Identification of genes responsible for the variation in facial and teeth morphology in Latin Americans

A Thesis submitted for the Degree of *Doctor of Philosophy*

Author:

Macarena Fuentes-Guajardo

Supervisors:

Prof. Andrés Ruiz-Linares

Prof. Claudio Stern

Department of Genetics, Evolution and Environment

University College London

London, United Kingdom

2019

Declaration

I, Macarena Fuentes-Guajardo confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Publications arising from this thesis:

- Adhikari K & **Fuentes-Guajardo** M et al., (2016). A genome-wide association scan implicates DCHS2, RUNX2, GLI3, PAX1 and EDAR in human facial variation. *Nature Communications.* 7:11616. doi: 10.1038/ncomms11616.

Publications not directly related to this thesis:

- Adhikari K, Mendoza-Revilla J, Chacón-Duque JC, **Fuentes-Guajardo** M & Ruiz-Linares A, (2016). Admixture in Latin America. *Current Opinion in Genetics & Development.* Volume 4, Pages 106-114. doi: 10.1016/j.gde.2016.09.003.

- Lorenzo Bermejo J, Boekstegers F, González Silos R, Marcelain K, **Fuentes-Guajardo** M et al., (2017). Subtypes of Native American ancestry and leading causes of death: Mapuche ancestry-specific associations with gallbladder cancer risk in Chile. *PLoS Genetics* 13:e1006756. doi: 10.1371/journal.pgen.1006756.

- Rothhammer F, Puddu G and **Fuentes-Guajardo M**, (2016). Can mitocondrial DNA provide information on the ethnogenesis of Chilean native populations? *Revista de Antropología Chilena Chungará.* Volume 49, Pages 635-642. doi:10.4067/S0717-73562017005000028.

- Adhikari K, Chacón-Duque JC, Mendoza-Revilla J, **Fuentes-Guajardo M** & Ruiz-Linares A, (2017). The Genetic Diversity of the Americas. *Annual Review of Genomics and Human Genetics.* Volume 18, Pages 277-296. doi: 10.1146/annurev-genom-083115-022331.

- Quinto-Sánchez M, Muñoz-Muñoz, Gomez-Valdes, Cintas, Navarro, **Fuentes-Guajardo** et al., (2018). Developmental pathways inferred from modularity, morphological integration and fluctuating asymmetry patterns in the human face. *Scientific Reports.* 8:963. doi:10.1038/s41598-018-19324-y.

- Chacón-Duque JC, Adhikari K, **Fuentes-Guajardo M**, Mendoza-Revilla J, Acuña-Alonzo V, Barquera Lozano R et al., (2018). Latin Americans show wide-spread Converso ancestry and imprint of local Native ancestry on

physical appearance. *Nature Communications.* 9:5389. doi:10.1038/s41467-018-07748-z.

- Carayon D, Adhikari K, Monsarrat P, Dumoncel J, Braga J, **Fuentes-Guajardo M** et al., (2019). A geometric morphometric approach to the study of shovel-shaped incisors variation. *American Journal of Physical Anthropology.* Volume 168, Pages 229-241. doi:10.1002/ajpa.23709.

- Delgado M, Ramírez LM, Adhikari K, **Fuentes-Guajardo M**, Zanolli C, Gonzalez-José R et al., (2019). Variation in dental morphology and inference of continental ancestry in admixed Latin Americans. *American Journal of Physical Anthropology.* Volume 168, Pages 438-447. doi:10.1002/ajpa.23756.

- Adhikari K, Mendoza-Revilla J, Sohail A, **Fuentes-Guajardo M**, Lampert, Chacón-Duque JC et al., (2019). A Genome Wide Association Scan in Latin Americans underlines the convergent evolution of lighter skin pigmentation in Eurasia. *Nature Communications.* 10:358. doi:10.1038/s41467-018-08147-0.

Statement of work

Any joint work with colleagues and supervisors reported in this thesis is clearly stated within each chapter and section. All the figures taken or adapted from published material are part of the paper in which I am first author.

Chapter 1:

Parts of the text presented in Chapter 1, Section 1.5 is based on and adapted from two reviews our group published. These scientific publications were written by Dr. Kaustubh Adhikari, Dr Juan Camilo Chacón-Duque, Dr. Javier Mendoza-Revilla, Professor Andrés Ruiz-Linares and me. This work is published in Adhikari et al. (2016) and Adhikari et al. (2017).

Chapter 2:

Methods refer to in this chapter were described exclusively by me.

Chapter 3:

As a result of the work in this chapter there is a scientific publication. Most of the work here was undertaken by Dr. Kaustubh Adhikari, Dr. Javier Mendoza-Revilla, Juan Camilo Chacón-Duque, Professor Andrés Ruiz-Linares and me. This work is published in Adhikari et al. (2016) and uses text adapted from it and some originally written by me and colleagues.

Chapter 4:

Work in Chapter 3 was mainly undertaken exclusively by me under the guidance of Dr. Kaustubh Adhikari.

Dataset contribution

Genotyping and phenotypig data from admixed Latin Americans were collected by me in Chile and colleagues from the Consortium for the Analysis and Diversity and Evolution (CANDELA) in Brazil, Colombia, Mexico and Perú. I am very grateful to those involved in the sample collection, documentation and processing of the genomic and phenotypic data. I am also deeply grateful to all the volunteers who agreed to take part in the CANDELA study. The rest of the samples were obtained from publicly available datasets as acknowledged in the text.

Acknowledgments

First, I would like to thank Professor Francisco Rothammer Engel, who opened doors for me to research and allowed me to be part of the Consortium for the Analysis and Diversity and Evolution (CANDELA) in Chile, giving me the opportunity to be part of an excellent research team. Second, I would like to thank my first supervisor, Professor Andres Ruiz-Linares, who gave me the opportunity to do my PhD at UCL, where I have obtained an unparalleled experience, not only in my academic life but also in my personal life. In addition, I would like to thank my third supervisor Professor Alison Hardcastle, who kindly offered her support from the beginning of my PhD and a special thanks to my second supervisor, Professor Claudio D. Stern, who has supervised me not only during my PhD, but also in my future as a researcher. He also received me in his laboratory, where I met great people, Dr. Nidia Oliveira, who offered me his unconditional support and help from the first time we saw each other. A great thank you to Dr. Kaustubh Adhikari, he was a key person in the development of my PhD, under his academic guidance I obtained most of the knowledge required to finish the PhD successfully. Also, I would like to thank my great friends in London, Daniela Giambruno, Dr. Juan Camilo Chacon-Duque and Javier Mendoza-Revilla, who made this experience unique. I would also like to thank a special person, Professor Ismelda Lobato, who has believed in me as a researcher since undergraduate. In addition, I would like to thank two institutions in Chile, University of Tarapaca and CONICYT (Programa de Becas de Docotrado en el extranjero – Becas Chile), thanks to the economic support of these two institutions I could study in London-UK. Finally, I would like to thank my family, especially my parents, because without their constant example and teaching I would not reach my goals in life and my husband for his eternal and unconditional support in each of the things I undertake. Muchas gracias!

Abstract

Facial and dental features are of considerable importance in biomedicine and forensics. Facial appearance has a strong genetic component and could have evolved to facilitate individual recognition. Teeth are the hardest and well-preserved parts of the body and they have been used to establish biological relatedness among past and current human populations and to identify individuals. Although genes have been identified for various facial and dental phenotypes, the genetic basis of normal variation for both traits are still poorly understood. I performed Genome Wide Association Studies (GWAS) using ~700,000 genome-wide markers from ~6,000 Latin American individuals (CANDELA cohort). Ordinal and quantitative facial traits were assessed in individual photographs. Single Nucleotide Polymorphisms (SNPs) situated in four gene regions showed associations with three ordinal and quantitative traits related to nose morphology. Quantitative analyses, in addition, detected an association of SNPs in the Ectodysplasin A receptor (*EDAR*) gene with chin protrusion. Subsequently, statistical and experimental follow-up analysis were performed to endorse the discovered significant associations. Consistently, *Edar* mouse mutants were characterized to observe alterations of mandible length. Subsequently, I conducted GWAS for dental traits using the same markers from a subgroup of ~500 volunteers from the same cohort. Eighty-six traits were scored using the Arizona State University Dental Anthropology System (ASUDAS) scale. In addition, inciso-cervical, mesiodistal and bucco-lingual distances were measured on the incisors, canines and premolars in the same photos. Eleven of the categorical traits examined showed genome-wide significant association with SNPs in at least one genomic region and seven measurements showed genome-wide significant association with SNPs in four genomic regions. Ten of the genomic regions detected have been associated for other dental GWAS.

Impact Statement

Over the last fifteen years, along with development of technology there has been an increase in the number of efforts to obtain more genomic data helping to broaden the knowledge in human populations. Genome-wide scans have played an important role in this task, it is an experimental design to detect associations between genetic variants and traits, primarily to better understand the biology behind of diseases and complex traits, assuming that increasing the knowledge will lead us to prevention or better treatment toward personalized medicine. In this PhD thesis I use the same rationale to detect new genomic variants associated to facial and dental features. I apply current methods to detect genetic associations related to physical appearance in a Latin American population. Moreover, I perform statistical and experimental analysis to deepen in the biology behind the genetic association and the traits. This work has several implications. First, most of the genome-wide studies have been conducted in European populations, this study has been performed in an underrepresented population of Latin Americans. Reflecting several advantages of working with a population so diverse in both phenotypic and genetic variability. Allowing to find not just novel genetic variants associated to the studied traits in Latin Americans, but also in the parental populations – Native American, European and African populations. Finally, detecting new genomic regions associated to normal physical appearance offers an opportunity to increase the knowledge of physical abnormalities, due to the shared role between normal physical features-associated loci and many type of syndromes and other anomalies, and to forensic applications through the development and application of facial and dental phenotype prediction based on DNA variants.

Contents

List of Tables

List of Figures

Acronyms

1KG 1000 Genomes Project.

aDNA ancient DNA.

BMI Body Mass Index.

CANDELA Consortium for the Analysis of the Diversity and Evolution of Latin America.

DNA DeoxyriboNucleic Acid.

GWAS Genome Wide Association Study.

HGDP Human Genome Diversity Panel.

HLA Human Leukocyte Antigen.

IBD Identity by Descent.

IBS Identity by State.

Kb Kilo-base pair.

Mb Mega-base

Kya Thousand years ago.

LD Linkage Disequilibrium.

LGM Last Glacial Maximum.

MAF Minor Allele Frequency.

MHC Major Histocompatibility Complex.

mtDNA mitochondrial DNA.

PCA Principal Component Analysis.

QC Quality Control.

SNP Single Nucleotide Polymorphism.

WC Waist Circumference.

WES Whole Exome Sequencing.

WHR Waist to Hip Ratio.

WGS Whole Genome Sequencing.

NCL Non-syndromic cleft-lip

NCL/P Non syndromic cleft lip and palate

CL Cleft lip

MZ Monozygotic twins

DZ Dizygotic Twins

NFBC Northern Finland Birth Cohorts

ALSPAC Avon Longitudinal Study of Parents and Children

DNBC Danish National Birth Cohort

ASUDAS Arizona State University Dental Anthropology System

CNCC Cranial neural crest cells

STR Short tandem repeats

CNV Copy number variants

EK Enamel knot

2D Two dimensions

3D Three dimensions

FS Forensic Science

FDP Forensic DNA phenotyping

AS-PCA Ancestry Specific Principal Component Analysis

WTCC Wellcome Trust Case Control

VIF Variance inflation factor

GM Geometric momrphometry

3D-CT Three dimensions computer tomography

ICC Intra class correlation

HWE Hardy Weinberg equilibrium

GRM Gentic relatedness matrix

HAPMAP Haplotype map project

PRS Poligenic risk score

HED Hypohidrotic Ectodermal Dysplasia

CEPH-HGDP

AI Artificial intelligence

MRI Magnetic resonance imaging

NAs Native Americans

AFR Africans

ACB African Caribbeans in Barbados

ASW Americans of African Ancestry in SW USA

ESN Esan in Nigeria

GWD Gambian in Western Divisions in the Gambia

LWK Luhya in Webuye, Kenya

MSL Mende in Sierra Leone

YRI Yoruba in Ibadan, Nigeria

AMR Americans

CLM Colombians from Medellin, Colombia

MXL Mexican Ancestry from Los Angeles USA

PEL Peruvians from Lima, Perú

PUR Puerto Ricans from Puerto Rico

EAS East Asians

CDX Chinese Dai in Xishuangbanna, China

CHB Han Chinese in Beijing, China

CHS Southern Han Chinese

JPT Japanese in Tokyo, Japan

KHV Kinh in Ho Chi Minh City, Vietnam

EUR Europeans

CEU Utah Residents (CEPH) with Northern and Western European Ancestry

FIN Finnish in Finland

GBR British in England and Scotland

IBS Iberian Population in Spain

TSI Toscani in Italia

Chapter 1: Introduction

1.1 Overview of this thesis

The human body has always captured humanity's attention, not just because it is the expression of what we are as human beings, but also because of our curiosity to understand how it works, how to fix it if necessary and, how to foresee how it will change.

Among human body features, facial and dental traits represent key elements. The development of the craniofacial region is a complex process that reflects the evolutionary forces that control morphogenesis. The craniofacial region in vertebrates protects the brain and houses the sensory organs and structures from the digestive system and, it also allows us to recognize unique individual traits [1]. Therefore, the survival of various species of vertebrates depends on the ability to perceive and devour their prey. Hence, the variations in anatomy and function of the craniofacial complex are one of the best evidences of evolutionary adaptation, thus, they are under permanent study [2].

Genetics is crucial in this endeavour because Deoxynucleic Acid (DNA) determines to a certain extent how craniofacial features will look like. The most emblematic example of this are monozygotic twins (MZ) [3], i.e. MZ twins despite being raised apart, still look very similar [4], reflecting that the genetic factor influence is stronger than the environmental effects. DNA will also control if there is genetic susceptibility to develop certain illnesses, or if there are advantages to adapt to different environments. Thus, if we can predict by using DNA analysis, we could help doctors with reconstructive surgeries, forensic scientists could draft the face of a criminal, and historians, anthropologists and archaeologists can reconstruct craniofacial traits from individuals from the past [5].

These challenges seem unattainable, but with the advent of new technologies, such as sequencing and large-scale genotyping platforms this may be possible in the future. Thousands of genome wide association scans have been conducted seeking genomic

regions associated with different diseases/traits [6], but not many related to facial and dental traits [7-10].

Although genes have been identified for various facial and dental phenotypes, the genetic basis of normal variation for both traits are still poorly understood. In this thesis I will present principally genome wide association analyses seeking genomic variants responsible for the variation in facial and dental morphology in Latin Americans, and statistical and experimental follow-up analysis to gain a deeper understanding of the genetic impact on the phenotypes. I have performed GWAS applying standard procedures used on genome-wide association studies. First, using genotyping data obtained from DNA microarrays, and later, utilizing imputed data to increase the power of the genome-wide scans [11].

Additionally, I have also performed a replication GWAS in the same population and other statistical, and experimental analysis to obtain deeper knowledge with regards of the biological mechanism behind some of the identified genetic associations.

1.2 The human face

Undoubtedly, one of the most significant parts of the body is the face, its features define the identity of a person. Facial physiognomies are developed from infancy until adulthood, but the main characteristics are retained, thus the face becomes the trademark of the human being.

The biological definition of face is the frontal part of the head, from the forehead until the chin, where the eyes, nose and mouth are situated. This concept sounds simple, but involves several aspects of the human being, in both biological and social functioning, such as human behavioural biology, genomics, life history, evolutionary psychology, biological anthropology, among other topics. Ultimately, all of them together became part of an evolutionary theory, that natural and social environments have modulated, by selective forces, the human body and face shapes.

1.2.1 Development embryology and anatomy of the face

1.2.1.1 Development embryology of the face

The beginning of facial development embryology occurs during the 4th week of the human embryo's development. This process will finish into the 12th week, with the completion of the soft palate [12].

Facial development comprises five prominences: frontonasal, two maxillary and two mandibular prominences. All five facial swellings form by the end of 4th week. These initially surround the primitive oral cavity, the stomodeum. During the 5th week, the paired maxillary prominences enlarge and grow ventrally and medially. Simultaneously, a pair of ectodermal thickenings called the nasal placodes form on the frontonasal process and begin to enlarge. In the 6th week, the ectoderm at the centre of each nasal placode invaginates to form an oval nasal pit, dividing the frontonasal prominence into the lateral and medial nasal prominence into the lateral and medial nasal processes [12].

During the 6th week, the medial nasal processes approximate toward the midline and join to form the primordium of the bridge and septum of the nose. By the end of the 7th week, the inferior tips of the medial nasal processes expand laterally and inferiorly and join to form the intermaxillary process [12].

Although the two mandibular prominences seem to be separated by a fissure midventrally, they form in continuity with each other, like the rest of the pharyngeal arches [12].

The formation of each region of the face and neck is mainly due to the migration of the neural crest cells which come from the ectoderm [12].

Neural crest cells (NCCs) are derived from dorsal midline ectoderm of developing vertebrate embryos. Neural crest cells migrate downwards beside the neural tube and laterally under the surface ectoderm at all axial levels. The cells formed at craniofacial levels migrate into the frontonasal process and pharyngeal arches and give rise to nearly all tissues of the face and neck. After cranial neural crest cells have completed their migration, facial growth is dominated by regional growth centres and the final differentiation of tissues occurs [13].

Sutures of the skull contain mesenchymal stem cells (MSCs), which will mature to osteoblasts at the adjacent bone fronts [14]. In mice the MSCs arise in response to sonic hedgehog (SHH) signalling from the adjacent notochord [15]. Later, the MSCs move to a centre situated above the developing eye. Mesenchymal stem cells are very important for skull growth as their ablation provokes suture fusion [16].

Facial prominence is shaped by cranial neural crest cells (CNCCs) that migrated from: (i) Diencephalon and anterior mesencephalon to the frontonasal and periocular regions. (ii) From posterior mesencephalon and rhombomeres 1 and 2 to the first pharyngeal arch (PA1) (Figure 1.1) [17].

The face is shaped by the fusion of 5 facial primordia, frontonasal prominence, right and left maxillar and right and left mandibular process around the oral cavity. The gene *IRF6* plays a key role in the fusion of facial prominences and secondary palates [18]. Disruption of *Ptch1* enhanced SHH activity and this leads to decrease *Wnt-p53-IRF6* signalling where the maxillary, medial nasal and lateral processes fuse generating epithelial seam persistence and cleft lip [19]. Also, *SHH* signalling pathway plays a role on the expression of NCC marker *Tfap2a* in the frontonasal process, it has been associated with CL/P [20]. The *TGFβ – IRF6* pathway is involved in palatal development; in mice, *Smad4* and *Irf6* interact controlling the disappearance during fusion of medial edge epithelia [13,21].

The jaws are derived from the maxillary and mandibular processes of PA1 [22]. Several signalling pathways are involved in this process. Endothelian 1 is encoded by *Edn1* gene in PA1, and it is expressed in ectodermal epithelium of the mandibular process. They communicate via endothelin receptor A (encoded by *Edar* gene) [23, 13].

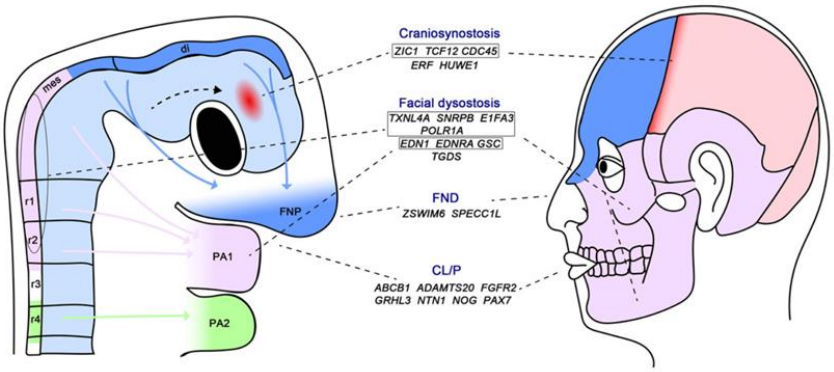


Figure 1.1. Development of the head and face and craniofacial genes involved in the process. An embryo is shown on the left, the patterns of migration are represented by the arrows. From diencephalic (di), anterior and posterior mesencephalic (mes), and rhombencephalic (r1–4) NCCs into the FNP and pharyngeal arches 1 and 2 (PA1, PA2). The dashed arrow indicates movement of mesoderm derived MSCs to the supra-orbital regulatory centre above the eye. The parietal and occipital bones from the head on the right are mesoderm-derived (pink). The coronal suture contains purely mesoderm derived cells (red). In the middle, the known genes implicated in the development of the head and face, with the embryological site underlying pathogenesis (if known—boxed genes), and the structures affected by mutation, indicated. This figure is from Twigg et al. (2015) [13].

1.2.1.2 Anatomy of the face

The skull has 22 bones, 8 cranial bones and 14 facial bones [24]. The bones and regions are explained below (Figure 1.2).

Cranium. The skull consists of paired parietal and temporal bones and a single frontal and posterior occipital bone (Figure 1.2). Within these bones there are five major sutures. Two coronals, one on each side, two lamboidals and two squamosals, and one sagittal and one metopic. In early life these bones are separated by five major sutures (Figure 1.2).

Forehead. Corresponds to the area above the eyebrows and below the hair line.

Glabella is the most prominent point on the frontal bone above the root of the nose.

Supra-orbital Ridge. The supraorbital portion of the frontal bones.

Midface, it is not a bone, it is a region of the face that goes from the lower part from the orbital margin to the nasal base. Including the upper jaw and zygoma.

Maxilla is shaped by the union of two bones and contains the upper teeth, it also form the boundaries of the palate, floor and lateral wall of the nose (malar process) and the

floor of the orbit. Towards the upper part each bone the zygomatic arch, malar, alveolar and palatine.

Zygomatic arch is formed at the junction of the zygomatic process of the temporal bone of the bones of the skull and the articulation of the malar process, typical of the bones of the face, located on one side of the orbital pits. It forms the prominence of the cheek and it is also part of the floor of the orbit.

Anterior nasal spine or maxilla, it is a pair of bones situated at the very tip of the upper jaw in which the four upper incisors develop and underlies the philtrum and upper lip.

Lower Face is located between the mouth and the lowest point of the chin.

Cheek are the soft tissues between the zygomatic arch and the mandible.

Mandible or lower jaw. It consists of a curved, horizontal portion, the body, and two perpendicular portions, the rami, which unite with the ends of the body nearly at right angles. The inferior teeth are situated in the lower jaw.

Chin or mental protuberance of the mandible it corresponds to the lowest central prominence of the lower jaw.

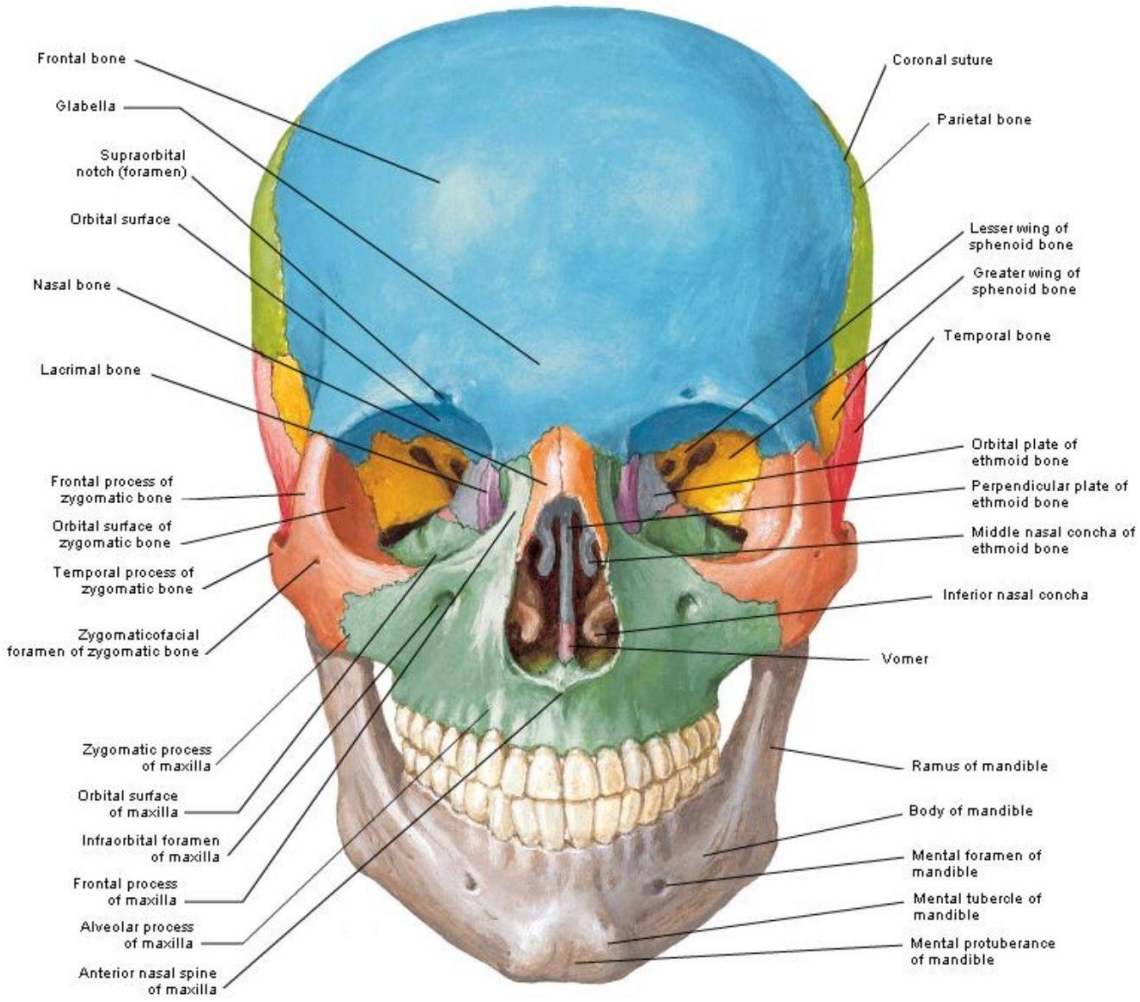


Figure 1.2. Anterior view of the skull. Figure from Atlas of Human Anatomy from Netter [24].

There are two groups of muscles on the face. The mimetic muscles and muscles of mastication [25].

Mimetic muscles are involved in the expression of feelings and thoughts by elevating or depressing the eyebrows and lips. Situated around the eyes and mouth [26]. Muscles of mastication are responsible for movement of the jaws to chew the food [25] (Figure 1.3).

Muscles upper face

Frontalis, orbicularis oculi, corrugator supercilia, procerus and depresor supercilii [25] (Figure 1.3).

Muscles of the middle face (most of them are mimetic muscles)

Orbicularis oris, buccinator, zygomaticus major, zygomaticus minor, levator labii superioris aleque nasi and levator labii superioris, levator anguli oris, risorius, nasalis and depressor septi [25] (Figure 1.3).

Muscles of the lower face

Depressor anguli oris, depressor labii inferioris, mentalis and platysma [25] (Figure 1.3).

Masticatory muscles (lateral part of the face)

Temporalis, masseter, lateral pterygoid and medial pterygoid [25] (Figure 1.3).

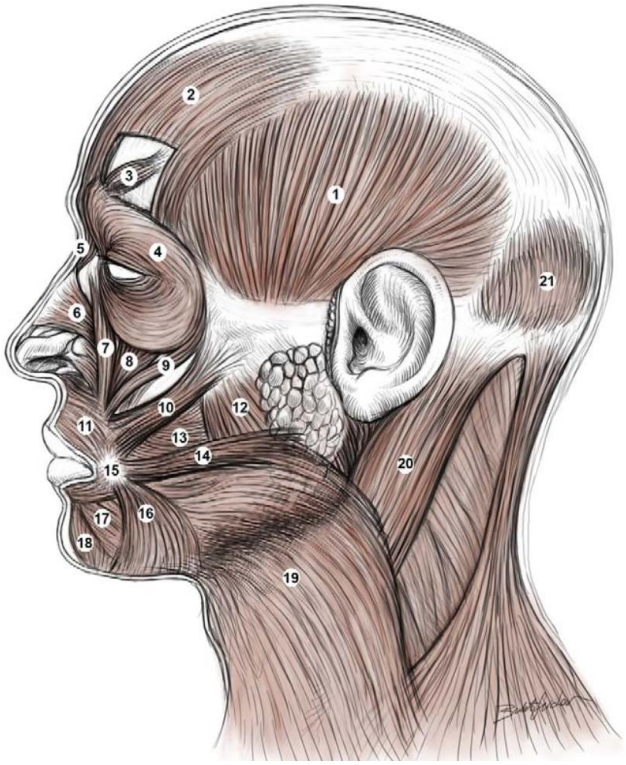


Figure 1.3. Mimetic muscles and masticatory muscles. (1) temporalis,(2) frontalis, (3) corrugator supercilii, (4) orbicularis oculi, (5) procerus, (6) nasalis, (7) levator labii superioris aleque nasi, (8) levator labii superioris, (9) zygomaticus minor, (10) zygomaticus major, (11) orbicularis oris, (12) masseter, (13) buccinator, (14) risorius, (15) modiolus, (16) depressor anguli oris, (17) depressor labii inferioris, (18) mentalis, (19) platysma, (20) sternocleidomastoid, (21) occipitalis. Figure from Marur et al. (2014) [25].

1.2.2 Evolution of the human face

Facial features show great variation amongst individuals and have been used by Physical Anthropologists to examine human population diversification. There is a possibility that this diversity in facial appearance could have evolved partially to simplify human recognition [27, 28]. The craniofacial region houses the brain, the sensory

organs and structures from the digestive system. Therefore, the variations in anatomy and function of the craniofacial complex are one of the best evidences of evolutionary adaptation [1].

Individual recognition (IR) is described as a complex form of communication, there are two parts, receivers and signalers [29]. Both perspectives are crucial, because selection can act on receivers and signalers, independently. The recipient perspective has been widely studied. Interestingly, the signalers perspective has not received enough attention and is of great importance due to selection of signalers to be memorably different, possibly providing an under-appreciated selective mechanism that increments on phenotypic diversity. Evolutionary biologists have shown huge interest in high phenotypic variability, because this is advantageous for IR. In fact, to disseminate individual identity, positive selection represents a possible explanation for high levels of phenotypic diversity in social individuals [30]. The face is shaped by 14 bones and there are around 43 muscles which are involved in the expression of more than 20 different emotions [31].

Certainly, the evolution of the face has involved several stages of morphological transformations of the face. From protruded and bulging forehead in a smaller brain case toward a shorter face located in a rounded and large brain case (Figure 1.4) [32, 33,34, 1]. These modifications implied an increase on the energy cost to maintain the bigger brain. The "expensive tissue" hypothesis suggests that the increase of the brain size happened simultaneously with the decrease of the gut size, to compensate for the increase in energy demands of the brain [35].

Figure 1.4 Crania of Homo sapiens and Middle-Late pleistocene hominins. a, La Ferrassie 1 Neanderthal (~60–40 ka.) **b,** Bodo (Ethiopia) (~600 ka.) **c,** Broken Hill 1 (Zambia) (~250–300 ka.) **d,** Nanjing, China, (~400 ka.) **e,** ATD6-69 maxilla, the holotype of H. antecessor (~850 ka.) **f,** H. sapiens from Jebel Irhoud 1 (Morocco) (~300 ka.) **g,** H. sapiens idaltu from Herto (Ethiopia), (~160 ka.) From ref. 95, SNL. **h,** H. sapiens from Abri Pataud, France (~20 ka.) Skulls are not to scale. Figure from Lacruz et al. 2019 [36, 37, 38, 39].

Lacruz et al. (2019) in their recent paper show an interesting approach in analyzing the evolution of the face, where they consider population history, palaeogenomics, environmental and social factors, in addition to the adaptative explanations, regarding the facial morphological changes over time. They assess the evolution of the face by following the hominin taxonomy (Figure 1.4). The first hominin found outside Africa was *H. erectus* (early Pleistocene) [40]. During the first dispersal out of Africa, in the Middle Pleistocene, *Homo* evolved and spread through different parts of Eurasia. Thus, the evolution of the face of *H. sapiens* is integrated by a group of possible ancestors. The authors focus the answer to which of these possible ancestors is the last common ancestor (LCA) on key morphological features, especially on the morphology of the modern human zygomaxillary, which is different to the zygomaxillary from the Middle Pleistocene group. They propose *H. heidelbergensis* and *H. antecessor* as the possible LCA, both share similar characteristics with *H. sapiens* but also differences. In the case of *H. heidelbergensis* they would have gone through many morphological changes in two different directions, such as a gracilization of the zygomaxillary morphology, as in H. sapiens and an increase of the midfacial projection and maxillary protrusion that lead to the Neanderthals facial shape. Another issue is the allometric

factors and the sexual dimorphism affecting the morphology of the face in the different individuals found, which is a-confounding factor. Under this scenario, *H. antecessor* seems a better candidate, because *H. sapiens* retained *H. antecessor*-like facial morphology, but the morphology was modified in Neanderthals.

Among the aspects considered in this review, an important one is how facial morphology is affected by population history and climatic changes. For instance, Neanderthals faces show a large nasal opening, a cushioned maxilla and a protuberant midface [41], this leads to the question if these changes are due to adaptation to a cold environment? [42] Or, are they because of genetic drift in a small population? [43]. They state that the answer is among the variation of modern human populations. Several studies have analyzed the effect of different factors using quantitative genetic approaches on facial features. The modern human cranium seems to be mainly affected by neutral evolutionary forces [44, 45] and the face has been influenced by both environmental and genetic factors [46, 47]. The nose is a good example because theit external morphology, as well as the nasal cavity shape are affected by temperature and humidity, especially in populations living in climates with extremely low temperatures [48, 45]. The air that passes through the air passages must be warm and humid and the nose is responsible for this, as the bigger the nose cavity the bigger the area in touch with the air entering the nose, and the air is warmed and moisturized [46].

The last analysed factor was how facial features have been affected by cultural and social components. The main change is the reduction of the face [34], a more vertical profile, a smaller midface and less protruded jaw. This could be associated with an increase of social tolerance as well as a decrease of androgen activity [49]. The reduction of eye bridge and midface was likely to increase the mobility of this area to show more variety and soft movements, contrary to the stiff features, commonly associated to an aggressive behaviour, found in Neanderthals [50]. All of this could have improved communication and interaction among human beings.

1.2.3 The genetics of facial morphology

It is evident, even for people without an understanding of Genetics, that facial morphology has a strong genetic component, due to a higher similarity between related individuals than unrelated individuals. Heritability of facial traits is in the range from moderate (0.4) to high (0.8) [27, 28, 51, 7, 13]. The first heritability studies were based on

qualitative facial data [52, 53]. Nevertheless, these studies were questioned [54] and quantitative methods were the option [55]. Heritability results from different quantitative studies are diverse, with regards estimates and conclusions, mainly due to different study designs, such as different methods to obtain the images (i.e. cephalograms, anthropometric measurements, 2D photographs, 3D images, etc.), differences in the phenotypes assessed (i.e. distances, angles, proportions, etc.), the sample (i.e. twins, families, unrelated individuals), a variety of statistical methods (i.e. parent-offspring regression, linear mixed model either pedigree-based or kinship matrix), among others [56]. The same trait shows a high correlation in heritability estimates among several studies [57].

The most emblematic example used to calculate heritability is monozygotic (MZ) twins. There are studies in which the contribution of heredity and environment were assessed in monozygotic, dizygotic twins and parent-offspring resemblance [8, 58, 59]. Different studies have found evidence of heritable shape variation in central and medial face structures in MZ and dizygotic (DZ) twins [13, 60, 61, 62] (Table 1.1). The concordance for vertical facial measurements of the middle and lower anterior segments of the face is greater in MZ twins [63]. Recent studies have assessed the association between the regional genetic variation across continents and physical appearance features, interestingly this contrast is significantly associated with variation in facial features, especially middle facial traits, such as, nose shape, nose protrusion, etc.[64].

Heritability	Liu 2014 Twin correlations		Adhikari 2016	Cole 2017	Cha 2018
Trait	h2 MZ	h2 DZ	h2	h2	h2
facial width	0.83	0.51	–	0.66	0.23
nasal width	0.77	0.26	0.9	0.62	0.29
nasal protrusion	0.66	0.1	0.84	0.34	0.42
upper lip height	–	–	0.63	0.31	0.21
lower lip height	–	–	0.46	0.28	–
Intercanthal width	–	–	–	0.57	0.26
Outercanthal width	–	–	–	0.41	0.24

Table 1.1. Heritability estimates (h2) of facial traits from different studies. The standard deviation for all the traits was between 0.05 and 0.07. More details about the studies presented here in section 3.1.1.

The environmental influences are key to understanding the genetic effects in facial traits. Interestingly, both factors seem to behave similarly, with respect to biological pathways and mechanisms [55]. Some traits are more affected by genetic influences, such as anterior face height [65], nasal height [62], allometry [66], among others. Environmental

influences estimates are higher in antero-posterior facial height, length of the mandibular body, ramus height, nasal width, etc. [66-68].The fetus is susceptible from a very early stage during pregnancy, many substances can cross through the placenta and affect the development of the fetus [68]. There are several substances which affect the development of the face *in utero*, such as certain drugs, alcohol [69], vitamin deficiencies [70], nicotine [71], solvents and pesticides [72], etc. There is evidence that high levels of alcohol consumption by the mother during pregnancy can affect facial morphology development [73]. Beaty et al. (2013), showed that maternal smoking may affect *GRID2* and *ELAVL2*, causing the fetus to have facial abnormalities, such as cleft lip and palate [74]. Climate also affects facial morphology, for instance individuals living in more sun-exposed latitudes on the planet will have narrower faces and people living in less sun-exposed areas will have broader faces [44,75]. Furthermore, as previously mentioned, low temperatures could have affected nose shape to adapt to extremely cold environments [46,47].

1.2.4 Genetics of craniofacial abnormalities

There are many genetic disorders in which physical abnormalities are present. Around two-thirds of the genetic conditions display some oral, dental or craniofacial malformations. These features have been frequently used by medical practitioners as a diagnostic tool. [76].

The head is the most complex structure of the body. The development of the human skull and face is a complicated process that involves many genes controlling each stage of the craniofacial morphology development [13]. Abnormalities in the head and face are one the most common anomalies present in children [77].

1.2.4.1 Craniofacial Abnormalities

The brain and sensory organs are hosted and protected by the skull and the bones of the face. Furthermore, the bones, muscles, connective tissue, blood vessels, associated innervation and teeth are responsible for breathing and feeding functions. These tissues are derived from ectoderm, mesoderm, endoderm and cranial neural crest cells (CNCCs) and their derivatives [78]. Primarily craniofacial tissue is formed by CNCCs and mesoderm and this mesenchyme is in constant communication with the rest of the cellular components providing positional signals and regulating the growth and differentiation of the different parts of the face (Figure 1.1). It is evident that the

craniofacial development is very complicated, and it depends on many factors acting in perfect coordination, therefore, it is susceptible to be easily damaged and generate malformations [79]. In this section I will present a summary of the most common craniofacial abnormalities, based on the area of disruption (craniosynostosis and cleft palate) and the possible genes responsible for it.

Craniosynostosis

Craniosynostosis is when one or more sutures are closed prematurely, and this leads to limit the growth of the brain and the skull is shapeless. Mutations in *ZIC1* were observed in patients with craniosynostosis, it is known that there is an epistatic relationship between *ZIC1* and Engrailed 1 (*EN1*) gene [13]. Experimental work in Drosophila and Xenopus have shown that disturbing the expression of encoding engrailed 1 (*En1*), which is part of an osteogenic path together with *Ms2* and *Twist1*, disturbs the early coronal suture patterning [13,15].

