Effects of High Altitude and Inorganic Lead (Pb) on DNA Methylation in Andeans

by

Ainash Childebayeva

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Anthropology and Toxicology)
in The University of Michigan
2019

Doctoral Committee:

      Assistant Professor Abigail W. Bigham, Co-Chair
      Professor Dana C. Dolinoy, Co-Chair
      Associate Professor Maureen Devlin
      Research Assistant Professor Jaclyn Goodrich

Ainash Childebayeva

ainash@umich.edu

ORCID iD 0000-0002-3144-6481

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This dissertation explores the relationship between the environment and the epigenome. Using gene-specific and genome-wide epigenetic approaches, this project answers two main questions: 1) Is there an epigenetic contribution to high-altitude adaptation in the Andes? 2) Does inorganic lead exposure affect the epigenomes of the indigenous Quechua?

Epigenetics is a study of changes to the DNA that do not change the sequence of the nucleotides, but can influence gene expression. Epigenetic modifications are mitotically, and sometimes meiotically heritable, and can be reversible. DNA methylation, histone tail modifications, and non-coding RNAs are the main epigenetic modifications. This study focuses on DNA methylation.

The following findings are reported. Global LINE-1 methylation is affected by exposure to high altitude at birth and current altitude exposure. Gene-specific methylation at the high-altitude adaptive gene, *EPAS1,* also is affected by current altitude exposure. This is the first study to show a decrease in its methylation associated with the altitude of recruitment. Moreover, we found significant DNA methylation changes associated with the altitude of birth.

We performed an epigenome-wide association study (EWAS) between the altitude of birth and DNA methylation, and found one significant position in the promoter region of *NPY1R* , a gene associated with pulmonary hypertension (PH), which is considered adaptive in Andeans. In addition, we identified 155 differentially methylated regions associated with hypoxic regulation, blood pressure, and pulmonary hypertension (PH).

DNA methylation changes in genes associated with neurological function and metal-ion binding have been reported with lead exposure. Additionally, a positive association between hemoglobin levels and inorganic lead, the distance from mining and lead exposure, and lead levels and the number of days since the last menstruation for women are reported. These findings demonstrate that even low levels of lead exposure can have a significant effect on the epigenome.

Overall, this dissertation research demonstrates that the environment, both past and present, can have a profound impact on epigenetic modifications. DNA methylation patterns can be affected by various exposures simultaneously making the study of epigenetics with regards to one specific exposure challenging. We have shown that both high-altitude hypoxia and lead are affecting the epigenome leading to changes that may be involved in adaptation to high altitude and buffering against the adverse effects of lead exposure.

CHAPTER 1

Introduction and Background

Epigenetics provides insight into understanding the extent to which the environment can affect individual biology (Mohn and Schubeler 2009). Epigenetic modifications provide the organism with a means to respond to external factors (Lv, et al. 2013), thus serving as a potential mechanism by which the body can adapt to ever-changing environmental conditions (Lam, et al. 2012). In this dissertation research, I determine how exposures to high-altitude hypoxia and inorganic lead (Pb) affect the epigenomes of high-altitude Andeans, and demonstrate that epigenetic changes contribute to high-altitude adaptation in Andeans by facilitating developmental adaptation.

Epigenetics is defined as the study of modifications to the DNA that do not affect the nucleotide sequence, but can modify gene expression. Different epigenetic mechanisms are at play during normal cell development, including, but not limited to, alterations of the chromatin structure, chemical alterations of nucleotides, post-translational and post-replicational modifications of DNA, and modifications to histones (Allis and Jenuwein 2016).

The most widely studied and best understood epigenetic modification is DNA methylation, a relatively accessible and stable marker (Mohn and Schubeler 2009; Lam, et al. 2012). DNA methylation occurs when a methyl group is added to the 5th carbon position of the base cytosine. This modification most commonly occurs when a cytosine is followed by a guanine in an arrangement called a cytosine-phosphate-guanine (CpG) site (Tollefsbol 2011). DNA methylation is carried out by a group of enzymes called DNA methyltransferases (DNMTs),

1

which selectively target specific DNA sequences to be either de novo or maintenance methylated (Fraga and Esteller 2002).

The period when the epigenome is most susceptible to environmental influences occurs very early in development, and it is characterized by periods of active and passive epigenetic reprogramming (Feil and Fraga 2011). For example, in mammals, this critical window is framed by waves of demethylation and reestablishment of global methylation levels occurring shortly after fertilization (Faulk and Dolinoy 2014). During this time, DNA methylation marks are completely removed from the genome and re-established de novo in a manner that is not fully understood (Messerschmidt, et al. 2014). However, mounting evidence suggests that early life events such as maternal stress levels, maternal nutrition, and exposure to heavy metals and endocrine disrupters such as bisphenol A (BPA), can affect the epigenome with outcomes predisposing the child to adverse health outcomes later in life (Dolinoy, et al. 2007; Senut, et al. 2012; Klengel, et al. 2013). Despite these recent advances in our understanding of epigenetic modifications and their effects on human health, the role of epigenetic change in modulating adaptation and adaptability remains to be explored. Human high-altitude adaptation provides a natural experimental design to study the effects of developmental exposure on changes to the epigenome.

Long-term high-altitude populations (e.g. Andeans, Tibetans, Ethiopians) can thrive in hypoxic regions with as little as 60% of the oxygen present at sea level, while sojourners often suffer acute or chronic mountain sickness (Moore 2001). Recent genome-wide studies have indicated a genetic basis for these adaptations in all three high-altitude populations (Beall, et al. 2010; Bigham, et al. 2010; Alkorta-Aranburu, et al. 2012). In fact, high-altitude adaptation is one of the strongest instances of natural selection acting on humans (Yi, et al. 2010). Yet, genetics alone does not fully explain the extent of variation we see in high-altitude adaptive phenotypes. Growth and development at high altitude are also hypothesized to contribute to these phenotypes

(Frisancho, et al. 1973; Frisancho, et al. 1997; Brutsaert 2001). We are using epigenetic methods to assess the role of high-altitude environment on high-altitude adaptation.

We assess the effect of lead (Pb) exposure on DNA methylation. Environmental contaminants are widespread, and humans are regularly exposed to a plethora of chemicals and toxicants, including Pb, BPA, arsenic, and phthalates (Marsit 2015). Pb is the world's most abundant heavy metal, and the most common neurotoxic pollutant in the environment (Tong, et al. 2000; Senut, et al. 2012). Occupational and environmental exposure to Pb is a serious health issue in both developed and developing countries (Tong, et al. 2000). Due to public health campaigns and removal of Pb from paint and gasoline, the abundance of Pb in the environment has decreased in many industrialized countries (Hernberg 2000). However, Pb is still present at high levels in many developing nations (Tong, et al. 2000). Several previous studies have shown that there is a relationship between Pb exposure and DNA methylation in humans (Hanna, et al. 2012; Engstrom, et al. 2015; Sen, et al. 2015). In this study, we determined the effects of Pb exposure on the epigenomes of the Quechua residents of Cerro de Pasco, a Peruvian high-altitude mining town, and Lima, the capital of Peru.

We determined DNA methylation associations between high-altitude exposure at birth and at time of recruitment, as well as the association between whole-blood Pb values and DNA methylation. DNA methylation was chosen as the marker of interest as it is the most frequently studied and best understood epigenetic modification (Klose and Bird 2006). We focused on three high-altitude developmental exposure groups: (1) Quechua individuals born and raised at high altitude and exposed to inorganic lead from mining (n = 301; recruited in Cerro de Pasco), (2) Quechua individuals born and raised at low altitude and not exposed to lead from mining (n = 152; recruited in Lima), and (3) Quechua individuals born in various high altitude locations who moved to low altitude as adolescents, and have not been exposed to lead from mining (n = 150; recruited in Lima).

We determined the effects of high-altitude and Pb exposure on gene-specific and genome-wide DNA methylation patterns in n=603 and n=87 individuals, respectively. We performed association analysis between high-altitude exposure at birth and at the time of recruitment and DNA methylation. We also determined associations between whole-genome Pb values and DNA methylation for a subset of participants with available Pb data.

BACKGROUND

Epigenetics and adaptation

Darwinian evolutionary theory in the age of the Modern Synthesis is based on the idea that natural selection acts on heritable genetic variation that contributes to phenotype. Epigenetics is the study of stable modifications to DNA that do not change the sequence itself. Rather, these modifications result in changes to gene expression. Genetically inherited predisposition for phenotypic variability can be modified through epigenetic changes that serve as an adaptive mechanism in changing environments. This ability to fine to genetic adaptation is particularly important given epigenetic modifications are reprogrammed in early development and can be acquired throughout lifetime (Feinberg and Irizarry 2010; Kanherkar, et al. 2014).

Epigenetics studies mitotically and, in some cases, meiotically heritable changes to gene expression that do not involve changes to DNA sequence (Wolffe and Guschin 2000; Feil and Fraga 2011). Epigenetic modifications undergo reprogramming during early development, which is when the epigenome is the most susceptible to environmental cues (Reik 2007). However, environmental exposures can influence the epigenome throughout life, albeit to a lesser extent (Alegria-Torres, et al. 2011; Kanherkar, et al. 2014). There is a constant interaction between the environment and the molecular processes inside the body (Lerner and Benson 2013). For example, exposures to toxicants, pharmaceutical agents, exercise, stress, and other factors have been shown to associate with changes in epigenetic marks, such as DNA methylation (Csoka and Szyf 2009; Senut, et al. 2012; Ronn, et al. 2013; Non, et al. 2016).

4

Epigenetic mechanisms have been proposed to occupy an intermediate role of "developmental adaptation" between physiological adaptation and genetic adaptation by natural selection (Kuzawa and Thayer, 2011). They mediate phenotypic variation in plants, mice, rats, and even in human monozygotic twins (Dolinoy et al., 2007; Weaver et al. 2004, Cubas et al., 1999). For example, toadflax (*Linaria*) flowers originally described by Linnaeus change their symmetry from bilateral to radial with excessive methylation that results in transcriptional silencing of the gene *Lcyc* (Cubas et al., 1999). Similarly, genetically identical mice develop different coat colors based on the methylation status of the *Agouti* locus, which can be affected by environmental exposures (Dolinoy et al., 2007; Dolinoy et al. 2008). In humans, monozygotic twins often display phenotypic differences including differential disease susceptibility that are linked to unique epigenetic modifications, and accumulate different epigenetic changes with age (Fraga, et al. 2005; Wong, et al. 2005).

The epigenome has been shown to play a role in adaptation of plants and animals. For example, *Arabidopsis* plants adapt to stressors like high salt, cold, and heat via changes in DNA methylation (Boyko, et al. 2010; Boyko and Kovalchuk 2010; Szyf and Bick 2013; Dubin, et al. 2015). In rats, maternal licking and grooming have been shown to alter DNA methylation levels of the glucocorticoid receptor gene promoter in the hippocampus of the offspring. This leads to behavioral changes in the offspring (Weaver, et al. 2004). Together, these studies demonstrate that epigenetic modifications contribute to phenotypic variation, and potentially mediate an organism's ability to adapt to various environments, including high altitude.

High-altitude adaptation

There are three main geographic regions where humans have adapted to life at high altitude: the Tibetan Plateau, the Ethiopian Plateau, and the Andean Altiplano (Moore 2001; Bigham 2016). Due to convergent evolution, populations from each of these zones exhibit distinct altitude-adaptive phenotypes compared to low-altitude populations, and each population presents

a unique suite of genetic adaptation (Bigham and Lee 2014). Out of the three high-altitude populations, Andeans are most suited for the study the epigenetic effects of high-altitude adaptation as previous work have demonstrated compelling evidence for a developmental component to their adaptive phenotype. This is particularly the case for several adaptive traits observed among Quechua residents of the Altiplano including changes to lung volume and arterial oxygen saturation that are not observed in Quechua residents who were born and raised at low altitude (Frisancho 1970, 1975, 1997; Kiyamu, et al. 2015). This suggests an epigenetic component in the Andean developmental adaptation to high altitude.

Andean high-altitude adaptation

The Andean Altiplano is a high-altitude plateau along the western coast of South America, stretching 50-100 miles from the Pacific coast to the Amazonian lowlands. According to archaeological data, humans have lived in the Andes for more than 10,000 years (Rademaker, et al. 2014). Despite a lack of evidence supporting genetic continuity between modern Andean populations and ancient inhabitants of the Altiplano, archaeological data show that the region has a long history of human occupation (Aldenderfer 2008).

Some of the first studies of human high-altitude adaptation were performed in the Andes. In 1923, Barcroft et al. published the results of the Cerro de Pasco expedition in "Observations upon the Effect of High Altitude on the Physiological Processes of the Human Body, Carried out in the Peruvian Andes, Chiefly at Cerro de Pasco" reporting increased pulmonary ventilation, decreased arterial oxygen saturation, and increased hemoglobin levels for the Andeans natives (Barcroft, et al. 1923).

In 1932, Alberto Hurtado reported smaller height and weight in Andeans compared to Europeans and Chileans living at low altitude (Hurtado 1932a). Despite small stature and body weight, chest circumference to body surface area ratio was noted to be higher in high-altitude born Andeans compared to the Andeans born at low altitude, as well as Europeans (Hurtado

1932a). Hurtado also reported higher red blood cell counts and decreased fragility of red blood cells in Andeans at high altitude (Hurtado 1932b). Forced vital capacity (FVC) in altitude-born Andeans is reportedly higher than that of sea level individuals acclimatized to altitude as adults, while the FVC of sea level individuals acclimatized to altitude in the developmental period is similar to high-altitude native FVC (Frisancho, et al. 1973). Moreover, high-altitude Andeans tested at altitude show significantly higher lung volume (RV) than non-altitude natives (Hurtado 1932a).

Overall, Andeans display population-specific adaptations to high altitude including enlarged chests, increased lung capacities, increased pulmonary diffusion, elevated hematocrit and hemoglobin, and preferential use of carbohydrates as fuel (Rupert and Hochachka 2001). They also exhibit a blunted (low) hypoxic ventilatory response (Beall, et al. 1997) and higher arterial oxygen saturation (SaO_2) at rest and during exercise (Brutsaert, et al. 2000). Like their Tibetan counterparts, Andean women are protected from intrauterine growth restriction (IUGR) caused by high altitude (Moore, et al. 2011). Consequently, infants born to Andean mothers do not exhibit altitude-associated reductions in birth weight that infants born to low-altitude mothers residing at high altitude experience. Emerging evidence suggests a genetic basis for many of these traits (Bigham, et al. 2014).

Genetics of High-Altitude Adaptation

Genome-wide screens for natural selection performed in Andeans have identified evidence for natural selection in many hypoxia-inducible factor (HIF) pathway genes (Bigham, et al. 2009; Bigham, et al. 2010). HIFs are a group of transcriptional factors that respond to environmental decreases in oxygen, and activate more than 100 genes implicated in various physiological processes (Ziello, et al. 2007). Unpublished follow-up work by Dr. Bigham has revealed significant associations with a SNP near the natural selection candidate locus endothelial PAS domain-containing protein 1 gene (*EPAS1*) and hemoglobin concentration.

7

Additionally, maternal polymorphisms in AMP-activated, alpha 1 catalytic subunit (*PRKAA1*) show significant associations with infant birth weight and metabolic homeostasis, suggesting that variation in this gene may influence Andean high-altitude adaptation (Bigham, et al. 2014). One of the most compelling cases of natural selection for Andeans is found in the cellular oxygen sensing gene egl-9 family hypoxia-inducible factor 1 (*EGLN1*) (Bigham, et al. 2010). Forthcoming work demonstrates an association with the maximal consumption of oxygen during oxygen (VO$_2$Max) for this locus (Bigham, et al. 2013).

Among Tibetans, two genes showing compelling evidence of natural selection contain SNPs that associate with lower hemoglobin levels. They include *EPAS1* and *EGLN1* (Beall, et al. 2010; Simonson, et al. 2010). Moreover, an *EGLN1* SNP has been found to be associated with increased erythropoiesis in normal oxygen (normoxic) conditions in the Tibetans (Lorenzo, et al. 2014).

Developmental adaptation hypothesis

Environmental influences during growth and development are capable of changing the expression of a person's genetic potential (Frisancho 1977). In late 1960s and early 1970s, based on studies of phenotypic variation associated with short- and long-term high-altitude adaptation in Peruvian and U.S. sea-level and altitude natives, Baker and Frisancho put forward the "developmental adaptation hypothesis." This hypothesis suggests that humans who have developed in hypoxic conditions present phenotypes that contribute to their adaptation to high altitude (Baker 1969; Frisancho 1970, 1975, 1977). Two major paradigms have come out of the developmental adaptation hypothesis: 1) the younger the organism, the greater the influence of the environment, and 2) the younger the organism, the more adaptable it is (Frisancho 2009) (Frisancho, 2009).

Empiric data support the developmental adaptation hypothesis of high-altitude adaptive phenotypes. Based on anthropomorphic data collected from residents of La Paz, Bolivia (3,750

m), Frisancho et al. (1997) showed that high-altitude exposure during growth and development is crucial for attaining the enlarged residual lung volume characteristic of long-term high-altitude inhabitants, and is related to both genetic and developmental factors (Frisancho, et al. 1997). Following the steps of Frisancho et al., (1997), other studies have shown the effects of developmental exposure to hypoxia on the adaptive phenotype. For example, genetic potential for increased lung capacity associated with high-altitude adaptation depends on exposure to hypoxia during critical periods in growth and development (Brutsaert, et al. 1999). In other words, individuals exposed to hypoxia during development have larger lung volumes than individuals not exposed to hypoxia. This suggests that there is a high degree of phenotypic plasticity in high-altitude adaptation, and that much of this plasticity is shaped early on in development.

Epigenetics of hypoxia

The epigenetic processes involved in human adaptation to high altitude remain unknown. However, epigenetic aspects of hypoxia have been well studied in relation to cancer, as hypoxia is a hallmark of most tumors that grow and proliferate at a rate that outpaces their blood supply (Wilson and Hay 2011). Thus, tumor cells must adapt to an environment with low oxygen concentrations, similar to humans living at high altitude. HIFs have been shown to work in concert with epigenetic mechanisms (Watson, et al. 2010). Not surprisingly, the HIF pathway is altered in cancer cells to adjust to low oxygen conditions (Rawluszko-Wieczorek, et al. 2014). Intriguingly, *EPAS1* transcription is mediated by DNA methylation in colorectal cancer (Rawluszko-Wieczorek, et al. 2014).

Given HIF is at the center of the body's hypoxia response, we expect to see high-altitude adaptive changes in the genes related to this pathway. Previous work by Alkorta-Aranburu et al. 2012 identified four CpG sites with significantly different methylation levels in the high- versus low-altitude Ethiopians. However, they were not able to unambiguously connect differentially methylated CpG loci with genes relevant to hypoxia response (Alkorta-Aranburu, et al. 2012).

Lead and epigenetics

Lead is a trace metal that has been associated with human civilization since the emergence of metallurgy (Mielke and Reagan 1998). Lead is the world's most abundant heavy metal, and exposure to it is estimated to account for 143,000 deaths per year, and 0.6% of the global disease burden (Tong, et al. 2000; Haefliger 2011).

Pb exposure can have negative effects on any organ in the body, including the neurological, hematological, cardiovascular, and renal systems (Kalia and Flora 2005; Sanders, et al. 2009; Haefliger 2011). However, the nervous system is the most susceptible to harmful effects of lead (Sanders, et al. 2009). Both central and peripheral nervous systems can be affected by Pb, with the central nervous system being more affected in adults and the peripheral nervous system being more affected in children (Flora, et al. 2012). In children, Pb exposure has been associated with encelopathy, which is characterized by a degeneration of certain parts of the brain (Flora, et al. 2012). Children exposed to lead may also experience delayed growth, decreased IQ, impairment in concentration ability, and other adverse outcomes (Needleman, et al. 1979; Needleman and Leviton 1979; Sanders, et al. 2009). Additionally, Pb has a direct effect on the hematopoietic system by inhibiting enzymes involved in the heme synthesis pathway, and thus leading to decreased hemoglobin concentration (Albahary 1972). Pb also increases the fragility of erythrocyte membranes, which reduces their lifespan (Lessler and Walters 1973; Quintanar-Escorza, et al. 2007). Thus, a common outcome of lead exposure is anemia (Landrigan, et al. 1976; Flora, et al. 2012). Lastly, renal dysfunction occurs at high levels of lead exposure, but even relatively low levels of lead exposure can be damaging (Grant 2009b).

Pb is predominantly stored in calcified tissues, and more than 94% of it is deposited in bone, which represents the largest lead storage site in the body (Barbosa, et al. 2005). Lead stored in bone is not homogeneous and static. The deposition of tissue lead is dependent on the density and structure of the underlying bone, and its turnover is high throughout life (Grant 2009a).

10

For example, in adult women 50 – 75% of blood lead is skeletal in origin rather than environmental (Gulson and Calder 1995; Gulson, et al. 1995). And the release of bone Pb is increased during periods of pregnancy and lactation, which are associated with lower calcium availability (Gulson, et al. 2003).

Lastly, environmental metals, including lead, can increase the production of reactive oxygen species (ROS) via redox cycling (Ahamed and Siddiqui 2007). Oxidative damage to the DNA can hamper with the ability of DNA methyltransferases to interact with the DNA, which may result in altered DNA methylation levels (Baccarelli and Bollati 2009). Previous studies have shown changes in DNA methylation associated with Pb exposure in humans and mice (Faulk, et al. 2013; Sen, et al. 2015).

The primary medium for human lead exposure is air. Pb particles can travel long distances before precipitating as wet, dry, or cloud deposition (Grant 2009a). Environmental lead exposure is mediated through inhalation and ingestion of lead particles (Davidson and Rabinowitz 1991). Soil is an important pathway for lead exposure in humans, and its Pb levels have been shown to associate with lead poisoning (Mielke and Reagan 1998). We studied the association between whole-blood lead levels and methylation, hemoglobin levels, proximity to mining, and soil lead values in the mining town of Cerro de Pasco, Peru.

Mining in Cerro de Pasco

One of the high-altitude exposure groups was recruited in Cerro de Pasco, Peru, an important Peruvian mining center (Helfgott 2013). Cerro de Pasco was established in the 17th century to support silver mining (Fischer 1977). In the late 18th century it was the second largest producer of silver in the Andes (Helfgott 2013). Upon the depletion of the city's silver reserves, efforts now focus on copper, zinc, and lead extraction (van Geen, et al. 2012). In 1952, the Cerro de Pasco Corporation claimed to be the largest South American producer of refined lead (Helfgott 2013). Interestingly, inorganic lead pollution in the region predates Spanish colonial extraction.

Early lead pollution is credited to pre-colonial silver smelting, and, based on the study of the sediments from Lake Llamacocha, the earliest evidence of anthropogenic lead enrichment was dated to ca. A.D. 600 (Cooke, et al. 2009).

Pb deposition increased through the twentieth century reaching the highest values in A.D. 1968 (Cooke, et al. 2009). A study conducted in 2007 found that whole blood lead levels (WBLL) in children and women from Cerro de Pasco were significantly above 5 µg/dL, which is the blood lead threshold set by the CDC (Conklin L., et al. 2008). Another study conducted in 2012 reported measures of lead concentration in the soil around the mine that were above 1200 mg/kg in the 1km diameter around the open pit, while the EPA-suggested threshold for toxicity in soil is 400 mg/kg (van Geen, et al. 2012).

Lead and Epigenetics

Previous studies of humans and mice have shown that lead exposure is associated with changes in DNA methylation. Early Pb exposure was associated with DNA methylation differences in retrotransposon genes in weanling mice (Faulk, et al. 2013). In humans, gene specific and genome-wide DNA methylation differences associated with Pb exposure have been reported. For example, DNA methylation differences in the repetitive element LINE-1 and imprinted gene *H19* were linked to Pb exposure (Faulk, et al. 2013; Goodrich, Dolinoy, Sanchez, et al. 2016). Moreover, a statistically significant negative correlation between blood lead levels (BLLs) and promoter methylation of the gene *COL1A2*, which has been found to contain polymorphisms that are associated with preterm birth in adult women undergoing in vitro fertilization, was observed (Hanna, et al. 2012). Previous studies using the Illumina HumanMethylation450 BeadChip found Pb-related changes in neuroepigenetic signaling pathways in human embryonic stem cells, as well as the transmission of DNA methylation changes associated with Pb exposure from pregnant women to their grandchildren (Senut, et al.

2014; Sen, et al. 2015). In this study, we determined DNA methylation changes associated with whole-blood lead levels in individuals exposed to lead from mining versus the unexposed.

STUDY DESIGN

Study population

Study participants consist of Peruvian Quechua previously recruited in Lima and Cerro de Pasco as part of NSF BCS-1132310 based on three high-altitude / developmental exposure patterns (Fig. 1.1). (1) High-altitude group. Born and raised at high altitude (301 individuals from Cerro de Pasco, Peru located at an elevation of 4300 meters (m) above sea-level). (2) Low-altitude group. Quechua individuals born and raised at low altitude (152 individuals from Lima, Peru at an elevation of 0 m). (3) High-to-low group. Quechua who were born in various high-altitude locations (>2,500 m) and moved to low altitude Lima as adults (n=150 individuals recruited in Lima, Peru).

Figure 1.1. Map of Peru with Lima and Cerro de Pasco indicated by stars

RESEARCH QUESTIONS AND CHAPTERS

In this thesis, I examine the relationship between high-altitude exposure and inorganic lead exposure and DNA methylation in the Andean Quechua. Specifically, I answer three broad questions associated with the effects of altitude and lead on the epigenetics: 1) Are there gene-specific changes association with high-altitude exposure in the Peruvian Quechua? (Chapter 2) 2) Are there genome-wide DNA methylation differences associated with high-altitude adaptation in the high-altitude Quechua? (Chapter 3) 3) Are there genome-wide DNA methylation changes associated with inorganic lead exposure in Quechua exposed to inorganic lead from mining versus the unexposed? Are associations between whole-blood lead levels and hemoglobin levels, proximity to mining, and soil lead values? (Chapter 4). I am the first author on all chapters,

however, each chapter represents a collaborative effort with multiple coauthors, including my dissertation co-chairs Drs. Abigail Bigham and Dana Dolinoy. I will use the pronoun "we" throughout this dissertation to reflect the collaborative nature of this work.

Question 1: Are there gene-specific changes association with high-altitude exposure in the Peruvian Quechua?

Chapter 2 is focused on the analysis of global methylation at the LINE-1 repetitive element and the high-altitude adaptive gene *EPAS1*. LINE-1 DNA methylation has been hypothesized to reflect the state of the overall methylome (Irahara, et al. 2010). It has been associated with environmental exposures, including exposure to toxicants and diet, as well as health and disease outcomes, such as cancer (Chalitchagorn, et al. 2004; Baccarelli, et al. 2009; Boeke, et al. 2012; Goodrich, Dolinoy, Sanchez, et al. 2016). We show that birth at high altitude is associated with higher LINE-1 methylation, and current altitude is associated with increased LINE-1 methylation. Moreover, LINE-1 methylation is associated with the number of years lived at high altitude. suggesting a potentially adaptive effect of increased methylation in Andeans, the amount of the Andean genetic ancestry is associated with higher LINE-1 methylation.

Focusing on gene-specific methylation changes, we focused our attention on *EPAS1.* This gene has been shown to be under selection in Tibetans and possibly Andeans (Beall, et al. 2010; Eichstaedt, et al. 2017). We show decreased *EPAS1* methylation is associated with current altitude of residence, with individuals residing at high-altitude showing the lowest levels of methylation. Based on previous work demonstrating decreased methylation increases *EPAS1* expression, we hypothesize that this decreased methylation may lead to activation of *EPAS1* at high altitude (Rawluszko-Wieczorek, et al. 2014).

Together, these findings suggest that exposure to high-altitude hypoxia during early life, as well as throughout life, is associated with changes in DNA methylation. Ours is the first study to report changes in LINE-1 and *EPAS1* methylation associated with high altitude.

Question 2: Are there genome-wide DNA methylation differences associated with high-altitude adaptation in the high-altitude Quechua?

In Chapter 3 we discuss the effects of high-altitude exposure on genome-wide DNA methylation signatures analyzed using the Illumina MethylationEPIC bead chip. In this study, we performed an epigenome-wide association study to identify DNA methylation differences associated with the altitude of birth. We found one significant differentially methylated CpG site in the promoter region of the gene *NPY1R* that has previously been associated with pulmonary hypertension and cold exposure (Zukowska-Grojec, et al. 1987; Raposinho, et al. 2004; Kuo, et al. 2007; Crnkovic, et al. 2014). We also identified differentially methylated regions associated with genes under hypoxic regulation, and with genes linked to pulmonary hypertension and blood pressure. We found differentially methylated regions associated with genes *PKLR*, *UPS29, and YAP1* involved in hypoxia regulation suggesting a potential role in high-altitude adaptation. We validated one significant differentially methylated region associated with Superoxide Dismutase 3 (*SOD3)* in the larger cohort of Quechua (n=499). Consistent with the EWAS finding, we identifyied higher methylation in individuals born at high altitude. Finally, we determined the epigenetic age based on the method developed by Horvath (Horvath 2013), and found accelerated aging associated with the altitude of birth, as well as higher hemoglobin levels.

Question 3: Are there genome-wide DNA methylation changes associated with inorganic lead exposure in Quechua exposed to inorganic lead from mining versus the unexposed? Are associations between whole-blood lead levels and hemoglobin levels, proximity to mining, and soil lead values?

Chapter 4 is focused on studying the effects of inorganic lead exposure on the body. Lead exposure is associated with changes in hemoglobin levels, since it is known to bind to erythrocytes (Bergdahl, et al. 1997). We found a positive association between WBLLs and hemoglobin

concentration in our Peruvian cohort. We performed a spatial analysis of lead exposure distribution, and observed a negative correlation between WBLLs and proximity to mining, with WBLL decreasing with distance from the mine. We interpolated soil lead values based on previously published data (van Geen et al., 2012) using kriging in ArcGIS. This analysis revealed a suggestive association between soil lead and blood lead. Lastly, we saw a positive relationship between lead levels and the number of days since the last menstruation for women.

Lead exposure is associated with both genome-wide and gene-specific DNA methylation changes (Sen, et al. 2015; Goodrich, Dolinoy, Sanchez, et al. 2016). We performed an EWAS between genome-wide DNA methylation levels and WBLLs. We identified significant differentially methylated positions and regions associated with genes linked to neurological function and disorders as well as heavy metal exposure and toxicity. Lead is a known neurotoxicant, and has been linked to decreased cognitive function in children (Needleman, et al. 1979; Needleman and Leviton 1979).

Overall, these finding demonstrate that low levels of lead exposure can have significant effects on the organism, including changes in hemoglobin levels and epigenetic changes in genes associated with neurological function and heavy metal exposure.

CHAPTER 2

LINE-1 and *EPAS1* DNA Methylation Associations with High-Altitude Exposure

INTRODUCTION

There are three main geographic regions where humans have adapted to life at high altitude including the Tibetan Plateau, the Ethiopian Plateau, and the Andean Altiplano (Moore 2001). Human populations from each of these zones exhibit unique circulatory, respiratory, and hematological adaptations with respect to low-altitude populations and with respect to each other. These adaptations include differences in resting ventilation, hypoxic ventilatory response, arterial oxygen saturation (SaO_2), and hemoglobin concentration among others (Bigham, et al. 2014). Genome-wide single-nucleotide polymorphism (SNP) and sequencing studies support a genetic contribution to many of these adaptations in all three high-altitude populations (Beall, et al. 2010; Bigham, et al. 2010; Simonson, et al. 2010). However, genetics alone does not fully explain the extent of variation observed for high-altitude adaptive phenotypes. Growth and development at high altitude also have been shown to contribute to high-altitude adaptive traits.

This is particularly the case among Andeans, for whom developmental adaptation has been shown to contribute to altitude-adaptive traits (Frisancho, et al. 1973; Frisancho, et al. 1997; Brutsaert 2001). For example, exposure to hypoxia during growth and development is crucial for attaining the enlarged residual lung volume characteristic of high-altitude Andeans, and is related to both genetic and developmental factors (Frisancho, et al. 1997; Brutsaert, et al. 2000). Likewise, SaO_2 exhibits both a genetic and a developmental component among Andean Quechua (Bigham, et al. 2008; Kiyamu, et al. 2015). The angiotensin-converting enzyme (*ACE*)

17

gene insertion/deletion (I/D) polymorphism explains ~4% of the total variance in resting and exercise SaO_2 measured in Quechua individuals born and raised at high altitude in Cerro de Pasco, Peru (4,388 meters [m]) and Quechua individuals born and raised at low altitude in Lima, Peru (0 m), but the group effect of being born at high versus low altitude accounts for most of the variance in SaO_2 (Bigham, et al. 2008). The developmental component to high-altitude adaptation may proceed through epigenetic modification. With the rise of epigenetic methods, we now have a direct way to examine mitotically heritable changes in gene expression that occur during growth and development.

Epigenetic regulation works hand-in-hand with genetic processes to 'prepare' the organism for the external environment where it will live and reproduce providing a potential mechanism by which it can adapt to ever-changing environmental conditions (Lam, et al. 2012; Lv, et al. 2013). Epigenetic modifications include DNA methylation, posttranslational histone tail modifications, and non-coding RNAs. DNA methylation is the most widely studied and best understood epigenetic modification (Mohn and Schubeler 2009; Lam, et al. 2012). DNA methylation occurs when a methyl group is added to the 5th carbon position of a cytosine. This modification most commonly occurs when cytosine is followed by guanine in an arrangement called a cytosine-phosphate-guanine (CpG) site (Tollefsbol 2011). DNA methylation is carried out by a group of enzymes called DNA methyltransferases (DNMT), which selectively target specific DNA sequences to be either *de novo* or maintenance methylated (Fraga and Esteller 2002).

We determined DNA methylation levels at the LINE-1 repetitive DNA element and a single candidate locus, the gene *EPAS1,* in a cohort of indigenous Quechua Andeans. The Long Interspersed Nuclear Element-1 (LINE-1) is a repetitive DNA element that constitutes about 21% of the human genome and is often used as a marker of global DNA methylation (Lander, et al. 2001). The LINE-1 retrotransposon is highly methylated in non-diseased states (Nelson, et al. 2011). LINE-1 methylation can affect the expression of nearby genes (Belancio, et al. 2009), and decreased methylation of LINE-1 has been shown to be associated with genomic instability, risk

of cancer, diet, and exposures to toxicants (Ogino, et al. 2008; Anderson, et al. 2012; Woo and Kim 2012). Its high genomic frequency and its association with environmental exposures and disease state make LINE-1 a suitable genetic marker to study the effects of high-altitude hypoxia on the epigenomes of indigenous Peruvians.

Endothelial PAS domain protein 1 (*EPAS1*), also known as HIF-2a, is a hypoxia-responsive transcription factor that has been shown to be under selection in high-altitude Tibetans and Andeans (Beall, et al. 2010; Foll, et al. 2014; Eichstaedt, et al. 2017). *EPAS1* is a hypoxia-activated gene. Its expression is higher in hypoxic conditions, and it was previously hypothesized to be involved in the high-altitude response (Semenza 2001a; Nanduri, et al. 2017). Furthermore, *EPAS1* methylation levels previously have been shown to associate with colorectal cancer, with increased methylation leading to decreased expression of the EPAS1 protein (Rawluszko-Wieczorek, et al. 2014). Together, these findings make *EPAS1* a natural a priori candidate for studying epigenetics of high-altitude adaptation.

We used quantitative bisulfite pyrosequencing to determine DNA methylation levels at four LINE-1 CpG sites in 572 Peruvian Quechua study participants of similar genetic ancestry with three different developmental exposures to high altitude. To explore gene-specific methylation outcomes of hypoxia exposure, we used quantitative bisulfite pyrosequencing to characterize methylation at *EPAS1* in this same cohort (n = 539). We identified significant DNA methylation differences among the three groups at LINE-1. Furthermore, we found that *EPAS1* methylation is decreased among Peruvian Quechua living at high altitude compared to those living at low altitude.

RESULTS

Study design and participant characteristics

We employed a migrant study design wherein we recruited a cross-sectional cohort of 603, healthy Quechua participants aged 18-35 from two locations in Peru differing by altitude:

Lima (0 meters [m] above sea level) and Cerro de Pasco (4,338 m) (Fig. 2.1A). From these two locations, we recruited three high-altitude exposure groups. 1) high-altitude Quechua (HAQ) (*n*=300) were multigenerational high-altitude Quechua individuals who were born and raised in Cerro de Pasco. None of the HAQ study participants were mine workers even though Cerro de Pasco is a mining town. 2) migrant Quechua (MQ) *(n=152)* were Quechua individuals who were born in highland Peru (> 2,500 m) and moved to low-altitude Lima (0 m) during their lifetime (Fig. 2.1B). They had resided in Lima, Peru for at least two months at the time of enrollment. Thus, this study group was composed of first generation down migrants who have had full developmental exposure to high altitude. 3) low-altitude Quechua (LAQ) (*n*=150) were Quechua individuals who were born, raised, and reside in in Lima (150 m). Thus, LAQ is composed of second-generation down-migrants who have had no developmental exposure to high altitude over the course of their lifetime, but their parents and both sets of grandparents were born at an altitude > 3,000 m (Fig. 2.1B). It is important to note the Quechua population is a highland group. The Quechua population has a history of life at high altitude, and all individuals in this study are genetically adapted to high-altitude conditions (Bigham et al., unpublished).

DNA was extracted from whole blood, and quantitative bisulfite pyrosequencing was performed. We excluded samples that repeatedly failed pyrosequencing or produced unreliable DNA methylation data. Unreliable data were defined as samples the repeatedly failed pyrosequencing or generated DNA methylation values outside 2 standard deviations of the entire cohort. After quality control (QC), statistical analysis was performed on 572 samples for LINE1 and 539 samples for *EPAS1*). Please see "Materials and Methods" for a more detailed description of our study cohort.

Participant characteristics are provided in Table 2.1. The three study groups were different by height, weight, and BMI (p<0.05). MQ study participants were slightly but significantly older than both HAQ and LAQ study participants (p<0.05). The study groups did not differ significantly by sex. Average age of down-migration for MQ was 15.21 ± 6.37 years of age. Principal

component analysis (PCA) was performed on Quechua DNA samples and three 1000 genomes populations (90 East Asians, 60 *Centre de Polymorphism Humain* [*CEPH*] Europeans, and 60 Yorubans). We used existing SNP microarray genotype data for the Peruvians in this study from the Affymetrix (Santa Clara, CA) Axiom Biobanking array that includes ~600,000 single nucleotide polymorphisms (SNPs). PCA was performed on 328,260 autosomal SNPs with genotyping rates > 95%, MAF > 0.0001, that passed Affy QC, and that were not in strong linkage disequilibrium (R^2<0.8). Together, all three, altitude-exposure groups formed a distinct cluster (Supp. Fig. 2.1A-C).

Sex and Age effects on DNA methylation

We performed a linear mixed effects regression to determine the effects of sex and age on LINE-1 and *EPAS1* methylation. Linear mixed models allow for multiple technical pyrosequencing replicates per individual. We considered average LINE-1 and *EPAS1* methylation across their four respective sites in addition to DNA methylation at each of the four sites independently. Males showed higher LINE-1 and *EPAS1* methylation than females on average (LINE-1: β=0.67, p-value=4.37e-08; *EPAS1*: β=1.13, p-value=2.47e-05) and at each CpG site when adjusted for age (S2.1 Table) consistent with previous research showing DNA methylation associations with sex (El-Maarri, et al. 2007). At the first LINE-1 CpG site (CpG1), older individuals displayed significantly lower levels of methylation (β=-0.037, p-value<0.05, adjusted for sex and altitude). All other LINE-1 CpG positions and the average LINE-1 methylation were not associated with age. *EPAS1* methylation levels were positively associated with age (β=0.09, p-value<0.05, adjusted for sex and city of recruitment).

Genetic ancestry contributes to LINE-1 methylation

DNA methylation signatures differ by human population and are known to contribute to variation between populations (Fraser, et al. 2012; Heyn, et al. 2013). Moreover, a previous study

has shown higher LINE-1 methylation in non-Hispanic whites compared to non-Hispanic blacks in the United States (Zhang, Cardarelli, et al. 2011). Peruvians today exhibit varying levels of European admixture as a result of the Spanish colonization of South America (Sandoval, et al. 2013). We implemented a linear mixed model analysis to test whether European admixture proportion contributed to the level of LINE-1 DNA methylation. We used the data from the first principal component (gPC1) (adjusted for the altitude of birth, age, sex, and BMI) that separated Quechua from Europeans to test for associations with LINE-1 methylation. We saw a significant positive association between LINE-1 DNA methylation and Andean ancestry (β=-45.05, p-value=0.0004) (Fig. 2.2) suggesting that individuals with less European admixture have higher LINE-1 methylation. The association between *EPAS1* and gPC1 was not significant, but approaching significance (β=51.27, p-value=0.052).

High-altitude exposure at birth and the time of recruitment influence global DNA methylation

Average LINE-1 methylation was significantly higher in HAQ compared to both LAQ (p-value=0.0010) and MQ (p-value=0.0038). LAQ and MQ did not differ in their levels of methylation (p-value=0.69). To isolate the effect of growth and development at altitude on global methylation patterns, we combined HAQ and MQ into a single group wherein all individuals were born at high altitude (HAQ + MQ; n = 429). We used linear mixed effects models to account for multiple technical pyrosequencing replicates per individual and to avoid averaging the LINE-1 CpG sites. Based on linear mixed models adjusted for sex, age, group, PC1, and BMI, individuals born at high altitude (HAQ+MQ) showed significantly higher levels of LINE-1 methylation compared to LAQ who were born and raised at low altitude (β=0.58, p-value= 0.00027, marginal R2 = 0.014, conditional R2 = 0.726) (Table 2.2, Fig. 2.3A). Furthermore, when considering only individuals who had lived at high altitude for some or all of their lifetime (HAQ+MQ), LINE-1 methylation significantly increased with the number of years an individual lived at high altitude (β=0.033, p-value=0.0047, marginal R2 = 0.012, conditional R2 = 0.737) and when considering all participants (HAQ, MQ, and LAQ) regardless of altitude of birth (β=0.024, p-value=4.73e-05, marginal R2 =

0.015, conditional R2 = 0.726) (Table 2.2, Supp. Fig. 2.2). The relationship between the altitude of birth and DNA methylation was not significant for *EPAS1* (Table 2.2).

We tested whether the city of recruitment in this study influenced LINE-1 and *EPAS1* methylation using linear mixed models adjusted for sex, age, BMI, and gPC1. LINE-1 methylation was higher in individuals recruited in Cerro de Pasco compared to Lima (β=-0.47, p-value=0.00026, marginal R2 = 0.014, conditional R2 = 0.727). Conversely, *EPAS1* methylation was significantly lower in individuals recruited in Cerro de Pasco (HAQ) compared to individuals recruited in Lima (LAQ+MQ) (β=-0.83, p-value=0.0025, marginal R2 = 0.044, conditional R2 = 0.887, Figure 4). When only considering individuals born at high altitude (HAQ+MQ), LINE-1 methylation was higher among HAQ (β=0.38, p-value=0.017, marginal R2 = 0.011, conditional R2 = 0.737) (Table 2.2, Fig. 2.3B), but *EPAS1* methylation was lower (β=-0.91, p-value=0.005, marginal R2 = 0.049, conditional R2 = 0.879, Fig. 2.4). These data suggest that current exposure to hypoxia could influence the epigenome. However, factors other than hypoxia might be at play in determining the degree of methylation in our cohort. One potential environmental difference between the recruitment sites is inorganic lead exposure.

Cerro de Pasco is a major mining center of Peru. High levels of inorganic lead in the soil have been reported as well as high whole-blood lead levels (WBLL) in women and children living there (Conklin L., et al. 2008; van Geen, et al. 2012). We determined WBLL in a random subset of the individuals (n = 294) (148 from Lima [LAQ (69) and MQ (79)] and 157 from Cerro de Pasco [HAQ]) using atomic absorption spectroscopy. Individuals recruited in Cerro de Pasco [HAQ] had significantly higher WBLLs (4.75$\pm$1.53 μg/dL) compared to the individuals recruited in Lima [LAQ (2.05$\pm$0.65 μg/dL) and MQ (2.01$\pm$0.60 μg/dL)] (p-value<2.2e-16) (Fig. 2.5). WBLLs were not associated with LINE-1 methylation (β=0.024, p-value=0.63) (Supp. Fig. 2.3). The association between WBLLs and *EPAS1* methylation was approaching significance when adjusted for sex, age, BMI, and gPC1 (β=-0.20, p-value=0.08, Sup. Fig. 2.4). When WBLLs were tested in

participants recruited in Lima and Cerro de Pasco separately, the relationship was not significant (Lima: β=-0.56, p-value=0.23; Cerro: β=-0.18, p-value=0.35).

