Journal of Occupational and Environmental Medicine* 16, no. 2 (2012): 59–62, https://doi.org/10.4103/0019-5278.107066.
[69] Paralikar.
[70] León-Velarde and Mejía, “Gene Expression in Chronic High Altitude Diseases.”
[71] León-Velarde and Mejía, 131.

living in high altitude regions without proper adaptations can lead to a range of health consequences. Fortunately, populations living at high altitude regions appear to have developed great adaptations to avoid these health complications and continue to increase biological success.

As a result of living amongst the associated health risks of the Andes Mountains, it is thought that humans developed evolutionary adaptations such as increased red blood cell production, increased lung capacity, improved circulation, enhanced heat retention, and increased melanin.[72] [73] [74] Reduced oxygen levels in the air led humans living in the Andes to produce more red blood cells—increasing oxygen delivery to the body's tissues.[75] Regarding the evolution of lung capacity, individuals living in the Andes developed larger lung capacities over time, increasing the amount of oxygen a person could transfer to blood capillaries in one breath.[76] [77] To further expedite oxygen delivery, populations in the Andes gradually adapted to have improved circulation, including a greater number of blood vessels and more efficient

[72] Jacob E. Crawford et al., "Natural Selection on Genes Related to Cardiovascular Health in High-Altitude Adapted Andeans," *The American Journal of Human Genetics* 101, no. 5 (November 2, 2017): 752–67, https://doi.org/10.1016/j.ajhg.2017.09.023.
[73] Abigail Bigham et al., "Identifying Signatures of Natural Selection in Tibetan and Andean Populations Using Dense Genome Scan Data," *PLoS Genetics* 6, no. 9 (September 9, 2010): e1001116, https://doi.org/10.1371/journal.pgen.1001116.
[74] Zhaohui Yang et al., "Genetic Adaptation of Skin Pigmentation in Highland Tibetans," *Proceedings of the National Academy of Sciences of the United States of America* 119, no. 40 (October 4, 2022): e2200421119, https://doi.org/10.1073/pnas.2200421119.
[75] Bigham et al., "Identifying Signatures of Natural Selection in Tibetan and Andean Populations Using Dense Genome Scan Data."
[76] Tom D. Brutsaert et al., "Effect of Developmental and Ancestral High Altitude Exposure on Chest Morphology and Pulmonary Function in Andean and European/North American Natives," *American Journal of Human Biology: The Official Journal of the Human Biology Council* 11, no. 3 (1999): 383–95, https://doi.org/10.1002/(SICI)1520-6300(1999)11:3<383::AID-AJHB9>3.0.CO;2-X.
[77] "Effects of Birthplace and Individual Genetic Admixture on Lung Volume and Exercise Phenotypes of Peruvian Quechua - Brutsaert - 2004 - American Journal of Physical Anthropology - Wiley Online Library," accessed April 10, 2023, https://onlinelibrary-wiley-com.proxy.library.emory.edu/doi/full/10.1002/ajpa.10319.

distribution of blood throughout the body.[78] [79] The 'reason' for these adaptations is ultimately to increase the amount of oxygen being delivered to cells throughout the human body.

There is also evidence that suggests high altitude populations have higher levels of melanin and greater skin pigmentation compared to low altitude populations.[80] People who live at high altitudes are exposed to higher levels of UV radiation due to the thinner atmosphere at higher altitudes, which allows more UV radiation to penetrate the atmosphere and reach the Earth's surface. As an example, a study of people living at different altitudes in Tibet found the high-altitude populations had darker skin and higher levels of melanin compared to lower altitude populations. [81] This study suggests evidence that people living in high altitudes evolved to produce more melanin to protect their skin from increased UV radiation exposure. Also, studies conducted on pigs and humans in Tibet and the Andes Mountains revealed high-altitude effect on the microbiome of skin; the reasons for this are still unclear.[82]

This chapter summarizes the chronological timeline of humans populating the Andes Mountains, the health risks associated with this shift is living conditions, and the consequent adaptations that have been observed in previous studies. To associate the adaptations within a specific environment, it is essential to understand the source of the selection pressure acting on the adaptation. Without an estimate of time, the source of selection pressure cannot be understood; a candidate adaptation could be the result of selection pressure from earlier evolutionary events rather than prolonged exposure to high altitude regions. This study seeks to

---

[78] Crawford et al., "Natural Selection on Genes Related to Cardiovascular Health in High-Altitude Adapted Andeans."
[79] Peter D. Wagner et al., "Pulmonary Gas Exchange and Acid-Base State at 5,260 m in High-Altitude Bolivians and Acclimatized Lowlanders," *Journal of Applied Physiology (Bethesda, Md.: 1985)* 92, no. 4 (April 2002): 1393–1400, https://doi.org/10.1152/japplphysiol.00093.2001.
[80] Yang et al., "Genetic Adaptation of Skin Pigmentation in Highland Tibetans."
[81] Yang et al.
[82] Bo Zeng et al., "High-Altitude Living Shapes the Skin Microbiome in Humans and Pigs," *Frontiers in Microbiology* 8 (October 6, 2017): 1929, https://doi.org/10.3389/fmicb.2017.01929.

understand human adaptations that developed as humans began populating high altitude regions. The next chapter outlines methods to detect recent selection, as opposed to previous moments in evolutionary history, in whole genomes. I also describe supplementary methods used to explore the genomic regions under putative positive selection.

Chapter 4: Methods

The data I investigate in this project originates from a computational study conducted primarily by Sophie Joseph and Dr. John Lindo. In their most recent paper, they analyzed the whole genomes of modern individuals living in the Loja Province and Kichwa people of Ecuador and discuss the putative positive selection associated with adaptations to tuberculosis.

The participants selected have an estimated 95% Indigenous ancestry, allowing the analyses to identify nuanced similarities and difference amongst the DNA of indigenous populations of the Andes. The whole genomes are computationally phased, meaning the two alleles at each genetic locus are assigned to the maternal or paternal chromosome. Phasing is an important step in genetic analysis because it separates and labels the two copies of a gene that an individual has inherited—tracking the inheritance of genetic traits across generations.[83] After the dataset is phased, the genomes are run through two types of selection scans.

The individuals of this study were grouped together by region and run through the Ohana and iHS selescans, so the seven individuals living in neighboring communities from the northern provinces of Ecuador were grouped and the eight individuals from the Loja province were grouped. These populations are separated in the selection scan analyses because the northern and southern regions of Ecuador are distinct; it is possible that different human adaptions arose in

[83] Brian L. Browning et al., "Fast Two-Stage Phasing of Large-Scale Sequence Data," *American Journal of Human Genetics* 108, no. 10 (October 7, 2021): 1880–90, https://doi.org/10.1016/j.ajhg.2021.08.005.

these regions. The analyses is more controlled if the two populations are kept separate. The first program used to analyze these two groups was the Rapid Evolutionary Haplotype Homozygosity or REHH program, which includes features called the extended haplotype homozygosity (EHH) and integrated haplotype scores (iHS). REHH is a haplotype method that is admixture-aware, using allele frequencies to produce a list of putatively selected gene positions across the population of individuals.[84] The two populations were also analyzed using the Ohana Program to further compare the putative selection in the phased dataset.

I hypothesize the positions on the chromosome under putative positive selection, made evident by the scans, are associated with genes that have functional consequences evident in the individuals' phenotypes. I validate hypothesized functional consequences of the genes under putative positive selection by surveying the chromosomal positions using GeneCard and GTex. These tools, described in further detail below, help me analyze the relationship between high altitude environments and its consequences on human adaptation.