The *TWIST1* gene encodes a transcription factor, which is expressed in the mid-sutural mesenchyme [80] controlling *RUNX2*-mediated osteogenesis [81]. If *TWIST1* decreases its activity this causes Saethre-Chotzen Syndrome (OMIM#101400) [82]. *TCF12* is a gene encoding protein, which works together with *TWIST1*. Mutations in the *Tcf12* gene in mice alters the combine dosage of *Twist1* and *Tcf12* generating severe coronal synostosis [83].

Fibroblast growth factor receptor (*FGFR*) genes are involved in the development of bone and cartilage [84]. *FGFR2* is mutated in several craniosynostotic syndromes, such as Crouzon, Apert, Pfeiffer, Antley-Bixler, Beare-Stevenson cutis gyrata, Jackson-Weiss, Bent Bone Dysplasia, and Seathre-Chotzen. Most of the mutations are missense mutations that activate the receptor and subsequently the downstream molecular pathways [85].

Cleft Lip and/or Palate

Cleft lip and/or palate (CLP) are birth defects resulting from failure of facial development to grow or fuse appropriately throughout the early embryological development. [86, 87]. The aetiology of this condition is complex and multifactorial, it includes both genetic and environmental factors [87]. Orofacial clefting is commonly classified as non-syndromic (isolated defect) or syndromic CLP, and it is present in approximately 150 chromosomal syndromes [88, 89]. There are several types of cleft lip

and/or palate, it is classified by laterality and completeness of the facial process during the embryologic development [88]. It varies from cleft lip to complete bilateral cleft lip and palate [90]. Cleft lip and/or palate is a relatively common anomaly and the rates vary depending on the population. Asian and native Americans have the highest incidence (0.79 to 4.04 per 1000) [91]. Even though, orofacial clefting has been widely studied, the aetiology is still unclear, and it is a problem for families having children with this condition, because they have feeding, hearing, speaking and social interaction problems, besides the economic problems that may result from expensive surgeries and dental treatments.

Through molecular methods and GWAS around 300 genes have been associated with cleft lip and/or palate, *MAFB*, *PAX7, ABCA4*, *THADA*, *VAX1*, *BMPR1B*, *FGFR2*, *IRF6, TBX1, COL2A1, COL11A1, POLR1C, TGFβ1, FOX31, GREM1*, etc. [92, 93, 94]. Nevertheless, the mechanisms and genetic variants behind these associations are still unknown. Some regions were selected and re-sequenced. A missense substitution within *PAX7* was detected in a non-syndromic CLP case. *PAX* gene family have been previously associated with facial traits, *PAX3* to nasion position [7, 8], *PAX1* to nose wing breadth [95] and *PAX7* has been considered as a risk factor by GWAS and linkage analyses [74]. Interestingly, mice with homozygous mutations in *Pax3* and *Pax7* have a reduction in the facial prominence and clefting [96]. *FOXE1* [97], *GREM1*[98], *CDH1* [99] and other associations have been validated [74, 100, 101, 102, 103]. An enhancer in a key non-syndromic CLP susceptibility locus (8q24) regulates *Myc* acting in cis over 1Mb. Mutations in this region downregulate *Myc* affecting the growth of the medionasal process [104].

A table summarizing the information above and more candidate genes is presented below (Table 1.2).

Gene	Locus	Clinical disorder	OMIM # disorder	Inheritance pattern	Mechanism/pathway/comments	References
Craniosynostosis						
CDC45	22q11.21	–	–	AR	Peturbation of DNA replication, probable that some function remains	Taylor et al. (2015)
ERF	19q13.2	ERF-related craniosynostosis	600 775	AD	Haploinsufficiency; FGFR/ERK signalling	Twigg et al.(2013)
HUWE1	Xp11.22	–	–	XLD (n)	Unknown	Taylor et al.(2015)
TCF12	15q21.3	TCF12-related craniosynsotosis	615 314	AD	Haploinsufficiency; heterodimerization with TWIST1; RUNX2, BMP and FGFR signalling	Sharma et al.(2013)
ZIC1	3q24	ZIC1-related craniosynostosis	–	AD (n)	WNT signalling	Taylor et al.(2015)
Selected CL/P candidates						
ABCB1	7q21.12	–	–	Complex	Control of foetal exposure to foreign chemical substances	Omoumi et al.(2013)
ADAMTS20	12q12	–	–	AR	Uknown—Extracellular matrix processing?	Wolf et al.(2015)
FGFR2	10q26.13	Crouzon Syndrome Pfeiffer Syndrome Apert Syndrome Antley-Bixler Syndrome Beare-Stevenson cutis gyrata syndrome Jackson-Weiss Syndrome Bent Bone Dysplasia syndrome Seathre-Chotzen-like syndrome	123500 101 600 101 200 207 410 123 790 123 150 614 592 101 400	Complex	254 kb downstream of FGFR2; risk allele disrupts NC enhancer activity. Several missense mutations.	Leslie et al.(2015)
GRHL3	1p36.11	Van der Woude syndrome 2	606 713	AD	Periderm development	Peyrard-Janvid et al.(2014)
NTN1	17p13.1	–	–	Complex	Expressed in palatal shelves	Leslie et al.(2015), Zalc et al.(2015)
NOG	17q22	–	–	Complex	Risk allele shows significantly	Leslie et al.(2015)
PAX7	1p36	–	–	Complex	Involved in NC induction and specification of NC derivatives	Leslie et al.(2015)
AFND			603 671	AD (n)	Gain-of-function; pathogenetic mechanism is unknown	Smith et al.(2014)
SPECC1L	22q11.2	Opitz G/BBB	145 410	AD	Haploinsufficiency; SPECC1L has role in cell adhesion and migration	Kuska et al.(2015)

Continue…

Gene	Locus	Clinical disorder	OMIM # disorder	Inheritance pattern	Mechanism/pathway/comments	References
Facial dysostoses						
EIF4A3	17q25.3	RCPS	268 305	AR	Reduced expression; component of exon junction complex; directly interacts with spliceosome components	Favaro et al.(2014)
EDN1	6p24.1	ACS; IQME	615 706 (ACS) & 612 798 (IQME)	AR (ACS)/AD (IQME)	Loss-of-function; likely that residual function remains with AR mutations	Gordon et al.(2013)
EDNRA	4q31.22	Mandibulofacial dysostosis with alopecia	616 367	AD (n)	Possible gain-of-function	Gordon et al.(2015)
GSC	14q31.23	Short stature, auditory canal atresia, mandibular hypoplasia, skeletal abnormalities (SAMS)	602 471	AR	Loss-of-function	Pany et al.(2013)
POLR1A	2p11.2	Acrofacial dysostosis, Cincinnati type	–	AD	Haploinsufficiency; ribosome biogenesis defect	Weaver et al.(2015)
SNRPB	20p13	Cerebrocostomandibular syndrome	117 650	AD (n)	Mutations affect the recognition and inclusion of the premature termination codon-containing alternative exon; spliceosomal component	Lynch et al.(2014), Bacrot et al.(2015)
TGDS	13q32.1	Catel-Manzke syndrome	616 145	AR	Loss-of-function; likely that residual function remains; pathogenetic mechanism unknown	Ehmke et al.(2014)
TXNL4A	18q23	Burn-McKeown syndrome	608 572	AR	Loss-of-function; likely that residual function remains—combination of null and hypomorphic allele	Wieczorek et al.(2014)

Table 1.2. Genes region known to be associated with craniofacial abnormalities. Most common craniofacial abnormalities organized by anatomic area affected. Table modified from Twigg et al. (2015) [13].

1.3 The human teeth

Teeth have been studied for a long time, because they are the best preserved part of a living creature, and this represents an important source of information in different areas. Dentition is the first part of the gastrointestinal tract and it also takes part in the shape of the face. It works as a perfect place to store DNA because it resists high temperatures and injuries. Furthermore, different ethnic groups have different features observed in teeth, which is useful to recognize different populations [105]. I will describe some of these aspects in this section to better understand how teeth have evolved and developed through time.

1.3.1 Development and anatomy of human teeth

1.3.1.1 The genetics of tooth development

The normal development of person's teeth depends on many factors, such as genetics, environmental and cultural factors. Teeth are shaped by a sequence of inductive signals communicating between the ephitelium and neural crest derived from mesenchyme. Each layer has a function and they influence each other to differentiate in a specific manner and this will culminate in the formation of well defined structures, such as incisors, canines, premolars and molars [106].

Tooth development begins with the formation of the dental lamina (Figure 1.5). This is achieved through the thickening of the dental epithelium [106]. Quickly, cells begin to multiply and to invaginate giving rise to different structures. First, the bud, which is a group of cells with no clear organization (Figure 1.5). Subsequently, the cap stage, invagination stops and the cells begin to compress each other (Figure 1.5)[107,108]. This leads to a new structure, which is found in several mammalian species [108], called the enamel knot (EK) (Figure 1.5). The EK, is a key structure because the epithelium develops surrounding it, but not in the knot itself. This results in the formation of the cervical loops, which are stem cells niches[109]. Enamel knots lay the foundation for the final size and shape of the tooth, these structures are the pattern for growing and folding of the inner enamel epithelium [110]. During the bell stage, there is a histodifferentiation of the epithelium into enamel-secreting ameloblasts and the mesenchyme into dentin-secreting odontoblasts[106]. In this stage, in multicusped teeth,

other knots appear at a certain distance from the first knot and, ultimately each knot will shape the tip of a cusp (Figure 1.5)[108].

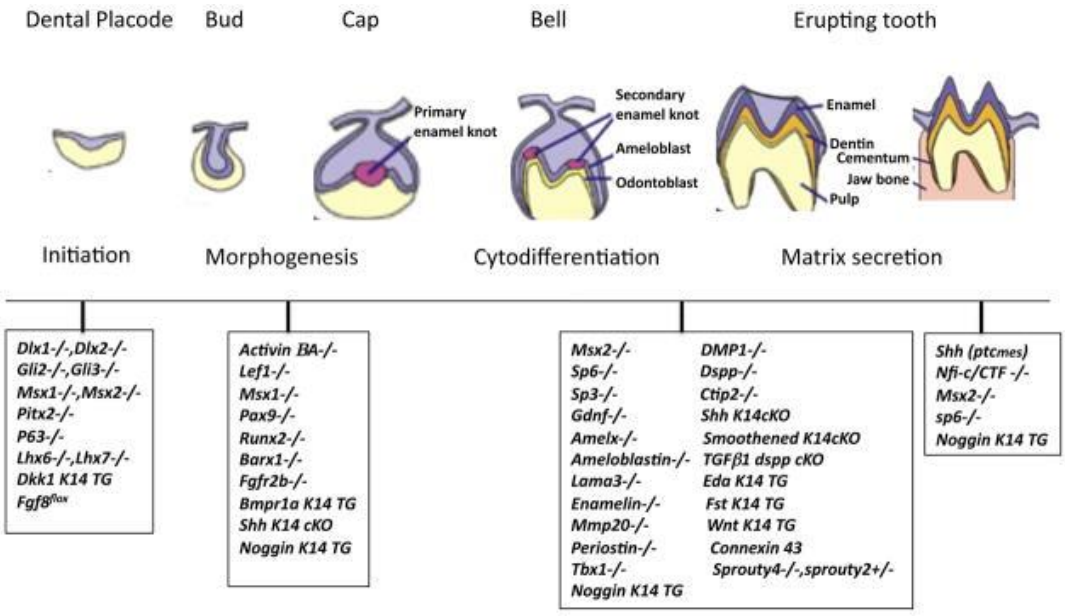


Figure 1.5. Development of teeth and genes known to be involved in the process. Most of the genes are expressed in one or several teeth, or specifically around the knots. Figure from Bei et al., 2009 [106].

1.3.1.2 Anatomy of teeth

Dental Anatomy is common to many fields, odontology, paleontology, dental anthropology, genetics, among others. In this thesis I will explain the Paleontologists terms based on the book from Scott and Turner (1997) [111].

General description the human teeth

Incisors, a narrow-edged tooth at the front of the mouth, used for cutting. In the human dentiton there are four incisors in the upper jaw and four incisors in the lower jaw.

Canines, a conical tooth between the incisors and premolars, there are two canines in the upper jaw and two canines in the lower jaw.

Premolars, bigger tooth than canine situated between canines and molars. There are four premolars in the upper jaw and four premolars in the lower jaw.

Molars, a grinding tooth at the back of the mouth. There are six molars in the upper jaw and 6 molars in the lower jaw (Figure 1.6).

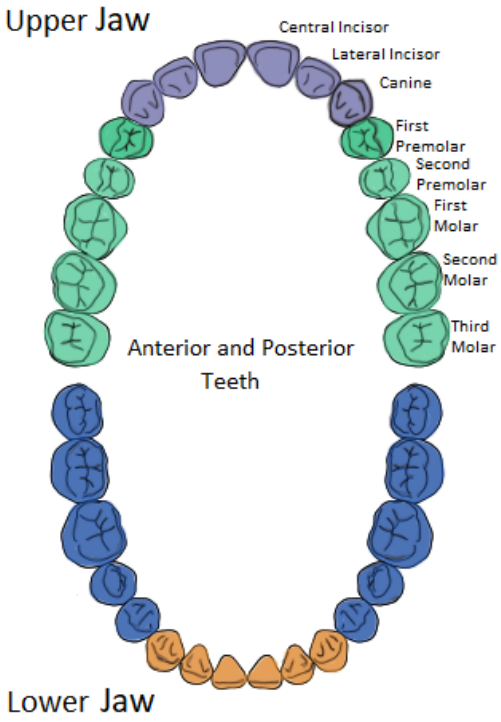


Figure 1.6. Anterior and Posterior teeth in upper and lower jaw.

Positional terms

Regarding the mandible and maxilla, there is a midline between the central incisors. The human jaw is parabolic with theeth shaping an arch in the canine region towards the midline, thus the use of the terms anterior and posterior are not the exact words to describe the front and back of a tooth. Therefore, the preferred words are *mesial* and *distal*, to denote the part of a tooth closest to or farthest to the midline, respectively. Anterior is used to refer to anterior teeth (incisors and canines), and posterior teeth (premolars and molars). The side of a tooth facing the tongue is called *lingual or medial* and the outer side of the tooth, depends on the type of the tooth, the surface in contact with the lips of incisors and canines is called *labial* side and the surface adjacent to the cheeks for premolars and molars is called *buccal*. In this thesis the term buccal will be used for both anterior and posterior teeth.

The vertical components of a tooth is divided in three sections, upper third (*incisal/occlusal*); middle third (*middle*); and lower third (*cervical*). Roots, the tip or apex of the root is called *apical* part. Toward the base (*Basal*) is applied for both either crowns (cervical third) or roots (apical thirds). Roots are divided equally to crowns, lower third (apical); middle third (middle); upper third (cervical) (Figure 1.7).

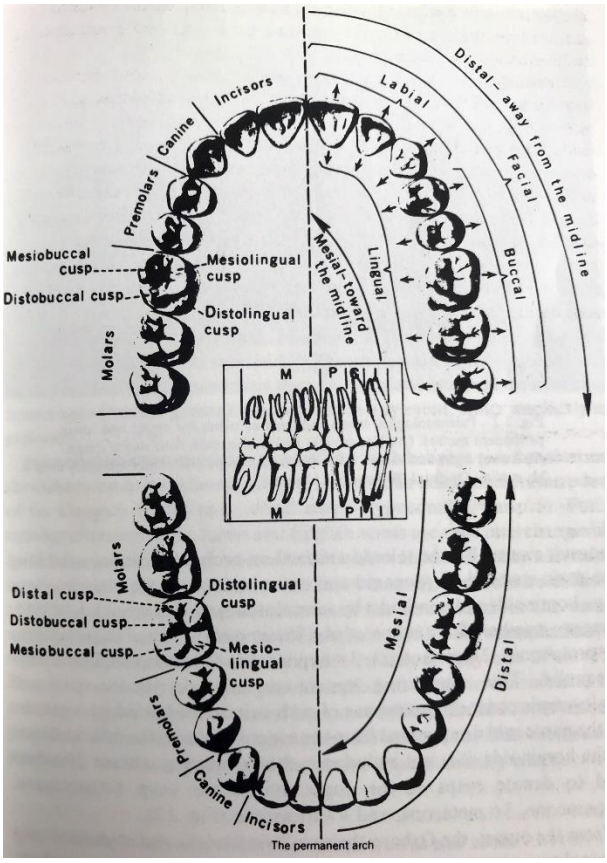


Figure 1.7. Positional terms of human permanent dentition. Figure from Scott and Turner (1997)[111].

Paleontological terms and cusp numbers

There have been several naming system for the cusp features and numbers based on different theories of dental evolution. Cope (1874, 1888) and Osborn (1888a, 1888b, 1897, 1907) developed the tritubercular theory. Osborne (1888b) proposed names for the three main cusps of upper molars (trigon) and lower molars (trigonid). This theory resulted in a naming system, *protocone, paracone*, and *metacone* forming the trigon with *protoconid, paraconid* and *metaconid* forming the trigonid. In primates there was an extra distolingual cusp called hypocone. In lower molars the paraconid was lost leaving the trigonid with two cusps, the protoconid and metaconid. A talonid in subocclusal position was additioned to the trigonid, and later the talonid attained an oclussal position, exhibiting two main cusps, the *hypoconid* and *entoconid*. Amplified in some taxa (honinoids and hominids), is the hypoconulid. In 1916, Gregory added numbers to the cusps of lower molar crown: cusp *1=protoconid, 2=metaconid,*

3=hypoconid, 4=entoconid, and *5=hypoconulid*. Numbers in the upper molars: 1=*protocone, 2= paracone, 3= metacone,* and *4=hypocone* (Figure 1.8) [111].

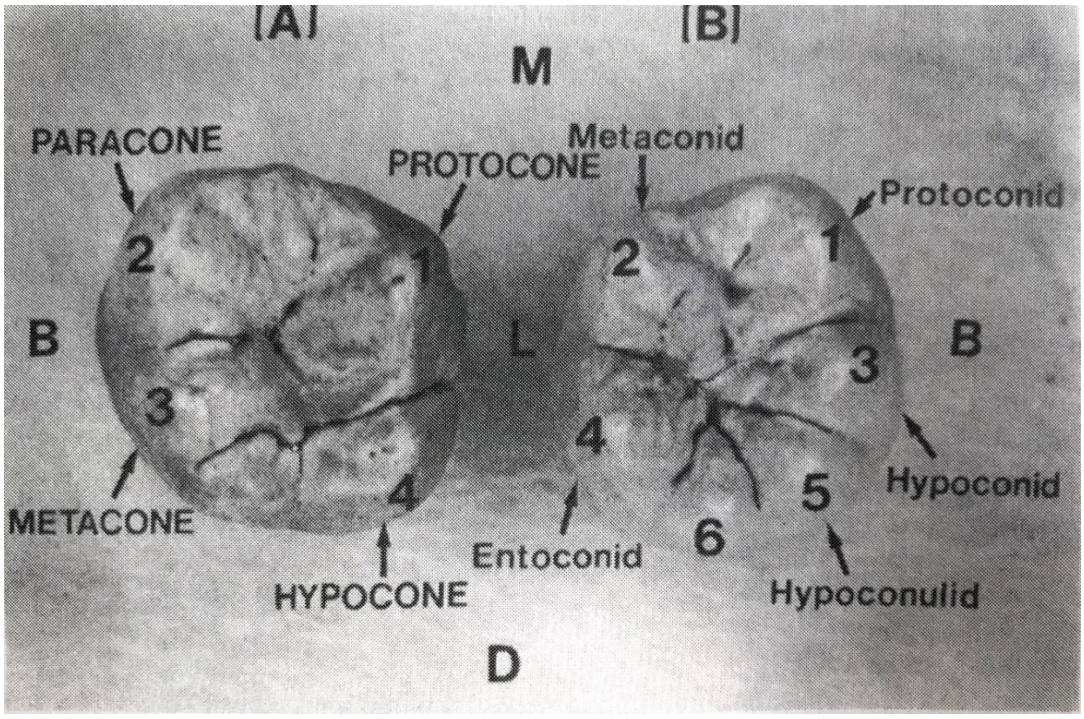


Figure 1.8. Paleontogical names and cusp numbers for the upper and lower molars. A) Upper right first molar, with four major cusps. B) Lower right first molar, with five major cusps and supernumerary cusp 6. M: mesial, D: distal, B: buccal, L: lingual. Figure from Scott and Turner (1997)[111].

1.3.2 Origin and evolution of the teeth

Understanding the structure of vertebrates is important to construct vertebrate taxonomy, thus to be able to know about the structure of human beings, since we are vertebrates [112].

Teething was for a long time the most analyzed feature of vertebrate skeletons. It focused on the development of the teeth, its number, and morphology. Dentition captured the attention of European biologists, paleontologists, and anatomists, such as George Cuvier and Richard Owen. Teeth reflect function and adaptation to the environment of the species, and they are the major structures found in fossils [113]. A very important aspect was the skeleton because it depicts serial homology among its parts, Owen defined serial homology as "representative or repetitive relation in the segments of the same organism," also called "homotypy" by Ernst Haeckel. This concept means the resemblance between different members of a single series of structures in an organism [113]. When Darwin proposed that similarities among species

could be explained in parallel with history, in other words, that homology is equated to common ancestry; teeth were the most important structures used in the building and understanding of vertebrates, particularly for mammalian phylogeny. Darwin hypothesized that the repetition of a trait was before regionalization. The existing diversity is the result of a common ancestoral variation [114].

The diversity of mammals was kept to depict the variation of common components, ancestrally shared, through different species and genetics accepted similar assumptions regarding the genetic mechanisms. This is still important for general biology and human oral biology, to justify the use of the mice as a model to explain the development and evolution of human dentition[113].

Darwin thought of the gradual evolution of the shape, meaning the shared traits by two species were also shared with their ancestors, having a common ancestor. But William Bateson, an important geneticist from the 19th century, believed that teeth had evolved qualitatively [115]. Many aspects of teeth are qualitative, i.e., the number of teeth, separated spaces, also discrete features of each tooth, such as cusps and roots. Bateson thought there was no need to share a common ancestor or stern homology within specific teeth or cusps in such circumstances [115]. Discrete variation cannot step back continuously. Nevertheless, Darwin's vision remains in modern biology[113].

In the past, the problem of serial homology evolution was considered by Darwin as "almost beyond investigation" [114]. But some years ago, new developmental and molecular methods have revealed interaction mechanisms among the tissues during odontogenesis (the generation of basic tissues and structures inside the teeth), these mechanisms are similar to other tissues in the body [113].

Several theories have attempted to explain the origin of dental morphology. One of the first theories was that dental crowns were multituberculates and there was a simplification throughout evolution. Nowadays, the most accepted theory is the opposite, the "Tritubercular Theory," that was described by Cope in 1871, the main idea of this theory was that the most complex tooth of mammals has derived from a single tubercle [113,116]. Subsequently, Cope, in 1883,announced that within fossils of mammals from the Eocene period, he had discovered upper molars with tritubercular

shape. Later, in 1907 Henry Fairfield Osborn published his theory about the evolution of the cusps dental system or "Tritubercular Theory of Cope.". In the beginning teeth were simple, and the evolution of the form was not because of the fusion of several simple teeth, but the addition of new parts.

Osborn considers as the first stage of this process, the existence of a primitive cone modified by lateral expansions. In the second stage, molars show three cusps in a row (Triconodon). When this stage ends, Osborn assumed that teeth would become triangular (trigonid), moving the protocone toward a vestibular direction. The third stage corresponds to the interdigitation between the upper and lower molars. During the fourth stage, the development begins from the distal side from the lower molar, from the hidden talonide under the upper molar protocone, when the jaws were closed. Finally, in the last stage additional cusps were emerging from the talonide; hypoconid or vestibular cusp, entoconid or lingual cusp and hypoconulid or distal cusp [117].
Currently, the name of the main cusps from molars are named based on Osborn nomenclature, facilitating the communication among anthropologists, paleontologists and odontologists that work in this field [118,119].

Outside-in or Inside-out?

The important role that teeth and jaws have in the evolutionary success of living jawed vertebrates is well understood. Nevertheless, the origin of them is still not clear. As mentioned previously, several theories are trying to explain this matter, but it still an object of controversy [120].

Around twenty years ago, the outside-in theory hypothesized that teeth were the result of the movement of tooth-like dermal denticles or "scales" from the external dermis to the oral cavity. But this hypothesis has been defied many times by fossil records and evidence of developmental differences of external dermal denticles and internal denticles and teeth. These elements led to a new hypothesis called "inside-out", which postulates that teeth and tooth-like structures inside and outside of the mouth evolved independently of one another [8,120]. Finally, based on different analysis and observations in different vertebrates, the "inside-out hypothesis" has been refuted, the "outside-in" hypothesis becomes stronger because it seems that internal odontodes

appeared several times in oral, pharyngeal and nasal areas of extint gnathostomes, as a contimuous part of the dermal scale cover. Therefore, this strengthens the argument that the best explanation for the evolution of the teeth is the “outside-in” hypothesis [120].

1.3.3 Dental phenotypes and the history of human populations

Currently, humans have 32 teeth. Two pairs of incisors, one pair of canines, two pairs of premolars, and three pairs of molars on each jaw. Incisors have the shape of a spatula and canines are single cusped, both have one root. Premolars have two cusps and molars have up to 5 cusps. The upper molars have three roots, while lower molars have two roots [111].

In the Hominidae family there are 8 living species in four genera, including humans, orangutans, gorillas, chimpanzees and bonobos [121]. The characterization and description of dental features in these different species has an important role in the study of the evolution of the dental archade.

The size of canines in Earliest Hominids were from large to mild, although they were still smaller than chimpanzees. They also have thicker enamel and a prognathic jaw [122, 123]. *Ardipithecus,* one of the youngest groups within the earliest hominids (5.6 to 4.4 million years ago), presented upper canines less sharp than chimpanzee’s and less sexual dimorphism in contrast to chimpanzees [121].

The archaic hominids *Australopithecus afarensis* have much smaller molars and canines, but larger than modern humans. A prognathic jaw, they were considered frugivorous, but there is evidence that they also ate meat, giving the advantage to survive in different environments [124].

The archaic megadont hominids presented the biggest reduction in the size of canines, but premolars were still larger than other species [125]. Their jaw was still bigger than the *Homo* species [126]. They also had thicker enamel than any other hominid specie. They showed signs of chewing hard, this may be due to nutritional stress. Therefore, they also presented bigger molars than modern humans [126].

Pre-modern humans presented smaller teeth and a reduction of facial prognathism [126]. In *Homo heidelbergensis* new traits appeared in the form of taurodont molars, reduced third molars, large buccal cusps in premolar three, larger jaw and smaller teeth [127].

Homo neanderthalensis had a much more prognathic face. And they also exhibited some peculiarities, a pronounced basal eminence and spines and ridges on the upper anterior teeth (i.e. tuberculum dentale) [111]. They have larger molars and well-developed shoveled incisors and canine teeth with no grooves [128].

The modern humans. *Homo sapiens* present lack of prognathism, parabola shaped mandible and maxilla, molars are the same size as front teeth, the crowns are small and with a small number of root and cusps [125].

Several non-metric traits were present in polymorphic frequencies in the teeth of fossil hominid. For example, some crown and root traits seen in modern humans have an evolutionary history in not only the hominid but also the primate lineage (e.g. hypocone, hypoconulid, cusp 6, cusp 7) [111].

Several archaeological and paleontological studies have employed dental phenotypic data to estimate biological relatedness among human populations from the past, in order to elucidate population histories, migration events, or hominin phylogenies [129] [130, 105,131]. Despite the numerous studies using dental phenotypes as proxies for genetic markers, not many studies have compared these two data types [132], attempting to assess the congruence between both [133, 134, 135]. The results obtained from these studies were not free of controversy, some of them reported weak to strong correlations between dental phenotypes and genetic distances, whereas others did not find a correlation at all. Although there were some limitations, the oldest studies used serological markers, however, most of the recent analysis use SNPs or short tandem repeat (STRs) because of the highly polymorphic conditions [136].

One of the latest studies regarding this topic was published last year by our group [135]. The main objective was to study variation amongst dental traits from a group of admixed Colombians (same data employed in this thesis) and assess if this variation is useful to predict genetic ancestry proportions. Thirty-four dental traits on each cast were scored following the protocol described (Section 4.4.2.1) and the individual ancestry was calculated using genotyping data (SNPs). We carried-out biodistances analysis between the Colombian sample and reference continental populations samples, and the inference of genetic ancestry was obtained using a regression approach. The frequency of the dental traits in Latin Americans is in-between Native-

Americans, Europeans, and Africans. The biodistance analysis showed that Colombians are closer to Europeans than to Native Americans and Africans, this was corroborated with the mean ancestry estimates from the dental data and genetic data, where the European component (59% vs. 63%, respectively) was higher than the rest of the paternal populations (Native American: 32% vs. 28% and African: 9% vs. 9%, respectively). This shows that certain aspects of tooth morphology are likely affected by alleles, which present a marked difference in its frequency among continental populations [134, 132]. Although, we reported that dental traits reflect Latin American history, the prediction of genetic ancestry using dental traits is more complicated and further studies on admixed populations could contribute to identify genomics regions associated with dental features.

1.3.4 Genes implicated in tooth morphogenesis

Around 300 genes are differentially expressed in different dental tissues [106], in one or more tissues, or in the knots, or around the knots. This led to the hypothesis that a single mechanism and gene-network determines the form and spatial shaping of the knots[108].

There are four main signaling pathways involved in tooth development. Different studies using transgenic animals have elucidated the gene-networks *BMP, FGF, SHH* and *WNT*, their ligands and receptors are key in this process [106,137]. In most of the studies, the disruption of these signalling pathways causes severe damage to tooth development, i.e. tooth agenesis, anodontia, among others [138-141]. Bei et al. (2009) present a complete summary of these genes, which I present in Table 1.3.

There are many studies on the genes involved in tooth development, but currently, the interest is focussed on the mechanisms of how the shape, size, and number of teeth are determined. The Turing-like or reaction-diffusion model, suggests that the development of new knots is inhibited by the existing knots, utilizing diffusible extracellular signals. This mechanism consists in diffusible molecules, activators and inhibitors. The strength on the interaction between these molecules and their diffusion rates, define the patterns of spots or lines with different amount of activators and

inhibitors arising in space from homogeneous conditions [142]. This model is accepted, with some variations. Some studies propose that all the cells react to incoming signals in the same manner, while others consider that there is no inhibitor production until the knots appear, and they secrete the inhibitors [108].

Gene	Mutation	Tooth phenotype	Reference
Msx1, Msx2	Double mutant	Initiation stage arrest	Bei and Maas -1998
Dlx1, Dlx2	Double mutant	Initiation stage arrest	Thomas et al. -1997
Fgf8	Fgf8flox	Initiation stage arrest	Trumpp et al. -1999
Lhx6/Lhx7	Double mutant	initiation stage arrest	Grigoriou et al., -1998
Pitx2	Null	Initiation stage arrest	Liu et al. -2003
Gli2, Gli3	Double mutant	Initiation stage arrest	Hardcastle et al. -1998
P63	Null	Initiation stage arrest	Yang et al. -1999
Dkk1	K14 transgenic	Initiation stage arrest	Andl T etal. -2002
Pax9	Null	Bud stage arrest	Peters et al. -1998
Lef1	Null	Bud stage arrest	Van genderen et al. -1994
Msx1	Null	Bud stage arrest	Satokata andMaas -1994
Runx2	Null	Bud stage arrest	Aberg et al. -2004
Barx1	Null	Bud stage arrest	Tucker et al. -1998
Bmpr1a	K14 transgenic	Bud stage arrest	Andl et al. -2004
Fgfr2b	Null	Budstagearrest	De Moerlooze et al. -2000
Shh	K14 conditional KOconditional	Bud stage arrest	Dassule et al. -2000
Noggin	K14TG	Bud stage arrest	Plikus et al. -2005
Activin Ba	Null	Bud stage arrest, lack incisors and mandibular molars	Matzuk et al. -1995
Ctip2	Null	Late bell stage defect	Golonzhka et al. -2009
Gli2	Null	Abnormal maxillary incisor	Hardcastle etal. -1998
Gli3	Heterozygous	Maxillary incisor development arrested as a rudimentery epithelium thickening	Hardcastle et al. -1998 and Mo et al . -1997
Eda	Tabby encode eda	Small enamel knot	Tucker et al. -2000
Edar	Downless	Absent enamel knot, disorgonized enamel rope	Headon andOverbeek -1999
Fgf10	Null	Smaller tooth germ, cervical loops of the incisors are hypoplastic	Harada et al. -2002 Liu et al. -2008
Wnt/β catenin	K14 conditional KO	Misshappen tooth bud, ectopic teeth	
Ectodin/Sostdc1/wise	Null	Supernumerary teeth, enlarge enamel knot, abnormal cusp	Kassai et al.-2005
Apc	K-14Cre;Apccko/cko	Supernumerary teeth	Kuraguchi et al. -2006
Sp6	Null	Supernumerary teeth	Nakamura et al. -2008
Lrp4	Null	Supernumerary teeth	Johnson et al. -2005
IFT88/polaris	Null	Supernumerary teeth	Liu et al., -2005
Gas1	Null	Supernumerary teeth	Ohazama et al., -2009
Osr2	Null	Supernumerary teeth	Zhang et al.-2009
Sprouty2,4	Null	Supernumerary teeth	Klein etal.-2006

Table 1.3. Gene regions known to be implicated in tooth morphogenesis. Mutations caused on these genes lead to several anomalies in tooth development.

1.3.5 Dental abnormalities

Similar to other process during embryonic development, morphogenesis and development of teeth depend on complex interactions between the ectoderm and the mesenchyme [143,144, 145]. Lately, different experimental studies of the molecular and cellular mechanism controlling the morphogenesis and differentiation of the teeth, and dental pathologies have been performed using mouse embryos [146] and birds [147]. There is a severe genetic control of position, number, size and shape of the teeth.

Tooth morphogenesis is controlled by important family factors, fibroblast growth factors (*FGF*) and transforming growth factors (*TGF*, including the bone morphogenetic protein 4, *BMP4*). Also, Sonic hedgehog (*Shh*) morphogenesis molecule and the family of *Wnt* (Wingless) [148]. *FGF8* and *FGF9* are expressed in the proximal area of the molars and *BMP4* is expressed in the distal region of the incisors [147,149,150].

Currently, more than 200 genes involved in morphogenesis and differentiation of teeth have been discovered [151]. Many of them have been detected by GWAS, linkage association analyses and experimental studies. The candidate genes associated to tooth development and dental pathologies found by GWAS are presented in Section 4.1.1.

Dental agenesis is the total or partial congenital absence of teeth. There are different levels, (i) hypodonthia, 1 to 6 teeth are missing (excluding third molar); (ii) oligodontia, more than 6 teeth are lost (excluding third molar); and (iii) anodontia, total absence of teeth [148]. The genes so far associated with dental agenesis are *MSX1* [152], *PAX9* [153], *AXIN2* [154] and *EDA1* [155]. Here, I present two tables summarizing the candidate genes, phenotypic features and causal mutations of dental agenesis found by GWAS and familial studies (Table 1.4 and Table 1.5).

Gene	Locus	Inheritance pattern	Mechanism	Phenotype	Reference
MSX1	4p16.1	AD	3 mutations in exon 1, 4 amutations in exon 2	Hypodontia Premolar 2 and molar 3	Vastardis et al.
MSX1	4p16.1	AR	One point mutation resulted in substitution A219T.	Oligodontia	Chishti et al.
PAX9	14q12	AD	G219 insertion in exon2 frameshit mutation	Lacking of most permanent molars.	Stockton et al.
PAX9	14q12	AD	Transversion A340T	Non-syndromic oligodontia	Nieminen et al.
PAX9	14q12	AD	Ins C793, frameshit mutation	Non-syndromic oligodontia	Frazier-Bowers et al.
PAX9	14q12	AD	3 different missense mutations Arg26Pro, Glu9Lys and Leu21Pro	Oligodontia First molar affected	Das et al.
PAX9	14q12	AD	Transition C76T and C139T	Oligodontia	Lammi et al.
AXIN2	17q23	AD	Nonsense mutation, it activates the Wnt signalling pathway	Oligodontia	Khabour et al.
EDA1	Xq12	XLD	Nonsense mutation cross.c.193C>G	Absence teeth, hypodontia	Li et al.
EDA1	Xq12	XLD	2 nonsense mutations Glu316Gly and Thr338Met	Lack of central and lateral incisors, and canine teeth in upper and lower jaw.	Han et al.
EDA1	Xq12	XLD	3 new mutations Ala259Glu, Arg289Cys and Arg334His	Non-syndromic oligodontia	Song et al.
WNT10A	2q35	AD	Mutation	Non-syndromic oligodontia	Kantaputra et al.

Table 1.4. Genomic regions that have been associated with tooth abnormalities.

Chr	Gene region/nearest candidate gene	SNP	Dental abnormality Phenotype	p-value	Population	Reference
1	*ASCL5/CACNA1S*	rs4498834	Tooth agenesis	2.90E-14	Icelandic discovery sample (n=1944)	Jonsson et al., 2018
3	*FOXI3*	rs35822372	Tooth agenesis	3.40E-13	Icelandic discovery sample (n=1944)	Jonsson et al., 2018
2	*EDAR*	chr2:108,896,996	Tooth agenesis	5.90E-09	Icelandic discovery sample (n=1944)	Jonsson et al., 2018
2	*ARHGAP15*	rs2034604	Tooth agenesis	2.10E-14	Icelandic discovery sample (n=1944)	Jonsson et al., 2018
2	*WNT10A*	rs121908120	Tooth agenesis	6.10E-40	Icelandic discovery sample (n=1944)	Jonsson et al., 2018
2	*WNT10A*	rs121908119	Tooth agenesis	4.90E-08	Icelandic discovery sample (n=1944)	Jonsson et al., 2018
8	*ZFHX4*	rs371555610	Tooth agenesis	4.40E-11	Icelandic discovery sample (n=1944)	Jonsson et al., 2018
4	*LEF1*	rs917412	Mandibular second premolars	2.50E-10	Icelandic discovery sample (n=1196)	Jonsson et al., 2018
17	*NOL11*	rs758468472	Maxillary second premolars	3.30E-08	Icelandic discovery sample (n=600)	Jonsson et al., 2018
3	*FOXP1*	rs35956082	Maxillary lateral incisors	1.10E-11	Icelandic discovery sample (n=600)	Jonsson et al., 2018
X	*EDA*	rs55846652	Maxillary lateral incisors	6.70E-11	Icelandic discovery sample (n=600)	Jonsson et al., 2018
2	*PXDN, MYT1L*	rs11681214	Severe erosive tooth wear	3.99E–08	Northern Finland Birth Cohorts 1966 (n=1944)	Alaraudanjok et al., 2019
1	*AJAP1*	rs3896439	Caries in premolars, canines	2.00E-08	920 self-reported white participants	Shaffer et al., 2013
10	*LYZL2*	rs399593	Caries in mandibular anterior	9.00E-09	920 self-reported white participants	Shaffer et al., 2013
7	*SYPL1, NAMPT*	**rs190395159**	Decay, missing and filled teeth	7.14E−9	Carribean, Central and South American (n=11754)	Morrison et al., 2016
20	*BMP7, MIR4325, SPO11*	rs72626594	Decay, missing and filled teeth	2.75E−8	Carribean, Central and South American (n=11754)	Morrison et al., 2016
3	*IGSF10, MIR5186, MIR548H2, AADACL2*	rs138769355	Decay, missing and filled surfaces	3.59E−8	Carribean, Central and South American (n=11754)	Morrison et al., 2016
7	*SYPL1, NAMPT*	**rs190395159**	Decay, missing and filled surfaces	5.97E−10	Carribean, Central and South American (n=11754)	Morrison et al., 2016

Table 1.5 continues…

Chr	Gene region/nearest candidate gene	SNP	Dental abnormality Phenotype	p-value	Population	Reference
1	*GNG4, LYST, B3GALNT2, TBCE, GGPS1, ARIB4D*	rs138642966	Decay, missing and filled surfaces	1.94E−8	Mexican (n=4578)	Morrison et al., 2016
10	*ANK3, CDK1, RHOBTB1*	rs116717469	Decay, missing and filled surfaces	3.23E−8	Mexican (n=4578)	Morrison et al., 2016
17	*CACNA1G, ABCC3, ANKRD40, LUC7L3, MIR8059, WFIKKN2, TOB1, SPAG9*	rs71381322	Decay, missing and filled surfaces	3.72E−8	Mexican (n=4578)	Morrison et al., 2016
18	*None*	rs16946661	Decay, missing and filled surfaces	4.02E−8	Mexican (n=4578)	Morrison et al., 2016
X	*ACOT9, PRDX4, SAT1, APOO*	rs141563584	Decay, missing and filled teeth	3.89E−8	Puerto Rican (2006)	Morrison et al., 2016
1	*RHOU*	rs9793739	Decay, missing and filled surfaces	5.27E-07	~7000 participants non-Hispanic whites	Wang et al., 2012
4	*ADAMTS3*	rs1383934	Decay, missing and filled surfaces	2.96E-07	~7000 participants non-Hispanic whites	Wang et al., 2012
6	*RPS6K2*	rs635808	Decay, missing and filled surfaces	1.06E-07	~7000 participants non-Hispanic whites	Wang et al., 2012
8	*PTK2B*	rs17057381	Decay, missing and filled surfaces	4.02E-07	~7000 participants non-Hispanic whites	Wang et al., 2012
14	*CNIH*	rs4251631	Decay, missing and filled surfaces	2.13E-07	~7000 participants non-Hispanic whites	Wang et al., 2012

Table 1.5. Hits reported in previous GWAS of dental abnormalities. Associations detected by genome-wide scans from previous studies.