One-carbon metabolism associations

The folate-mediated one-carbon (1C) metabolism pathway is involved in the biosynthesis of amino acids, metabolism of nucleotides, maintenance of epigenetic modifications, and mediation of redox mechanisms (Ducker and Rabinowitz 2017). S-adenosylmethionine (SAM) is the primary methyl donor for DNA methylation. It is generated from methionine with the help of DNA methyltransferases in the 1C metabolism pathway (Anderson, et al. 2012; Bleich, et al. 2014). Previous studies have shown that polymorphisms in the 1C metabolism pathway can affect LINE-1 methylation (Wernimont, et al. 2011; Tajuddin, et al. 2013). We identified 44 SNPs in the 1C metabolism pathway that have been found to associate with changes in LINE-1 methylation and for which we had genotype data from the Affymetrix (Santa Clara, CA) Axoim Biobanking array (Hazra, et al. 2010; Gibson, et al. 2011; Wernimont, et al. 2011; Tajuddin, et al. 2013; Bleich, et al. 2014; Chang, et al. 2014; Llanos, et al. 2015). We first tested each SNP for an association with average LINE-1 methylation using an additive model of inheritance in R (R Core Group, 2018) (Supplementary Table 2.2). Models were adjusted for the altitude of birth (high vs. low), sex, BMI, and admixture measured as the first PC of the genome-wide SNP data. In the filtering step, out of the 44 SNPs tested four (*MTHFD1* rs2236225, *TYMS* rs502396, *FOLH1* rs202676, *GLDC* rs10975681), were significant at p-value<0.05 (Table 2.3). These genes are involved in the 1C pathway generating methyl groups for DNA methylation, which is why the we hypothesized that they would act in a cumulative fashion. We included the four SNPs in the same model. This four-SNP-model explained 11.29% variation in the LINE-1 methylation, while models with each SNP alone explained 9% or less of the variation in the LINE-1 methylation. *EPAS1* methylation was not associated with the one-carbon metabolism SNPs.

DISCUSSION

Genetic and epigenetic factors have been hypothesized to contribute to high-altitude adaptive phenotypes in Andeans (Frisancho 2013). However, it is not understood whether epigenetic processes play a role in their establishment. We hypothesized that epigenetic modification may underlie Andean developmental adaptation to high-altitude hypoxia. Our results suggest that high-altitude hypoxia increases global levels of DNA methylation in Peruvian Quechua measured at the LINE-1 repetitive element. Additionally, they show that high-altitude residence decreases gene-specific *EPAS1* methylation. Our study marks the first time that epigenetic modifications have been explored in the context of differential high-altitude exposure among Andean Quechua. Furthermore, our findings support the role of epigenetic processes in contributing to high-altitude adaptive phenotypes in the Andes.

We chose to assess global levels of DNA methylation using the LINE-1 element as its methylation state may reflect the overall state of methylation in the epigenome (van Bemmel, et al. 2012; Wangsri, et al. 2012; Li, et al. 2013a). We found that hypoxia exposure was associated with increased global methylation. This was supported by three observations: 1) HAQ have significantly higher levels of DNA methylation compared to MQ and LAQ 2) Individuals born at high altitude (HAQ+MQ) have significantly higher levels of DNA methylation compared to individuals born at low altitude (LAQ), and 3) the number of years an individual has lived at high altitude significantly increases DNA methylation. The differences in LINE-1 methylation we observed between altitude exposure groups were small in magnitude. However, LINE-1 elements are present at a high copy number in the genome (~17%), and thus even a small difference in DNA methylation may reflect significant global epigenomic differences (Breton, et al. 2017).

DNA methylation signatures differ between populations (Fraser, et al. 2012; Heyn, et al. 2013). We observed a significant association between LINE-1 methylation and the amount of European admixture (β=0.29, p-value=0.0011) reflected by the first PC from a PCA performed using genome-wide SNPs indicating that underlying genetic variation may significantly impact

25

global LINE-1 methylation levels, and thus must be considered in DNA methylation analyses.

We identified lower *EPAS1* promoter region methylation associated with current altitude of residence in Andean Quechua. *EPAS1* is a transcription factor involved in the body's response to oxygen levels that is active in hypoxic condition (Tian, et al. 1997; Semenza 2001b). Mutations in *EPAS1* are associated with erythrocytosis (Percy, et al. 2008), and *EPAS1* SNPs are associated with high-altitude adaptation in Tibetans (Beall, et al. 2010). Lower methylation in the *EPAS1* promoter region has been shown to correlate with higher expression of *EPAS1* (Rawluszko-Wieczorek, et al. 2014)*.* Therefore, in our cohort, this suggests that high-altitude residence is potentially associated with increased *EPAS1* expression (Rawluszko-Wieczorek, et al. 2014). An observation consistent with *EPAS1* being activated in hypoxic conditions (Semenza 2001a).

A previous study conducted in Cerro de Pasco demonstrated an increase in oxidative stress markers like plasma glutathione and lipid peroxidation products (Jefferson, et al. 2004), and increased levels of ROS have been found to associate with lower DNA methylation levels (Pogribny, et al. 2007). In our cohort, individuals from Cerro de Pasco have higher levels of LINE-1 methylation compared to the individuals from Lima.

DNA methylation and oxidative stress come together in the 1C metabolism pathway, as it stimulates synthesis of glutathione and reduces homocysteine, thus interfering with methylation (Menezo, et al. 2016). We found associations between average LINE-1 methylation and four one-carbon metabolism SNPs (*MTHFD1* rs2236225, *TYMS* rs502396, *FOLH1* rs202676, *GLDC* rs10975681). These genes are involved in various aspects of the 1C pathway, including the conversion of tetrahydrofolate (THF) into various methylenetetrahydrofolate derivatives (*MTHFD1*), conversion of deoxyuridylate to thymidylate (*TYMS*), the ability to absorb dietary folate (*FOLH1*), and the degradation of glycine (*GLDC*) (Brody, et al. 2002; Trinh, et al. 2002; Ji, et al. 2011; Guo, et al. 2013).

The effects of high-altitude adaptation on DNA methylation have been explored in other

studies of Andeans and in high-altitude Ethiopians. Previous studies of DNA methylation in high-altitude adaptation have shown that increased methylation of a different high-altitude adaptive gene, *EGLN1,* was associated with excessive erythrocytosis in Andean men suggesting that it plays a role in the development of chronic mountain sickness characterized by increased levels of erythrocytes, which is commonly seen among Andean highlanders (Julian 2017). Differences in DNA methylation have also been explored between Ethiopian highlanders and lowlanders, however, the study did not find significant methylation differences in hypoxia-associated genes (Alkorta-Aranburu, et al. 2012).

Although epigenetic processes involved in human adaptation to high altitude remain largely unknown, epigenetic aspects of hypoxia have been well studied in relation to cancer, as hypoxia is a hallmark of most tumors (Hanahan and Folkman 1996). Tumor cells grow and proliferate at a rate that outpaces their blood supply. Therefore, they must adapt to an environment with low oxygen concentrations (Vaupel and Mayer 2007; Wilson and Hay 2011), similar to humans living in the hypobaric environment of high altitude. In cancerous cells, hypoxia inducible factors (HIFs) have been shown to work in concert with epigenetic mechanisms to respond to hypoxia (Watson, et al. 2010). For example, in renal carcinomas and multiple myelomas the HIF pathway gene *von-hippel lindau* (*VHL)* is silenced via hypermethylation of its 5' CpG island (Herman, et al. 1994; Hatzimichael, et al. 2009). Additionally, transcription of a known altitude-adaptive candidate gene, *Endothelial PAS Domain Protein 1* (*EPAS1),* is mediated by DNA methylation in colorectal cancer (Rawluszko-Wieczorek, et al. 2014). These findings highlight the ability of DNA methylation to play a role in mediating the cellular response to hypoxia. Furthermore, they support the notion that human populations may adapt to high-altitude hypoxia through epigenetic mechanisms.

In our cohort, the city of recruitment (Lima vs. Cerro de Pasco) had a significant effect on DNA methylation. We hypothesized that this could be a result of environmental exposure to lead given that Cerro de Pasco is a mining town and lead exposure has been shown to decrease LINE-

1 methylation in previous studies (Wright, et al. 2010b; Li, et al. 2013b; Goodrich, Dolinoy, Sánchez, et al. 2016). We did not find strong associations between DNA methylation and lead exposure in our study. However, WBLLs are representative of recent lead exposure, and it is unknown what the lifetime exposure levels were for all participants. Another potential difference between the two recruitment sites is diet, which is known to affect DNA methylation (Zhang, Morabia, et al. 2011). Lima is an urban center and Cerro de Pasco is a remote, Andean town. Any potential differences in diet between participants recruited in Lima and Cerro de Pasco may contribute to the differences in DNA methylation that we observed.

DNA methylation signatures differ by tissue and cell type (Kitamura, et al. 2007). We used DNA extracted from whole blood, which is comprised of a mixture of DNA-containing cells. It is possible that these DNA-containing blood cell types are different between high- and low-altitude individuals. For example, high-altitude Andeans have elevated hemoglobin levels (Beall, et al. 1998) compared to low-altitude individuals suggesting a higher number of red blood cells associated with altitude. We tested the association between hemoglobin levels and LINE-1 and *EPAS1*, which was not significant (data not shown), suggesting that differences in blood cell types are potentially not affecting our results. However, hemoglobin is in red-blood cells, so it is possible that DNA containing white-blood cells show variation between low and high altitude.

The environment affects DNA methylation signatures, and exposures to toxicants, pharmaceutical agents, exercise, stress, and other factors have been shown to associate with changes in epigenetic marks, such as DNA methylation (Csoka and Szyf 2009; Senut, et al. 2012; Ling and Ronn 2014; Non, et al. 2016). Our findings suggest that being born at high altitude leaves a persistent mark on the epigenome through adulthood by leaving lasting effects on global levels of DNA methylation measured at LINE-1. We also show that altitude of residence affects *EPAS1* potentially activating it. Based on our findings, genome-wide DNA methylation studies of high-altitude adaptation are a natural next step in determining the individual genes and pathways that may undergo changes in DNA methylation in association with adaptation to high altitude.

MATERIALS AND METHODS

Study Population

Study participants consisted of Peruvian Quechua individuals recruited based on three high-altitude/developmental exposure patterns (Table 2.1): 1) high-altitude Quechua (HAQ) *(n=300):* Quechua individuals born and raised at high altitude who were recruited from the city of Cerro de Pasco, Peru located at an elevation of 4,338 m; 2) *migrant Quechua (MQ) (n=150)*: Quechua individuals who were born in various high-altitude locations (>2,500 m) and moved to low-altitude Lima at some point in their lifetime; 3) low-altitude Quechua (LAQ) (n=152)*:* Quechua individuals born and raised at low altitude who were recruited in Lima, Peru at an elevation of 0 m. Male and female participants were unrelated, healthy, non-pregnant/lactating, non-smokers, between 18 and 35 years old (Table 2.1). Relatedness was determined during the recruitment process using a pre-screening questionnaire. Genetic relatedness was estimated with the program King (Manichaikul, et al. 2010) using genome-wide SNP data from the Affymetrix Axiom Biobanking array. All individuals were in good health at the time of enrollment with no history of chronic disease or recent illness. All study participants were screened for anemia using altitude specific cut-offs. None of the study participants recruited in Cerro de Pasco were mine workers even though Cerro de Pasco is a mining town.

Height and weight were collected at the time of enrollment, and BMI was calculated using the kg/m^2 equation. Participants provided a blood sample drawn from an antecubital vein into a vacutainer collection tube containing EDTA for DNA extraction. Whole blood was field stabilized in lysis buffer and transported to the University of Michigan. Genomic DNA was extracted using the Puregene DNA purification system (Qiagen, Valencia, CA) according to the manufacturer's instructions.

DNA methylation

Quantitative pyrosequencing was performed to assess DNA methylation levels at four

CpG sites of the LINE-1 repetitive element and *EPAS1*. One μg of DNA from each sample was bisulfite converted using the EZ-96 DNA Methylation™ Kit (Zymo Research, Irvine, CA). Bisulfite converted DNA was amplified in duplicate using *EPAS1* and LINE-1 primers and HotstarTaq plus Master Mix (Qiagen, Valencia, CA). LINE-1 forward primer sequence: 5'-TTGAGTTAGGTGTGGGATATAGTT-3', LINE-1 reverse primer sequence: 5'-[biotin]-CAAAAAATCAAAAAATTCCCTTTCC-3', LINE-1 sequencing primer: 5'-AGGTGTGGATATAGT-3'. The following *EPAS1* region was targeted chr2:46,526,762-46,527,105 (hg19). *EPAS1* forward primer sequence: 5'-TGGGAGTAGGGGAAAAATGAT-3', *EPAS1* reverse primer sequence: 5'-[biotin]- ACAATCCCCACAATAATTCTTAA-3', *EPAS1* sequencing primer: 5'-AGTAGGGGAAAAATGATT-3'.

DNA methylation levels at each CpG site for *EPAS1* and LINE-1 were assessed using the Pyromark Q96 pyrosequencer (Qiagen, Valencia, CA), for more detail see (Virani, et al. 2012). A bisulfite conversion check ensured full bisulfite conversion of the DNA. The assay was validated in duplicate using a DNA methylation scale of DNA with known values of methylation (0%, 20%, 40%, 60%, and >80%). The DNA methylation scale was created from whole genome amplified DNA (representing 0% DNA methylation), and DNA treated with CpG methyltransferase *M.SssI* (New England Biolabs, Ipswich, MA) (representing >80% DNA methylation). PCR amplification followed by pyrosequencing was performed on each bisulfite treated sample in duplicate for LINE-1. For *EPAS1*, 10% of the samples were duplicated. The data were checked via built in quality control measures, and samples with coefficients of variance (CV), which was calculated as standard deviation for all samples over the average for all samples multiplied by 100, exceeding 10 were excluded from the further analyses. Of the 603 samples, 64 either failed pyrosequencing or had high CV for *EPAS1* and 31 for LINE-1 and were excluded from the statistical analyses. Statistical analysis was performed on 572 samples (282 HAQ, 147 MQ, and 143 LAQ).

Microarray Genotyping

Microarray genotype data were generated for all Andean Quechua DNA samples using the Affymetrix (Santa Clara, CA) Axoim Biobanking array. The Axiom array features approximately 610,000 markers including haplotype tagging SNPs as well as exonic and eQTL variants. We performed a PCA on autosomal variants with call rates > 95%, major allele frequency > 0.0001 that passed the QC of Affymetrix (n=373,260).

To identify candidate genes from the one-carbon metabolism pathway we performed a literature search for publications reporting one-carbon metabolism SNPs in association with DNA methylation, cancer, and diet among others (Hazra, et al. 2010; Aneiros-Guerrero, et al. 2011; Gibson, et al. 2011; Wernimont, et al. 2011; Tajuddin, et al. 2013; Bleich, et al. 2014; Chang, et al. 2014; Llanos, et al. 2015; Lucock, et al. 2015; Kim, et al. 2016). Based on the previously published one-carbon metabolism SNP associations, we selected 81 SNPs that were present on the Affymetrix (Santa Clara, CA) Axoim Biobanking array. After removing SNPs that were monomorphic in our samples, were in high LD, and had MAF less than 0.05, we ended up with 44 SNPs used in the analysis. The following model was used to determine the combined effect of 1C SNPs on LINE-1 methylation. Yi = B00 + B01(rs2236225) + B02(rs202676) + B03(rs502396) + B04(rs10975681) + Altitude + Sex + Age + PC1 + BMI + E. where Yi = average LINE-1 methylation.

Statistical analyses

All statistical analyses were conducted using R version 3.4.0 (R Core Team, 2018). Packages lme4 (Bates, et al. 2015), lmerTest (Kuznetsova, et al. 2017), ggplot2 (Wickham 2009), and MuMIn (Barton, 2009) were employed. We used both linear mixed effects models and simple linear models to be able to show results while accounting for multiple technical pyrosequencing replicates per individual and to also test the effects on average LINE-1 or *EPAS1* methylation. The following model was tested for linear mixed models:

Yi (% methylation) ~ B00 + B01(X) + B03(Sex) + B04(Age) + (1|Pos) + (1|ID) + ei,

Where X = altitude, city of recruitment, years lived at high altitude, One-carbon metabolism SNPs, Pos = Individual CpG site position, and ID = sample ID. Position and ID were coded as random effects.

Author Contributions: A.W.B., A.C., T.D.B., F.L.V., M.R., M.K., D.C.D., and J.M.G., conceived and designed the research; A.C. and T.R.J performed experiments; A.W.B., A.C., J.M.G. analyzed data; A.W.B., A.C., and wrote the manuscript with contributions from all authors.

TABLES

Table 2.1 Participant Characteristics

Variable	All (n=572)	Low-altitude Quechua (n=143)	Migrant Quechua (n=147)	High-altitude Quechua (n=282)
Average LINE-1 methylation	69.99 ± 1.45	69.74 ± 1.18	69.82 ± 1.28	70.20 ± 1.63
Average EPAS1 methylation	14.42 ± 3.08	14.65 ± 3.06	15.06 ± 2.89	14.00 ± 3.12
% Female	49.65 (45.55, 53.75)	48.95 (40.76, 57.14)	53.06 (44.99, 61.13)	48.23 (42.40,54.06)
Age, yr*†	24.67 ± 5.03	24.38 ± 4.55	25.52 ± 5.49	24.37 ± 4.98
Age migration, yr	NA	NA	15.21 ± 6.37	NA
Height (cm)*†	158.91 ± 8.35	161.40 ± 8.72	158.01 ± 7.84	158.11 ± 8.19
Weight (kg)*†	61.95 ± 9.83	66.36 ± 11.69	61.76 ± 9.57	59.82 ± 8.07
BMI**	24.52 ± 3.28	25.43 ± 3.79	24.69 ± 3.12	23.96 ± 2.96

Data are means ± SD, 95% confidence interval (CI) for proportions in brackets; NA, not applicable
* $p < 0.05$ for born low vs. high-to-low based on t-test
† $p < 0.05$ for born high vs high-to-low based on t-test
** $p < 0.05$ based on ANOVA

Table 2.2 Significant linear mixed effects models

	Beta	p-value	R2 marginal	R2 conditional
LINE1 Entire cohort (n = 572)				
Altitude of Birth*	0.598 (High)	0.0002	0.014	0.729
City of Recruitment**	0.474 (Cerro)	0.0002	0.014	0.729
Years at high altitude**	0.025 (years HA)	3.51E-05	0.015	0.729
LINE1 Individuals born at high altitude (n= 429)				
City of Recruitment**	0.387 (Cerro)	0.0170	0.011	0.740
Years at high altitude**	0.033 (years HA)	0.0047	0.012	0.740
LINE1 Quechua Low-Altitude Migrants MQ (n=146)				
Years at high altitude**	0.029 (years HA)	0.071	0.012	0.729
EPAS1 Entire cohort (n = 539)				
Altitude of Birth ***	-0.392 (High)	0.2119	0.036	0.887
City of Recruitment ***	-0.828 (Cerro)	0.0025	0.044	0.887
Years at high altitude ***	-0.029 (years HA)	0.0193	0.040	0.887
EPAS1 Individuals born at high altitude (n= 380)				
City of Recruitment***	-0.912 (Cerro)	0.0054	0.049	0.879
Years at high altitude***	-0.059 (years HA)	0.0092	0.047	0.879
EPAS1 Quechua Low-Altitude Migrants MQ (n=123)				
Years at high altitude***	-0.049 (years HA)	0.1895	0.021	0.862

Altitude refers to the altitude of birth (High vs. low)
City refers to the city of recruitment (Cerro de Pasco vs Lima)
Years HA refers to the number of years one lived at high altitude
LA, Low altitude
HA, high altitude
REML, restricted maximum likelihood
Marginal R2 represents the variance explained by the fixed factors, and conditional R2 represents the variance explained by both fixed and random factors
* Model was adjusted for sex, age, group, BMI, and PC1
** Models were adjusted for sex, age, BMI, and PC1
*** Models were adjusted for sex, age, and PC1

Table 2.3. Significant one-carbon metabolism SNPs

Gene	SNP	Model	MAF	Minor allele	Function	β (95% CI)	P-value
FOLH1	rs202676	Additive	0.39	G	Loss of stop	0.21 (0.04, 0.38)	0.018
GLDC	rs10975681	Additive	0.30	C	Intronic	-0.17 (-0.33, -0.01)	0.042
MTHFD1	rs2236225	Additive	0.34	A	Missense	0.22 (0.05, 0.40)	0.014
TYMS	rs502396	Additive	0.40	T	Intronic	0.23 (0.06, 0.40)	0.008

All models were adjusted for the altitude of birth, sex, age, BMI, and PC1

FIGURES

A.

B.
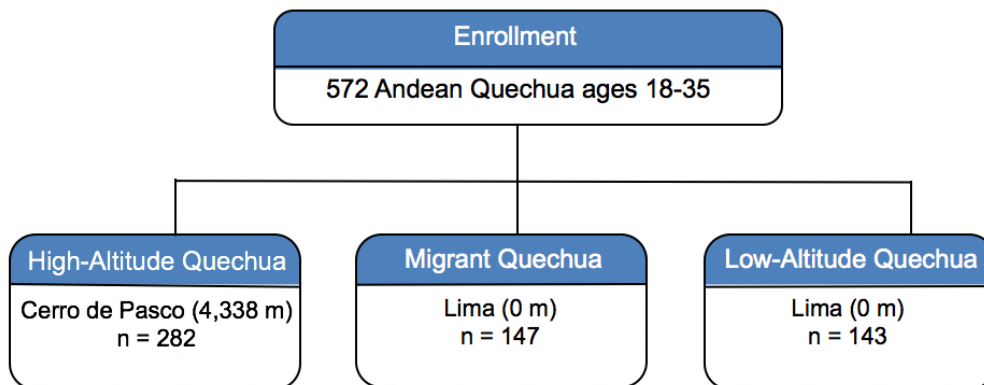
Figure 2.1. Study Design. (A) Map of Peru depicting high-altitude, Cerro de Pasco (4,338 m), and low-altitude, Lima (0m), participant recruitment locations; (B) *LINE1* DNA methylation was determined in 572 Peruvian Quechua with different exposures to high-altitude hypoxia over their lifetime. High-altitude Quechua (HAQ) were born, raised, and reside in or near Cerro de Pasco, Peru. Migrant Quechua (MQ) were born and raised at high altitude (>3,000 m), but moved to low altitude Lima (0m) during the lifetime. Low-altitude Quechua (LAQ) were born, raised, and reside at sea-level in Lima.
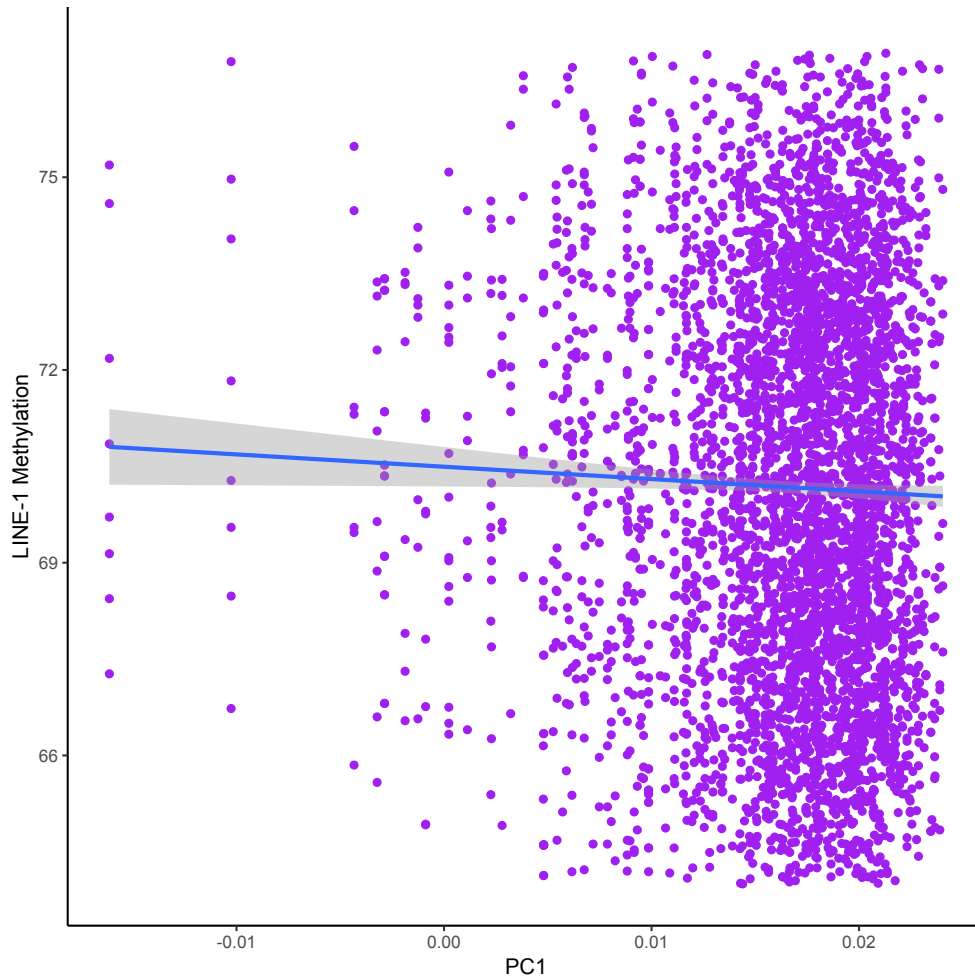
Figure 2.2. Association between LINE-1 methylation and the first principle component (PC1) from the genome-wide PCA analysis (n=572). PCA was performed on genome-wide SNP data generated using the Affymetrix Axiom Biobanking Array. All study participants are shown as filled circles. The blue line represents X. X is shown by the grey shading. Lower levels of European admixture are associated with higher levels of LINE-1 methylation (p-value= 0.0004). Individuals with lower levels of European admixture have higher values for PC1.
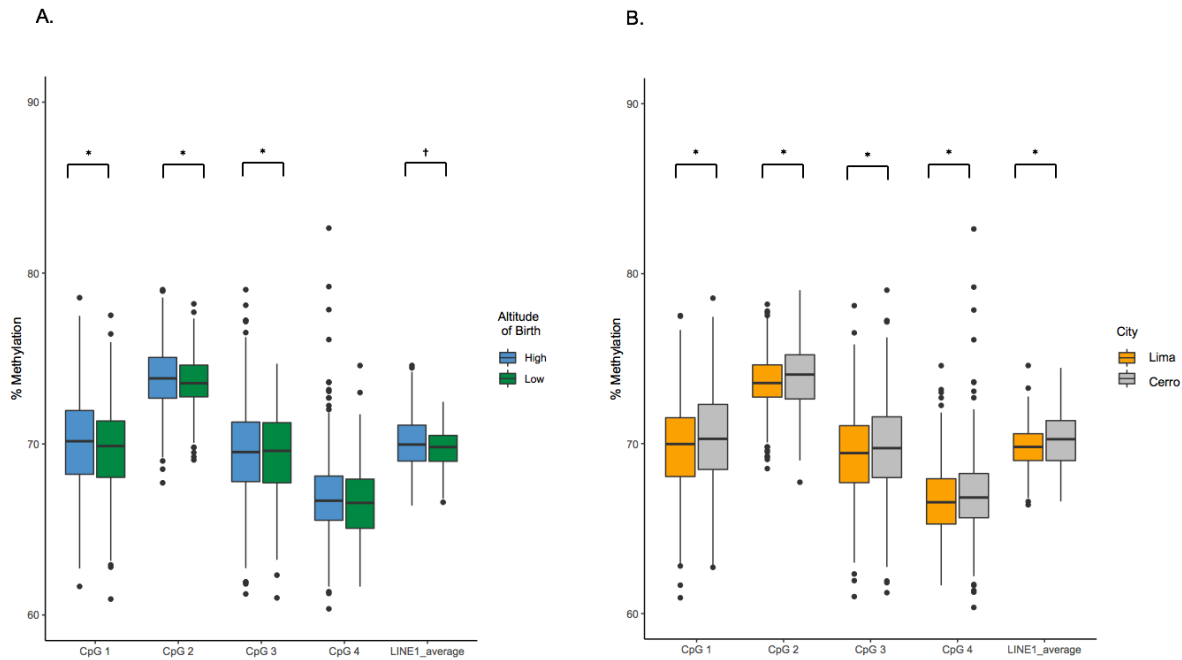
Figure 2.3. (A) LINE-1 methylation level comparison by CpG position between individuals born at high altitude (HAQ+MQ) and individuals born at low altitude (LAQ). LINE-1 methylation is significantly higher at $p<0.05$ among born at high-altitude participants for CpG positions 1 and 2 and the average CpG, and at $p<0.10$ for position 4. DNA methylation is not significantly different at position 3. Linear mixed models were corrected for sex, age, recruitment group (i.e. HAQ, MQ, or LAQ), PC1, and BMI. * significant at $p<0.05$, † significant at $p<0.10$ (B) LINE-1 methylation level comparison between individuals recruited in Lima, Peru (LAQ+MQ) and individuals recruited in Cerro de Pasco, Peru (HAQ). LINE-1 methylation is significantly higher in participants recruited in Cerro de Pasco for all 4 LINE-1 CpG positions and the average CpG. Linear mixed models were corrected for sex, age, recruitment group (i.e. HAQ, MQ, or LAQ), PC1, and BMI. * Significant at $p<0.05$.

Average CpG was calculated over 4 LINE-1 CpG sites including the technical replicates.
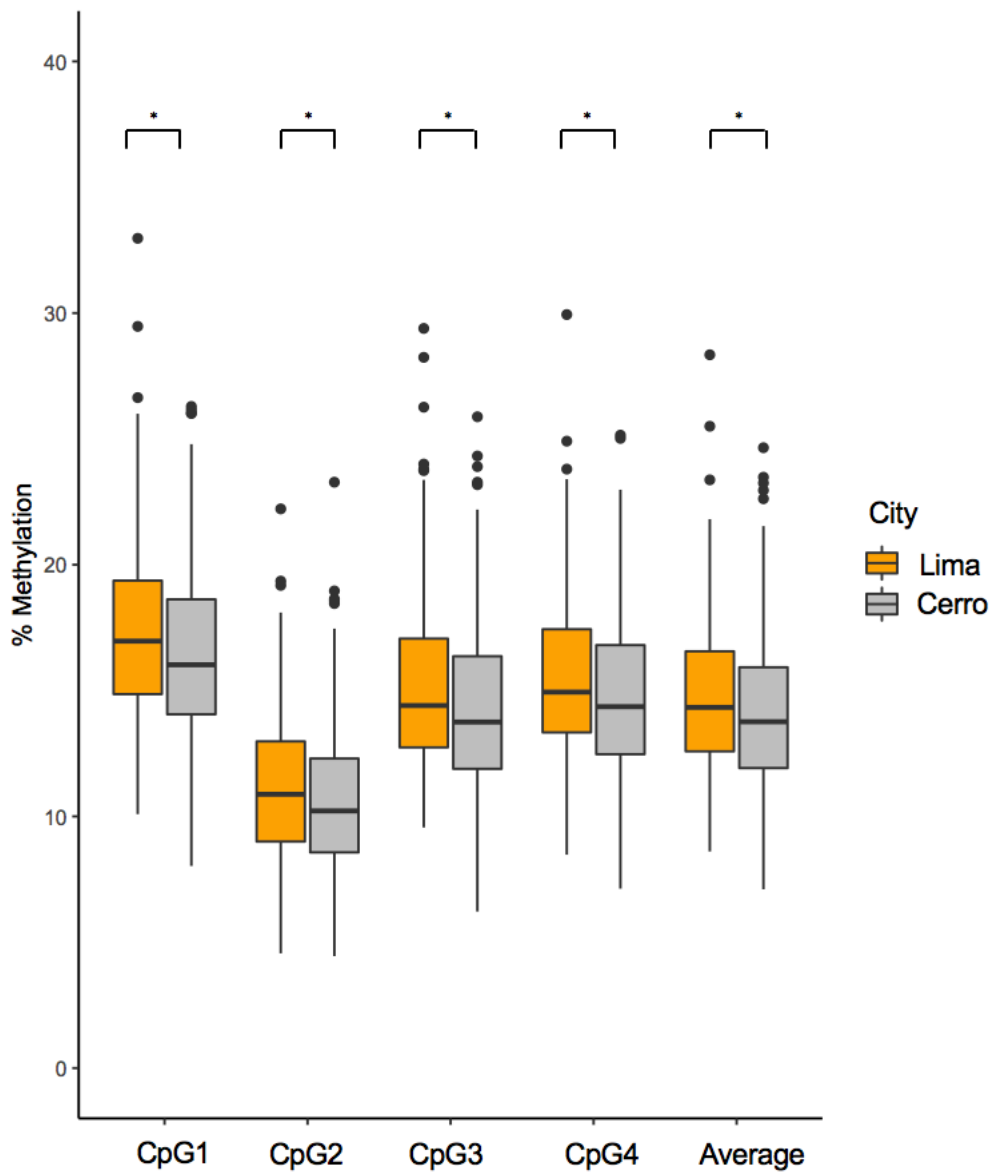
Figure 2.4. EPAS1 methylation level comparison between individuals recruited in Lima, Peru (LAQ+MQ) and individuals recruited in Cerro de Pasco, Peru (HAQ). *EPAS1* methylation is significantly higher in participants recruited in Lima for all 4 EPAS1 CpG positions and the average methylation. * Significant at p<0.05.
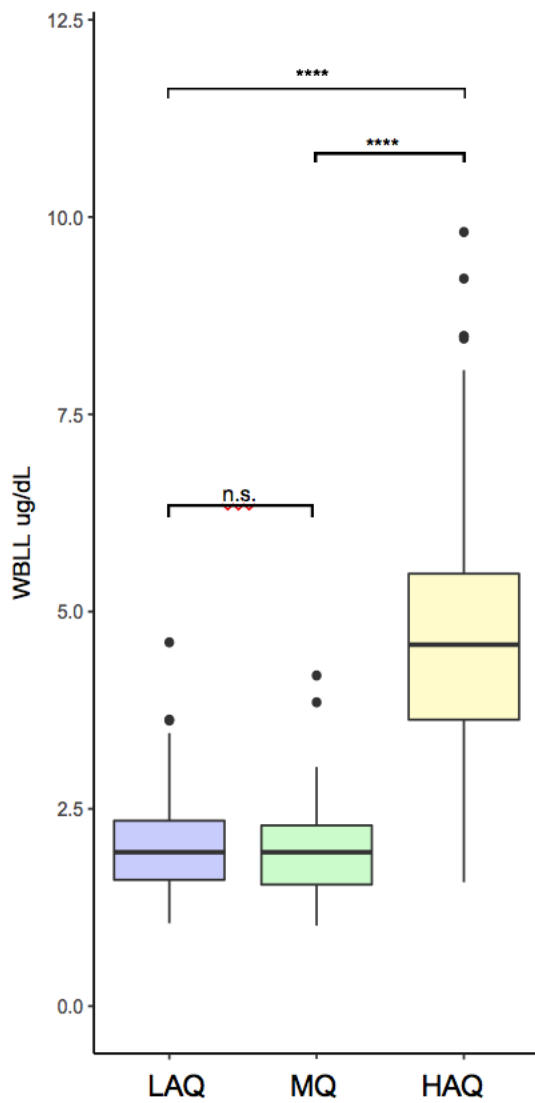
Figure 2.5. Whole-blood lead level WBLL (μg/dL) comparison between the participants from the three different groups: LAQ, MQ, and HAQ. WBLLs are significantly lower in LAQ (p-value=3.56E-45) and MQ (p-value=3.33E-48) compared to HAQ. WBLLs do not differ significantly between LAQ and MQ (p-value=0.76).

CHAPTER 3

Genome-Wide Epigenetic Signatures of High-Altitude Adaptation in the Andes

Introduction

High altitude (HA) environments are defined as regions of the earth lying above 2,500 meters (m) sea level. They are characterized by decreased oxygen concentration, high levels of UV radiation, low ambient temperatures, and high caloric demands, making them a challenging environment for human life (Baker and Little 1976). Despite this, more than 140 million people worldwide permanently live at high altitude and 40 million more visit high altitudes annually (Ward, et al. 2000; Moore 2001). High-altitude sojourners often suffer from conditions like acute mountain sickness (AMS), high-altitude pulmonary (HAPE) or cerebral edema (HACE) (Hackett and Roach, 2001), while high-altitude natives including Andeans, Tibetans, and Ethiopians display unique evolutionary adaptations allowing them to overcome hypoxic stress (Moore 2001; Beall 2006). In recent years, population-specific genetic adaptations have been discovered in high-altitude natives that contribute to their unique physiological responses to chronic hypoxia (Beall, et al. 2010; Bigham, et al. 2010; Simonson, et al. 2010). However, altitude-adaptive phenotypes likely are not determined solely by genetic sequence variation. Rather, genetic variation works in concert with epigenetic modifications, which in turn takes cues from environmental stimuli to fine-tune gene expression and ultimately impact phenotypic expression. Therefore, epigenetic modifications, due their plastic nature and ability to affect gene expression, have been proposed to contribute the process of adaptation, including adaptation to high altitude (Frisancho 2009; Szyf and Bick 2013; Maccari, et al. 2014; Julian 2017)

41

Epigenetics is the study of mitotically and meiotically heritable changes to gen expression that do not involve changes to the DNA sequence (Wolffe and Guschin, 2000; Feil and Fraga, 2012). Epigenetic marks include histone modification, DNA methylation, and non-coding RNAs, among others. Of these, DNA methylation is the most widely studied and best understood (Mohn and Schubeler 2009; Lam, et al. 2012). DNA methylation most often occurs at cytosines within CpG dinucleotides wherein a methyl group is added to the 5th carbon position. Clusters of CpG sites exist known as CpG islands. They are most often associated with promoter regions where CpG hypomethylation promotes gene expression through transcription factor binding and CpG hypermethylation results in a transcriptionally inactive state (Mohn and Schubeler 2009; Lam, et al. 2012).

DNA methylation undergoes reprogramming during early development making this period a critical window of susceptibility to environmental exposures (Morgan, et al. 2005). Exposures during this time can have lasting effects on the epigenome and potentially health and disease throughout the lifetime (Bollati and Baccarelli 2010; Devlin, et al. 2010; Goodrich, Reddy, et al. 2016). Therefore, it follows that early hypoxia exposure could lead to epigenetic changes that contribute to high-altitude adaptation. Indeed, growth and development at high altitude have been shown to contribute to altitude-adaptive traits among Andeans (Frisancho, et al. 1973; Frisancho, et al. 1997; Brutsaert 2001). For example, Peruvian Quechua born and raised at high altitude maintain higher arterial oxygen saturation (SaO_2) during exercise than Peruvian Quechua born and raised at sea-level (Kiyamu, et al. 2015). Similarly, high-altitude Andeans exhibit greater lung volumes than sea-level residents (Frisancho, et al. 1997). Despite the potential importance of epigenetic regulation in determining high-altitude adaptive phenotypes, remarkably few epigenetic studies have been published to date on this topic.

Previous research in the same population of Andean Quechua considered here has shown that DNA methylation measured at LINE-1, a marker of global DNA methylation, was higher in individuals born at high altitude and that current altitude of residence affected the methylation

42

status of the oxygen-sensing gene, *EPAS1* (Childebayeva, et al. 2019). These findings suggest that high-altitude exposure affects the epigenome. However, it is unknown whether Andeans display genome-wide DNA methylation changes and if these changes contribute to high-altitude adaptation or buffering against high-altitude associated health conditions.

Despite being generally adapted to high altitude, Andeans also display certain signs of incomplete adaptation to this environment, such as pulmonary hypertension (PH) and chronic mountain sickness (CMS), the latter a more severe condition (Monge 1942; Peñaloza, et al. 1963). PH is characterized by a mean pulmonary arterial pressure of ≥25 mm Hg and associated with vasoconstriction, structural changes in the pulmonary arteries, and increased arterial blood pressure occur in PH (Peñaloza, et al. 1963; Antezana, et al. 1998; Beall 2007). High-altitude Andeans have increased muscularity in their small lung arteries, allowing them to have elevated pulmonary resistance to blood flow and mild pulmonary hypertension (PH) as a result (Peñaloza, et al. 1963; Heath and Williams 1991). PH plays an adaptive role in Andeans, allowing for greater oxygenation in hypoxic conditions, and goes away in normoxic conditions (Reeves and Grover 2005). With age, some high-altitude Andeans lose their capacity for adaptation and develop CMS, which is characterized by excessive red blood cell and hemoglobin production, and severe pulmonary hypertension (Leon-Velarde, et al. 2003). The incidence of CMS varies based on altitude, affecting approximately 20% of the Peruvian Quechua male population (Leon-Velarde, et al. 2005). CMS is a significant cause of morbidity and mortality in the Andean population, and thus an important public health issue.

Epigenetic modifications have been shown to play a role in the development of PH at low altitude (Xu, Cheng, et al. 2011). However, it is unclear if epigenetic modifications play a protective role against the development of moderate or severe PH at altitude, especially in Andeans who are known to have genetic adaptations to this environment (Kim, et al. 2011; Stenmark, et al. 2012).

To understand the effect of high-altitude exposure at birth and early development on the epigenome, we performed an epigenome-wide association study (EWAS) on 87 individuals of Peruvian Quechua ancestry who were born and raised in hypoxia or normoxia. We identified differentially methylated regions (DMRs) associated with birth and early development at high-altitude, including genes involved in hypoxic regulation and pulmonary hypertension. We also identified one significant differentially methylated position (DMP) in the promoter region of *NPY1R* that may potentially be protective against PH. Lastly, we found that individuals born at high-altitude age faster compared to the individuals born at low altitude suggesting that hypoxia influences cellular machinery at large. Accelerated epigenetic aging was positively associated with hemoglobin levels suggesting a link between epigenetic age and CMS. Overall, our findings suggest that epigenetics is involved in high-altitude adaptation and the protection against chronic hypoxia-associated conditions like PH.

Results

Study design and participant characteristics

We generated DNA methylation data for ~850,000 CpG sites using the Illumina Infinium® MethylationEPIC BeadChip assay for 90 Quechua individuals between the ages of 18 and 35. These 90 DNA samples were randomly selected from a larger cohort of Peruvian Quechua participants (Childebayeva, et al. 2019) that were recruited using a migrant study design. Our analysis included individuals from three high-altitude exposure groups: 1) high-altitude Quechua (HAQ) (n=28) were multigenerational highland individuals who were born, raised, and reside at high altitude and were recruited in the city of Cerro de Pasco, Peru (4,300 m); 2) migrant Quechua (MQ) (n=28) were individuals who were born in highland Peru (> 2,500 m) and moved to low-altitude Lima, Peru (150 m) during their lifetime; and 3) low-altitude Quechua (LAQ) (n=31) individuals were born and raised in Lima, Peru, but their parents and both sets of grandparents were born and raised in highland Peru (Fig. 3.1 A and B, the study design is described in greater

detail in Childebayeva et al., 2019). This migrant study design allowed us to separate the effects of developmental exposure to high altitude from underlying shared genetic ancestry. At the time of enrollment, all study participants provided written informed consent approved by Syracuse University, Universidad Peruana Cayetano Heredia, and The University of Michigan Institutional Review Boards.

Participant characteristics are provided in Table 3.1. We excluded three HAQ participants whose DNA samples failed initial quality control (QC). All statistical analyses were performed on 87 samples with reliable DNA methylation data. Height, weight, and hemoglobin levels were collected at the time of recruitment. Hemoglobin levels were measured using a HemoCue Hb201+ analyzer (Angelholm Sweden). The three study groups were significantly different by height and weight ($p<0.05$), but not BMI. The study groups did not differ significantly by sex or age. Average age of down-migration for MQ was 14.30 ± 6.62 years of age. To explore the effects of hypoxia exposure in early life, we combined HAQ and MQ into a single born at high altitude group (born high-altitude Quechua [BHAQ]). We compared DNA methylation data from BHAQ to LAQ who were all born and raised at sea level. Cerro de Pasco is a major mining center of Peru with high levels of inorganic lead reported in the soil as well as high whole-blood lead levels (WBLL) among women and children living there (Conklin L., et al. 2008; van Geen, et al. 2012). Given that lead exposure has been shown to affect DNA methylation (Wright, et al. 2010a; Senut, et al. 2012; Senut, et al. 2014), we measured participant's WBLLs using graphite furnace atomic absorption spectrometry. Individuals recruited in Cerro de Pasco [HAQ] had significantly higher WBLLs (4.91 ± 1.72 μg/dL) compared to the individuals recruited in Lima [LAQ (2.11 ± 0.67 μg/dL) and MQ (1.99 ± 0.61 μg/dL)] (β=3.05, p-value<2e-16) (Table 3.1). All statistical models were corrected for WBLLs to account for the potential effects of lead exposure on the epigenome.