REHH and its integrated mechanisms utilize extended haplotype homozygosity or EHH. This EHH feature, as the name suggests, uses probability to measure the length of DNA segments, or haplotypes, that are inherited from a common ancestor (the 'homozygosity' factor).[85] When two individuals have an identical long haplotype segment from the same line of descent, this indicates that they share an ancestor relatively recently. And when the shared segment is short, it suggests that the common ancestor is more distant. This is a result of continuous recombination of genetic information over generations—the haplotype lengths gradually become shorter. Thus, the statistical method uses probability to estimate the likelihood

[84] Mathieu Gautier, Alexander Klassmann, and Renaud Vitalis, "Rehh 2.0: A Reimplementation of the R Package Rehh to Detect Positive Selection from Haplotype Structure," *Molecular Ecology Resources* 17, no. 1 (January 2017): 79, https://doi.org/10.1111/1755-0998.12634.
[85] Gautier, Klassmann, and Vitalis, 79.

of a particular DNA sequence being shared by two individuals due to chance versus being inherited from a common ancestor.[86] This helps us determine the length of a haplotype and how many individuals in the tested population share that haplotype. Ultimately, the EHH allows us to infer the genetic history of the population and identify regions of the genome that have undergone positive selection or experienced genetic drift.

The REHH program also calculates the integrated haplotype scores (iHS) for each single nucleotide polymorphism (SNP) in the genome. The iHS statistic is a selection test that is essentially an extension of EHH; it measures the decay of EHH on one side of a selected allele to the decay of EHH on the other side of the allele.[87] The program then compares the decay measurements to a reference population's decay—in this case, the reference population are individuals from Europe and East Asia selected from the Simons Genome Diversity Project. In other words, the package measures severity of recent positive selection in a haplotype, relative to the background level of selection in the population.[88] A high iHS score indicates a recent positive selection occurred: the haplotypes with a 'candidate' allele may have a longer EHH, indicating a rapid increase in frequency across a population. The iHS statistic is then standardized and translated into a standardized normal distribution to identify the most extreme values in the distribution.[89] The program then generates a ranked list of loci based on the magnitude of the iHS statistic, with the most extreme values at the top of the list—the loci candidates most pertinent to my analyses.

---

[86] Gautier, Klassmann, and Vitalis, 84.
[87] Gautier, Klassmann, and Vitalis, 79–81.
[88] Mathieu Gautier and Renaud Vitalis, "Rehh: An R Package to Detect Footprints of Selection in Genome-Wide SNP Data from Haplotype Structure," *Bioinformatics (Oxford, England)* 28, no. 8 (April 15, 2012): 1176–77, https://doi.org/10.1093/bioinformatics/bts115.
[89] Gautier, Klassmann, and Vitalis, "Rehh 2.0," 89.

In summary, the REHH program combines iHS and EHH for each SNP in the genome to identify regions of the genome that have likely undergone recent positive selection.[90] Regions with a higher iHS score are more likely to have experienced recent positive selection. These statistical methods are used to analyze patterns of genetic variation across populations, helping to identify candidate genes that may have played a critical role in local adaptation to high altitude regions and the associated evolutionary changes.

Ohana, named after the Hawaiian word for 'family', is another selection scan program used to identify potential regions of the genomes experiencing recent positive selection.[91] The program compares the frequency distribution of single nucleotide polymorphisms (SNPs) using two populations: the high-altitude populations and an outgroup (a population that lacks selection pressures of high-altitude environments). The outgroup population of this study included fourteen individuals from East Asian and nineteen European individuals from the Simons Genome Diversity Project.[92] If one of the populations contains an excess of high-frequency derived alleles, it is likely that this population experienced recent positive selection. The Ohana addresses limitations posed by other selection scan programs; for example, other programs may require a priori assumptions regarding the mode of selection and 'appropriate' null models.[93]

The Ohana program operates by first genotyping the genetic variants of two populations—in this case, two geographically distinct human populations. Genotyping means the program is scoring the type of variant present at a given chromosomal position (locus) in the

[90] Gautier, Klassmann, and Vitalis, 90.
[91] Jade Yu Cheng, Thomas Mailund, and Rasmus Nielsen, "Fast Admixture Analysis and Population Tree Estimation for SNP and NGS Data," *Bioinformatics* 33, no. 14 (July 15, 2017): 2148–55, https://doi.org/10.1093/bioinformatics/btx098.
[92] Swapan Mallick et al., "The Simons Genome Diversity Project: 300 Genomes from 142 Diverse Populations," *Nature* 538, no. 7624 (October 13, 2016): 201-207H, https://doi.org/10.1038/nature18964.
[93] Cheng, Mailund, and Nielsen, "Fast Admixture Analysis and Population Tree Estimation for SNP and NGS Data."

genome. Ohana then compares the frequency distributions of each variant between the two populations and selects the variants showing significant differences in allele frequency.[94] Variants identified as 'significant' are clustered based on their positions in the genome to help identify the regions that likely experienced selection. Lastly, Ohana checks the accuracy of the clustering statistic by utilizing a permutation test, effectively anonymizing which variants are associated with the two populations to create a null distribution for the clustering statistic.[95] The program compares the null distribution, theoretically neutral evolution, to the observed distribution to confirm statistical significance. Once these steps are complete, the program produces a ranked list of variants under putative positive selection—I cover these results in the upcoming section.

Both Ohana and iHS identify the regions of the genomes associated with genes under putative positive selection. During this process, I wanted to focus my research on the candidate loci with particularly strong putative positive selection scans. To do this, I selected loci that were associated with genes frequently present in the top 20% of the ranked lists produced by the scans. I did not select loci that were below this threshold because I am interested in identifying particularly *strong* positive selection. Once I identify the candidates, I use GeneCards, a database of human genes, to explore the functional roles of the genes associated with the loci.[96] GeneCard provides information about gene function, structure, expression, and associated diseases. Each gene in the database is uniquely identified and summarized, including the gene's position on a chromosome, function, biological processes, and interactions with other genes and proteins. I

---

[94] Jade Yu Cheng et al., "Detecting Selection in Multiple Populations by Modeling Ancestral Admixture Components," *Molecular Biology and Evolution* 39, no. 1 (January 1, 2022): 8–9, https://doi.org/10.1093/molbev/msab294.
[95] Cheng et al., "Detecting Selection in Multiple Populations by Modeling Ancestral Admixture Components."
[96] "GeneCards - Human Genes | Gene Database | Gene Search," accessed April 10, 2023, https://www.genecards.org/.

combine the GeneCard findings with known genes associated with high-altitude human adaptations to produce a filtered list of candidate genes associated with high-altitude adaptations.

I validate the loci and associated genes have functional consequences prevalent in the populations' phenotypes using public genome browsers such GTEx Portal, UCSC genome browers, and Ensembl.[97] [98] [99] GTEx, for example, details the expression levels of a specific gene across different tissues and cell types and the relationships between genetic variation and gene expression. Each feature helps me understand the tissue-specific functions that are most strongly regulated by the candidate genes identified in the selection scan. By analyzing the GTEx data, I identify the mechanisms underlying tissue-specific gene expression and the association between genetic variation and the environment.

Chapter 5: Results

This chapter reports an overview of the ranked lists of candidate loci produced by the REHH program and Ohana, discusses these results, and delivers the findings of my genome browser research. First, I identified notable genetic positions ranked in the top 20% of the selection scans. The selected candidates continually reported comparatively high LLE ratios and were often associated with genes *repeatedly* listed amongst the candidate gene positions under putative selection produced by iHS scores and Ohana selection scans. The LLE ratios are a metric used by the Ohana scan to determine the potential role of positive selection in a specific region of the genome. The ratio is calculated by comparing the number of base pairs, or length, of a candidate region to the length of the original region prior to positive selection. If a region undergoes positive selection, the adaptive mutations responsible for the selection likely have

[97] "Ensembl Genome Browser 109," accessed April 10, 2023, https://useast.ensembl.org/index.html.
[98] "UCSC Genome Browser Home," accessed April 10, 2023, https://genome.ucsc.edu/.
[99] "GTEx Portal," accessed April 10, 2023, https://gtexportal.org/home/.

spread throughout the region, effectively extending the length of a genomic region that will likely be inherited together. Regions with a higher LLE ratio and iHS score indicate the locus experienced stronger positive selection.