1.4 The significance of facial and dental features in Forensic Science

Facial and dental features show great variation in humans and are of considerable importance in forensics. In order to understand why it is important and beneficial to find new genomics regions associated with these traits I will explain the current situation on Forensic Science with regards to these features and how it relates to this study.

Nowadays, Forensic Science (FS) is a protagonist, both in the real world, helping to identify important figures from the history and victims from terrorism or accidents, finding the culprits on a crime scene, and in the fantastic world, such as television programs, where they show how exciting and intriguing FS is and how it is applied to criminal and civil laws.

Forensic DNA phenotyping (FDP) is used as a generic term for such DNA-based intelligence approaches. Its main objective is the prediction of externally visible features from DNA samples [156-158]. FDP also contributes with bio-geographic ancestry and estimation of age, using epigenetics. Furthermore, utilizing DNA can give information regarding the physical appearance of a person (missing individuals, skeletal remains, etc.) [159-161].

Over the years several externally visible characteristics-predictive genetic markers have been discovered by GWAS and prediction modules have been proposed, the most remarkable example are human pigmentation traits [158,162-165]. For instance, the prediction of categorical eye and hair colour phenotypes from SNPs is now available, with different validated systems [158,166,167].

1.4.1 Forensic Science and the human face

Forensic Science tries to establish both the legal requirement and social right identity; applying scientific principles to assess the data collected by different sources. The main goal is to prove or disprove the guilt of a suspect with high reliability for the legal system. Forensic science tries to find the guilty using any source available, eyewitnesses, closed-circuit television (CCTV) systems, applying automatic facial recognition, amongst others.

All the methods described previously are about the recognition of the face by witnesses or machines, but what happens when there are no witnesses or cameras, or there is no certainty regarding the criminal evidence or the recognition of the body. DNA is the option, but the current methods need a database or DNA from a family member to compare and obtain the results. Therefore, Forensic DNA Phenotyping is the solution. Some studies have suggested to predict the face of a person from DNA [5,168], this has been called 'Reverse Genetics for forensic prediction'[68]. They propose that reverse genetics is about a known function of one or several genes which will participate on the shaping of a particular phenotype with a degree of certainty [168] similar to the eyes and hair colour FDP model prediction. And this could be utilized to predict the shape or size of the phenotype, i.e. height, several SNPs have been found associated to this trait. Weedon et al. (2008), showed that there is a difference of 5 cm between people carrying 17 or less "tall" alleles and individuals carrying 27 or more alleles associated with bigger size [169].

This sounds promising, but DNA prediction is still far from an accurate forecast of the human face, albeit some authors have claimed that currently there is the potential to identify an individual from a few people group of candidates [168,170,171]. Using different approaches such as bootstrapped response-based imputation modelling to generate imputed predictor variables, which help to model different co-variables (sex, ancestry, etc), thus the independent effect of specific alleles on facial traits is unveiled [168]. Other authors, more daringly, utilizing whole-genome sequencing, detailed phenotyping and statistical modelling, attempted to identify individuals, but the results were limited with regards individuals, each predictive model added little to the identity of a person [5].

1.4.2 Forensic Science and dental traits

Following the same approach as with facial features, the teeth and its traits are very important for Forensic Sciences. Forensic dentistry deals with the examination of dental evidence and dental findings to identify an individual or to find the culprit of a crime [172]. Teeth are also important because they remain unaltered even when exposed to high temperatures, which allows to preserve the DNA inside the tooth [173]. Teeth are used to determine age and sex. Age is assessed based on developmental and degenerative changes. For instance, if the third molar has emerged, this tends to happen

around 17-19 years of age or once the teeth have erupted, the wear down starts, there is an intuitive connection between age and wear [174, 175, 172]. Sex determination has an important role in the identification of individuals in natural catastrophes, terrorist attacks, plane accidents, among others [176]. It is well known that teeth show sexual dimorphism, this feature is used to differentiate between men and women, the dimensions of the teeth are a reliable method to determine sex [172]. The morphology is also utilized to identify gender, distal accessory ridge in canines is more pronounced in male than female [177]; and women have a smaller number of cusps in the mandibular first molar [172].

1.5 Latin American genetic history and its implications for genetic studies of human phenotypes

In order to comprehend how diverse populations, both phenotypically and genetically, may increase the power to detect novel genomic regions and improve the fine-mapping resolution in association studies, it is particularly important to understand how the current Latin American population was shaped. In this section I will briefly describe the settlement of the American continent, mainly Central and South-America, and subsequently I will refer to the results of sub-continental ancestry and population structure in the samples used on this thesis, belonging to the Consortium for the Analysis of the Diversity and Evolution of Latin America (CANDELA). These analyses are described in Chacón-Duque et al. (2018). Finally, I will describe how beneficial it is in using the variable admixture in the Latin America population to identify genomic regions in which variation in ancestry is correlated to phenotypic variation.

1.5.1 Genetic history of the Native American

There are many theories about the time, number of migratory waves and the speed of the settlement of the American continent. The theory that Amerindians descend from at least three streams of migration, coming from East-Asia through the Bering Strait and most of the native American groups descend from the first stream of Asian gene flow has been widely supported by several studies [178]. The Beringian Standstill hypothesis suggests that the first settlers arrived in present day Siberia around 30,000

years ago [179] and they were isolated for extended periods of time in Northern Siberia, due to the Last Glacial Maximum (LGM) [180]. Subsequently, in 2015, Raghavan comparing genomes of modern people, supported this hypothesis, although they proposed that the arrival was no earlier than 23,000 years ago [181]. Interestingly, a recent study based on Y chromosome sequences suggested the entry south of the ice sheet was after 19,500 years ago and the star-like expansion was 15,000 years ago, they estimated that Beringian Standstill duration was around 4,600 years, shorter than the time previously proposed [182].

These groups started to settle from the northwest until the southernmost tip of America, Monteverde-Chile [183, 184, 182]. And the definitive establishment in South America was 12,000 years ago [182]. Whole-genome studies have documented that ancestral Native Americans (NA) come from Siberians and East Asians, and around 2,000 years later, there was a split between later NAs and Ancient Beringians (AB). North Native Americans (NNA) and South Native Americans (SSA) diverged from the later Native Americans (~17,500 and 14,600 years ago) [178, 181, 185, 186]. This movement was the precursor of several migratory waves through and within South America. Mesoamericans, and east and west Andeans were the most divergent groups [187].

Recent studies using genome sequences have also attempted to explain other matters, such as earlier migrations into Americas by Australasians, the amount of population splits and the velocity with which this occurred. Moreno-Mayar et al. (2018) found that the initial peopling was more extensive and faster than the later ones [186]. This suggests the access to the different geographic areas was unrestricted. Interestingly, a clear signal linking NA in the Amazonian region of Brazil to present-day Australasians has been detected [188, 187], although how early NA lineage were related remains unclear [187]. The intricate geography, the rapid expansion and social structures gave rise to a complex population. In North America a strong population structure, and in South America, due to multiple uneven migrations, Australasian and Mesoamerican signals can be detected in present-day Native Americans.

1.5.2 Demographic history of America

On the arrival of the Spanish conquerors during the XVI century, the American continent had been settled by native American groups, descendants of the first inhabitants of America [189]. The number of native people during the arrival of Europeans has been widely discussed; from ten to one hundred million, ten million being the most likely [190] (Figure 1.9). They were spread throughout the continent and exposed to different geographical and environmental conditions. Thus, when the Europeans landed in the Bahamas in 1492, there were different native American groups, using different languages, garments, means of subsistence and social organizations. There were groups living from agriculture (mostly from Mesoamerica and the Andes) and the rest were hunter-gatherer communities (mainly from Patagonia and North America) [191]. The number of Europeans who arrived in the New World was approximately half a million Spanish and approximately the same number of Portuguese [192, 193]. The conquerors founded colonies in the Caribbean and on the coastal mainland (including the Pacific coast), the inland colonization was usually following Native American colonies [184]. African slaves were introduced to the continent by the Spanish and Portuguese, especially during the collapse of the NA groups – it has been suggested that ~90% of the Native American population was extinguished – after the arrival of the Europeans [194] (Figure 1.9), although this has been questioned for Peruvians and Mexicans by some recent studies [195,196]. The African slave trade was an established institution by the settlers to exploit the lands taken from the natives. Approximately, 10 million African slaves were introduced to America, ~ 50% were taken to Brazil and the rest to British, Spanish and French colonies [197] (Figure 1.10).

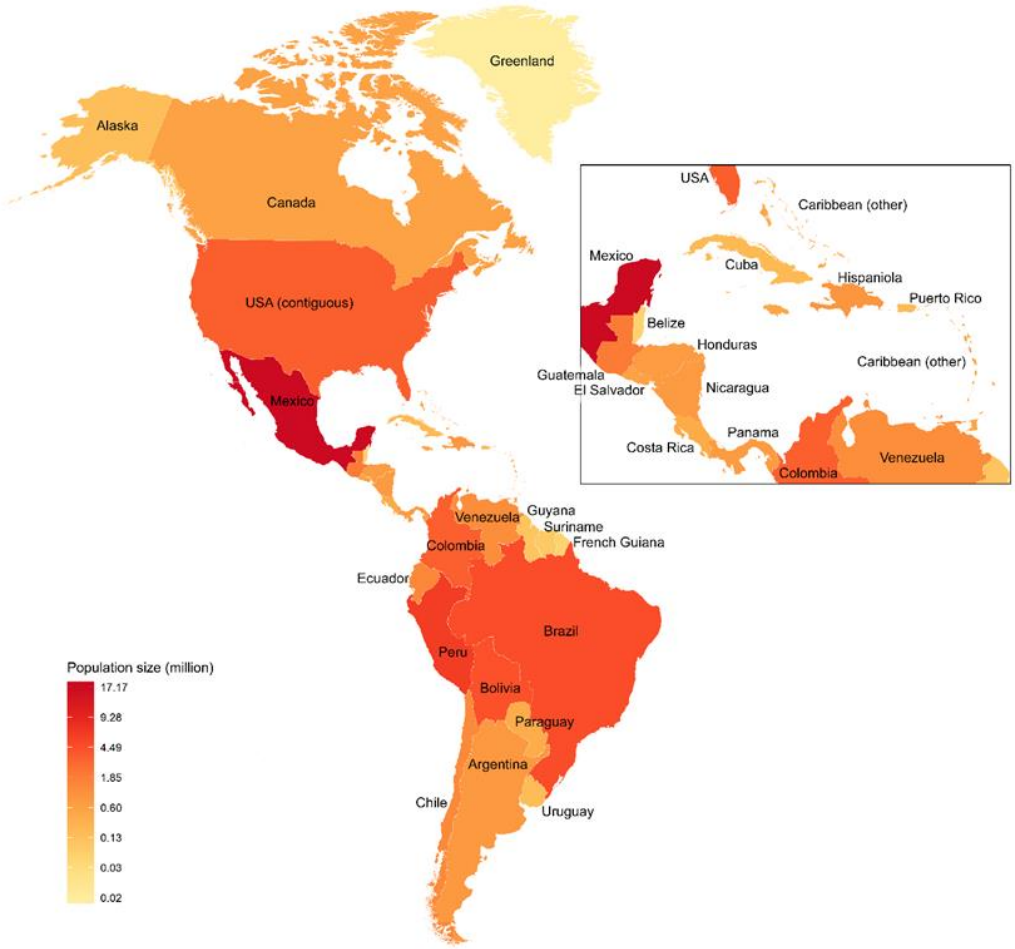


Figure 1.9. Estimated size of the Native American population at the arrival of Europeans. From Adhikari et al. (2017).

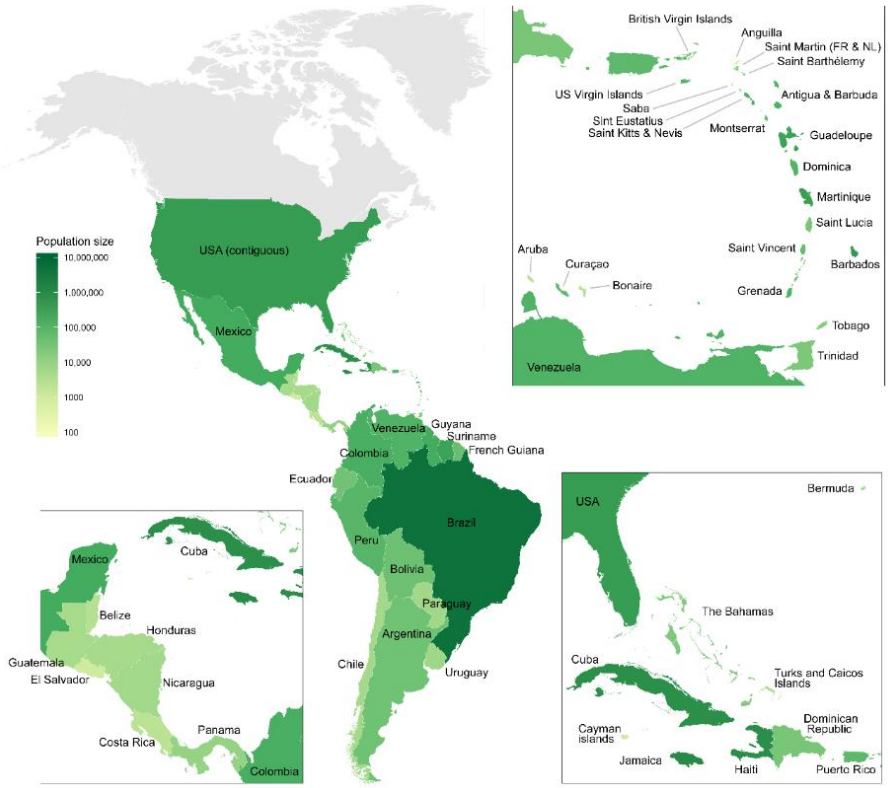


Figure 1.10. Estimated number of African slaves that were transported to America. From Adhikari et al. (2017).

1.5.3 Genetic history of Latin American

At the arrival of the Spaniards around 80% were men, leading to a marked sex bias during the admixture, that involved European men and native women [198]. This was first pointed out by early mt-DNA and Y-chromosome studies, where Y-chromosome traces the paternal ancestry to European and mt-DNA to maternal Native Americans [199, 200, 201]. These studies also show that the most common haplogroups (mtDNA) currently present in admixed Latin American populations correspond to Native American population living in the same region [199, 202], this suggests 'genetic continuity', despite the complexity of the history in Latin America, native American groups are totally incorporated to the current populations [199,203]. Autosomal markers and X-chromosome studies have been compared with uniparental marker analysis and they have found that estimates of European ancestry based on autosomal markers are higher than X-chromosome estimates, as expected due to women contributing with two chromosomes, similar to the previous findings based on mtDNA/Y-chromosome [203-207] (Figure 1.11).

Advances in genomic analysis have allowed an increase in the resolution and accuracy of population and individual ancestry estimates – contrary to what it might seem – Latin American populations present an extensive population structure, individual admixture shows great variation in Africa/European/Native ancestry through different geographic regions [59, 208, 209] (Figure 1.12). This means there have been several complex changes within the continent, such as population substructure and other migratory flows from different regions of the world after the main population admixture during the first colonial settlements [210 195].

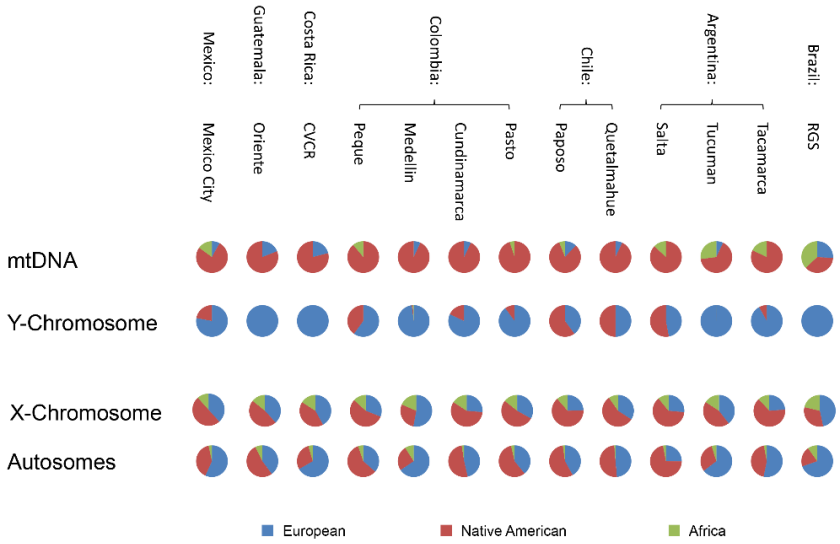


Figure 1.11. Proportion of African, European and Native American ancestry based on mtDNA, Y-chromosome, X-chromosome and autosomal markers estimated from thirteen Latin American populations. The proportion of Native American ancestry estimated from mtDNA is higher, contrary to European ancestry which is bigger when analysed in the Y-chromosome. This pattern points-out the sex-bias during the European colonization of Latin America. From Adhikari et al. (2017). Modified from Wang et al. (2008).

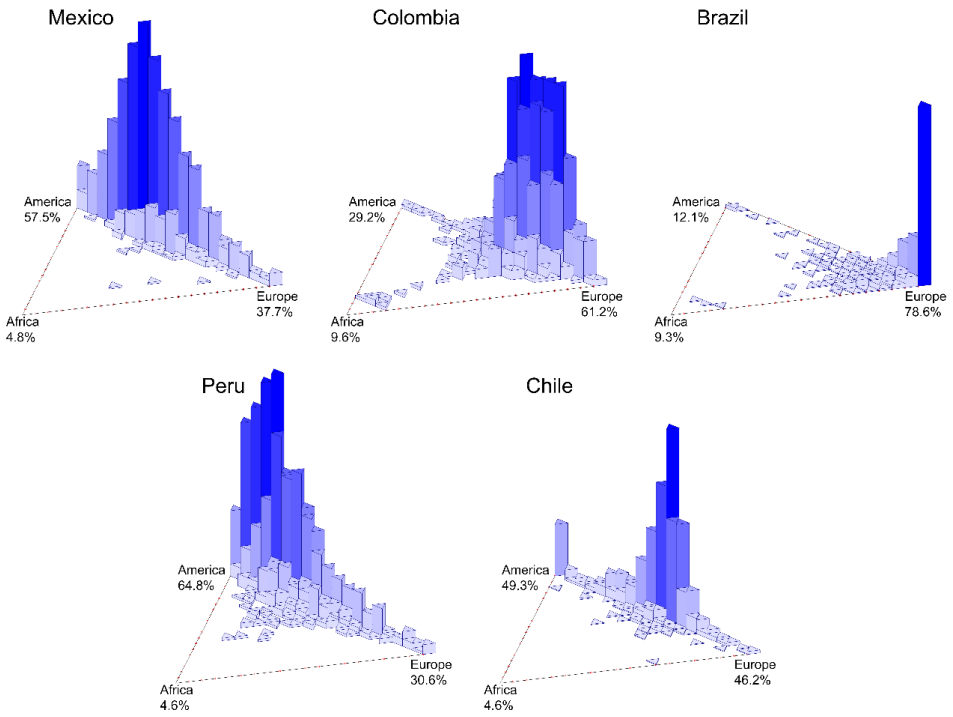


Figure 1.12. Individual ancestry proportions in Latin Americans. Regarding the three main paternal populations (Africa, Europe and Native America) estimated from 93,328 SNPs genotyped in 6357 Latin Americans from Brazil, Chile, Colombia, Mexico and Peru. Mean admixture ancestry is on the corner of the triangle plots. Adapted from Adhikari et al. (2016) [211].

1.5.4 Sub-continental ancestry

In addition to calculating continental ancestry, high-density genetic data has allowed to explore the sub-continental ancestry in Latin American populations, i.e. haplotype-based analysis instead of independent SNPs. Moreno-Estrada et al. (2014) using Ancestry-Specific Principal components (AS-PCA), confirmed that the native American component from individuals from central Mexico correspond to the native American group Nahua, whereas individuals from south-east Mexico present Maya native American ancestry [208]. This coincides entirely with the geographic location of these native groups. A similar study that explored the subcontinental ancestry in South America, found that Peruvians carry native American genomic segments that corresponds to Quechuas and Aymaras, whereas Colombians show a higher affinity with native American groups from Amazonia and coastal tribes situated in northern South America. Regarding the European ancestry, they found that as expected, most of the European component corresponds to the Iberian Peninsula, but there are also signals of Italian ancestry, especially in Argentina [206 195].

Chacon-Duque et al. (2018), utilizing haplotype-based methods, analysed the genetic ancestry from the CANDELA cohort (Section 1.7), a present-day admixed population from Latin America [64]. The findings of this study not only tie in with previous studies mentioned above [206,208], but also extends these results inferring 25 Native American ancestry components, enabling a deeper and more accurate estimation of the native ancestry in the CANDELA sample. The native American component varies considerably across countries, and there is a strong differentiation among the native ancestry of each country related to the geography and native American reference groups, highlighting the genetic continuity between pre-Columbian groups and the current admixed populations in Latin America. The Nahua sub-component is found in northern and central Mexico, and two smaller sub-components, one related to Mayans (Yucatan area) and another to natives of south Mexico. The Quechua component is present in Peruvians in central Peru. In Northern Peru there appears a sub-component related to Natives from the Andes; the Aymara sub-component is present in Southern Peru. In Chile, the Mapuche sub-component is present in the southern regions, and in the north of Chile, bordering with southern Peru, the Aymara component is present. In Colombians, there is higher affinity with the Chibzan-Paezan Natives from Colombia and lower Central America, they are also related to Central American Maya, and in the south with the Peruvian Andean component. Brazil CANDELA samples shows lower native American ancestry in comparison to the rest of the countries, because most of the samples were obtained in an area with high recent European immigration [59], however the native American component shows affinity with populations from the Amazon basin.

The patterns of European ancestry in the CANDELA samples match perfectly with the history of Latin America. There is a pronounced differentiation between Brazil, in which the predominant component is Portugal/West Spain, and the Spanish American countries, where Central/South-Spanish ancestry is the most significant. In Brazil there is also an important genetic component closely related to Italian and German reference groups, which reflects the documented migration to Southern Brazil from Italy and Germany, at the end of the 19th century [192].

Sub-Saharan, East Asian and East/South Mediterranean ancestry in CANDELA individuals were found in less amount than Native and European ancestry. Sub-Saharan components in the full CANDELA sample is <4%, and around 22% of the

volunteers show more than 5% of Sub-Saharan African ancestry. East Asian ancestry is very low (<1%) in all the countries, except for Peru (1.4%). This coincides with historical documents on the arrival of Chinese and Japanese laborers (mid-19th century to early 20th century) to Peru and Brazil, respectively. Interestingly, the Sephardic/East/South Mediterranean component is detectable in all countries (~1% to 4%). Around 23% of the individuals represent more than 5% of such ancestry, an average of 12.2%. These findings show that this signal is much more prevalent than it has been suggested by other studies, where they mention that the East/South Mediterranean component is present in certain Latin American populations [212, 213].

Ongaro et al. (2019) in their recent study used a genome-wide database from individuals from 12 countries (~12,000 people) and ~ 6,000 individuals from different populations from the world as reference samples, to perform haplotype-based methods in order to analyse how historical movements coming from outside of America affected several aspects of the New World, such as the admixture profile and the time when this admixture happened [195].

Regarding the admixture events this study found similar results to those presented in the CANDELA cohort [64]and previous studies [208]. With regard the time when the biggest mixing event occurred, Ongaro et al. also showed a great admixture event among Native Americans, Europeans and African populations, between 6 to 12 generations back [64].

Furthermore, they present an interesting approach, using the haplotype information from single individuals in their database, they can expose other ancestries that are not detected when using population haplotypes data. Thus, they detect the signal of waves of migration occurred during the last 500 years, as reported in in anthropological and historical documents [195]. This is because the degree of variation in the haplotype distribution is better detected. The admixture time in single individuals showed the complex process of admixture present in America. In the case of African ancestry, a general North to South temporal pattern [214] was observed, reflecting that the slaves from Senegal and Gambia were deported to America earlier during the colonization, until 1640 [215], this region of Africa was the main provider of slaves for Spain. A high proportion of ancestry coming from Angola and Namibia was detected in the analysed Brazilian populations, according with the Portuguese settlement in Angola at the

beginning of 18th century [214, 216]. The gene flow from different European sources also showed different moments of admixture. First, from the Iberic Peninsula, later from France and Great Britain sources. And during 1850 onwards, there were several admixture events from Italian sources, that contributed mainly to European Americans from USA, Argentinian and Brazilian populations [195]. Interestingly, Ongaro et al. (2019) revealed that the Native-American component did not show a decrease in Mexicans and Peruvians, such as the experienced by the rest of the continental ancestries [196], differing with historical records that have reported a dramatic reduction of Native Americans once the colonization started [194].

Ongaro et al. (2019), Chacón-Duque et al. (2018), Moreno-Estrada et al. (2014), among others have done a great contribution to the genetic history of America, by giving support to the historical records and showing how and when the genetic admixture of America has been shaped, and this is crucial for the development of medical approaches, such as genome-wide scans to improve the fine-mapping resolution, which in the future could lead to personalized medicine and other benefits in the Latin-American and other populations worldwide.

1.5.5 Genetic diversity and physical appearance in Latin America

Latin American populations show huge phenotypic diversity [59], this is due to the relatively high genetic and phenotypic diversity of Africans, Europeans and Native Americans. There are many studies that have shown that the risk of disease in Latin Americans correlates with the continental and sub-continental ancestry [217, 218, 219]. Except in cases where the correlation between the disease risk with the ancestry is due to environmental covariates, the correlation seems to be due to underlying susceptibility alleles with a highly variable frequency among Africans, Europeans and Native Americans [220, 221]. This is the base to 'Admixture-mapping', which identifies trait loci with highly different allele frequencies in the populations that contributed to the admixed population [222,223]. The same approach is used to seek genomic regions associated with physical features. Many of these traits are highly heritable [224,225]. There is a significant correlation between the ancestry and the phenotypes. In the case of Latin Americans, a wide range of features can be observed, similar to the parental populations [59].

1.6 Overview of Genome Wide Association Studies (GWAS) for facial features

More than ten years have passed since GWAS statistical analysis was developed. In the beginning, there was a misconception and mistrust about this analysis, because there was not enough empirical evidence to support GWAS discoveries. But currently, it has been widely accepted because there are robust and outstanding results, which provide support to these findings [226] .

The main objective of GWAS is to contribute to a better understanding of the biology behind the trait or disease, thereby leading to prevention and better treatment. Genome wide analyses investigate the entire genome, using genetics variants, such as SNPs. Mainly focused in associations between these variants and a trait or major illnesses. Genome-wide association studies intend this by using linkage disequilibrium (LD) analysis to map genomic loci that have an effect on complex traits or diseases [227]. Its success relies on (1) the segregation of the loci affecting the trait on the population, (2) the combination of the allele frequency and the distribution of effect size, (3) genetic architecture, (4) the coverage of the panel of genetics variants used in the GWAS and (5) the heterogeneity of the studied illness or feature [226].

Genetics has been searching for associations between genetic variants and phenotypes for decades. But association signals between genotypes and complex traits are distributed across the entire genome, unlike Mendelian diseases, which are caused by mutations in one gene in the genome. This important difference was supposed to be solved by GWAS, when it was developed, and to a certain extent it was, but many of the discovered variants only explain a small fraction of the phenotypic variability due to genotypic variation, this was called the mystery of the "missing heritability" [228-230]. An explanation for this issue is that SNPs, which do not reach genome wide significance, contribute largely to "missing heritability" [231,232].

Another difference between complex traits and Mendelian disorders is that while the latter one is caused mainly by protein-coding changes [228,233], complex traits are caused by noncoding variants that may affect gene regulation, i.e. promoters, enhancers, etc. [234]. These are conclusions that have arisen from the past decade, of GWAS analysis.

There are also specific problems that researchers doing genome wide studies have to face, such as ascertainment bias, this is a systematic distortion in measuring the allele frequency in the studied population, it depends on the allele frequency in the population for a specific allele and hence, the haplotypes. The older the studied population, the higher the variability of the allele variants and haplotypes are smaller due to the recombination rate in this population. It is therefore necessary to try to cover as many genotyping variants as possible [235]. This can be partly solved with imputation, a method used to infer the missing SNPs in a haplotype based on reference population panels, such as 1000 genomes (1KG) and HapMap projects [236]. Population stratification is another important issue involved in GWAS, which is also due to the systematic difference in allele frequencies in different subpopulations from a population, probably due to different ancestries. Because of this difference, these subgroups will cluster apart and they will be over or underrepresenting the associated variants, causing a false positive association between SNPs and the trait or disease analysed. Principal Components Analysis (PCAs), are used to adjust genotypes and phenotypes by amounts of assigned ancestry through the top Principal Components (PCs) [237].

One of the greatest problems that has become evident, with GWAS is the issue of diversity of the populations utilized to perform the analysis not reflecting the level of variability that currently exists in the world [238]. The incomplete representation of ethnically diverse populations leads to a lack of capacity to understand the genetic architecture behind the assessed phenotypes. These studies cannot be used in clinical practice or public health policies [239].

Contrary to complex traits, a pathogenic variant in a Mendelian disease will cause the sickness, despite the population. Contrastingly, in complex traits a causative variant will differ from one population to another [239] (Figure 1.13A).

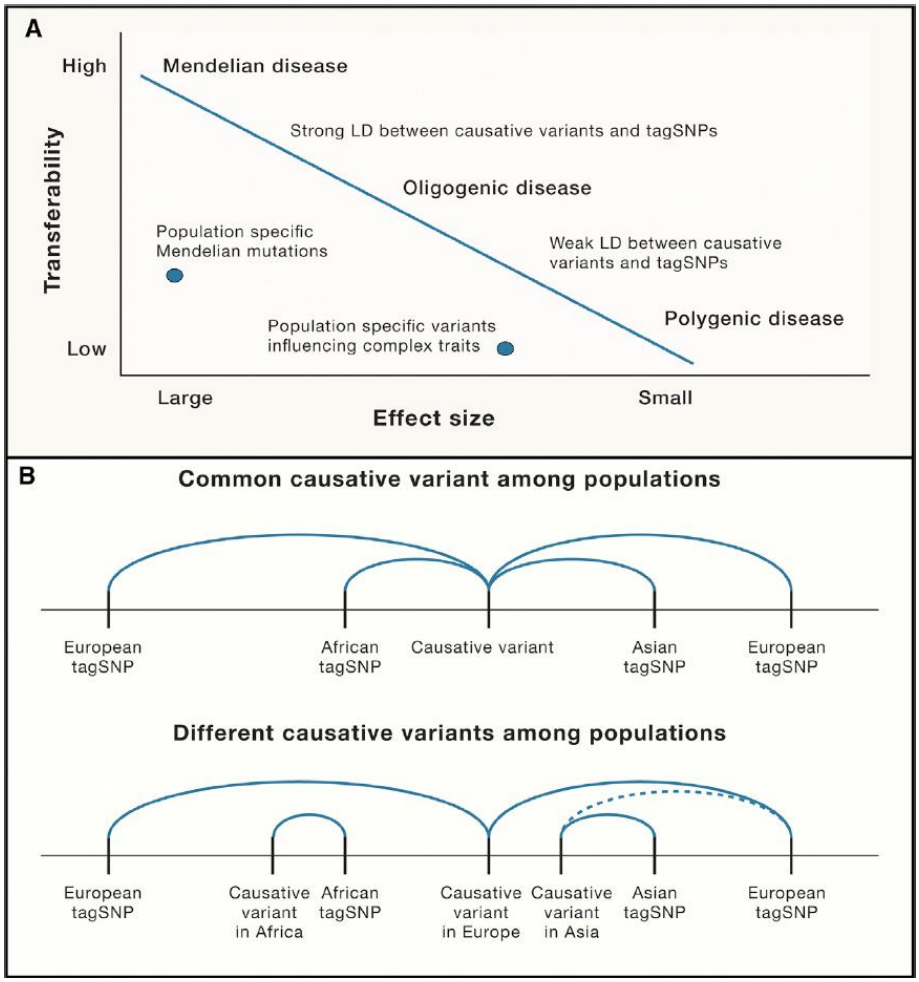


Figure 1.13. Factors affecting replication of GWAS associations. A) Causative variants of Mendelian diseases have large effect sizes. Thus, the disease will occur, regardless of the population. In contrast to complex traits, where there are a few genetic associations with the trait, but each of them has a moderate or small effect size, therefore the transferability through generations may be limited. B) Linkage disequilibrium plays an important role in the replication of association studies. In the example at the top, the causative SNP is tagged by different SNPs across different populations. In the bottom example, there could be several causative variants in different populations, which are tagged by different SNPs, not allowing replication across genome-wide studies. Figure from Sirugo et al., 2019 [239].

The diverse evolutionary history of the different populations is an important reason why different mutations on the same gene could cause the same disease (allelic heterogeneity), and this variation on the causative mutation may confound the diagnoses and treatment [239]. For instance, cystic fibrosis is a Mendelian disease, the prevalence is 1 in 2000-3000 births in Europe [239], however in African Americans this number drops to 1 in 17000 births. This is because the causative mutation is different in both populations, and the African American variation is underdiagnosed [240]. This takes on far greater significance when talking about complex traits. Due to many variants affecting the same trait/disease, each of them causes a small effect and the

mutations affecting the phenotype vary across populations. This highlights the necessity of wider variation on GWAS populations utilized. By 2018, 78% of the volunteers participating in GWAS were Europeans, 10% were Asian, 2% were African, 1% were Hispanic, around 6% were not reported and 3% were from multiple ethnicities [239]. These differences must be solved, especially considering genome-wide associations may not replicate across different ethnicities.

The replication of genetic associations varies across diverse ethnic groups due to several reasons: 1) Linkage disequilibrium (LD) reflects the demographic histories of the populations around the world, thus the LD between a causative variant and tag SNPs will be different from one ethnic group to another [241] (Figure 1.13B). 2) The genetic architecture also affects the lack of replication of GWAS because of population specific variation and changes in allele frequency, which is the result of genetic drift, local selection or both, and subsequent local adaptation events [239]. 3) Another issue is the epistasis due to variances in genetic backgrounds, as well as gene-environment interactions present in different populations [242].

Genome-wide association study have discovered more than 10,000 significant associations (p-value threshold of $5x10^{-8}$) between complex traits, diseases and genetics variants [243]. The first "official" GWAS paper was published in 2007, this publication from Wellcome Trust Case Control Consortium (WTCCC) is considered as the first one, because it was a large, well-designed GWAS for complex diseases with a chip that covered a wide range of genetics variants [244] . After this starting-point, many GWAS have been developed searching for different associations. Not many GWAS regarding facial morphology have been published (Section 3.1.1), only a few mostly in European and Asian populations, following the dominant trend. And just a reduced number of genes have been linked with traits located in the upper, middle and lower face, such as *PAX3* associated with nasion position [7,8], *PARK2* related to midface height and *FREM1* linked to upper lip height [245], all of them found in European populations. The *EDAR* gene, significantly associated with type of chin [246], *ENPP1* and *FGFR1* genes influencing lower and anterior face height [247] were discovered in an East-Asian population, called Uyghur, taking advantage of the admixed population of East-Asian and European ancestry [246].

1.7 Consortium for the analysis of the diversity and evolution of Latin America – CANDELA

The consortium for the analysis of the diversity and evolution of Latin-America (CANDELA) seeks to characterize the population of Latin America, with regards to physical appearance and ancestry from genetic and social perspectives [59]. More than 6000 volunteers form part of this joint effort from five countries in Latin-America; Brazil (Porto Alegre), Chile (Arica), Colombia (Medellin), Mexico (Mexico City) and Peru (Lima) (Figure 1.14). Ethical approval was obtained in each country and also in The University College London, UK. Adult subjects of both sexes aged between 18 and 45 years old were invited to participate mainly through public lectures and media presentations. A wide range of data was obtained from each participant, family information, different phenotypes, such as skin, hair and eye color, different anthropological measurements, and questionaires to assess socio-economic status and self-ancestry perception. More than 730,000 SNPs were genotyped on the Illumina Omni-express bead chip. A full description of the samples can be found in Ruiz-Linares et al. (2014).

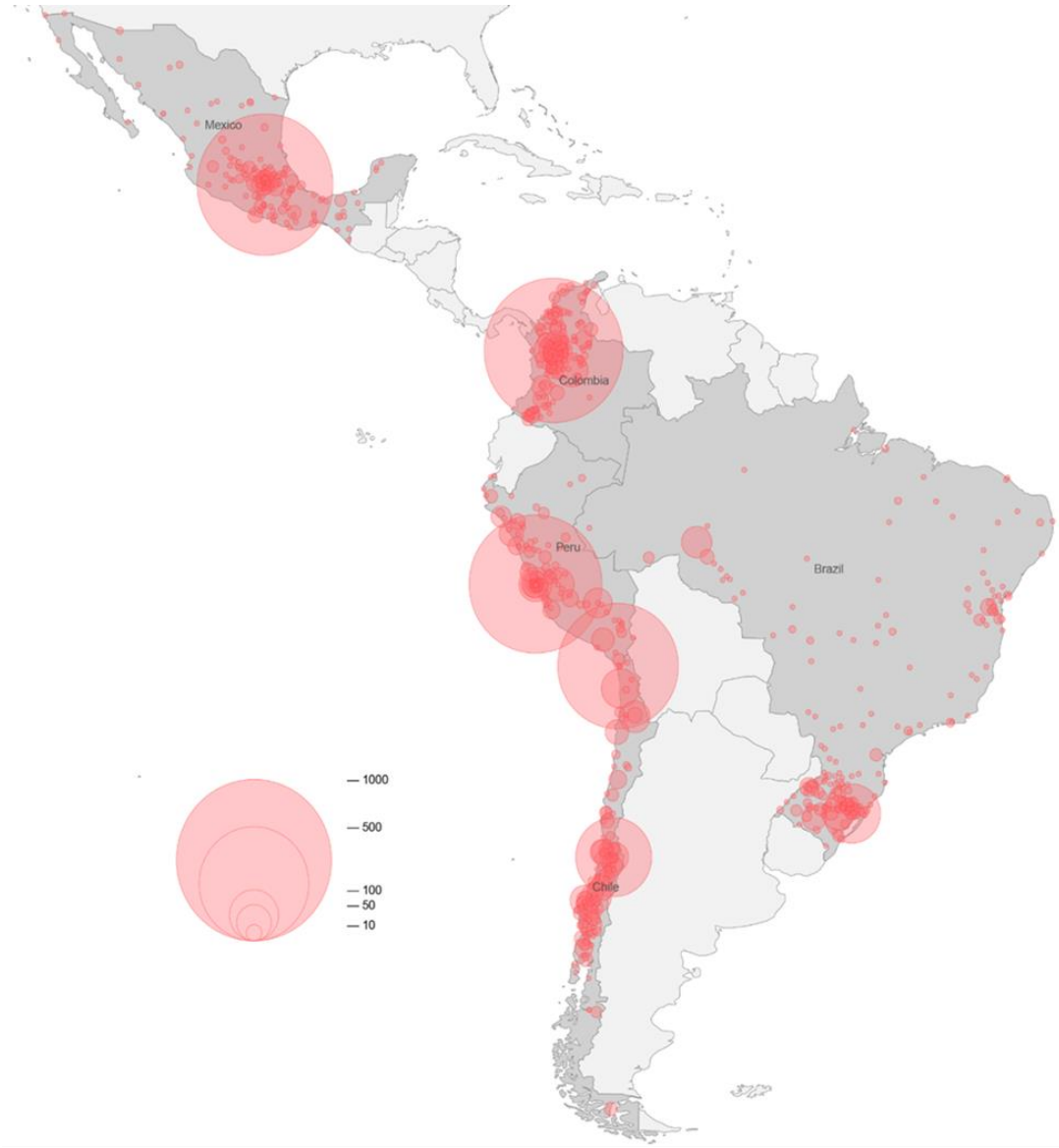


Figure 1.14. Geographical origin of CANDELA individuals. Circles depicts the number of people belonging to specific geographic origins. Figure adapted from Ruiz-Linares et al., (2014).