Differentially methylated positions

We identified significantly different DNA methylation positions (DMPs) between individuals born at high (n = 56) versus low altitude (n = 31) using linear modeling in limma (Fig. 3.2) (Ritchie, et al. 2015). We identified one significant DMP at q-value < 0.10 and genome-wide inflation factor λ=1.103 (Table 3.2, Sup Fig 3.1) located in the promoter region of the gene *NPY1R* (Sup. Fig. 3.2). *NPY1R* or Neuropeptide Y Receptor Y1 is involved in nerve cell activity and blood pressure control (Wang, et al. 2009) and has been shown to have increased expression in patients with PH (Crnkovic, et al. 2014).

Differentially methylated regions

We determined differentially methylated regions (DMRs) associated with birth at high altitude. DMRs, defined as genomic regions with three or more CpG sites significant at unadjusted p-value<0.05 within 3,000 base pairs (bp) of each other, reflect several nearby CpG sites changing in a similar fashion. Therefore, they may be more representative of actual biologically relevant epigenetic changes than single methylation sites. Again, individuals born at high (BHAQ = 56) versus low altitude (LAQ = 31) were compared. We identified 155 DMRs significant after Sidak's multiple testing correction (p<0.05). Of these 155, five were hypomethylated in BHAQ individuals, and 150 were hypermethylated in BHAQ (Supp. Table 3.1). Twenty out of the 155 significant DMRs were either under HIF regulation or known to be differentially expressed in hypoxic conditions (Table 3.3). Each of these DMRs were hypermethyalted among BHAQ compared to LAQ. Moreover, ten of the significant DMRs were associated with PH (*RYR3, HTR1B, AVPR1A, SLIT3, KCNA5, ALDH1A2,* and *SOD3)* and three were associated with genes potentially involved in high-altitude adaptation (*PKLR*, *USP29*, and *YAP1)*, including one that has been shown to be under selection for high-altitude adaptation (Yi, et al. 2010).

We chose to validate the differential methylation findings associated with the HIF pathway gene *SOD3* (Figure 3) given its role in protecting cells against oxidative stress (Suliman, et al. 2004). To do so, we performed quantitative pyrosequencing using a Pyromark Q96

pyrosequencer (Qiagen, Valencia, CA) in the larger cohort of Peruvian Quechua (n = 498 where HAQ=240, MQ=124, LAQ=134). These findings replicated the results of our EWAS analysis, demonstrating that *SOD3* methylation was significantly increased among Quechua born at high altitude (BHAQ = 364) compared to those born at low altitude (LAQ=134) (β=1.83, p-value<0.01, Figure 4).

Accelerated Epigenetic Aging at High Altitude

Epigenetic aging has been hypothesized to reflect the overall state of the epigenome (Horvath 2013; Boks, et al. 2015; Perna, et al. 2016). Accelerated epigenetic aging has been associated with cardiovascular mortality, obesity, and lifetime stress, and thus can serve as a marker of epigenome "health" (Horvath, et al. 2014; Marioni, et al. 2015; Zannas, et al. 2015). In previous research, we demonstrated differences in LINE-1 DNA methylation in Andeans born at high versus low altitude suggesting that high-altitude is affecting the epigenome globally (Childebayeva, Jones et al. 2019). To further establish how the altitude of birth is affecting the epigenome overall, we estimated epigenetic age using the method developed by S. Horvath (Horvath 2013). This method relies on 353 CpG sites that previously have been shown to be associated with age. We used linear models to determine if DNA methylation age difference (ΔDNAmAge, where ΔDNAmAge = Horvath Epigenetic Age – Chronological Age) is associated with the altitude of birth. ΔDNAmAge was significantly higher among individuals born at high altitude (β=1.96, p-value = 0.036, Fig. 3.5) when adjusting for chronological age, sex, group (HAQ, LAQ, MQ), and the first two blood PC).

Discussion

Andean high-altitude adaptation is hypothesized to involve both genetic and developmental factors. However, to date, little attention has been focused on understanding the epigenomic

mechanisms that may be responsible for developmental adaptation. We focused on the epigenome to understand the role of developmental adaptation to high-altitude adaptation. We determined genome-wide DNA methylation changes associated with birth at high altitude among Peruvian Quechua, and performed an EWAS exploring the relationship between altitude of birth and DNA methylation status. We identified one significant DMP in the gene *NPY1R* and 155 DMRs associated with altitude of birth. Ours is the first study to show genome-wide DNA methylation changes associated with early life at high altitude. These findings further support the notion that the environment plays a role in adaptation and highlight epigenetics as a potential mediator of environmental effects on the body.

Three of the 155 significant DMRs were associated with genes indicated in high-altitude adaptation including *PKLR*, *USP29,* and *YAP1*. *PKLR,* encoding the R/L-type pyruvate kinase, is involved in red blood cell production and maintenance, has been associated with hereditary hemolytic anemia, and is upregulated in the rat fetus in hypoxic conditions (Kanno, et al. 1992; Huang, et al. 2004). Importantly, *PKLR* is under selection in Tibetans and high-altitude Mongolians (Yi, et al. 2010; Xing, et al. 2013; Wuren, et al. 2014). Ubiquitin specific protease 29 (USP29) stabilizes hypoxia-inducible factor 1 - alpha (HIF-1α), a hypoxia-inducible factor 1 subunit that is involved in regulating the molecular response to hypoxia, by de-ubiquitinating it and thus preventing it from degradation (Schober 2016). Previous studies found increased *USP29* DNA methylation in hypoxic conditions in prostate cells and altered *ESP29* expression in mouse brain upon chronic hypoxic exposure (Zhou, et al. 2008; Watson, et al. 2009). YAP1 is a transcription factor involved in oxygen stress resistance (Liu and Barrientos 2013). Its expression is increased by HIF-2α also known as EPAS1, a hypoxia-inducible factor that's under selection in Tibetans and Andeans (Beall, et al. 2010; Ma, et al. 2017).

Healthy high-altitude natives have altered heart and pulmonary circulation characterized by pulmonary hypertension (PH), right ventricular hypertrophy (RVH) and increased number of smooth muscle cells (SMCs) (Reeves and Grover 2005; Penaloza and Arias-Stella 2007). The

degree of PH at high altitude is affected by the number of generations a population has been exposed to this environment. Interestingly, Tibetans have decreased levels of PH, while the European residents of Leadville, Colorado have high levels of PH compared to high-altitude Andeans (Penaloza and Arias-Stella 2007). In Andeans, PH is considered adaptive, and subsides after going down to sea level (Reeves and Grover 2005). We found one differentially methylated position and four significant DMRs that may be involved in epigenetic buffering against PH. The significant differentially methylated position (cg17500103) is in the promoter region of Neuropeptide Y Receptor Y1 (*NPY1R*), a transmembrane protein that mediates the function of the neurotransmitter neuropeptide Y (NPY). *NPY1R* is a potent vasoconstrictor (Zukowska-Grojec, et al. 1987). Its expression is increased in lungs of mice exposed to chronic hypoxia and in lung samples of patients with idiopathic pulmonary arterial hypertension suggesting a role in pathogenesis of PH (Crnkovic, et al. 2014). We have shown increased methylation of cg17500103. Elevated DNA methylation in promoter regions is often associated with gene repression, and a previous study of cancer epigenetics has shown a negative correlation between *NPY1R* expression and cg17500103 methylation (corr.coeff =-2.70E-01; p-value=1.11E-04; q-value=2.21E-05) (TCGA Research Network, 2012). Given that overexpression of *NPY1R* is associated with PH, hypermethylation of its promoter may constitute an adaptive response in the individuals born at high altitude that reduces the occurance of PH at altitude.

We identified nine DMRs associated with genes related to PH, four of which may serve a protective role against PH (*BMP4*, *RYR3*, *HTRB1*, and *AVPR1A*), while others may be associated with the development of PH and excessive erythrocytosis (*SOD3* and *KCNA5*). *BMP4* or bone morphogenic protein 4 is upregulated in hypoxia-induced pulmonary hypertension (Frank, et al. 2005; Miriyala, et al. 2006). We found increased methylation upstream of *BMP4*, which may be associated with lower expression of *BMP4* and have a protective role against PH. Type-3 Ryanodine Receptor (*RYR3*) is under hypoxic regulation and implicated in PH via mediating the hypoxia-induced calcium release in pulmonary artery smooth muscle cells (Zheng, et al. 2005;

49

Zhang, et al. 2012). We found increased methylation in the CpG island in the promoter region of *RYR3* suggesting downregulation of this gene in individuals born at high altitude and possible protective effects against PH. *HTR1B* encodes a lung 5-hydroxytryptamine receptor 1B that mediates the contraction of human pulmonary arteries, and *HTR1B* knockout mice are protected against hypoxic PH (Eddahibi, et al. 2000; Keegan, et al. 2001). We found increased methylation in a CpG island in the promoter region of *HTR1B* suggesting it might be downregulated in the high-altitude born Quechua and thus potentially adaptive against pulmonary hypertension. *AVPR1A* (Arginine Vasopressin Receptor 1A) belongs to the group of genes in the angiotensin aldosterone system (RAAS) that can influence blood pressure and risk of myocardial infarction, and has been previously associated with increased blood pressure, as well as hypertension (Nossent, et al. 2011; Natekar, et al. 2014). The expression of *AVPR1A* is increased in human heart failure (Zhu, et al. 2014). Our results suggest downregulation of *AVPR1A* in individuals born at high altitude, suggesting a potential adaptive role of birth at high altitude.

In addition to epigenetic changes that may be protective against PH, we also found differential methylation in genes that may be involved in the pathology of PH and CMS, such as *SOD3* and *KCNA5*. *SOD3* encodes the superoxide dismutase-3, an enzyme that converts ROS superoxide anion O_2^- to H_2O_2 and O_2, and is a major extracellular scavenger of superoxide anion O_2^- (Zelko, et al. 2002; Suliman, et al. 2004). *SOD3* promotes the expression of erythropoietin (EPO) in the kidney, which in turn increases the red blood cell production, and *SOD3* knockout mice have a lagging hematopoietic response to hypoxia (Suliman, et al. 2004). We identified higher DNA methylation levels in the promoter region of *SOD3*, which might be associated with decreased expression of *SOD3*. Decreased levels of SOD3 were shown in high-altitude natives of India, the Ladakhi, compared to acclimatized lowlanders (Padhy, et al. 2016). Decreased levels of *SOD3* may be related to the elevated hemoglobin levels in Andeans who are known to develop chronic mountain sickness with age (Monge 1942). Moreover, *SOD3* is implicated in the

development of PH. Several studies have proposed that SODs are associated with pulmonary hypertension due to playing a role in lung function and vascular tone (Xu, Cheng, et al. 2011). Animals with loss of function of *SOD3* have exacerbated pulmonary arterial hypertension and hypoxia-induced right ventricular hypertrophy (Xu, Guo, et al. 2011). It is unclear whether the increased methylation of *SOD3* promoter region is associated with adaptation or "maladaptation" to high altitude.

Kv1.5 channel gene (*KCNA5*) polymorphisms are associated with idiopathic pulmonary arterial hypertension (Remillard, et al. 2007; Wipff, et al. 2010). *KCNA5* expression is decreased in hypoxia, which may be associated with PH (Wang, et al. 1997). We found increased methylation of the CpG island in the promoter region of *KCNA5* suggesting decreased *KCNA5* expression in the high-altitude Quechua.

We saw a significant difference in the rate of epigenetic aging between individuals born at high and low altitudes (β=1.65, p-value=0.019), with participants born at high altitude having higher rates of epigenetic aging (Fig. 3.5). We also saw a positive association between the rate of epigenetic aging and hemoglobin levels (β=0.81, p-value=0.002, Fig. 3.6) suggesting a potential link between epigenetic age and CMS. The epigenome is not static throughout life and is altered with age in the process called 'epigenetic clock' (Teschendorff 2013). Moreover, epigenetic patterns get increasingly deregulated with age, and can serve as markers of the epigenetic maintenance system, as well as of premature biological ageing, especially as a result of previous environmental and toxicant exposures (Horvath 2013; Teschendorff, West, et al. 2013; Kochmanski, et al. 2017). For examples, a 5-year Δage between epigenetic and chronological age has been shown to associate with higher mortality risk (Marioni, et al. 2015). The acceleration of epigenetic age at high altitude may be explained by increased levels of ROS at high altitude, since hypoxia induces the production of ROS and ROS have been hypothesized to affect the epigenetic clock (Weitzman, et al. 1994; Guzy 2005; Cencioni, et al. 2013; Mitteldorf 2013).

Moreover, the association between accelerated epigenetic age and increased hemoglobin levels could reflect the potential loss of adaptation in Andeans, some of whom develop excessive erythrocytosis and CMS as they age (Leon-Velarde, et al. 2005).

Limitations

In this study, we are using DNA methylation data from whole blood, which is composed of various cell types, each with its own methylation landscape. However, we are controlling for blood cell types bioinformatically following the procedure proposed by Houseman et al. (2012). Moreover, individuals recruited in Cerro de Pasco have higher levels of lead exposure than the individuals recruited in Lima. Lead is known to affect DNA methylation, which is why our statistical models of differential methylation are controlled for lead exposure (Goodrich, Dolinoy, Sanchez, et al. 2016). Lastly, we were not able to recruit a fourth developmental group in our migrant design study, i.e. individuals born at low altitude who moved to high altitude, since it is not common for Peruvians from generally more urban low-altitude locations to move to high-altitude locations that tend to be less developed and remote.

NPY1R that we identified as potentially protective against PH is also involved in cold exposure response, and potentially cold adaptation. NPY receptors also are involved in controlling lipolysis and leptin secretion in human adipocytes (Serradeil-Le Gal, et al. 2000). Furthermore, increased activity of *NPY1R* is associated with obesity, and NPY levels are increased upon cold exposure (Zukowska-Grojec and Vaz 1988) . *NPY1R* could have a role cold adaptation (Kuo, et al. 2007), which is significant because high-altitude environments are known for low ambient temperatures and the epigenetic changes in *NPY1R* may also reflect the effects of cold, and not only hypoxia.

Conclusion

Overall, our findings show that being born at high altitude can have lasting effects on the epigenome, at the level of individual positions and regions. More importantly, we found a differentially methylated position and several differentially methylated regions associated with buffering against pulmonary hypertension. We also found significant differentially methylated regions and positions that are associated with hypoxic regulation, as well as developmental processes. Lastly, we show that the rate of epigenetic aging is faster in individuals born at high altitude.

Materials and Methods

Study Population

The full study cohort analyzed in this study is described in greater detail elsewhere (Childebayeva, et al. 2019). For EWAS, study participants consisted of multigenerational Peruvian Quechua individuals (Table 3.1): 1) high-altitude Quechua (HAQ) (n=56): Quechua individuals born and raised at high altitude who were recruited from the city of Cerro de Pasco, Peru (4,338 m) and Quechua individuals who were born in various high altitude locations (>2,500 m) and moved to low-altitude Lima at some point in their lifetime; 2) low-altitude Quechua (LAQ) (n=31)*:* Quechua individuals born and raised at low altitude who were recruited in Lima (150 m). Participants were unrelated, healthy, non-pregnant/lactating, non-smokers, between 18 and 35 years old (Table 3.1). At the time of enrollment, hemoglobin levels, height, weight, BMI, and phenotypes associated with high-altitude adaptation, including forced vital capacity (FVC), forced expiratory volume (FEV), arterial oxygen saturation (SaO_2), exhaled nitric oxide (FeNO), and others were collected. All study participants were screened for anemia using altitude specific cut-offs (11 g/dl for females and 13 g/dl for males).

DNA methylation analysis

We analyzed a total of 90 samples for genome-wide DNA methylation signatures using the Illumina Infinium® MethylationEPIC (EPIC) BeadChip. We used the EZ-96 DNA Methylation™ Kit (Zymo Research, Irvine, CA) to bisulfite convert each DNA sample following the standard protocol with alternative incubation conditions optimized for the Illumina Infinium® MethylationEPIC BeadChip assay. Raw data was loaded into R in the form of idat files using the minfi package (version 1.20.0) (Ayree et al., 2014). QC was performed using minfi and ENmix to assess the quality of the bisulfite conversion and probe hybridization for each sample. In minfi, we plotted log2 unmethylated median intensities over log2 methylated median intensities, compared predicted sex based on X and Y methylation to recorded sample sex, and generated density bean plots of beta values. We also visualized methylated and unmethylated intensities of control probes using ENmix. Based on QC metrics, three HAQ samples were excluded from the analysis.

Sample Preprocessing and Normalization. Data preprocessing and normalization are crucial steps in the array analysis. Background correction, dye-bias adjustment, and within-array normalization are commonly performed to adjust for technical variation in the sample, (Marabita et al., 2013). We performed probe correction and normalization using preprocessIllumina followed by preprocessSWAN in minfi. Probes that were above the 0.001 detection p-value threshold in more than 10% of the samples, i.e. failed to hybridize, were excluded from the analysis (n = 4,169). Cross-reactive probes (n = 43,254), probes associated with sex chromosomes (n = 16,290), probes containing SNPs with MAF > 5% at target CpG sites (n = 12,510), single base extension sites of type I probes (n = 414), and in the body of the probe (n = 110,445) were removed (Chen, et al. 2013; Pidsley, et al. 2016). Samples were tested for batch effects using singular value decomposition (SVD) analysis in champ.SVD performed on n= 656,183 beta values. Batch was determined to play a significant role on the variation of the genome-wide DNA methylation data based on the output from SVD, and corrected for using champ (Supp. Fig. 3.3).

Blood cell type heterogeneity correction. DNA methylation is tissue and cell-type specific. Therefore, it is crucial to correct for cell-type composition, since cell-type variation can be correlated with phenotype (Liu et al., 2013). Blood is composed of seven to eight cell types: neutrophils, eosinophils, basophils, CD14+ monocytes, CD4+ T cells, CD8+T cells, CD19+ B cells, and CD56+ natural killer cells (Houseman et al., 2012; Reinius et al., 2012; Teschendorff et al., 2017). Thus, we estimated cell counts in blood computationally using minfi following the method by Houseman et al. 2014 (Houseman et al., 2014). Houseman's is a reference-based method that is based on an *a priori* defined DNA methylation reference for the cell types present in blood (Houseman et al., 2012). We performed a PCA on whole-blood cell counts, and used the first two PCs and PCs 5 and 6 in our statistical models to correct for differential methylation associated with cell-type heterogeneity in blood (Jaffe et al., 2014). We chose to only include PCs that were among the most significant contributors of variation based on our singular value decomposition analysis of genome-wide methylation (Sup. Fig. 3.3).

Differential Methylation Analysis. Fully-processed M-values were tested for differential methylation using the package limma. The following model was used to test for the differentially methylated positions: Y_i (DNA methylation) <- B00 + B01(Altitude of birth) + Sex + Lead levels + PC1 (blood cell types) + PC2 (blood cell types) + PC5 (blood cell types) + PC6 (blood cell types) + gPC5 (genome-wide PC5). The covariates to include in the model were determined based on the analysis of overall variation in DNA methylation based on singular value decomposition (SVD) analysis (Sup. Fig. 3.3). Models adjusted for sex, WBLLs, genome-wide principal component five, and for blood cell principal components one, two, five, and six reflecting the blood cell types (Houseman, et al. 2012). The resulting p-values were adjusted for multiple testing using the false-discovery rate (FDR) following the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995).

Comb-p was used on the results of the limma analysis to test for differentially methylated regions (DMRs) (Pedersen, et al. 2012a). Regions were determined based on differentially

methylated probes at the unadjusted p-value<0.05 adjacent to one another within 3,000 base pairs of each other. DMRs were defined as three or more significant differentially methylated positions in close proximity.

Pyrosequencing. Quantitative pyrosequencing was performed to assess DNA methylation levels of *SOD3* in n=603 samples. The following primer sequences were used to amplify the region of interest: *SOD3* forward primer sequence: 5'-[biotin]-TTGTGTGTTGAAGGTTATTGGTTATAAG-3'; *SOD3* reverse primer sequence: 5'-CCTCCTCTACCCCTCCCATTTTT-3'; *SOD3* sequencing primer: 5'-ACCCCTCCCATTTTTA -3'. We assessed the DNA methylation levels at four CpG sites in *SOD3* using the Pyromark Q96 pyrosequencer (Qiagen, Valencia, CA), for more detail see (Virani, et al. 2012). Of the 603 samples analyzed using pyrosequencing, 104 either failed pyrosequencing or had high coefficient of variance between duplicates (>10%) and were excluded from the statistical analyses.

Microarray genotyping. We generated microarray genotype data for all Andean Quechua participants using the Affymetrix (Santa Clara, CA) Axoim Biobanking array that features approximately 610,000 markers including haplotype tagging SNPs, exonic SNPs, and eQTL variants. We performed a PCA on autosomal variants that passed the QC of Affymetrix (n=373,260) with call rates > 95%, major allele frequency > 0.0001. A principal components analysis (PCA) was performed on genome-wide SNP data for the participants in this study together with three populations from 1000 genomes (90 East Asians including 45 Han Chinese from Beijing and 45 Japanese from Tokyo, 60 *Centre de Polymorphism Humain* [*CEPH*] Europeans, and 60 Yorubans from Ibadan, Nigeria).

TABLES

Table 3.1. Participant Characteristics

Variable	All (n=90)	Low-altitude Quechua (LAQ, n=31)	Migrant Quechua (MQ, n=31)	High-altitude Quechua (HAQ, n=28)
Average methylation	0.60 ± 0.39	0.60 ± 0.39	0.60 ± 0.38	0.60 ± 0.39
% Female	50.44 (41.34, 59.54)	52.78 (36.47, 69.09)	46.15 (30.51, 61.80)	52.63 (36.76,68.51)
Age, yr	24.38 ± 4.76	24.07 ± 4.46	25.23 ± 5.44	23.78 ± 4.25
Age migration, yr	NA	NA	14.30 ± 6.62	NA
Height (cm)*	158.91 ± 7.84	162.38 ± 7.16	158.08 ± 7.29	158.47 ± 7.84
Weight (kg)†	63.51 ± 10.28	67.73 ± 11.51	63.25 ± 10.40	59.79 ± 7.23
BMI**	25.13 ± 3.52	25.61 ± 3.60	25.28 ± 3.73	24.51 ± 3.21
WBLL (ug/dL)***	3.08 ± 1.83	2.11 ± 0.67	1.99 ± 0.61	4.91 ± 1.72
Hemoglobin (g/dL)***	15.37 ± 2.40	14.04 ± 1.51	14.17 ± 1.50	18.03 ± 1.65

Data are means ± SD, 95% confidence interval (CI) for proportions in brackets; NA, not applicable
* $p < 0.05$ for LAQ vs. MQ and LAQ vs. HAQ based on t-test
† $p < 0.05$ between all three groups based on t-test
** no significant difference between groups
*** $p < 0.05$ between LAQ and HAQ and MQ and HAQ

Table 3.2. Top differentially methylated positions identified by EWAS

CpG site	Gene	Chr**	Pos**	P-value***	Q-value	% Δ (High-Low)
cg17500103	NPY1R	4	164254082	1.23E-07	0.08	0.67
cg13866747	MIR17HG*	13	91965808	1.19E-06	0.21	-8.39
cg21380842	KCNJ6	21	39288830	1.25E-06	0.21	0.57
cg07676961	ADGRB2	1	32222490	1.30E-06	0.21	-1.31
cg27376707	A2BP1	16	6533187	1.83E-06	0.24	4.07
cg15567297	RHCG	15	90029474	2.96E-06	0.25	3.88
cg06895218	AKAP12*	6	151558246	3.26E-06	0.25	-2.05
cg14445518	LGALS8	1	236686751	3.56E-06	0.25	3.31
cg18405765	ZNF541	19	48052357	4.32E-06	0.25	-0.67
cg16980637	IGLON5	19	51815359	4.96E-06	0.25	2.52

*Gene is within 30kb of the CpG site

** Hg19 genome was used for coordinates

*** Genome-wide inflation factor for the model comparing observed p-values to expected p-values was

Table 3.3

Gene	Gene full name	Region	Δ Methylation	CpG Island	N sites in region	P-value region	Corrected p-value**	References
Blood Pressure								
EFCAB1	EF-Hand Calcium Binding Domain 1	chr8:49647579-49648520	+	YES	7	6.61E-07	4.61E-04	Levy et al., 2009; Ho et al., 2011
RBFOX1	RNA Binding Fox-1 Homolog 1	chr16:6532643-6534318	+	YES	6	2.64E-09	1.03E-06	He et al., 2017; Ruso et al., 2018
FAM84A	Family With Sequence Similarity 84 Member A	chr2:14773973-14776097	+	YES	9	7.46E-07	2.30E-04	Mei et al., 2012
CNTN4	Contactin 4	chr3:2140256-2142142	+	YES	4	4.43E-06	1.54E-03	Levy et al., 2017
AVPR1A	Arginine Vasopressin 1A Receptor	chr12:63543831-63545289	+	YES	8	3.08E-07	1.39E-04	Nossent et al., 2011; Batkai et al., 2012
GPR35	G Protein-Coupled Receptor 35	chr2:241563347-241564625	+	NO	3	2.72E-05	1.39E-02	Divorty et al., 2018
High-Altitude Pulmonary Edema								
MACROD2	MACRO Domain Containing 2	chr20:13975279-13977220	+	YES	14	6.41E-11	2.17E-08	Li et al., 2017
Pulmonary Hypertension								
SLIT3	Slit guidance ligand 3	chr5:168727088-168728271	+	YES	5	1.22E-06	6.77E-04	Phillips et al., 2014; Sa et al., 2017
ALDH1A2	Aldehyde Dehydrogenase 1 Family Member A2	chr15:58356761-58357892	+	YES	5	8.97E-06	5.19E-03	Ji et al., 2017
KCNA5	Potassium Voltage-Gated Channel Subfamily A Member 5	chr12:5152380-5153289	+	YES	4	2.09E-06	1.51E-03	Remillard et al., 2007; Pousada et al., 2014
HTR1B	5-Hydroxytryptamine Receptor 1B	chr6:78172337-78174287	+	YES	6	1.07E-07	3.60E-05	Thomas et al., 2013; West et al., 2016

Gene	Name	Location						Reference
NRG1	Neuregulin 1	chr8:32404691-32406151	+	YES	6	1.04E-07	4.68E-05	Mendes-Ferreira et al., 2016; de Jesus Perez et al., 2014
BMP4	Bone morphogenetic protein 4	chr14:54412780-54413932	+	YES	4	1.82E-05	1.03E-02	Frank et al., 2005; Anderson et al., 2010
FAAH	Fatty acid amide hydrolase	chr1:46859671-46860512	+	YES	5	4.48E-07	3.49E-04	*Wenzel et al., 2013; Baranowska-Kuczko et al., 2014*
SOD3	Superoxide Dismutase 3	chr4:24796514-24797177	+	NO	5	1.51E-06	1.50E-03	Xu et al., 2011; Wedgwood et al., 2011
ZNF776	Zinc Finger Protein 776	chr19:58220080-58221623	+	YES	9	3.72E-06	1.58E-03	Hoffman et al., 2014
RYR3	Ryanodine Receptor 3	chr15:33601870-33603961	+	YES	8	1.49E-08	4.68E-06	Kinnear et al., 2008; Zheng et al., 2005

Preeclampsia

Gene	Name	Location						Reference
NAPRT1	Nicotinate phosphoribosyltransferase	chr8:144659492-144661052	+	YES	9	1.61E-10	6.78E-08	Leslie et al., 2015; Yeung et al., 2016
ADAM12	ADAM Metallopeptidase Domain 12	chr10:128076260-128078011	+	YES	8	4.96E-06	1.86E-03	Gack et al., 2005

Differentially expressed in hypoxic conditions

Gene	Name	Location						Reference
PAX3	Paired Box 3	chr2:223163588-223166095	+	YES	7	1.22E-05	3.19E-03	Li et al., 2005
LGALS8	Galectin 8	chr1:236685524-236687899	+	YES	10	6.12E-08	1.69E-05	Boso et al., 2006
PCDH7	Protocadherin 7	chr4:30720877-30723984	+	YES	8	6.41E-08	1.35E-05	Koklanaris et al., 2006; Irigoyen et al. 2007

61

Gene	Name	Location	Strand		Count			Reference
GRM7	Glutamate Metabotropic Receptor 7	chr3:6902100-6903922	+	YES	8	1.71E-06	6.17E-04	Felfly et al., 2011
NEFM	Neurofilament Medium	chr8:24771466-24772514	+	YES	6	2.92E-07	1.83E-04	Prasad et al., 2017
TBX18*	T-Box 18	chr6:85472949-85474804	+	YES	10	3.22E-10	1.14E-07	Jing et al., 2016; Kocabas et al., 2012
TJP1	Tight Junction Protein 1	chr15:30114291-30115300	+	YES	8	2.24E-07	1.46E-04	Hsu et al., 2017
IRS2	Insulin Receptor Substrate 2	chr13:110437759-110440180	+	YES	7	1.21E-07	3.28E-05	Mardilovich and Shaw, 2009; Wei et al., 2013
LYPD1	Ly6/PLAUR Domain-Containing Protein 1	chr2:133426653-133428653	+	YES	7	2.45E-07	8.04E-05	Frump et al., 2017
SIX3*	SIX Homeobox 3	chr2:45155408-45156938	+	YES	5	8.72E-08	3.74E-05	Garita-Hernandez et al., 2013
SOD3	Superoxide Dismutase 3	chr4:24796514-24797177	+	NO	5	1.51E-06	1.50E-03	Fike et al., 2013
EPHA5	EPH Receptor A5	chr4:66535145-66536773	+	YES	9	1.78E-06	7.17E-04	Xiaoli et al., 2013
ADAM12	ADAM Metallopeptidase Domain 12	chr10:128076260-128078011	+	YES	8	4.96E-06	1.86E-03	Cui et al., 2015
USP29*	Ubiquitin Specific Peptidase 29	chr19:57617602-57618437	+	YES	4	4.04E-06	3.17E-03	Watson et al., 2009
YAP1*	Yes Associated Protein 1	chr11:102139342-102140602	+	YES	4	1.12E-05	5.81E-03	Involved in oxygen stress response

HIF regulated

Gene	Name	Location	Strand		Count			Reference
NPTX2	Neuronal Pentraxin 2	chr7:98246006-98247143	+	YES	3	1.75E-05	1.00E-02	Expression regulated by HIF-1a
KITLG	KIT Ligand	chr12:88973398-88975429	+	YES	8	5.23E-07	1.69E-04	Bartoszewska et al., 2915
RBPJL	Recombination Signal Binding Protein For Immunoglobulin Kappa J Region Like	chr20:43944741-43945736	+	YES	4	3.87E-08	2.55E-05	Okumura et al., 2017

TNXB	Tenascin-X	chr6:32063501-32065703	+	YES	19	9.22E-08	2.75E-05	Koizume et al., 2015; Warnecke et al., 2008
SLC35F3	Solute Carrier Family 35 Member F3	chr1:234367145-234367494	+	YES	3	1.40E-06	2.63E-03	Wan et al., 2009
Angiogenesis								
S100A13	S100 Calcium Binding Protein A13	chr1:153599479-153600598	+	NO	14	4.14E-08	2.43E-05	Landriscina et al., 2006

Table 3.4. Pyrosequencing results for *SOD3*

	Gene	Sample Size	β	Std. Error	p-value
Altitude of birth*	*SOD3*	499	1.83	0.50	0.00029

*adjusted for group, sex, age, BMI

FIGURES
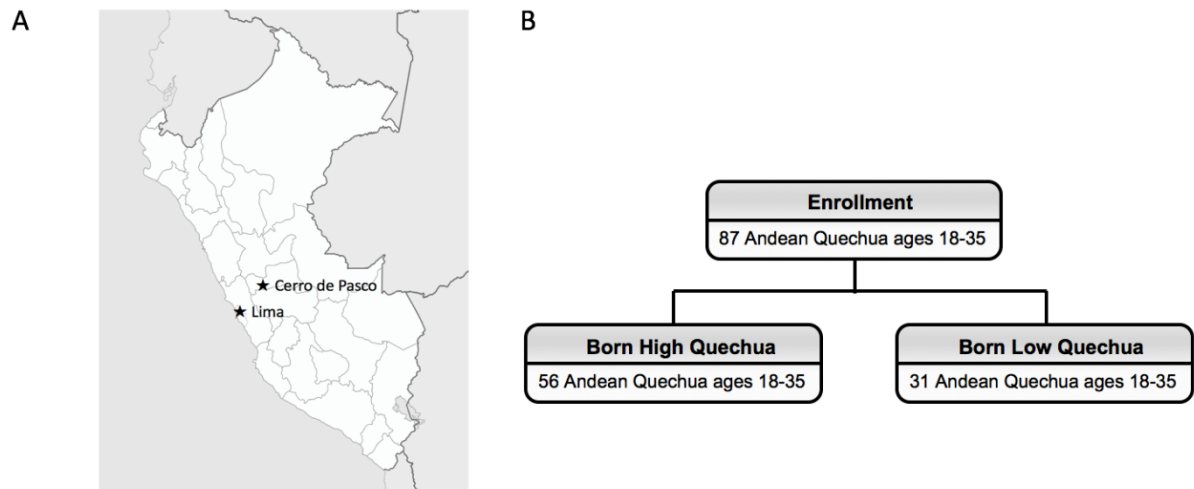
Figure 3.1. Study Design. (A) Map of Peru depicting high-altitude, Cerro de Pasco (4,300 m), and low-altitude, Lima (150 m), participant recruitment locations; (B) Genome-wide DNA methylation was determined in 87 Peruvian Quechua with different exposures to high-altitude hypoxia over their lifetime. Born High Quechua consist of High-Altitude Quechua (HAQ), i.e. individuals who were born, raised, and reside in or near Cerro de Pasco, Peru and Migrant Quechua (MQ) were born and raised at high altitude (>3,000 m), but moved to low altitude Lima (0m) during the lifetime. Born Low Quechua (LAQ) were born, raised, and reside at sea-level in Lima.
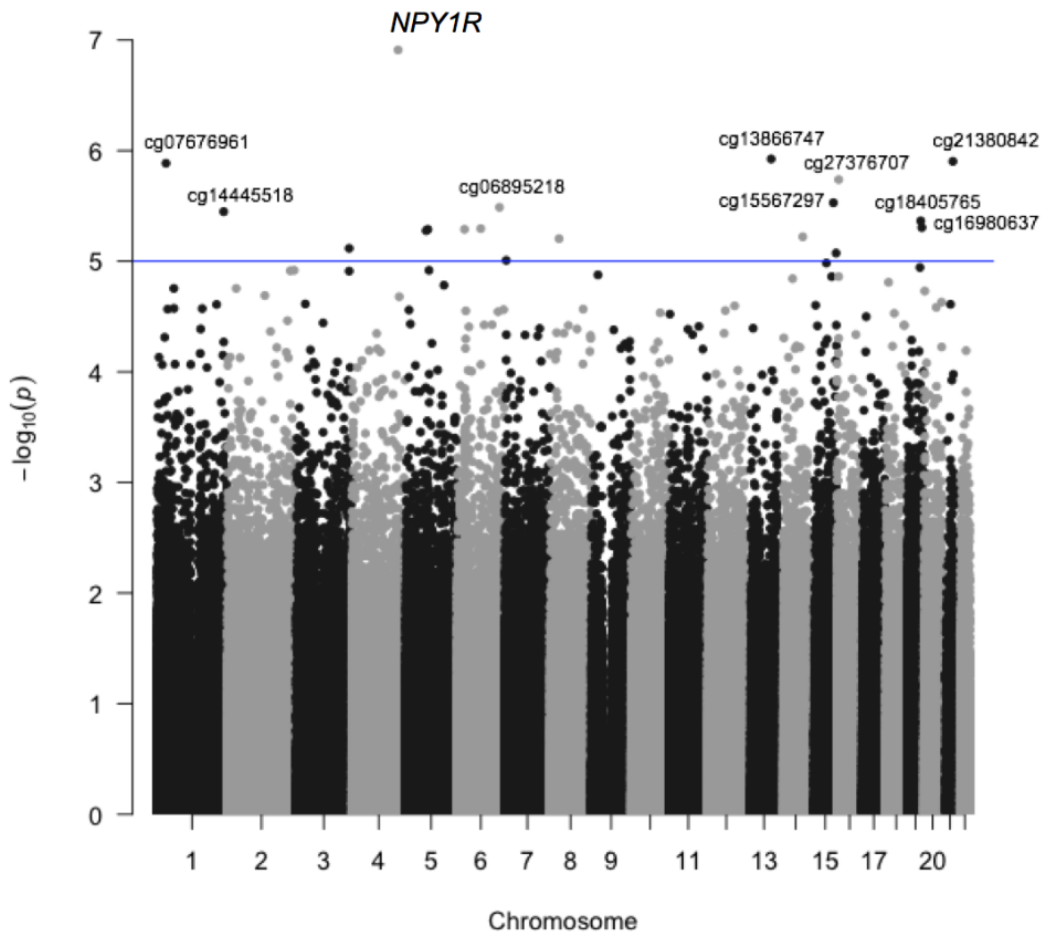
Figure 3.2. Manhattan plot of differentially methylated positions. Top hits are marked by their CpG site probe ID. The –log(P value) is shown on the y axis for each CpG site according to its genomic location. The significant site (q<0.1) is marked with its gene ID.
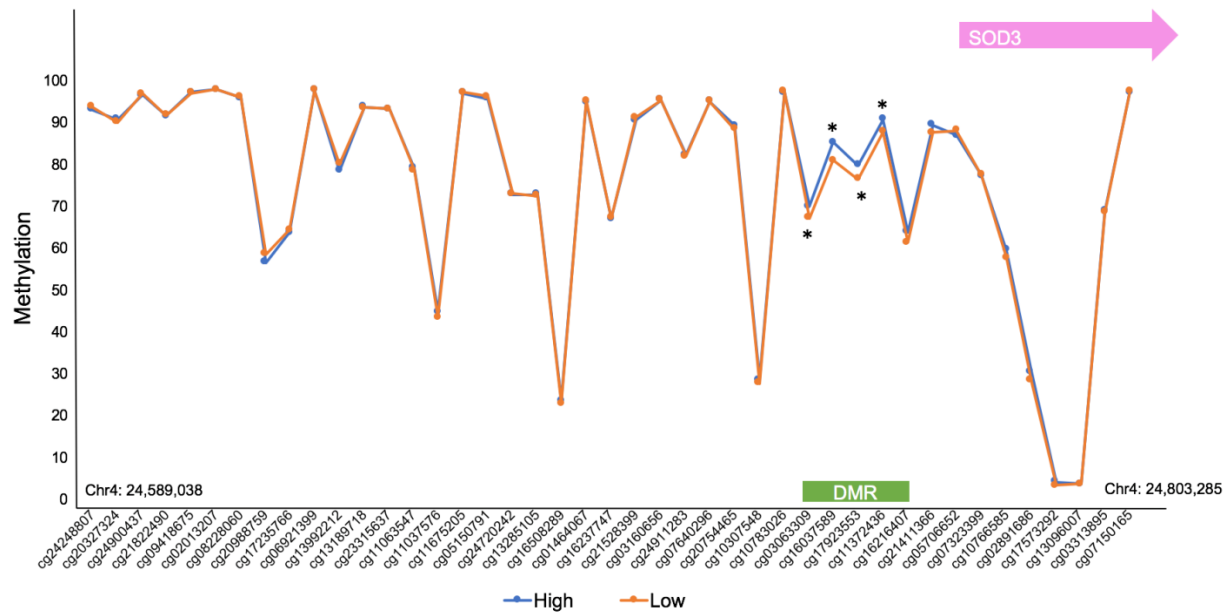
Figure 3.3. *SOD3* differentially methylated region. Genetic coordinates correspond to hg19 build.

Significant DMRs are marked with *
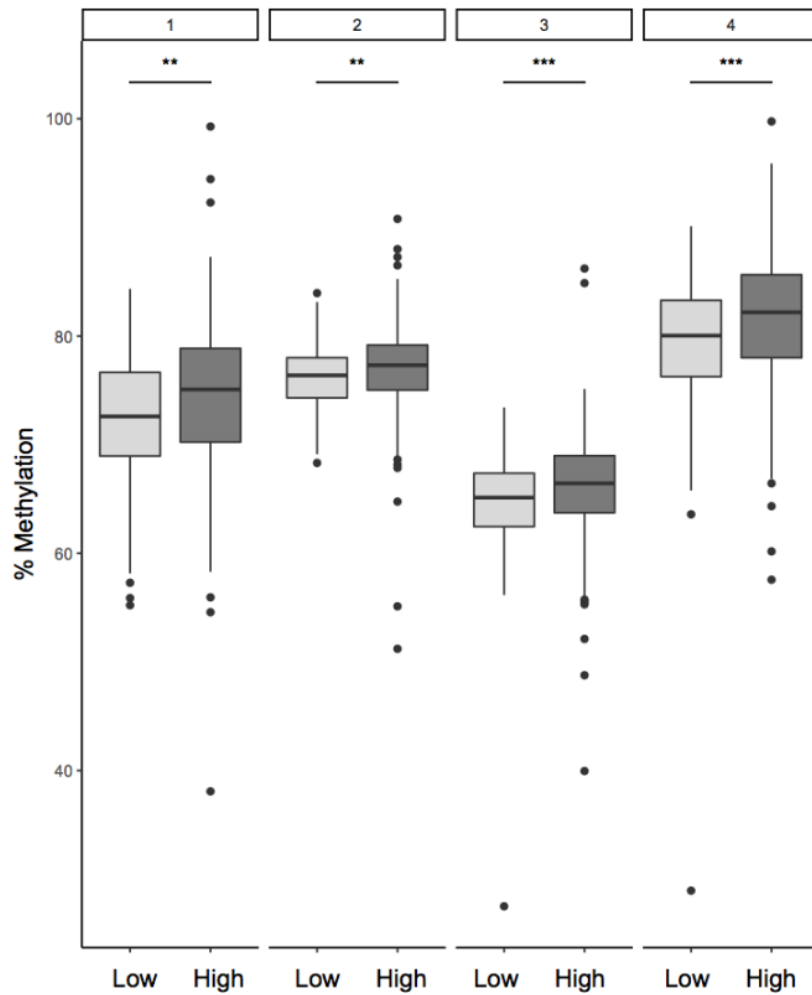
Figure 3.4. Boxplot of *SOD3* methylation by altitude of birth for each CpG position analyzed in the gene. The x axis represents the altitude of birth (low vs. high), while y axis represents DNA methylation levels of each *SOD3* CpG site. Significance levels are shown for each CpG site separately.