I selected notable candidate loci from the ranked lists provided by the iHS and Ohana selection scans. I combined this observation with the GeneCard and genome browser to determine genetic regions relevant to high-altitude human adaptation. Unfortunately, there was very little information about the functional consequences associated with indivudal loci, making it difficult to investigate the candidates separately. Instead, I investigated the functional consequences of the genes associated with the loci. The intergenic loci were the most difficult to investigate because they were associated with two genes. Even still, these loci could be downstream enhancers and silencers without proximity to the genes they regulate. Ultimately, I identified notable genetic positions associated with high-altitude populations: *ANO7, PIK3CB, GALNT13, HAND1,* and *OCA2.* The functional role of these genetic positions and associated genes in the following sections, including tables to visualize the results of the selection scans.

5.1 ANO7

The Ohana scan revealed strong selection of *ANO7* in individuals living in the Loja Province, with two positions on the gene having an LLE ratio of 19.1 and four positions having an LLE ratio of 17.2.

Ranked Pos	Chr	Pos*	Variant type	nnotated Ger	lle ratio
4	2	242148764	exonic	ANO7	19.1
5	2	242156320	intronic	ANO7	19.1
31	2	242147703	intronic	ANO7	17.2
32	2	242151675	intronic	ANO7	17.2
33	2	242155229	intronic	ANO7	17.2
34	2	242162273	intronic	ANO7	17.2

Table 1. *ANO7* loci of interest due to high probability of selection in the Ohana selection scan. These individuals live in the Loja province. The "Ranked Position" indicates the position of the loci in comparison to the 250 total loci suggested by the Ohana scan. The loci were in the top 14% of the selection scan. *The chromosomal positions are listed in as Hg19 coordinates.

The first locus listed in table 1 is the only exonic position observed amongst the loci included in this study. Exonic positions are a part of the segments of DNA that code for a peptide sequence or protein. Interestingly, the selection scans rarely produce loci located in the exonic regions of the genome—most of the loci throughout the selection scans are intronic. I expected the consequences of not have intact exons would result in a nonfunctioning protein. As it turns out, exons have fewer somatic mutations compared introns, and this is *not* a result of purifying selection, but, rather, a higher rate of mutation repair mechanisms.[100] This is not because introns are inconsequential; these regions still play a significant role in forming the structural integrity of a protein. Rather, it suggests that introns cannot rely on repair mechanisms to the same degree that exons can. The results suggest that the introns are more frequently insulated against mutations through selection such as purifying selection. This might explain why the selection scans produce majority intronic loci.

ANO7 is a gene that encodes a transmembrane protein that is involved in calcium ion transport, and it is most frequently expressed in prostate tissues.[101] [102] A study suggested that the *ANO7* gene may play a role in regulating calcium homeostasis, which is crucial for maintaining cellular function in low-oxygen environments.[103] The study also found that variants in the *ANO7* gene are associated with increased hemoglobin levels, which is a common adaptation to high-

[100] Joan Frigola et al., "Reduced Mutation Rate in Exons Due to Differential Mismatch Repair," *Nature Genetics* 49, no. 12 (December 2017): 1684, https://doi.org/10.1038/ng.3991.
[101] "GeneCards - Human Genes | Gene Database | Gene Search."
[102] "Ensembl Genome Browser 109."
[103] Kuai Yu et al., "Explaining Calcium-Dependent Gating of Anoctamin-1 Chloride Channels Requires a Revised Topology," *Circulation Research* 110, no. 7 (March 30, 2012): 990–99, https://doi.org/10.1161/CIRCRESAHA.112.264440.

altitude environments. One intronic SNPs of this study is associated with *ANO7* is a binding site for the transcription factor *GATA-2*, which interacts with the hypoxia-inducible transcription factor (HIF) to modulate its transcriptional activity in response to hypoxia.[104] This pathway is induced in humans when exposed to environments with low levels of oxygen.[105]

5.2 PIK3CB

The Ohana scan also revealed chromosomal positions associated with *PIK3CB* in individuals from the Loja Province, including three intronic positions and three intergenic positions located between *PIK3CB* and *LINC01391* on chromosome 3, all with an LLE ratio of 17.2.

Ranked Pos	Chr	Pos*	Variant type	Annotated Gene	lle ratio
45	3	138419788	intronic	PIK3CB	17.2
46	3	138502058	intronic	PIK3CB	17.2
47	3	138539680	intronic	PIK3CB	17.2
48	3	138588071	intergenic	PIK3CB;LINC01391	17.2
49	3	138620026	intergenic	PIK3CB;LINC01391	17.2
50	3	138640171	intergenic	PIK3CB;LINC01391	17.2

Table 2. *PIK3CB* loci of interest due to high probability of selection in the Ohana selection scan. These individuals live in the Loja province. The "Ranked Position" indicates the position of the loci in comparison to the 250 total loci suggested by the Ohana scan. The loci were in the top 20% of candidate loci produced in the selection scan. *The chromosomal positions are listed in as Hg19 coordinates.

PIK3CB is a gene that encodes an enzyme that is involved in the phosphatidylinositol 3-kinase (PI3K) signaling pathway, which plays a critical role in regulating cellular responses to various stimuli, including growth factors and oxygen levels.[106] One study suggested that the

[104] Masahiko Tabata et al., "Stimulation of GATA-2 as a Mechanism of Hydrogen Peroxide Suppression in Hypoxia-Induced Erythropoietin Gene Expression," *Journal of Cellular Physiology* 186, no. 2 (2001): 260–67, https://doi.org/10.1002/1097-4652(200002)186:2<260::AID-JCP1025>3.0.CO;2-K.
[105] Bigham et al., "Identifying Signatures of Natural Selection in Tibetan and Andean Populations Using Dense Genome Scan Data," 7.
[106] "PIK3CB Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Beta [Homo Sapiens (Human)] - Gene - NCBI," accessed April 10, 2023, https://www.ncbi.nlm.nih.gov/gene/5291.

PIK3CB gene may play a role in regulating erythropoiesis, or the production of red blood cells, which is another common adaptation to high-altitude environments.[107] The *PIK3CB* gene has also been found to be positively selected in Andean populations; the study found that variants in the *PIK3CB* gene are associated with increased red blood cell count and hemoglobin levels in the Andean population.[108] *PIK3CB* is also involved in the regulation of HIF-1α, a key transcription factor involved in hypoxia adaptation.[109] One previous study suggested that *PIK3CB* may be associated with positive selection in high-altitude populations due to its role in hypoxia adaptation.[110]

5.3 GALNT13

Ohana and iHS identified notable genetic positions associated with *GALNT13*, a gene involved in protein glycosylation.[111] The iHS program identified one candidate locus amongst the southern Loja populations, but the candidate was within the top 30% rather than 20% threshold. Amongst the individuals living in the northern region of Ecuador, the Ohana scan identified two candidate loci within the top 16% of candidate loci in the selections scan. This exact scan indicated twenty positions with an LLE ratio of 13.7, which are not included in the