1.8 Summary and guidelines of this thesis

In chapter 1, the Introduction, I have explained the importance of facial and dental features in different disciplines. I have also described the current knowledge about the evolution of the face and teeth. I have presented a summary about the genetics of the face and teeth, summarising the existing information. I have also described the genetic history of the settlement of America, focusing in Central and South America. I have presented a comprehensive overview of GWAS, advantages and limitations. And finally, I have described the studied sample. All of these supports the main aim of this thesis, to identify genetic loci associated with human facial and dental traits in a Latin American population.

In chapter 2, I describe the core statistical analysis I have used in this work, the genome-wide association study. I also explain all the stages involved in this analysis, based on the pipeline used in our group. In some points I have briefly mentioned the theory involved in that specific step of the analysis. Subsequently, I describe quantitative phenotyping method, the geometric morphometry. For both methods I briefly describe the history, and then I explain the method themselves.

In chapter 3, I conduct a GWAS on ordinal facial features. Subsequently, I report both novel variants associated with facial features and the replication of genetic variants previously associated to face traits in a Latin-American population. Additionally, I explain the methods and results of the statistical and experimental analysis conducted to understand the biology behind the associations and to ratify the associations between the candidate genes and facial features. I also compare the associations obtained by other GWAS conducted after the results of this study were published.

In chapter 4, I perform GWAS on ordinal and quantitative dental features. Then, I describe the genomic regions where the significant genetic variants belong. I also conduct a local-ancestry inference analysis to verify the origin of the haplotypes of some of the significant SNPs. Finally, I discuss about the associations and how these analyses may be improved, increasing the number and the genetic variability of the samples.

In chapter 5 I discuss and draw my conclusions on the results obtained in this thesis and I describe the possible directions and significance and advantages of genome-wide studies now and in the future. I briefly discuss about the importance of the associations found in this thesis and how they are related to the development of the face and teeth and finally I discuss about the prediction of complex traits based on genetics data.

In each chapter presented in this thesis, there are tables and figures. In the cases where additional tables and figures contribute to a better understanding of the chapter, they are in the thesis appendices and referred to by letters (A-B).

Chapter 2: Methods

2.1 Introduction

Over the years, technology has improved rapidly, contributing to the development of stronger analysis tools, such as genome wide association studies (GWAS). More than 15 years ago, this type of analysis would be impossible to perform due to the number of samples that need to be processed and the amount of data that needs to be analysed. Regarding this, GWAS appears as a powerful tool to elucidate genomics variants behind different phenotypes. Linkage has been used to map genomic loci associated with diseases or complex traits, but this relies on the co-segregation of causative variants (new mutations that affect the increase or decrease of the presence of a disease/trait in a population) and marker alleles within pedigrees. Although the number of recombinations per meiosis is relatively low, making it easier to track a causative variant within a few genetic markers, the weakness is the low mapping resolution (how close the causal variants is located with regard the linked markers) [227]. Mapping genomic loci associated to Mendelian diseases have been successful [248], but it has not been the same for complex traits, one of the reasons is that each causative variant has a low effect (penetrance) on the trait, therefore it is more difficult to detect the co-segregation within pedigrees [227]. Genome-wide association studies appears as a powerful analysis that can be used to scan a big amount of candidate markers against different phenotypes (diseases and/or traits). Currently, several commercial companies produce chips that allows genotyping from hundreds to millions of SNPs in hundreds of samples at the same time. Nevertheless, genome-wide association scans represent one side of the analysis, which is affected by several variables, within these variables, phenotyping plays an important role. The type, accuracy of phenotyping and how the phenotype is assessed are key in a successful analysis [249].

In this chapter I will describe the main methods I used in this thesis, genome-wide association analysis and one phenotyping method, geometric morphometry, which is used to analyse facial traits

In order to describe the methods used to perform the genome-wide scans I will use the standard pipeline our group uses to run the GWAS, which was developed by the statistician of the group K. Adhikari. Regarding the geometric morphometry method, I will contextualize the history of the method and how it is developed in order to better explain the use of this analysis.

2.2 Genome- wide Association Analyses

The genotyping data is transformed by GenomeStudio Genotyping Module v1.0 software from Illumina into another format. This software takes the raw intensity files, subsequently makes genotype calls (score of a quality metric that indicates the reliability of the genotypes called) [250] and then converts them into ped/map format (format used by the software PLINK [251]).

Chip genotyping performance is not a totally perfect technology. SNP genotype calling from fluorescence intensity signal is done by a computer mathematical clustering algorithm, which, for a given individual and for a given SNP locus, estimates the likelihood that their genotype is aa, Aa and AA. But this sometimes fails. Therefore, QC is necessary to remove markers or individuals with high errors rates [252] (Figure 2.1 and Figure 2.2).

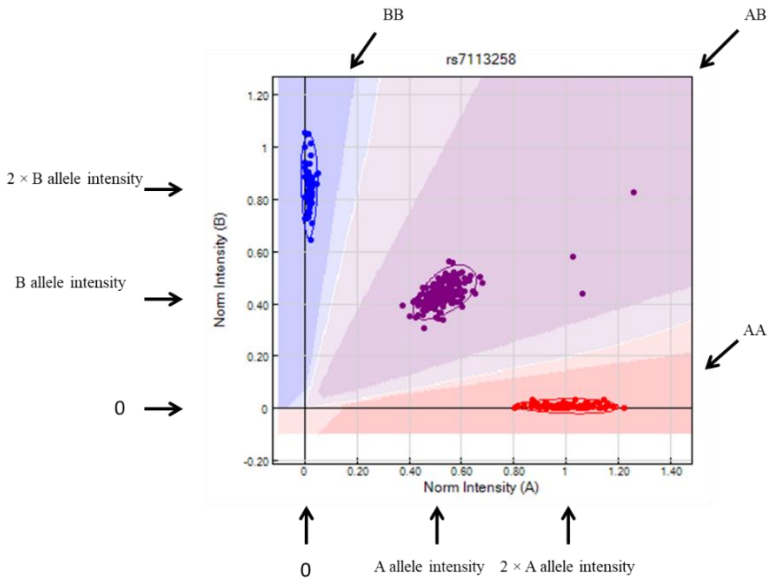


Figure 2.1. Example of clustering of SNPs genotyped correctly. In this case the genotyping worked very well. The 3 clusters are totally separated.

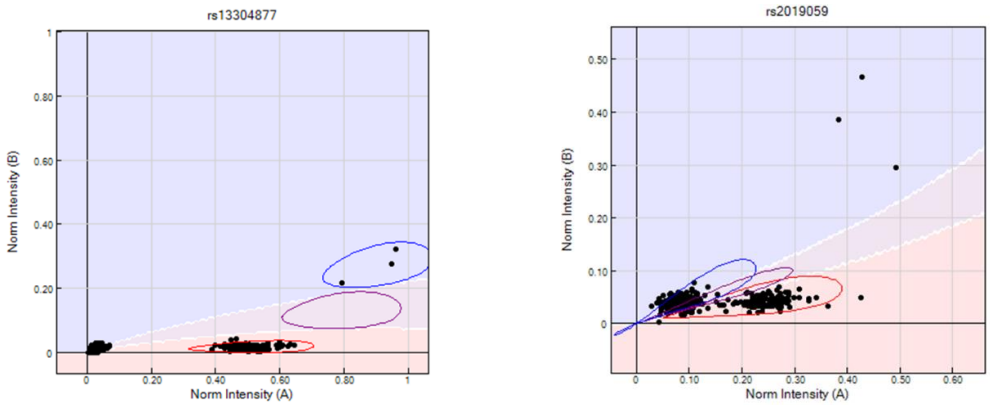


Figure 2.2. Examples of clustering of SNPs genotyped incorrectly. In this case the genotyping did not work well, the algorithm failed to identify the 3 different clusters.

2.2.1 Quality Control of the genotyping data

There are two types of quality control (QC) involved with genotyping data; individual QC and marker QC.

2.2.1.1 Individual Quality Control

1. Checking if there were individuals with discordant sex information. X-chromosome homozygosity is used for this. Male genotypes are exported as homozygous calls, so a male with G allele will be marked as GG genotype. Thus, males will have high

homozygosity. Since PLINK measures excess of homozygosity, females will have values close to 0 and males will have values close to 1.

2. Individuals with missing genotype. We used the missingness proportion >0.05. Samples with missingness proportion more than 0.05 were removed.

3. Heterozygosity rate, in this case we checked the excess of homozygosity, equivalent to the inbreeding coefficient. If samples were outliers, they were removed.

4. Pruning of SNPs in Linkage disequilibrium (LD). In order to check identity by state (IBS), first, the SNPs in LD, must be removed because LD will distort relatedness. The parameters we used here are the following, from a window of size 50 SNPs, checking multiple correlation R^2 of each marker with remaining 49 SNPs. Measuring variance inflation factor VIF = 1/ (1- R^2). VIF = 2 corresponds to R^2 = 0.5. SNPs showing R^2>0.5 with neighbouring SNPs are removed. Then this window moves 5 SNPs to the right and this procedure is repeated.

5. Unexpected relatedness among individuals. A standard population-based association study is based on a sample of unrelated individuals [252], hence this needs to be checked. This was performed using IBS measuring excess of relatedness. For each locus, we looked at similarity between two peoples' genotypes, if GG & GG, score 2, if GG & GA, score 1, if GG & AA, score 0. This of course depends on the minor allele frequency (MAF), if MAF = 0 so everyone is GG, pairwise scores will always be 2. Then, subtract average score from each, and divide by the standard deviation (SD) of the score. Later, add this over all loci, to get genome wide scores for each pair of individuals. People with IBD values >0.1 were removed.

6. Population structure. In this scenario, a signal of association will appear, not because of a real association with the phenotype, but because of allele frequency differences between the different 'hidden' populations within the analysed sample [252]. One of the most used methods to assess for population stratification is Principal Component Analysis [253], this is a dimensionality-reduction method used to lower the number of variables measured on a data set, in this case the genotyping data, to a smaller one that still contains most of the information of the original data set. Principal components are the result of linear combinations of the original genotypes, by observing the location of these individuals along the major axes of variation on the plots, should show just one cluster. If any additional cluster or spread samples appear far from the main

cluster, they must be removed. Mathematically, from each value of each variable the mean is subtracted and then divided by the SD, in order to have zero mean and a unit variance. Subsequently, *Z* denotes a matrix with n rows corresponding to the number of individuals and *l* corresponding to the number of SNPs, with i^{th}, l^{th} elements, for the i^{th} individual and l^{th} SNP, i.e.

$$Z_{il} = \frac{g_i^l - \hat{p}_l}{\sqrt{\hat{p}_l\,(1 - \hat{p})}}$$

where $\hat{p}_l$ is the population allele frequency at SNP *l*. Later, the kinship matrix $\widehat{K}$obtained by,

$$\widehat{K} = \frac{1}{l} ZZ^T$$

It is decomposed by eigenvalues and eigenvectors to identify the principal components, where eigenvectors correspond to the directions of the axes where there is most variance (most information) and they are called PCs. Eigenvalues are the coefficients attached to eigenvectors, corresponding to the amount of variance carried in each PC. Then, PCs will be used as covariables on the GWAS, to account for the underlaying population structure.

2.2.1.2 Markers Quality Control

1. Removing SNPs with high missing data, i.e. low genotyping rate. Missingness proportion >0.05.

2. Removing rare SNPs. These markers are not very useful due to low variation in the data. They are more likely to contain genotyping errors. I used the Minor allele frequency threshold <0.01. This threshold is used in CANDELA data because it is a large dataset.

2.2.2 Quality Control of the phenotyping data

Similarly, to genotyping quality control. Phenotyping data needs to be checked, depending on the studied trait, different filters were applied.

1. Normal distribution of the data set.

2. Rareness. Very low frequency of a category within a trait. I used the threshold <10%. If that was the case, the trait was removed.

3. Constant. Only one category was present in a phenotype. If that was the case, the trait was removed.

4. Missingness proportion of data in an individual. Threshold >15% of data lost. If that was the case, the trait was removed.

5. Missingness proportion of data in a specific phenotype. Threshold >15% of data lost. If that was the case, the trait was removed.

2.2.3 Association testing

Once the QC is completed the genome-wide association test can be performed. The GWAS seeks for associations between genetic markers throughout the genome and different phenotypes (diseases and/or complex traits). At this stage of the analysis, three different types of files are required, the genotype files, the phenotype file and the covariate file. They have been through in-depth QC and now the analysis can be completed. The association test was performed using multiple linear regression with an additive genetic model incorporating age, sex, BMI and genetic PCs (the number of PCs utilized depends on how many of them explain most of the variation of the genotyping data) as covariates.

The objective of Multiple Linear Regression is to model the relationship between explanatory variables (independent variables) and a response variable (dependent variable) by adjusting a linear equation to the observed data.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + .. + .. + \beta_p x_{ip} + \epsilon ,$$

where for $i = n$ observations:

y_i = dependant variable, the phenotype

x_i = explanatory variable, the genotype and other variables, such as age, BMI, PCs, etc.

β_0 = y-intercept (constant term)

β_p = slope coefficients for each explanatory variable, the effect of each variable.

ϵ = the error term of the model, also known as residual.

2.3 Overview of geometric morphometry: the study of form and its application in facial morphology

In this section I will describe the geometric morphometry method used to obtain the quantitative facial phenotypes for the genome-wide association analysis. This method comprises a quantitative method based on landmarks located on 2D photographs of the volunteers. Subsequently, coordinates are generated and processed to be used as phenotypes. In order to help on the understanding of the method I will first describe the history of the method and later the theory in which this method is based on.

Since the dawn of biology, anatomical and morphological description of living organisms have been carried out, with the main objective of differentiating inter- and intra- species[254]. In the beginning, the organisms were compared with previously identified forms, thus the description was qualitative [255].

During the first part of the twentieth century, there was a transition from qualitative to quantitative analysis [256], which resulted in a concept known as morphometrics; fusion of two terms of the classical Greek words morphé: 'form', and metron: 'that by which anything is measured' [257]. Thus, morphometrics is the quantitative study of the variation of biological forms [258]. Traditional morphometrics was based on linear variables, such as measurements, distances, ratios, and angles, and this data was analysed by applying multivariate statistical methods. The results showed the variability of size and form through a set of coefficients and graphs, but they were difficult to interpret, which is one of the main reasons why Geometric Morphometrics emerged (GM) [255,259].

This new branch of morphometrics analyses the shape of living organisms or their structures, regarding the geometrical space using multivariate statistical methods [260]. One important element in GM is the *form*: it is the geometrical feature from an object and it does not consider scale, rotational and translational movement [258,261] (Figure 2.3) . It has been impossible to assess the form separately from the size, due to the ontogenetics of living organisms [255]. *Biological homology* refers to the corresponding

structures or parts between individuals [262,263] and its *localization* (spatial localization in two or three dimensions of these parts or structures [264]) as the main sources of GM

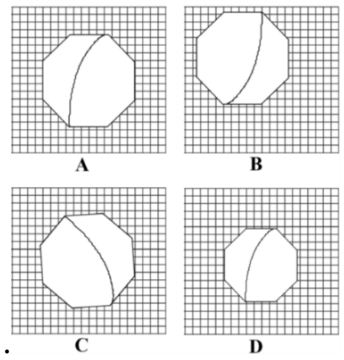


Figure 2.3.The form of an object (A) is not affected by the translation (B), the orientation of the plane, (C) rotation or (D) its scale or size. This figure was published in Bookstein, 1986 [262].

To localize these homologous structures, two variables are used: *outlines* and *landmarks*, anatomical points located according to some anatomical references, which do not alter its topological position, regarding other landmarks. They can be easily situated in the same position between one individual to another [255]. There are 3 types of landmarks: type I are discrete appositions of tissue (i.e. insertion between three sutures), type II are minimum or maximum curvature and type III are distant points to the place of interest (Figure 2.4). Occasionally, the structures are flat or smooth, hence landmarks are difficult to place, to solve this problem, equally distributed points are placed across the contour; these points are known as semi-landmarks [265].

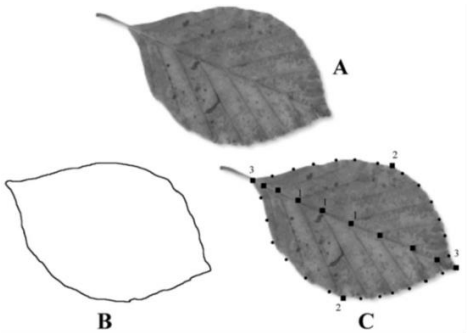


Figure 2.4. Object captured in two dimensions (2D) , (A) can be analysed using Outlines, (B) adding landmarks and (C) in the same structure utilizing the three types of landmarks (squares): type I (1), type II (2) and type III (3). Also, using semi-landmarks (circles). This figure was published in Bookstein, 1997 [265].

Geometric Morphometrics methods capture Cartesian coordinates of landmarks, semi-landmarks and outlines. These landmarks provide information about shape, size, orientation and position of the object. In order to remove the noise caused by size and the rest of the parameters and obtain just the shape, a Generalized Procrustes Analysis is performed [266]. This analysis consists of 3 steps: (1) the configuration of the landmarks is scaled to the same centroid size (the square root of the sum of squared distances of a set of landmarks from their centroid, it is a measure of size that quantifies the extent of the landmark around its centre of gravity), (2) effect of position is removed and (3) the configuration of the landmarks is rotated in order to reduce the deviation among them (Figure 2.5) [267]. The new superimposed landmark configuration coordinates only contain information about form, they are called Procrustes shape coordinates [268].

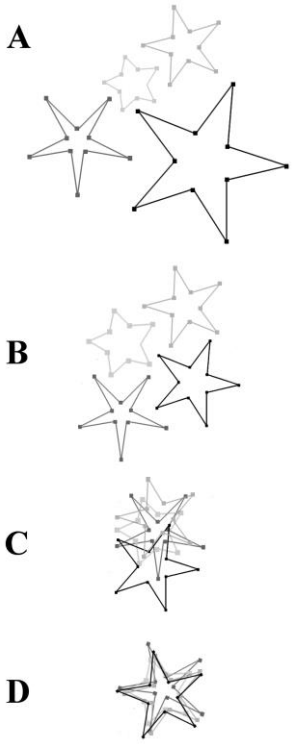


Figure 2.5. Steps of Procrustes superimposition illustrated with star shapes. (A) Raw landmark-configuration, (B) they were translated, so all of them have the same centroid, (C) after that they are scaled to the same centroid size and (D) constantly rotated until the summed squared distances between the landmarks and the sample mean position is a minimum [268,269]. This figure was published in Bookstein, 1997 [265].

Once the Procrustes shape coordinates are obtained, these can be used to perform multivariate analysis, such as regression and PCAs which are used to reduce the

dimensionality of the data and some of these PCs may be utilized as shape variables [270].

The methods described above has been commonly used with facial traits. A large number of studies have investigated how facial features leads to sexual dimorphism, social inference, attractiveness, intelligence, dominance, psychological prediction, etc. [271-274]. Over the past 5 years, scientific interest has been focused in better understanding the molecular genetic basis behind the variation of the face appearance and GM has been widely used for this purpose. In the beginning the data was captured from 2-dimensional (2D) images and landmarks were manually added to the photographs [7,8]. Currently, 2D is still used, but advances in image technology has allowed to use 3D scanning, then these images are imported to softwares that refine them, fill the missing parts of the scan and they can automatically recognize some facial landmarks. Thus, facilitating and improving the acquisition and accuracy of the data [247].

Although genes have been identified for various abnormal phenotypes and genome-wide studies in European populations have been conducted primarily using GM methods, the genetic basis of normal variation in human facial traits is still poorly understood. Therefore, this study intends to contribute to a better understanding of this issue in a sample of Latin American volunteers.

Chapter 3: Genome-Wide Association Study of facial morphology in the CANDELA cohort

3.1 Introduction

Over the years, several genome-wide scans have been performed to find genetic variants involved in the shape of the face. These studies have been successful in their aim to find genetic markers associated with facial traits, but still the genetics of normal variation of the Human face is poorly understood. In this chapter I perform a primary GWAS of facial traits on ~ 6,000 Latin-Americans belonging to CANDELA Cohort for facial traits using categorical phenotyping. Then, I also present follow-up analysis, a replication GWAS using quantitative phenotyping and experimental work to endorse the discovery GWAS findings. The work in this chapter places in evidence the necessity and the advantage of conducting GWAS in human populations under-represented in previous genome-wide scans, and how using quantitative phenotypes increase the power of the analysis.

3.1.1 Previous genome-wide association studies of facial morphology

At present, eleven GWAS have been carried-out in different populations, eight in Europeans [7,8, 3,275-278], and the remaining in Latin-Americans [279], Chinese (Uyghurs) [280], Koreans[281] and Africans [282] (Table 3.1).

Facial feature	Facial phenotype	Population	Sample N	Chr position	SNP	Genes	p value	Author
Ala aperture	Area around the nose wings and size of nostrils	Europeans	2329	4q31.3 (154828366)	rs9995821	***DCHS2***	3.67E-18	Claes et al., 2018
Chin prominence	Centralized prominence of chin extending to lower lip associated to mandibular recession in line with the commissures.	Europeans	2329	1q31.3 (197329041)	rs2821116	*ASPM*	2.63E-19	Claes et al., 2018
Chin prominence	Prominence of chin button associated with lateral recession.	Europeans	2329	7q21.3 (96124975)	rs10238953	*DLX6, DYNC1L1*	1.06E-22	Claes et al., 2018
Forehead	Centralized prominence of forehead with vertical depression above the orbits	Europeans	2329	1p12 (119762175)	rs72691108	*TBX15*	5.81E-14	Claes et al., 2018
Forehead	Recessive central portion of forehead with prominence laterally.	Europeans	2329	6q23.2 (133615646)	rs5880172	*RPS12, EYA4*	6.1E-12	Claes et al., 2018
Lip prominence	Prominent lips, lateral retrusive to upper and lower lips; slight narrowing of nasolabial crease to nasal sidewalls with small prominence superiorly.	Europeans	2329	2q31.1 (177111819)	**rs970797**	*HOXD cluster*	6.17E-11	Claes et al., 2018
Mental fold	Prominence of mental fold, with subtle retrusive effects on labio-mandibular crease.	Europeans	2329	2p21 (42181679)	rs6740960	*PKDCC*	3.03E-14	Claes et al., 2018

Continue…

Facial Feature	Facial phenotype	Population	Sample N	Chr position	SNP	Genes	p value	Author
Nose bridge	Small depression superior to tip of the nose, associated with small depressed areas in the middle of the nasal side walls. Increased prominence bridge of nose.	Europeans	2329	1p32.1 (60997570)	rs4916068	intergenic	4.81E-10	Claes et al., 2018
Nose bridge	Associated with small transverse ridges across the bridge of the nose prominent midway and less prominent across the alae.	Europeans	2329	3q27.1 (184333169)	rs58022575	*EPHB3, DVL3*	7.99E-10	Claes et al., 2018
Nose prominence	Recessive tip of nose with increased width along the sidewall of the nose.	Europeans	2329	17q24.3 (69139583)	rs11655006	*BC039327 /CASC17*	5.24E-21	Claes et al., 2018
Nose prominence	Prominent superior to tip of nose associated with a localized area of recession just above the alae.	Europeans	2329	19q13.11 (34290995)	rs287104	*KCTD15*	1.26E-10	Claes et al., 2018
Nose width	Prominent nose with small nasal wings	Europeans	2329	17q24.3 (70036479)	rs5821892	***SOX9***	2.3E-09	Claes et al., 2018
Nose position	Prominence at nasion and tip/alae width if nose with reduced width along the sidewalls of the nose.	Europeans	2329	2q36.1 (223039052)	rs10176525	***PAX3***	3.76E-14	Claes et al., 2018
Nose width	Deeper nasion position, wider nasal side walls and deeper subnasale.	Europeans	2329	6p21.1 (44681840)	rs227833	***SUPT3H***	3.38E-14	Claes et al., 2018
Philtrum	Depression of philtrum with prominent philtral pillars; reduced ale and prominent subnasale.	Europeans	2329	3q21.3 (128106267)	rs2977562	*RAB7A, ACAD9*	1.39E-17	Claes et al., 2018

Continue…

Facial Feature	Facial phenotype	Population	Sample N	Chr position	SNP	Genes	p value	Author
Nose Size	Nose Size and Nose area	Europeans	2329	1q41 (219642187)	rs5781117	*LYPLAL1*	1.90E-10	Claes et al., 2018
Eye width and depth	Eyes height width and depth	Europeans	1832		rs7560738	*TMEM163*		Crouch et al., 2018
Upper facial profile height	Skeletal pattern/mandibular profile	Europeans	1832		rs11642644	*MBTPS1*	2.7 × 10−3	Crouch et al., 2018
Upper facial profile prominence	Skeletal pattern/mandibular profile	Europeans	1832		rs2045145	*PCDH15*	2.85 × 10−3	Crouch et al., 2018
Eye shape	Curvature of eyelid	Koreans	7569	2:176,820,065	**rs970797**	*HOXD1-MTX2*	7.40E-09	Cha et al., 2018
Eye shape	Eye tail length	Koreans	7569	6:169,699,889	rs3736712	*WDR27*	5.89E-08	Cha et al., 2018
Face height/depth	Right gonion to exR and enR	Koreans	7569	2:19,595,772	rs7567283	*OSR1-WDR35*	1.72E-07	Cha et al., 2018
Nose prominence	Profile nasal angle and nose prominence	Koreans	7569	17:67,537,404	rs2193054	***SOX9***	1.43E-11 5.34E-09	Cha et al., 2018
Nose width	Alae width	Koreans	7569	20:37,426,155	rs2206437	*DHX35*	4.75E-07	Cha et al., 2018
Eye width	Dsitance between endocanthion and exocanthion	Uyghurs China	865	11:122391442	rs1868752	*UBASH3B*	1.00E-10	Qiao et al., 2018
Nose Size	Distance between Nasion point-Pronasale-Subnasale	Uyghurs China	865	5:177922198	rs118078182	*COL23A1*	8.00E-09	Qiao et al., 2018
Mouth	Mouth principal component with strongest segregation between EUR and HAN-TZ men	Uyghurs China	865	4:31120752	rs60159418	*PCDH7*	9.00E-11	Qiao et al., 2018
Cheek	Cheek partial least square with strongest segregation between EUR and HAN-TZ women	Uyghurs China	865	2:52032773	rs17868256		7.00E-09	Qiao et al., 2018

Continue…

Facial Feature	Facial phenotype	Population	Sample N	Chr position	SNP	Genes	p value	Author
Nose	Nose partial least square with strongest segregation between EUR and HAN-TZ women	Uyghurs China	865	20:7067738	rs3920540	*BMP2*	3.00E-08	Qiao et al., 2018
Side-Face	Side-faces partial least square with strongest segregation between EUR and HAN -TZ	Uyghurs China	865	3:82196528	rs61672954		2.00E-08	Qiao et al., 2018
Face height	Mid-face height	Europeans	2187	6q26	rs9456748	*PARK2*	5 × 10−8	Lee et al., 2017
Lip (upper)	Central upper lip height	Europeans	2187	9p22	rs72713618	*FREM1*	2 × 10−8	Lee et al., 2017
Endocanthion	Intercanthal width	Europeans	2187	Xq13	**rs11093404**	*PABP1 - C1L2A, HADC8*	1.07E-09	Lee et al., 2017 **Shaffer et al., 2016**
Nose Size	Nose Size and	Europeans	67000	3p13 (219642187)	3:71227306:T_ TGA	*FOXP1*	8.2×10−26	Pickrell et al.,2016
Chin and shape	Chin Dimple	Europeans	58000	7q21.3	rs11768577	*C7orf76*	5.3×10−56	Pickrell et al.,2016
Nasion prominence	Prominence and vertical position of nasion	Latinamerican	6275	2q35	**rs7559271**	***PAX3***	4 × 10−7	Adhikari et al., 2016 **Paternoster et al., 2012**
Nose prominence	Columella inclination, nose protrusion, nose tip angle	Latinamerican	6275	4q31 (155235392)	rs12644248	***DCHS2***	4 × 10−3	Adhikari et al., 2016
Nose width	Nose wing breadth	Latinamerican	6275	20p11 (22041577)	rs927833	***PAX1***	4 × 10−3	Adhikari et al., 2016
Nose width	Nose bridge breadth	Latinamerican	6275	6p21 (45329656)	rs1852985	***SUPT3H/*** *RUNX2*	5 × 10−3	Adhikari et al., 2016
Nose width	Nose wing breadth	Latinamerican	6275	7p13 (42131390)	rs17640804	*GL13*	6 × 10−3	Adhikari et al., 2016
Eye width	Intercanthal width	Europeans	3118	Xq13.2 (72289467)	**rs11093404**	*PABP1 - C1L2A, HADC8*	**4.16 × 10−8**	Shaffer et al., 2016
Face width	Cranial base width	Europeans	3118	14q21.1 (38038468)	rs17106852	*PAX9, SLC25A2, MIPOL1, FOXA1*	**1.01 × 10−8**	Shaffer et al., 2016
Face width	Cranial base width	Europeans	3118	20q12 (38904203)	rs6129564	*MAFB*	**1.65 × 10−9**	Shaffer et al., 2016
Facial depth	Left tragus to nasion	Europeans	3118	11q22.1 (101394765)	rs12786942	*TRPC6,*	**4.59 × 10−8**	Shaffer et al., 2016

Continue…

Facial Feature	Facial phenotype	Population	Sample N	Chr position	SNP	Genes	p value	Author
Nose prominence	Left ala to tip of nose	Europeans	3118	14q11.2 (21365801)	rs21365801	*ZNF219, CHD8*	**3.36 × 10−8**	Shaffer et al., 2016
Nose prominence and width	Nose width, nose height and left and right alae to prn.	Europeans	3118	1p36.32	**rs4648379 Liu et al., 2012**	***PRDM16***	**1.70 × 10−5**	Shaffer et al., 2016
Nose width	Nasal width	Europeans	3118	20p11.22 (21632545)	rs2424399	***PAX1***	**2.62 × 10−8**	Shaffer et al., 2016
Eye width	Intercanthal width	Europeans	3118	1p13.3 (110218761)	rs619686	*GSTM2, GNAI3, ALX3*	**4.7 × 10−8**	Shaffer et al., 2016
Nasion position and zygomatic arch	Prominence and vertical position of nasion	Europeans	2185	2q35	**rs7559271**	***PAX3***	2.2E-10	Paternoster et al., 2012
	Endocanthion prominence	Europeans	2185	12 (81946438)	rs108662567	*TMTC2*	4.40E-08	Paternoster et al., 2012
prominence	Distance pronasale-left alare	Europeans	2185	3 (55039780)	rs1982862	*CACNA2D3*	1.80E-08	Paternoster et al., 2012
	Distance pronasale-left alare to zygomatic arch and eyeballs	Europeans	2185	5 (61046695)	rs11738462	*C5orf64*	1.80E-08	Paternoster et al., 2012
Nose width	Nose width and nose height	Europeans	5388	1p36.23-p33 (3251376)	**rs4648379**	***PRDM16***	1.13E-08	Liu et al., 2012

Table 3.1. Genomic regions previously associated with facial traits. Genes and SNPs that have been replicated are presented in bold. P-value corresponds to the discovery sample, unless the authors only present meta-analysis results.

These studies have been successful in finding new associations between genome wide significant SNPs and facial traits, but the overlap across traits is very small. The association between nasion position and rs7559271 was found by Paternoster et al. (2012) in Europeans [8]. Our group replicated the exact same SNP in Latin-Americans [279]. Other groups, Liu et al. (2012) and Claes et al. (2018) [7], [275], have also found *PAX3* gene implicated in nasion position in Europeans, but not the same SNP (rs974448 and rs10176525, respectively). Shaffer et al. (2016), discovered an association between the SNP rs11093404, located in Xq13.2 and the distance between endocanthions [278]. Subsequently, another GWAS replicated the association between the same SNP and trait [276]. Usually, the associations found are replicated between the same SNP and the same or similar trait, depending on the phenotyping methods used by the different studies. The SNP rs970797 is an exception because it was discovered in two genome-wide studies, but it was associated with two different features, lip prominence and

curvature of the eyelid [275, 281]. Apart from *PAX3* gene, other genes have been replicated, but not the same SNP, such as *DCHS2* associated with nose traits (columella inclination and area of nose wings), in Latin-Americans and Europeans [275, 279]. Claes et al. (2018) discovered the association between *SOX9* (rs5821892), and the protrusion of the nose [275], Cha et al. (2018) replicated the association between the gene (rs2193054) and the same trait [281]. We detected an association between *SUPT3H/RUNX2* (rs1852985) overlapping region and nose bridge breadth in Latin Americans [279]. Claes et al. (2018) found an association between *SUPT3H* (rs227833) gene and features related to the width of the nose [275]. Finally, two genome-wide association studies discovered an association between *PAX1* gene and nasal width, in Latin Americans and later replicated in Europeans [279, 278].

Not only have GWAS attempted to find new genes implicated in human facial shape but also whole-exome sequencing studies, such as Wu et al. (2019), where they identified 4 genetic loci associated with facial features. They obtained 3D CT scanning from 50 volunteers, and 48 landmarks were used to run the linear mixed model and calculate the association between these traits and the genotypes. Four SNPs in three genes (*RGPD3, IGSF3* and *USP40*) were associated with skull shape, and three SNPs (rs6215530, rs647711 and rs10868138) were associated with facial features [283].

The SNP rs3827760 has been associated with several physical features, such as shovel shape incisors [284], hair shape [211, 285] and ear morphology [286]. The chromosome position of this marker is 2q12, corresponding to the *EDAR* gene. Peng et al. (2016) tested if this variant EDARV370A was associated with facial traits and they discovered that chin protrusion was associated with this SNP in a Uyghur population [287]. Our group replicated the association with a GWAS in Latin Americans [95].

One of the last studies attempting to replicate previous genetic variants associated to facial traits is from Li et al. (2019), recently published. They applied 17 facial landmarks to 3dMD facial images from 612 volunteers from a European-Asian admixed population. The genetic association tests between 125 previously reported SNPs and the facial phenotypes, showed that 8 SNPs were significantly associated with different facial features (Table 3.2) [288].

Facial feature	Population	Sample N	Chr position	SNP	Gene	p-value	Author	GWAS
Distance between the right endocanthion and the inferious part of the right ear lobe	Eurasians	612	6p21.1 (44681840)	rs227833	*SUPT3H*	9.89E-04	Li et al., 2019	Claes et al. 2018
Distance from right alare to left inferior part of ear lobe	Eurasians	612	1q41 (219642187)	rs5781117	*LYPLAL1*	1.43E-04	Li et al., 2019	Claes et al. 2018
Subnasale-Right side chelion	Eurasian	612	1p36.32	rs4648379	*PRDM16*	8.55E-04	Li et al., 2019	Liu et al., 2012 Shaffer et al., 2016
Distance between left exocanthion and the inferious part of the left ear lobe	Eurasian	612	2q12	rs3827760	*EDAR*	2.39E-05	Li et al., 2019	Adhikari et al., 2016
Distance between left endocanthion and right alare	Eurasian	612	2q35	rs7559271	*PAX3*	7.88E-04	Li et al., 2019	Paternoster et al. 2012 Adhikari et al., 2016
Distance between the inferious part of the right and left ear lobe	Eurasian	612	3 (55039780)	rs1982862	*CACNA2D3*	5.29E-04	Li et al., 2019	Paternoster et al. 2012

Table 3.2. Associations replicated from previous facial traits GWAS. This table was modified from Li et al., (2019) [288].

3.2 Materials and Methods

3.2.1 Study subjects

In total, 6,275 volunteers from 5 countries (Colombia, N=1,402; Brasil, N= 658; Chile, N=1,760; Mexico, N=1,200; and Peru, N=1,255) were involved in the GWAS. Except for Chile, most subjects recruited to participate in the project were students and staff from universities, but in Chile 2/3 of the individuals were professional soldiers. Adult subjects of both sexes aged between 18 and 45 years were invited to participate (Table 3.3 and Figure 1.14). This is part of the CANDELA consortium sample [59].

	Total	Colombia	Brazil	Chile	Mexico	Peru
Sample size	6275	1402	658	1760	1200	1255
Percentage	100	22.3	10.5	28.1	19.1	20
% Female	54.1	56.3	68.9	39.4	60.3	58.5
Age (years)						
Min	18	18	18	18	18	18
Mean	24.2	24	25.8	25.2	24.4	22.2
Max	45	40	45	45	44	44
S.D.	5.7	5.3	6.3	5.8	5.6	5.2
Age, for Males						
Min	18	18	18	18	18	18
Mean	24.9	24.7	25.8	25.3	25.1	23
Max	45	40	45	45	44	44
S.D.	5.7	5.5	6.4	5.5	5.6	5.7
Age, for Females						
Min	18	18	18	18	18	18
Mean	23.8	23.5	25.4	25.2	24	21.6
Max	45	40	44	45	41	42
S.D.	5.7	5	4.2	6.2	4.7	4.7

Table 3.3. Features of the study sample.

Ethics approval was obtained from: Universidad Nacional Autónoma de México (México), Universidad de Antioquia (Colombia), Universidad Peruana Cayetano Heredia (Perú), Universidad de Tarapacá (Chile), Universidade Federal do Rio Grande do Sul (Brasil) and University College London (UK). All participants provided written informed consent.

Volunteers with antecedents of craniofacial dysmorphologies, facial surgery, severe facial trauma and BMI over 33 were excluded from this study.

Blood samples were collected by a certified phlebotomist and DNA extracted following standard laboratory procedures.

Five digital photographs of the face: left side (-90), left angle (-45), frontal (0), right angle (45°), right side (90°) were taken from ~1.5 meters at eye level using a Nikon D90 camera fitted with a Nikkon 50mm fixed focal length lens (Figure 3.1).

Left side (-90°) Left angle (-45°) Frontal (0°) Right angle (45°) Right side (90°)

Figure 3.1. Face angles in the digital photographs taken from the volunteers. Fifteen photo shoots were taken per individual in five different positions, as described above. Therefore, three photos per position were obtained in order to choose the best shoot of each position. Photography from M.F-G.