** p-value <0.01

*** p-value < 0.001
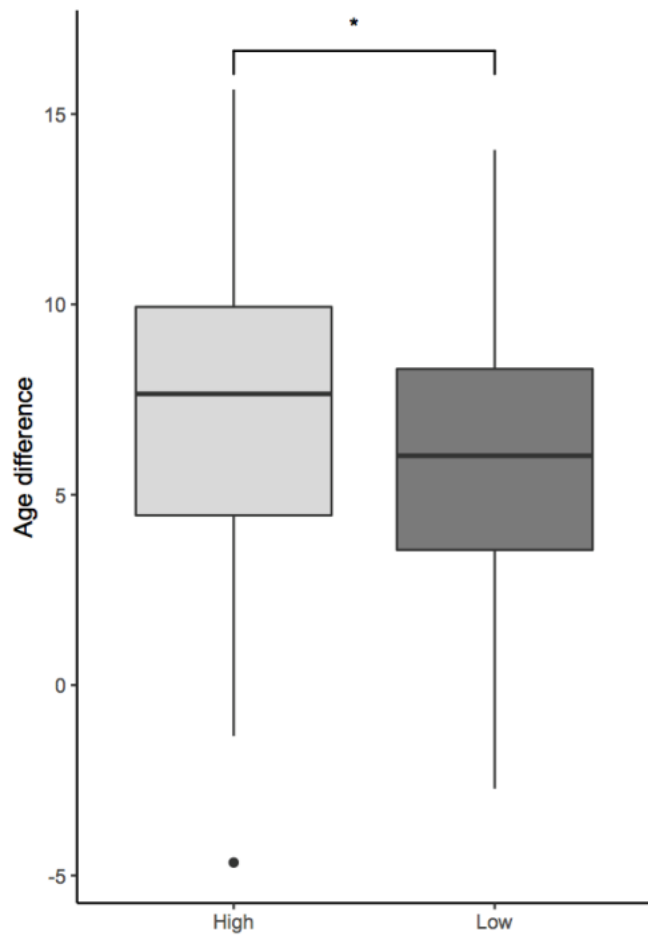
Figure 3.5. Boxplot of ΔDNAmAge (Horvath Age – Chronological Age) in individuals born at high versus low altitude. Individuals born at high altitude have significantly higher Δage (β=1.65, p-value=0.019) when adjusted for chronological age, group, the first 2 blood PCs (batch, unique name, and slide were coded as random effects).
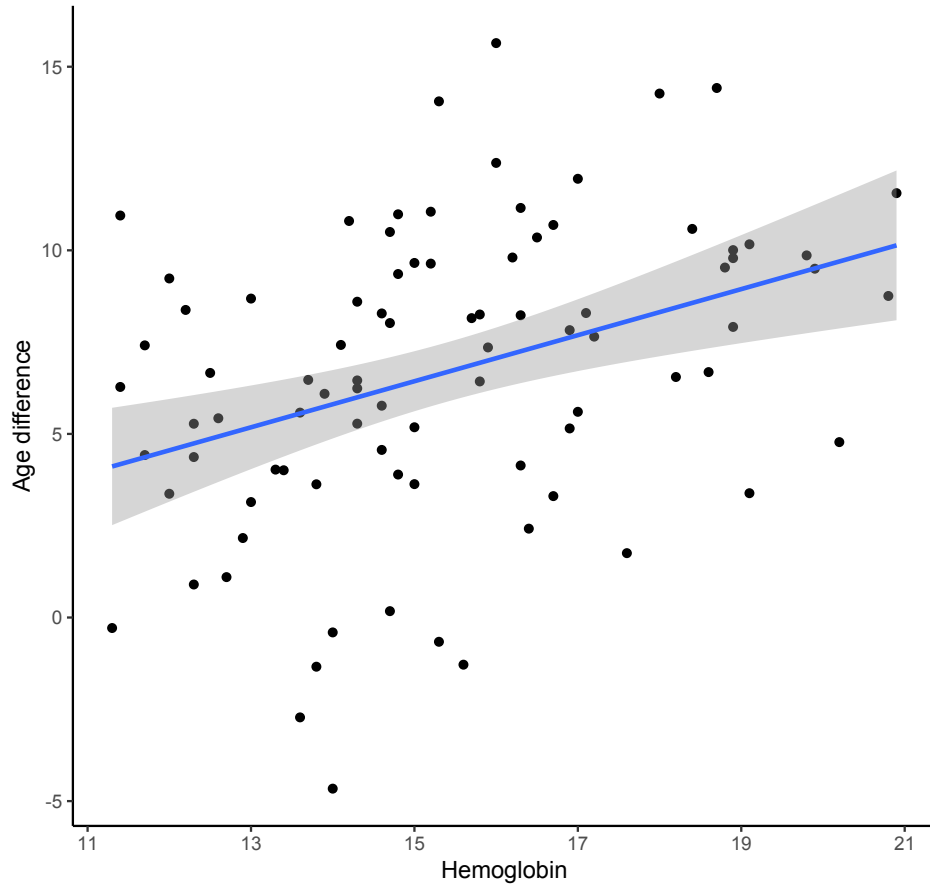
Figure 3.6. ΔDNAmAge (Horvath Age – Chronological Age) is plotted against hemoglobin levels.

There is a positive association between Δage and hemoglobin levels (β=0.81, p-value=0.002)

when adjusted for chronological age, group, the first 2 blood PCs.

CHAPTER 4

Blood Lead Levels in Peruvian Adults Are Associated with Proximity to Mining, Soil Lead, and

DNA Methylation

1. Introduction

Lead (Pb) is a trace metal that has been associated with human civilization since the emergence of metallurgy (Mielke and Reagan 1998). Pb is the world's most abundant heavy metal, and exposure to it is estimated to account for 143,000 deaths per year, and 0.6% of the global disease burden (Tong, et al. 2000; Haefliger 2011). Miners and mining communities are at a higher risk of exposure to Pb and its effects, including higher levels of Pb in both indoor and outdoor air, soil, and self-produced vegetables in areas close to Pb mining (Qu, et al. 2012).

Pb exposure can have negative effects on any organ in the body, including the neurological, hematological, cardiovascular, and renal systems, with the nervous system being the most susceptible to its harmful effects (Kalia and Flora 2005; Sanders, et al. 2009; WHO 2011). Both central and peripheral nervous systems can be affected by Pb, with the central nervous system being more affected in adults, and the peripheral nervous system in children (Flora et al., 2012). In children, Pb exposure has been associated with encephalopathy characterized by a degeneration of certain parts of the brain (Sanders, et al. 2009). Moreover, children exposed to Pb may experience delayed growth, decreased IQ, impairment in concentration ability, and other adverse outcomes (Sanders, et al. 2009; Flora, et al. 2012). Additionally, Pb has a direct

effect on the hematopoietic system by inhibiting enzymes involved in the heme synthesis pathway, thus leading to decreased hemoglobin concentration (Barltrop and Smith 1971; Hu, et al. 1994). Pb also increases the fragility of erythrocyte membranes, which reduces their lifespan, and results in anemia (Flora, et al. 2012). Lastly, renal dysfunction occurs at high levels of Pb exposure, but even relatively low levels of Pb exposure can be damaging to the kidneys (Grant 2009a).

Pb is predominantly stored in calcified tissues, and more than 94% of it is deposited in bone, which represents the largest Pb storage site in the body (Barbosa, et al. 2005; Bernal, et al. 2013). Pb stored in bone is dynamic and heterogeneous, and its deposition is dependent on the density and structure of the underlying tissue (Grant 2009a). In adult women, 50–75% of blood Pb is immediately skeletal in origin rather than environmental and its release is increased during pregnancy and lactation, periods which are associated with lower calcium availability (Gulson and Calder 1995; Elias and Gulson 2003; Gulson, et al. 2003; Grant 2009a).

There is growing evidence that Pb can impact the epigenome. Epigenetic modifications can have an effect on gene expression without changing the sequence of the nucleotides. DNA methylation is the most commonly studied epigenetic modification and is generally associated with gene repression (Fraga and Esteller, 2002). Inorganic Pb may affect DNA methylation via increased production of reactive oxygen species (ROS) (Adonaylo and Oteiza 1999). Increased oxidative stress is known to result in decreased DNA methylation (Pogribny, et al. 2007), and it can inhibit binding of the methyl-CpG binding proteins that promote chromatin inactivation (Valinluck et al., 2004).

Previous studies have shown associations between Pb exposure and both global LINE-1 and genome-wide DNA methylation signatures in humans and human embryonic stem cells (Wright, et al. 2010a; Hanna, et al. 2012; Li, et al. 2013a; Senut, et al. 2014; Li, et al. 2016). However, little is known about the effects of recent Pb exposure on the genome-wide DNA methylation in adults exposed to mining who are not miners themselves. We determined whole-blood lead levels

72

(WBLLs) and DNA methylation using the Illumina EPIC DNA methylation chip to run epigenome-wide association studies. We determined the relationship between Pb and hemoglobin, proximity to mining, soil Pb, and menstrual cycle in women. We found significant associations between Pb and hemoglobin, proximity to mining, soil lead, menstrual cycle, and DNA methylation.

2. Materials and Methods

2.1 Study Population

Study participants consist of Peruvian Quechua individuals stratified based on location (Table 4.1): 1) Cerro de Pasco (n=157): Quechua individuals who were born and raised at high altitude and recruited in Cerro de Pasco, a Peruvian mining town and continued to reside there at the time of the study; 2) Lima (n=148): Quechua individuals living in Lima as young adults at the time of study, some of whom were born and raised in Lima and some who were born in various locations in Peru and moved to Lima sometime before recruitment. Male and female participants were recruited to participate if they were unrelated, healthy, non-pregnant/lactating, non-smokers, non-miners, and between 18 and 35 years old (Table 4.1). At the time of enrollment, all study participants provided written informed consent approved by Syracuse University, Universidad Peruana Cayetano Heredia, and The University of Michigan Institutional Review Boards.

2.2 Phenotypic Assessment and Pb Exposure Assessment

Height, and weight were collected at the time of enrollment, and the participants provided a venous blood sample into a collection tube containing EDTA (Lavender-top, BD Vacutainer, Franklin Lakes, NJ) for DNA extraction and Pb analysis. The Puregene DNA purification system (Qiagen, Valencia, CA) was used to extract the genomic DNA according to the manufacturer's instructions. Hemoglobin levels were determined in venous blood using the HemoCue 201+ analyzer (Angelholm Sweden). All study participants were screened for anemia using altitude specific cut-offs (11 g/dl for females and 13 g/dl for males). Days since the last menstrual cycle were self-reported. WBLLs were measured using graphite furnace atomic absorption

spectrometry (GFAAS) (Parsons and Slavin 1993). GFAAS measurement was performed in duplicate at the Blufstein Laboratory located in Lima, Peru (coefficient of variance between the duplicates for the entire cohort < 2).

2.3 Geocoding and Soil Pb Interpolation

Geographic coordinates for the participants from Cerro de Pasco were determined using a portable Garmin GPS (Garmin, Olathe, Kansas) based on the address information received during recruitment. We determined the Euclidean distance from the mine (mine-distance) for each participant's current place of residence from Cerro de Pasco using the function distm (function Harvesine) in the R package geosphere (Hijmans, et al. 2017).

Van Geen et al. (2012) recorded soil Pb values at 74 locations throughout Cerro de Pasco at varying distances from the mine in May 2009, using a hand-held X-ray fluorescence analyser (Innov-X Alpha, Olympus Corporation, Tokyo, Japan) (van Geen, et al. 2012). Using the soil Pb values recorded by van Geen et al. (2012), we interpolated the soil Pb levels at the household locations for participants from Cerro de Pasco, using ordinary kriging (Graler, et al. 2016). Kriging is a geostatistical interpolation method that produces a prediction surface based on points with known values, and incorporates a semivariogram model to account for the estimated spatial autocorrelation between these points (Krige 1966; Burgess and Webster 1980). We created the predicted surface of soil Pb values using ordinary kriging in ArcGIS 10.4, with a spherical semivariogram model and a cell size of 30 meters. In addition to a prediction surface, kriging also creates a surface of estimated prediction error based on the distance between points. This approach has been previously used to interpolate blood Pb level risk with Pb in water for children in Flint, MI (Hanna-Attisha, et al. 2016).

2.4 DNA methylation

A random subset of participants was selected from both Lima (n=59) and Cerro de Pasco (n=31) for DNA methylation analysis. EZ-96 DNA Methylation™ Kit (Zymo Research, Irvine, CA) was used for bisulfite conversions of all DNA samples following the standard protocol with

alternative incubation conditions optimized for the Illumina Infinium® MethylationEPIC BeadChip (Zymo, Irvine, California). Raw methylation intensity data was loaded into R using the package minfi (Aryee, et al. 2014). Eighty-seven samples passed QC, and their beta values were imputed, normalized using functional normalization (FunNorm) in minfi followed by the beta-mixture quantile normalization (BMIQ) (Teschendorff, Marabita, et al. 2013; Fortin, et al. 2014; Morris, et al. 2014). We removed probes with detection p-values>10e-5 in more than 10% of samples, probes identified as cross-reactive, probes containing common SNPs, and probes on X and Y chromosomes (for more information on the pipeline see Supp. Fig. 4.2) (Chen, et al. 2013; Pidsley, et al. 2016). The resulting beta values (n=656,183) were tested for batch effects using Singular Value Decomposition (SVD) (Supp. Fig. 4.3), and the batch effects were corrected using Combat in ChAMP (Johnson, et al. 2007; Teschendorff, et al. 2009). Following batch correction, beta values were converted into M-values, and differentially methylated positions were determined using in limma (Ritchie, et al. 2015).

Differentially methylated regions (DMRs) were determined based on the uncorrected p-values<0.05 from the limma analysis using comb-p with the following options (comb-p pipeline -c 4 --seed 10e-4 --dist 3000) (Pedersen, et al. 2012a). DMRs were defined as three or more differentially methylated probes at the unadjusted p-value<0.05 adjacent to one another within 3,000 base pairs of each other.

Differentially methylated regions were corrected for multiple comparisons using the Sidak correction function in comb-p (Pedersen, et al. 2012b). Pathway analysis was also performed on significant (p-value<0.05) unadjusted sites from the limma analysis using Consensus PathDB (Kamburov, et al. 2011; Herwig, et al. 2016).

2.5 SNP Genotyping

Microarray genotype data were generated for all Peruvian Quechua DNA samples using the Affymetrix (Santa Clara, CA) Axoim Biobanking array. The Axiom array features approximately 610,000 markers including haplotype tagging SNPs as well as exonic and

expression quantitative trait loci (eQTL) variants. We performed a principal components analysis (PCA) of genome-wide SNP data for the participants in this study together with three 1000 genomes populations (90 East Asians including 45 Han Chinese from Beijing and 45 Japanese from Tokyo, 60 *Centre de Polymorphism Humain* [*CEPH*] Europeans, and 60 Yorubans from Ibadan, Nigeria). PCA was performed on 328,260 autosomal SNPs with genotyping rates > 95%, MAF > 0.0001, that passed Affymetrix QC, and that were not in strong linkage disequilibrium (r^2<0.8) (Supp. Figs 5A-C).

2.6 Statistical Analyses

All statistical analyses on WBLLs and phenotypes of interest were conducted using R version 3.4.0 (R Core Team, 2018). Package ggplot2 (Wickham 2009) was used to plot the data. We ran linear models to assess the relationship between interpolated soil Pb values and WBLLs in a subset of participants from Cerro de Pasco whose geographical location we were able to determine based on the home address they provided (n=152). We used interpolated soil Pb values for each participant in the following regression model to find associations between soil Pb and WBLLs: Y_i (WBLLs) ~ B00 + B01(interpolated soil Pb) + B02(Sex) + B03 (Education) + B04(Kriging standard error), where Y_i is the dependent variable, whole-blood lead level; "interpolated soil Pb" = interpolated soil Pb value.

Differentially methylated positions were determined in limma using the following model: DNA methylation ~ WBLLs + Altitude of birth + Sex + Blood PCs (1,2,3,5,6) + Genetic PCs (1-3). Blood PCs were determined based on the estimated cell counts function proposed by Houseman et al. (2012). Covariates in the model were chosen based on a singular variable decomposition analysis (Sup. Fig. 3.3). Genetic principal components (PCs) were determined based on the genome-wide SNP data (see previous section) using PLINK and were included to account for the effects of genetics on DNA methylation (Purcell, et al. 2007). Benjamini-Hochberg correction was used to control for multiple testing and a false discovery rate of 10% was considered significant (Ritchie, et al. 2015). Distribution of expected and observed p-values was used to validate the

model fit (Supp. Fig. 4.4).

3. Results

3.1 Study design and participant characteristics

We recruited 305 Peruvian individuals of self-reported multigenerational Quechua ancestry between ages 18-35 (Figure 1A). The individuals were recruited in two locations with expectations of higher Pb exposure in the first location: 1) Mining community: 157 individuals recruited in Cerro de Pasco, a mining town in the highlands of Peru; 2) Urban community: 148 individuals were from Lima, the capital of Peru and the biggest city in Peru. None of the Cerro de Pasco participants were mine workers even though Cerro de Pasco is a mining town (Fig. 4.1B, the study design is described in greater detail in Childebayeva et al., 2019). Participant characteristics are provided in Table 4.1. The study groups did not differ significantly by sex or age.

3.2 Lead levels are associated with study group, sex, and hemoglobin

In this study, WBLLs were associated with study group (Fig. 4.2A, Table 4.1). Participants from Cerro de Pasco had significantly higher WBLLs (4.75 ±1.53 μg/dL) compared to the participants recruited in Lima (2.03±0.62 μg/dL) (β=2.72, p-value=2e-16). Sex was significantly associated with WBLLs in both Lima and Cerro de Pasco, with males having higher Pb levels than females (Fig. 4.2B) in each group (Lima: β=0.31, p-value=2.40E-03, Cerro de Pasco: β=1.15, p-value=1.13e-06).

Interestingly, we saw a positive association between WBLLs and hemoglobin levels when adjusted for sex and sample group in the entire cohort (β=0.17, p-value= 8.60E-04) (Fig. 4.3, Table 4.2). However, when considering the groups separately, WBLLs were only significantly associated with hemoglobin levels in participants from Lima (β=0.32, p-value=0.007), but not in Cerro de Pasco, although the model was approaching significance (β=0.16, p-value=0.069)

Moreover, WBLLs were associated with the number of days since the last menstruation (dlmp) for women in the entire cohort (Fig. 4.4, Table 4.2; β=0.02, p-value=0.002) and in Cerro de Pasco (Fig. 4.4, Table 4.2; β=0.04, p-value=0.003). When the same model was tested for the women recruited in Lima, WBLLs were not associated with dlmp (Fig. 4.4, Table 4.2; β=0.003, p-value=0.53).

3.3 Blood lead is negatively correlated with distance from the mine and soil lead

The distance from the mine was negatively associated with WBLLs when adjusted for sex and education (Fig. 4.5, β=-0.0005, p-value=0.032) (Fig. 4.4). Association between WBLLs and interpolated soil Pb (mg/kg) values in Cerro de Pasco adjusting for sex, education, and Kriging standard error was approaching significance (β=0.0003, p-value=0.08) (Table 4.3, Figs. 4.6-7).

3.4 DNA methylation and Blood Lead

We performed a singular value decomposition (SVD) analysis of genome-wide DNA methylation data to capture the most salient factors of variation in the data with a small number of components (Teschendorff et al., 2009). Based on SVD, WBLLs were among the significant components of DNA methylation variation (p<0.05) before and after experimental batch effects adjustment (Supp. Fig. 4.3).

Moreover, we found five significant differentially methylated positions (DMPs), i.e. single CpG sites, associated with WBLLs at q-value<0.10 (Table 4.4). These DMPs were within and/or in regulatory regions of genes: *ZMIZ1, TERT, ADO, TGM7, and SOX18* (Fig. 4.8). We performed a pathway analysis of the probes significant at the unadjusted p-value<0.05 over all probes as background using ConsensusPathDB, based on which we found n=134 enriched gene ontology-based sets under the q-value 0.05 cutoff, including "neuronal part" (GO:0097458, q-value=7.99E-15), "neuron to neuron synapse" (GO:0098984, q-value=3.67E-08), and "neurotransmitter secretion" (GO:0007269 , q-value= 0.00161) (Supp. Table 4.2).

Lastly, we identified twenty-one significant differentially methylated regions (DMRs), i.e. regions with three or more significant differentially methylated positions within a 3000 base-pair window, using comb-p (Table 4.5). Out of the significant DMRs, eight are associated with heavy metal exposure and toxicity and seven with neurological function and conditions (Table 4.5).

4. Discussion

We tested the effects of Pb exposure in a cohort of adults from two locations in Peru: Lima, the capital, and Cerro de Pasco, a high-altitude mining town. Cerro de Pasco is an important Peruvian mining center that was established in the 17th century to support silver mining, and was the second largest producer of silver in the Andes (Fisher, 1977; Helfgott, 2013). Upon the depletion of the city's silver reserves, silver mining has been replaced by copper, zinc, and Pb extraction (van Geen et al. 2012). In 1952, the Cerro de Pasco Corporation claimed to be the largest South American producer of refined Pb (Helfgott, 2013).

A study conducted in 2007 (Conklin et al., 2008) found that WBLLs in children and women from Cerro de Pasco were significantly above the 5 µg/dL U.S. Centers for Disease Control and Prevention (CDC) blood Pb reference level for children. Another study conducted in 2012 reported measures of Pb concentration in the soil around the mine that were above 1200 mg/kg within 1km from the open pit, while the U.S. Environmental Protection Agency-suggested threshold for soil Pb toxicity is 400 mg/kg (van Geen et al., 2012). We show higher WBLLs in participants from Cerro de Pasco compared to the participants from Lima. Moreover, we show a positive association between WBLLs and proximity to the Cerro de Pasco mine, and interpolated soil Pb values that also tend to be higher close to the mine.

Most of Pb is stored in calcified tissues with the majority of it being eliminated via kidney filtration. However, Pb also binds to hemoglobin in the whole blood, and WBLLs are representative of recent (~40 day) Pb exposure (Barltrop and Smith 1972; Barbosa, et al. 2005). The relationship between Pb and hemoglobin remains inconclusive. Two previous studies in

79

children and adults did not find significant associations between blood Pb and hemoglobin (Huo, et al. 2007; Roy, et al. 2011), whereas a third has shown a negative association between Pb and hemoglobin (Counter, et al. 2000).

In adults, a previous study has shown a significant association between bone Pb levels in patella and decreased hemoglobin levels, but no correlation between blood Pb levels and hemoglobin (Hu et al., 1994). In Andeans, a study by Counter et al. 2001 showed increased Pb exposure associated with lowered hemoglobin, however, the relationship was not statistically significant ($r = -0.289$, p-value= 0.103) (Counter, et al. 2001). Another study by Counter et al. (2000) on Andean children reported a negative association between hemoglobin and blood Pb.

Contrary to most of the previous studies of hemoglobin and blood Pb levels, we show a statistically significant positive association between them. One potential reason why we see a positive association could be that the levels of Pb in our cohort are low compared to previous studies. For example, Counter et al.'s (2000) mean Pb values in a cohort of Andeans involved in Pb-glazing was 30.2 μg/dL (SD: 20.4, range: 2.0 – 80.7 μg/dL) compared to the mean of 4.75 μg/dL (SD: 1.53, range: 1.57-9.81 μg/dL) in Cerro de Pasco and 2.03 μg/dL in Lima (SD: 0.62, range: 1.02-4.61 μg/dL).

Another explanation for the positive relationship between Pb and Hb levels could be that high-altitude adaptation is playing a role in alleviation of the effects of inorganic Pb on the individuals in our study. Our participants are of self-reported high-altitude Quechua ancestry, and indigenous Andeans possess known genetic and physiological adaptations to life at high altitude (Brutsaert 2001; Bigham, et al. 2010; Bigham and Lee 2014). Increased hemoglobin levels are common in high-altitude Andeans, and considered an adaptive mechanism. However, Andeans are known to be at a higher risk for Chronic Mountain Sickness, or excessive erythrocytosis characterized by a loss of adaptation to high altitude (Moore, 2000). In our cohort, the relationship between Hb and Pb is stronger in Quechua participants recruited at low altitude, suggesting that

at lower levels of Pb exposure, increased Hb is a potentially adaptive mechanism. This is especially true since Pb is known to cause anemia in children and adults (Papanikolaou, et al. 2005).

We found an association between WBLLs and the number of days since last menstruation in women in the entire cohort and in Cerro de Pasco, but not in Lima. Pb is a known reproductive toxicant. Previous studies of the effects of Pb exposure on sexual maturation in girls from the U.S.-based Third National Health and Nutrition Examination Survey (NHANES) showed an association between higher blood Pb levels and delayed attainment of menarche and pubic hair suggesting that Pb exposure may influence sexual maturation (Wu, et al. 2003). Moreover, women who were occupationally exposed to Pb show lower age at the onset of menopause (Popovic, et al. 2005). Studies of women working in lead-based industries have reported cases of sterility, frequent abortions, disturbances in ovulation, and abnormal menses (Winder, 1993). In animals, high levels of Pb exposure affect sexual development and menstrual function (Der, et al. 1974; Laughlin, et al. 1987). To our knowledge, this is the first report of a positive association between Pb and the menstrual cycle in post-pubertal women.

Pb exposure is known to affect DNA methylation. Previous studies in humans have shown effects of developmental Pb exposure on the expression of DNA methyltransferases *DNMT1* and *DNMT2,* and the methyl cytosine binding protein *MeCP2* in rat hipoccampi (Schneider, et al. 2013). Hanna et al. detected a statistically significant negative correlation between WBLLs and promoter methylation of the gene *COL1A2* (Hanna et al., 2012). A different study showed a positive correlation between WBLLs and DNA methylation of an important tumor suppressor *p16* in Pb-exposed individuals (Kovatsi, et al. 2010). Lastly, maternal Pb levels were shown to have an effect on child's DNA methylation profile (Sen, et al. 2015; Goodrich, Dolinoy, Sanchez, et al. 2016) (Sen et a., 2015).

Environmental metals, including Pb, can increase the production of reactive oxygen species (ROS) via redox cycling (Adonaylo and Oteiza 1999; Ahamed and Siddiqui 2007). This

can influence epigenetic processes, and specifically DNA methylation. Oxidative damage to the DNA can hamper with the ability of DNA methyltransferases to interact with the DNA, which may result in altered DNA methylation levels (Baccarelli and Bollati 2009).

We found differences in DNA methylation, both at the level of individual CpG sites and in regions of multiple neighboring CpG sites, associated with WBLLs (Tables 4.4 and 4.5). We found three DMPs significant at a q-value < 0.10 associated with genes *ZMIZ1, TERT*, and *SOX18* that previously have been shown to be related to neurological development. This is important since exposure to Pb has been linked with neurodegeneration in adults and neurodevelopmental effects in children (Needleman, et al. 1979; Stewart, et al. 2006).

ZMIZ1 (Zinc Finger MIZ-Type Containing 1) is a transcriptional co-activator known to interact with the androgen receptor. *ZMIZ1* variants have been associated with neurodevelopmental disorders (Fewings, et al. 2017; Carapito, et al. 2019). We saw decreased DNA methylation at a CpG site in a CpG island of *ZMIZ1* (Table 4.4, probe cg20543544). We saw decreased methylation of a CpG site in the second exon of *SOX18* (SRY [sex determining region Y] box 18; probe cg0135592). *SOX18* is a transcription factor involved in the regulation of embryonic development and cell fate determination, and when mutated can lead to cardiovascular dysfunction (Pennisi, et al. 2000; Downes and Koopman 2001). Previous research has shown downregulation of *Sox18* in male rat hippocampi associated with postnatal exposure to 250-ppm Pb acetate (Schneider, et al. 2012). We also report increased methylation of a CpG site in the CpG island spanning with the promoter region of the gene *TERT* (Table 4.4; probe cg17166338). *TERT* is a telomerase reverse transcriptase, a catalytic subunit of telomerase, an enzyme involved in telomere lengthening. The expression of *TERT* is positively associated with promoter hypermethylation in human-induced pluripotent stem cells (Takasawa, et al. 2018). Studies have shown decreased expression of *TERT* associated with the methylation status of its promoter (Jiang, et al. 2012). While others reported no association between the methylation status of the *TERT* promoter and its expression (Devereux, et al. 1999). Pb exposure is associated with

telomere shortening and instability (Wu, et al. 2012; Pottier, et al. 2013), and thus it is possible that Pb exposure affects the telomerase activity via DNA methylation.

We performed a pathway analysis among top DMPs, and identified three gene ontology pathways associated with neurological functioning: "neuron part", "neuron to neuron synapse", and "neurotransmitter secretion". Pb is a known neurotoxicant, and exposure to it has been associated with neuronal death, intraneuronal regulatory mechanisms, and neurotransmission, among others (Lidsky and Schneider 2003). A previous study, which used the Illumina HumanMethylation450 BeadChip, found Pb-related changes in neurogenetic signaling pathways in human embryonic stem cells (Senut et al., 2014).

We found eighteen significant differentially methylated regions (DMR) associated with WBLLs. All DMRs associated had less methylation at each CpG site with increasing Pb exposure, which has been shown previously with Pb exposure (Faulk, et al. 2013). Out of the eighteen significant DMRs, five were associated with genes linked to heavy metal exposure and toxicity *NNAT, S100A13, AGAP2, SLC43A2, LPCAT1* (Sivaraja, et al. 2006; Bigagli, et al. 2010; Seow, et al. 2014; Vidal, et al. 2015; Kisko, et al. 2018). Furthermore, seven of the significant DMRs were associated with genes linked to neurological function and disorders: *NNAT, SYNGAP1, PRRT1, MCF2L, AGAP2, CEACAM21* and *TYROBP* (Joseph, et al. 1995; Alkelai, et al. 2012; Carvill, et al. 2013; Li, et al. 2015; Ma, et al. 2015; Gasparoni, et al. 2018; Matt, et al. 2018). Interestingly, *NNAT* or neuronatin is an imprinted gene that is paternally expressed and known to play an important role in brain development (Joseph, et al. 1995; Evans, et al. 2001). A previous study has shown decreased methylation of *NNAT* in newborns associated with maternal cadmium exposure (Vidal, et al. 2015).

Our study contributes to the body of knowledge of Pb toxicity and epigenetics by showing an association between Pb and DNA methylation at individual CpG sites and in regions in a cross-sectional study of adults exposed to Pb due to proximity to mining.

5. Conclusions

In this study, we demonstrated an association between low levels of Pb exposure and various phenotypic measurements, and DNA methylation. In our cohort of Peruvian adults, Pb was associated with hemoglobin levels, proximity to mining, soil Pb, and menstrual cycle in women. We also showed differences in DNA methylation, mainly decreased methylation, in individuals with higher levels of Pb exposure. Our data indicate that proximity to mining can have profound effects on human physiology and epigenetics, even when the levels of exposure are low. We suggest future research be done on Pb exposure in children living in Cerro de Pasco, as well as pregnant women, since these two groups are especially vulnerable to the negative effects of Pb exposure.

6. Limitations

DNA methylation signatures differ by tissue and cell type (Kitamura, et al. 2007). In this study, we used DNA extracted from whole blood representing a mixture of DNA-containing cells. However, our statistical models of DNA methylation were adjusted for the principal components based on the blood cell types determined bioinformatically using the method by Houseman et al. (2012).

Diet is also known to affect DNA methylation signatures (Zhang, et al. 2011). Lima is an urban center and Cerro de Pasco is a remote, Andean town. Any potential differences in diet between participants recruited in Lima and Cerro de Pasco may contribute to the differences in DNA methylation that we observed.

Lastly, Lima and Cerro de Pasco differ by altitude. Lima is at an altitude of 0 meters above sea level (masl), and Cerro de Pasco is at 4,300 masl. Our previous work has shown that exposure to high-altitude hypoxia is associated with differences in LINE-1 global DNA methylation in this cohort (ADD CHILD 2019). It is possible that the DNA methylation differences seen between participants from Lima and Cerro de Pasco are due to the decreased oxygen availability observed at high altitude. We expect this effect to be mitigated however by our use of WBLLs in our analyses, which diminishes the confounding between Pb exposure and high-altitude hypoxia.

TABLES

Table 4.1 Participant characteristics

Phenotype	Total (n=305)	Lima (n=148)	Cerro (n=157)
Age*	24.75 ± 5.00	24.83 ± 5.05	24.67 ± 4.96
%female*	52%	52%	52%
Hemoglobin, g/dL (mean±SD)**	15.66 ± 2.66	13.74 ± 1.40	17.50 ± 2.24
WBLL, µg/dL (mean±SD)**	3.43 ± 1.53	2.03 ± 0.62	4.75 ± 1.53

*No significant differences in sex or age between the two locations found

** p-value=2.00E-16 comparing the two locations

Table 4.2 Significant associations between WBLLs and location and phenotypes

	Sample size	β	SE	p-value
Group *	305	2.72	0.13	2.00E-16
Hemoglobin**	305	0.22	0.07	8.60E-04
Lima: Hemoglobin*	148	0.32	0.12	7.00E-03
Cerro: Hemoglobin*	157	0.16	0.09	6.90E-02
Sex (M)	305	0.75	0.20	2.36E-04
Lima: Sex (M)	147	0.31	0.01	2.40E-03
Cerro de Pasco: Sex (M)	158	1.15	0.23	1.13E-06
Days since last menstruation***	159	0.02	0.007	1.75E-03
Lima: Days since last menstruation†	77	0.004	0.0059	5.30E-01
Cerro: Days since last menstruation†	82	0.04	0.01	2.90E-03

* Adjusted for sex, Lima was modeled as a reference

** Adjusted for sex and group

† Adjusted for hemoglobin

*** Adjusted for hemoglobin and group

Table 4.3 Associations between WBLLs and distance from mine and interpolated soil lead values

for participants from Cerro de Pasco

	β	SE	p-value	Adj. R2
Distance from mine (m)*	-0.0005	0.0002	0.04	0.176
Interpolated soil lead mg/kg**	0.0003	0.0001	0.08	0.181

* Adjusted for sex and education
** Adjusted for sex, education, and Kriging standard error

Table 4.4 Significant differentially methylated positions associated with WBLLs

Gene	CpG site	P-value	Q-value	Position	% Δ *	Relation to CpG Island
ZMIZ1	cg20543544	2.68E-08	0.02	chr10:81003657	-4.73	Island
TERT	cg17166338	3.54E-07	0.06	chr5:1295969	3.52	Island
ADO	cg01005506	4.33E-07	0.06	chr10:64565768	-3.07	Island
TGM7	cg06437171	4.79E-07	0.06	chr15:43571912	0.59	OpenSea
SOX18	cg01355392	4.81E-07	0.06	chr20:62679255	-7.01	N_Shore

* % methylation difference between Cerro de Pasco and Lima

Table 4.5 Significant differentially methylated regions associated with WBLLs

Gene	Region	Type	DMR location	N	P-value region	Corrected p-value	Δ Meth	Relevant function
NNAT	chr20:36147549-36150136	CpG island	Promoter region	41	1.76E-16	5.63E-14	-	Neurological function and disorders; heavy metal exposure and toxicity
PRRT1	chr6:32117281-32119686	CpG island	Gene body	25	2.99E-17	8.16E-15	-	Neurological function and disorders
WFIKKN2	chr17:48911313-48913414	CpG shore	Promoter region	20	1.37E-09	4.29E-07	-	
S100A13	chr1:153599479-153600973	CpG shore	Gene body; promoter region	19	4.33E-15	1.90E-12	-	Heavy metal exposure and toxicity
AGAP2	chr12:58131681-58133184	CpG island	Promoter region	10	1.08E-11	4.72E-09	-	Neurological function and disorders; heavy metal exposure and toxicity
CEACAM21	chr19:42069883-42071241	Open sea	Gene body	10	3.08E-08	1.49E-05	-	Neurological function and disorders
SYNGAP1	chr6:33400477-33401543	CpG shelf	Gene body	9	2.66E-10	1.64E-07	-	Neurological function and disorders
ZMYND8	chr20:45980204-45982167	Open sea	Gene body	9	7.52E-09	2.51E-06	-	
MCF2L	chr13:113655884-113656425	CpG island	Promoter region	9	4.29E-09	5.21E-06	-	Neurological function and disorders
TYROBP	chr19:36399185-36400438	Open sea	Promoter region	8	2.65E-09	1.39E-06	-	Neurological function and disorders
PEBP4	chr8:22734960-22735995	CpG shore	Gene body	6	2.27E-08	1.44E-05	-	
A1BG	chr19:58861502-58862399	CpG island	Gene body	6	1.87E-07	1.37E-04	-	
PRR5	chr22:45072724-45073149	non CpG island	Gene body	6	1.10E-07	1.69E-04	-	Neurological function and disorders
LCE3C	chr1:152571909-152572931	Open sea	5' TSS	5	8.51E-08	5.46E-05	-	

Gene	Location	Region	Feature	n	p-value	FDR		Pathway
SLC43A2	chr17:1507643-1508449	CpG shore	Gene body	5	5.49E-05	4.37E-02	-	Heavy metal exposure and toxicity
LPCAT1	chr5:1494980-1495357	CpG island	Gene body	4	1.12E-07	1.95E-04	-	Heavy metal exposure and toxicity
MAN1C1	chr1:26096954-26098215	CpG island	Gene body	4	1.48E-06	7.67E-04	-	
ZDHHC19	chr3:195913801-195914114	Open sea	~11 kbs upstream	3	3.09E-08	6.47E-05	-	
ZNF678	chr1:227748424-227749378	CpG island	~3 kbs upstream	3	3.08E-07	2.12E-04	-	
FBXL18	chr7:5528044-5528267	CpG island	Gene body	3	8.37E-08	2.46E-04	-	
FKBP1A	chr20:1317600-1317747	CpG shore	Gene body	3	4.56E-06	2.01E-02	-	