[107] Mahjoobeh Jafari et al., "PI3k/AKT Signaling Pathway: Erythropoiesis and beyond: JAFARI et Al.," *Journal of Cellular Physiology* 234 (September 7, 2018), https://doi.org/10.1002/jcp.27262.
[108] Priti Azad et al., "High-Altitude Adaptation in Humans: From Genomics to Integrative Physiology," *Journal of Molecular Medicine (Berlin, Germany)* 95, no. 12 (December 2017): 1269–82, https://doi.org/10.1007/s00109-017-1584-7.
[109] B. H. Jiang et al., "Phosphatidylinositol 3-Kinase Signaling Controls Levels of Hypoxia-Inducible Factor 1," *Cell Growth & Differentiation: The Molecular Biology Journal of the American Association for Cancer Research* 12, no. 7 (July 2001): 363–69.
[110] Emilia Huerta-Sánchez et al., "Genetic Signatures Reveal High-Altitude Adaptation in a Set of Ethiopian Populations," *Molecular Biology and Evolution* 30, no. 8 (August 1, 2013): 1877–88, https://doi.org/10.1093/molbev/mst089.
[111] Bahram Namjou et al., "EMR-Linked GWAS Study: Investigation of Variation Landscape of Loci for Body Mass Index in Children," *Frontiers in Genetics* 4 (December 3, 2013): 268, https://doi.org/10.3389/fgene.2013.00268.

table because they do not meet the threshold. Even so, the number of loci ranked in positions 165-185 out of 250 is significant.

Ranked Pos	Chr	Pos*	Variant type	nnotated Ger	lle ratio
40	2	154857288	intronic	GALNT13	15.6
41	2	154866073	intronic	GALNT13	15.6

Table 3. *GALNT13* loci of interest due to high probability of selection in the Ohana selection scan. These individuals live in northern provinces of Ecuador. The "Ranked Position" indicates the position of the loci in comparison to the 250 total loci suggested by the Ohana scan. The loci were in the top 16% of candidate loci produced in the selection scan. *The chromosomal positions are listed in as Hg19 coordinates.

GALNT13 has been linked to several diseases, including cancer, obesity, and Alzheimer's disease, which suggests that positive selection acting on this gene may have implications for human health.[112] [113] Recent research has also suggested that *GALNT13* was identified upstream of hypoxia and inflammation-related pathways.[114] Interestingly, the gene has also been identified as a potential risk factor for sickle cell disease.[115]

5.4 HAND1

The iHS scan of the Loja province identified one intergenic locus associated with *HAND1*, a transcription factor known to play a role in heart development and function.[116] Because the locus is intergenic, it is not entirely associated with the *HAND1* gene, but the locus's ranking is the highest of the iHS scan of the Loja individuals.

[112] "Whole Exome Sequence-Based Association Analyses of Plasma Amyloid-β in African and European Americans; the Atherosclerosis Risk in Communities-Neurocognitive Study | PLOS ONE," accessed April 10, 2023, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0180046.
[113] Namjou et al., "EMR-Linked GWAS Study."
[114] Min Yang et al., "Establishing a Prediction Model of Severe Acute Mountain Sickness Using Machine Learning of Support Vector Machine Recursive Feature Elimination," *Scientific Reports* 13 (March 21, 2023): 4633, https://doi.org/10.1038/s41598-023-31797-0.
[115] Bradley A. Maron, Roberto F. Machado, and Larissa Shimoda, "Pulmonary Vascular and Ventricular Dysfunction in the Susceptible Patient (2015 Grover Conference Series)," *Pulmonary Circulation* 6, no. 4 (December 2016): 426–38, https://doi.org/10.1086/688315.
[116] Kristiana Rood et al., "Gestational Hypoxia and Programing of Lung Metabolism," *Frontiers in Physiology* 10 (November 29, 2019): 1453, https://doi.org/10.3389/fphys.2019.01453.

Ranked Pos	Chr	Pos*	Variant type	Annotated Gene	Ile ratio
1	5	153954518	intergenic	HAND1;MIR3141	7.9802082

Table 4. HAND1 locus of interest due to high probability of selection in the iHS selection scan. These individuals live in the Loja province of Ecuador. The "Ranked Position" indicates the position of the locus in comparison to the 250 total loci suggested by the iHS scan. The locus was in the top 1% of candidate loci produced in the selection scan. *The chromosomal positions are listed in as Hg19 coordinates.

HAND1 has been shown to have a protective effect against myocardial ischemia, the reduction in the heart muscle's ability to pump blood efficiently, especially in early development[117] Failure to turn off this gene results in rapid death of the fetus.[118] Additional studies have shown that *HAND1* continues to play a role in regulating functions of the heart beyond early stages of development. The functional roles of *HAND1*, combined with the fact that the heart is particularly vulnerable to hypoxia, support the idea that *HAND1* may be a human adaptation to high-altitude environments of the Andes Mountains. The positive selection on *HAND1* may have played a role in the adaptation of populations living at high altitudes.

5.5 OCA2

The Ohana scan of the individuals living in northern Ecuador identified strong putative positive selection acting on *OCA2*, a gene involved in pigmentation. The Ohana scan of the individuals living northern Ecuador suggest 22 different loci associated with the *OCA2* gene, and each of these loci were in the top fifteen percent of candidate loci. The high frequency (22 separate 'hits') and level of putative suggestion compared to the other loci produced does suggest strong putative selection amongst the population.[119] This aligns with previous research, which

[117] Rood et al.
[118] Heinrich Taegtmeyer, Shiraj Sen, and Deborah Vela, "Return to the Fetal Gene Program," *Annals of the New York Academy of Sciences* 1188 (February 2010): 191–98, https://doi.org/10.1111/j.1749-6632.2009.05100.x.
[119] Bigham et al., "Identifying Signatures of Natural Selection in Tibetan and Andean Populations Using Dense Genome Scan Data," 7.

suggests that positive selection on *OCA2* may be linked to the evolution of skin pigmentation in humans.[120] A table of the results is listed below.

Ranked Pos	Chr	Pos*	Variant type	Annotated Gene	Ile ratio
6	15	28268990	intronic	OCA2	17.1
7	15	28282741	intronic	OCA2	17.1
8	15	28285284	intronic	OCA2	17.1
9	15	28288121	intronic	OCA2	17.1
10	15	28288748	intronic	OCA2	17.1
11	15	28289533	intronic	OCA2	17.1
12	15	28290228	intronic	OCA2	17.1
13	15	28290882	intronic	OCA2	17.1
14	15	28290960	intronic	OCA2	17.1
15	15	28291059	intronic	OCA2	17.1
16	15	28291061	intronic	OCA2	17.1
17	15	28292178	intronic	OCA2	17.1
29	15	28269008	intronic	OCA2	15.8
30	15	28277683	intronic	OCA2	15.8
31	15	28279941	intronic	OCA2	15.8
32	15	28281765	intronic	OCA2	15.8
33	15	28285967	intronic	OCA2	15.8
34	15	28287344	intronic	OCA2	15.8
35	15	28288549	intronic	OCA2	15.8
36	15	28291812	intronic	OCA2	15.8
37	15	28293788	intronic	OCA2	15.8
38	15	28295026	intronic	OCA2	15.8

Table 5. *OCA2* loci of interest due to high probability of selection. The "Ranked Position" indicates the positon of the loci in comparision to the 250 total loci suggested by the Ohana scan. *The chromosomal positions are listed in as Hg19 coordinates.