Other phenotypes including height, weight, BMI, age and sex were also recorded for each participant [59].

After the discovery GWAS, an additional 501 individuals were recruited to serve as a replication sample. These individuals were recruited, and the phenotyping was carried out following the same procedures as for the sample included in the discovery GWAS.

3.2.2 Phenotyping

3.2.2.1 Ordinal traits:

Right side and frontal photographs (Figure 3.1) were used to assess 14 facial features (Table 3.4). The scored traits were chin shape and protrusion, cheekbone and brow-ridge protrusion, forehead profile, upper and lower lip thickness and seven nose features (breadth of nasal root, bridge and wing, columella inclination, nose protrusion, nose profile and nose tip shape) (Figure 3.2). These features were selected based on their reported variation in European populations [289].

Face section	Trait	Categories		
		0	1	2
Upper	Forehead profile	Steep (vertical)	Sloping	
	Brow ridge protrusion	None	Slightly pronounced	Strongly pronounced
Middle	Cheekbone protrusion	None	Slightly pronounced	Strongly pronounced
	Nasal root breadth	Narrow	Broad	
	Nose bridge breadth	Short	Average	Long
	Nose wing breadth	Narrow	Average	Broad
	Nose profile	Convex	Straight	Concave
	Nose protrusion	Low	Slightly	Strong
	Nose tip shape	Pointed	Round	Bulbous
	Columella inclination	Up	Straight	Down
Lower	Upper lip thickness	Thin	Average	Full
	Lower lip thickness	Thin	Average	Full
	Chin shape	Pointed	Round	Square
	Chin protrusion	Receding	Normal	Pronounced

Table 3.4. Ordinal facial traits assessed in the CANDELA sample. The traits were selected based on the part of the face where they are located (upper, middle and lower part of the face) and they were classified in 3 categories (0, 1 and 2).

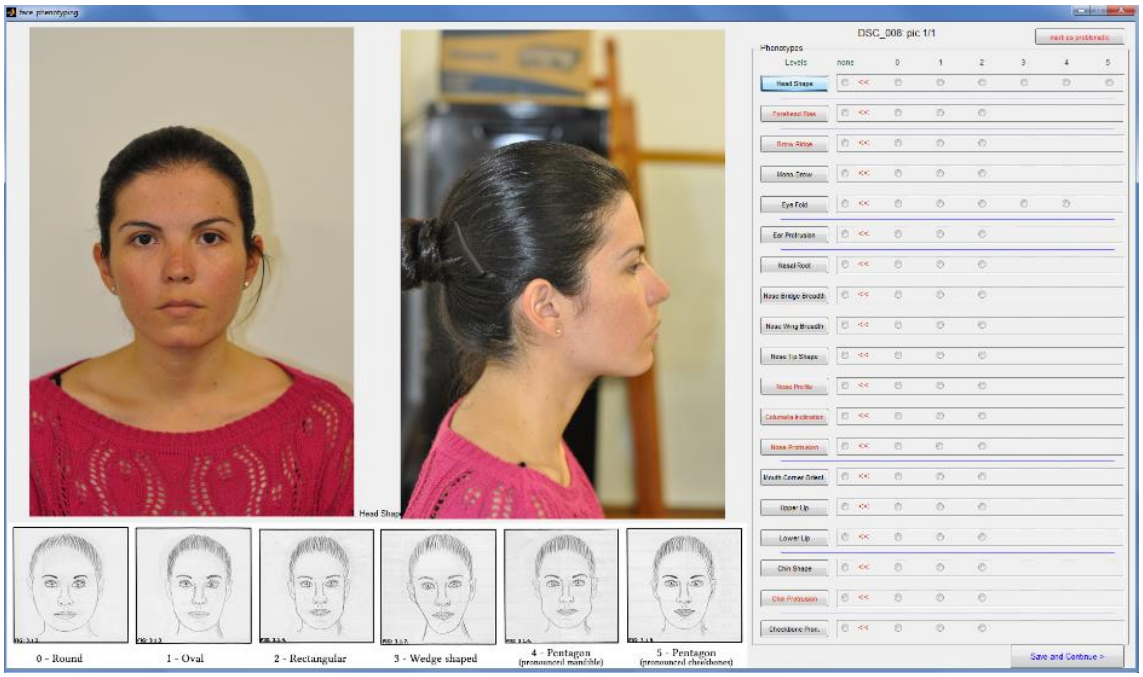


Figure 3.2. Computer interphase used in the scoring of photographs. This tool was developed by K. Adhikari in MATLAB [290]. Frontal and right facial photographs are displayed in the left part of the screen, and the rater should assess based on a range of categories of different facial features, i.e. Nose wing breadth consists of 3 categories (0=narrow, 1=moderate and 2=wide breadth). Photography from M.F-G.

3.2.2.1.1 Rater Reliability for face trait scores

Intraclass correlation coefficients (ICCs) [291] were used to evaluate the rater reliability for facial features scores. This approach uses a two-way mixed effects ANOVA model, with two different scorings (from repeated scoring by one rater or from scores by two raters) as a fixed effect, and variation across subjects as a random effect. Scores from a set of photographs for 450 individuals (>7% of total sample size, combining all 5 constituent countries) were used for calculating ICCs for each facial trait. The photographs were scored twice by two raters (M.F-G. and I.P.A.), independently, two weeks apart. Photographs for all the volunteers were scored by the same rater (M.F.-G.).

3.2.2.2 Quantitative traits

3.2.2.2.1 3D Phenotyping

Quantitative phenotypes were obtained using Procrustes-adjusted 3D facial landmark coordinates available for 2,955 of the individuals included in the ordinal trait GWAS. These coordinates were obtained for 34 anatomical landmarks (Figure 3.3). Briefly, landmarks were placed in five facial photographs described before (Section 3.2.1, Figure 3.1), and raw 3D coordinates obtained using the Photomodeler software. The

raw 3D landmark coordinates were Procrustes-adjusted using the MorphoJ software [292].

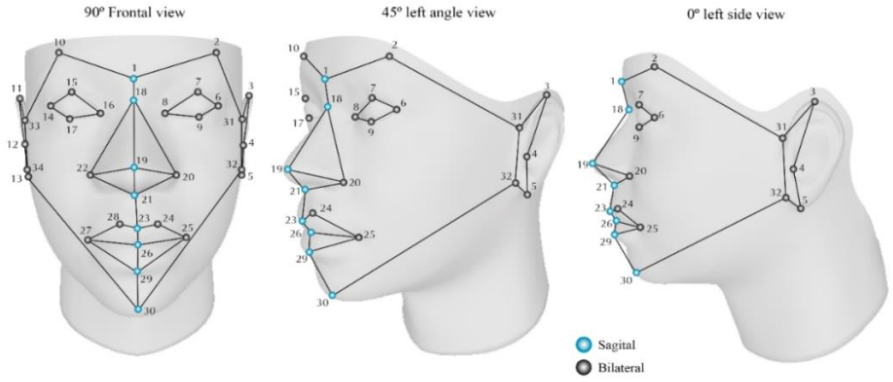


Figure 3.3. Anatomical landmarks used for the analysis of shape variation in the CANDELA volunteers. 34 anatomical landmarks were located in the frontal, fronto-lateral and lateral left and right photographs of each volunteer, using photomodeler software, to obtain the raw 3D coordinates. Figure from Quinto-Sanchez et al. (2015) [274].

Quantitative measurements (distances and angles) were defined corresponding to seven of the ordinal traits initially examined (Table 3.5).

Face section	Trait	Quantitative definition[a,b]
Upper	Forehead profile	–
	Brow ridge protrusion	–
Middle	Nasion position	Distance from point 18 to the mid-point of a line joining points 8 and 16
	Cheekbone protrusion	–
	Nasal root breadth	–
	Nose bridge breadth	–
	Nose wing breadth	Distance between points 20 and 22
	Nose profile	–
	Nose protrusion	Distance of point 19 to a line joining points 18 and 21
	Nose tip shape	Angle between points 18-19-21
	Columella inclination	Angle between points 19-21-23
Lower	Upper lip thickness	Distance between points 23 and 26
	Lower lip thickness	Distance between points 26 and 29
	Chin shape	–
	Chin protrusion	Ratio of distance between points 30-32 to distance between points 29-32

Table 3.5.Quantitative facial traits. The traits were selected based on the part of the face where they are located (upper, middle and lower part of the face). The details of the landmarks are explained in Figure 3.3.

3.2.2.2.2 2D Phenotyping:

Since 3D landmarks allowing quantitative proxies for nose root and bridge breadth were not available, I placed 2D landmarks on the frontal photographs of the same individuals with 3D landmarks (Figure 3.4, Table 3.6): two landmarks were added each for nasal root and for nose bridge width, in addition to the major frontally visible 3D landmarks. Since the 3D coordinates are free of head tilts and rotations (thus allowing more accurate measurements) the 2D coordinates were calibrated with reference to the 3D coordinates using corresponding frontal landmarks (having both 2D and 3D coordinates) (Figure 3.3 and Figure 3.4).

Face section	Trait	Quantitative definition[a]
Middle	Nasal root breadth	Distance between points 3 and 4
	Nose bridge breadth	Distance between points 5 and 6

Table 3.6.Morphology measurements of nose.

[a] The quantitative definitions refer to coordinate numbers as shown in Figure 3.3.

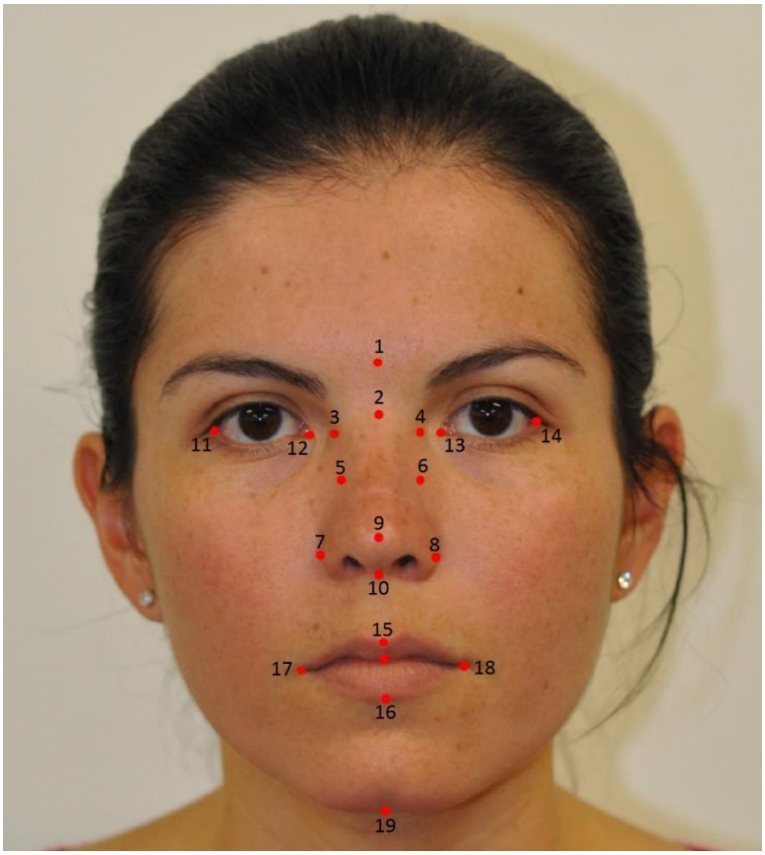


Figure 3.4. 2D Frontal Landmarks examined in the CANDELA sample. 19 landmarks and semi-landmarks were positioned in the frontal take of each volunteer. Photography from M.F-G.

3.2.2.3 Replication sample:

The replication sample was genotyped in the same way and submitted to the same quality controls as for the GWAS sample (Section 3.2.4).

3.2.3 DNA genotyping and quality control

DNA samples from participants were genotyped on the Illumina HumanOmniExpress chip including 730,525 SNPs. PLINK v1.9 [251] was used to exclude SNPs and individuals with >5% missing data, markers with minor-allele frequency <1%, related individuals (Plink IBD estimate >0.1), and those who failed the X-chromosome sex concordance check (sex estimated from X-chromosome heterozygosity not matching recorded sex information). After applying these filters 671,038 SNPs and 5,958 individuals (1,303 from Colombia, 608 from Brasil, 1,651 from Chile, 1,165 from Mexico, 1,231 from Peru) were retained for further analysis. Due to the admixed nature of the study sample (Figure 3.5) there is an inflation in Hardy–Weinberg P values (Figure 3.6). We therefore did not exclude markers based on Hardy–Weinberg deviation, but performed stringent quality controls at software and biological levels and checked the genotyping cluster plots for each index SNP manually (Figures 2.1 and 2.1).

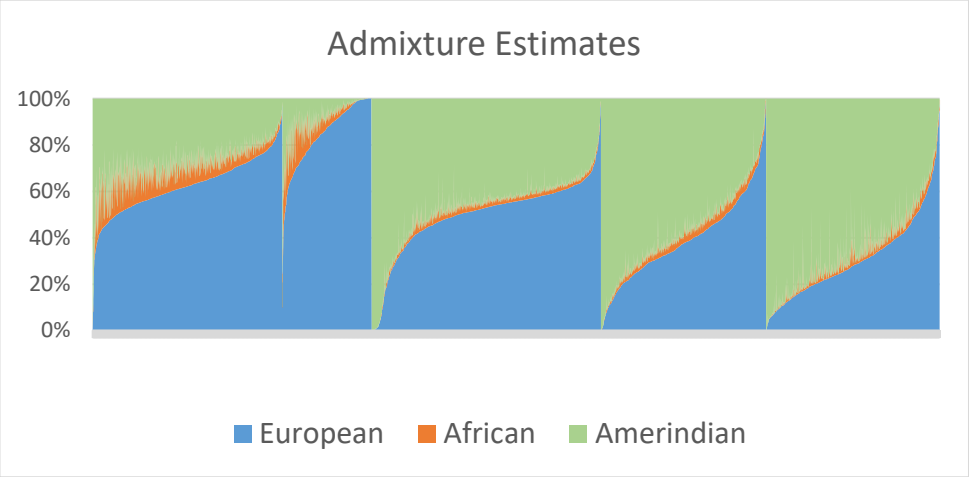


Figure 3.5. Estimated individual African, European and Native American ancestry (%) in the CANDELA individuals included in the GWAS for ordinal face traits (N=5.958). Based on genome-wide SNP data, average autosomal admixture proportions for the full sample were estimated as: 50% European, 45% Native American and 5% African.

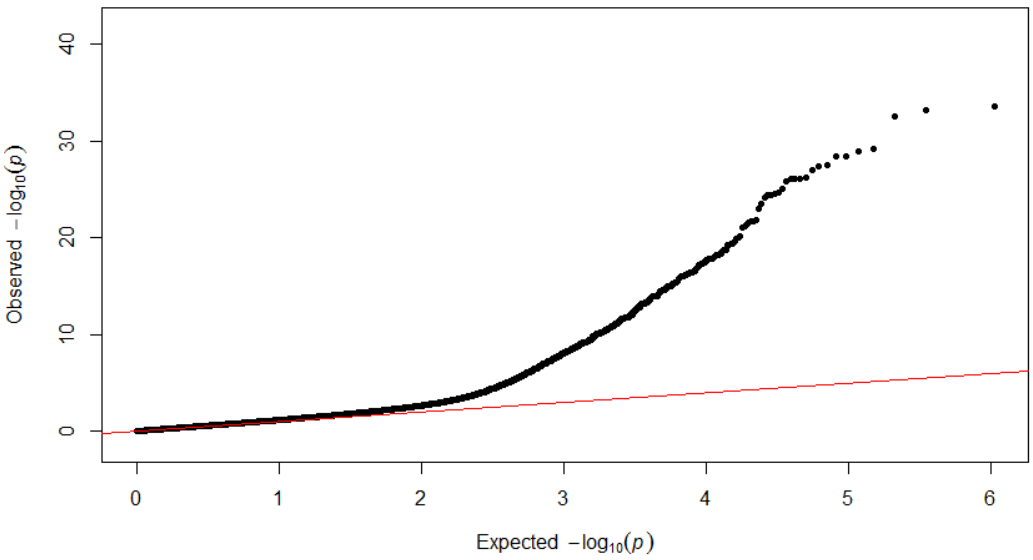


Figure 3.6. Genome-wide testing of deviation from Hardy-Weinberg equilibrium in Latin American samples from 1000 Genomes Phase 3 [293]. To demonstrate that our dataset of post-QC SNPs present deviation from HWE not due to genotyping error but as a result to population stratification, HWE testing was performed for Latin American samples from the 1000 Genomes Phase 3 release, this population is commonly used as a reference panel. 328 Latin Americans individuals from this panel (after removing related individuals) were utilized for the HWE P-value calculation in Plink 1.9. The resulting Q-Q plot is shown above.

3.2.4 SNP Genotype Imputation

The chip genotype data was phased using SHAPEIT2 [294]. IMPUTE2 [295] was then used to impute genotypes at untyped SNPs using variant positions from the 1000 Genomes Phase I data [293]. The 1000 Genomes reference data set includes haplotype information for 1,092 individuals across the world for 36,820,992 variant positions. Positions that are monomorphic in 1000 Genomes Latin American samples (CLM, MXL and PUR) were excluded, leading to 11,025,002 SNPs being imputed in our data set. Of these, 48,695 had imputation quality scores <0.4 and were excluded. Chip genotyped SNPs having a low concordance value (<0.7) or a large gap between info and concordance values (info_type0—concord_type0>0.1), which might be indicators of poor genotyping, were also removed, both from the imputed and chip data set. The IMPUTE2 genotype probabilities at each locus were converted into best-guess genotypes using PLINK [251] (at the default setting of <0.1 uncertainty). SNPs with a proportion of samples with uncalled genotypes >5% and minor-allele frequency <1% were excluded. The final imputed data set contained genotypes for 9,117,642 SNPs.

3.2.5 Statistical genetic analyses

3.2.5.1 Narrow-sense heritability (h2) estimation

Narrow-sense heritability (defined as the additive phenotypic variance explained by a Genetic Relatedness Matrix, GRM, computed from the SNP data) was estimated using GCTA[61] by fitting an additive linear model with a random-effect term whose variance is given by the GRM, with age, sex and BMI as covariates. The GRM was obtained using the LDAK approach [60], which accounts for LD between SNPs.

3.2.5.2 Continental Ancestry Estimation

An LD-pruned set of 93,328 autosomal SNPs was used to estimate European, African and Native American ancestry using supervised runs of ADMIXTURE [296]. Reference parental populations included in the ADMIXTURE analyses consisted of Africans and Europeans from HAPMAP [297] and selected Native Americans, as described in Ruiz-Linares et al. [59]

3.2.5.3 Genome Wide Association Analyses

PLINK 1.9 [251] was used to perform the primary genome-wide association tests for each phenotype using multiple linear regression with an additive genetic model incorporating age, sex, BMI and five genetic PCs as covariates. Association analyses were performed on the imputed data set with two approaches: using the best-guess imputed genotypes in PLINK and using the IMPUTE2 [295] genotype probabilities in SNPTEST v2.5 [298]. Both were consistent with each other and with the results from the chip genotype data. For analysis of the X chromosome an inactivation model was used (male genotypes encoded as 0/2 and female genotypes as 0/1/2). The genetic PCs were obtained using PLINK 1.9 [251] from an LD-pruned dataset of 93,328 SNPs. They were selected by inspecting the proportion of variance explained and checking scree and PC scatter plots (Figure 3.7). Individual outliers were removed, and PCs recalculated after each removal. The top PCs appear to be a good proxy for continental ancestry (Table 3.7).

Correlations of the top 5 genetic PCs with the three continental ancestries (Figure 3.7 and Table 3.7) are presented in the table below. Since European-Amerindian admixture is the main genetic component in our samples, PC1 is highly correlated with these ancestries with opposite signs. PC2 is highly correlated with African ancestry.

Ancestry	PC1	PC2	PC3	PC4	PC5
European	0.951	-0.307	0.001	-0.025	-0.008
Amerindian	-0.998	0.056	0.012	0.005	0.01
African	0.379	0.92	-0.053	0.072	-0.013

Table 3.7. Correlation of 5 genetic PCs and three continental ancestries. Due to CANDELA samples belonging to Latin America. The mix between Amerindian and European is the principal genetic component in this sample. PC1 shows a high correlation with these two ancestries with opposite signs, while PC2 is highly correlated with African ancestry.

Using these PCs, the Q–Q plots (Figure 3.8) for all association tests showed no sign of inflation, the genomic control factor lambda being <1.02 in all cases (Figures A.1-A.13. and Table A.1), thus confirming that we are appropriately accounting for population stratification [253]. Similar analyses were applied for association testing of the index SNPs followed-up in the replication sample. To account for multiple testing, we also applied a global false-discovery rate test using the Benjamini–Hochberg procedure across all traits and SNPs (Table 3.16 and Table 3.17). To account for the correlations between traits, a multivariate GWAS was also performed, testing for association with all facial traits simultaneously using a Wald test conditioned on all covariates (Table 3.19). A meta-analysis was carried out for the index SNPs identified in the primary analyses by testing for association separately in each country sample and combining the results (using the PLINK [251] implementation of the meta-analysis software METAL [299]). Forest plots were produced with MATLAB [290]. Cochran's Q-statistic was computed for each trait to test for effect-size heterogeneity across country samples. The fraction of trait variance explained by the covariates, by each index SNP, and by all index SNPs altogether, were estimated from linear regression models implemented using R^2 (Tables A.2). To evaluate the role of non-syndromic cleft lip and palate (NSCL/P) loci on the facial traits examined we selected index SNPs in the 15 associated regions reported in the literature and performed individual SNP associations, global Kolmogorov–Smirnov tests and Polygenic Risk Score tests using PLINK.

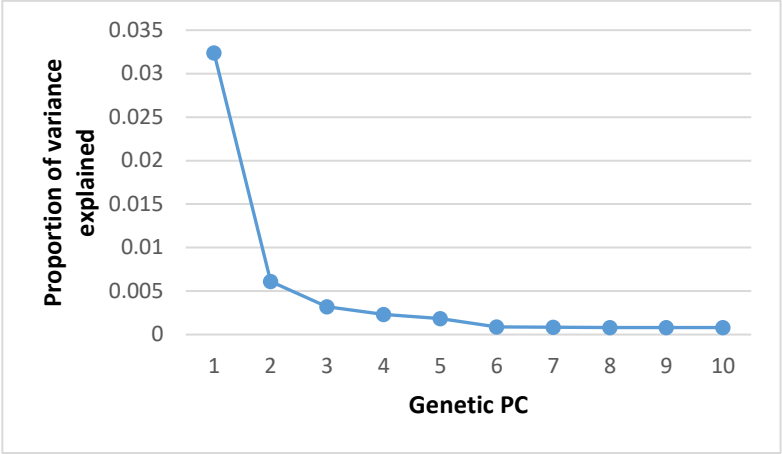


Figure 3.7. Selection of genetic Principal components to be used as covariables in GWAS analysis. The proportion of the variance explained by each PC. PC 1 explains most of the variation in the sample.

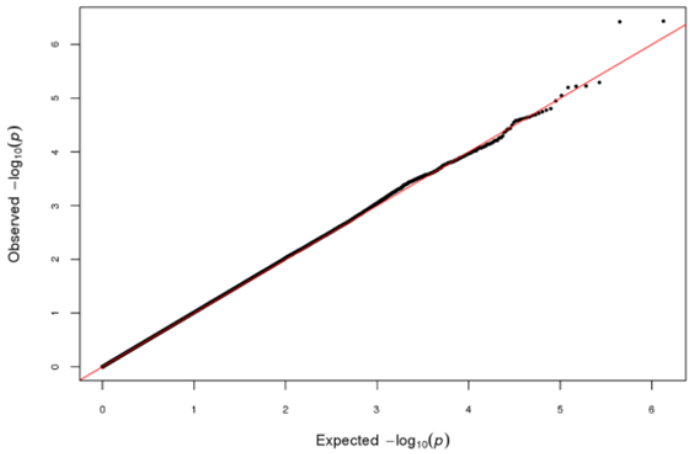


Figure 3.8. GWAS Q-Q Plots of Nose Profile. The remaining Q-Q Plots are shown in Figures A.1-A.13., Table A.1. This plot does not show sign of inflation between the expected and observed P-values. All the traits show similar pattern, the genomic control factor lambda <1.02, demonstrating there is no population stratification.

3.2.6 Mouse analyses

Animal studies were reviewed and approved by The Roslin Institute Animal Welfare and Ethical Review Body (AWERB). The human care and use of mice (*Mus musculus*) in this study was carried out under the authority of the appropriate UK Home Office Project License. The mouse samples and head photographs examined are from the same set described fully in Adhikari et al.[58] Briefly, we included 14 and 15-day-old animals (17 males and 23 female). The mouse genotypes were *Edar*^{dlJ} (a loss of function EDARp.E379K mutation [300]) as either homozygote or heterozygote, wild-

type (+/ +) and the homozygous *Edar*^{Tg951} line (which has ~16 extra copies of *Edar* per haploid genome [301]). Thirteen 2D anatomical landmarks were placed on lateral photographs of the mouse heads, using TPSDig [302] and TPSUtil [303] (Figure 3.9). Generalized procrustes analysis was carried out using the software MorphoJ [292] to check whether the distribution of landmarks was homogeneous. There were no outliers. Mouse mandible length was measured using the landmark coordinates (as detailed in Figure 3.9 and Figure 3.10).

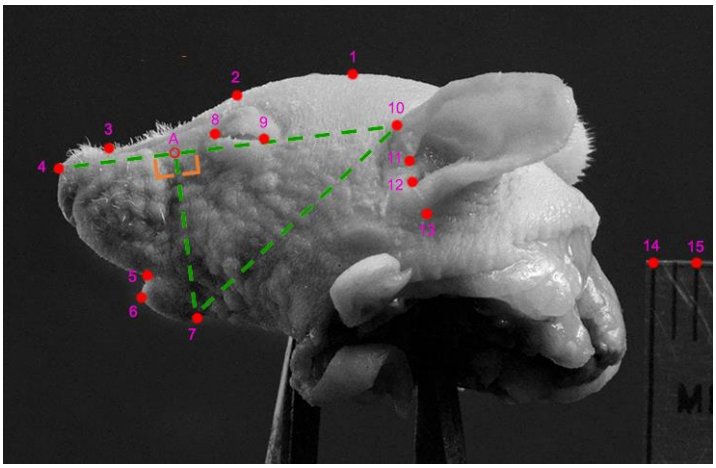


Figure 3.9. Mouse head landmarking protocol to measure mandible length in *Edar* mutant mice. Thirteen landmarks were added in the face of the mouse and two landmarks in a scale to calibrate the measurements.

Each mouse head was mounted horizontally on spikes and the camera was positioned parallel to this plane. Colour photographs were taken from the left side. A scale was included in each photo for calibration. Thirteen landmarks were placed across the mouse head and two on the scale, as indicated by filled red circles. Three landmarks were placed on the chin – points 5 to 7.

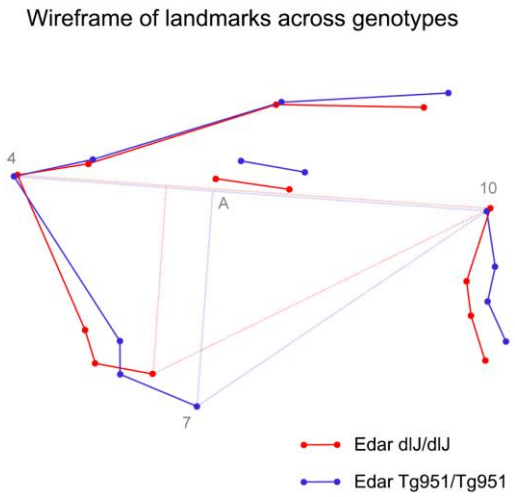


Figure 3.10. Mouse head shape variation and Edar genotype.

Landmark point 10 at the top of the ear was taken as the reference point. Direct distances were computed between this point and each chin landmark. A front-to-back line of reference was constructed by joining reference landmark point 10 to the most frontal landmark, point 4 on the nose. The projection of each chin landmark onto this line was calculated. E.g. in the example shown above, chin landmark 7 is perpendicularly projected onto the line joining 4 and 10, the projection point denoted by hollow red circle A. The projected distance is the distance between this constructed landmark A and reference point 10. These distances as obtained in pixel units are then converted to millimetres by calibrating with the distance on the scale (between points 14 and 15). As facial distances will be variable depending on the overall head size, chin protrusion was converted into a proportion by dividing with head length.

Average wireframes for homozygous *Edar*$^{dlJ/dlJ}$ (red lines) and *Edar* $^{Tg951/Tg951}$ (blue lines) homozygous mice Landmark points 4, 7, 10 and A are shown (Figure 3.10). Measured direct distance between landmarks 7 and 10 is the length of the dotted line between them. Projected distance between landmarks 7 and 10 is the length of the dotted line between points A and 10. This diagram shows that the lengths are smaller in the case of blue i.e. *Edar* $^{Tg951/Tg951}$ mice.

Mandible length (as a proportion of head size, measured directly on the heads) was regressed onto age, sex and *Edar* genotype. In this regression *Edar* genotype was coded as 1–4 based on increasing *Edar* expression: 1- *Edar*$^{dlJ/dlJ}$ homozygotes, 2-*Edar*$^{dlJ/+}$ heterozygotes, 3-wild-type $^{+/+}$ mice and 4-*Edar*$^{Tg951/Tg951}$ homozygotes.

3.3 Results

3.3.1 Study Sample

The cohort used in this study was collected in Latin America and is part of CANDELA [59] (Table 3.3 and Figure 1.14).

3.3.2 Ordinal traits

Fourteen facial traits were assessed on an ordered categorical scale reflecting the distinctiveness of each trait (Figure 3.11, Table 3.8) using facial photographs of 6,275 individuals. Features included of the lower face: chin shape, chin protrusion, and upper/lower lip thickness, the middle face: cheekbone protrusion, breadth of nasal root, bridge and wing, columella inclination, nose protrusion, nose profile, and nose tip shape, and the upper face: brow-ridge protrusion and forehead profile.

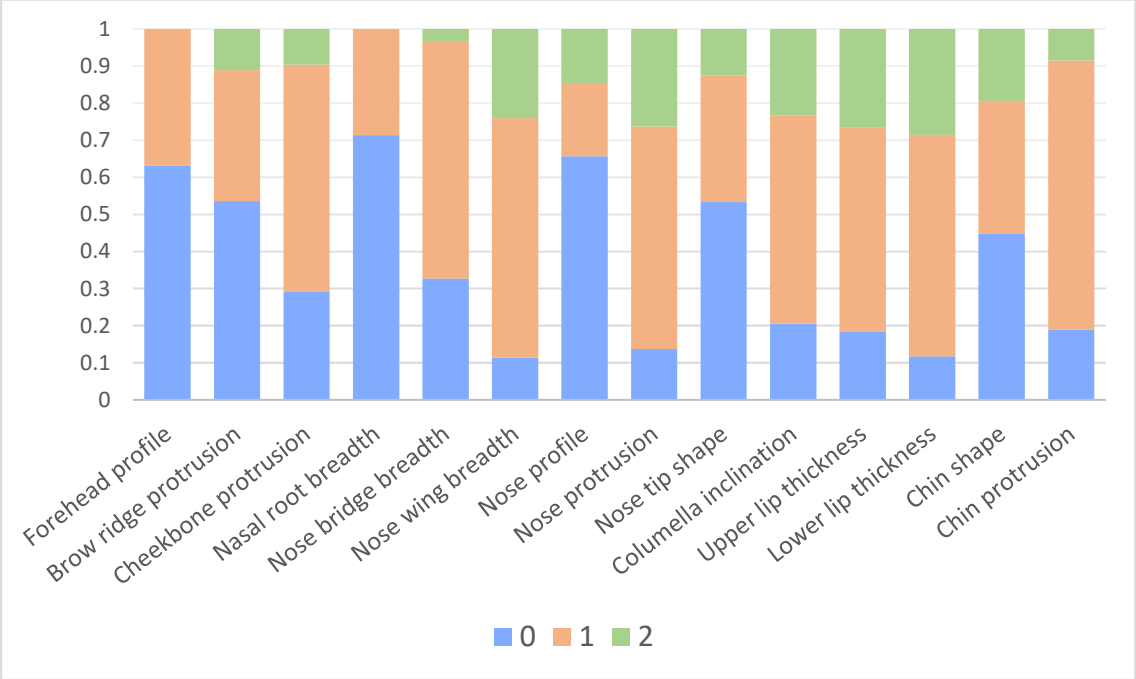


Figure 3.11. Frequency distribution of 14 ordinal face traits in the CANDELA samples. The facial traits selected, showed an extensive variation in the CANDELA samples.

Face section	Trait	Categories		
		0	1	2
Upper	Forehead profile	Steep (vertical)	Sloping	-
	Brow ridge protrusion	None	Slightly pronounced	Strongly pronounced
Middle	Cheekbone protrusion	None	Slightly pronounced	Strongly pronounced
	Nasal root breadth	Narrow	Broad	
	Nose bridge breadth	Short	Average	Long
	Nose wing breadth	Narrow	Average	Broad
	Nose profile	Convex	Straight	Concave
	Nose protrusion	Low	Slightly	Strong
	Nose tip shape	Pointed	Round	Bulbous
	Columella inclination	Up	Straight	Down
Lower	Upper lip thickness	Thin	Average	Full
	Lower lip thickness	Thin	Average	Full
	Chin shape	Pointed	Round	Square
	Chin protrusion	Receding	Normal	Pronounced

Table 3.8. Ordinal Face traits assessed. Each trait was analysed using three categories from 0 to 2.

3.3.2.1 Rater reliability for facial features scores

The rater reliability for face traits scores were assessed by calculating intra-class correlation coefficients (ICC) [291]. They indicate a moderate–to-high intra-rater reliability of the trait scores (Table 3.9), with relatively lower inter-rater reliability for certain traits.

Face section	Trait	Rater 1 (M.F.)	Rater 2 (I.P.)	Inter-rater
Upper	Brow ridge protrusion	0.77	0.78	0.69
	Forehead profile	0.76	0.67	0.48
Middle	Cheekbone protrusion	0.71	0.50	0.44
	Nasal root breadth	0.77	0.46	0.42
	Nose bridge breadth	0.61	0.58	0.48
	Nose wing breadth	0.68	0.72	0.65
	Nose profile	0.66	0.73	0.65
	Nose protrusion	0.75	0.63	0.53
	Nose tip shape	0.80	0.63	0.51
	Columella inclination	0.72	0.81	0.57
Lower	Upper lip thickness	0.80	0.85	0.68
	Lower lip thickness	0.75	0.66	0.69
	Chin shape	0.76	0.66	0.65
	Chin protrusion	0.75	0.59	0.57

Table 3.9. Rater reliability for face traits scores. There were 2 raters who each scored the same 200 people. M.F-G, Macarena Fuentes-Guajardo and I.P., Iván Pulgar.

3.3.3 Continental Ancestry Estimation

Individuals were genotyped on Illumina's Omni Express BeadChip and imputation performed using 1000 Genomes data [293]. After quality control filters, final analyses were carried out on 671,038 genotyped SNPs and 9,117,642 imputed SNPs in 5,958 individuals (Figure 3.5).

3.3.4 Correlations of ordinal face traits

Significant correlations were observed between the ordinal phenotypes (using a Bonferroni-adjusted permutation *P*-value threshold for significance of 6×10^{-4}, Table 3.10). Strongest correlation was observed between upper and lower lip thickness (r = 0.72), followed by forehead profile and brow ridge protrusion (r = 0.57). The three traits related to nose width (root, bridge and wing breadth) show positive correlations among them (r = 0.16 to 0.37) and negative correlations with nose protrusion (r = -0.08 to r = -0.25) (Table 3.10).

Trait	FP	BR	CP	NR	NB	NW	NP	NPR	NT	CI	ULT	LLT	CS	CP
Forehead profile		**<5E-4**	1E-02	**<5E-4**	**<5E-4**	**<5E-4**	2E-02	**<5E-4**	**<5E-4**	**<5E-4**	2E-01	9E-01	**<5E-4**	4E-01
Brow ridge protrusion	**0.57**		**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	2E-03	3E-01	**<5E-4**	7E-02	6E-01	**<5E-4**	4E-01
Cheekbone protrusion	-0.03	**-0.10**		2E-02	**<5E-4**	**<5E-4**	**<5E-4**	1E-02	5E-01	**<5E-4**	6E-02	4E-01	3E-02	**<5E-4**
Nasal root breadth	**-0.10**	**-0.08**	0.03		**<5E-4**	**<5E-4**	8E-03	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	3E-01	7E-03
Nose bridge breadth	**0.12**	**0.18**	**-0.05**	**0.29**		**<5E-4**	3E-02	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	4E-02	1E-01
Nose wing breadth	**0.10**	**0.17**	**-0.05**	**0.16**	**0.37**		2E-02	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	3E-03	7E-03
Nose profile	-0.03	**-0.05**	**0.05**	0.03	-0.03	-0.03		9E-03	2E-01	**<5E-4**	6E-01	5E-01	9E-01	1E-01
Nose protrusion	**0.11**	0.04	0.03	**-0.16**	**-0.08**	**-0.14**	0.03		**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	7E-01	1E-01
Nose tip shape	**-0.05**	0.01	0.01	**0.18**	**0.16**	**0.29**	0.02	**-0.25**		**<5E-4**	**<5E-4**	**<5E-4**	4E-01	1E+00
Columella inclination	**0.08**	**0.11**	**-0.09**	**-0.06**	**0.05**	**0.05**	**-0.20**	**0.10**	**-0.14**		8E-04	**<5E-4**	2E-01	5E-02
Upper lip thickness	0.02	0.02	-0.02	**0.07**	**0.08**	**0.11**	-0.01	**-0.21**	**0.12**	-0.04		**<5E-4**	3E-01	**<5E-4**
Lower lip thickness	0.00	-0.01	-0.01	**0.05**	**0.06**	**0.06**	0.01	**-0.13**	**0.08**	**-0.06**	**0.72**		5E-01	**<5E-4**
Chin shape	**0.06**	**0.08**	0.03	-0.01	0.03	0.04	0.00	0.00	0.01	-0.02	0.01	-0.01		2E-03
Chin protrusion	-0.01	-0.01	**0.05**	0.03	0.02	0.03	0.02	0.02	0.00	-0.02	**-0.13**	**-0.17**	0.04	

Table 3.10. Correlation between face traits. Correlation values are presented in the lower left triangle, with corresponding permutation P values in the upper right triangle. Correlations with significant P values (<0.0006, Bonferroni-adjusted threshold) and their corresponding P values are highlighted in bold.

Several of the facial traits examined also show moderate (and significant) correlations with age, sex, body mass index (BMI) and genetic ancestry (Table 3.11). The strongest correlation with sex was seen for brow ridge protrusion and forehead profile (r = -0.62 and r = -0.47 respectively). Age correlates most strongly with upper and lower lip thickness (r = -0.19 and r = -0.24 respectively), while the strongest correlation for BMI

was seen with brow-ridge protrusion (r = 0.17). Genetic ancestry has strongest correlation with lip thickness (European ancestry being negatively correlated with upper and lower lip thickness, r= -0.25 and r= -0.16, respectively). European ancestry is also significantly correlated with all the nose features examined, particularly with nose protrusion (r= 0.18) and nose wing breadth (r= -0.15).