A   B
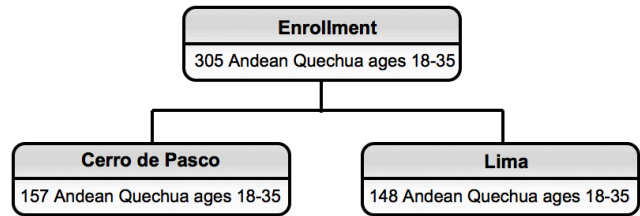
Figure 4.1. Study Design (A) Map of Peru depicting high-altitude, Cerro de Pasco (4,338 m), and low-altitude, Lima (0m), participant recruitment locations; (B) Whole-blood lead levels were determined in 305 Peruvian Quechua recruited in or near Cerro de Pasco, Peru and in Lima.
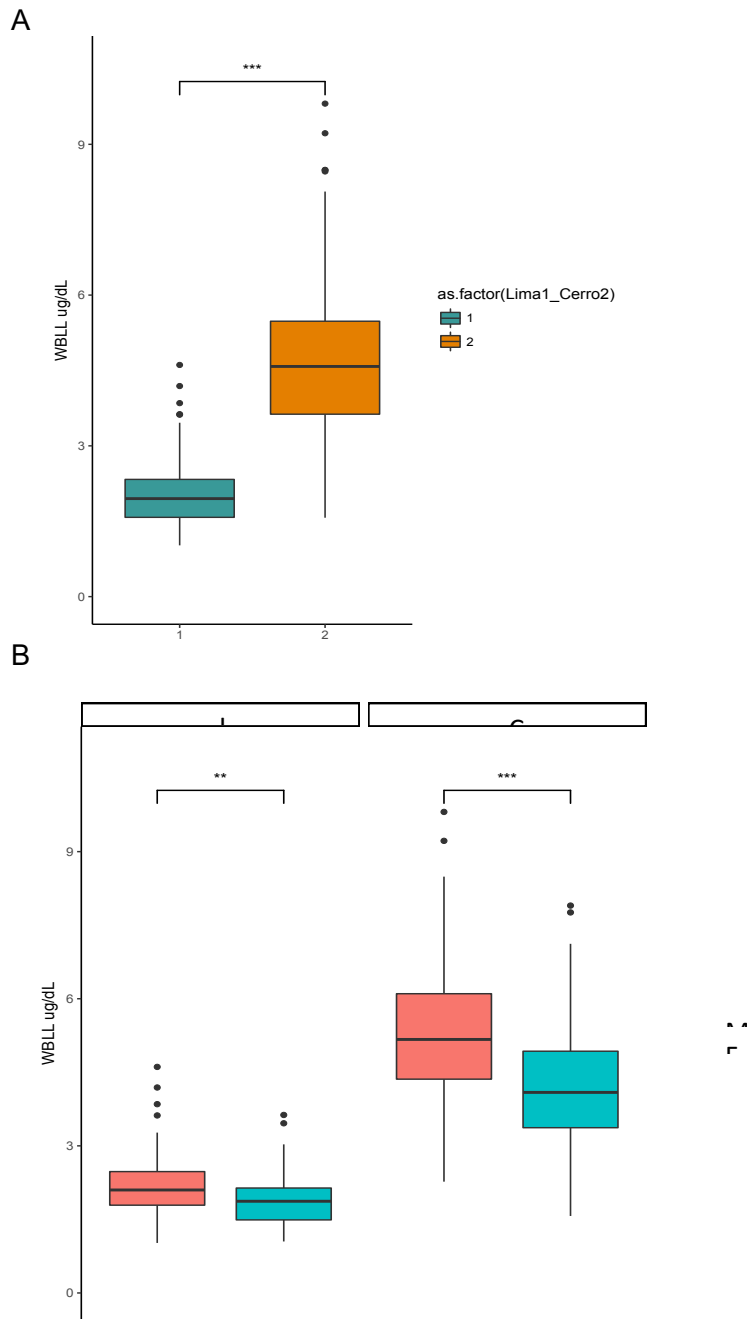
Figure 4.2. (A) WBLLs in Lima and Cerro de Pasco. Participants from Cerro de Pasco have higher WBLLs than the participants from Lima (p-value<2.2e-16).(B) WBLLs between Cerro de Pasco and Lima by sex. Males have higher WBLLs than females in both cities (Lima: p-value=0.0024; Cerro de Pasco: p-value=1.13E-06).
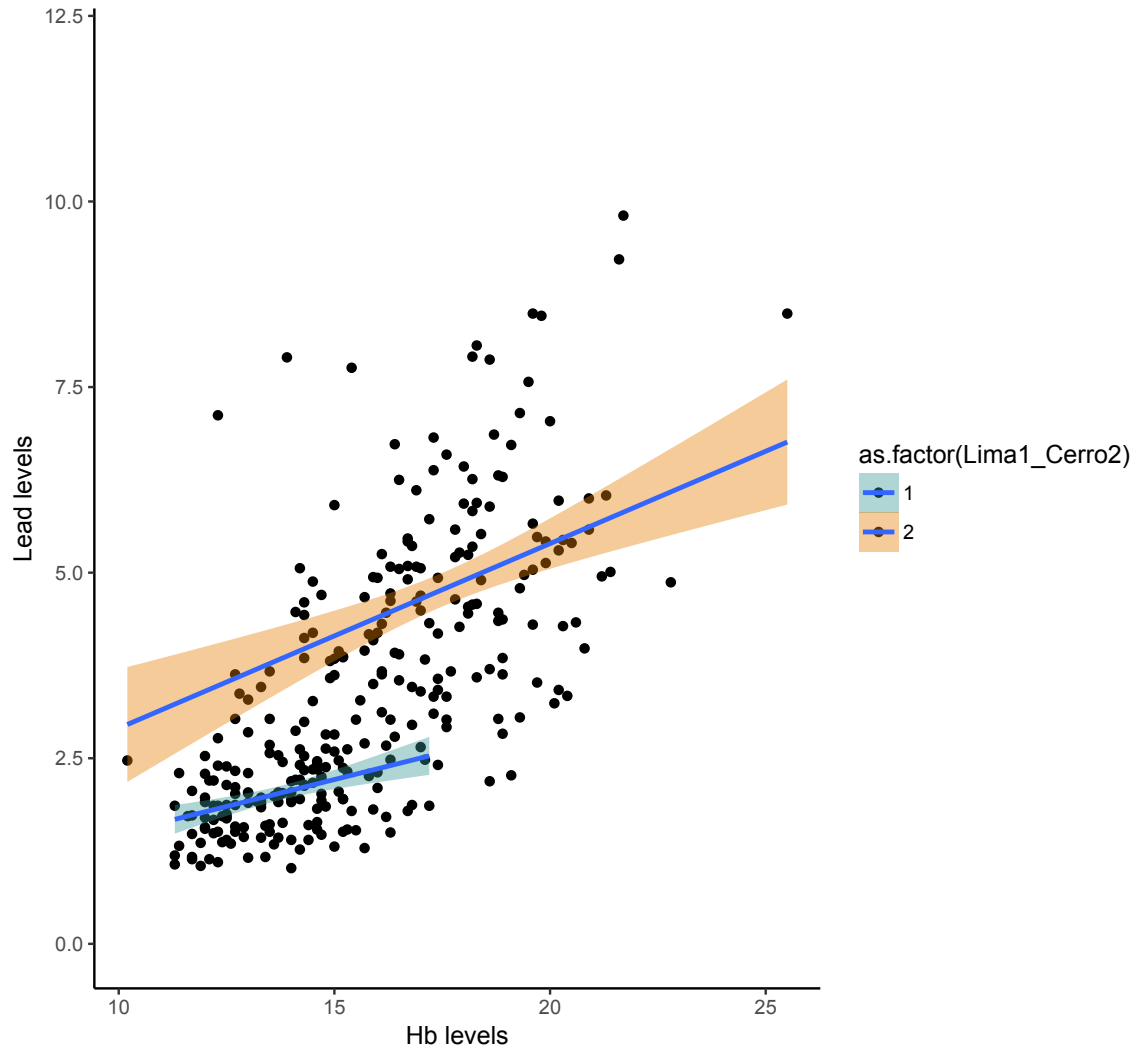
Figure 4.3. Association between hemoglobin and whole-blood lead levels (WBLLs) in participants from Lima and Cerro de Pasco separately. Statistical models were adjusted for sex.
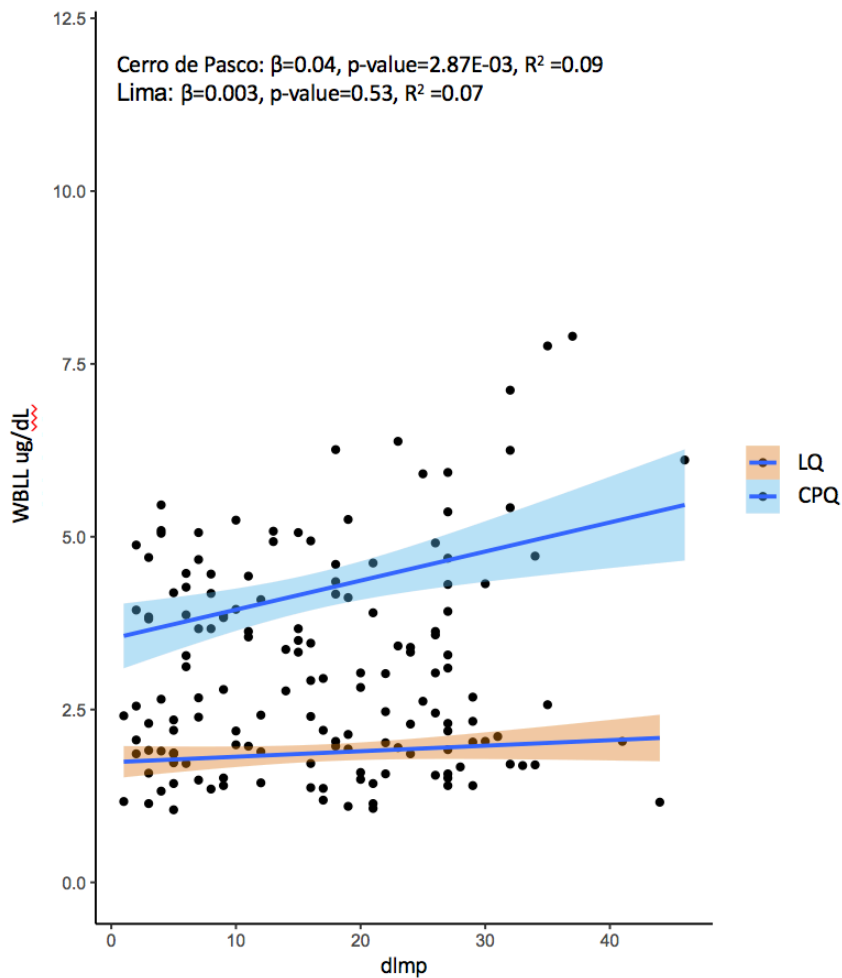
Figure 4.4. Association between days since the last menstruation and whole-blood lead levels (WBLLs) in women from Lima and Cerro de Pasco separately. WBLLs were significantly associated with the number of days since the last menstruation (dlmp) in the participants from Cerro de Pasco (β=0.04, p-value=2.87E-03, R2 =0.09), but not in Lima individuals (β=0.003, p-value=0.53, R2 =0.07). Statistical models were adjusted for hemoglobin.
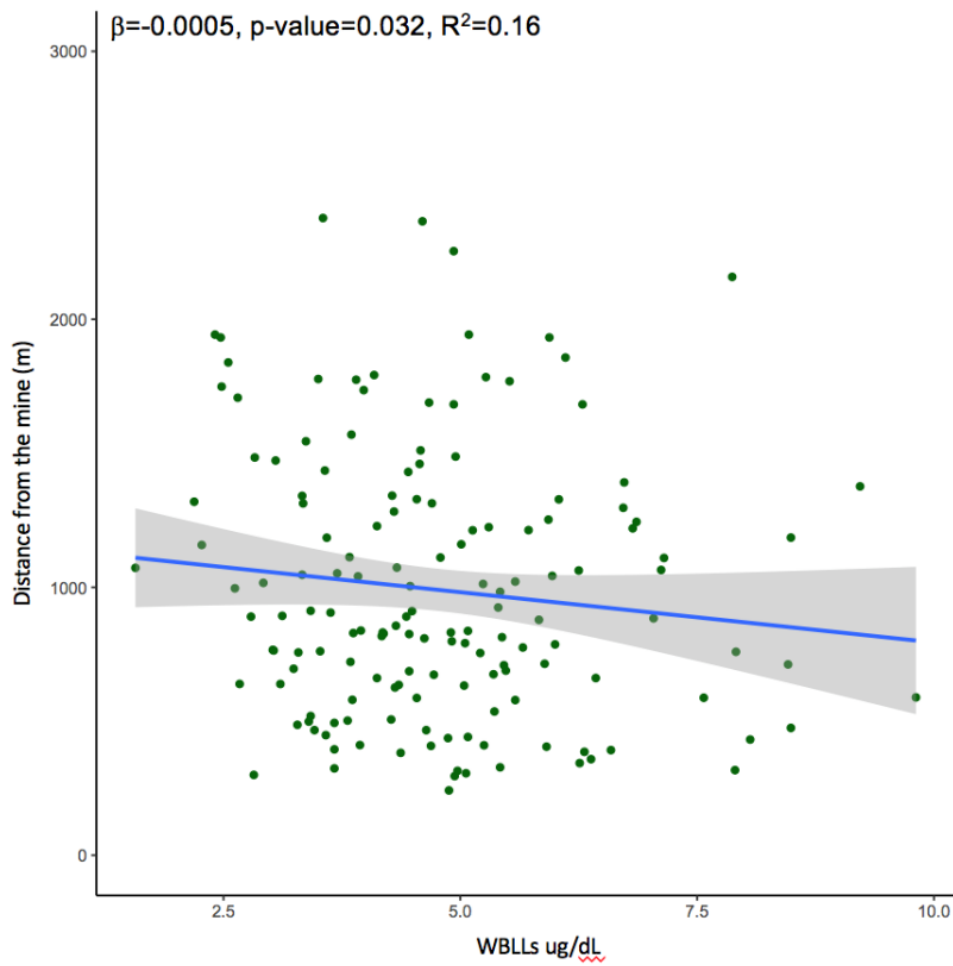
Figure 4.5. Whole-blood lead levels plotted against the distance from the mine. Whole-blood lead levels (WBLL) are negatively associated with distance from the mine in the participants from Cerro de Pasco (β=-0.0005, p-value=0.032, R2=0.16, model adjusted for age and sex).
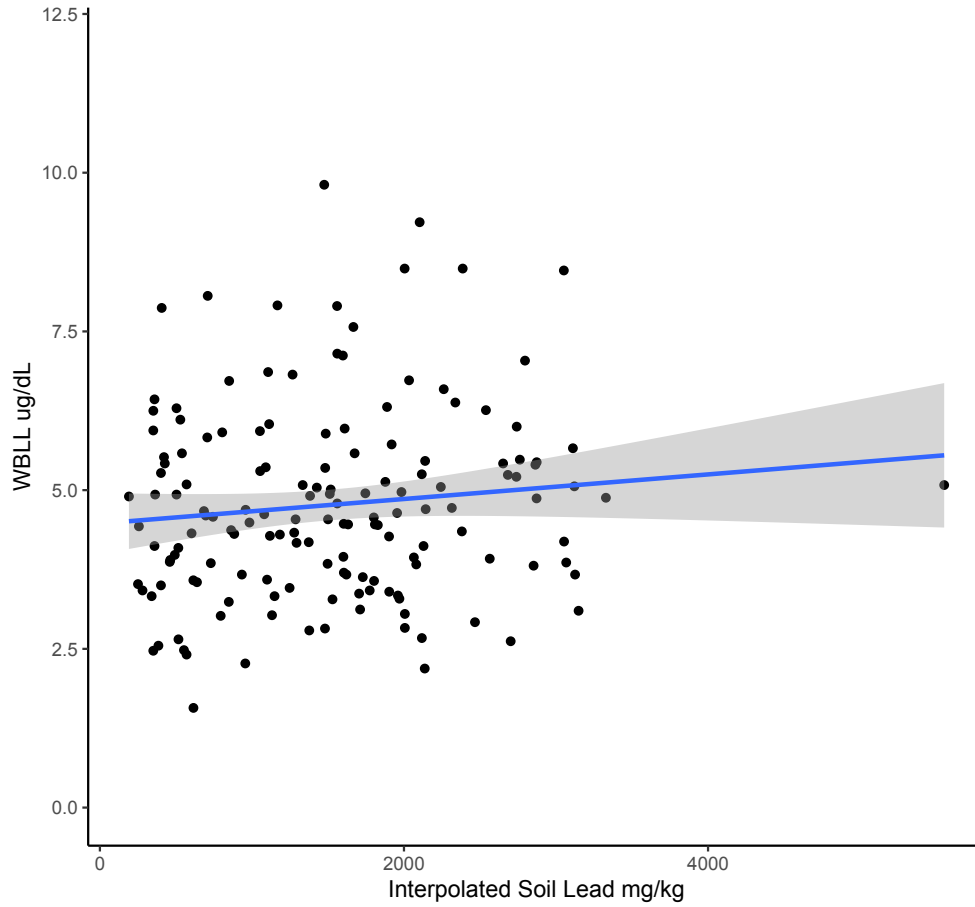
Figure 4.6. Relationship between interpolated soil lead and blood lead in participants from Cerro de Pasco. The model was adjusted for sex and education level (β=0.0003, p-value=0.02, R2(model)=0.18).
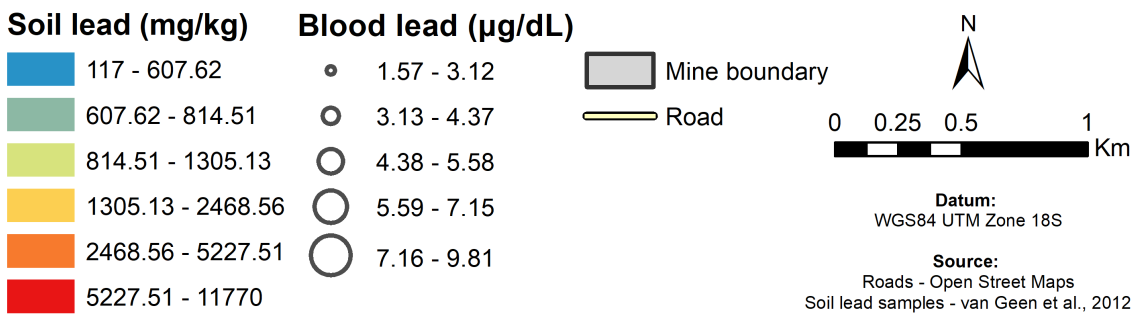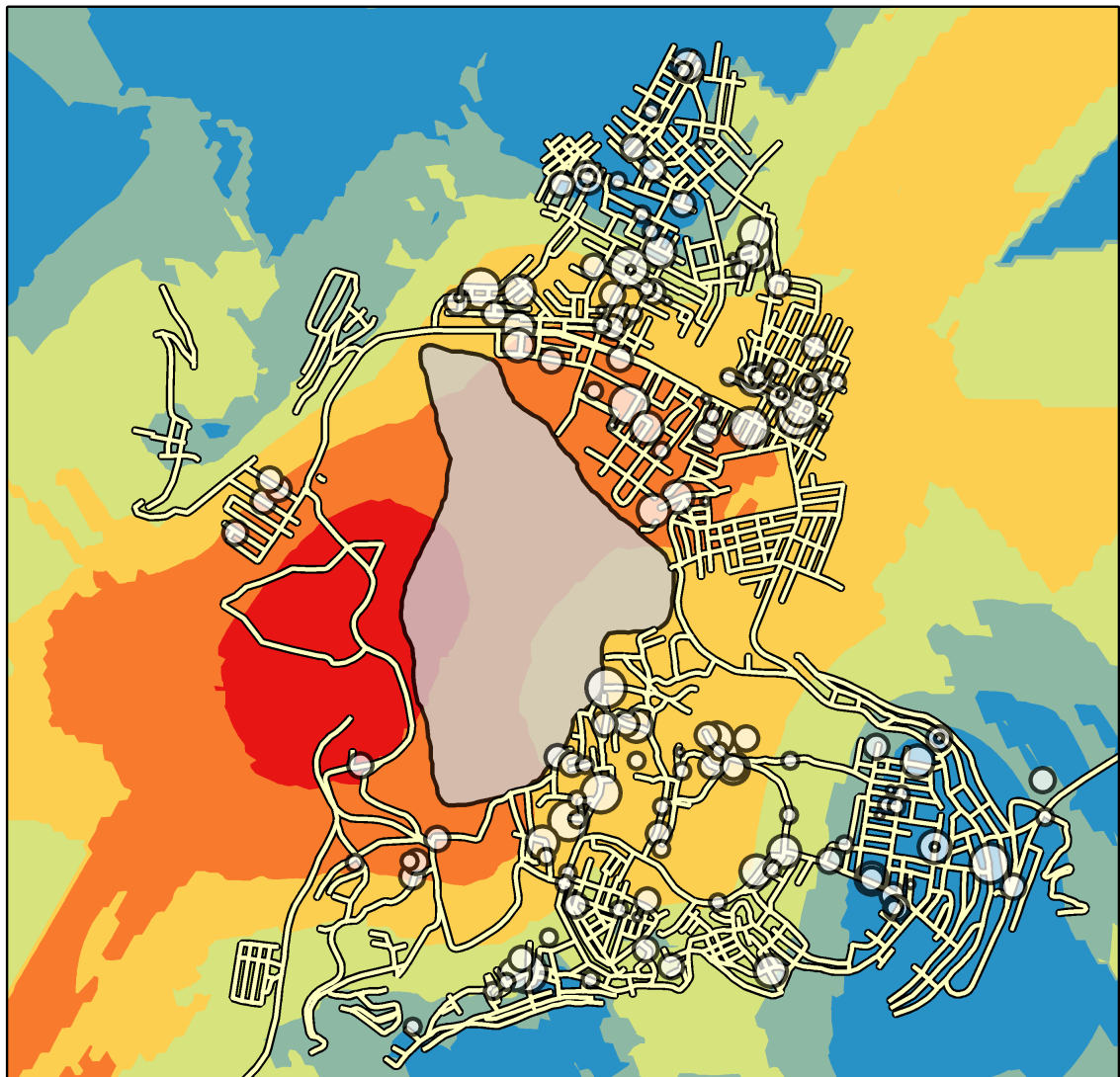
Figure 4.7. Interpolated soil lead values for participants from Cerro de Pasco. Soil lead values were interpolated for participants from Cerro de Pasco using the kriging option in ArcGIS.
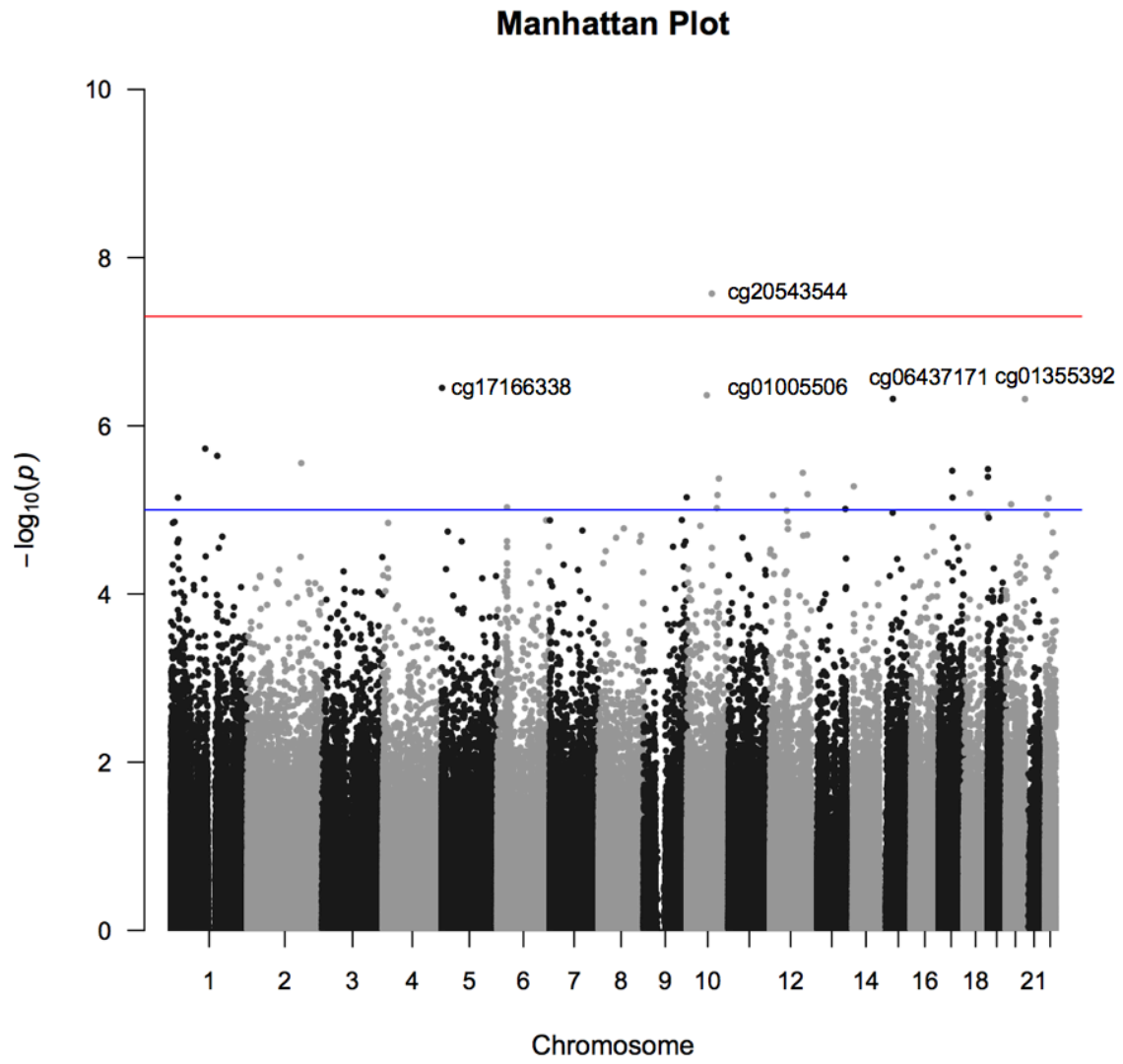
Figure 8. Manhattan plot with top hits from EWAS between WBLLs and DNA methylation. Significance levels are shown by lines, and the top hits from EWAS are indicated. Blue line indicates the significance level of q<0.10, and red line represents the significance level q<0.05

CHAPTER 5

Conclusion

This dissertation focused on the effects of high-altitude and lead exposure on the epigenomes of the Peruvian Quechua. We used a variety of epigenetic methods to test the following hypotheses: 1) Does high-altitude exposure at birth and during life affect DNA methylation signatures of adults? 2) Does Pb exposure in adulthood affect DNA methylation signatures of adults? We were interested in both questions due to the nature of one of the research sites where we conducted participant recruitment. High-altitude epigenetics work was a natural extension of decades of scholarly work in the Peruvian Andes focusing on the effects of hypoxia on physiology and genetics of the indigenous Andeans (Hurtado 1932a; Baker and Little 1976; Frisancho 1983). To interrogate this question, we recruited participants according to a migrant study design, i.e. individuals from three different exposure groups were recruited: 1) individuals born and raised at low altitude who were recruited in Lima, Peru; 2) individuals born at high altitude and moved to low altitude sometime during life who were recruited in Lima, Peru; 3) individuals born and raised at high altitude who were recruited in Cerro de Pasco, Peru.

Although it is well known that Andeans possess unique physiological and genetic adaptations to high-altitude conditions, it is not well understood whether epigenetic mechanisms play a role in this adaptation. However, epigenetic processes have been hypothesized to play a role in this adaptation process, due to their flexible nature and the ability to be affected by environmental change (Frisancho 2009).

The lead exposure question arose because Cerro de Pasco, our high-altitude recruitment locale, is a lead mining center of Peru and individuals living in Cerro de Pasco are known to have elevated lead levels (Conklin L., et al. 2008; van Geen, et al. 2012). It is known that lead exposure can affect the epigenome in young children and adolescents (Sen, et al. 2015; Goodrich, Dolinoy, Sanchez, et al. 2016), and here we assessed the effects on inorganic lead on the epigenomes of the Peruvian adults.

We were curious to study two seemingly disparate questions together, since it is known that the epigenome can be affected by various environmental exposures (Feil and Fraga 2011), and it was crucial for us be able to disentangle the effects of different environmental effects on the epigenome.

RESEARCH QUESTIONS

Question 1: Are there gene-specific changes association with high-altitude exposure in the Peruvian Quechua?

This question is inspired by the work of Roberto Frisancho and Paul Baker on developmental adaptation to high-altitude in the Andeans (Frisancho 1970; Frisancho and Baker 1970). Frisancho's work has shown that birth and development at high altitude are crucial for the attainment of the enlarged chest capacity characteristic of the Andean mode of high-altitude adaptation (Frisancho 1977). We extended the question of developmental adaptation to focus on epigenetics as a potential actor in mediating the effects of the high-altitude environment on the body. We hypothesized that high-altitude hypoxia, both during development and throughout life, would have an effect on the epigenome, since hypoxic conditions are known to affect epigenetics in case of cancer, which is often characterized by hypoxia (Wilson and Hay 2011; Rawluszko-Wieczorek, et al. 2014).

In Chapter 2 we analyzed global methylation of the LINE-1 repetitive element and the high-altitude adaptive gene *EPAS1*. We chose LINE-1 since it has been shown to reflect global genomic methylation (Irahara, et al. 2010). *EPAS1* was chosen because its polymorphisms are associated with high-altitude adaptation in Tibetans and potentially Andeans (Beall, et al. 2010; Eichstaedt, et al. 2017). Moreover, *EPAS1* is a hypoxia-inducible gene and its expression is known to increase in hypoxic conditions (Tian, et al. 1997). We targeted a region in the CpG island of *EPAS1* associated with its expression (Rawluszko-Wieczorek, et al. 2014).

We found higher LINE-1 methylation associated with birth at high altitude, as well as the altitude of recruitment. Surprisingly, we found decreased *EPAS1* methylation associated with current altitude, which may indicate increased *EPAS1* expression (Rawluszko-Wieczorek, et al. 2014). This finding is consistent with our current understanding of EPAS1 as a hypoxia-inducible gene (Tian, et al. 1997).

Our results presented in Chapter 2 show that early life as well as current high-altitude exposure can both have effects on the epigenome. This this a first study to report changes in LINE-1 and *EPAS1* methylation associated with high altitude.

Question 2: Are there genome-wide DNA methylation differences associated with high-altitude adaptation in the high-altitude Quechua?

To answer this question, we analyzed a subset of 87 samples with the Illumina MethylationEPIC bead chip covering more than 850k DNA methylation sites across the genome. We performed an epigenome-wide association study (EWAS), and presented our findings in Chapter 3. Based on the EWAS, we found one significant differentially methylated CpG site at q-value < 0.10 in the promoter region of the gene *NPY1R,* which has been associated with pulmonary hypertension and fat metabolism (Zukowska-Grojec, et al. 1987; Raposinho, et al. 2004; Kuo, et al. 2007; Crnkovic, et al. 2014). Moreover, we identified two differentially methylated regions associated with genes *PKLR*, *USP29, and Yap1* involved in hypoxia regulation suggesting

a potential role in high-altitude adaptation, as well as ten differentially methylated regions associated with pulmonary hypertension. We validated the differentially methylated region associated with *SOD3*, using quantitative pyrosequencing in the entire cohort. We found that exposure to high-altitude at birth is associated with increased epigenetic aging suggesting that hypoxia is acting as a cellular burden. Accelerated aging was also associated with increased hemoglobin levels suggesting a possible link with CMS.

Overall, our findings show that high altitude is a strong influence on the epigenome affecting various cellular pathways, including the hypoxia-inducible factor pathway. We also show DNA methylation changes in genes associated with pulmonary hypertension, which is common at high altitude and considered adaptive (Naeije and Vanderpool 2013).

Question 3: Are there genome-wide DNA methylation changes associated with inorganic lead exposure in Quechua exposed to inorganic lead from mining versus the unexposed? Are associations between whole-blood lead levels and hemoglobin levels, proximity to mining, and soil lead values?

We answer these questions in Chapter 4 which is focused on the effects of inorganic lead exposure on the body. One of our study sites, Cerro de Pasco is a mining town with high lead levels in the soil surrounding the mine (van Geen, et al. 2012). People living in Cerro de Pasco are known to have elevated lead levels (Conklin L., et al. 2008). We determined the relationship between whole-blood lead levels (WBLLs) and hemoglobin levels, proximity to mining, and DNA methylation in a subset of our participants with available lead data.

We found a positive association between WBLLs and hemoglobin levels in out Peruvian cohort. We performed a spatial analysis of lead exposure distribution, and saw a negative association between WBLLs and proximity to mining, with the lead levels going down further away from the mine. We interpolated soil lead values based on the already published data (van Geen, et al. 2012) using kriging in ArcGIS, and found a suggestive association between soil lead and

blood lead. We saw a positive relationship between lead levels and the number of days since the last menstruation for women. Lastly, we found significant differentially methylated positions and regions associated with genes linked to neurological function and disorders and heavy metal exposure and toxicity, which is significant since lead exposure has been linked to decreased neurological changes in children (Needleman and Leviton 1979; Ahamed, et al. 2008).

Our findings show that low levels of lead exposure can have significant effects, including changes in hemoglobin levels and epigenetic changes in genes associated with neurological function and heavy metal exposure.

SIGNIFICANCE OF RESEARCH

This research is significant for four reasons. First, no study to date has researched the epigenetic signatures of Andean high-altitude adaptation. Without knowledge of the epigenetic modifications that likely contribute to hypoxia-adaptive phenotypes, it is challenging to fully understand the process of high-altitude adaptation. And determining the epigenetic changes associated with adaptation to hypoxia will aid our understanding of the relationship between genetic and epigenetic markers as they contribute to the high-altitude phenotypes. By investigating this interaction between genetic and epigenetic changes, we will be able to connect genetic adaptation with developmental adaptation in high-altitude Andeans. More broadly, our research will contribute to the growing understanding of how genetics influences epigenetic modifications, and vice versa.

Second, more than 140 million people live at high altitude (Moore, 2001). High-altitude environments are challenging for humans, and some of the most drastic conditions associated with being exposed to it manifest early in human development (e.g., intrauterine growth restriction (IUGR), preeclampsia). In populations with a long history of living at high altitude, such as the Andeans, the effects of low oxygen on fetal development are not nearly as dramatic as in populations with a shorter history of living above 2,500 m. Long-term high-altitude occupants have

genetic mutations that protect them from altitude-associated IUGR and preeclampsia. However, the epigenetic mechanisms that may play a role in the development of these conditions are currently unknown. Determining epigenetic changes that help humans adapt to high altitude will aid our understanding of how low-oxygen availability affects fetal growth and development for treatable public health issues such as fetal growth restriction and preeclampsia, and knowing more about role of hypoxia in these conditions will help us find ways to alleviate their effects.

Third, this study will investigate the effects of lead exposure at high altitude. The influence of lead on human health and the epigenome has been extensively studied, however, its exposure has not been well-characterized in high-altitude populations. This is significant because lead exposure is known to decrease hemoglobin and red blood cell levels in humans (Goyer 1973; Hegazy et al. 2010), both of which are important players in body's response to chronic hypoxia (Moore 2001; Heinicke et al., 2003). People of Cerro de Pasco are chronically exposed to Pb, as well as other contaminants, originating from mining. This research will determine whether the residents of Cerro de Pasco have lasting DNA methylation differences that are associated with their whole-blood lead exposure and proximity to mining.

Fourth, the results of this study will be shared with the people of Cerro de Pasco and with certain non-governmental organizations (NGOs). Brochures with the main findings will be translated into Spanish and donated to the public library of Cerro de Pasco. And the results of this study will be shared with environmental justice NGOs working in Peru, such as CooperAccion and Source International, to aid their efforts with relocating the residents of Cerro de Pasco. The Peruvian Ministry of Health declared "state of environmental emergency" in Cerro de Pasco (Ministerial Resolution N° 094-2013-MINAM) in May 2012. However, due to weak mining regulations and the lack of resources, little has been done to alleviate the health burden of Pb exposure. Our findings will show whether there is a correlation between the proximity to the mine, and WBLLs and DNA methylation. This data will aid the efforts of the Peruvian government with

relocating the people of Cerro de Pasco and imposing fines on the mining company working in Cerro de Pasco.

CONCLUDING REMARKS AND FUTURE DIRECTIONS

In this dissertation, we analyzed the effects of environmental exposures on the epigenome. The epigenome can be simultaneously influenced by many exposures, including but not limited to diet, stress, toxicants, and others (Goodrich, Dolinoy, Sanchez, et al. 2016; Non, et al. 2016). We were interested in identifying associations between DNA methylation and hypoxia exposure at birth and during recruitment, as well recent inorganic lead exposure (Figure 5.1).
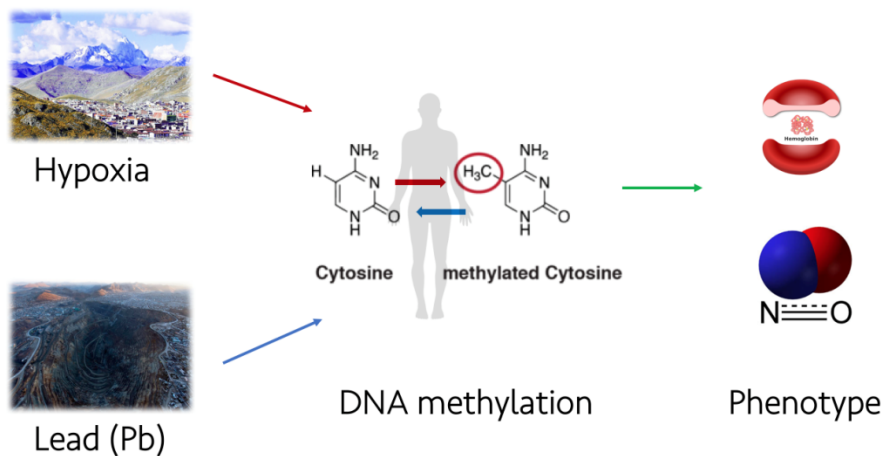
Figure 5.1. Study design.

This dissertation was first to study the role of the epigenome in high-altitude adaptation and protection against altitude-associated conditions in high-altitude Quechua. Developmental adaptation has been considered an important mode of adaptation among high-altitude Andeans, and we determined the role of epigenetics in this process. Based on this research, several conclusions and new questions emerged:

First, both the altitude at birth and the current altitude affect the epigenome, but it remains unclear whether they have an adaptive function. The role of epigenetics in adaptation is just

beginning to be explored. Continued work will be necessary to determine whether any of the epigenetic changes we found are associated with changes in gene expression, protein levels, and specific pathways. Moreover, epigenetic studies in other altitude native populations are necessary to determine to what degree the epigenome plays a role in high-adaptation globally.

Second, we show that small levels of lead exposure are associated with epigenetic changes in adults. We determined lead levels in whole blood, but bone is the main storage site for inorganic lead and better reflect lifetime exposure (Barbosa, et al. 2005). Therefore, we were unable to separate the effects of lead exposure at birth and throughout life from current lead exposure. This is especially critical for the participants from Cerro de Pasco, since children born there are known to have elevated lead levels (Conklin L., et al. 2008). Future work using bone lead measurements in adults and their association with epigenetic changes is needed to render a detailed understanding of the effects of lead exposure on epigenetics.

Third, it is still unclear to what degree small epigenetic changes can affect gene expression. Small-magnitude effect sizes are often reported in epigenetic studies (Breton, et al. 2017), and more research is needed to understand the true biological significance of small effect sizes in epigenetic research.

There are several limitations in this dissertation study: use of whole-blood for DNA methylation analysis, relatively small sample size for EWAS, and confounding between lead exposure and high-altitude exposure. We used bioinformatic tools to correct for different cell types in the EWAS analysis, however, the pyrosequencing results could not be corrected for cell types. Our sample size was relatively limited for an EWAS, allowing us to only identify changes with large effect sizes. A larger study is needed to find smaller effect size changes. It is unclear to what degree high-altitude hypoxia exposure and lead exposure are interacting, and one of our study groups, Cerro de Pasco Quechua, is characterized by the longest lifetime hypoxia exposure and lead exposure, which may confound each other.

Overall, what we have shown is that both altitude and lead are affecting DNA methylation at single sites and regionally. We identified changes in the HIF pathway genes, genes associated with pulmonary hypertension, neurological development and function.

APPENDICES

Appendix A – Supplementary Figures and Tables

Supplementary Table 2.1. Supplementary Table 2.1. EPAS1 and LINE-1 methylation differences between males and females

CpG	Males	Females	β	p-value
LINE-1				
Average	70.33 ± 1.43	69.64 ± 1.40	0.67	4.37E-08
1	70.32 ± 2.68	69.71 ± 2.65	0.48	0.007
2	74.12 ± 1.69	73.45 ±1.62	0.63	2.15E-07
3	70.02 ± 2.55	68.99 ± 2.55	1.03	3.55E-09
4	67.03 ± 2.10	66.56 ± 2.22	0.53	0.000
EPAS1				
Average	13.80 ± 2.97	15.06 ± 3.03	1.13	2.47E-05
1	16.32 ± 3.38	17.73 ± 3.40	1.27	3.56E-05
2	10.35 ± 2.76	11.42 ± 2.87	0.88	4.70E-04
3	14.07 ± 3.29	15.27 ± 3.28	1.06	4.02E-04
4	14.22 ± 3.16	15.92 ± 3.33	1.56	0.000

Data are means ± SD
All models have been adjusted for a

Supplementary Table 2.2. One-carbon metabolism SNPs that were tested

Gene	SNP rs id	chr	pos
MTRR	rs10380	5	7897078
MAT2B	rs10515861	5	163503741
MAT2A	rs1078004	2	85542588
CTH	rs10889869	1	70412281
GLDC	rs10975681	9	6606746
ALDH1L1-AS2	rs13082246	3	126192112
MTRR	rs1532268	5	7878066
AHCYL2	rs1665105	7	129429137
MTHFR	rs1801131	1	11794419
MTHFR	rs1801133	1	11796321
TCN2	rs1801198	22	30615623
FASTKD3, MTRR	rs1801394	5	7870860
MTR	rs1805087	1	236885200
SHMT1	rs1979277	17	18328782
FOLH1	rs202676	11	49206068
MTHFD1	rs2236225	14	64442127
MTHFR	rs2274976	1	11790870
FOLR2	rs2298444	11	72221370
ATIC	rs2372536	2	215325297
DNMT3B	rs2424922	20	32798643
DMGDH	rs28326	5	79000031
TYMS	rs2853533	18	658064
SLC25A32	rs3134297	8	103415350
BHMT	rs3733890	5	79126136
ENOSF1	rs3744962	18	674320
ATIC	rs3821353	2	215333056
FPGS, LOC102723566	rs4451422	9	127814318
TYMS	rs502396	18	659236
TCN1	rs526934	11	59866020
DMGDH	rs532964	5	79044463
TCN1	rs557564	11	59861832
DNMT3A	rs6546045	2	25264789
CBS	rs6586282	21	43058387
MTHFS	rs7177659	15	79865700
ATIC	rs7563206	2	215325931
DNMT3A	rs7581217	2	25302075
MTRR	rs7730643	5	7875850