This gene is involved in the production of melanin, the pigment that gives color to the skin, hair, and eyes. Melanin provides protection against UV radiation, which can cause DNA damage and increase the risk of skin cancer.[121] Past studies have found that *OCA2* is positively

[120] Piyu Parth Naik and Syed Nadir Farrukh, "Influence of Ethnicities and Skin Color Variations in Different Populations: A Review," *Skin Pharmacology and Physiology* 35, no. 2 (2022): 72–73, https://doi.org/10.1159/000518826.
[121] Yang et al., "Genetic Adaptation of Skin Pigmentation in Highland Tibetans," 1.

selected for in Tibet and Han Chinese populations living in high-altitude regions.[122] Given what is known about *OCA2*, its tendency to protect against skin cancer and past studies' results, the gene may play a role in human adaptations to high levels of UV radiation in high-altitude regions. Because *OCA2* also appears to be positively selected for the high-altitude individuals in this study, it may be an example of *convergent evolution*; wherein unrelated populations evolve similar phenotypes independently because they adapted to similar environmental pressures. Therefore, the Ohana Selection scan suggests, in alignment with previous studies, that *OCA2* might play a significant role in the evolution of darker skin pigmentation in high altitude populations.

5.6 Regulating the Hypoxia-Inducible Transcription Factor (HIF)

One of the most notable observations is the significant involvement of introns in regulating HIF and its downstream targets. Regulation of *HIF* activity is critical for survival at high altitudes, as it is a key regulator of the body's response to low oxygen levels.[123] [124]In the presence of low oxygen levels, HIF activates a set of genes that are involved in adapting to hypoxia, including genes involved in erythropoiesis, angiogenesis, and metabolism.[125] [126] Many of the candidate genes including in my project are direct or indirect participants in regulating HIF, including *ANO7, HAND1,* and *PIK3BC*.

Populations that have lived at high altitudes for thousands of years have evolved a range of adaptations to help them cope with this hypoxic environment. These adaptations include

---

[122] Yang et al., 3.
[123] Bigham et al., "Identifying Signatures of Natural Selection in Tibetan and Andean Populations Using Dense Genome Scan Data," 7.
[124] León-Velarde and Mejía, "Gene Expression in Chronic High Altitude Diseases," 131–32.
[125] Bigham et al., "Identifying Signatures of Natural Selection in Tibetan and Andean Populations Using Dense Genome Scan Data," 2.
[126] Deepika Watts et al., "Hypoxia Pathway Proteins Are Master Regulators of Erythropoiesis," *International Journal of Molecular Sciences* 21, no. 21 (October 30, 2020): 8131, https://doi.org/10.3390/ijms21218131.

changes in hemoglobin concentration, increased ventilation, and changes in metabolism. Many of the genes listed above are involved in these adaptations, and their regulation by HIF is likely an important factor in their role in high-altitude adaptation.[127] Understanding the regulation of HIF and its downstream targets is important not only for understanding high-altitude adaptation but it could also be helpful for developing treatments for a range of hypoxia-related diseases.[128]

Chapter 6: Limitations & Future Directions

There are, of course, limitations to the interpreted results of the high-altitude adaptations codified in human DNA, which, in turn, limits our ability to generalize about these specific adaptations. Despite accounting for geographic variation in the current sample size, we cannot make broad claims that certain genetic variations were ancestral adaptations widely adopted by human populations across the Andes—our findings are limited to the fifteen participants involved in this study. The smaller sample size was partially due to the necessary exclusion of admixed persons, as accurate selection scans rely on participants having a high percentage of indigenous ancestry.

In addition to incorporating a larger modern sample size, future studies should also incorporate whole genomes of ancient individuals living in the high-altitude regions across a range of time periods. Studying a population's genetics over multiple time periods offers insight into which genes were most selected during specific time periods. The relevance of a gene may shift over time due to drastic changes in environmental conditions such as contact with outside populations or major environmental events like increased volcanic activity in an area. Thus,

---

[127] León-Velarde and Mejía, "Gene Expression in Chronic High Altitude Diseases," 131.
[128] Abigail W. Bigham et al., "Andean and Tibetan Patterns of Adaptation to High Altitude," *American Journal of Human Biology: The Official Journal of the Human Biology Council* 25, no. 2 (2013): 196, https://doi.org/10.1002/ajhb.22358.

increasing the sample size of both modern and ancient whole genomes could further clarify high altitude selection pressures evident throughout human history.

Additionally, research limitations are baked into the Ohana and iHS selection scans and the consequent interpretations.[129] I first became interested in these limitations after observing minimal overlap between the results produced by the two selection scans. I expected the regions of the genome under putative selection to be similar between the two selection scans, but several factors contribute to the variation in genetic regions identified as under selection amongst the two selection scans.

The scans use different methods and assumptions for detecting selection, resulting in different putative positive selections. The Ohana scan compares the frequency distribution of SNPs in a particular genomic region to the expected distribution under neutral evolution, while the iHS scan analyzes the length of haplotype homozygosity around a selected variant.[130] [131] The distinction between these two methods leads to differences in genomic regions identified as under selection. The programs also use different assumptions about selection and its underlying genetic model to drive the analyses. The Ohana scan, for example, assumes that the selection in question is ongoing and recent, while the iHS scan assumes that the selection is older and has reached fixation in the population.[132] [133] The different uses of methods and assumptions leads to different regions being identified as under selection by iHS and Ohana, even when each program receives the same input.

---

[129] Benjamin F. Voight et al., "A Map of Recent Positive Selection in the Human Genome," *PLoS Biology* 4, no. 3 (March 2006): e72, https://doi.org/10.1371/journal.pbio.0040072.
[130] Gautier, Klassmann, and Vitalis, "Rehh 2.0."
[131] Cheng, Mailund, and Nielsen, "Fast Admixture Analysis and Population Tree Estimation for SNP and NGS Data."
[132] Gautier, Klassmann, and Vitalis, "Rehh 2.0."
[133] Cheng, Mailund, and Nielsen, "Fast Admixture Analysis and Population Tree Estimation for SNP and NGS Data."

The selection scans can produce false positives and negatives because of the population structure, errors in genotyping, and human errors or "noise" underlying the collected dataset.[134] Additionally, there are also population-specific effects that may influence the selection a population undergoes. Not every population, even within proximity to one another, undergo the same selection pressures. The effects of selection can vary between populations due to environmental pressures, demographic history, and genetic variation. Thus, regions of the genome under putative selection in one population may not be under selection in another population. We mitigate population-specific effects by splitting up the samples into two groups based on where they live: the Loja Province and Northern Ecuador. To summarize, the differences in the methods, assumptions, noise, and population-specific effects can all influence the differences observed in the regions identified as under selection by the Ohana scan and the REHH scan. The selection scan results should be interpreted in the context of the limitations and multiple lines of evidence should be incorporated, such as GTEx, to identify genomic regions under selection.

But resources like GTEx also have limitations; for one, these kinds of databases do not provide information specific to the functional consequences of the candidate chromosomal positions pertinent to this study. Much of what is known about the positions, as I discovered through my research, is heavily understood through the lens of cancer research, not human evolution. Thus, the exact mechanisms related to these adaptations are difficult to analyze in depth until more of the human genome is understood.

---

[134] "A Beginner's Guide to Low-coverage Whole Genome Sequencing for Population Genomics - Lou - 2021 - Molecular Ecology - Wiley Online Library," accessed April 10, 2023, https://onlinelibrary-wiley-com.proxy.library.emory.edu/doi/full/10.1111/mec.16077.

Lastly, in future studies, I hope to study population-specific adaptations related to the onset of agriculture, consequent changes in human diet, and, ultimately, its role in changing our genome and epigenome over time. During my time in Ecuador last summer, I wanted to explore the potential genetic impact of incorporating starch into the diet of ancient individuals living in a coastal town called Salango. To include an archaeological context to the project, I planned to analyze and detect foodstuffs found on the surface of pottery found at the archaeological site in Salango—I even set up a research project with Renee Stein of the conservation lab of the Carlos Museum to gain structure for this project. As is the case with most research, it did not go as planned. But the premise of the project still excites me, and I hope to explore historical shifts in diet and its impact on our genetics someday.