	Upper		Middle								Lower			
Covariate	FP	BR	CB	NR	NB	NW	NP	NPR	NT	CI	ULT	LLT	CS	CP
Sex	**-0.47**	**-0.62**	**0.24**	**0.16**	**-0.23**	**-0.17**	**0.10**	**-0.05**	**0.05**	**-0.20**	-0.01	0.02	**-0.08**	**0.08**
Age	0.09	**0.06**	0.02	-0.04	-0.01	-0.02	-0.01	**0.08**	**-0.06**	**0.08**	**-0.19**	**-0.24**	0.04	**0.06**
BMI	**0.11**	**0.17**	**-0.16**	**0.07**	**0.10**	**0.10**	-0.02	**-0.15**	**0.11**	0.04	-0.01	**-0.08**	**-0.08**	**0.05**
African anc.	-0.02	-0.01	0.02	**0.06**	**0.06**	**0.06**	0.02	-0.03	0.03	0.00	0.02	**0.06**	0.03	0.03
European anc.	0.02	0.02	0.04	**-0.08**	**-0.05**	**-0.15**	**0.05**	**0.18**	**-0.11**	**-0.06**	**-0.25**	**-0.16**	0.01	**0.12**
American anc.	-0.01	-0.02	-0.04	**0.06**	0.03	**0.13**	**-0.05**	**-0.16**	**0.09**	**0.06**	**0.23**	**0.13**	-0.02	**-0.13**

Table 3.11a. Correlation between ordinal facial traits and covariates. Correlation values are presented in Table 3.11a, with corresponding P values in Table 3.11b. Correlations with significant P values (<0.0005, Bonferroni-adjusted threshold), obtained by permutation, are highlighted in bold. Anc.= Continental ancestry estimated from the genetic data. Sex coded as female=1, male=0.

	Upper		Middle								Lower			
Covariate	BR	FP	CB	NR	NB	NW	NP	NPR	NT	CI	ULT	LLT	CS	CP
Sex	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	5E-01	6E-02	**<5E-4**	**<5E-4**
Age	**<5E-4**	**<5E-4**	1E-01	6E-01	3E-03	2E-01	4E-01	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	7E-04
BMI	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	2E-01	**<5E-4**	**<5E-4**	5E-04	4E-01	**<5E-4**	**<5E-4**	**<5E-4**
African anc.	4E-01	8E-02	1E-01	**<5E-4**	5E-02	**<5E-4**	6E-02	8E-03	2E-02	9E-01	2E-01	9E-06	5E-02	**<5E-4**
European anc.	7E-02	2E-01	6E-03	**<5E-4**	4E-01	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**
American anc.	2E-01	4E-01	2E-03	2E-02	1E-01	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**

Table 3.11b. Corresponding P values correlation between ordinal facial traits and covariates.

3.3.5 Narrow-sense heritability (h2)

Based on a kinship matrix derived from the SNP data [60], the narrow-sense heritability for the facial traits using GCTA [61] was estimated. Moderate (and significant) values for all traits were found, with the highest heritability being estimated for nose protrusion (0.47) and the lowest for columella inclination (0.20) (Table 3.12). Similar (or higher) heritabilities have been estimated for a range of facial traits using family data [28,51,304].

Face section	Trait	Heritability	S.E.	*P* value
Upper	Forehead profile	0.28	0.06	2.6E-14
	Brow ridge protrusion	0.44	0.06	0.0E+00
Middle	Cheekbone protrusion	0.28	0.05	1.8E-15
	Nasal root breadth	0.23	0.05	2.2E-16
	Nose bridge breadth	0.23	0.06	8.5E-13
	Nose wing breadth	0.41	0.05	0.0E+00
	Nose profile	0.22	0.06	4.8E-10
	Nose protrusion	0.47	0.05	0.0E+00
	Nose tip shape	0.27	0.05	0.0E+00
	Columella inclination	0.20	0.06	6.9E-12
Lower	Upper lip thickness	0.46	0.06	0.0E+00
	Lower lip thickness	0.40	0.06	0.0E+00
	Chin shape	0.31	0.06	3.1E-14
	Chin protrusion	0.22	0.05	0.0E+00

Table 3.12.Heritability estimates for the 14 ordinal face traits examined. Low to moderate significant values of heritability are shown for ordinal traits. The lowest value was for Columella inclination (0.20) and the highest heritability was Nose protrusion (0.47) among ordinal facial features.

3.3.6 GWAS for ordinal phenotypes

We performed genome-wide association tests using multivariate linear regression, as implemented in PLINK [251], using an additive genetic model adjusting for: age, sex, BMI and the first five principal components (Figure 3.7) computed from the SNP data. The resulting statistics showed no evidence of residual population stratification for any of the traits (Figure 3.8). Three of the nose traits examined (columella inclination, nose bridge and wing breadth) showed genome-wide significant association (*P* values < 5×10^{-8}) with SNPs in at least one genomic region (Figure 3.12 and Table 3.13). Columella inclination and nose bridge breadth show association with SNPs in a single region (4q31 and 6p21 respectively), while nose wing breadth shows association with SNPs in two genomic regions (7p13 and 20p11).

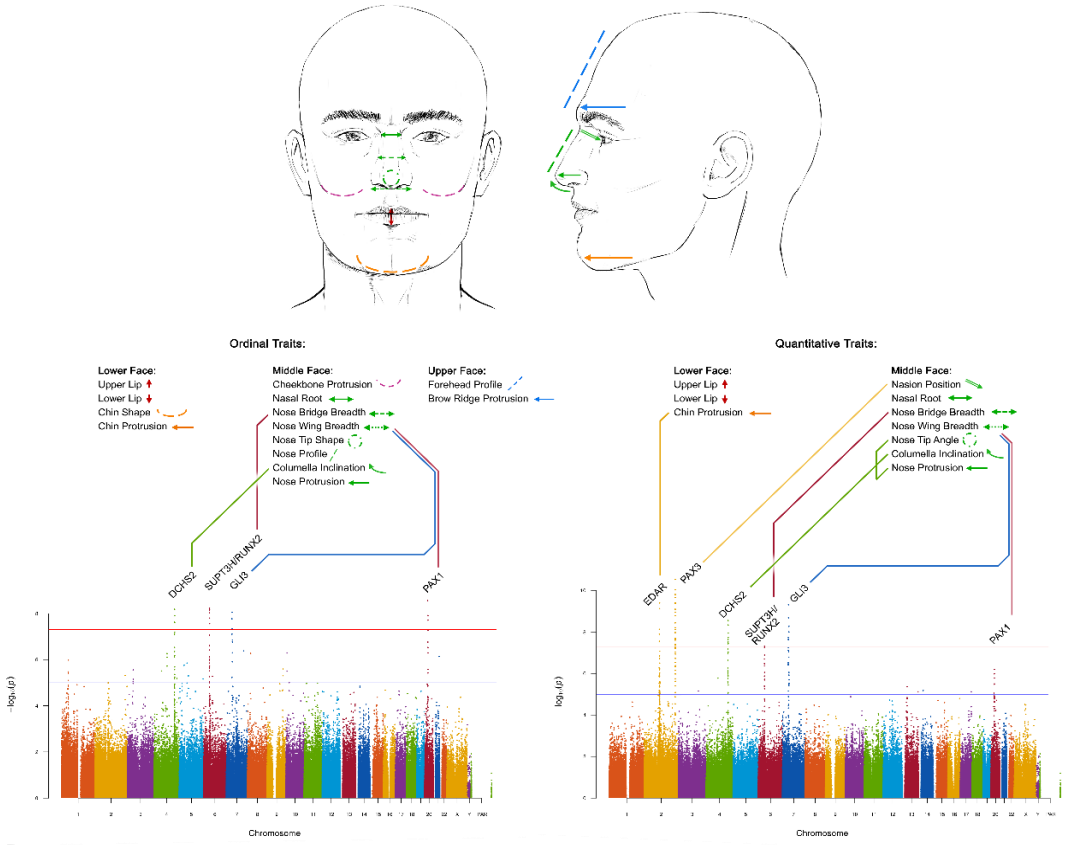


Figure 3.12. GWAS for facial traits in CANDELA samples. The first GWAS was performed with 14 ordinal traits from the face in 5,958 volunteers. To follow-up, quantitative proxies were obtained for 9 out 14 ordinal traits initially examined (and obtained a measure of nasion position) in a subset of 2,955 individuals and performed another GWAS. A 'composite' of Manhattan plot shows (on the left side, ordinal traits and on the right side, quantitative traits) the results across traits. All the SNPs with P values exceeding thresholds genome-wide suggestive (10^{-5}) are over the blue line and P values reaching the threshold genome-wide significant ($5x10^{-8}$) are above the red line.

Chromosomal Region	**Index SNP**	**Associated Trait**	***P* value**	**Candidate gene[1]**	**Alleles[2]**	**Effect size**	**% Variance explained**	**Replication *P* value**
4q31	rs12644248	Columella inclination	7×10^{-9}	***DCHS2***	A>G	-8.40×10^{-2}	0.49	4×10^{-3}
6p21	rs1852985	Nose bridge breadth	6×10^{-10}	***SUPT3H/RUNX2***	C>T	6.90×10^{-2}	0.71	5×10^{-3}
7p13	rs17640804	Nose wing breadth	9×10^{-9}	***GLI3***	C>T	-6.50×10^{-2}	0.62	6×10^{-3}
20p11	rs927833	Nose wing breadth	1×10^{-9}	*PAX1*	T>C	-7.70×10^{-2}	0.66	4×10^{-3}

Table 3.13. Properties of index SNPs in chromosomal regions showing genome-wide significant association with ordinal facial traits.

[1] For intragenic SNPs, gene names are shown in bold.

[2] Derived alleles are shown after ancestral alleles.

To account for the multiple phenotypes tested we performed a global false discovery rate (FDR) [305] test across all traits and SNPs and identified the same significantly associated regions (Table 3.14).

Rank	Chromosome	SNP	Trait	*P* value
1	**6**	**rs1852985**	**Nose Bridge Breadth**	**6.15E-10**
2	**20**	**rs927833**	**Nose Wing Breadth**	**1.05E-09**
3	6	rs1285029	Nose Bridge Breadth	5.80E-09
4	**4**	**rs12644248**	**Columella Inclination**	**6.64E-09**
5	6	rs6458435	Nose Bridge Breadth	6.79E-09
6	6	rs6458432	Nose Bridge Breadth	7.14E-09
7	**7**	**rs17640804**	**Nose Wing Breadth**	**8.85E-09**
8	6	rs1284964	Nose Bridge Breadth	9.40E-09
9	4	rs10517589	Columella Inclination	1.25E-08
10	6	rs12529907	Nose Bridge Breadth	1.59E-08
11	7	rs846312	Nose Wing Breadth	1.70E-08
12	4	rs10029359	Columella Inclination	2.40E-08
13	4	rs12651681	Columella Inclination	2.47E-08
14	6	rs35565233	Nose Bridge Breadth	2.73E-08
15	4	rs12506449	Columella Inclination	6.57E-08
16	4	rs6821649	Columella Inclination	7.90E-08
17	6	rs12528232	Nose Bridge Breadth	8.71E-08

Table 3.14. False Discovery Rate multiple testing correction. False Discovery Rate (FDR) method of multiple testing correction [305] was applied to the combined set of all 14 categorical trait GWAS P-values. The FDR method controls the overall Type I error rate (false rejection of true null hypotheses) at a specified level 0.05, while not being as overly conservative as the Bonferroni correction method. The Benjamini–Hochberg procedure for FDR was applied, giving a significance threshold of 9.05E-08 for all tests. Using this procedure, 17 SNPs from 4 regions were significant, corresponding to the same reported associated regions in Table 3.13. The FDR results are presented below, with the four index SNPs from Table 3.14 highlighted in bold (Table 3.15).

We examined association for each index SNP (the variant with the lowest *P* value in a chromosomal region, Table 3.13) in all countries sampled separately and combined results as a meta-analysis using METAL (Table 3.16) [299]. For all associations, significant effects were in the same direction in all countries (Table 3.16), the variability of effect size across countries reflecting sample size (Figure 3.13). There was no significant effect size heterogeneity across countries for any of the associations.

Region	SNP	Allele	Trait	*P* value	Beta	S.E.	Q Stat. *P* value
4q31	rs12644248	G	Columella Inclination	1.92E-07	-0.0741	0.0142	0.3139
6p21	rs1852985	T	Nose Bridge Breadth	5.64E-10	0.0658	0.0110	0.1238
7p13	rs17640804	C	Nose Wing Breadth	3.69E-09	0.0612	0.0104	0.9355
20p11	rs927833	T	Nose Wing Breadth	5.84E-12	0.0831	0.0121	0.7114

Table 3.15. Overall meta-analysis P-values.

SNP	Trait	Composite *P* value (n=5958)	Colombia (n=1303)	Brazil (n=608)	Chile (n=1651)	Mexico (n=1165)	Peru (n=1231)
rs12644248	Columella Inclination	1.92E-07	6.16E-01	1.52E-01	1.64E-03	5.87E-02	6.66E-05
rs1852985	Nose Bridge Breadth	5.64E-10	2.55E-03	3.78E-01	3.63E-06	2.26E-03	1.40E-02
rs17640804	Nose Wing Breadth	3.69E-09	1.69E-02	7.22E-02	8.62E-04	3.39E-03	8.88E-03
rs927833	Nose Wing Breadth	5.84E-12	8.96E-05	1.86E-01	5.52E-03	6.39E-04	3.50E-04

Table 3.16. Country-wise breakdown of P-values.

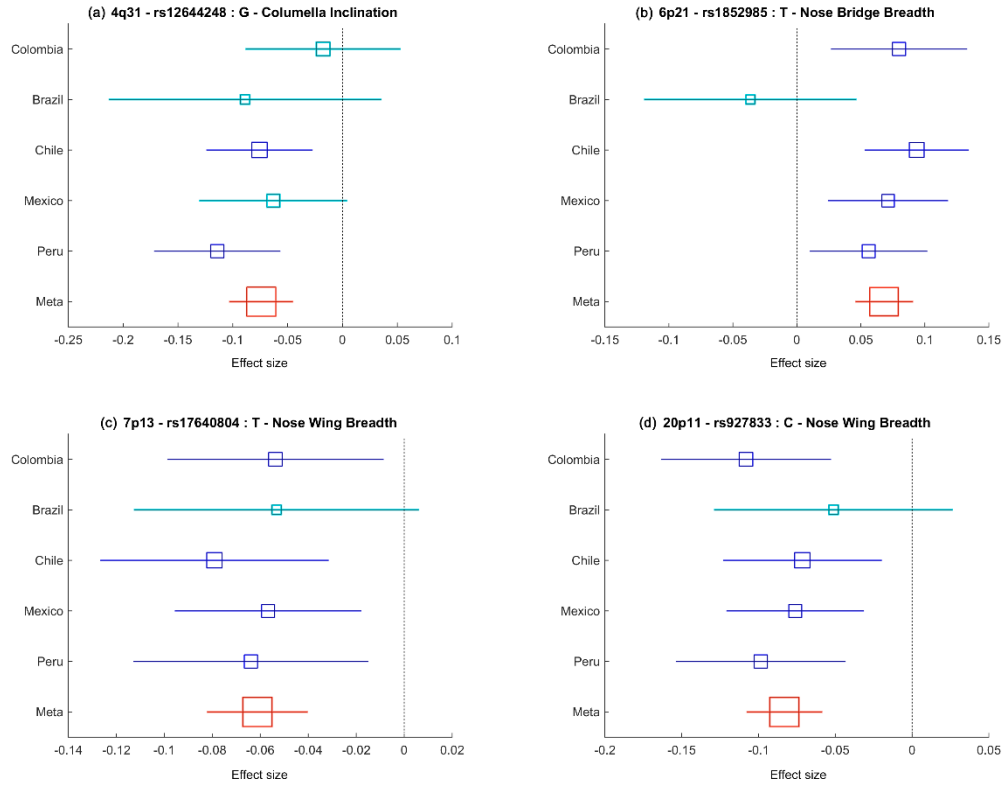


Figure 3.13. Effect sizes (regression coefficients) for the derived allele at index SNPs in the genome regions associated with ordinal face traits. (a) 4q31 rs12644248, (b) 6p21 rs1852985, (c) 7p13 rs17640804, (d) 20p11 rs927833. Estimates obtained in each country are shown as blue boxes. Red boxes indicate estimates obtained in the meta-analysis. Box size is proportional to sample size. Horizontal bars indicate confidence intervals representing 2 x standard errors. Intervals that include zero (that is, non-significant effects) are shown in light blue.

In order to exploit the correlations observed between various facial traits a multivariate GWAS was performed [306], but this approach did not identify any additional associated regions (Table 3.17). Multivariate regression analysis, an extension of the single-trait regression performed in GWAS, was performed to check if correlations between traits can lead to new or stronger genetic signals. The regression analysis was performed for each SNP to test for association with all traits combined while adjusting for covariates age, sex, BMI and genetic PCs. The method provides a regression coefficient for each trait, and the vector of coefficients is tested jointly using the Wald test [307] to check if it deviates significantly from a zero vector. The test employed here is nearly identical to the Wald test performed in the multivariate linear regression model analysis used in Zhou and Stephens (2014) [308], the only difference being in the use of genetic PCs instead of genetic kinship matrix; the difference is minor since the PCs are top eigenvectors of the kinship matrix.

P-values from this test that exceed the suggestive significance threshold of 10^{-5} are reported below, with the four index SNPs from Table 3.13 highlighted in bold. No new gene regions were found to be significantly associated. This is expected in cases when the correlation between traits is relatively low and effects of SNPs are not shared across phenotypes [309].

Chromosome	SNP	Position	Nearest gene	Wald *P* value	-log10(P)
4	rs9995821	154828366	DCHS2	1.97E-07	6.70
4	**rs12644248**	155235392	DCHS2	5.42E-06	5.27
6	rs542444	44800015	SUPT3H	1.80E-08	7.75
6	rs12528232	44982593	SUPT3H	1.56E-08	7.81
6	rs1285007	45102025	SUPT3H	5.34E-09	8.27
6	rs1285029	45122247	SUPT3H	8.87E-09	8.05
6	rs1284964	45176540	SUPT3H	9.39E-09	8.03
6	rs6458432	45205017	SUPT3H	5.98E-09	8.22
6	rs6458435	45222674	SUPT3H	5.96E-09	8.22
6	rs12529907	45255230	SUPT3H	1.39E-08	7.86
6	**rs1852985**	45329656	SUPT3H| RUNX2	8.65E-10	9.06
6	rs35565233	45349497	RUNX2	2.76E-08	7.56
7	rs846312	42128896	GLI3	2.88E-07	6.54
7	**rs17640804**	42131390	GLI3	2.54E-07	6.59
20	**rs927833**	22041577	PAX1	1.77E-08	7.75

Table 3.17. Multivariate association analysis combining all traits. P-values that exceed the suggestive significance threshold of 10^{-5}. Four index SNPs from Table 3.13 are highlighted in bold. There were no new genomics regions showing significant associations.

3.3.7 Follow-up analyses

After the GWAS described above, data from an additional set was obtained from 501 individuals from the same countries as for the GWAS and used as a replication sample

(descriptive features of this sample are presented in Figure 3.14 and Figure 3.15). These individuals were phenotyped and genotyped as for the GWAS sample. Association tests for the four index SNPs in Table 3.13 were performed using the same regression model as for the discovery GWAS, with a Bonferroni-adjusted threshold for significance of 0.05/4 = 0.0125. All tests were found to be significant in this replication sample (Table 3.13).

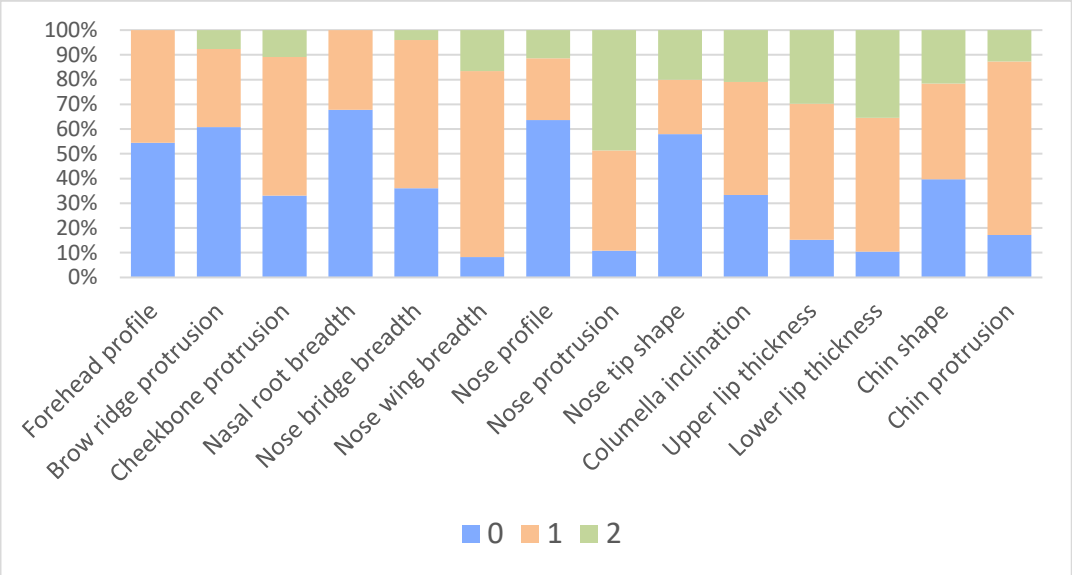


Figure 3.14. Frequency distribution of categorical phenotypes in the replication sample.

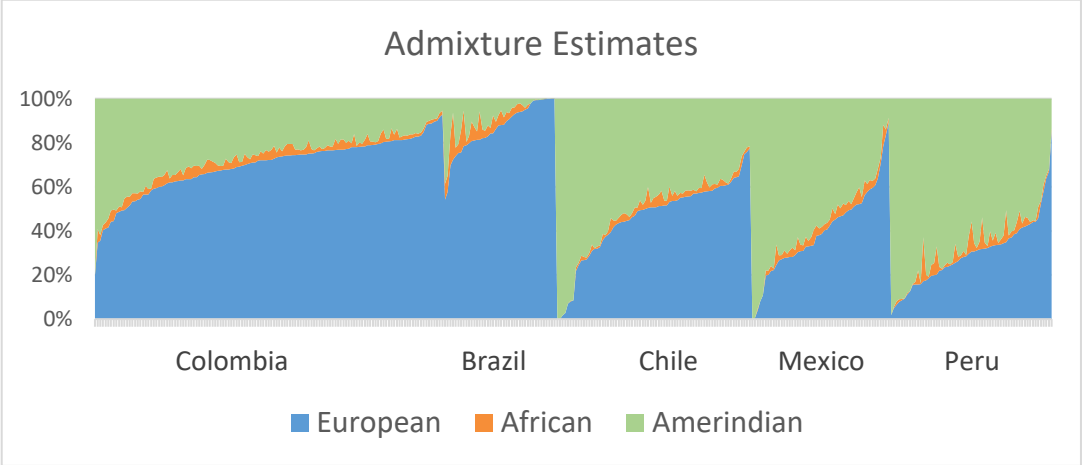


Figure 3.15. Individual ancestry in the replication sample. Three main continental ancestries: European, African and Amerindian by country.

The ordinal facial trait GWAS were also followed-up by obtaining facial measurements (distances and angles) related to the ordinal traits initially examined and performing a GWAS on these quantitative data. These measurements were obtained mainly using 3D anatomical landmark coordinates available for 2,955 of the

individuals included in the ordinal trait GWAS [274] (Figure 3.3). These landmarks allowed to define quantitative proxies for seven of the ordinal facial traits, the other traits having no appropriate 3D landmarks allowing related measurements to be obtained (Table 3.18).

Face section	Trait	Quantitative definition[a,b]
Upper	Forehead profile	–
	Brow ridge protrusion	–
Middle	Nasion position	Distance from landmark 18 to the mid-point of a line joining landmarks 8 and 16
	Cheekbone protrusion	–
	Nasal root breadth	–
	Nose bridge breadth	–
	Nose wing breadth	Distance between landmarks 20 and 22
	Nose profile	–
	Nose protrusion	Distance of landmark 19 to a line joining landmarks 18 and 21
	Nose tip shape	Angle between landmarks 18-19-21
	Columella inclination	Angle between landmarks 19-21-23
Lower	Upper lip thickness	Distance between landmarks 23 and 26
	Lower lip thickness	Distance between landmarks 26 and 29
	Chin shape	–
	Chin protrusion	Ratio of distance between landmarks 30-32 to distance between landmarks 29-32

Table 3.18. Measurements defined using 3D facial landmarks. [a]The quantitative definitions refer to coordinate numbers as shown in Figure 3.3. [b]A dash (-) indicates that a measurement corresponding to this ordinal phenotype could not be defined based on the 3-D landmarks.

Since the ordinal assessment of nose root and bridge breadth produced genome-wide significant associations (but could not be measured with the 3D landmarks available), I carried out 2D landmarking of the frontal photographs of these 2,955 individuals and also obtained measurements for these two traits (Table 3.6 and Figure 3.4). In addition, we used the 3D landmark coordinates to obtain a measure of nasion position so as to evaluate in our sample the reported association of this feature with SNPs in the *PAX3* gene region [7,8].

The ordinal variables showed a moderate to high (and significant) correlation with the quantitative variables (all permutation P values < 0.0005; Table 3.19 and Figure 3.16). Correlation between ordinal and quantitative traits was strongest for nose wing breadth and lower lip thickness (both with $r = 0.70$) and lowest for columella inclination ($r = 0.16$).

Trait	Correlation	*P* value
Nasal root breadth	0.48	<0.0005
Nose bridge breadth	0.37	<0.0005
Nose wing breadth	0.70	<0.0005
Nose protrusion	0.58	<0.0005
Nose tip angle	0.17	<0.0005
Columella inclination	-0.16[a]	<0.0005
Upper lip thickness	0.60	<0.0005
Lower lip thickness	0.70	<0.0005
Chin protrusion	0.58	<0.0005

Table 3.19. Correlations between categorical and quantitative face traits. [a]Correlation for Columella inclination is expected to be negative. Since the categorical trait is coded as 0-1-2 = up-straight-down, which corresponds to the columella angle being wide-moderate-narrow, hence it is negatively correlated with angle magnitude.

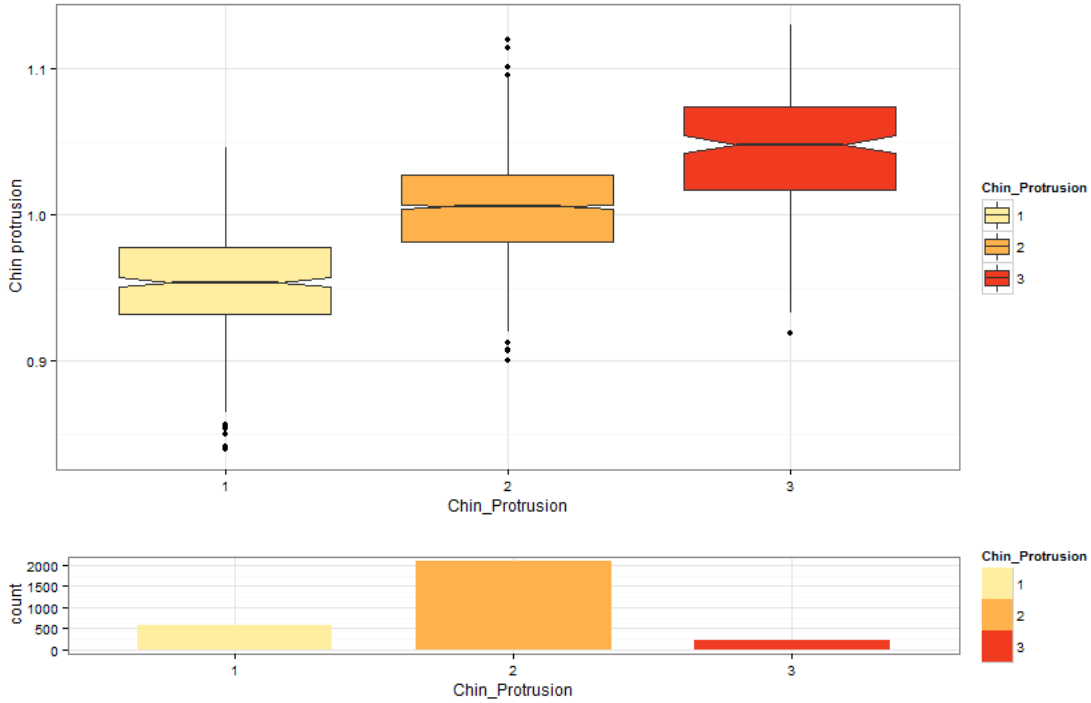


Figure 3.16. Correlation between ordinal Chin Protrusion and quantitative Chin Protrusion. This box-plot is showing the correlation between ordinal Chin Protrusion and quantitative Chin Protrusion as an example, remaining box plots are in Appendix A (Figures A.14 - A.21.).

The pattern of correlation amongst quantitative traits was similar to that observed for the ordinal traits, as was the correlation between quantitative traits and covariates (Table 3.20 and Table 3.21).

	NRB	NBB	NWB	NPR	NTA	CI	ULT	LLT	CP	NAP
Nasal root breadth		**<0.0005**	**<0.0005**	**<0.0005**	**<0.0005**	0.008	**<0.0005**	**<0.0005**	0.729	**<0.0005**
Nose bridge breadth	**0.568**		<0.0005	0.005	0.002	**<0.0005**	**<0.0005**	**<0.0005**	0.02	**<0.0005**
Nose wing breadth	**0.213**	0.389		0.011	**<0.0005**	**<0.0005**	0.034	0.053	0.005	**<0.0005**
Nose protrusion	**-0.179**	-0.052	-0.047		**<0.0005**	0.053	**<0.0005**	**<0.0005**	**<0.0005**	**<0.0005**
Nose tip angle	**0.198**	0.057	**-0.068**	**-0.635**		**<0.0005**	**<0.0005**	**<0.0005**	**<0.0005**	**<0.0005**
Columella inclination	-0.049	**-0.111**	**-0.217**	-0.036	**0.443**		**<0.0005**	**<0.0005**	0.008	**<0.0005**
Upper lip thickness	**0.161**	**0.122**	0.039	**-0.115**	**0.119**	**-0.272**		**<0.0005**	**<0.0005**	**<0.0005**
Lower lip thickness	**0.121**	**0.073**	-0.036	**-0.075**	**0.152**	**-0.097**	**0.573**		**<0.0005**	0.001
Chin protrusion	-0.006	0.043	0.052	**0.081**	**-0.126**	0.049	**-0.366**	**-0.499**		**<0.0005**
Nasion position	**-0.17**	**-0.094**	**-0.183**	**0.292**	**-0.148**	**0.138**	**-0.151**	-0.064	**0.093**	

Table 3.20. Correlation for quantitative traits. Correlation values are presented in the lower left triangle, with corresponding permutation P values in the upper right triangle. Correlations with significant P values (<0.0006, Bonferroni-adjusted threshold) and their corresponding P values are highlighted in bold.

	Middle							Lower		
Covariate	NAP	NRB	NBB	NWB	NPR	NTA	CI	ULT	LLT	CP
Sex	**-0.35**	**0.28**	0.06	**-0.11**	**-0.06**	**0.16**	0.00	**0.13**	**0.13**	**-0.08**
Age	0.01	**-0.12**	-0.02	**0.16**	**0.26**	**-0.27**	0.00	**-0.43**	**-0.47**	**0.21**
BMI	**-0.07**	-0.04	-0.02	**0.15**	**-0.06**	**-0.09**	-0.05	**-0.20**	**-0.23**	**0.23**
African anc.	**-0.06**	**0.23**	**0.25**	**0.19**	-0.03	0.06	-0.05	**0.17**	**0.19**	**-0.07**
European Anc.	**0.38**	**-0.12**	**-0.14**	**-0.24**	**0.36**	**-0.19**	**0.19**	**-0.22**	**-0.07**	**0.20**
American anc.	**-0.32**	0.05	**0.07**	**0.17**	**-0.33**	**0.16**	**-0.16**	**0.15**	0.01	**-0.17**

Table 3.21a. Correlation between quantitative face traits and sex, age and ancestry.

	Middle							Lower		
Covariate	NAP	NRB	NBB	NWB	NPR	NTA	CI	ULT	LLT	CP
Sex	**<5E-4**	**<5E-4**	2.E-03	**<5E-4**	**<5E-4**	**<5E-4**	9.E-01	**<5E-4**	**<5E-4**	**<5E-4**
Age	6.E-01	**<5E-4**	2.E-01	**<5E-4**	**<5E-4**	**<5E-4**	9.E-01	**<5E-4**	**<5E-4**	**<5E-4**
BMI	**<5E-4**	4.E-02	4.E-01	**<5E-4**	**<5E-4**	**<5E-4**	2.E-03	**<5E-4**	**<5E-4**	**<5E-4**
African anc.	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	8.E-02	7.E-04	9.E-03	**<5E-4**	**<5E-4**	**<5E-4**
European Anc.	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**
American anc.	**<5E-4**	4.E-03	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	**<5E-4**	6.E-01	**<5E-4**

Table 3.21b. Correlation between quantitative face traits and covariates corresponding *P*- values.

As expected for continuous variables, heritability estimates based on the quantitative phenotypes (Table 3.22) are higher than obtained for the ordinal phenotypes and more in line with published estimates [28,51,304].

Face region	Trait	Heritability	S.E.	*P* value
Mid-face	Nasion Position	0.80	0.10	0.0E+00
	Nasal root breadth	0.80	0.09	0.0E+00
	Nose bridge breadth	0.89	0.09	0.0E+00
	Nose wing breadth	0.90	0.10	0.0E+00
	Nose protrusion	0.84	0.10	0.0E+00
	Nose tip shape	0.67	0.10	0.0E+00
	Columella inclination	0.77	0.11	0.0E+00
Lower face	Upper lip thickness	0.63	0.10	0.0E+00
	Lower lip thickness	0.46	0.10	2.3E-15
	Chin protrusion	0.45	0.09	0.0E+00

Table 3.22. Heritability of quantitative facial traits. It was moderate and significant for all quantitative traits. The lowest value was for Chin protrusion (0.45) and the highest heritability was Nose wing breadth (0.90) among ordinal facial features.

As before, a GWAS for the quantitative traits was performed using an additive multivariate regression model adjusting for age, sex, BMI and the first five PCs. Previous associations were replicated, nasion position with SNPs in 2q35 overlapping the *PAX3* gene region, with strongest association seen for rs7559271 (*P* value of 4×10^{-11}, Figure 3.12, Table 3.23, Figure 3.17). This is the same SNP producing strongest association in the Paternoster et al. (2012) GWAS[8].

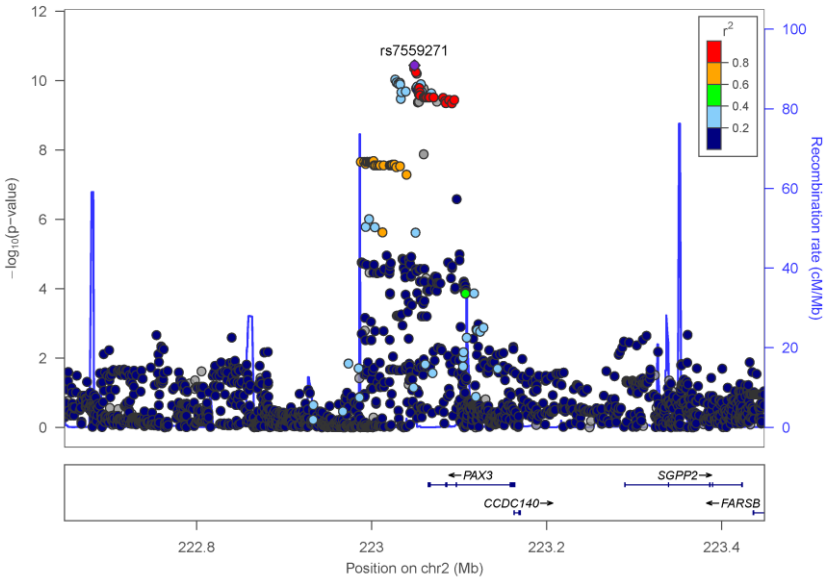


Figure 3.17. Replicated regional association plot in 2q35 and nasion position. This plot was produced in Locus Zoom [310].

In addition, genome-wide significant association was observed for six of the nine quantitative proxies of the ordinal traits initially examined (Figure 3.12 and Table 3.23). As for the ordinal assessments, the quantitative analysis of columella inclination, nose bridge breadth and nose wing breadth produced genome-wide significant associations with SNPs in 4q31, 6p21 and 7p13, respectively (Figure 3.12, Table 3.13 and Table 3.23). In addition, the 4q31 region also showed genome-wide significant association with two other measurements related to nose morphology: nose protrusion and nose tip angle, with strongest *P* values for SNPs rs2045323 of 1×10^{-8} and 2×10^{-8}, respectively. SNPs in 4q31 produced small but not genome-wide significant *P* values in the ordinal assessment of nose protrusion and nose tip angle (strongest *P* values of 4×10^{-4} and 3×10^{-4}, respectively). The 20p11 region, showing genome-wide significant association in the ordinal assessment of nose wing breadth, showed genome-wide suggestive association in the quantitative trait GWAS (strongest *P* value of 6×10^{-7} for SNP rs927833). Other than reproducing the associations detected with ordinal traits, the quantitative analyses detected a genome-wide significant association with chin protrusion for markers in 2q12 (strongest *P* value of 4×10^{-10}, for rs3827760; Figure 3.12 and Table 3.23). This marker had an association *P*-value of 1×10^{-4} in the ordinal assessment of chin protrusion.

A regression model similar to the one used in the GWAS analyses explains up to ~30% of the phenotypic variation for the traits with significant SNP associations, with each of the associated SNPs explaining about 1% of variation in the trait (Table 3.13, Table 3.23 and Table 3.24). The estimates of trait variance explained by associated SNPs are similar to those calculated for other anthropometric traits and are very close to the estimates obtained in a previous GWAS for facial features [8]

Chr. region	Index SNP	Associated Trait	*P* value	Candidate gene[1]	Alleles[2]	Effect size	% Variance explained
2q12	rs3827760	Chin protrusion	4×10^{-10}	***EDAR***	A>G	-7.60×10^{-3}	1.32
2q35	rs7559271	Nasion position	4×10^{-11}	***PAX3***	A>G	8.20×10^{-2}	1.33
4q31	rs2045323	Columella inclination[3]	3×10^{-9}	*DCHS2*	G>A	1.80×10^{-2}	0.63
4q31	rs2045323	Nose protrusion	1×10^{-8}	*DCHS2*	G>A	-5.90×10^{-4}	0.95
4q31	rs2045323	Nose tip angle	2×10^{-8}	*DCHS2*	G>A	1.60×10^{-2}	1.08
6p21	rs1852985	Nose bridge breadth	2×10^{-8}	***SUPT3H/ RUNX2***	C>T	4.40×10^{-4}	1.18
7p13	rs17640804	Nose wing breadth	5×10^{-10}	***GLI3***	C>T	-4.90×10^{-4}	1.15

Table 3.23. Properties of index SNPs in regions showing genome-wide significant association to quantitative facial traits.

[1] For intragenic SNPs, gene names are shown in bold.

[2] Derived alleles are shown after ancestral alleles.

[3] Columella inclination was measured as an angle which decreases at greater ordinal columella inclination are of opposite sign. rs12644248 (Table 3.13, Table 3.4 and Table 3.5), the index SNP associated with categorical columella inclination, has a P value of 4×10^{-8} for association with the quantitative assessment of columella inclination.