MTHFD1L	rs7765521	6	150883407
MTHFD1L	rs803447	6	150943642
MTHFD1L	rs803455	6	150926303
GCSH	rs8177940	16	81083789
ALDH2	rs886205	12	111766623
GART	rs8971	21	33511311
TCN2	rs9621049	22	30617432
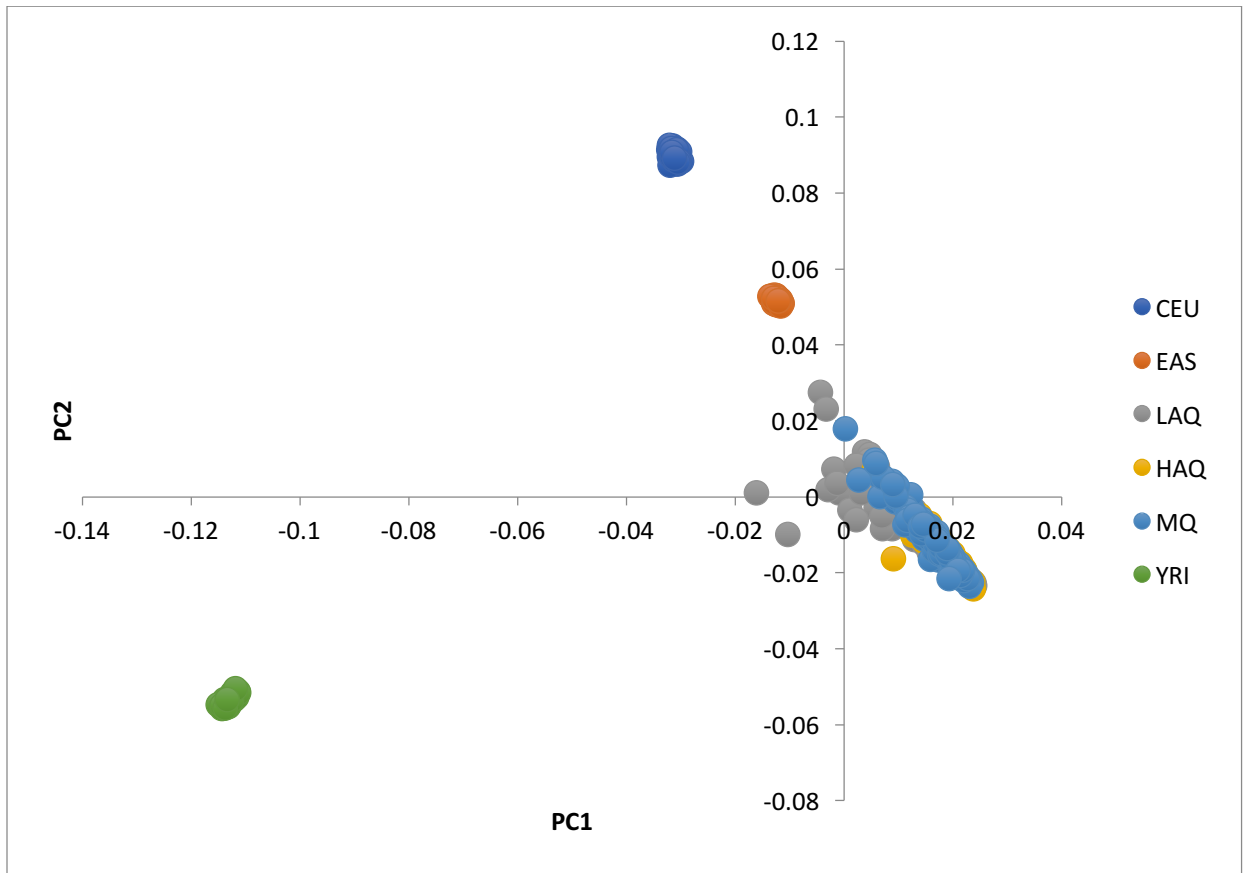
Supplementary Figure 2.1A. Principal analysis of genome-wide SNP data from the Peruvians in this study (LAQ, MQ, HAQ), and CEU Europeans, EAS East Asians, and YRI Yorubans from HapMap. Principal components 1 and 2 are plotted against each other.
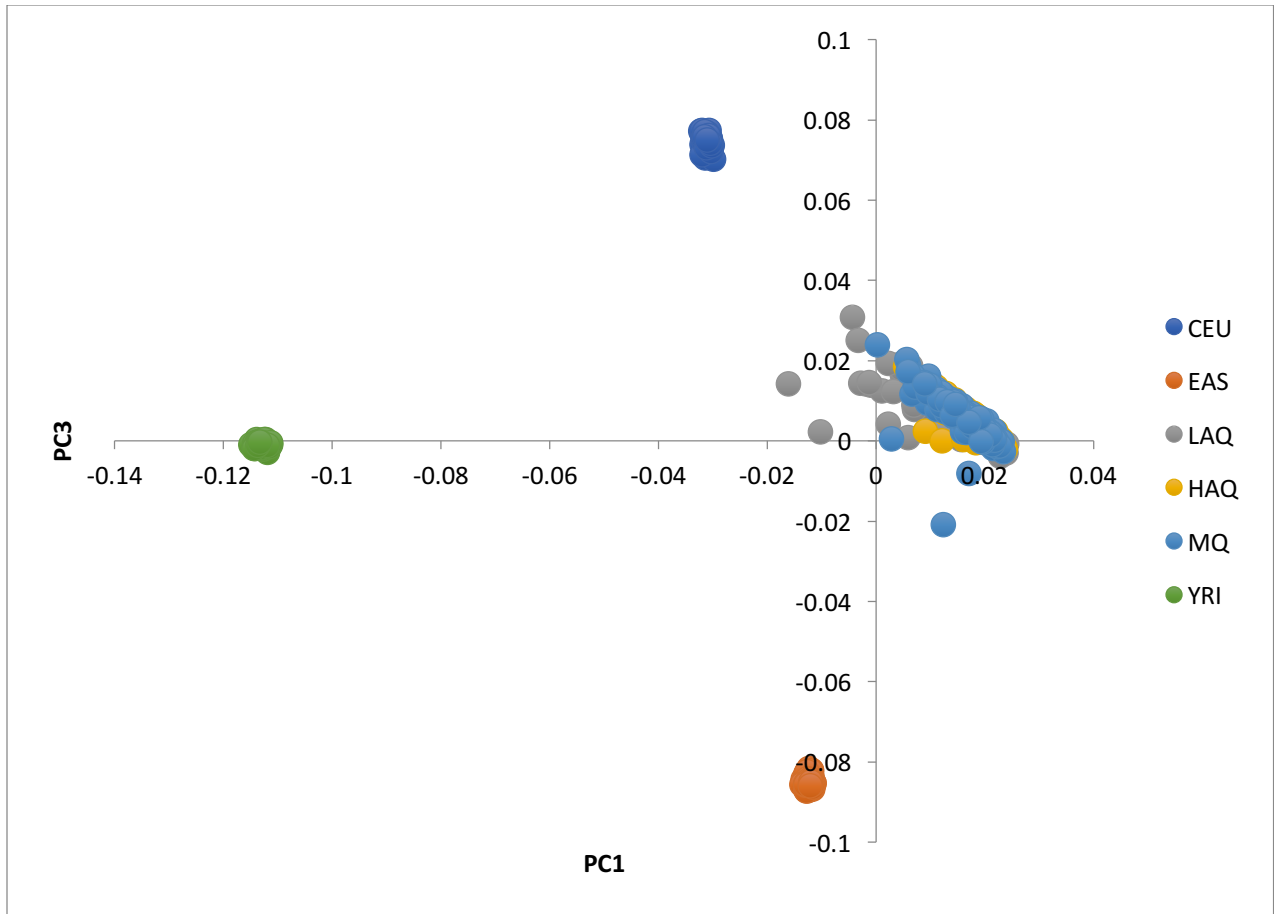
Supplementary Figure 2.1B. Principal analysis of genome-wide SNP data from the Peruvians in this study (LAQ, MQ, HAQ), and CEU Europeans, EAS East Asians, and YRI Yorubans from HapMap. Principal components 1 and 3 are plotted against each other.
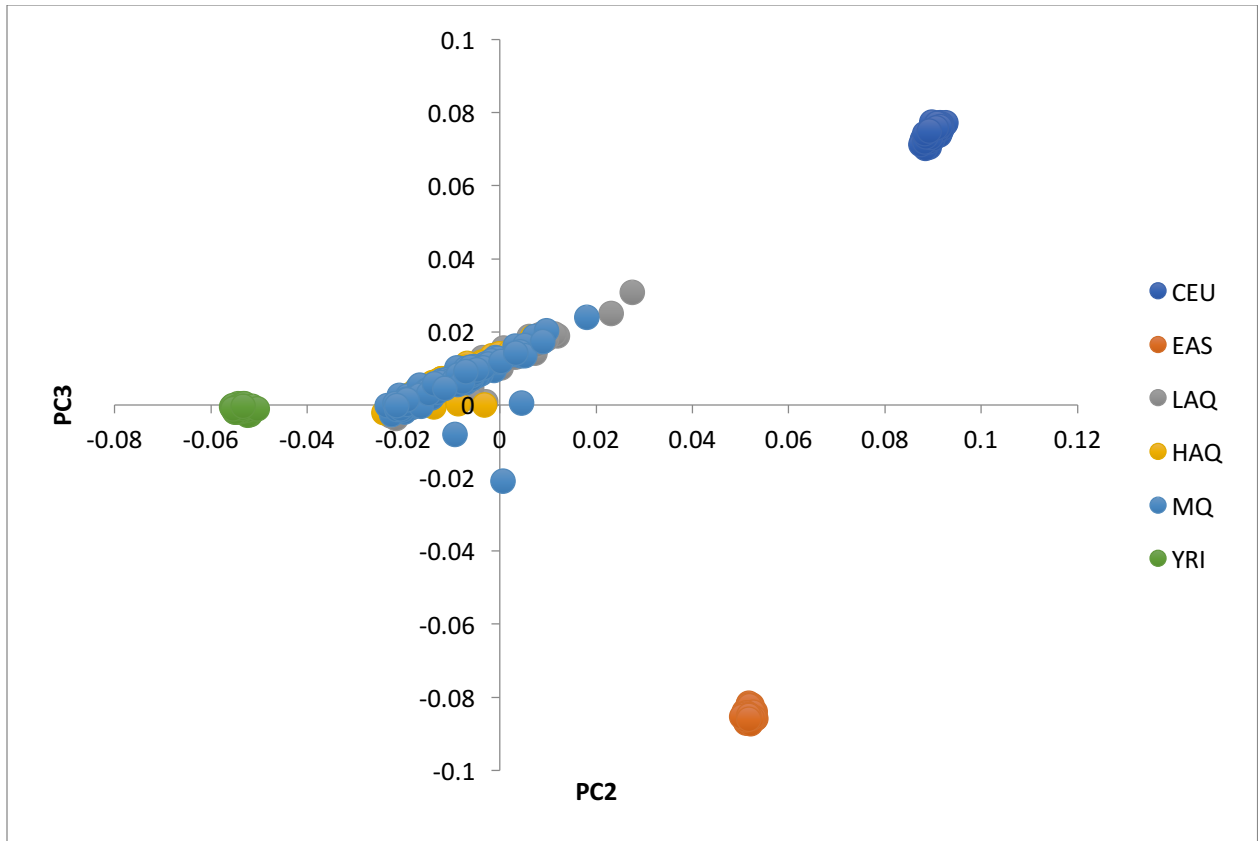
Supplementary Figure 2.1C. Principal analysis of genome-wide SNP data from the Peruvians in this study (LAQ, MQ, HAQ), and CEU Europeans, EAS East Asians, and YRI Yorubans from HapMap. Principal components 2 and 3 are plotted against each other.
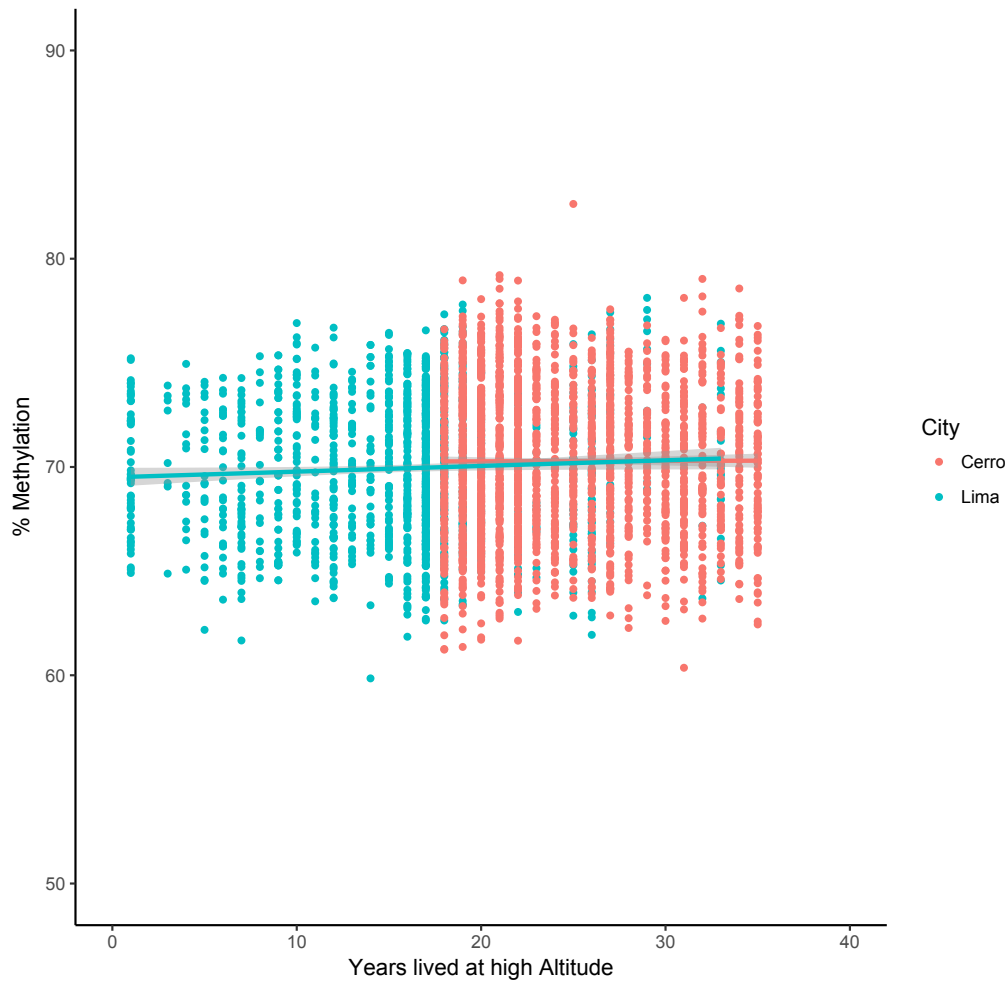
Supplementary Figure 2.2. Scatter plot of DNA methylation by years lived at high altitude in participants born at high altitude [HAQ and MQ]. Years lived at high altitude are positively associated with LINE-1 methylation for HAQ and MQ combined (p-value=0.0047), but not for HAQ alone (p-value=0.72). The association between years lived at high altitude and LINE-1 methylation is approaching significance for MQ alone (p-value=0.073).
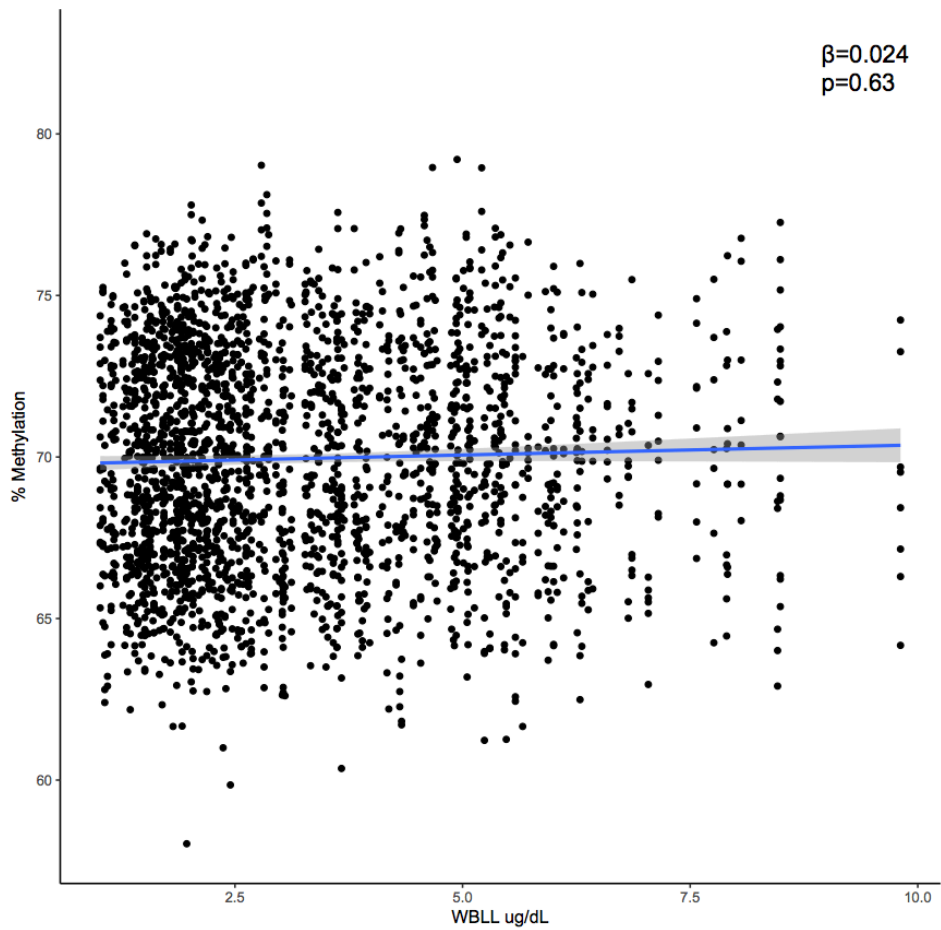
Supplementary Figure 2.3. Association between LINE-1 methylation and whole-blood lead values for a subset of participants with lead data available (n=294). The association is not significant (β=0.024, p-value=0.63).
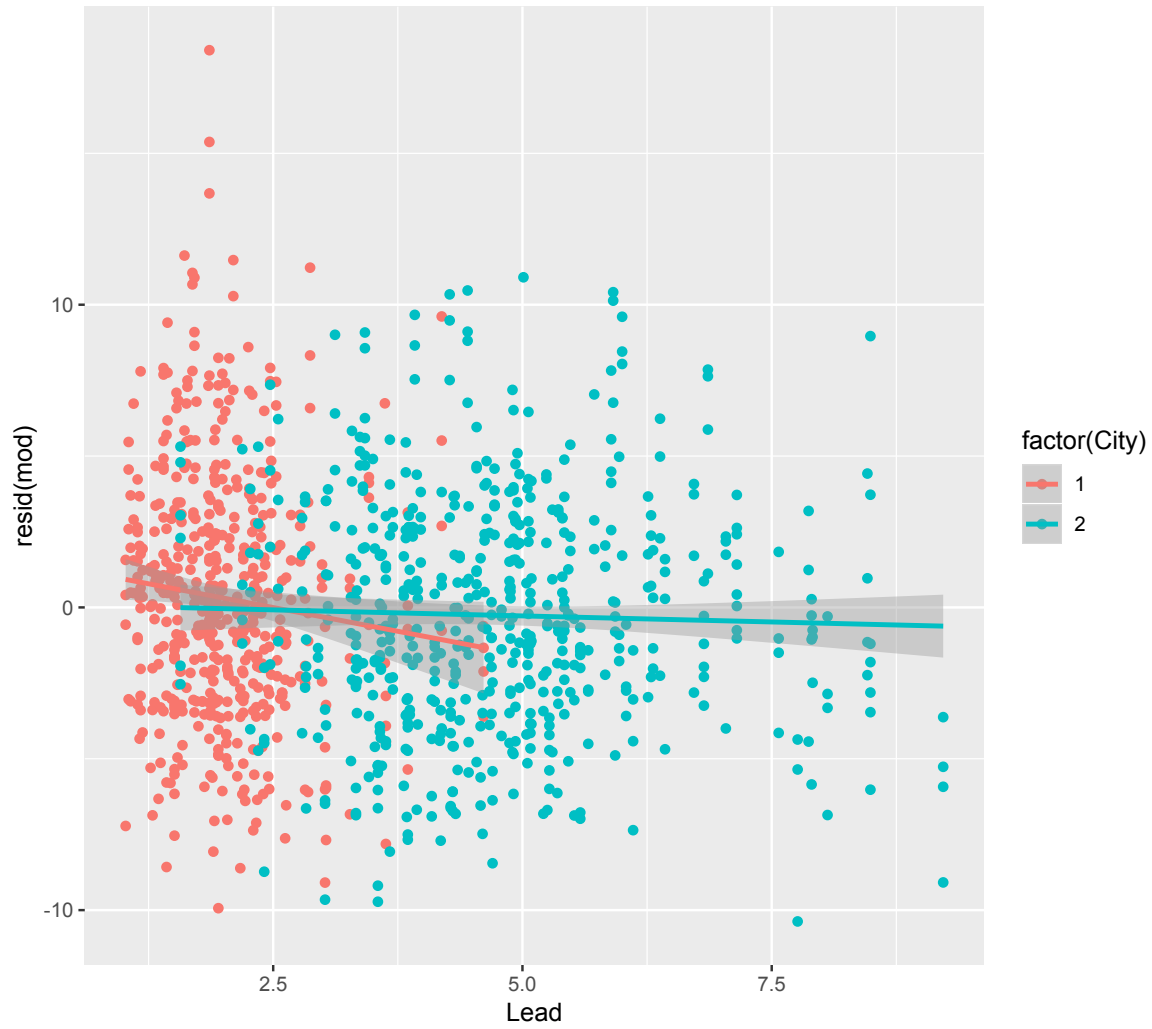
Supplementary figure 2.4. Residuals of the EPAS1 model adjusted for age, sex, BMI, and genome-wide PC1 plotted against WBLLs.

Supplementary Table 3.1. All significant DMRs

Gene	Region	Δ Meth	Island	N sites	P-value region	Corrected p-value**
TNXB; ATF6B	chr6:32063501-32065703	+	YES	19	9.22E-08	2.75E-05
SOX11	chr2:5831147-5836367	+	YES	15	1.19E-10	1.49E-08
MACROD2; SEL1L2	chr20:13975279-13977220	+	YES	14	6.41E-11	2.17E-08
S100A13	chr1:153599479-153600598	+	NO	14	4.14E-08	2.43E-05
UNC5C	chr4:96469286-96471144	+	YES	13	3.00E-07	1.06E-04
SCAND3*	chr6:28583655-28584465	+	YES	12	5.47E-13	4.43E-10
AGBL4	chr1:50489240-50490707	+	YES	12	2.48E-07	1.11E-04
CDH2	chr18:25754857-25758435	+	YES	11	3.07E-09	5.63E-07
PHYHIPL	chr10:60935910-60937258	+	YES	11	1.10E-07	5.33E-05
CDH8	chr16:62069317-62071194	+	YES	11	2.14E-06	7.48E-04
TBX18	chr6:85472949-85474804	+	YES	10	3.22E-10	1.14E-07
SCAND3*	chr6:28601271-28602640	+	NO	10	5.74E-10	2.75E-07
RFPL2	chr22:32598479-32601343	+	NO	10	1.53E-09	3.51E-07
CCDC105;SLC1A6	chr19:15121199-15122225	+	YES	10	1.34E-09	8.59E-07
DPP6	chr7:153582929-153585369	+	YES	10	1.94E-08	5.21E-06
LGALS8	chr1:236685524-236687899	+	YES	10	6.12E-08	1.69E-05
SLC38A11	chr2:165811765-165812479	+	NO	10	4.83E-08	4.43E-05
EPHA5	chr4:66535145-66536773	+	YES	9	1.78E-06	7.17E-04
LRIG3	chr12:59313254-59315040	+	YES	9	2.86E-05	1.04E-02
NAPRT1	chr8:144659492-144661052	+	YES	9	1.61E-10	6.78E-08
FAM196A; DOCK1	chr10:128993810-128995479	+	YES	9	1.69E-09	6.65E-07
GATM	chr15:45669010-45671709	+	YES	9	3.46E-09	8.41E-07
LOC145757	chr15:101389272-101390351	+	YES	9	1.61E-08	9.76E-06
DUSP22	chr6:291687-292824	+	YES	9	3.71E-08	2.14E-05
KCNN2	chr5:113697487-113698848	+	YES	9	4.69E-07	2.26E-04
FAM84A	chr2:14773973-14776097	+	YES	9	7.46E-07	2.30E-04
MROH6	chr8:144653855-144655680	+	YES	9	1.01E-06	3.62E-04
GRIK1	chr21:31311797-31312906	+	YES	9	1.65E-06	9.73E-04
ISL1	chr5:50678451-50679966	+	YES	9	2.63E-06	1.14E-03

Gene	Location	Strand		Score	p-value	q-value
TCRVB	chr7:142493865-142495099	+	YES	9	2.85E-06	1.51E-03
ZNF154; ZNF776	chr19:58220080-58221623	+	YES	9	3.72E-06	1.58E-03
SPAG6	chr10:22633916-22635144	+	YES	9	6.17E-06	3.29E-03
RYR3	chr15:33601870-33603961	+	YES	8	1.49E-08	4.68E-06
AVPR1A	chr12:63543831-63545289	+	YES	8	3.08E-07	1.39E-04
TJP1	chr15:30114291-30115300	+	YES	8	2.24E-07	1.46E-04
WNT2	chr7:116962950-116964473	+	YES	8	2.94E-06	1.26E-03
SRRM4	chr12:119418004-119419787	+	YES	8	6.38E-07	2.35E-04
PCDH7	chr4:30720877-30723984	+	YES	8	6.41E-08	1.35E-05
EPHA6	chr3:96532832-96533825	+	YES	8	2.23E-08	1.48E-05
DKK1; PRKG1-AS1	chr10:54073642-54075529	+	YES	8	2.66E-07	9.24E-05
UNC5D	chr8:35092188-35094054	+	YES	8	3.59E-07	1.26E-04
KITLG	chr12:88973398-88975429	+	YES	8	5.23E-07	1.69E-04
GRM7	chr3:6902100-6903922	+	YES	8	1.71E-06	6.17E-04
GALNTL6	chr4:172733300-172734844	+	YES	8	1.51E-06	6.40E-04
VWA7	chr6:31733434-31735475	+	NO	8	2.26E-06	7.27E-04
ADAM12	chr10:128076260-128078011	+	YES	8	4.96E-06	1.86E-03
ARPP21	chr3:35680290-35681884	+	YES	8	1.16E-05	4.76E-03
FOXG1	chr14:29234981-29236536	+	YES	8	1.40E-05	5.89E-03
FEZF1	chr7:121944755-121946609	+	YES	8	1.87E-05	6.60E-03
GPR19	chr12:12848977-12850083	-	YES	8	1.52E-06	8.99E-04
IRS2	chr13:110437759-110440180	+	YES	7	1.21E-07	3.28E-05
LYPD1	chr2:133426653-133428653	+	YES	7	2.45E-07	8.04E-05
CTNND2	chr5:11903651-11904913	+	YES	7	9.88E-10	5.14E-07
ALOX12P2	chr17:6796745-6797772	+	YES	7	1.29E-09	8.27E-07
TBX18*	chr6:85482570-85484718	+	YES	7	7.32E-09	2.23E-06
CRTAC1	chr10:99734081-99735203	+	NO	7	3.93E-09	2.30E-06
ZSCAN1	chr19:58545001-58546308	+	NO	7	1.44E-08	7.24E-06
ST6GAL2	chr2:107502615-107504170	+	NO	7	2.22E-08	9.35E-06
LOC643802	chr16:53406197-53407595	+	YES	7	8.38E-08	3.93E-05
CBLN4	chr20:54579355-54581399	+	YES	7	2.52E-07	8.08E-05

Gene	Location	Strand		Count	p-value	FDR
CADPS	chr3:62860103-62861312	+	YES	7	2.95E-07	1.60E-04
EFCAB1	chr8:49647579-49648520	+	YES	7	6.61E-07	4.61E-04
STK32C	chr10:134043362-134043898	+	NO	7	1.08E-06	1.33E-03
RANBP17	chr5:170288606-170289849	+	YES	7	3.19E-06	1.68E-03
ELAVL2	chr9:23821428-23822391	+	YES	7	2.51E-06	1.71E-03
PAX3; CCDC140	chr2:223163588-223166095	+	YES	7	1.22E-05	3.19E-03
BHMT	chr5:78406928-78408254	+	NO	7	1.77E-05	8.74E-03
FBXO39; XAF1	chr17:6678929-6679565	+	YES	7	1.02E-05	1.04E-02
KHDRBS2	chr6:62995523-62996698	+	YES	7	2.14E-05	1.19E-02
ZNF135	chr19:58570077-58571639	+	NO	7	0.00012	4.92E-02
PKLR	chr1:155264372-155265379	+	YES	6	3.89E-10	2.53E-07
HTR1B	chr6:78172337-78174287	+	YES	6	1.07E-07	3.60E-05
SRD5A2	chr2:31805218-31806899	+	YES	6	1.49E-06	5.82E-04
PENK	chr8:57358165-57359415	+	YES	6	1.80E-06	9.42E-04
RFX6	chr6:117197545-117199068	+	YES	6	4.23E-05	1.81E-02
RBFOX1	chr16:6532643-6534318	+	YES	6	2.64E-09	1.03E-06
LPPR1	chr9:103790851-103792173	+	YES	6	1.04E-08	5.14E-06
TRIOBP	chr22:38091825-38093080	+	NO	6	1.11E-08	5.82E-06
ANKRD20A19P	chr13:24519920-24520508	+	NO	6	5.87E-09	6.55E-06
SYT9	chr11:7271955-7274235	+	YES	6	6.66E-08	1.92E-05
NRG1	chr8:32404691-32406151	+	YES	6	1.04E-07	4.68E-05
NEFM	chr8:24771466-24772514	+	YES	6	2.92E-07	1.83E-04
RGMA; CHD2*	chr15:93580022-93580847	+	NO	6	2.59E-07	2.06E-04
ZNF835	chr19:57182127-57183493	+	YES	6	5.62E-07	2.70E-04
IRX2	chr5:2751041-2752546	+	YES	6	1.03E-06	4.49E-04
HOOK1	chr1:60280088-60281078	+	YES	6	1.79E-06	1.19E-03
WBSCR17	chr7:70597058-70598283	+	YES	6	3.19E-06	1.71E-03
CCKBR	chr11:6291339-6292897	+	YES	6	5.04E-06	2.12E-03
PDZRN3	chr3:73673231-73674364	+	YES	6	5.26E-06	3.04E-03
GRM1	chr6:146350041-146351045	+	YES	6	7.23E-06	4.71E-03
KCNJ2	chr17:68164468-68166022	+	YES	6	1.51E-05	6.37E-03

Gene	Location	Strand		Count	p-value	q-value
MDGA2	chr14:48143618-48144766	+	YES	6	1.29E-05	7.32E-03
ZIC2	chr13:100636183-100637561	+	YES	6	1.86E-05	8.83E-03
ADRA1A	chr8:26722151-26723366	+	YES	6	2.50E-05	1.34E-02
POU2AF1	chr11:111250093-111250745	+	NO	6	2.61E-05	2.59E-02
TENM3	chr4:183728196-183729462	-	YES	6	7.19E-08	3.73E-05
SLIT3	chr5:168727088-168728271	+	YES	5	1.22E-06	6.77E-04
SOD3	chr4:24796514-24797177	+	NO	5	1.51E-06	1.50E-03
ALDH1A2	chr15:58356761-58357892	+	YES	5	8.97E-06	5.19E-03
DSC3	chr18:28621985-28623391	+	YES	5	4.17E-05	1.93E-02
RSPH3*	chr6:159359940-159360985	+	YES	5	1.11E-09	6.99E-07
TTYH1	chr19:54945959-54947257	+	NO	5	4.49E-09	2.27E-06
SIX3*	chr2:45155408-45156938	+	YES	5	8.72E-08	3.74E-05
PRMT8*	chr12:3475063-3476273	+	YES	5	1.20E-07	6.52E-05
NA	chr5:101119084-101119767	+	NO	5	1.54E-07	1.48E-04
MIR124-2	chr8:65289139-65290849	+	YES	5	4.77E-07	1.83E-04
ZSCAN23*	chr6:28446794-28447116	+	NO	5	1.49E-07	3.04E-04
ZNF415	chr19:53635327-53637179	+	YES	5	9.31E-07	3.30E-04
FAAH	chr1:46859671-46860512	+	YES	5	4.48E-07	3.49E-04
ADAM33	chr20:3661910-3663187	+	YES	5	8.87E-07	4.56E-04
CDH20	chr18:59221320-59222051	+	YES	5	5.16E-07	4.63E-04
RHCG	chr15:90029474-90030680	+	YES	5	2.32E-06	1.26E-03
ALDH3A1	chr17:19627735-19628975	+	YES	5	4.72E-06	2.49E-03
VWC2	chr7:49813031-49813764	+	YES	5	3.38E-06	3.02E-03
ZBTB16	chr11:113933513-113934356	+	YES	5	4.01E-06	3.12E-03
ESRRG	chr1:217309906-217311235	+	YES	5	9.20E-06	4.53E-03
CPXM2*	chr10:125731617-125732843	+	YES	5	1.25E-05	6.68E-03
CTNNA2; LRRTM1	chr2:80529547-80530949	+	YES	5	2.27E-05	1.06E-02
GALNT15	chr3:16215248-16217128	-	NO	5	4.38E-05	1.52E-02
KCNA5	chr12:5152380-5153289	+	YES	4	2.09E-06	1.51E-03
RBPJL	chr20:43944741-43945736	+	YES	4	3.87E-08	2.55E-05
COL20A1	chr20:61953801-61954988	+	YES	4	8.50E-08	4.70E-05

Gene	Location	Strand		Count		
FAM181B	chr11:82443961-82445171	+	YES	4	3.43E-07	1.86E-04
NA	chr5:72598865-72599787	+	YES	4	1.05E-06	7.48E-04
FGF12	chr3:192126825-192127992	+	YES	4	1.79E-06	1.01E-03
ZNF541	chr19:48048129-48049234	+	YES	4	2.16E-06	1.28E-03
CNTN4	chr3:2140256-2142142	+	YES	4	4.43E-06	1.54E-03
GJA3*	chr13:20711042-20711638	+	YES	4	2.19E-06	2.41E-03
USP29*	chr19:57617602-57618437	+	YES	4	4.04E-06	3.17E-03
YAP1*	chr11:102139342-102140602	+	YES	4	1.12E-05	5.81E-03
KLHL23; PHOSPHO2-KLHL23	chr2:170589808-170590491	+	YES	4	7.42E-06	7.10E-03
ZIC2*	chr13:100649800-100650848	+	YES	4	1.65E-05	1.03E-02
BMP4	chr14:54412780-54413932	+	YES	4	1.82E-05	1.03E-02
TRPM3	chr9:74061323-74062097	+	NO	4	2.09E-05	1.75E-02
GLB1L; STK16	chr2:220107847-220108408	+	YES	4	2.48E-05	2.86E-02
PITX2	chr4:111542444-111543402	+	YES	4	6.26E-05	4.19E-02
AGAP1	chr2:236923322-236924311	-	NO	4	9.26E-08	6.14E-05
CRIP2	chr14:105943043-105944656	-	NO	4	6.97E-05	2.79E-02
TPRG1	chr3:188850137-188851244	+	NO	3	1.32E-08	7.81E-06
IGLON5	chr19:51815359-51815659	+	YES	3	7.77E-09	1.70E-05
EBF4	chr20:2730191-2731177	+	YES	3	8.07E-08	5.37E-05
NA	chr6:26323437-26323587	+	NO	3	1.58E-07	6.90E-04
EPHA10	chr1:38200920-38201819	+	YES	3	2.25E-06	1.64E-03
SDK1*	chr7:4456202-4456404	+	NO	3	6.89E-07	2.24E-03
SLC35F3	chr1:234367145-234367494	+	YES	3	1.40E-06	2.63E-03
TMEM225B	chr7:99195655-99195932	+	YES	3	1.54E-06	3.63E-03
XKR4	chr8:56014406-56015577	+	YES	3	1.70E-05	9.50E-03
C1QL4	chr12:49726500-49727111	+	YES	3	9.28E-06	9.92E-03
NPTX2	chr7:98246006-98247143	+	YES	3	1.75E-05	1.00E-02
PPP1R13L	chr19:45884900-45886079	+	NO	3	1.88E-05	1.04E-02
GPR35	chr2:241563347-241564625	+	NO	3	2.72E-05	1.39E-02
TTBK1	chr6:43214396-43214851	+	YES	3	1.52E-05	2.17E-02
LPA/PLG*	chr6:161100122-161100673	+	YES	3	2.24E-05	2.63E-02

FAM189A2	chr9:71940287-71940579	+	YES	3	1.66E-05	3.65E-02
ZNF764	chr16:30572739-30573014	+	YES	3	1.73E-05	4.05E-02

* Gene within 50 kb of the CpG site
** Sidak corrected p-value
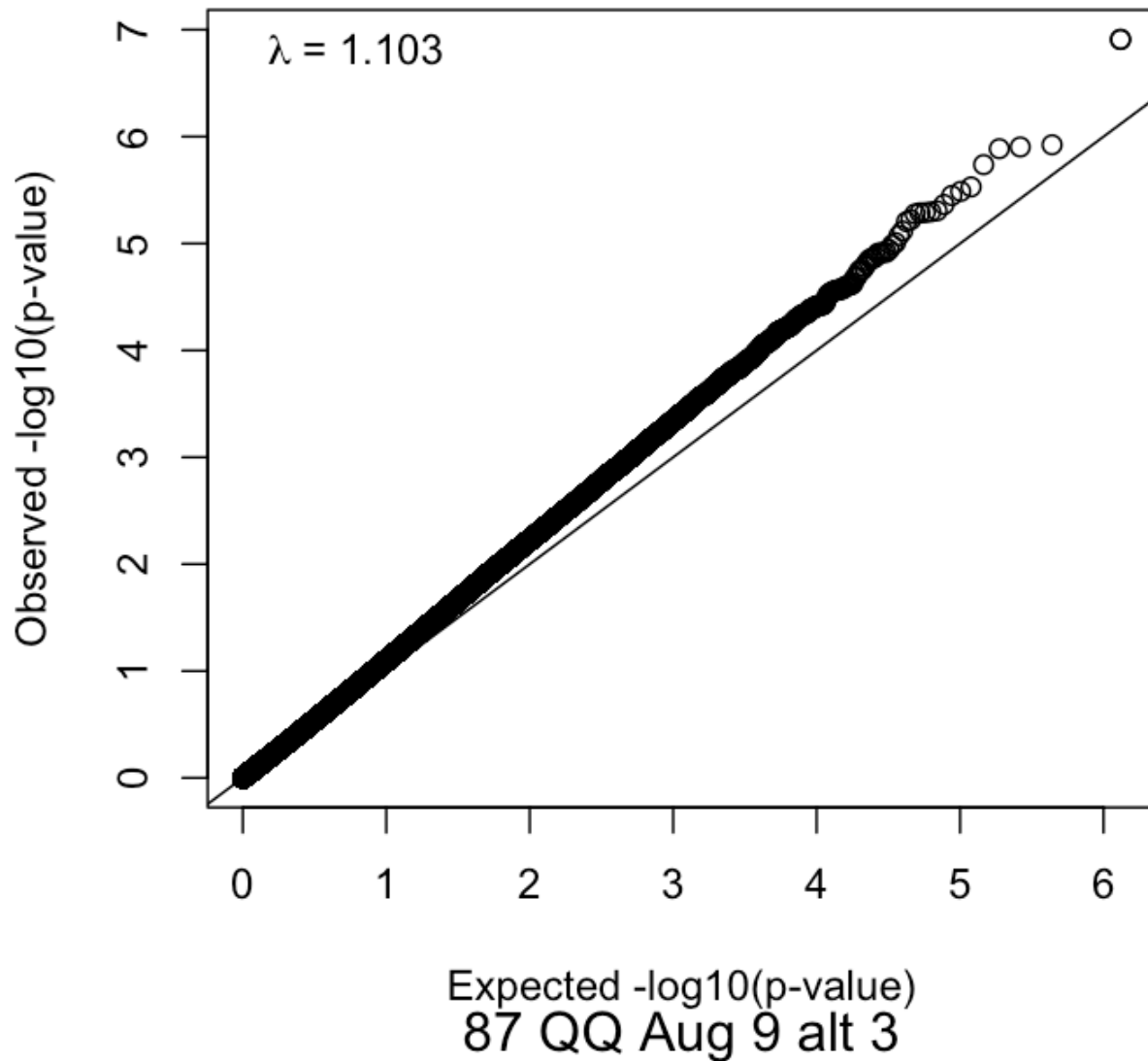
Supplementary Figure 3.1. Genome-wide quantile-quantile plot based on expected and observed p-values from the analysis of differentially methylated positions using limma.
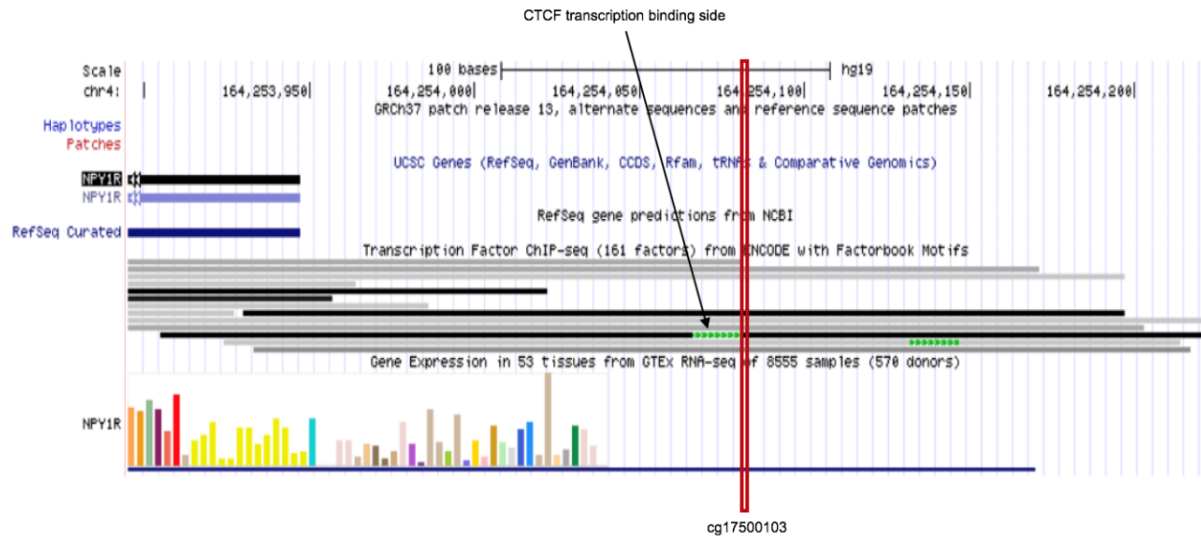
Supplementary Figure 3.2. *PLG* CpG site in UCSC Genome Browser

Singular Value Decomposition Analysis (SVD

Supplementary figure 3.3. Singular value decomposition analysis output after correcting for batch.

Supplementary Table 4.1. Pathway enrichment analysis

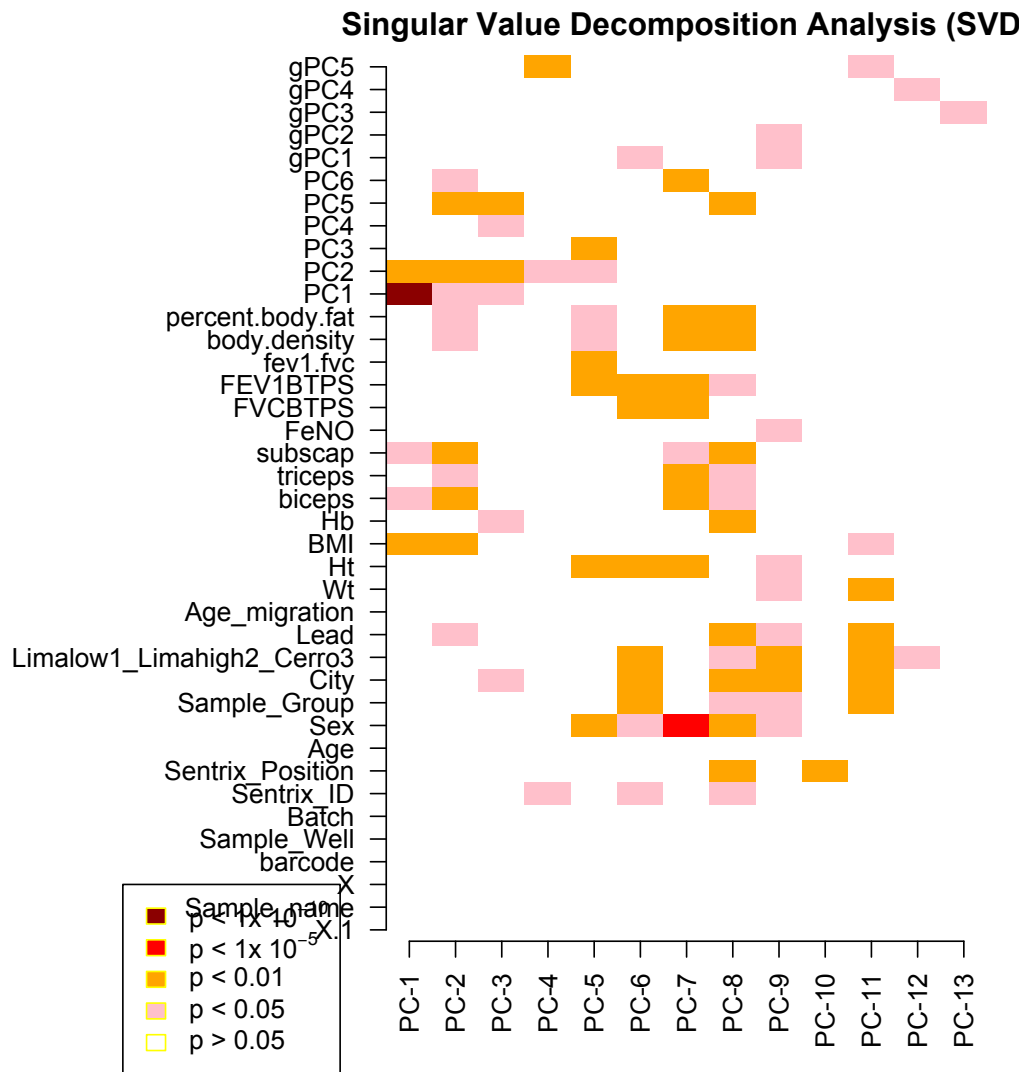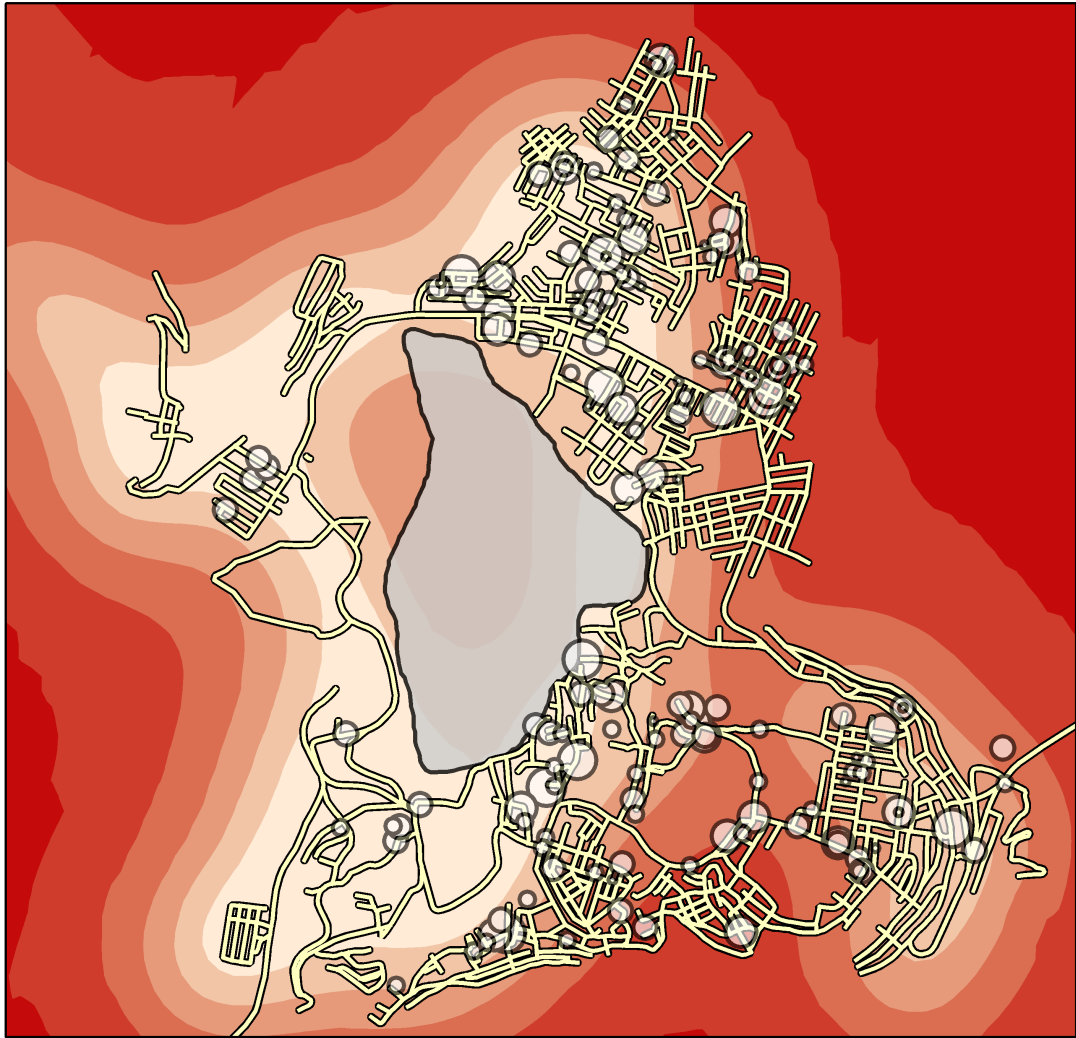|	Gene ontology term	Category, level	Set size	Candidates contained	P-value	Q-value
GO:0044424	intracellular part	CC 2	14559(12901)	8193 (63.5%)	7.66E-49	2.19E-46
GO:0005622	intracellular	CC 2	14838(13142)	8326 (63.4%)	5.98E-48	8.55E-46
GO:0005515	protein binding	MF 2	11696(10465)	6754 (64.5%)	1.01E-43	9.27E-42
GO:0043229	intracellular organelle	CC 2	12763(11325)	7213 (63.7%)	1.20E-34	1.14E-32
GO:0009653	anatomical structure morphogenesis	BP 2	2564(2406)	1720 (71.5%)	5.40E-34	1.30E-31
GO:0016043	cellular component organization	BP 2	6340(5693)	3791 (66.6%)	1.67E-31	2.01E-29
GO:0048856	anatomical structure development	BP 2	5749(5222)	3493 (66.9%)	1.31E-30	1.05E-28
GO:0007275	multicellular organism development	BP 2	5261(4780)	3210 (67.2%)	2.20E-29	1.32E-27
GO:0044446	intracellular organelle part	CC 2	9172(8189)	5306 (64.8%)	4.47E-29	3.19E-27
GO:0043227	membrane-bounded organelle	CC 2	12746(11368)	7186 (63.2%)	7.87E-26	4.50E-24
GO:0048869	cellular developmental process	BP 2	4298(3879)	2612 (67.3%)	8.03E-24	3.87E-22
GO:0043233	organelle lumen	CC 2	5241(4773)	3129 (65.6%)	1.25E-17	5.96E-16
GO:0097458	neuron part	CC 2	1538(1431)	1008 (70.4%)	1.95E-16	7.99E-15
GO:0044237	cellular metabolic process	BP 2	10884(9814)	6190 (63.1%)	2.99E-16	1.20E-14
GO:0042995	cell projection	CC 2	2067(1913)	1316 (68.8%)	9.30E-16	3.32E-14
GO:0036094	small molecule binding	MF 2	2501(2299)	1561 (67.9%)	1.85E-15	8.49E-14
GO:1901363	heterocyclic compound binding	MF 2	6038(5464)	3536 (64.7%)	4.07E-15	1.25E-13
GO:0050789	regulation of biological process	BP 2	11645(10423)	6544 (62.8%)	4.70E-15	1.62E-13
GO:0051234	establishment of localization	BP 2	5192(4671)	3044 (65.2%)	7.16E-15	2.16E-13
GO:0051641	cellular localization	BP 2	2780(2524)	1697 (67.2%)	2.20E-14	5.89E-13
GO:0048646	anatomical structure formation involved in morphogenesis	BP 2	1039(989)	708 (71.6%)	4.41E-14	1.06E-12
GO:0065008	regulation of biological quality	BP 2	3761(3504)	2311 (66.0%)	4.87E-14	1.07E-12
GO:0043167	ion binding	MF 2	6233(5738)	3695 (64.4%)	5.01E-14	1.06E-12
GO:0097159	organic cyclic compound binding	MF 2	6124(5547)	3577 (64.5%)	5.77E-14	1.06E-12
GO:0006928	movement of cell or subcellular component	BP 2	2015(1823)	1248 (68.5%)	5.91E-14	1.19E-12
GO:0044238	primary metabolic process	BP 2	10852(9693)	6095 (62.9%)	8.67E-14	1.61E-12
GO:0006807	nitrogen compound metabolic process	BP 2	10333(9222)	5811 (63.0%)	1.04E-13	1.79E-12
GO:0065009	regulation of molecular function	BP 2	3374(3112)	2059 (66.2%)	3.42E-13	5.50E-12
GO:0071704	organic substance metabolic process	BP 2	11202(10013)	6280 (62.7%)	4.63E-13	6.98E-12