Human adaptations arise over time that are specific to populations. If an isolated population is exposed to an environment over a significant period, these adaptations may become incorporated into a population's DNA. In the case of Andean highlands, human populations settled in the Andes adapted to their environment. The Ohana and iHS scans produce putative positive selection candidates, combined with information about genetic functional consequences from public databases and contextualized environmental factors, to identify specific genes contributing to adaptations still prevalent and utilized in the DNA of Indigenous individuals living in the Andes Mountains today.

Bibliography

"A Beginner's Guide to Low-coverage Whole Genome Sequencing for Population Genomics - Lou - 2021 - Molecular Ecology - Wiley Online Library." Accessed April 10, 2023. https://onlinelibrary-wiley-com.proxy.library.emory.edu/doi/full/10.1111/mec.16077.

American Academy of Arts & Sciences. "Genetic Ancestry Testing with Tribes: Ethics, Identity & Health Implications," March 22, 2018. https://www.amacad.org/publication/genetic-ancestry-testing-tribes-ethics-identity-health-implications.

"Ancient Mitogenomes Reveal the Evolutionary History and Biogeography of Sloths - ScienceDirect." Accessed April 10, 2023. https://www-sciencedirect-com.proxy.library.emory.edu/science/article/pii/S096098221930613X?via%3Dihub.

"Andes Mountains | Definition, Map, Plate Boundary, & Location | Britannica." Accessed April 10, 2023. https://www.britannica.com/place/Andes-Mountains.

Azad, Priti, Tsering Stobdan, Dan Zhou, Iain Hartley, Ali Akbari, Vineet Bafna, and Gabriel G Haddad. "High-Altitude Adaptation in Humans: From Genomics to Integrative Physiology." *Journal of Molecular Medicine (Berlin, Germany)* 95, no. 12 (December 2017): 1269–82. https://doi.org/10.1007/s00109-017-1584-7.

Bauer, Brian S., Javier Fonseca Santa Cruz, and Miriam Aráoz Silva. *Vilcabamba and the Archaeology of Inca Resistance*, 2015. https://escholarship.org/uc/item/3xk2n538.

Bigham, Abigail, Marc Bauchet, Dalila Pinto, Xianyun Mao, Joshua M. Akey, Rui Mei, Stephen W. Scherer, et al. "Identifying Signatures of Natural Selection in Tibetan and Andean Populations Using Dense Genome Scan Data." *PLoS Genetics* 6, no. 9 (September 9, 2010): e1001116. https://doi.org/10.1371/journal.pgen.1001116.

Bigham, Abigail W., Megan J. Wilson, Colleen G. Julian, Melisa Kiyamu, Enrique Vargas, Fabiola Leon-Velarde, Maria Rivera-Chira, et al. "Andean and Tibetan Patterns of Adaptation to High Altitude." *American Journal of Human Biology: The Official Journal of the Human Biology Council* 25, no. 2 (2013): 190–97. https://doi.org/10.1002/ajhb.22358.

Browning, Brian L., Xiaowen Tian, Ying Zhou, and Sharon R. Browning. "Fast Two-Stage Phasing of Large-Scale Sequence Data." *American Journal of Human Genetics* 108, no. 10 (October 7, 2021): 1880–90. https://doi.org/10.1016/j.ajhg.2021.08.005.

Brutsaert, Tom D., Rudy Soria, Esperanza Caceres, Hilde Spielvogel, and Jere D. Haas. "Effect of Developmental and Ancestral High Altitude Exposure on Chest Morphology and Pulmonary Function in Andean and European/North American Natives." *American Journal of Human Biology: The Official Journal of the Human Biology Council* 11, no. 3

(1999): 383–95. https://doi.org/10.1002/(SICI)1520-6300(1999)11:3<383::AID-AJHB9>3.0.CO;2-X.

Chang, Dan, Michael Knapp, Jacob Enk, Sebastian Lippold, Martin Kircher, Adrian Lister, Ross D. E. MacPhee, et al. "The Evolutionary and Phylogeographic History of Woolly Mammoths: A Comprehensive Mitogenomic Analysis." *Scientific Reports* 7 (March 22, 2017): 44585. https://doi.org/10.1038/srep44585.

Cheng, Jade Yu, Thomas Mailund, and Rasmus Nielsen. "Fast Admixture Analysis and Population Tree Estimation for SNP and NGS Data." *Bioinformatics* 33, no. 14 (July 15, 2017): 2148–55. https://doi.org/10.1093/bioinformatics/btx098.

Cheng, Jade Yu, Aaron J Stern, Fernando Racimo, and Rasmus Nielsen. "Detecting Selection in Multiple Populations by Modeling Ancestral Admixture Components." *Molecular Biology and Evolution* 39, no. 1 (January 1, 2022): msab294. https://doi.org/10.1093/molbev/msab294.

Crawford, Jacob E., Ricardo Amaru, Jihyun Song, Colleen G. Julian, Fernando Racimo, Jade Yu Cheng, Xiuqing Guo, et al. "Natural Selection on Genes Related to Cardiovascular Health in High-Altitude Adapted Andeans." *The American Journal of Human Genetics* 101, no. 5 (November 2, 2017): 752–67. https://doi.org/10.1016/j.ajhg.2017.09.023.

"Effects of Birthplace and Individual Genetic Admixture on Lung Volume and Exercise Phenotypes of Peruvian Quechua - Brutsaert - 2004 - American Journal of Physical Anthropology - Wiley Online Library." Accessed April 10, 2023. https://onlinelibrary-wiley-com.proxy.library.emory.edu/doi/full/10.1002/ajpa.10319.

"Ensembl Genome Browser 109." Accessed April 10, 2023. https://useast.ensembl.org/index.html.

Espinosa, Carlos. "Los indígenas Karanki, Natabuela, Otavalo y Kayambis danzaron en Ibarra." El Comercio, November 30, 1AD. https://www.elcomercio.com/actualidad/ecuador/indigenas-karanki-natabuela-otavalo-kayambis.html.

Frigola, Joan, Radhakrishnan Sabarinathan, Loris Mularoni, Ferran Muiños, Abel Gonzalez-Perez, and Núria López-Bigas. "Reduced Mutation Rate in Exons Due to Differential Mismatch Repair." *Nature Genetics* 49, no. 12 (December 2017): 1684–92. https://doi.org/10.1038/ng.3991.

Garrido, E., J. Botella de Maglia, and O. Castillo. "Acute, Subacute and Chronic Mountain Sickness." *Revista Clínica Española (English Edition)* 221, no. 8 (October 1, 2021): 481–90. https://doi.org/10.1016/j.rceng.2019.12.009.

Gautier, Mathieu, Alexander Klassmann, and Renaud Vitalis. "Rehh 2.0: A Reimplementation of the R Package Rehh to Detect Positive Selection from Haplotype Structure." *Molecular*

Ecology Resources 17, no. 1 (January 2017): 78–90. https://doi.org/10.1111/1755-0998.12634.

Gautier, Mathieu, and Renaud Vitalis. "Rehh: An R Package to Detect Footprints of Selection in Genome-Wide SNP Data from Haplotype Structure." *Bioinformatics (Oxford, England)* 28, no. 8 (April 15, 2012): 1176–77. https://doi.org/10.1093/bioinformatics/bts115.

"GeneCards - Human Genes | Gene Database | Gene Search." Accessed April 10, 2023. https://www.genecards.org/.

"Genetic Researcher Uses Nuu-Chah-Nulth Blood for Unapproved Studies in Genetic Anthropology | Ha-Shilth-Sa Newspaper," July 22, 2013. https://hashilthsa.com/news/2013-07-22/genetic-researcher-uses-nuu-chah-nulth-blood-unapproved-studies-genetic.