	R^2 explained by (%)								
Trait	Covariates	rs3827760	rs7559271	rs2045323	rs12644248	rs1852985	rs17640804	rs927833	All SNPs
Forehead profile	23.60	0.00	0.01	0.14	0.12	0.02	0.05	0.04	0.29
Brow ridge protrusion	39.81	0.01	0.02	0.02	0.22	0.24	0.00	0.00	0.46
Cheekbone protrusion	8.75	0.03	0.00	0.01	0.01	0.02	0.08	0.02	0.09
Nasal root breadth	5.84	0.07	0.17	0.02	0.02	0.21	0.02	0.00	0.45
Nose bridge breadth	7.08	0.01	0.01	0.00	0.09	**0.71**	0.06	0.01	0.92
Nose wing breadth	9.61	0.00	0.08	0.01	0.14	0.01	**0.62**	**0.66**	1.53
Nose profile	4.77	0.00	0.01	0.14	0.11	0.02	0.04	0.00	0.26
Nose protrusion	1.93	0.00	0.10	0.01	0.06	0.03	0.00	0.05	0.27
Nose tip shape	11.57	0.03	0.02	0.01	0.00	0.01	0.01	0.03	0.11
Columella inclination	5.29	0.00	0.07	0.44	**0.49**	0.09	0.03	0.00	0.86
Upper lip thickness	12.09	0.02	0.14	0.09	0.22	0.01	0.01	0.03	0.51
Lower lip thickness	9.61	0.08	0.09	0.01	0.19	0.04	0.00	0.01	0.41
Chin shape	2.77	0.00	0.03	0.03	0.09	0.02	0.01	0.03	0.14
Chin protrusion	4.83	0.15	0.01	0.03	0.03	0.02	0.03	0.09	0.36

Table 3.24. Ordinal facial traits. For all traits, it was calculated the fraction of trait variance explained by the covariates and by all index SNPs. A similar regression model to that used in the GWAS was used to estimate the R^2 of the model (representing the fraction of the variance of the trait explained by all regressors). The following models were applied:

Model 1: Trait ~ age + sex + BMI + PC1 … PC5

Model 2: Trait ~ age + sex + BMI + PC1 … PC5 + index SNP

Model 3: Trait ~ age + sex + BMI + PC1 … PC5 + all index SNPs

Model 2 was applied separately for each index SNP.
From models 2 and 3, the fraction of trait variance explained by the covariates (i.e. Model 1) was subtracted to get the additional contribution of the SNP(s).
SNPs showing genome-wide significant associations to a trait are highlighted in bold.

To assess independent evidence of association for the regions implicated here we examined SNPs that produced at least genome-wide suggestive *P* values in the two GWAS for facial features that have been published [7,8]. We found that SNP rs2108166 in high LD ($r^2 = 0.77$, D' = 1) with the index SNP of the 7p13 region was associated with nose wing breadth (rs17640804), produced an association *P* value of 5×10^{-7} with the same trait in the study of Liu et al. (2012)[7]. In addition, evidence of association between rs3827760 and chin shape has been reported in a candidate gene study of a Central Asian population [287].

It has been suggested that gene regions associated with non-syndromic cleft lip and palate might impact on normal variation in facial morphology [7,311]. Although the regions reported to be associated with NSCL/P do not overlap with those identified here, index SNPs in each NSCL/P region were selected and tested for association of these SNPs with the ordinal and quantitative facial traits examined here (Table 3.25

and 3.26). Few tests survived Bonferroni correction, mostly involving SNPs associated with quantitative nose-breadth traits (nose root, nose bridge and nose wing breadth; Table 3.26). A global one-sided Kolmogorov-Smirnoff test was significant both for ordinal and quantitative traits (*P* value ~10^{-3}; Figure 3.18 and 3.19) and a polygenic risk score test combining all 15 index SNPs was significant for the nose-breadth traits (Table 3.27 and 3.28).

Region	Chr.	Gene	SNP	Strongest associated trait	min P	min P (adjusted)
1	1p36.13	*PAX7*	rs560426[1,3]	Nasal root breadth	0.0275	0.3846
2	1p22.1	*ABCA4*	rs742071[1,2]	Nose wing breadth	0.0285	0.3983
3	1q32.2	*IRF6*	rs2235371[4,6]	Brow ridge protrusion	0.0019	0.0269
4	2p21	*THADA*	rs7590268[1,7]	Nose protrusion	0.0744	1.0000
5	3p11.1	*EPHA3*	rs7632427[1]	Nose protrusion	0.0342	0.4792
6	3p12.3	*COL8A1/FILIP1L*	rs793464[2]	Nose wing breadth	0.0281	0.3933
7	8q21.3	*DCAF4L2*	rs12543318[1]	Brow ridge protrusion	0.0287	0.4012
8	8q24.21	--	rs987525[1,3,8,9]	Nose protrusion	0.0727	1.0000
9	10q25.3	*VAX1*	rs7078160[1,6,7]	Nose bridge breadth	0.0316	0.4425
10	13q31.1	*SPRY2*	rs8001641[1]	Nasal root breadth	0.0361	0.5051
11	15q22	*TMP1*	rs1873147[1]	Chin shape	0.0416	0.5817
12	16p13.3	*ADCY9*	rs8049367[6]	Nose wing breadth	0.0122	0.1708
13	17p13.1	*NTN1*	rs479177[6]	Nose protrusion	0.0207	0.2901
14	17q22	*NOG*	rs227731[1,7]	Chin protrusion	0.0131	0.1838
15	20q12	*MAFB*	rs13041247[1,3,6]	Nasal root breadth	0.0335	0.4693

Table 3.25. Testing for association of NSCL/P loci with the ordinal facial traits assessed here. The smallest P value for each SNP across ordinal traits is presented in this table, in addition to the Bonferroni-corrected (adjusted by the number of traits) P value [312-320].

Region	Chr.	Gene	SNP	Trait	min P	min P (adjusted)
1	1p36.13	*PAX7*	rs560426[1,3]	Nose wing breadth	0.00082	0.00816
2	1p22.1	*ABCA4*	rs742071[1,2]	Nasion position	0.08715	0.87150
3	1q32.2	*IRF6*	rs2235371[4,6]	Upper lip thickness	0.00661	0.06612
4	2p21	*THADA*	rs7590268[1,7]	Nose bridge breadth	0.02832	0.28320
5	3p11.1	*EPHA3*	rs7632427[1]	Chin protrusion	0.10080	1.00000
6	3p12.3	*COL8A1/FILIP1L*	rs793464[2]	Lower lip thickness	0.02898	0.28980
7	8q21.3	*DCAF4L2*	rs12543318[1]	Columella inclination	0.01585	0.15850
8	8q24.21	--	rs987525[1,3,8,9]	Nose wing breadth	0.05817	0.58170
9	10q25.3	*VAX1*	rs7078160[1,6,7]	Nasal root breadth	0.24310	1.00000
10	13q31.1	*SPRY2*	rs8001641[1]	Nasion position	0.23940	1.00000
11	15q22	*TMP1*	rs1873147[1]	Nasal root breadth	0.00089	0.00889
12	16p13.3	*ADCY9*	rs8049367[6]	Nasal root breadth	0.02382	0.23820
13	17p13.1	*NTN1*	rs479177[6]	Columella inclination	0.03354	0.33540
14	17q22	*NOG*	rs227731[1,7]	Nasion position	0.02255	0.22550
15	20q12	*MAFB*	rs13041247[1,3,6]	Nasal root breadth	0.00002	0.00017

Table 3.26. **Testing for association of NSCL/P loci with the quantitative facial traits assessed here.** The smallest P value for each SNP across quantitative traits is presented in this table, in addition to the Bonferroni-corrected (adjusted by the number of traits) P value.

Combining all SNPs and all traits, a one-sided Kolmogorov-Smirnov test for deviation from the null hypothesis of no association was significant both for ordinal and for quantitative face traits. For both types of traits, a plot of the empirical cumulative distribution function (eCDF, in blue) of all individual P-values is shown next to the expected CDF (in purple) of P-values under the null of no association (Figure 3.18 and 3.19).

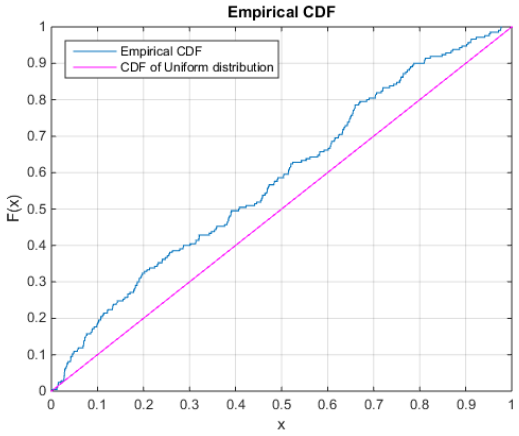


Figure 3.18. Kolmorov-Smirnov test for ordinal facial traits. The one-sided Kolmogorov-Smirnov test had a P-value of 0.0011.

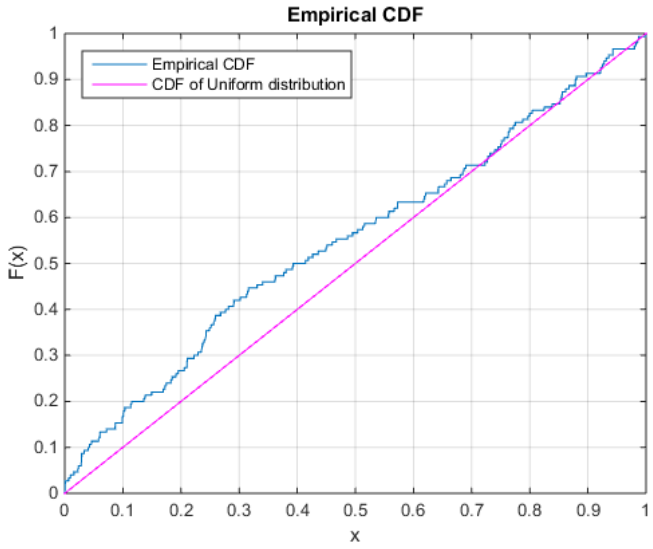


Figure 3.19.Kolmorov-Smirnov test for quantitative facial traits.The one-sided Kolmogorov-Smirnov test had a P-value of 0.0046.

After regressing each SNP with a trait (controlling for covariates), regression coefficients from the 15 SNPs were combined into a composite Polygenic Risk Score (PRS). Each trait was then regressed against their corresponding PRS (controlling for covariates). The P-value from this regression is referred to below as the unadjusted P-value. Since the second regression uses a PRS score that combines the individual best-fit regression coefficients from the first regression, using this PRS score as a single variable in the same regression model results in overfitting. As a remedy, 2,000 simulations were performed for each trait, in which phenotypes were randomized across genotypes of each SNP (to maintain the same genotype distribution but under the null of no association with the phenotypes). These simulated genotypes and phenotypes were used to calculate PRSs and obtain an empirical distribution of the PRS test statistic. For each trait, the observed PRS test statistic was compared to this empirical distribution to obtain an 'adjusted' P-value (Table 3.27 and 3.28).

Trait	Unadjusted P-value	Adjusted P-value
Forehead profile	9.26E-04	0.2310
Brow ridge protrusion	9.91E-07	0.0790
Cheekbone protrusion	8.98E-06	0.1985
Nasal root breadth	6.59E-08	0.0115
Nose bridge breadth	5.91E-06	0.1655
Nose wing breadth	2.13E-08	0.0070
Nose profile	3.44E-05	0.3105
Nose protrusion	2.15E-08	0.0110
Nose tip shape	3.65E-06	0.1195
Columella inclination	9.04E-03	0.9635
Upper lip thickness	4.16E-05	0.3485
Lower lip thickness	5.12E-05	0.3520
Chin shape	4.73E-06	0.1350
Chin protrusion	1.64E-06	0.0945

Table 3.27. Polygenic risk score for ordinal traits.

Trait	Unadjusted P-value	Adjusted P-value
Nasal root breadth	5.25E-12	<0.0005
Nose bridge breadth	1.15E-08	0.0060
Nose wing breadth	3.24E-07	0.0365
Nose protrusion	1.97E-04	0.5125
Nose tip shape	1.41E-03	0.8110
Columella inclination	1.88E-05	0.2420
Upper lip thickness	4.22E-05	0.3315
Lower lip thickness	3.16E-04	0.6015
Chin protrusion	3.10E-04	0.5925
Nasion Position	4.22E-05	0.3420

Table 3.28. **Polygenic risk score for quantitative traits**.

3.3.8 Candidate genes in regions associated with facial morphology

Chin protrusion showed significant genome-wide association with several SNPs in the region 2q13, the index SNP was rs3827760 (4×10^{-10}), which is situated in the Ectodysplasin A receptor gene (*EDAR*, Figure 3.20a). Genome-wide scans detected a mutation in this variant (V370A), which has been suggested to have been positively selected in East-Asian populations ~ 30.000 years ago [321,322]. Interestingly, mutations in the *EDA* pathway, which is comprised of three genes, *EDA-EDAR-EDARADD*, cause Hypohidrotic ectodermal dysplasia (HED) [323-326]. HED is a disease characterized by being born with a lower number of sweat glands, oligodontia, decrease in the amount of hair and facial dysmorphia, such as frontal prominence, growth modification of the cranial base and more protrusive chin and mandible [323,327]. By contrast, the V370A mutation acts as a protector allele in people with this disease [328]

and it has been demonstrated in vitro that the derived allele (EDAR370A) produces stronger signalling than the ancestral allele [329,330], individuals who carry this allele have thicker hair and shovel shape incisors[211,322,328]. In this study, it was found that the effect of EDAR370A produces chin less protruded contrary to HED.

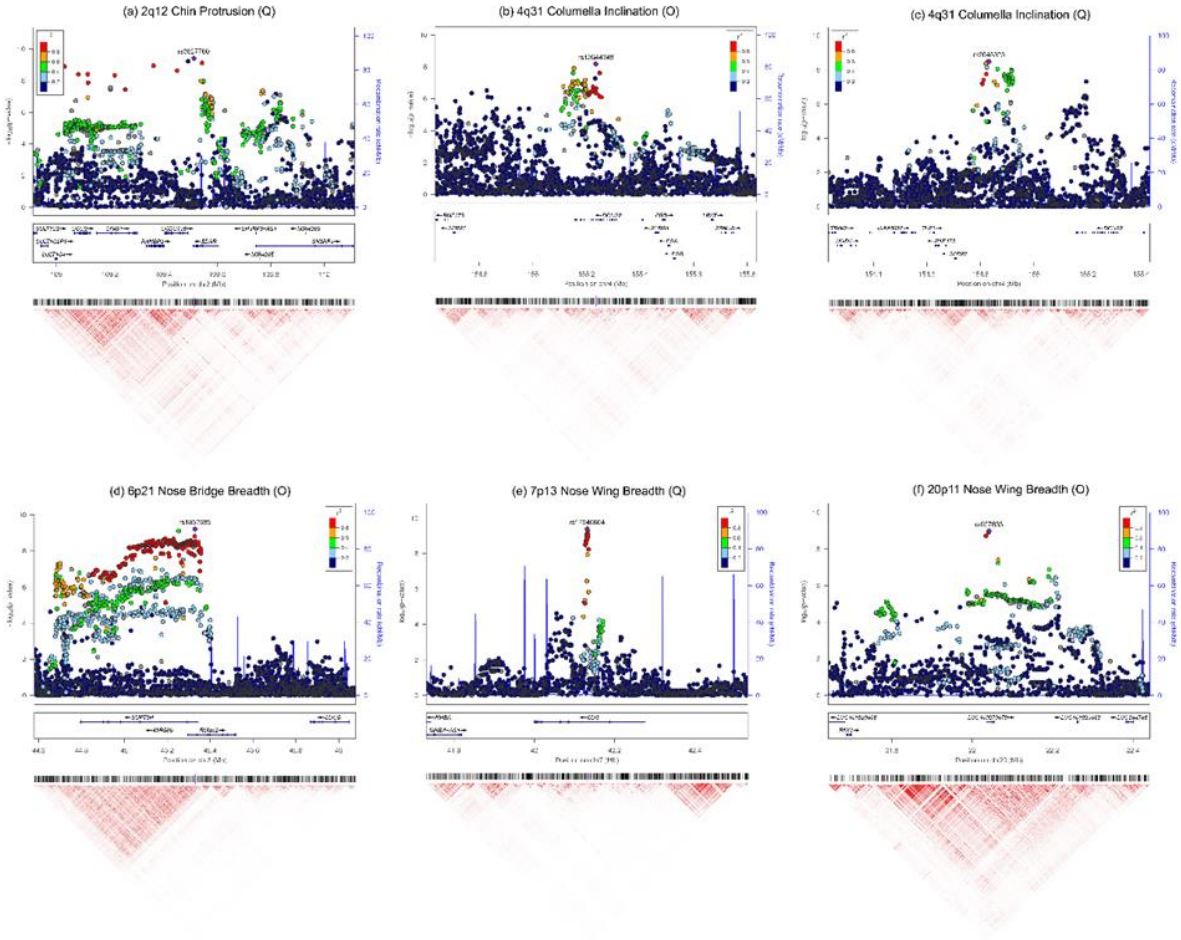


Figure 3.20. Genomic regions showing genome-wide significant association with face traits. For each facial feature, we show the results that achieved strongest statistical significance regardless of the type of variable analysed (ordinal, O; or quantitative, Q). (a) 2q12 (Q), (b) 4q31 (O), (c) 4q31 (Q), (d) 6p21 (O), (e) 7p13(Q), (f) 20p11 (O). Association results (on a _log10 P scale; left y-axis) are shown for SNP 500 kb on either side of the index SNP (purple diamond; Table 3.13 and Table 3.23) with the marker (dot) colour indicating the strength of LD (r2) between the index SNP and that SNP in the 1000 genomes AMR data set. Local recombination rate in the AMR data is shown as a continuous blue line (scale on the right y-axis). Genes in each region, their intron–exon structure, direction of transcription and genomic coordinates (in Mb, using the NCBI human genome sequence, Build 37, as reference) are shown at the bottom. Plots were produced with LocusZoom [310]. Below each region we also show an LD heatmap (using r2, ranging from red indicating r^2=1 to white indicating r^2=0) produced using a MATLAB [290] implementation similar to Haploview [331].

To perform functional replication of the effect of *EDAR* seen on chin protrusion, lower jaw protrusion was examined, as a proxy to chin protrusion in *Edar* mouse mutants. In addition to wild-type mice, $Edar^{dlJ}$ and $Edar^{Tg951}$ mouse lines were examined, which have a loss and a gain of *Edar* function respectively (Section 3.2.7). Lower jaw protrusion was measured from landmarking of side photographs (Figure 3.9 and Figure 3.10) of 14-15 days old mice for four genotype groups: $Edar^{dlJ/dlJ}$ mutant and heterozygous $Edar^{dlJ/+}$ littermates, wild type mice and $Edar^{Tg951}$ transgenic mice, arranged according to increasing *Edar* expression. A multivariate linear regression of lower jaw protrusion on genotype, age and sex provided strongly significant association with genotype, with jaw protrusion becoming lower with increased *Edar* expression, replicating the direction of *EDAR* effect seen in CANDELA samples (Figure 3.21 and Table 3.29).

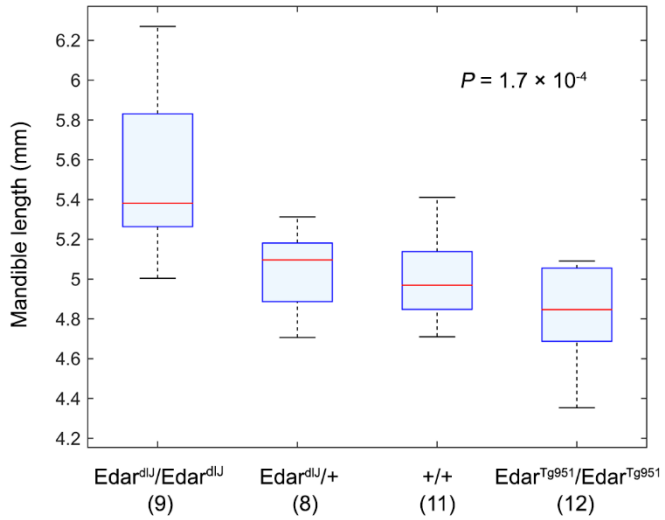


Figure 3.21. Effect of *Edar* genotype on mouse mandible length. The boxplots show the mandible length (y-axis) in mice with different *Edar* genotypes (X-axis). The measure of mandible length shown is the projected distance between head landmarks 5 and 10 (Figure 3.9 and Figure 3.10). Boxplot whiskers extend to data points within 1.5 times the interquartile range on both sides. The numbers in parenthesis below genotypic categories refer to the number of mice examined for each genotype.

P values and regression coefficients (beta) for the linear regression between mandible length (as a proportion of head size) and *Edar* genotype are given below (controlling for age and sex). For each landmark on the mandible (5-7, Figure 3.9) two types of distances were measured – direct and projected – with reference to landmark 10 (Figure 3.9 and Figure 3.10). These distances were divided by head size to convert into a proportion (Table 3.29). All the P values were significant, suggesting that overall

mandible length is affected by *Edar* genotype. The negative beta implies that mandible length becomes shorter with increased *Edar* function, consistent with the effect we see in humans.

Type	Distance	beta	*P* value
Projected	5–10	-0.01916	1.7E-04
Projected	6–10	-0.01543	2.6E-03
Projected	7–10	-0.01809	1.1E-04
Direct	5–10	-0.01680	1.3E-03
Direct	6–10	-0.01255	1.8E-02
Direct	7–10	-0.01136	1.3E-02

Table 3.29. Effect of *Edar* genotype on mouse mandible length. Regression analysis indicates a significant effect of *Edar* genotype on mandible length (P value $1.7x10^{-4}$). All the P values were significant, suggesting that overall mandible length is affected by *Edar* genotype. The negative beta implies that mandible length becomes shorter with increased *Edar* function, consistent with the effect we see in humans.

SNPs in the 4q31 region with *P* values above the suggestive association threshold in the ordinal trait assessment of columella inclination extend over ~400Kb from the 3' half of the Dachsous Cadherin-Related 2 gene (*DCHS2*) into the *DCHS2-SFRP2* (Secreted Frizzled-related protein 2) intergenic region (Figure 3.20 b and Table 3.23), with strongest association seen for SNP rs12644248 within *DCHS2* (*P* value 7×10^{-9}). Noticeably, although association analyses based on the quantitative assessment of columella inclination also show genome-wide significant association for rs12644248 (*P* value of 4×10^{-8}), the quantitative analyses show that SNPs in the *DCHS2-SFRP2* intergenic region have an even stronger association, peaking at rs2045323 (*P* value of 3×10^{-9}, Table 3.23, Figure 3.20). A similar pattern of association is seen for the quantitative assessments of nose protrusion and nose tip angle, with strongest association for both traits being observed for rs2045323 (*P* values of 1×10^{-8} and 2×10^{-8} respectively, Table 3.23, Figure 3.22 a and b), association with rs12644248 only exceeding the genome-wide suggestive threshold (respective *P* values of 8×10^{-6} and of 6×10^{-6} for nose protrusion and nose tip angle).

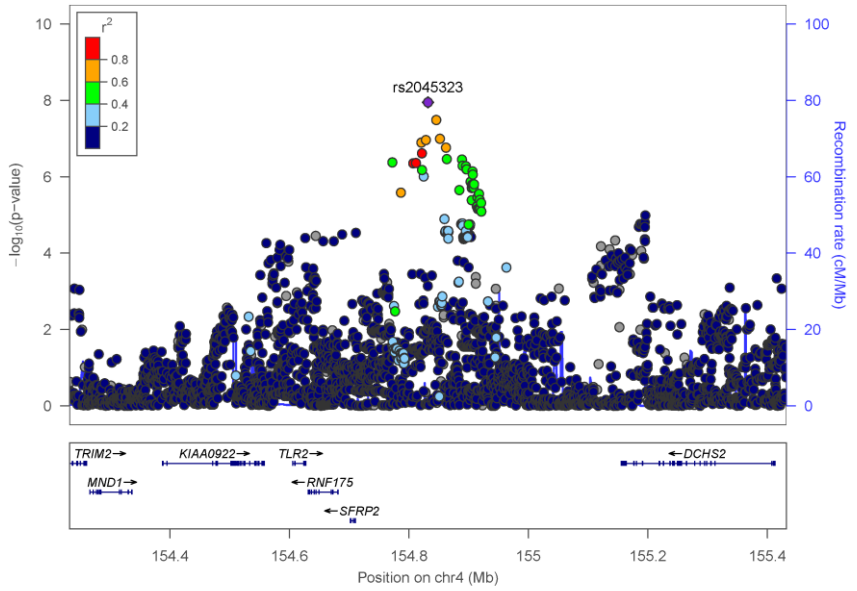


Figure 3.22a: 4q31 and quantitative assessment of nose protrusion. This plot was produced in Locus Zoom [310].

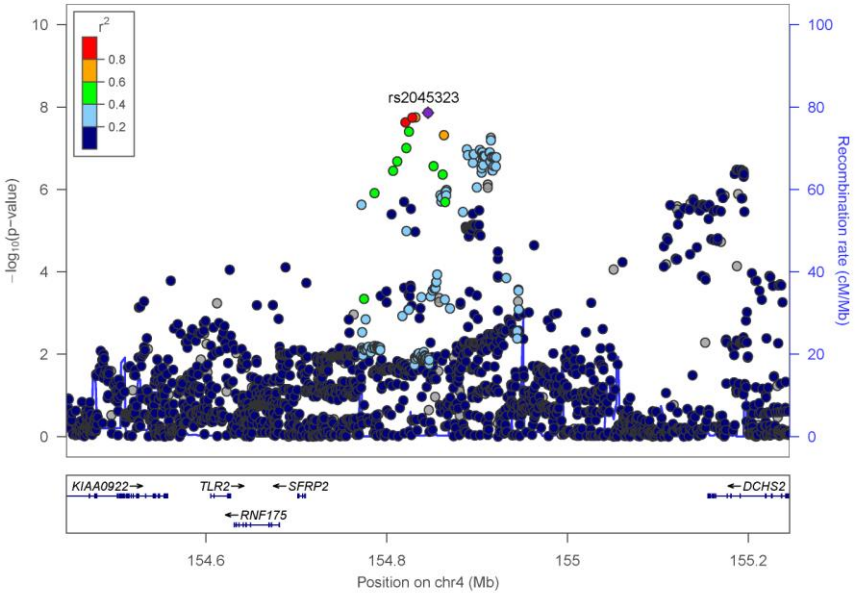


Figure 3.22b: 4q31 and quantitative assessment of nose tip angle. This plot was produced in Locus Zoom [310].

As seen in Figure 3.20 and Figure 3.22 a and b, the 4q31 region shows two peaks of association for measurements of columella inclination, nose protrusion and nose tip angle. Strongest association at one peak occurs at SNP rs2045323 in the *DCHS2-SFRP2* intergenic region. The second peak has strongest association at rs12644248 in the *DCHS2* gene. This is the only genome-wide significant association peak for the

ordinal assessment of columella inclination (Table 3.13, Figure 3.20). To evaluate the independence of these two signals we performed regional association test conditioned on rs12644248 or on rs2045323 on the quantitative nose traits. Columella inclination, nose protrusion and nose tip angle all showed the same behaviour. Below I present the plots for nose protrusion, those for the other two traits being very similar. SNP rs2045323 is not in strong LD with rs12644248 and tests conditioned on either SNP attenuate the signal of association at the other SNP but do not abolish it entirely (Figure 3.23 a and b). These observations suggest that the signal of association around rs2045323 in the *DCHS2-SFRP2* intergenic region is somewhat independent from that peaking at rs12644248 within *DCHS2*.

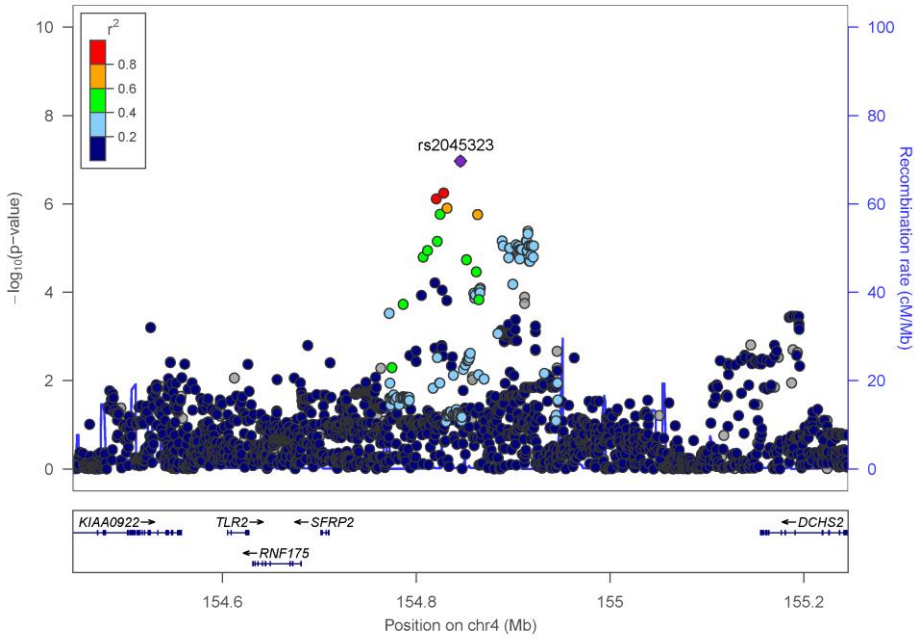


Figure 3.23a: Regional association plot for nose protrusion conditioned on rs12644248. This plot was produced in Locus Zoom [310].

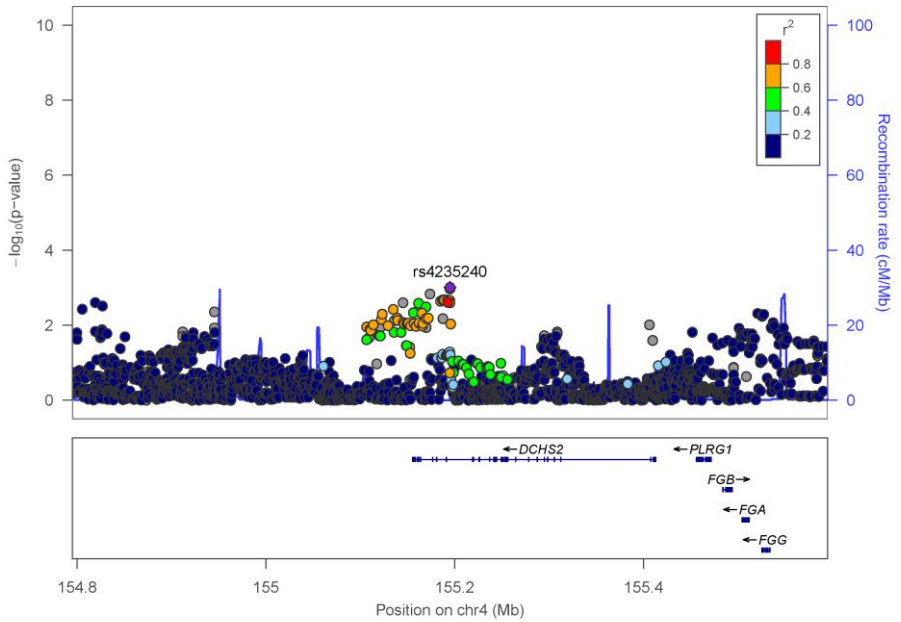


Figure 3.23b: Regional association plot for nose protrusion conditioned on rs2045323. This plot was produced in Locus Zoom [310].

Intergenic SNP rs2045323 is in an evolutionarily conserved region (Figure 3.26), suggesting that this SNP could play a role in the regulation of genes in the region.

The 6p21.1 region associated with nose bridge breadth extends across ~500 KB overlapping the suppressor of Ty 3 homolog (*S. cerevisiae*) (*SUPT3H*) gene and the 5' half of the Runt-related transcription factor 2 (*RUNX2*) gene (Figure 3.20 and Figure 3.25). Strongest association is seen for SNPs in the region of *SUPT3H/RUNX2* overlap, peaking at SNP rs1852985 for both the ordinal and the quantitative assessment of nose bridge breadth (Figure 3.20 and Figure 3.24).

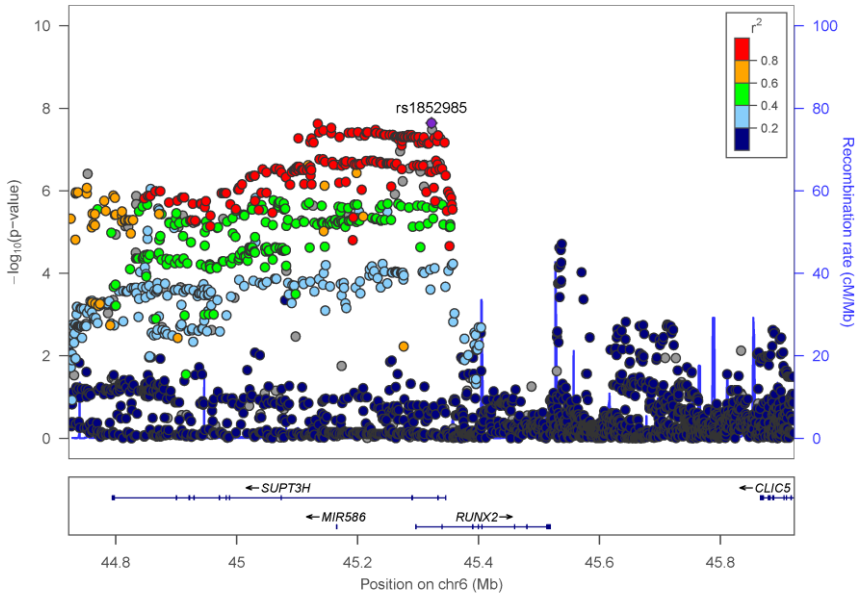


Figure 3.24. Regional association plot for 4q31 and quantitative assessment of nose bridge breadth. This plot was produced in Locus Zoom [310].

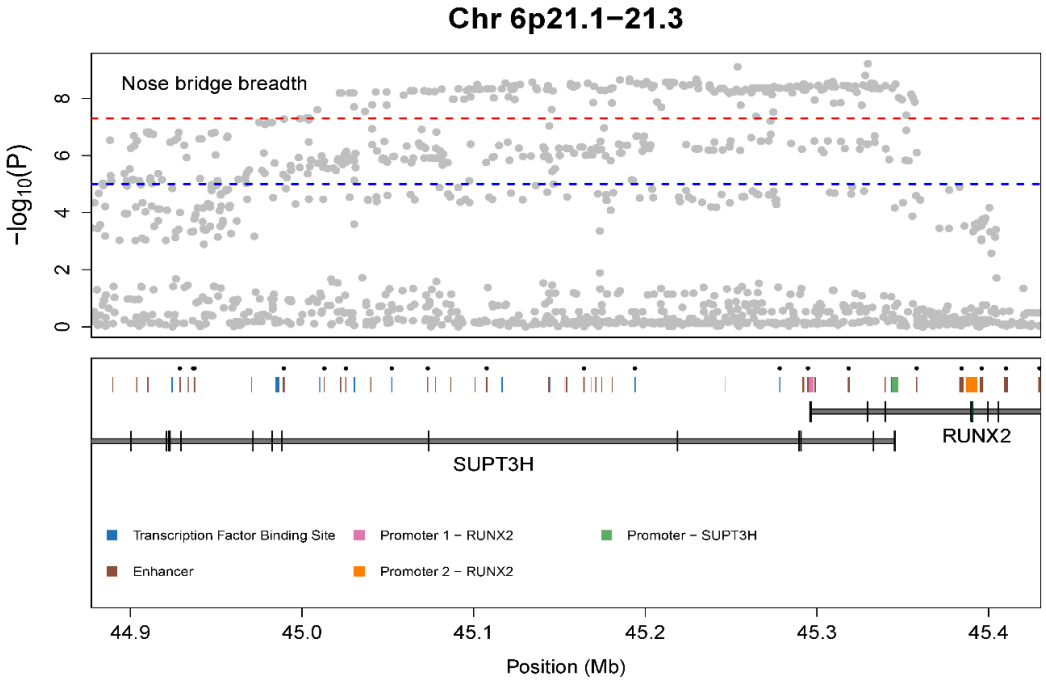


Figure 3.25. SNPs in 6p21 associated with nose bridge breadth and regulatory elements in the *SUPT3H/RUNX2* gene region. The top panel shows $-\log_{10}$ association P values (dashed lines indicate the genome-wide significance threshold of 7.3, and a suggestive threshold of 5). The bottom panel shows the location of exons/introns and of various regulatory elements. Genome annotations were obtained from the Ensembl Genome Bioinformatics database (build GrCh37). Regulatory annotations supported by experimental evidence [332-339] are indicated by black dots.

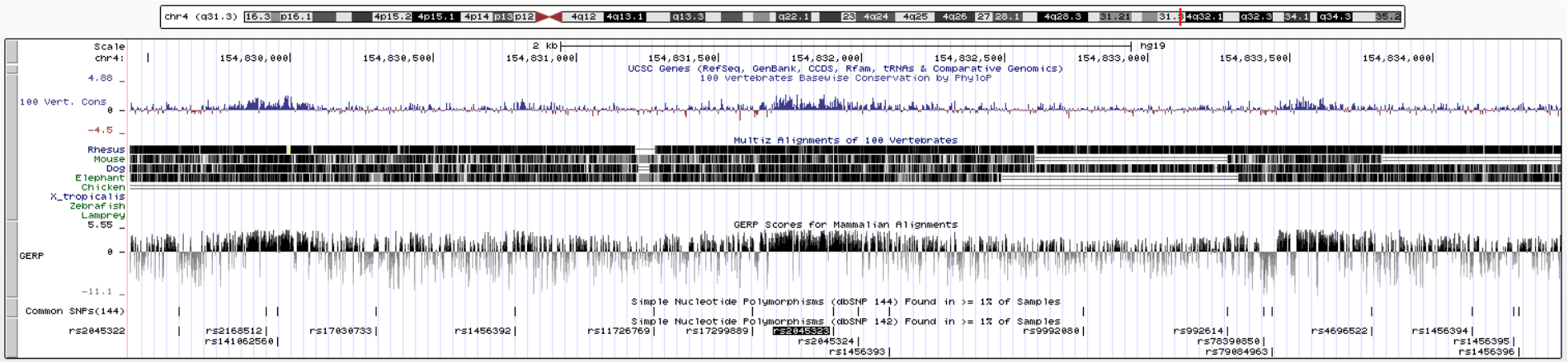


Figure 3.26. Sequence conservation around rs2045323 in the *DCHS2-SFRP2* intergenic region. A UCSC genome-browser screenshot from the (GRCh37/hg19) Assembly for the *DCHS2-SFRP2* intergenic region around rs2045323 (highlighted in black at the bottom), showing enriched values of various conservation scores including GERP.

SNPs in 7p13 region showed strong genome-wide association with nose wing breadth. This region corresponds to GLI Family Zinc Finger 3 gene (*GLI3*; Figure 3.2020 and Figure 3.27), which is composed of fifteen exons [340], within the third exon is positioned the index SNP rs17640804 (Table 3.13 and Table 3.23, Figure 3.12), located in a genomic region with strong evolutionary conservation (Figure 3.28).

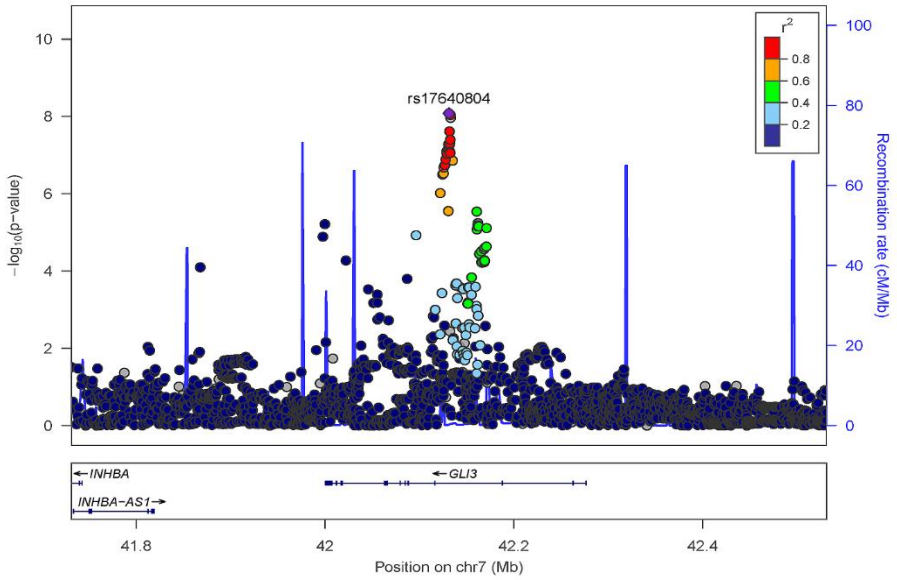


Figure 3.27. Regional association plot for 7p13 and ordinal assessment of nose wing breadth. This plot was produced in Locus Zoom [310].