GO:0030029						
	actin filament-based process	BP 2	732(686)	501 (73.0%)	1.57E-12	2.22E-11
GO:0044463	cell projection part	CC 2	1385(1274)	885 (69.5%)	3.15E-12	1.00E-10
GO:0098590	plasma membrane region	CC 2	1074(1001)	707 (70.6%)	4.01E-12	1.15E-10
GO:0009058	biosynthetic process	BP 2	6596(5993)	3833 (64.0%)	4.48E-12	5.99E-11
GO:0097367	carbohydrate derivative binding	MF 2	2205(2031)	1368 (67.4%)	6.20E-12	9.51E-11
GO:0012505	endomembrane system	CC 2	4481(4066)	2643 (65.0%)	7.32E-12	1.90E-10
GO:0033036	macromolecule localization	BP 2	2961(2732)	1810 (66.3%)	7.86E-12	9.97E-11
GO:0140096	catalytic activity, acting on a protein	MF 2	2426(2136)	1433 (67.1%)	1.06E-11	1.40E-10
GO:0043228	non-membrane-bounded organelle	CC 2	4194(3729)	2432 (65.2%)	1.11E-11	2.66E-10
GO:0098794	postsynapse	CC 2	452(422)	320 (75.8%)	1.25E-11	2.74E-10
GO:0007049	cell cycle	BP 2	1882(1713)	1161 (67.8%)	3.14E-11	3.78E-10
GO:0009628	response to abiotic stimulus	BP 2	1202(1126)	781 (69.4%)	1.08E-10	1.24E-09
GO:0098805	whole membrane	CC 2	1645(1495)	1017 (68.0%)	1.77E-10	3.61E-09
GO:0007155	cell adhesion	BP 2	1346(1253)	861 (68.7%)	2.38E-10	2.61E-09
GO:0007267	cell-cell signaling	BP 2	1663(1561)	1056 (67.6%)	5.43E-10	5.68E-09
GO:0008144	drug binding	MF 2	1753(1630)	1099 (67.4%)	7.52E-10	8.65E-09
GO:0048870	cell motility	BP 2	1548(1376)	935 (68.0%)	1.50E-09	1.45E-08
GO:0051674	localization of cell	BP 2	1548(1376)	935 (68.0%)	1.50E-09	1.45E-08
GO:0005911	cell-cell junction	CC 2	440(410)	305 (74.4%)	1.54E-09	2.94E-08
GO:0044459	plasma membrane part	CC 2	2877(2567)	1687 (65.7%)	2.00E-09	3.57E-08
GO:0016049	cell growth	BP 2	471(436)	322 (73.9%)	2.14E-09	1.99E-08
GO:0098984	neuron to neuron synapse	CC 2	236(224)	177 (79.0%)	2.18E-09	3.67E-08
GO:0007154	cell communication	BP 2	6755(6055)	3842 (63.5%)	2.30E-09	2.05E-08
GO:0008219	cell death	BP 2	2208(2045)	1358 (66.4%)	2.55E-09	2.19E-08
GO:0099572	postsynaptic specialization	CC 2	230(219)	173 (79.0%)	3.44E-09	5.47E-08
GO:0007389	pattern specification process	BP 2	426(399)	296 (74.2%)	4.38E-09	3.64E-08
GO:0048589	developmental growth	BP 2	584(544)	392 (72.1%)	6.98E-09	5.61E-08
GO:0070161	anchoring junction	CC 2	550(518)	374 (72.2%)	1.07E-08	1.61E-07
GO:0000988	transcription factor activity, protein binding	MF 2	655(600)	428 (71.3%)	1.16E-08	1.19E-07
GO:0031252	cell leading edge	CC 2	387(367)	272 (74.1%)	2.16E-08	3.09E-07
GO:0005085	guanyl-nucleotide exchange factor activity	MF 2	326(310)	233 (75.2%)	2.95E-08	2.71E-07

GO:0044085	cellular component biogenesis	BP 2	3259(2955)	1918 (64.9%)	3.43E-08	2.66E-07
GO:0009719	response to endogenous stimulus	BP 2	1596(1500)	1002 (66.8%)	7.96E-08	6.00E-07
GO:0007165	signal transduction	BP 2	6219(5548)	3510 (63.3%)	1.44E-07	1.05E-06
GO:0003700	DNA binding transcription factor activity	MF 2	1568(1437)	960 (66.8%)	1.50E-07	1.25E-06
GO:0031090	organelle membrane	CC 2	2986(2673)	1734 (64.9%)	2.39E-07	3.25E-06
GO:0097060	synaptic membrane	CC 2	318(293)	218 (74.4%)	3.21E-07	4.18E-06
GO:0022402	cell cycle process	BP 2	1392(1253)	840 (67.0%)	4.08E-07	2.90E-06
GO:0003682	chromatin binding	MF 2	544(494)	351 (71.1%)	4.62E-07	3.46E-06
GO:0043235	receptor complex	CC 2	375(358)	261 (72.9%)	4.67E-07	5.81E-06
GO:0016740	transferase activity	MF 2	2486(2266)	1477 (65.2%)	4.89E-07	3.46E-06
GO:0022857	transmembrane transporter activity	MF 2	1079(983)	667 (67.9%)	5.45E-07	3.58E-06
GO:0007163	establishment or maintenance of cell polarity	BP 2	207(196)	151 (77.0%)	6.17E-07	4.25E-06
GO:1902494	catalytic complex	CC 2	1433(1296)	865 (66.7%)	8.27E-07	9.85E-06
GO:0002520	immune system development	BP 2	931(871)	594 (68.2%)	8.45E-07	5.66E-06
GO:0098793	presynapse	CC 2	380(356)	258 (72.5%)	1.22E-06	1.40E-05
GO:0005667	transcription factor complex	CC 2	349(328)	239 (72.9%)	1.51E-06	1.66E-05
GO:0051716	cellular response to stimulus	BP 2	7567(6785)	4250 (62.6%)	1.73E-06	1.13E-05
GO:0005057	signal transducer activity, downstream of receptor	MF 2	169(162)	125 (77.2%)	4.87E-06	2.99E-05
GO:0030055	cell-substrate junction	CC 2	405(385)	274 (71.2%)	6.67E-06	7.07E-05
GO:0044877	protein-containing complex binding	MF 2	906(841)	568 (67.5%)	9.08E-06	5.22E-05
GO:0001763	morphogenesis of a branching structure	BP 2	198(190)	143 (75.3%)	1.18E-05	7.50E-05
GO:0044297	cell body	CC 2	514(485)	338 (69.7%)	1.26E-05	0.000129
GO:0001503	ossification	BP 2	368(349)	249 (71.3%)	1.29E-05	7.99E-05
GO:0090130	tissue migration	BP 2	268(254)	185 (72.8%)	2.32E-05	0.00014
GO:1990351	transporter complex	CC 2	331(310)	222 (71.6%)	2.51E-05	0.000248
GO:0099081	supramolecular polymer	CC 2	941(793)	534 (67.3%)	2.75E-05	0.000262
GO:0033002	muscle cell proliferation	BP 2	173(168)	126 (75.0%)	5.09E-05	0.000299
GO:0003006	developmental process involved in reproduction	BP 2	669(610)	415 (68.0%)	5.38E-05	0.000308
GO:0007626	locomotory behavior	BP 2	191(182)	135 (74.2%)	6.95E-05	0.000385
GO:0009056	catabolic process	BP 2	2566(2368)	1517 (64.1%)	7.04E-05	0.000385

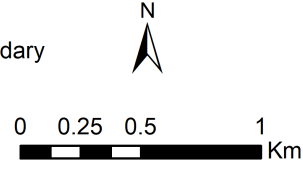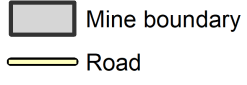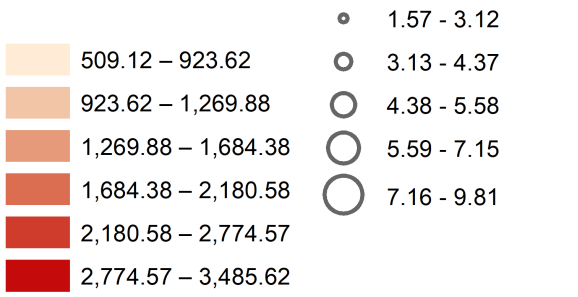| GO ID	Description	Category	Count	Percentage	P1	P2
GO:0000075	cell cycle checkpoint	BP 2	224(213)	155 (72.8%)	0.000111	0.000597
GO:0016247	channel regulator activity	MF 2	139(126)	96 (76.2%)	0.000142	0.000767
GO:0051301	cell division	BP 2	612(551)	374 (67.9%)	0.000165	0.000862
GO:0005905	clathrin-coated pit	CC 2	67(65)	53 (81.5%)	0.00023	0.00212
GO:0061919	process utilizing autophagic mechanism	BP 2	487(452)	309 (68.4%)	0.000277	0.00142
GO:0031012	extracellular matrix	CC 2	482(454)	310 (68.3%)	0.000307	0.00275
GO:0016787	hydrolase activity	MF 2	2706(2367)	1508 (63.7%)	0.000311	0.00158
GO:0007269	neurotransmitter secretion	BP 2	146(135)	101 (74.8%)	0.000321	0.00161
GO:0001775	cell activation	BP 2	1408(1229)	799 (65.0%)	0.000417	0.00205
GO:0044420	extracellular matrix component	CC 2	119(112)	85 (75.9%)	0.000421	0.00365
GO:0032153	cell division site	CC 2	82(66)	53 (80.3%)	0.000481	0.00399
GO:0098589	membrane region	CC 2	317(298)	208 (69.8%)	0.000488	0.00399
GO:0006950	response to stress	BP 2	3998(3551)	2234 (62.9%)	0.000514	0.00248
GO:0098796	membrane protein complex	CC 2	1176(1041)	680 (65.3%)	0.000547	0.00434
GO:0035637	multicellular organismal signaling	BP 2	197(188)	135 (71.8%)	0.000761	0.0036
GO:0019827	stem cell population maintenance	BP 2	154(149)	109 (73.2%)	0.000804	0.00372
GO:0099634	postsynaptic specialization membrane	CC 2	34(32)	28 (87.5%)	0.000843	0.00652
GO:0098727	maintenance of cell number	BP 2	156(151)	110 (72.8%)	0.000983	0.00447
GO:0098636	protein complex involved in cell adhesion	CC 2	35(35)	30 (85.7%)	0.00115	0.00864
GO:0044419	interspecies interaction between organisms	BP 2	893(830)	544 (65.5%)	0.00125	0.00557
GO:0016079	synaptic vesicle exocytosis	BP 2	89(82)	63 (76.8%)	0.00132	0.0058
GO:0032155	cell division site part	CC 2	65(59)	47 (79.7%)	0.00139	0.0102
GO:0030234	enzyme regulator activity	MF 2	1050(968)	630 (65.1%)	0.00144	0.00699
GO:0042330	taxis	BP 2	601(559)	372 (66.5%)	0.00159	0.00686
GO:0051235	maintenance of location	BP 2	311(294)	202 (68.7%)	0.00202	0.00856
GO:0071804	cellular potassium ion transport	BP 2	200(191)	135 (70.7%)	0.0021	0.00875
GO:0045321	leukocyte activation	BP 2	1258(1087)	702 (64.6%)	0.00241	0.00984
GO:0007611	learning or memory	BP 2	240(230)	160 (69.6%)	0.00257	0.0103
GO:0099604	ligand-gated calcium channel activity	MF 2	29(29)	25 (86.2%)	0.00259	0.0119
GO:0030534	adult behavior	BP 2	142(137)	99 (72.3%)	0.00264	0.0104

130

GO ID	Term	Category				
GO:0030496	midbody	CC	2	166(152)	108 (71.1%)	0.00428	0.0306
GO:0031519	PcG protein complex	CC	2	51(45)	36 (80.0%)	0.00441	0.0308
GO:0030427	site of polarized growth	CC	2	159(145)	103 (71.0%)	0.00526	0.0358
GO:0008021	synaptic vesicle	CC	2	177(168)	118 (70.2%)	0.0054	0.0359
GO:0032940	secretion by cell	BP	2	1469(1374)	876 (63.8%)	0.00545	0.0212
GO:0017053	transcriptional repressor complex	CC	2	85(79)	59 (74.7%)	0.00578	0.0376
GO:0006903	vesicle targeting	BP	2	87(82)	61 (74.4%)	0.00586	0.0224
GO:0007568	aging	BP	2	287(269)	183 (68.0%)	0.00611	0.0228
GO:0097242	amyloid-beta clearance	BP	2	19(19)	17 (89.5%)	0.00616	0.0228
GO:0019898	extrinsic component of membrane	CC	2	293(270)	183 (67.8%)	0.0077	0.049
GO:0035264	multicellular organism growth	BP	2	143(137)	97 (70.8%)	0.0077	0.0281
GO:0002262	myeloid cell homeostasis	BP	2	138(131)	93 (71.0%)	0.00796	0.0286
GO:0019725	cellular homeostasis	BP	2	914(853)	550 (64.5%)	0.00806	0.0286
GO:0032039	integrator complex	CC	2	27(14)	13 (92.9%)	0.00893	0.0555

Prediction Std. Error **Blood lead (μg/dL)**

509.12 – 923.62	∘ 1.57 - 3.12
923.62 – 1,269.88	○ 3.13 - 4.37
1,269.88 – 1,684.38	○ 4.38 - 5.58
1,684.38 – 2,180.58	○ 5.59 - 7.15
2,180.58 – 2,774.57	○ 7.16 - 9.81
2,774.57 – 3,485.62	

Mine boundary

Road

N

0   0.25   0.5   1
                          Km

Datum:
WGS84 UTM Zone 18S

Source:
Roads - Open Street Maps
Soil lead samples - van Geen et al., 2012

Supplementary Figure 4.1. Prediction standard error for Kriging analysis in ArcGIS
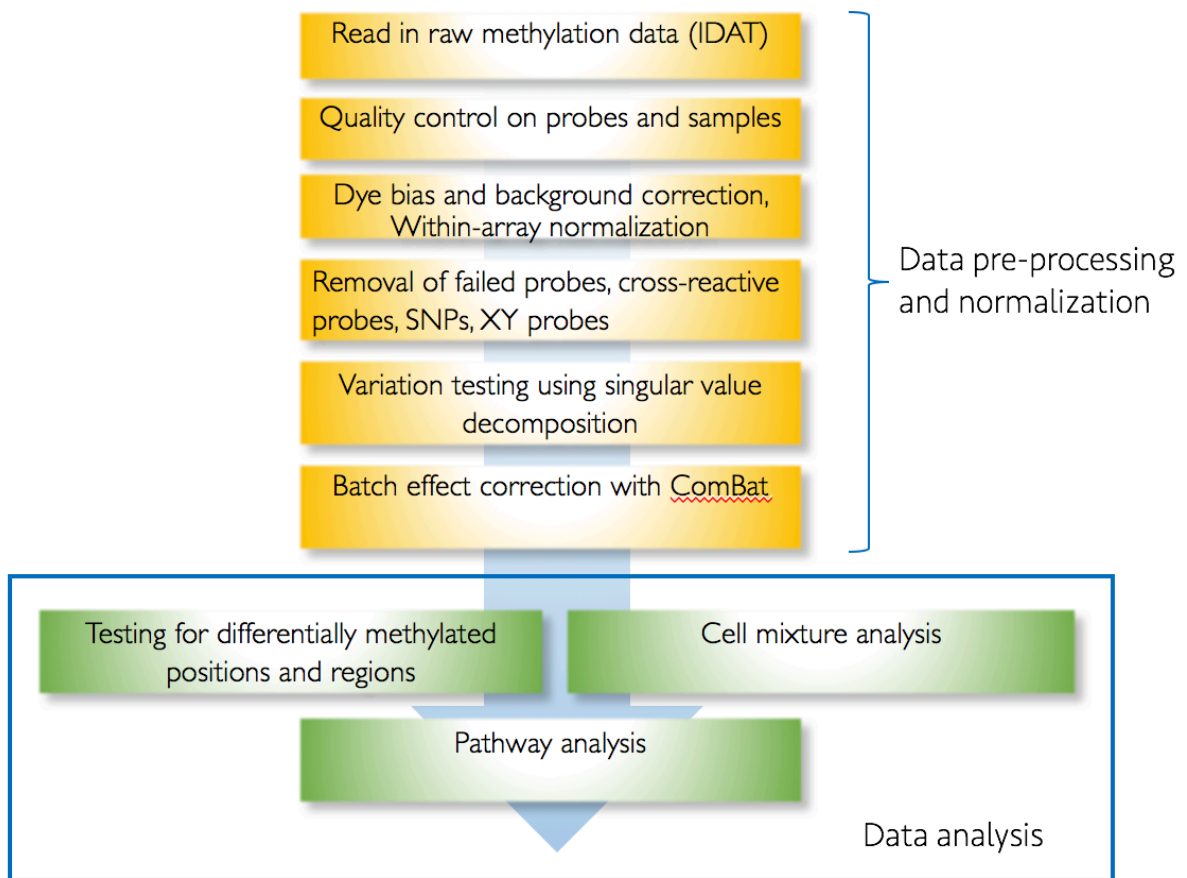
Supplementary Figure 4.2. DNA methylation analysis pipeline
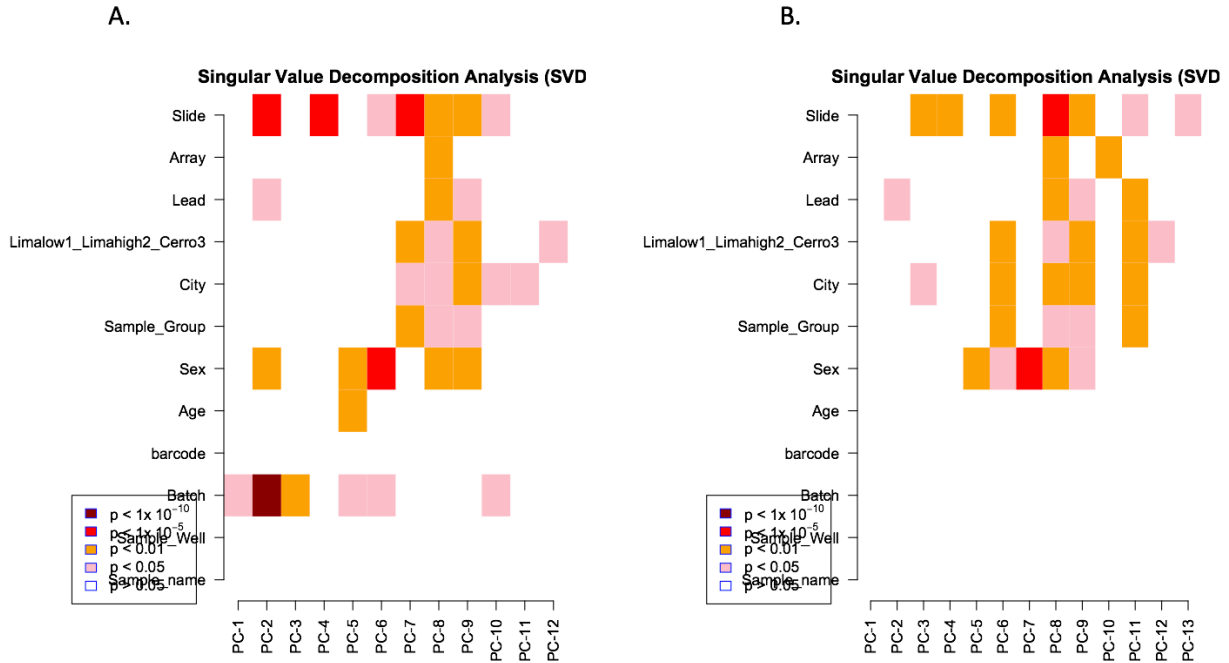
Supplementary Figure 4.3. Singular value decomposition analysis before and after batch effects correction. (A) Singular value decomposition analysis of uncorrected DNA methylation β values. (B) Singular value decomposition analysis of DNA methylation β values that were corrected for batch using Combat. WBLLs are among the significant contributors to genome-wide DNA methylation variation (p-value<0.05) before and after correction for batch effects.

QQ plot

$\lambda = 1.048$

Observed -log10(p-value)

Expected -log10(p-value)
87 Q-Q limma Dec 10 5
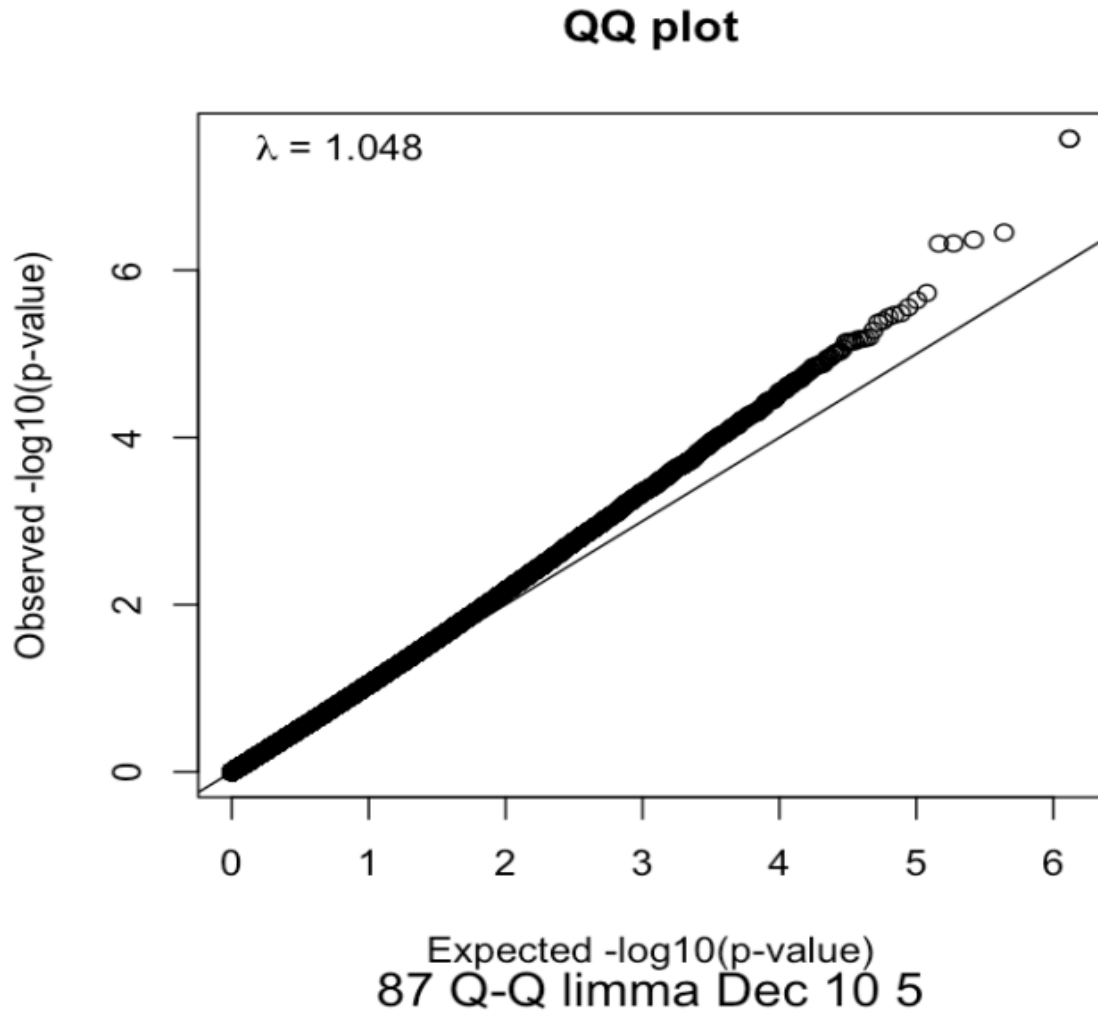
Supplementary Figure 4.4. QQ plot of expected and observed p-values based on the analysis of differential methylation in limma

REFERENCES

<Frisancho, Velasquez, Sanchez - 1973 - Influence of Developmental Adaptation on Lung Function at High Altitude.pdf>.

Adonaylo VN, Oteiza PI. 1999. Lead intoxication: antioxidant defenses and oxidative damage in rat brain. Toxicology 135:77-85.

Ahamed M, Fareed M, Kumar A, Siddiqui WA, Siddiqui MK. 2008. Oxidative stress and neurological disorders in relation to blood lead levels in children. Redox Rep 13:117-122.

Ahamed M, Siddiqui MK. 2007. Low level lead exposure and oxidative stress: current opinions. Clin Chim Acta 383:57-64.

Albahary C. 1972. Lead and hemopoiesis. The mechanism and consequences of the erythropathy of occupational lead poisoning. Am J Med 52:367-378.

Aldenderfer MS. 2008. High Elevation Foraging Societies. Handbook of South American Archaeology:131-143.

Alegria-Torres JA, Baccarelli A, Bollati V. 2011. Epigenetics and lifestyle. Epigenomics 3:267-277.

Alkelai A, Lupoli S, Greenbaum L, Kohn Y, Kanyas-Sarner K, Ben-Asher E, Lancet D, Macciardi F, Lerer B. 2012. DOCK4 and CEACAM21 as novel schizophrenia candidate genes in the Jewish population. Int J Neuropsychopharmacol 15:459-469.

Alkorta-Aranburu G, Beall CM, Witonsky DB, Gebremedhin A, Pritchard JK, Di Rienzo A. 2012. The genetic architecture of adaptations to high altitude in Ethiopia. PLoS Genet 8:e1003110.

Allis CD, Jenuwein T. 2016. The molecular hallmarks of epigenetic control. Nat Rev Genet 17:487-500.

Anderson OS, Sant KE, Dolinoy DC. 2012. Nutrition and epigenetics: an interplay of dietary methyl donors, one-carbon metabolism and DNA methylation. J Nutr Biochem 23:853-859.

Aneiros-Guerrero A, Lendinez AM, Palomares AR, Perez-Nevot B, Aguado L, Mayor-Olea A, Ruiz-Galdon M, Reyes-Engel A. 2011. Genetic polymorphisms in folate pathway enzymes, DRD4 and GSTM1 are related to temporomandibular disorder. BMC Med Genet 12:75.

Antezana AM, Antezana G, Aparicio O, Noriega I, Velarde FL, Richalet JP. 1998. Pulmonary hypertension in high-altitude chronic hypoxia: response to nifedipine. Eur Respir J 12:1181-1185

Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA. 2014. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics 30:1363-1369.

Baccarelli A, Bollati V. 2009. Epigenetics and environmental chemicals. Curr Opin Pediatr 21:243-251.

Baccarelli A, Wright RO, Bollati V, Tarantini L, Litonjua AA, Suh HH, Zanobetti A, Sparrow D, Vokonas PS, Schwartz J. 2009. Rapid DNA methylation changes after exposure to traffic particles. Am J Respir Crit Care Med 179:572-578.

Baker PT. 1969. Human adaptation to high altitude. Science 163:1149-1156.

Baker PT, Little MA. 1976. Man in the Andes : multidisciplinary study of high-altitude Quechua. Stroudsburg, Pa.
[New York]: Dowden
Exclusive distributor, Halsted Press.

Barbosa F, Tanus-Santos JE, Gerlach RF, Parsons PJ. 2005. A critical review of biomarkers used for monitoring human exposure to lead: Advantages, limitations, and future needs. Environmental Health Perspectives 113:1669-1674.

Barcroft J, Binger C, Bock AV, Doggart J, Forbes H, Harrop G, Meakins JC, Redfield AC, Davies H, Scott JD. 1923. Observations upon the effect of high altitude on the physiological processes of the human body, carried out in the Peruvian Andes, chiefly at Cerro de Pasco. Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character:351-480.

Barltrop D, Smith A. 1971. Interaction of lead with erythrocytes. Experientia 27:92-93.

Barltrop D, Smith A. 1972. Lead binding to human haemoglobin. Experientia 28:76-77.

Bates D, Machler M, Bolker BM, Walker SC. 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 67:1-48.

Beall CM. 2006. Andean, Tibetan, and Ethiopian patterns of adaptation to high-altitude hypoxia. Integr Comp Biol 46:18-24.

Beall CM. 2007. Two routes to functional adaptation: Tibetan and Andean high-altitude natives. Proc Natl Acad Sci U S A 104 Suppl 1:8655-8660.

Beall CM, Brittenham GM, Strohl KP, Blangero J, Williams-Blangero S, Goldstein MC, Decker MJ, Vargas E, Villena M, Soria R, et al. 1998. Hemoglobin concentration of high-altitude Tibetans and Bolivian Aymara. Am J Phys Anthropol 106:385-400.

Beall CM, Cavalleri GL, Deng L, Elston RC, Gao Y, Knight J, Li C, Li JC, Liang Y, McCormack M, et al. 2010. Natural selection on EPAS1 (HIF2alpha) associated with low hemoglobin concentration in Tibetan highlanders. Proc Natl Acad Sci U S A 107:11459-11464.

Beall CM, Strohl KP, Blangero J, Williams-Blangero S, Almasy LA, Decker MJ, Worthman CM, Goldstein MC, Vargas E, Villena M, et al. 1997. Ventilation and hypoxic ventilatory response of Tibetan and Aymara high altitude natives. Am J Phys Anthropol 104:427-447.

Belancio VP, Deininger PL, Roy-Engel AM. 2009. LINE dancing in the human genome: transposable elements and disease. Genome Med 1:97.

Bergdahl IA, Schutz A, Gerhardsson L, Jensen A, Skerfving S. 1997. Lead concentrations in human plasma, urine and whole blood. Scand J Work Environ Health 23:359-363.

Bernal AJ, Dolinoy DC, Huang D, Skaar DA, Weinhouse C, Jirtle RL. 2013. Adaptive radiation-induced epigenetic alterations mitigated by antioxidants. FASEB J 27:665-671.

Bigagli E, Luceri C, Bernardini S, Dei A, Dolara P. 2010. Extremely low copper concentrations affect gene expression profiles of human prostate epithelial cell lines. Chem Biol Interact 188:214-219.

Bigham A, Bauchet M, Pinto D, Mao X, Akey JM, Mei R, Scherer SW, Julian CG, Wilson MJ, Lopez Herraez D, et al. 2010. Identifying signatures of natural selection in Tibetan and Andean populations using dense genome scan data. PLoS Genet 6:e1001116.

Bigham AW. 2016. Genetics of human origin and evolution: high-altitude adaptations. Curr Opin Genet Dev 41:8-13.

Bigham AW, Julian CG, Wilson MJ, Vargas E, Browne VA, Shriver MD, Moore LG. 2014. Maternal PRKAA1 and EDNRA genotypes are associated with birth weight, and PRKAA1 with uterine artery diameter and metabolic homeostasis at high altitude. Physiol Genomics 46:687-697.

Bigham AW, Kiyamu M, Leon-Velarde F, Parra EJ, Rivera-Ch M, Shriver MD, Brutsaert TD. 2008. Angiotensin-converting enzyme genotype and arterial oxygen saturation at high altitude in Peruvian Quechua. High Alt Med Biol 9:167-178.

Bigham AW, Lee FS. 2014. Human high-altitude adaptation: forward genetics meets the HIF pathway. Genes Dev 28:2189-2204.

Bigham AW, Mao X, Mei R, Brutsaert T, Wilson MJ, Julian CG, Parra EJ, Akey JM, Moore LG, Shriver MD. 2009. Identifying positive selection candidate loci for high-altitude adaptation in Andean populations. Hum Genomics 4:79-90.

Bigham AW, Wilson MJ, Julian CG, Kiyamu M, Vargas E, Leon-Velarde F, Rivera-Chira M, Rodriquez C, Browne VA, Parra E, et al. 2013. Andean and Tibetan patterns of adaptation to high altitude. Am J Hum Biol 25:190-197.

Bleich S, Semmler A, Frieling H, Thumfart L, Muschler M, Hillemacher T, Kornhuber J, Kallweit U, Simon M, Linnebank M. 2014. Genetic variants of methionine metabolism and DNA methylation. Epigenomics 6:585-591.

Boeke CE, Baccarelli A, Kleinman KP, Burris HH, Litonjua AA, Rifas-Shiman SL, Tarantini L, Gillman M. 2012. Gestational intake of methyl donors and global LINE-1 DNA methylation in

maternal and cord blood: prospective results from a folate-replete population. Epigenetics 7:253-260.

Boks MP, van Mierlo HC, Rutten BP, Radstake TR, De Witte L, Geuze E, Horvath S, Schalkwyk LC, Vinkers CH, Broen JC, et al. 2015. Longitudinal changes of telomere length and epigenetic age related to traumatic stress and post-traumatic stress disorder. Psychoneuroendocrinology 51:506-512.

Bollati V, Baccarelli A. 2010. Environmental epigenetics. Heredity (Edinb) 105:105-112.

Boyko A, Blevins T, Yao Y, Golubov A, Bilichak A, Ilnytskyy Y, Hollunder J, Meins F, Jr., Kovalchuk I. 2010. Transgenerational adaptation of Arabidopsis to stress requires DNA methylation and the function of Dicer-like proteins. PLoS One 5:e9514.

Boyko A, Kovalchuk I. 2010. Transgenerational response to stress in Arabidopsis thaliana. Plant Signal Behav 5:995-998.

Breton CV, Marsit CJ, Faustman E, Nadeau K, Goodrich JM, Dolinoy DC, Herbstman J, Holland N, LaSalle JM, Schmidt R, et al. 2017. Small-Magnitude Effect Sizes in Epigenetic End Points are Important in Children's Environmental Health Studies: The Children's Environmental Health and Disease Prevention Research Center's Epigenetics Working Group. Environ Health Perspect 125:511-526.

Brody LC, Conley M, Cox C, Kirke PN, McKeever MP, Mills JL, Molloy AM, O'Leary VB, Parle-McDermott A, Scott JM, et al. 2002. A polymorphism, R653Q, in the trifunctional enzyme methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase/formyltetrahydrofolate synthetase is a maternal genetic risk factor for neural tube defects: report of the Birth Defects Research Group. Am J Hum Genet 71:1207-1215.

Brutsaert TD. 2001. Genetic and environmental adaptation in high altitude natives. Conceptual, methodological, and statistical concerns. Adv Exp Med Biol 502:133-151.

Brutsaert TD, Araoz M, Soria R, Spielvogel H, Haas JD. 2000. Higher arterial oxygen saturation during submaximal exercise in Bolivian Aymara compared to European sojourners and Europeans born and raised at high altitude. Am J Phys Anthropol 113:169-181.

Brutsaert TD, Soria R, Caceres E, Spielvogel H, Haas JD. 1999. Effect of developmental and ancestral high altitude exposure on chest morphology and pulmonary function in Andean and European/North American natives. Am J Hum Biol 11:383-395.

Burgess TM, Webster R. 1980. OPTIMAL INTERPOLATION AND ISARITHMIC MAPPING OF SOIL PROPERTIES. Journal of Soil Science 31:315-331.

Carapito R, Ivanova EL, Morlon A, Meng L, Molitor A, Erdmann E, Kieffer B, Pichot A, Naegely L, Kolmer A, et al. 2019. ZMIZ1 Variants Cause a Syndromic Neurodevelopmental Disorder. Am J Hum Genet 104:319-330.

Carvill GL, Heavin SB, Yendle SC, McMahon JM, O'Roak BJ, Cook J, Khan A, Dorschner MO, Weaver M, Calvert S, et al. 2013. Targeted resequencing in epileptic encephalopathies identifies de novo mutations in CHD2 and SYNGAP1. Nat Genet 45:825-830.

Cencioni C, Spallotta F, Martelli F, Valente S, Mai A, Zeiher AM, Gaetano C. 2013. Oxidative stress and epigenetic regulation in ageing and age-related diseases. Int J Mol Sci 14:17643-17663.

Chalitchagorn K, Shuangshoti S, Hourpai N, Kongruttanachok N, Tangkijvanich P, Thong-ngam D, Voravud N, Sriuranpong V, Mutirangura A. 2004. Distinctive pattern of LINE-1 methylation level in normal tissues and the association with carcinogenesis. Oncogene 23:8841-8846.

Chang SC, Chang PY, Butler B, Goldstein BY, Mu L, Cai L, You NC, Baecker A, Yu SZ, Heber D, et al. 2014. Single nucleotide polymorphisms of one-carbon metabolism and cancers of the esophagus, stomach, and liver in a Chinese population. PLoS One 9:e109235.

Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, Gallinger S, Hudson TJ, Weksberg R. 2013. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics 8:203-209.

Childebayeva A, Jones TR, Goodrich JM, Leon-Velarde F, Rivera-Chira M, Kiyamu M, Brutsaert TD, Dolinoy DC, Bigham AW. 2019. LINE-1 and EPAS1 DNA methylation associations with high-altitude exposure. Epigenetics 14:1-15.

Conklin L., Carlos A. Sánchez MD, Antonio Neri MD, Paula Staley RN MPA, Wendy Blumenthal MPH, Jeffrey M Jarrett MS, Rebecca LePrell MPH, James Durant MSPH CIH, MSEnvE RJS-S. 2008. Exposiciones a metales pesados en niños y mujeres en edad fértil en tres comunidades mineras Cerro de Pasco, Perú. In: CDC.

Cooke CA, Wolfe AP, Hobbs WO. 2009. Lake-sediment geochemistry reveals 1400 years of evolving extractive metallurgy at Cerro de Pasco, Peruvian Andes. Geology 37:1019-1022.

Counter SA, Buchanan LH, Ortega F. 2001. Gender differences in blood lead and hemoglobin levels in Andean adults with chronic lead exposure. Int J Occup Environ Health 7:113-118.

Counter SA, Buchanan LH, Ortega F, Rifai N. 2000. Blood lead and hemoglobin levels in Andean children with chronic lead intoxication. Neurotoxicology 21:301-308.

Crnkovic S, Egemnazarov B, Jain P, Seay U, Gattinger N, Marsh LM, Balint Z, Kovacs G, Ghanim B, Klepetko W, et al. 2014. NPY/Y(1) receptor-mediated vasoconstrictory and proliferative effects in pulmonary hypertension. Br J Pharmacol 171:3895-3907.

Csoka AB, Szyf M. 2009. Epigenetic side-effects of common pharmaceuticals: a potential new field in medicine and pharmacology. Med Hypotheses 73:770-780.

Davidson C, Rabinowitz M. 1991. Lead in the environment: from sources to human. Human lead exposure:65.

Der R, Fahim Z, Hilderbrand D, Fahim M. 1974. Combined effect of lead and low protein diet on growth, sexual development, and metabolism in female rats. Res Commun Chem Pathol Pharmacol 9:723-738.

Devereux TR, Horikawa I, Anna CH, Annab LA, Afshari CA, Barrett JC. 1999. DNA methylation analysis of the promoter region of the human telomerase reverse transcriptase (hTERT) gene. Cancer Res 59:6087-6090.

Devlin AM, Brain U, Austin J, Oberlander TF. 2010. Prenatal exposure to maternal depressed mood and the MTHFR C677T variant affect SLC6A4 methylation in infants at birth. PLoS One 5:e12201.

Dolinoy DC, Huang D, Jirtle RL. 2007. Maternal nutrient supplementation counteracts bisphenol A-induced DNA hypomethylation in early development. Proc Natl Acad Sci U S A 104:13056-13061.

Downes M, Koopman P. 2001. SOX18 and the transcriptional regulation of blood vessel development. Trends Cardiovasc Med 11:318-324.

Dubin MJ, Zhang P, Meng D, Remigereau MS, Osborne EJ, Paolo Casale F, Drewe P, Kahles A, Jean G, Vilhjalmsson B, et al. 2015. DNA methylation in Arabidopsis has a genetic basis and shows evidence of local adaptation. Elife 4:e05255.

Ducker GS, Rabinowitz JD. 2017. One-Carbon Metabolism in Health and Disease. Cell Metab 25:27-42.

Eddahibi S, Hanoun N, Lanfumey L, Lesch KP, Raffestin B, Hamon M, Adnot S. 2000. Attenuated hypoxic pulmonary hypertension in mice lacking the 5-hydroxytryptamine transporter gene. J Clin Invest 105:1555-1562.

Eichstaedt CA, Pagani L, Antao T, Inchley CE, Cardona A, Morseburg A, Clemente FJ, Sluckin TJ, Metspalu E, Mitt M, et al. 2017. Evidence of Early-Stage Selection on EPAS1 and GPR126 Genes in Andean High Altitude Populations. Sci Rep 7:13042.

El-Maarri O, Becker T, Junen J, Manzoor SS, Diaz-Lacava A, Schwaab R, Wienker T, Oldenburg J. 2007. Gender specific differences in levels of DNA methylation at selected loci from human total blood: a tendency toward higher methylation levels in males. Hum Genet 122:505-514.

Elias RW, Gulson B. 2003. Overview of lead remediation effectiveness. Sci Total Environ 303:1-13.

Engstrom K, Rydbeck F, Kippler M, Wojdacz TK, Arifeen S, Vahter M, Broberg K. 2015. Prenatal lead exposure is associated with decreased cord blood DNA methylation of the glycoprotein VI gene involved in platelet activation and thrombus formation. Environ Epigenet 1:dvv007.

Evans HK, Wylie AA, Murphy SK, Jirtle RL. 2001. The neuronatin gene resides in a "micro-imprinted" domain on human chromosome 20q11.2. Genomics 77:99-104.

Faulk C, Barks A, Liu K, Goodrich JM, Dolinoy DC. 2013. Early-life lead exposure results in dose- and sex-specific effects on weight and epigenetic gene regulation in weanling mice. Epigenomics 5:487-500.

Faulk C, Dolinoy DC. 2014. Timing is everything. Epigenetics 6:791-797.

Feil R, Fraga MF. 2011. Epigenetics and the environment: emerging patterns and implications. Nat Rev Genet 13:97-109.

Feinberg AP, Irizarry RA. 2010. Evolution in health and medicine Sackler colloquium: Stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease. Proc Natl Acad Sci U S A 107 Suppl 1:1757-1764.

Fewings N, Gatt PN, McKay FC, Parnell GP, Schibeci SD, Edwards J, Basuki MA, Goldinger A, Fabis-Pedrini MJ, Kermode AG, et al. 2017. Data characterizing the ZMIZ1 molecular phenotype of multiple sclerosis. Data Brief 11:364-370.

Fischer JR. 1977. Silver Mines and Silver Miners in Colonial Peru, 1776 - 1824: University of Liverpool. Centre for Latin American Studies.

Flora G, Gupta D, Tiwari A. 2012. Toxicity of lead: A review with recent updates. Interdiscip Toxicol 5:47-58.

Foll M, Gaggiotti OE, Daub JT, Vatsiou A, Excoffier L. 2014. Widespread signals of convergent adaptation to high altitude in Asia and america. Am J Hum Genet 95:394-407.

Fortin JP, Labbe A, Lemire M, Zanke BW, Hudson TJ, Fertig EJ, Greenwood CM, Hansen KD. 2014. Functional normalization of 450k methylation array data improves replication in large cancer studies. Genome Biol 15:503.

Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, Heine-Suner D, Cigudosa JC, Urioste M, Benitez J, et al. 2005. Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci U S A 102:10604-10609.

Fraga MF, Esteller M. 2002. DNA methylation: a profile of methods and applications. Biotechniques 33:632, 634, 636-649.

Frank DB, Abtahi A, Yamaguchi DJ, Manning S, Shyr Y, Pozzi A, Baldwin HS, Johnson JE, de Caestecker MP. 2005. Bone morphogenetic protein 4 promotes pulmonary vascular remodeling in hypoxic pulmonary hypertension. Circ Res 97:496-504.