González Andrade, Fabricio, and Dora Sánchez. "Genetic Profile of the Kichwas (Quichuas) from Ecuador by Analysis of STR Loci." *Human Biology* 76, no. 5 (2004): 723–30. https://doi.org/10.1353/hub.2005.0007.

Greely, Henry T. "The Future of DTC Genomics and the Law." *Journal of Law, Medicine & Ethics* 48, no. 1 (ed 2020): 151–60. https://doi.org/10.1177/1073110520917003.

"GTEx Portal." Accessed April 10, 2023. https://gtexportal.org/home/.

Haas, Randall, Ioana C. Stefanescu, Alexander Garcia-Putnam, Mark S. Aldenderfer, Mark T. Clementz, Melissa S. Murphy, Carlos Viviano Llave, and James T. Watson. "Humans Permanently Occupied the Andean Highlands by at Least 7 Ka." *Royal Society Open Science* 4, no. 6 (June 2017): 170331. https://doi.org/10.1098/rsos.170331.

Hayes, Matthew. "Into the Universe of the Hacienda: Lifestyle Migration, Individualism and Social Dislocation in Vilcabamba, Ecuador." *Journal of Latin American Geography* 14, no. 1 (2015): 79–100. https://doi.org/10.1353/lag.2015.0001.

Hoopes, John W., and Maria Masucci. "Evaluating Pre-Columbian Contact between Ecuador and Costa Rica: A Ceramic Approach (Masucci and Hoopes 2022)." *Waves of Influence: Pacific Maritime Networks Connecting Mexico, Central America, and Northwestern South America*, January 1, 2022. https://www.academia.edu/79246890/Evaluating_Pre_Columbian_Contact_between_Ecuador_and_Costa_Rica_A_Ceramic_Approach_Masucci_and_Hoopes_2022_.

Huerta-Sánchez, Emilia, Michael DeGiorgio, Luca Pagani, Ayele Tarekegn, Rosemary Ekong, Tiago Antao, Alexia Cardona, et al. "Genetic Signatures Reveal High-Altitude Adaptation in a Set of Ethiopian Populations." *Molecular Biology and Evolution* 30, no. 8 (August 1, 2013): 1877–88. https://doi.org/10.1093/molbev/mst089.

Jafari, Mahjoobeh, Elham Ghadami, Tahereh Dadkhah, and Haleh Akhavan. “PI3k/AKT Signaling Pathway: Erythropoiesis and beyond: JAFARI et Al.” *Journal of Cellular Physiology* 234 (September 7, 2018). https://doi.org/10.1002/jcp.27262.

Jiang, B. H., G. Jiang, J. Z. Zheng, Z. Lu, T. Hunter, and P. K. Vogt. “Phosphatidylinositol 3-Kinase Signaling Controls Levels of Hypoxia-Inducible Factor 1.” *Cell Growth & Differentiation: The Molecular Biology Journal of the American Association for Cancer Research* 12, no. 7 (July 2001): 363–69.

Julian McFarlane, Soroya, Aurora Occa, Wei Peng, Oluwatumininu Awonuga, and Susan E. Morgan. “Community-Based Participatory Research (CBPR) to Enhance Participation of Racial/Ethnic Minorities in Clinical Trials: A 10-Year Systematic Review.” *Health Communication* 37, no. 9 (August 2022): 1075–92. https://doi.org/10.1080/10410236.2021.1943978.

Klein, Richard G. “Profile of Svante Pääbo: 2022 Nobel Laureate in Physiology or Medicine.” *Proceedings of the National Academy of Sciences* 120, no. 1 (January 3, 2023): e2217025119. https://doi.org/10.1073/pnas.2217025119.

Knapp, Gregory. “The Changing Kichwa Language Map in Ecuador,” 1731–42, 2020. https://doi.org/10.1007/978-3-319-73400-2_51-1.

León-Velarde, Fabiola, and Olga Mejía. “Gene Expression in Chronic High Altitude Diseases.” *High Altitude Medicine & Biology* 9, no. 2 (2008): 130–39. https://doi.org/10.1089/ham.2007.1077.

Lindo, John, Randall Haas, Courtney Hofman, Mario Apata, Mauricio Moraga, Ricardo A. Verdugo, James T. Watson, et al. “The Genetic Prehistory of the Andean Highlands 7000 Years BP Though European Contact.” *Science Advances* 4, no. 11 (November 8, 2018): eaau4921. https://doi.org/10.1126/sciadv.aau4921.

“Loja | Ecuador | Britannica.” Accessed April 10, 2023. https://www.britannica.com/place/Loja-Ecuador.

Mallick, Swapan, Heng Li, Mark Lipson, Iain Mathieson, Melissa Gymrek, Fernando Racimo, Mengyao Zhao, et al. “The Simons Genome Diversity Project: 300 Genomes from 142 Diverse Populations.” *Nature* 538, no. 7624 (October 13, 2016): 201-207H. https://doi.org/10.1038/nature18964.

Marciniak, Stephanie, and George H. Perry. “Harnessing Ancient Genomes to Study the History of Human Adaptation.” *Nature Reviews Genetics* 18, no. 11 (November 2017): 659–74. https://doi.org/10.1038/nrg.2017.65.

Maron, Bradley A., Roberto F. Machado, and Larissa Shimoda. “Pulmonary Vascular and Ventricular Dysfunction in the Susceptible Patient (2015 Grover Conference Series).”

Pulmonary Circulation 6, no. 4 (December 2016): 426–38. https://doi.org/10.1086/688315.

Matsunaga, D. S., R. Enos, C. C. Gotay, R. O. Banner, H. DeCambra, O. W. Hammond, N. Hedlund, E. K. Ilaban, B. F. Issell, and J. A. Tsark. "Participatory Research in a Native Hawaiian Community. The Wai'anae Cancer Research Project." *Cancer* 78, no. 7 Suppl (October 1, 1996): 1582–86.

Mueller, Benjamin. "Nobel Prize Awarded to Scientist Who Sequenced Neanderthal Genome." *The New York Times*, October 3, 2022, sec. Health. https://www.nytimes.com/2022/10/03/health/nobel-prize-medicine-physiology-winner.html.

Muñoz, Felipe Esteban Acosta. "Shunguhuan Yuyai: The Battle for Kichwa Language and Culture Revitalization in Ecuador As," n.d.

Naik, Piyu Parth, and Syed Nadir Farrukh. "Influence of Ethnicities and Skin Color Variations in Different Populations: A Review." *Skin Pharmacology and Physiology* 35, no. 2 (2022): 65–76. https://doi.org/10.1159/000518826.

Namjou, Bahram, Mehdi Keddache, Keith Marsolo, Michael Wagner, Todd Lingren, Beth Cobb, Cassandra Perry, et al. "EMR-Linked GWAS Study: Investigation of Variation Landscape of Loci for Body Mass Index in Children." *Frontiers in Genetics* 4 (December 3, 2013): 268. https://doi.org/10.3389/fgene.2013.00268.

Ninacuri, C.P. "The influence of Kichwa in Ecuadorian Andean Spanish: the case of the morpheme -ka." *Boletin de Filologia* 57, no. 1 (2022): 209–31. https://doi.org/10.4067/S0718-93032022000100209.

Palkopoulou, Eleftheria, Mark Lipson, Swapan Mallick, Svend Nielsen, Nadin Rohland, Sina Baleka, Emil Karpinski, et al. "A Comprehensive Genomic History of Extinct and Living Elephants." *Proceedings of the National Academy of Sciences* 115, no. 11 (March 13, 2018): E2566–74. https://doi.org/10.1073/pnas.1720554115.