Fraser HB, Lam LL, Neumann SM, Kobor MS. 2012. Population-specificity of human DNA methylation. Genome Biol 13:R8.

Frisancho AR. 1977. Developmental adaptation to high altitude hypoxia. Int J Biometeorol 21:135-146.

Frisancho AR. 2009. Developmental adaptation: where we go from here. Am J Hum Biol 21:694-703.

Frisancho AR. 2013. Developmental functional adaptation to high altitude: review. Am J Hum Biol 25:151-168.

Frisancho AR. 1970. Developmental responses to high altitude hypoxia. Am J Phys Anthropol 32:401-407.

Frisancho AR. 1975. Functional adaptation to high altitude hypoxia. Science 187:313-319.

Frisancho AR. 1983. Perspectives on functional adaptation of the high altitude native. Prog Clin Biol Res 136:383-407.

Frisancho AR. 1997. Reduction of birth weight among infants born to adolescents: maternal-fetal growth competition. Ann N Y Acad Sci 817:272-280.

Frisancho AR, Baker PT. 1970. Altitude and growth: a study of the patterns of physical growth of a high altitude Peruvian Quechua population. Am J Phys Anthropol 32:279-292.

Frisancho AR, Frisancho HG, Albalak R, Villain M, Vargas E, Soria R. 1997. Developmental, genetic, and environmental components of lung volumes at high altitude. Am J Hum Biol 9:191-203.

Frisancho AR, Martinez C, Velasquez T, Sanchez J, Montoye H. 1973. Influence of developmental adaptation on aerobic capacity at high altitude. J Appl Physiol 34:176-180.

Gasparoni G, Bultmann S, Lutsik P, Kraus TFJ, Sordon S, Vlcek J, Dietinger V, Steinmaurer M, Haider M, Mulholland CB, et al. 2018. DNA methylation analysis on purified neurons and glia dissects age and Alzheimer's disease-specific changes in the human cortex. Epigenetics Chromatin 11:41.

Gibson TM, Brennan P, Han S, Karami S, Zaridze D, Janout V, Kollarova H, Bencko V, Navratilova M, Szeszenia-Dabrowska N, et al. 2011. Comprehensive evaluation of one-carbon metabolism pathway gene variants and renal cell cancer risk. PLoS One 6:e26165.

Goodrich JM, Dolinoy DC, Sanchez BN, Zhang Z, Meeker JD, Mercado-Garcia A, Solano-Gonzalez M, Hu H, Tellez-Rojo MM, Peterson KE. 2016. Adolescent epigenetic profiles and environmental exposures from early life through peri-adolescence. Environ Epigenet 2:dvw018.

Goodrich JM, Dolinoy DC, Sánchez BN, Zhang Z, Meeker JD, Mercado-Garcia A, Solano-González M, Hu H, Téllez-Rojo MM, Peterson KE. 2016. Adolescent epigenetic profiles and environmental exposures from early life through peri-adolescence. Environ Epigenet 2.

Goodrich JM, Reddy P, Naidoo RN, Asharam K, Batterman S, Dolinoy DC. 2016. Prenatal exposures and DNA methylation in newborns: a pilot study in Durban, South Africa. Environ Sci Process Impacts 18:908-917.

Graler B, Pebesma E, Heuvelink G. 2016. Spatio-Temporal Interpolation using gstat. The R Journal 8:204-218.

Grant LD. 2009a. Lead and Compounds. In: Lippmann M, editor. Environmental Toxicants: Wiley & Sons, Inc.

Grant LD. 2009b. Lead and Compounds. In: Lippmann M, editor. Environmental Toxicants: Human exposure and their health effects: Wiley-Interscience.

Gulson BL, Calder IC. 1995. Lead mobilisation during pregnancy. Med J Aust 163:447.

Gulson BL, Mahaffey KR, Mizon KJ, Korsch MJ, Cameron MA, Vimpani G. 1995. Contribution of tissue lead to blood lead in adult female subjects based on stable lead isotope methods. J Lab Clin Med 125:703-712.

Gulson BL, Mizon KJ, Korsch MJ, Palmer JM, Donnelly JB. 2003. Mobilization of lead from human bone tissue during pregnancy and lactation--a summary of long-term research. Sci Total Environ 303:79-104.

Guo J, Xie H, Wang J, Zhao H, Wang F, Liu C, Wang L, Lu X, Bao Y, Zou J, et al. 2013. The maternal folate hydrolase gene polymorphism is associated with neural tube defects in a high-risk Chinese population. Genes Nutr 8:191-197.

Guzy RD. 2005. Role of the mitochondrial electron transport chain in oxygen sensing and hypoxic HIF-1 alpha stabilization. In.

Haefliger P. 2011. Brief guide to analytical methods for measuring lead in blood. WHO Library Cataloguing-in-Publication Data.

Hanahan D, Folkman J. 1996. Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis. Cell 86:353-364.

Hanna CW, Bloom MS, Robinson WP, Kim D, Parsons PJ, vom Saal FS, Taylor JA, Steuerwald AJ, Fujimoto VY. 2012. DNA methylation changes in whole blood is associated with exposure to the environmental contaminants, mercury, lead, cadmium and bisphenol A, in women undergoing ovarian stimulation for IVF. Hum Reprod 27:1401-1410.

Hanna-Attisha M, LaChance J, Sadler RC, Champney Schnepp A. 2016. Elevated Blood Lead Levels in Children Associated With the Flint Drinking Water Crisis: A Spatial Analysis of Risk and Public Health Response. Am J Public Health 106:283-290.

Hatzimichael E, Dranitsaris G, Dasoula A, Benetatos L, Stebbing J, Crook T, Bourantas KL. 2009. Von Hippel-Lindau methylation status in patients with multiple myeloma: a potential predictive factor for the development of bone disease. Clin Lymphoma Myeloma 9:239-242.

Hazra A, Fuchs CS, Kawasaki T, Kirkner GJ, Hunter DJ, Ogino S. 2010. Germline polymorphisms in the one-carbon metabolism pathway and DNA methylation in colorectal cancer. Cancer Causes Control 21:331-345.

Heath D, Williams D. 1991. Pulmonary vascular remodelling in a high-altitude Aymara Indian. Int J Biometeorol 35:203-207.

Helfgott FM. 2013. Transformations in Labor, Land and Community: Mining and Society in Pasco, Peru, 20th Century to the Present. [University of Michigan.

Herman JG, Latif F, Weng Y, Lerman MI, Zbar B, Liu S, Samid D, Duan DS, Gnarra JR, Linehan WM, et al. 1994. Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. Proc Natl Acad Sci U S A 91:9700-9704.

Hernberg S. 2000. Lead poisoning in a historical perspective. Am J Ind Med 38:244-254.

Herwig R, Hardt C, Lienhard M, Kamburov A. 2016. Analyzing and interpreting genome data at the network level with ConsensusPathDB. Nat Protoc 11:1889-1907.

Heyn H, Moran S, Hernando-Herraez I, Sayols S, Gomez A, Sandoval J, Monk D, Hata K, Marques-Bonet T, Wang L, et al. 2013. DNA methylation contributes to natural human variation. Genome Res 23:1363-1372.

Hijmans RJ, Williams E, Vennes C. 2017. Geosphere: spherical trigonometry. R.

Horvath S. 2013. DNA methylation age of human tissues and cell types. Genome Biol 14:R115.

Horvath S, Erhart W, Brosch M, Ammerpohl O, von Schonfels W, Ahrens M, Heits N, Bell JT, Tsai PC, Spector TD, et al. 2014. Obesity accelerates epigenetic aging of human liver. Proc Natl Acad Sci U S A 111:15538-15543.

Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, Wiencke JK, Kelsey KT. 2012. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics 13:86.

Hu H, Watanabe H, Payton M, Korrick S, Rotnitzky A. 1994. The relationship between bone lead and hemoglobin. JAMA 272:1512-1517.

Huang ST, Vo KC, Lyell DJ, Faessen GH, Tulac S, Tibshirani R, Giaccia AJ, Giudice LC. 2004. Developmental response to hypoxia. FASEB J 18:1348-1365.

Huo X, Peng L, Xu X, Zheng L, Qiu B, Qi Z, Zhang B, Han D, Piao Z. 2007. Elevated blood lead levels of children in Guiyu, an electronic waste recycling town in China. Environ Health Perspect 115:1113-1117.

Hurtado A. 1932a. Respiratory adaptation in the Indian natives of the peruvian andes. Studies at high altitude. Am J Phys Anthropol 17:137-165.

Hurtado A. 1932b. Studies at high altitude: Blood Observations on the Indian Natives of the Peruvian Andes. American Journal of Physiology-Legacy Content 100:487-505.

Irahara N, Nosho K, Baba Y, Shima K, Lindeman NI, Hazra A, Schernhammer ES, Hunter DJ, Fuchs CS, Ogino S. 2010. Precision of pyrosequencing assay to measure LINE-1 methylation in colon cancer, normal colonic mucosa, and peripheral blood cells. J Mol Diagn 12:177-183.

Jefferson JA, Simoni J, Escudero E, Hurtado ME, Swenson ER, Wesson DE, Schreiner GF, Schoene RB, Johnson RJ, Hurtado A. 2004. Increased oxidative stress following acute and chronic high altitude exposure. High Alt Med Biol 5:61-69.

Ji Y, Hebbring S, Zhu H, Jenkins GD, Biernacka J, Snyder K, Drews M, Fiehn O, Zeng Z, Schaid D, et al. 2011. Glycine and a glycine dehydrogenase (GLDC) SNP as citalopram/escitalopram response biomarkers in depression: pharmacometabolomics-informed pharmacogenomics. Clin Pharmacol Ther 89:97-104.

Jiang J, Zhao LJ, Zhao C, Zhang G, Zhao Y, Li JR, Li XP, Wei LH. 2012. Hypomethylated CpG around the transcription start site enables TERT expression and HPV16 E6 regulates TERT methylation in cervical cancer cells. Gynecol Oncol 124:534-541.

Johnson WE, Li C, Rabinovic A. 2007. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 8:118-127.

Joseph R, Dou D, Tsang W. 1995. Neuronatin mRNA: alternatively spliced forms of a novel brain-specific mammalian developmental gene. Brain Res 690:92-98.

Julian CG. 2017. Epigenomics and human adaptation to high altitude. J Appl Physiol (1985) 123:1362-1370.

Kalia K, Flora SJ. 2005. Strategies for safe and effective therapeutic measures for chronic arsenic and lead poisoning. J Occup Health 47:1-21.

Kamburov A, Pentchev K, Galicka H, Wierling C, Lehrach H, Herwig R. 2011. ConsensusPathDB: toward a more complete picture of cell biology. Nucleic Acids Res 39:D712-717.

Kanherkar RR, Bhatia-Dey N, Csoka AB. 2014. Epigenetics across the human lifespan. Front Cell Dev Biol 2:49.

Kanno H, Fujii H, Hirono A, Omine M, Miwa S. 1992. Identical point mutations of the R-type pyruvate kinase (PK) cDNA found in unrelated PK variants associated with hereditary hemolytic anemia. Blood 79:1347-1350.

Keegan A, Morecroft I, Smillie D, Hicks MN, MacLean MR. 2001. Contribution of the 5-HT(1B) receptor to hypoxia-induced pulmonary hypertension: converging evidence using 5-HT(1B)-receptor knockout mice and the 5-HT(1B/1D)-receptor antagonist GR127935. Circ Res 89:1231-1239.

Kim GH, Ryan JJ, Marsboom G, Archer SL. 2011. Epigenetic mechanisms of pulmonary hypertension. Pulm Circ 1:347-356.

Kim W, Woo HD, Lee J, Choi IJ, Kim YW, Sung J, Kim J. 2016. Dietary folate, one-carbon metabolism-related genes, and gastric cancer risk in Korea. Mol Nutr Food Res 60:337-345.

Kisko M, Bouain N, Safi A, Medici A, Akkers RC, Secco D, Fouret G, Krouk G, Aarts MG, Busch W, et al. 2018. LPCAT1 controls phosphate homeostasis in a zinc-dependent manner. Elife 7.

Kitamura E, Igarashi J, Morohashi A, Hida N, Oinuma T, Nemoto N, Song F, Ghosh S, Held WA, Yoshida-Noro C, et al. 2007. Analysis of tissue-specific differentially methylated regions (TDMs) in humans. Genomics 89:326-337.

Kiyamu M, Leon-Velarde F, Rivera-Chira M, Elias G, Brutsaert TD. 2015. Developmental Effects Determine Submaximal Arterial Oxygen Saturation in Peruvian Quechua. High Alt Med Biol 16:138-146.

Klengel T, Mehta D, Anacker C, Rex-Haffner M, Pruessner JC, Pariante CM, Pace TW, Mercer KB, Mayberg HS, Bradley B, et al. 2013. Allele-specific FKBP5 DNA demethylation mediates gene-childhood trauma interactions. Nat Neurosci 16:33-41.

Klose RJ, Bird AP. 2006. Genomic DNA methylation: the mark and its mediators. Trends Biochem Sci 31:89-97.

Kochmanski J, Montrose L, Goodrich JM, Dolinoy DC. 2017. Environmental Deflection: The Impact of Toxicant Exposures on the Aging Epigenome. Toxicol Sci 156:325-335.

Kovatsi L, Georgiou E, Ioannou A, Haitoglou C, Tzimagiorgis G, Tsoukali H, Kouidou S. 2010. p16 promoter methylation in Pb2+ -exposed individuals. Clin Toxicol (Phila) 48:124-128.

Krige DG. 1966. Two-Dimensional Weighted Moving Average Trend Surfaces For Ore Valuation Journal of South African Institute of Mining and Metallurgy.

Kuo LE, Kitlinska JB, Tilan JU, Li L, Baker SB, Johnson MD, Lee EW, Burnett MS, Fricke ST, Kvetnansky R, et al. 2007. Neuropeptide Y acts directly in the periphery on fat tissue and mediates stress-induced obesity and metabolic syndrome. Nat Med 13:803-811.

Kuznetsova A, Brockhoff PB, Christensen RHB. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. Journal of Statistical Software 82:1-26.

Lam LL, Emberly E, Fraser HB, Neumann SM, Chen E, Miller GE, Kobor MS. 2012. Factors underlying variable DNA methylation in a human community cohort. Proc Natl Acad Sci U S A 109 Suppl 2:17253-17260.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, et al. 2001. Initial sequencing and analysis of the human genome. Nature 409:860-921.

Landrigan PJ, Baker EL, Jr., Feldman RG, Cox DH, Eden KV, Orenstein WA, Mather JA, Yankel AJ, Von Lindern IH. 1976. Increased lead absorption with anemia and slowed nerve conduction in children near a lead smelter. J Pediatr 89:904-910.

Laughlin NK, Bowman RE, Franks PA, Dierschke DJ. 1987. Altered menstrual cycles in rhesus monkeys induced by lead. Fundam Appl Toxicol 9:722-729.

Leon-Velarde F, Maggiorini M, Reeves JT, Aldashev A, Asmus I, Bernardi L, Ge RL, Hackett P, Kobayashi T, Moore LG, et al. 2005. Consensus statement on chronic and subacute high altitude diseases. High Alt Med Biol 6:147-157.

Leon-Velarde F, McCullough RG, McCullough RE, Reeves JT, Group CMSCW. 2003. Proposal for scoring severity in chronic mountain sickness (CMS). Background and conclusions of the CMS Working Group. Adv Exp Med Biol 543:339-354.

Lerner RM, Benson JB. 2013. Introduction. In. Embodiment and Epigenesis: Theoretical and Methodological Issues in Understanding the Role of Biology within the Relational Developmental System - Part B: Ontogenetic Dimensions. p. 1-20.

Lessler MA, Walters MI. 1973. Erythrocyte osmotic fragility in the presence of lead or mercury. Proc Soc Exp Biol Med 142:548-553.

Li C, Yang X, Xu M, Zhang J, Sun N. 2013a. Epigenetic marker (LINE-1 promoter) methylation level was associated with occupational lead exposure. Clin Toxicol (Phila) 51:225-229.

Li C, Yang X, Xu M, Zhang J, Sun N. 2013b. Epigenetic marker (LINE-1 promoter) methylation level was associated with occupational lead exposure. Clinical toxicology (Philadelphia, Pa.) 51:225-229.

Li M, Quan C, Toth R, Campbell DG, MacKintosh C, Wang HY, Chen S. 2015. Fasting and Systemic Insulin Signaling Regulate Phosphorylation of Brain Proteins That Modulate Cell Morphology and Link to Neurological Disorders. J Biol Chem 290:30030-30041.

Li Y, Xie C, Murphy SK, Skaar D, Nye M, Vidal AC, Cecil KM, Dietrich KN, Puga A, Jirtle RL, et al. 2016. Lead Exposure during Early Human Development and DNA Methylation of Imprinted Gene Regulatory Elements in Adulthood. Environ Health Perspect 124:666-673.

Lidsky TI, Schneider JS. 2003. Lead neurotoxicity in children: basic mechanisms and clinical correlates. Brain 126:5-19.

Ling C, Ronn T. 2014. Epigenetic adaptation to regular exercise in humans. Drug Discov Today 19:1015-1018.

Liu J, Barrientos A. 2013. Transcriptional regulation of yeast oxidative phosphorylation hypoxic genes by oxidative stress. Antioxid Redox Signal 19:1916-1927.

Llanos AA, Chandwani S, Bandera EV, Hirshfield KM, Lin Y, Ambrosone CB, Demissie K. 2015. Associations between sociodemographic and clinicopathological factors and breast cancer subtypes in a population-based study. Cancer Causes Control 26:1737-1750.

Lorenzo FR, Huff C, Myllymaki M, Olenchock B, Swierczek S, Tashi T, Gordeuk V, Wuren T, Ri-Li G, McClain DA, et al. 2014. A genetic mechanism for Tibetan high-altitude adaptation. Nat Genet 46:951-956.

Lucock M, Yates Z, Martin C, Choi JH, Beckett E, Boyd L, LeGras K, Ng X, Skinner V, Wai R, et al. 2015. Methylation diet and methyl group genetics in risk for adenomatous polyp occurrence. BBA Clin 3:107-112.

Lv J, Xin Y, Zhou W, Qiu Z. 2013. The epigenetic switches for neural development and psychiatric disorders. J Genet Genomics 40:339-346.

Ma J, Jiang T, Tan L, Yu JT. 2015. TYROBP in Alzheimer's disease. Mol Neurobiol 51:820-826.

Ma X, Zhang H, Xue X, Shah YM. 2017. Hypoxia-inducible factor 2alpha (HIF-2alpha) promotes colon cancer growth by potentiating Yes-associated protein 1 (YAP1) activity. J Biol Chem 292:17046-17056.

Maccari S, Krugers HJ, Morley-Fletcher S, Szyf M, Brunton PJ. 2014. The consequences of early-life adversity: neurobiological, behavioural and epigenetic adaptations. J Neuroendocrinol 26:707-723.

Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. 2010. Robust relationship inference in genome-wide association studies. Bioinformatics 26:2867-2873.

Marioni RE, Shah S, McRae AF, Chen BH, Colicino E, Harris SE, Gibson J, Henders AK, Redmond P, Cox SR, et al. 2015. DNA methylation age of blood predicts all-cause mortality in later life. Genome Biol 16:25.

Marsit CJ. 2015. Influence of environmental exposure on human epigenetic regulation. J Exp Biol 218:71-79.

Matt L, Kirk LM, Chenaux G, Speca DJ, Puhger KR, Pride MC, Qneibi M, Haham T, Plambeck KE, Stern-Bach Y, et al. 2018. SynDIG4/Prrt1 Is Required for Excitatory Synapse Development and Plasticity Underlying Cognitive Function. Cell Rep 22:2246-2253.

Menezo YJ, Silvestris E, Dale B, Elder K. 2016. Oxidative stress and alterations in DNA methylation: two sides of the same coin in reproduction. Reprod Biomed Online 33:668-683.

Messerschmidt DM, Knowles BB, Solter D. 2014. DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. Genes Dev 28:812-828.

Mielke HW, Reagan PL. 1998. Soil is an important pathway of human lead exposure. Environ Health Perspect 106 Suppl 1:217-229.

Miriyala S, Gongora Nieto MC, Mingone C, Smith D, Dikalov S, Harrison DG, Jo H. 2006. Bone morphogenic protein-4 induces hypertension in mice: role of noggin, vascular NADPH oxidases, and impaired vasorelaxation. Circulation 113:2818-2825.

Mitteldorf JJ. 2013. How does the body know how old it is? Introducing the epigenetic clock hypothesis. Biochemistry (Mosc) 78:1048-1053.

Mohn F, Schubeler D. 2009. Genetics and epigenetics: stability and plasticity during cellular differentiation. Trends Genet 25:129-136.

Monge C. 1942. Life in the Andes and Chronic Mountain Sickness. Science 95:79-84.

Moore LG. 2001. Human genetic adaptation to high altitude. High Alt Med Biol 2:257-279.

Moore LG, Charles SM, Julian CG. 2011. Humans at high altitude: hypoxia and fetal growth. Respir Physiol Neurobiol 178:181-190.

Morgan HD, Santos F, Green K, Dean W, Reik W. 2005. Epigenetic reprogramming in mammals. Hum Mol Genet 14 Spec No 1:R47-58.

Morris TJ, Butcher LM, Feber A, Teschendorff AE, Chakravarthy AR, Wojdacz TK, Beck S. 2014. ChAMP: 450k Chip Analysis Methylation Pipeline. Bioinformatics 30:428-430.

Naeije R, Vanderpool R. 2013. Pulmonary hypertension and chronic mountain sickness. High Alt Med Biol 14:117-125.

Nanduri J, Semenza GL, Prabhakar NR. 2017. Epigenetic changes by DNA methylation in chronic and intermittent hypoxia. Am J Physiol Lung Cell Mol Physiol 313:L1096-L1100.

Natekar A, Olds RL, Lau MW, Min K, Imoto K, Slavin TP. 2014. Elevated blood pressure: Our family's fault? The genetics of essential hypertension. World J Cardiol 6:327-337.

Needleman HL, Gunnoe C, Leviton A, Reed R, Peresie H, Maher C, Barrett P. 1979. Deficits in psychologic and classroom performance of children with elevated dentine lead levels. N Engl J Med 300:689-695.

Needleman HL, Leviton A. 1979. Lead and neurobehavioural deficit in children. Lancet 2:104.

Nelson HH, Marsit CJ, Kelsey KT. 2011. Global methylation in exposure biology and translational medical science. Environ Health Perspect 119:1528-1533.

Non AL, Hollister BM, Humphreys KL, Childebayeva A, Esteves K, Zeanah CH, Fox NA, Nelson CA, Drury SS. 2016. DNA methylation at stress-related genes is associated with exposure to early life institutionalization. Am J Phys Anthropol 161:84-93.

Nossent AY, Hansen JL, Doggen C, Quax PH, Sheikh SP, Rosendaal FR. 2011. SNPs in microRNA binding sites in 3'-UTRs of RAAS genes influence arterial blood pressure and risk of myocardial infarction. Am J Hypertens 24:999-1006.

Ogino S, Nosho K, Kirkner GJ, Kawasaki T, Chan AT, Schernhammer ES, Giovannucci EL, Fuchs CS. 2008. A cohort study of tumoral LINE-1 hypomethylation and prognosis in colon cancer. J Natl Cancer Inst 100:1734-1738.

Padhy G, Gangwar A, Sharma M, Himashree G, Singh K, Bhaumik G, Bhargava K, Sethy NK. 2016. Plasma kallikrein-bradykinin pathway promotes circulatory nitric oxide metabolite availability during hypoxia. Nitric Oxide 55-56:36-44.

Papanikolaou NC, Hatzidaki EG, Belivanis S, Tzanakakis GN, Tsatsakis AM. 2005. Lead toxicity update. A brief review. Med Sci Monit 11:RA329-336.

Parsons PJ, Slavin W. 1993. A rapid Zeeman graphite furnace atomic absorption spectrometric method for the determination of lead in blood. Spectrochimica Acta Part B: Atomic Spectroscopy 48:925-939.

Pedersen BS, Schwartz DA, Yang IV, Kechris KJ. 2012a. Comb-p: software for combining, analyzing, grouping and correcting spatially correlated P-values. Bioinformatics 28:2986-2988.

Pedersen BS, Schwartz DA, Yang IV, Kechris KJ. 2012b. Comb-p: software for combining, analyzing, grouping and correcting spatially correlated P-values. Bioinformatics 28:2986-2988.

Penaloza D, Arias-Stella J. 2007. The heart and pulmonary circulation at high altitudes: healthy highlanders and chronic mountain sickness. Circulation 115:1132-1146.

Peñaloza D, Sime F, Banchero N, Gamboa R, Cruz J, Marticorena E. 1963. Pulmonary hypertension in healthy men born and living at high altitudes. American Journal of Cardiology 11:150-157.

Pennisi D, Gardner J, Chambers D, Hosking B, Peters J, Muscat G, Abbott C, Koopman P. 2000. Mutations in Sox18 underlie cardiovascular and hair follicle defects in ragged mice. Nat Genet 24:434-437.

Percy MJ, Beer PA, Campbell G, Dekker AW, Green AR, Oscier D, Rainey MG, van Wijk R, Wood M, Lappin TR, et al. 2008. Novel exon 12 mutations in the HIF2A gene associated with erythrocytosis. Blood 111:5400-5402.

Perna L, Zhang Y, Mons U, Holleczek B, Saum KU, Brenner H. 2016. Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort. Clin Epigenetics 8:64.

Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P, Van Djik S, Muhlhausler B, Stirzaker C, Clark SJ. 2016. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. Genome Biol 17:208.

Pogribny IP, Tryndyak VP, Woods CG, Witt SE, Rusyn I. 2007. Epigenetic effects of the continuous exposure to peroxisome proliferator WY-14,643 in mouse liver are dependent upon peroxisome proliferator activated receptor alpha. Mutat Res 625:62-71.

Popovic M, McNeill FE, Chettle DR, Webber CE, Lee CV, Kaye WE. 2005. Impact of occupational exposure on lead levels in women. Environ Health Perspect 113:478-484.

Pottier G, Viau M, Ricoul M, Shim G, Bellamy M, Cuceu C, Hempel WM, Sabatier L. 2013. Lead Exposure Induces Telomere Instability in Human Cells. PLoS One 8:e67501.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559-575.

Qu CS, Ma ZW, Yang J, Liu Y, Bi J, Huang L. 2012. Human exposure pathways of heavy metals in a lead-zinc mining area, Jiangsu Province, China. PLoS One 7:e46793.

Quintanar-Escorza MA, Gonzalez-Martinez MT, Navarro L, Maldonado M, Arevalo B, Calderon-Salinas JV. 2007. Intracellular free calcium concentration and calcium transport in human erythrocytes of lead-exposed workers. Toxicol Appl Pharmacol 220:1-8.

Rademaker K, Hodgins G, Moore K, Zarrillo S, Miller C, Bromley GR, Leach P, Reid DA, Alvarez WY, Sandweiss DH. 2014. Paleoindian settlement of the high-altitude Peruvian Andes. Science 346:466-469.

Raposinho PD, Pedrazzini T, White RB, Palmiter RD, Aubert ML. 2004. Chronic neuropeptide Y infusion into the lateral ventricle induces sustained feeding and obesity in mice lacking either Npy1r or Npy5r expression. Endocrinology 145:304-310.

Rawluszko-Wieczorek AA, Horbacka K, Krokowicz P, Misztal M, Jagodzinski PP. 2014. Prognostic potential of DNA methylation and transcript levels of HIF1A and EPAS1 in colorectal cancer. Mol Cancer Res 12:1112-1127.

Reeves JT, Grover RF. 2005. Insights by Peruvian scientists into the pathogenesis of human chronic hypoxic pulmonary hypertension. J Appl Physiol (1985) 98:384-389.

Reik W. 2007. Stability and flexibility of epigenetic gene regulation in mammalian development. Nature 447:425-432.

Remillard CV, Tigno DD, Platoshyn O, Burg ED, Brevnova EE, Conger D, Nicholson A, Rana BK, Channick RN, Rubin LJ, et al. 2007. Function of Kv1.5 channels and genetic variations of KCNA5 in patients with idiopathic pulmonary arterial hypertension. Am J Physiol Cell Physiol 292:C1837-1853.

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43:e47.

Ronn T, Volkov P, Davegardh C, Dayeh T, Hall E, Olsson AH, Nilsson E, Tornberg A, Dekker Nitert M, Eriksson KF, et al. 2013. A six months exercise intervention influences the genome-wide DNA methylation pattern in human adipose tissue. PLoS Genet 9:e1003572.

Roy A, Hu H, Bellinger DC, Mukherjee B, Modali R, Nasaruddin K, Schwartz J, Wright RO, Ettinger AS, Palaniapan K, et al. 2011. Hemoglobin, lead exposure, and intelligence quotient: effect modification by the DRD2 Taq IA polymorphism. Environ Health Perspect 119:144-149.

Rupert JL, Hochachka PW. 2001. Genetic approaches to understanding human adaptation to altitude in the Andes. J Exp Biol 204:3151-3160.

Sanders T, Liu Y, Buchner V, Tchounwou PB. 2009. Neurotoxic effects and biomarkers of lead exposure: a review. Rev Environ Health 24:15-45.

Sandoval JR, Salazar-Granara A, Acosta O, Castillo-Herrera W, Fujita R, Pena SD, Santos FR. 2013. Tracing the genomic ancestry of Peruvians reveals a major legacy of pre-Columbian ancestors. J Hum Genet 58:627-634.

Schneider JS, Anderson DW, Talsania K, Mettil W, Vadigepalli R. 2012. Effects of developmental lead exposure on the hippocampal transcriptome: influences of sex, developmental period, and lead exposure level. Toxicol Sci 129:108-125.

Schneider JS, Kidd SK, Anderson DW. 2013. Influence of developmental lead exposure on expression of DNA methyltransferases and methyl cytosine-binding proteins in hippocampus. Toxicol Lett 217:75-81.

Schober A. 2016. Establishment and characterisation of USP29 as a novel regulator of Hypoxia Inducible Factor α. [University of Liverpool.

Semenza GL. 2001a. Hypoxia-inducible factor 1: control of oxygen homeostasis in health and disease. Pediatr Res 49:614-617.

Semenza GL. 2001b. Hypoxia-inducible factor 1: oxygen homeostasis and disease pathophysiology. Trends Mol Med 7:345-350.

Sen A, Heredia N, Senut MC, Land S, Hollocher K, Lu X, Dereski MO, Ruden DM. 2015. Multigenerational epigenetic inheritance in humans: DNA methylation changes associated with maternal exposure to lead can be transmitted to the grandchildren. Sci Rep 5:14466.

Senut MC, Cingolani P, Sen A, Kruger A, Shaik A, Hirsch H, Suhr ST, Ruden D. 2012. Epigenetics of early-life lead exposure and effects on brain development. Epigenomics 4:665-674.

Senut MC, Sen A, Cingolani P, Shaik A, Land SJ, Ruden DM. 2014. Lead exposure disrupts global DNA methylation in human embryonic stem cells and alters their neuronal differentiation. Toxicol Sci 139:142-161.

Seow WJ, Kile ML, Baccarelli AA, Pan WC, Byun HM, Mostofa G, Quamruzzaman Q, Rahman M, Lin X, Christiani DC. 2014. Epigenome-wide DNA methylation changes with development of arsenic-induced skin lesions in Bangladesh: a case-control follow-up study. Environ Mol Mutagen 55:449-456.

Serradeil-Le Gal C, Lafontan M, Raufaste D, Marchand J, Pouzet B, Casellas P, Pascal M, Maffrand JP, Le Fur G. 2000. Characterization of NPY receptors controlling lipolysis and leptin secretion in human adipocytes. FEBS Lett 475:150-156.

Simonson TS, Yang Y, Huff CD, Yun H, Qin G, Witherspoon DJ, Bai Z, Lorenzo FR, Xing J, Jorde LB, et al. 2010. Genetic evidence for high-altitude adaptation in Tibet. Science 329:72-75.

Sivaraja V, Kumar TK, Rajalingam D, Graziani I, Prudovsky I, Yu C. 2006. Copper binding affinity of S100A13, a key component of the FGF-1 nonclassical copper-dependent release complex. Biophys J 91:1832-1843.

Stenmark KR, Frid MG, Yeager M, Li M, Riddle S, McKinsey T, El Kasmi KC. 2012. Targeting the adventitial microenvironment in pulmonary hypertension: A potential approach to therapy that considers epigenetic change. Pulm Circ 2:3-14.

Stewart WF, Schwartz BS, Davatzikos C, Shen D, Liu D, Wu X, Todd AC, Shi W, Bassett S, Youssem D. 2006. Past adult lead exposure is linked to neurodegeneration measured by brain MRI. Neurology 66:1476-1484.

Suliman HB, Ali M, Piantadosi CA. 2004. Superoxide dismutase-3 promotes full expression of the EPO response to hypoxia. Blood 104:43-50.

Szyf M, Bick J. 2013. DNA methylation: a mechanism for embedding early life experiences in the genome. Child Dev 84:49-57.

Tajuddin SM, Amaral AF, Fernandez AF, Rodriguez-Rodero S, Rodriguez RM, Moore LE, Tardon A, Carrato A, Garcia-Closas M, Silverman DT, et al. 2013. Genetic and non-genetic predictors of LINE-1 methylation in leukocyte DNA. Environ Health Perspect 121:650-656.

Takasawa K, Arai Y, Yamazaki-Inoue M, Toyoda M, Akutsu H, Umezawa A, Nishino K. 2018. DNA hypermethylation enhanced telomerase reverse transcriptase expression in human-induced pluripotent stem cells. Hum Cell 31:78-86.

Teschendorff AE. 2013. Epigenetic aging: insights from network biology. Aging (Albany NY) 5:719-720.

Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, Beck S. 2013. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics 29:189-196.

Teschendorff AE, Menon U, Gentry-Maharaj A, Ramus SJ, Gayther SA, Apostolidou S, Jones A, Lechner M, Beck S, Jacobs IJ, et al. 2009. An epigenetic signature in peripheral blood predicts active ovarian cancer. PLoS One 4:e8274.

Teschendorff AE, West J, Beck S. 2013. Age-associated epigenetic drift: implications, and a case of epigenetic thrift? Hum Mol Genet 22:R7-R15.

Tian H, McKnight SL, Russell DW. 1997. Endothelial PAS domain protein 1 (EPAS1), a transcription factor selectively expressed in endothelial cells. Genes Dev 11:72-82.

Tollefsbol TO. 2011. Advances in epigenetic technology. Methods Mol Biol 791:1-10.

Tong S, von Schirnding YE, Prapamontol T. 2000. Environmental lead exposure: a public health problem of global dimensions. Bull World Health Organ 78:1068-1077.

Trinh BN, Ong CN, Coetzee GA, Yu MC, Laird PW. 2002. Thymidylate synthase: a novel genetic determinant of plasma homocysteine and folate levels. Hum Genet 111:299-302.

van Bemmel D, Lenz P, Liao LM, Baris D, Sternberg LR, Warner A, Johnson A, Jones M, Kida M, Schwenn M, et al. 2012. Correlation of LINE-1 methylation levels in patient-matched buffy coat, serum, buccal cell, and bladder tumor tissue DNA samples. Cancer Epidemiol Biomarkers Prev 21:1143-1148.

van Geen A, Bravo C, Gil V, Sherpa S, Jack D. 2012. Lead exposure from soil in Peruvian mining towns: a national assessment supported by two contrasting examples. Bull World Health Organ 90:878-886.

Vaupel P, Mayer A. 2007. Hypoxia in cancer: significance and impact on clinical outcome. Cancer Metastasis Rev 26:225-239.

Vidal AC, Semenova V, Darrah T, Vengosh A, Huang Z, King K, Nye MD, Fry R, Skaar D, Maguire R, et al. 2015. Maternal cadmium, iron and zinc levels, DNA methylation and birth weight. BMC Pharmacol Toxicol 16:20.

Virani S, Dolinoy DC, Halubai S, Jones TR, Domino SE, Rozek LS, Nahar MS, Padmanabhan V. 2012. Delivery type not associated with global methylation at birth. Clin Epigenetics 4:8.

Wang J, Juhaszova M, Rubin LJ, Yuan XJ. 1997. Hypoxia inhibits gene expression of voltage-gated K+ channel alpha subunits in pulmonary artery smooth muscle cells. J Clin Invest 100:2347-2353.

Wang L, Rao F, Zhang K, Mahata M, Rodriguez-Flores JL, Fung MM, Waalen J, Cockburn MG, Hamilton BA, Mahata SK, et al. 2009. Neuropeptide Y(1) Receptor NPY1R discovery of naturally occurring human genetic variants governing gene expression in cella as well as pleiotropic effects on autonomic activity and blood pressure in vivo. J Am Coll Cardiol 54:944-954.

Wangsri S, Subbalekha K, Kitkumthorn N, Mutirangura A. 2012. Patterns and possible roles of LINE-1 methylation changes in smoke-exposed epithelia. PLoS One 7:e45292.

Ward M, Milledge JS, West JB. 2000. High altitude medicine and physiology. London: Arnold.

Watson JA, Watson CJ, McCann A, Baugh J. 2010. Epigenetics, the epicenter of the hypoxic response. Epigenetics 5:293-296.

Watson JA, Watson CJ, McCrohan AM, Woodfine K, Tosetto M, McDaid J, Gallagher E, Betts D, Baugh J, O'Sullivan J, et al. 2009. Generation of an epigenetic signature by chronic hypoxia in prostate cells. Hum Mol Genet 18:3594-3604.

Weaver IC, Cervoni N, Champagne FA, D'Alessio AC, Sharma S, Seckl JR, Dymov S, Szyf M, Meaney MJ. 2004. Epigenetic programming by maternal behavior. Nat Neurosci 7:847-854.

Weitzman SA, Turk PW, Milkowski DH, Kozlowski K. 1994. Free radical adducts induce alterations in DNA cytosine methylation. Proc Natl Acad Sci U S A 91:1261-1264.

Wernimont SM, Clark AG, Stover PJ, Wells MT, Litonjua AA, Weiss ST, Gaziano JM, Tucker KL, Baccarelli A, Schwartz J, et al. 2011. Folate network genetic variation, plasma homocysteine, and global genomic methylation content: a genetic association study. BMC Med Genet 12:150.

WHO. 2011. Brief guide to analytical methods for measuring lead in blood.

Wickham H. 2009. ggplot2: Elegant Graphics for Data Analysis. Ggplot2: Elegant Graphics for Data Analysis:1-212.

Wilson WR, Hay MP. 2011. Targeting hypoxia in cancer therapy. Nat Rev Cancer 11:393-410.

Wipff J, Dieude P, Guedj M, Ruiz B, Riemekasten G, Cracowski JL, Matucci-Cerinic M, Melchers I, Humbert M, Hachulla E, et al. 2010. Association of a KCNA5 gene polymorphism with systemic sclerosis-associated pulmonary arterial hypertension in the European Caucasian population. Arthritis Rheum 62:3093-3100.

Wolffe AP, Guschin D. 2000. Review: chromatin structural features and targets that regulate transcription. J Struct Biol 129:102-122.

Wong AH, Gottesman, II, Petronis A. 2005. Phenotypic differences in genetically identical organisms: the epigenetic perspective. Hum Mol Genet 14 Spec No 1:R11-18.

Woo HD, Kim J. 2012. Global DNA hypomethylation in peripheral blood leukocytes as a biomarker for cancer risk: a meta-analysis. PLoS One 7:e34615.

Wright RO, Schwartz J, Wright RJ, Bollati V, Tarantini L, Park SK, Hu H, Sparrow D, Vokonas P, Baccarelli A. 2010a. Biomarkers of lead exposure and DNA methylation within retrotransposons. Environ Health Perspect 118:790-795.

Wright RO, Schwartz J, Wright RJ, Bollati V, Tarantini L, Park SK, Hu H, Sparrow D, Vokonas P, Baccarelli A. 2010b. Biomarkers of lead exposure and DNA methylation within retrotransposons. Environmental health perspectives 118:790-795.

Wu T, Buck GM, Mendola P. 2003. Blood lead levels and sexual maturation in U.S. girls: the Third National Health and Nutrition Examination Survey, 1988-1994. Environ Health Perspect 111:737-741.

Wu Y, Liu Y, Ni N, Bao B, Zhang C, Lu L. 2012. High lead exposure is associated with telomere length shortening in Chinese battery manufacturing plant workers. Occup Environ Med 69:557-563.

Wuren T, Simonson TS, Qin G, Xing J, Huff CD, Witherspoon DJ, Jorde LB, Ge RL. 2014. Shared and unique signals of high-altitude adaptation in geographically distinct Tibetan populations. PLoS One 9:e88252.

Xing J, Wuren T, Simonson TS, Watkins WS, Witherspoon DJ, Wu W, Qin G, Huff CD, Jorde LB, Ge RL. 2013. Genomic analysis of natural selection and phenotypic variation in high-altitude mongolians. PLoS Genet 9:e1003634.

Xu D, Guo H, Xu X, Lu Z, Fassett J, Hu X, Xu Y, Tang Q, Hu D, Somani A, et al. 2011. Exacerbated pulmonary arterial hypertension and right ventricular hypertrophy in animals with loss of function of extracellular superoxide dismutase. Hypertension 58:303-309.

Xu XF, Cheng F, Du LZ. 2011. Epigenetic regulation of pulmonary arterial hypertension. Hypertens Res 34:981-986.

Yi X, Liang Y, Huerta-Sanchez E, Jin X, Cuo ZX, Pool JE, Xu X, Jiang H, Vinckenbosch N, Korneliussen TS, et al. 2010. Sequencing of 50 human exomes reveals adaptation to high altitude. Science 329:75-78.

Zannas AS, Arloth J, Carrillo-Roa T, Iurato S, Roh S, Ressler KJ, Nemeroff CB, Smith AK, Bradley B, Heim C, et al. 2015. Lifetime stress accelerates epigenetic aging in an urban, African American cohort: relevance of glucocorticoid signaling. Genome Biol 16:266.

Zelko IN, Mariani TJ, Folz RJ. 2002. Superoxide dismutase multigene family: a comparison of the CuZn-SOD (SOD1), Mn-SOD (SOD2), and EC-SOD (SOD3) gene structures, evolution, and expression. Free Radic Biol Med 33:337-349.

Zhang FF, Cardarelli R, Carroll J, Fulda KG, Kaur M, Gonzalez K, Vishwanatha JK, Santella RM, Morabia A. 2011. Significant differences in global genomic DNA methylation by gender and race/ethnicity in peripheral blood. Epigenetics 6:623-629.

Zhang FF, Morabia A, Carroll J, Gonzalez K, Fulda K, Kaur M, Vishwanatha JK, Santella RM, Cardarelli R. 2011. Dietary patterns are associated with levels of global genomic DNA methylation in a cancer-free population. J Nutr 141:1165-1171.

Zhang J, Zhou J, Cai L, Lu Y, Wang T, Zhu L, Hu Q. 2012. Extracellular calcium-sensing receptor is critical in hypoxic pulmonary vasoconstriction. Antioxid Redox Signal 17:471-484.

Zheng YM, Wang QS, Rathore R, Zhang WH, Mazurkiewicz JE, Sorrentino V, Singer HA, Kotlikoff MI, Wang YX. 2005. Type-3 ryanodine receptors mediate hypoxia-, but not neurotransmitter-induced calcium release and contraction in pulmonary artery smooth muscle cells. J Gen Physiol 125:427-440.

Zhou D, Wang J, Zapala MA, Xue J, Schork NJ, Haddad GG. 2008. Gene expression in mouse brain following chronic hypoxia: role of sarcospan in glial cell death. Physiol Genomics 32:370-379.

Zhu W, Tilley DG, Myers VD, Tsai EJ, Feldman AM. 2014. Increased vasopressin 1A receptor expression in failing human hearts. J Am Coll Cardiol 63:375-376.

Ziello JE, Jovin IS, Huang Y. 2007. Hypoxia-Inducible Factor (HIF)-1 regulatory pathway and its potential for therapeutic intervention in malignancy and ischemia. Yale J Biol Med 80:51-60.

Zukowska-Grojec Z, Marks ES, Haass M. 1987. Neuropeptide Y is a potent vasoconstrictor and a cardiodepressant in rat. Am J Physiol 253:H1234-1239.

Zukowska-Grojec Z, Vaz AC. 1988. Role of neuropeptide Y (NPY) in cardiovascular responses to stress. Synapse 2:293-298.