Paralikar, Swapnil J. "High Altitude Pulmonary Edema-Clinical Features, Pathophysiology, Prevention and Treatment." *Indian Journal of Occupational and Environmental Medicine* 16, no. 2 (2012): 59–62. https://doi.org/10.4103/0019-5278.107066.

"PIK3CB Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Beta [Homo Sapiens (Human)] - Gene - NCBI." Accessed April 10, 2023. https://www.ncbi.nlm.nih.gov/gene/5291.

"Quechua." Accessed April 10, 2023. https://plc.sas.upenn.edu/quechua.

"Quechua: The Surviving Language of the Inca Empire | GVI." Accessed April 10, 2023. https://www.gviusa.com/blog/quechua-the-surviving-language-of-the-inca-empire/.

Rademaker, Kurt, Gregory Hodgins, Katherine Moore, Sonia Zarrillo, Christopher Miller, Gordon R. M. Bromley, Peter Leach, David A. Reid, Willy Yépez Álvarez, and Daniel H. Sandweiss. “Paleoindian Settlement of the High-Altitude Peruvian Andes.” *Science (New York, N.Y.)* 346, no. 6208 (October 24, 2014): 466–69. https://doi.org/10.1126/science.1258260.

Rademaker, Kurt, Gregory Hodgins, Katherine Moore, Sonia Zarrillo, Christopher Miller, Gordon R. M. Bromley, Peter Leach, David Reid, Willy Yépez Álvarez, and Daniel H. Sandweiss. “Cuncaicha Rockshelter, a Key Site for Understanding Colonization of the High Andes: Reply to Capriles et Al.” *Current Anthropology* 57, no. 1 (February 2016): 101–3. https://doi.org/10.1086/684826.

Rood, Kristiana, Vanessa Lopez, Michael R. La Frano, Oliver Fiehn, Lubo Zhang, Arlin B. Blood, and Sean M. Wilson. “Gestational Hypoxia and Programing of Lung Metabolism.” *Frontiers in Physiology* 10 (November 29, 2019): 1453. https://doi.org/10.3389/fphys.2019.01453.

Santos, Lorrieann. “Genetic Research in Native Communities.” *Progress in Community Health Partnerships : Research, Education, and Action* 2 (December 1, 2008): 321–27. https://doi.org/10.1353/cpr.0.0046.

Schmidt, C W. “Indi-Gene-Ous Conflict.” *Environmental Health Perspectives* 109, no. 5 (May 2001): A216–19.

Sikora, Martin, Vladimir V. Pitulko, Vitor C. Sousa, Morten E. Allentoft, Lasse Vinner, Simon Rasmussen, Ashot Margaryan, et al. “The Population History of Northeastern Siberia since the Pleistocene.” *Nature* 570, no. 7760 (June 2019): 182–88. https://doi.org/10.1038/s41586-019-1279-z.

“Speaking for a State: Standardized Kichwa Greetings and Conundrums of Commensuration in Intercultural Ecuador.” Accessed April 10, 2023. https://doi.org/10.1086/708164.

Sterling, Robyn L. “Genetic Research among the Havasupai: A Cautionary Tale.” *AMA Journal of Ethics* 13, no. 2 (February 1, 2011): 113–17. https://doi.org/10.1001/virtualmentor.2011.13.2.hlaw1-1102.

Tabata, Masahiko, Takahisa Tarumoto, Ken Ohmine, Yusuke Furukawa, Kiyohiko Hatake, Keiya Ozawa, Yuichi Hasegawa, Harumi Mukai, Masayuki Yamamoto, and Shigehiko Imagawa. “Stimulation of GATA-2 as a Mechanism of Hydrogen Peroxide Suppression in Hypoxia-Induced Erythropoietin Gene Expression.” *Journal of Cellular Physiology* 186, no. 2 (2001): 260–67. https://doi.org/10.1002/1097-4652(200002)186:2<260::AID-JCP1025>3.0.CO;2-K.

Taegtmeyer, Heinrich, Shiraj Sen, and Deborah Vela. “Return to the Fetal Gene Program.” *Annals of the New York Academy of Sciences* 1188 (February 2010): 191–98. https://doi.org/10.1111/j.1749-6632.2009.05100.x.

"UCSC Genome Browser Home." Accessed April 10, 2023. https://genome.ucsc.edu/.

Voight, Benjamin F., Sridhar Kudaravalli, Xiaoquan Wen, and Jonathan K. Pritchard. "A Map of Recent Positive Selection in the Human Genome." *PLoS Biology* 4, no. 3 (March 2006): e72. https://doi.org/10.1371/journal.pbio.0040072.

Wagner, Peter D., Mauricio Araoz, Robert Boushel, José A. L. Calbet, Birgitte Jessen, Göran Rådegran, Hilde Spielvogel, Hans Søndegaard, Harrieth Wagner, and Bengt Saltin. "Pulmonary Gas Exchange and Acid-Base State at 5,260 m in High-Altitude Bolivians and Acclimatized Lowlanders." *Journal of Applied Physiology (Bethesda, Md.: 1985)* 92, no. 4 (April 2002): 1393–1400. https://doi.org/10.1152/japplphysiol.00093.2001.

Watts, Deepika, Diana Gaete, Diego Rodriguez, David Hoogewijs, Martina Rauner, Sundary Sormendi, and Ben Wielockx. "Hypoxia Pathway Proteins Are Master Regulators of Erythropoiesis." *International Journal of Molecular Sciences* 21, no. 21 (October 30, 2020): 8131. https://doi.org/10.3390/ijms21218131.

"Whole Exome Sequence-Based Association Analyses of Plasma Amyloid-β in African and European Americans; the Atherosclerosis Risk in Communities-Neurocognitive Study | PLOS ONE." Accessed April 10, 2023. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0180046.

Xu, Wenhao, Yu Lin, Keliang Zhao, Haimeng Li, Yinping Tian, Jacob Njaramba Ngatia, Yue Ma, et al. "An Efficient Pipeline for Ancient DNA Mapping and Recovery of Endogenous Ancient DNA from Whole-Genome Sequencing Data." *Ecology and Evolution* 11, no. 1 (2021): 390–401. https://doi.org/10.1002/ece3.7056.

Yang, Min, Yang Wu, Xing-biao Yang, Tao Liu, Ya Zhang, Yue Zhuo, Yong Luo, and Nan Zhang. "Establishing a Prediction Model of Severe Acute Mountain Sickness Using Machine Learning of Support Vector Machine Recursive Feature Elimination." *Scientific Reports* 13 (March 21, 2023): 4633. https://doi.org/10.1038/s41598-023-31797-0.

Yang, Zhaohui, Caijuan Bai, Youwei Pu, Qinghong Kong, Yongbo Guo, Ouzhuluobu, Gengdeng, et al. "Genetic Adaptation of Skin Pigmentation in Highland Tibetans." *Proceedings of the National Academy of Sciences of the United States of America* 119, no. 40 (October 4, 2022): e2200421119. https://doi.org/10.1073/pnas.2200421119.

Yu, Kuai, Charity Duran, Zhiqiang Qu, Yuan-Yuan Cui, and H. Criss Hartzell. "Explaining Calcium-Dependent Gating of Anoctamin-1 Chloride Channels Requires a Revised Topology." *Circulation Research* 110, no. 7 (March 30, 2012): 990–99. https://doi.org/10.1161/CIRCRESAHA.112.264440.

Zeng, Bo, Jiangchao Zhao, Wei Guo, Siyuan Zhang, Yutong Hua, Jingsi Tang, Fanli Kong, et al. "High-Altitude Living Shapes the Skin Microbiome in Humans and Pigs." *Frontiers in Microbiology* 8 (October 6, 2017): 1929. https://doi.org/10.3389/fmicb.2017.01929.