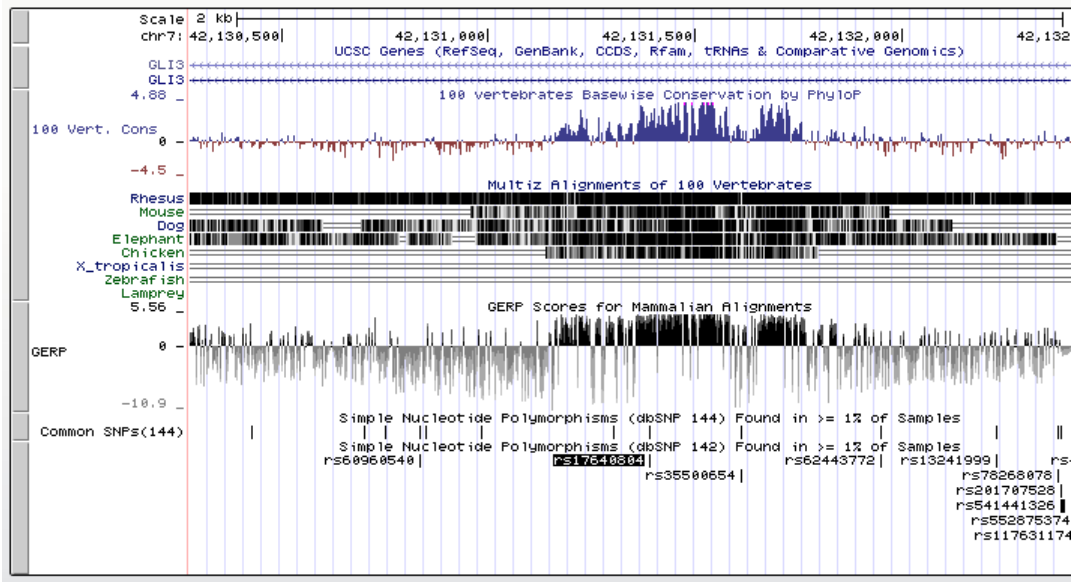


Figure 3.28. Strong sequence conservation around rs17640804 in *GLI3*. A UCSC genome-browser screen shot from the (GRCh37/hg19) Assembly for the *GLI3* region around rs17640804 (highlighted in black at the bottom), showing enriched values of various conservation scores including GERP (Cropped view).

Interestingly, it has been shown experimentally that *Gli3* interacts with *Runx2* in the regulation of mouse osteoblast differentiation [341]. Therefore, statistical interaction between the *GLI3* & *RUNX2* index SNPs was tested on nose bridge breadth and it was found to be significant (*P* value 0.004, Figure 3.29), even though the *GLI3* index SNP by itself does not have a significant effect on nose bridge breadth.

```
Call:
lm(formula = Nose_Bridge_Breadth ~ age + BMI + SEX + PC1 + PC2 +
PC3 + PC4 + PC5 + rs1852985 + rs17640804 + interaction,
data = big)

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)       1.841654   0.063317  29.086  < 2e-16 ***
age              -0.002469   0.001263  -1.955 0.050653 .
BMI               0.010840   0.002157   5.026 5.18e-07 ***
SEX              -0.241279   0.014267 -16.911  < 2e-16 ***
PC1               0.066260   0.578029   0.115 0.908742
PC2              -4.352393   0.547472  -7.950 2.25e-15 ***
PC3              -0.022513   0.522886  -0.043 0.965660
PC4               0.148741   0.542587   0.274 0.783992
PC5              -1.805611   0.533361  -3.385 0.000716 ***
rs1852985         0.066973   0.010853   6.171 7.27e-10 ***
rs17640804        0.016687   0.009929   1.681 0.092898 .
interaction       0.041660   0.014320   2.909 0.003638 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 3.29. R regression output of interaction of *GLI3* with *RUNX2* and nose bridge breadth. The interaction term was highly significant, even though the GLI3 index SNP was not, providing support that the effect of GLI3 on nose bridge breadth is through interaction only.

Strongest association in 20p11 with the ordinal assessment of nose wing breadth was observed for SNP rs927833 located in *LOC100270679*, a long intergenic non-protein coding RNA (LINC01432). There is substantial LD around this SNP and suggestive evidence of association (i.e. *P* values $<10^{-5}$), for SNPs over a region of ~400kb extending to the Paired box gene 1 (*PAX1*; Figure 3.20), a strong candidate gene in this region.

It has also been reported that mouse embryos with *Gli3* null mutations display drastically reduced *Pax1* expression, possibly mediated through *Gli3*’s involvement in the *Shh* signalling pathway [342]. Consistent with these experimental findings, a significant statistical interaction is observed between the *GLI3* and *PAX1* index SNPs on nose wing breadth (*P* value 0.005, Figure 3.30)

```
Call:
lm(formula = Nose_Wing_Breadth ~ age + BMI + SEX + PC1 + PC2 +
PC3 + PC4 + PC5 + rs927833 + rs17640804 + interaction,
data = big)

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)     2.114152   0.067138  31.489  < 2e-16 ***
PC1            -4.761126   0.587589  -8.103 6.49e-16 ***
PC2            -8.774973   0.562619 -15.597  < 2e-16 ***
PC3             0.661149   0.556373   1.188  0.23476
PC4             0.895311   0.576295   1.554  0.12034
PC5            -1.666798   0.564107  -2.955  0.00314 **
SEX            -0.213757   0.015171 -14.090  < 2e-16 ***
age            -0.001481   0.001342  -1.104  0.26976
BMI             0.012685   0.002292   5.535 3.25e-08 ***
rs927833        0.077470   0.012282   6.308 3.04e-10 ***
rs17640804      0.064821   0.010621   6.103 1.11e-09 ***
interaction     0.048300   0.017203   2.808  0.00501 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 3.30. R regression output of interaction of *GLI3* with *PAX1* and nose bridge breadth. The interaction effect between the *GLI3* and *PAX1* index SNPs (rs17640804 and rs927833 respectively) was significant, in addition to their main effects.

3.4 Discussion

A genome-wide scan for 14 facial features phenotyped on an ordinal scale showed associations in four independent loci at 4q31 (columella inclination), 6p21 (nose bridge breadth), 7p13 and 20p11 (nose wing breadth), consisting of common DNA variants associated with normal facial shape phenotypes in ~6,000 Latin-Americans.

False Discovery Rate and Meta-analysis

In addition to the previous results presented here, a False Discovery Rate analysis was performed, and the same associations were found. Combining the results from all the countries a meta-analysis was run, and for all the associations, the significant effects were in the same direction in all countries. A multivariate GWAS was performed to see if correlations between traits could lead to new or stronger associations, but as expected, due to the relatively low correlation between traits, there were no additional associations.

Replication analysis

As follow-up analyses, a replication GWAS in a subsample of 501 Latin-Americans was done. This sample was genotyped and phenotyped (ordinal phenotyping) similarly to the sample from the discovery GWAS, and the associations found were the same. Subsequently, in a subsample of ~3,000 individuals, 9 quantitative traits related to 9 of the ordinal phenotypes were obtained, in addition to the measure of nasion position. Quantitative analyses not only confirmed the ordinal-based associations, but also identified SNPs in 2q12 associated to chin protrusion and replicated the reported association of nasion position with SNPs in *PAX3*, previously found by Paternoster et al. (2012) [8]. Strongest association in 2q12, 4q31, 6p21 and 7p13 was observed for SNPs in the *EDAR*, *DCHS2*, *RUNX2* and *GLI3* genes, respectively. Associated SNPs in 20p11 extend to *PAX1*.

Heritability

The heritability of quantitative traits was calculated and is higher than the heritability of ordinal traits, and more similar to published estimates (Table 1.1). The highest heritability was for nose wing breadth and the lowest heritability was for chin protrusion. In comparison to other studies heritability results varies across traits, upper lip thickness seems to be more similar among different studies (0.2 -0.6), this probably may depend on the phenotype used (Table 1.1).

Association between previously reported SNPs and traits in this study

In order to assess independent evidence of association for the regions implicated in this study, SNPs that were at least suggestive in previously published GWAS were checked [7,8]. The SNP rs2108166, which is in high LD with the SNP rs17640804 (7p13) associated to nose wing breadth, produced a suggestive association with the same trait in Liu et al. (2012) [7]. Also, rs3827760, located in the *EDAR* gene, associated to chin protrusion, was reported in a previous study of a Central Asian population [287].

Association between Cleft-lip and palate and traits in this study

In these analyses, SNPs previously associated to NSCL/P were selected and tested for associations with the facial features assessed here. Few tests passed Bonferroni correction, mostly SNPs associated with quantitative nose-breadth traits (nose root, nose bridge and nose wing breadth). Subsequently, a Kolmorov-Smirnoff test was

significant for both ordinal and quantitative traits and polygenic risk score test combining all 15 index SNPs was significant for the nose breadth traits. Although these few traits related to the nose breadth were significant, similar to the found link between NSCL/P associated SNPs and normal face shape traits, specifically nose width, found by Liu et al. (2012) [7]. More precise evaluation of the impact of NSCL/P associated variants on facial variation in the general population is required.

3.4.1 Candidate regions associated with facial traits

Chin protrusion

Chin protrusion showed significant association with several SNPs in region 2q13, the index SNP was rs3827760 (*EDAR* gene). Strikingly, this variant is a missense mutation (V370A), which was proposed to have been positively selected in East-Asian populations [321]. In addition, it is known that mutations in the *EDA* pathway, which is composed of the *EDA-EDAR-EDARADD* genes, cause Hypohidrotic ectodermal dysplasia (HED). This disease is characterized by several features, especially facial dysmorphia, within these, a more protrusive chin and mandible. By contrast, the V370A mutation acts as a protector allele in people with the disease. Here, it was found that the effect of EDAR370A generates chin less protruded, contrary to HED. Then, to perform functional replication of the effect of *EDAR* seen on chin protrusion, lower jaw protrusion was assessed, as a proxy to chin protrusion in *Edar* mouse mutants. Interestingly, consistent with effect of *EDAR* on chin protrusion, we documented alterations of mandible length in mice with modified *Edar* funtion.

Columella inclination

Single nucleotide polymorphisms in the 4q31 region extended ~400Kb between *DCHS2* and *SFRP2* were associated with nose shape traits, columella inclination, nose protrusion and nose tip angle. *DCHS2* is a calcium-dependent cell-adhesion protein which has recently been shown to participate in a regulatory network controlling cartilage differentiation and polarity during vertebrate craniofacial development [343]. This network includes *SOX9*, a well-known regulator of cartilage differentiation, mutations of which lead in humans to Campomelic Dysplasia (OMIM #114290) a disorder characterized by a range of craniofacial defects. Although *DCHS2* seems the strongest candidate in the 4q31 region, *SFRP2* is also an interesting candidate in that it has been shown to be expressed in osteoblasts, participates in the regulation of *Wnt*

signalling [344] and craniofacial malformations have been reported in *Sfrp2* mutant mice[345].

Nose bridge breadth

The 6p21 region associated with nose bridge width falls in an overlapping region *SUPT3H-RUNX2* genes, by both quantitative and ordinal traits. This region is known to contain key *RUNX2* regulatory elements [338] (Figure 3.25). Rare mutations in *RUNX2* cause Cleidocranial dysplasia (CCD), an autosomal dominant disorder involving alterations of cranial ossification (OMIM #119600). *Runx2* has been shown to participate in the differentiation of mouse osteoblasts, chondrocyte and mesenchymal stem cells and bone development [346], null *Runx2* mutants show a range of chondrocyte proliferation and maturation defects [347]. Interestingly, the length of a functional glutamine/alanine repeat in *RUNX2* has been shown to correlate strongly with the evolution of facial length in dog breeds and, more broadly, in Carnivora [348].

Nose wing breadth

Nose wing breadth showed association with SNPs in the *GLI3* gene, situated in the 7p13 region. Chromatin immunoprecipitation (ChIP) with enhancer-associated protein P300 followed by high-density microarrays have shown that the mouse orthologous gene *Gli3* is overexpressed in forebrain mice embryonic tissue [349]. *GLI3* is known to act both as activator and repressor in the sonic hedgehog signalling pathway (*SHH*), a key regulatory pathway of chondrocyte differentiation [350]. Blockade of *SHH* signalling in the forebrain leads to a narrow and truncated face [351,352] . The index SNP rs17640804 is, according to ChIP-Seq analysis from the Roadmap Epigenomics Consortium [333,353], an active enhancer in Osteoblast primary cells, and is conserved according to both GERP and SiPhy conservation scores (Figure 3.28).

Mutations in *GLI3* have been shown to cause several Mendelian disorders associated with craniofacial and limb abnormalities, including Greig cephalopolysyndactyly syndrome (GCPS), Pallister-Hall syndrome, preaxial polydactyly type IV, and postaxial polydactyly types A1 and B. GCPS is characterized by several craniofacial abnormalities including macrocephaly, scaphocephaly (and narrow cranial vault) and a broad nose [354]. A mouse null *Gli3* mutant has been reported to show a range of craniofacial abnormalities, including a wider nose [355]. A mutant mouse ($Gli3^{Xt\text{-}J/Xt\text{-}J}$) with an intragenic deletion has been reported to show polydactyly and a range of

craniofacial abnormalities [356,357]. Hurst et al. (2011) described mice with deletions in *Gli3* displaying several malformations in the skull, face and body; in particular, two mice with deletions in our region of interest show broad nasal base as a characteristic [358]. Deletions in the genomic region (7p13) where this SNP is located, provoke a large range of changes in facial morphology, similar to other enhancers implicated in craniofacial morphology [359].

Strong association was detected between the ordinal trait nose wing breadth and the marker rs92783, situated in *LOC100270679* gene. This SNP is in LD with genetic markers in an extended region close to *PAX1*. This gene is a key developmental transcription factor which has been shown experimentally to affect chondrocyte differentiation through its participation in a regulatory pathway that also includes *RUNX2* and *SOX9* [360]. More broadly, a *Pax-Six-Eya-Dach* (*Dachshund*) network (PSEDN), involving protein-protein and protein-DNA interactions impacting on a range of basic developmental processes has been described [361]. As indicated above, another *PAX* gene (*PAX3*) has been twice reported to impact on nasion position [7,8] and we replicated that association here. A missense mutation in *PAX1* has been shown to cause autosomal recessive oto-facio-cervical syndrome, a disorder characterized by various skeletal and facial abnormalities [362]. A significant statistical interaction between *GLI3* and *PAX1* index SNPs on nose wing breadth was shown here. Possibly due to *GLI3* reducing the expression of *PAX1* because it is part of *Shh* signalling pathway.

3.4.2 Replications of associations between genes and facial features reported in this study

As a result of these analyses a scientific publication was published in 2016. Before that, there were only 3 GWAS on facial features [7,8,277]. In our genome-wide scans on facial traits we found new genetics variants associated with facial features and, we replicated the SNP rs7559271 associated with nasion position, previously reported by Paternoster et al. (2012) [8]. This SNPs is situated in the gene *PAX3*. The same association, between *PAX3* gene and nasion position was detected by Liu et al. (2012), earlier the same year, but with a different genetic variant [7]. In 2018, Claes et al. found a different SNP in the same gene associated to nasion position [275].

Subsequently, three genes detected by our GWAS on facial features were replicated. A SNP in *DCHS2* gene was associated with the area around the nose wings [275], we found the same gene, but not the same marker associated with columella inclination. We reported a new association between nose wing breadth and the *PAX1* gene, later in the same year, Shaffer et al. (2016) found an association between another SNP in the same gene and nasal width [278].

We detected an association between nose bridge breadth and an extensive area in the *SUPT3H-RUNX2* overlapping region. Interestingly, Claes et al. (2018) detected a SNP in the *SUPT3H* gene associated to nose width [275].

3.4.3 The advantage of genetically diverse populations

An interesting outcome of these analysis was to observe how physical appearance is affected by the genetic diversity. As an example of this, I will present a trait analysed in this thesis, Nose protrusion. The heritability of this trait is 84%, when this trait is observed through the CANDELA samples, a clear difference is seen between people with high European ancestry, they have a greater protrusion of the nose in comparison to people with high native American component [95] (Figure 3.31A). The variation of the trait is partially explained by genetic Principal components (continental and sub-continental ancestry), basic covariates (sex, age, etc) and around 1% by a SNP (rs2045323), which shows significant association in the *DCHS2* gene region (Section 3.3.8)[95]. The difference in allele frequencies between European and Native Americans is around 57% (Figure 3.31B).

These genetically diverse populations (i.e. Latin Americans) present an advantage compared to less diverse populations (i.e. Europeans) [363, 364, 365]. Thus, the mixed ancestry of Latin Americans offers the opportunity to find novel trait loci that could not be found in the un-admixed parental groups of Latin American populations by increasing the power of detection in association analyses due to this difference between allele frequencies between parental groups. In addition, this is helping to understand the genetic architecture of complex traits [203].

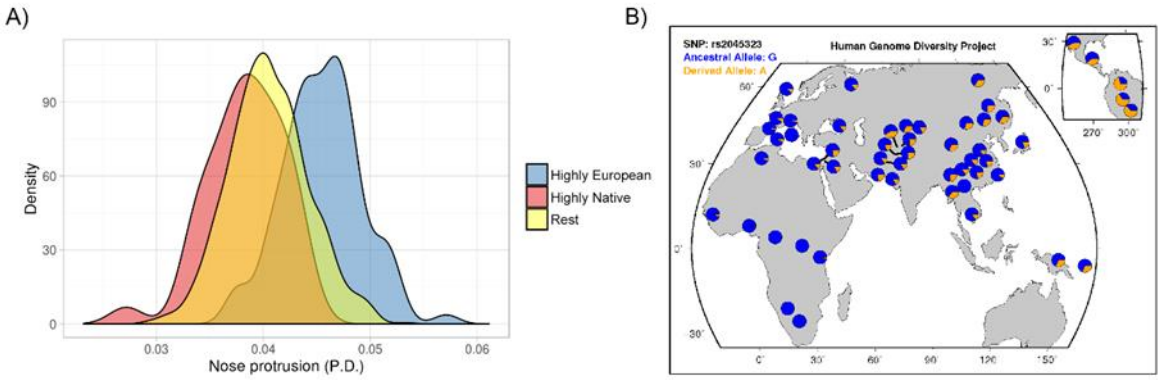


Figure 3.31. Distribution of nose protrusion in the CANDELA sample. A) Density plots of nose protrusion for individuals from CANDELA sample. B) Allele frequencies in the CEPH-HGDP population panel for the index SNP (rs2045323). Adapted from Adhikari et al. (2016) [203].

3.4.4 Limitations

Despite several new genetic variants associated with facial traits being found in this study, it has also showed some limitations. First, the GWAS with facial features categorized in an ordinal way, lacking information within categories. Later, the GWAS performed with quantitative traits demonstrated how the lack of information in the phenotypes decrease the power of detection because the associations were stronger, but there are still limitations, i.e. 34 landmarks are not enough to represent all the details of the facial traits well. This is partly because the software Photomodeler, used to transform the 2D images to 3D images cannot totally capture the three dimensions of the facial traits, also because it depends on the expertise of the operator to add the landmarks to the photos, another factor is the 2D photographs , unavoidably they present differences in face orientation and different images sizes also affect the phenotyping process and this cannot be improved.

This work also showed the importance of performing genome-wide scans in different populations. Either to replicate these findings or to find new genetic variants that could not be present in the parental populations of the Latin-American people. Future work should be done to improve the phenotyping method and to increase the diversity of the populations assessed in these kinds of studies.

3.5 Summary

In this chapter I reported both novel variants associated with facial features and the replication of genetic variants previously associated to face traits in a Latin-American population. The results presented in this thesis showed the complex genetic architecture behind the facial traits and is the importance of the phenotyping method used to characterize the traits, as evidenced by the associations using quantitative phenotypes that showed greater statistical power than the ordinal phenotypes. Further, some of the reported genomic regions associated with facial traits are potential candidate genes that should be followed up by functional analyses.

Chapter 4: Genome-Wide Association Study of dental morphology in the CANDELA Cohort

4.1 Overview

Teeth are the hardest and well-preserved parts of the body; they resist high temperatures, decomposition of the body and their disintegration occurs after that of bones. Therefore, dental pieces are of great importance, and they represent the last remnants of humans and animals after death. Dental features have been widely used by anthropologists and archaeologists to establish biological relatedness among past and current human populations, they have also been used by the forensic sciences in incidents such as fires, terrorist attacks, natural disasters, airplane crashes, train and car accidents, etc. to identify human remains [176].

Dentition has long been analyzed from different locations, because it is unique and offers an unparalleled opportunity for a better understanding of the origin, evolution, and phylogenesis of vertebrates, including humans.

The aim of this chapter is to contribute to the understanding of the genetic basis of normal dental variation, even though some genes have previously been identified for various dental traits, it is still poorly understood. In this chapter, through GWAS of different dental phenotypes (ordinal and quantitative) performed in a subset of volunteers from Colombia (CANDELA Cohort [59]), I obtained a group of possible genes associated with these phenotypes. Subsequently, I assessed the local ancestry of the haplotypes where the candidate SNPs are situated, using RFMix modeling approach [366] to verify if the haplotyes correspond to the population where the allele frequency of the candidate markers was higher. Finally, I described the candidate genomic regions and discussed them comparing with results of previous GWAS and linkage studies.

4.1.1 Previous studies

To the best of my knowledge there are several GWAS related to teeth, summarised in Table 4.1; the main objective of these studies has been the development and disease of teeth. Pillas et al. in (2010) and Fatemifar et al. (2013), performed a similar GWAS of teeth eruption time and number of teeth on the same populations of Northern Finland Birth Cohorts (NFBC) 1966 and Avon Longitudinal Study of Parents and Children England (ALSPAC) [9, 367]. The main difference between the studies was the imputation and the number of samples, which was bigger in 2013. Both studies found significant associations in similar genes, i.e. the *IGF2BP1* gene associted to number of teeth in both studies. This gene is part of the Hedgedog family, and it is well known that this pathway in involved in tooth development [141]. Interestingly, Geller et al. (2011) found an association between the gene *ADK* and the permanent eruption of the teeth between age 6 and 14 years old in women from the Danish National Birth Cohort (DNBC) [10]. This was also replicated by Fatemifar et al. (2013); *ADK* was also associated with facial width distances [367]. Furthermore, there have been at least three GWAS seeking genes involved in the appearance of dental caries. The findings of these studies do not overlap among each other, perhaps because they were applied to different ethnic groups [368-370]. The *ADAMTS3* gene, situated in chromosome 4, was associated with decay, missing and filled surfaces of the teeth in a cohort from the U.S., which did not include people of Hispanic origin. This gene is highly expressed during tooth development in the dental papilla in mice [371].

Chr	Gene	SNP	Phenotype	p-value	Reference	Population
17	*KCNJ2*	**rs8079702**	Teeth eruption time	3.77E-22	Pillas et al., 2010	NFBC 1966 and ALSPAC
X	*EDA*	**rs4844096**	Teeth eruption time	2.61E-08	Pillas et al., 2010	NFBC 1966 and ALSPAC
X	*EDA*	**rs5936487**	Teeth eruption time	6.18E-11	Pillas et al., 2010	NFBC 1966 and ALSPAC
12	*MSRB3*	rs10506525	Teeth eruption time	6.17E-09	Pillas et al., 2010	NFBC 1966 and ALSPAC
17	*IGF2BP1*	rs9674544	Number of teeth	1.56E-08	Pillas et al., 2010	NFBC 1966 and ALSPAC
14	*RAD51L1*	rs1956529	Number of teeth	2.60E-08	Pillas et al., 2010	NFBC 1966 and ALSPAC
17	*KCNJ2*	**rs8079702**	Number of teeth	1.25E-14	Pillas et al., 2010	NFBC 1966 and ALSPAC
X	*EDA*	**rs4844096**	Number of teeth	4.58E-11	Pillas et al., 2010	NFBC 1966 and ALSPAC
X	*EDA*	**rs5936487**	Number of teeth	3.37E-10	Pillas et al., 2010	NFBC 1966 and ALSPAC
6	*6q21*	rs6568401	Age at first tooth	1.50E-10	Fatemifar et al., 2013	NFBC 1966 and ALSPAC
10	*ADK VCL AP3M1*	**rs7924176**	Age at first tooth	1.80E-08	Fatemifar et al., 2013	NFBC 1966 and ALSPAC
11	*CDON*	rs4937076	Age at first tooth	4.00E-08	Fatemifar et al., 2013	NFBC 1966 and ALSPAC
12	*MSRB3*	**rs12229918**	Age at first tooth	7.30E-14	Fatemifar et al., 2013	NFBC 1966 and ALSPAC
12	*HMGA2*	**rs17101923**	Age at first tooth	6.30E-11	Fatemifar et al., 2013	NFBC 1966 and ALSPAC
14	*BMP4*	rs17563	Age at first tooth	9.10E-17	Fatemifar et al., 2013	NFBC 1966 and ALSPAC
17	*IGF2BP1*	**rs1994969**	Age at first tooth	2.30E-14	Fatemifar et al., 2013	NFBC 1966 and ALSPAC
17	*TEX14/RAD51C*	rs412000	Age at first tooth	1.70E-09	Fatemifar et al., 2013	NFBC 1966 and ALSPAC
17	*KCNJ2 KCNJ16*	**rs8080944**	Age at first tooth	7.60E-34	Fatemifar et al., 2013	NFBC 1966 and ALSPAC
X	*FAM155E - EDA*	**rs11796357**	Age at first tooth	3.10E-22	Fatemifar et al., 2013	NFBC 1966 and ALSPAC
2	*2q35*	rs10932688	Number of teeth	2.50E-08	Fatemifar et al., 2013	NFBC 1966 and ALSPAC
7	*CALU/OPN1SW*	rs1799922	Number of teeth	4.00E-09	Fatemifar et al., 2013	NFBC 1966 and ALSPAC
10	*CACNB2*	rs10740993	Number of teeth	1.70E-09	Fatemifar et al., 2013	NFBC 1966 and ALSPAC
10	*ADK VCL AP3M1*	**rs7924176**	Number of teeth	7.80E-16	Fatemifar et al., 2013	NFBC 1966 and ALSPAC
12	*MSRB3*	**rs12229918**	Number of teeth	2.30E-09	Fatemifar et al., 2013	NFBC 1966 and ALSPAC
12	*HMGA2*	**rs17101923**	Number of teeth	1.10E-10	Fatemifar et al., 2013	NFBC 1966 and ALSPAC
13	*DLEU7*	rs9316505	Number of teeth	3.40E-08	Fatemifar et al., 2013	NFBC 1966 and ALSPAC
14	*AJUBA/C14orf93*	rs997154	Number of teeth	2.60E-08	Fatemifar et al., 2013	NFBC 1966 and ALSPAC
17	*IGF2BP1*	**rs1994969**	Number of teeth	7.20E-16	Fatemifar et al., 2013	NFBC 1966 and ALSPAC
17	*KCNJ2 KCNJ16*	**rs8080944**	Number of teeth	1.50E-19	Fatemifar et al., 2013	NFBC 1966 and ALSPAC
X	*FAM155E - EDA*	**rs11796357**	Number of teeth	6.90E-19	Fatemifar et al., 2013	NFBC 1966 and ALSPAC
1	*ASCL5/CACNA1S*	rs4498834	Tooth agenesis	2.90E-14	Jonsson et al., 2018	Icelandic discovery sample (n=1944)
3	*FOXI3*	rs35822372	Tooth agenesis	3.40E-13	Jonsson et al., 2018	Icelandic discovery sample (n=1944)
2	*EDAR*	chr2: 108,896,996	Tooth agenesis	5.90E-09	Jonsson et al., 2018	Icelandic discovery sample (n=1944)
2	*ARHGAP15*	rs2034604	Tooth agenesis	2.10E-14	Jonsson et al., 2018	Icelandic discovery sample (n=1944)
2	*WNT10A*	rs121908120	Tooth agenesis	6.10E-40	Jonsson et al., 2018	Icelandic discovery sample (n=1944)
2	*WNT10A*	rs121908119	Tooth agenesis	4.90E-08	Jonsson et al., 2018	Icelandic discovery sample (n=1944)

Continue…

Chr	Gene	SNP	Phenotype	p-value	Reference	Population
8	*ZFHX4*	rs371555610	Tooth agenesis	4.40E-11	Jonsson et al., 2018	Icelandic discovery sample (n=1944)
4	*LEF1*	rs917412	Mandibular second premolars	2.50E-10	Jonsson et al., 2018	Icelandic discovery sample (n=1196)
17	*NOL11*	rs758468472	Maxillary second premolars	3.30E-08	Jonsson et al., 2018	Icelandic discovery sample (n=600)
3	*FOXP1*	rs35956082	Maxillary lateral incisors	1.10E-11	Jonsson et al., 2018	Icelandic discovery sample (n=600)
X	*EDA*	rs55846652	Maxillary lateral incisors	6.70E-11	Jonsson et al., 2018	Icelandic discovery sample (n=600)
2	*PXDN, MYT1L*	rs11681214	Severe erosive tooth wear	3.99E–08	Alaraudanjok et al., 2019	Northern Finland Birth Cohorts 1966 (n=1944)
1	*AJAP1*	rs3896439	Caries in premolars, canines and maxillary mid-dentition	2.00E-08	Shaffer et al., 2013	920 self-reported white participants
10	*LYZL2*	rs399593	Caries in mandibular anterior (incisors, canines, first pre-molar)	9.00E-09	Shaffer et al., 2013	920 self-reported white participants
12	*HMGA2*	rs12424086	Permanent teeth erupted between age 6 and 14 years	1.11E-08	Geller et al., 2011	5099 women from DNBC (Danish National Birth Cohort)
2	*TNP1*	rs4491709	Permanent teeth erupted between age 6 and 14 years	6.51E-10	Geller et al., 2011	5088 women from DNBC (Danish National Birth Cohort)
1	*CACNA1S*	rs2281845	Permanent teeth erupted between age 6 and 14 years	1.31E-10	Geller et al., 2011	5097 women from DNBC (Danish National Birth Cohort)
10	*ADK*	**rs7924176**	Permanent teeth erupted between age 6 and 14 years	6.58E-09	Geller et al., 2011	5100 women from DNBC (Danish National Birth Cohort)
7	*SYPL1, NAMPT*	**rs190395159**	Decay, missing and filled teeth	7.14E−9	Morrison et al., 2016	Cuban, Dominican, Mexican, Puerto Rican, Central and South American (n=11754)
20	*BMP7, MIR4325, SPO11*	rs72626594	Decay, missing and filled teeth	2.75E−8	Morrison et al., 2016	Cuban, Dominican, Mexican, Puerto Rican, Central and South American (n=11754)
3	*IGSF10, MIR5186, MIR548H2, AADACL2*	rs138769355	Decay, missing and filled surfaces	3.59E−8	Morrison et al., 2016	Cuban, Dominican, Mexican, Puerto Rican, Central and South American (n=11754)
7	*SYPL1, NAMPT*	**rs190395159**	Decay, missing and filled surfaces	5.97E−10	Morrison et al., 2016	Cuban, Dominican, Mexican, Puerto Rican, Central and South American (n=11754)
1	*GNG4, LYST, B3GALNT2, TBCE, GGPS1, ARIB4D*	rs138642966	Decay, missing and filled surfaces	1.94E−8	Morrison et al., 2016	Mexican (n=4578)
10	*ANK3, CDK1, RHOBTB1*	rs116717469	Decay, missing and filled surfaces	3.23E−8	Morrison et al., 2016	Mexican (n=4578)
17	*CACNA1G, ABCC3, ANKRD40, LUC7L3, MIR8059, WFIKKN2, TOB1, SPAG9*	rs71381322	Decay, missing and filled surfaces	3.72E−8	Morrison et al., 2016	Mexican (n=4578)
18	*None*	rs16946661	Decay, missing and filled surfaces	4.02E−8	Morrison et al., 2016	Mexican (n=4578)

Continue…

Chr	Gene	SNP	Phenotype	p-value	Reference	Population
X	*ACOT9, PRDX4, SAT1, APOO*	rs141563584	Decay, missing and filled teeth	3.89E−8	Morrison et al., 2016	Puerto Rican (2006)
1	*RHOU*	rs9793739	Decay, missing and filled surfaces, proportion of decay filled surfaces and caries severity	5.27E-07	Wang et al., 2012	~7000 participants non-Hispanic whites
4	*ADAMTS3*	rs1383934	Decay, missing and filled surfaces, proportion of decay filled surfaces and caries severity	2.96E-07	Wang et al., 2012	~7000 participants non-Hispanic whites
6	*RPS6K2*	rs635808	Decay, missing and filled surfaces, proportion of decay filled surfaces and caries severity	1.06E-07	Wang et al., 2012	~7000 participants non-Hispanic whites
8	*PTK2B*	rs17057381	Decay, missing and filled surfaces, proportion of decay filled surfaces and caries severity	4.02E-07	Wang et al., 2012	~7000 participants non-Hispanic whites
14	*CNIH*	rs4251631	Decay, missing and filled surfaces, proportion of decay filled surfaces and caries severity	2.13E-07	Wang et al., 2012	~7000 participants non-Hispanic whites

Table 4.1. Genes regions that have been associated with tooth development and disease. This summary corresponds to GWAS testing for associations with dental phenotypes, such as number of teeth, eruption time, agenesis and presence of caries. SNPs replicated in different studies and different traits are highlighted in bold.

As far as I know, there are no GWAS analyzing dental morphology.

4.1.2 The Arizona State University Dental Anthropology System (ASUDAS)

The Turner-Scott Dental Anthropology System dental plaques are a series of 24 reference plaques that highlight human tooth morphology and variation. Created by Drs. Christy G. Turner II, Christian R. Nichol, and G. Richard Scott. This standardized collection of plaques showcases non-metric tooth crown and root traits that are found in a given human population, as well as showing the degree of expression of each trait [372]. This is a widely used non-metric method to categorize the different dental features assessed in this study to perform the GWAS. Consequently, a complete description of each phenotype assessed in this thesis is presented below.

4.1.2.1 Development of ASUDAS

The observation of teeth morphology has been widely used by physical anthropologists, frequently as the presence or absence of certain features or the description of the whole tooth. But Ales Hrdlicka in his work on shoveling shape from 1920 [373], established a difference by noticing and highlighting only the shovel shape

feature of the tooth. He realized that when the trait was present, different forms and grades of expression could be observed. Later, in 1940, Dahlberg starts making the first effort to standardize this method for the observation of dental morphology. In order to achieve this idea, he created the plaster plaque system, with plaques depicting different teeth characteristics and different grades of expression of the trait. Subsequently, he distributed them among his colleagues. The interest in developing a standardized and replicable method did not stop there. The Dental Anthropology Laboratory of Arizona State University continued working on the development of this method, and the result is the Arizona State University Dental Anthropology System [372].

The main goals of this system are to go beyond the dichotomy between the presence or absence of a feature in the teeth, and to be able to replicate the results among observers. The plaques devised show a physical representation of a scale, from the minimal to the maximal expression of a trait, and different grades in between. Not all the traits described in the literature were depicted on the plaques. Only the features that were most easily and reliably observed, with low sex dimorphism, and the more resistant features to the permanent and aggressive daily use of the teeth. Finally, this method is relatively easy to apply, it provides useful information within a short time and at low cost [372]. Consequently, I will present the traits assessed using the ASUDAS method [372] in this thesis.

4.1.2.2 Description of the dental features assessed using ASUDAS

Human dentition has 32 teeth, 8 incisors, 4 canines, 8 premolars and 12 molars [111] (Figure 1.6). The anatomy and general description of these teeth has been explained previously (Section 1.3.1.2). In this section I explain each feature examined using ASUDAS scoring method [111] on each tooth. The list is organized by type of tooth, incisors, canines, premolars and molars and next to each trait and between brackets is the acronym used in this thesis. Most of the descriptions belong to Scott and Turner (1997) and Scott and Irish (2017), the traits that were not described in this book have the reference attached to the tittle.

Incisors:

Shoveling (SS-U, SS-U2, SS-L1 and SS-L2).

This trait is observed in upper incisors, canine, and lower incisors. The lingual surface of these teeth is scooped by the presence of marginal ridges (Figure 4.1). The scale of this trait goes from 0 to 7.

0. *None:* Lingual surface is flat.

1. *Faint*: Elevations of mesial and distal aspects of lingual surface are barely seen and palpated.

2. *Trace*: Elevations are easily seen.

3. *Semishovel:* The ridges are strongly pronounced and there is a tendency to converge toward the cingulum.

4. *Semishovel:* The characteristics of the trait are stronger than in grade 3.

5. *Shovel:* Ridges are highly developed, and they almost contact at the cingulum.

6. *Marked Shovel:* Strongest development. In some cases, the ridges are in contact with cingulum.

7. *Barrel:* Shovelling is more pronounced than grade 6. It is considered as barrel-shaped. This should not be confused with a hypertrophied tuberculum dentale.

In this sample, shovelling was assessed in upper and lower central and lateral incisors.

Comments: In this thesis, this trait was assessed only in upper and lower incisors, not in canines [111].

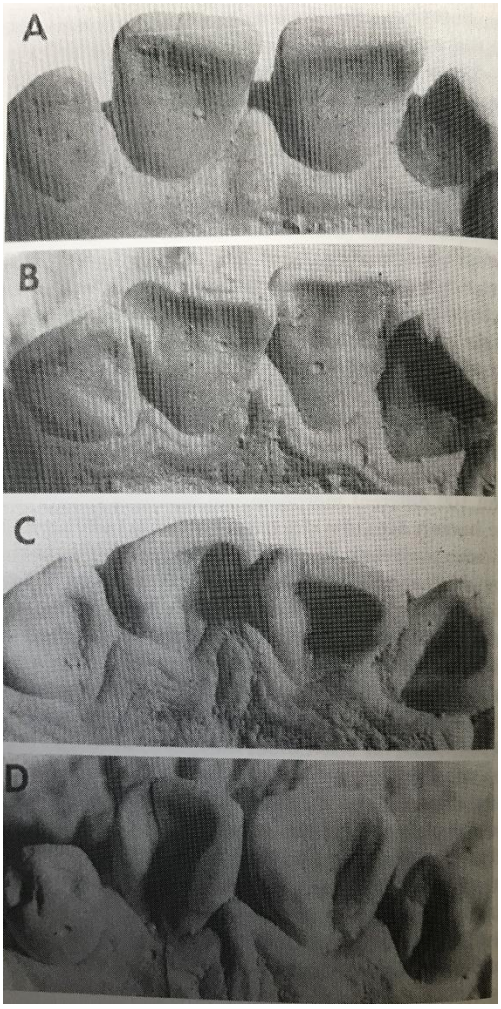


Figure 4.1. Shovel-shape incisors. Different levels of expression across the trait. A and B almost no expression and C and D exhibit pronounced shoveling. Figure from Scott and Turner (1997) [111].

Winging (WIN-U1)

Winging corresponds to the rotation of the upper central incisors. This procedure was modified by Turner in 1970, and it does not use a reference plaque (Figure 4.2). It goes from 1 to 4.

1. *Bilateral winging:* from incisal view central incisors are rotated mesiolingually, shaping a V between both central incisors. When the angle is more than 20 degrees, it is classified as 1A; when it is less than 20 degrees, 1B.

2. *Unilateral winging:* Only one incisor is rotated.

3. *Straight:* Both central incisors follow the arcade curvature or form a straight labial surface.

4. *Counter-winging:* One or both central incisors are rotated distolingually.

Comments: this trait is only evaluated in upper central incisors [111].

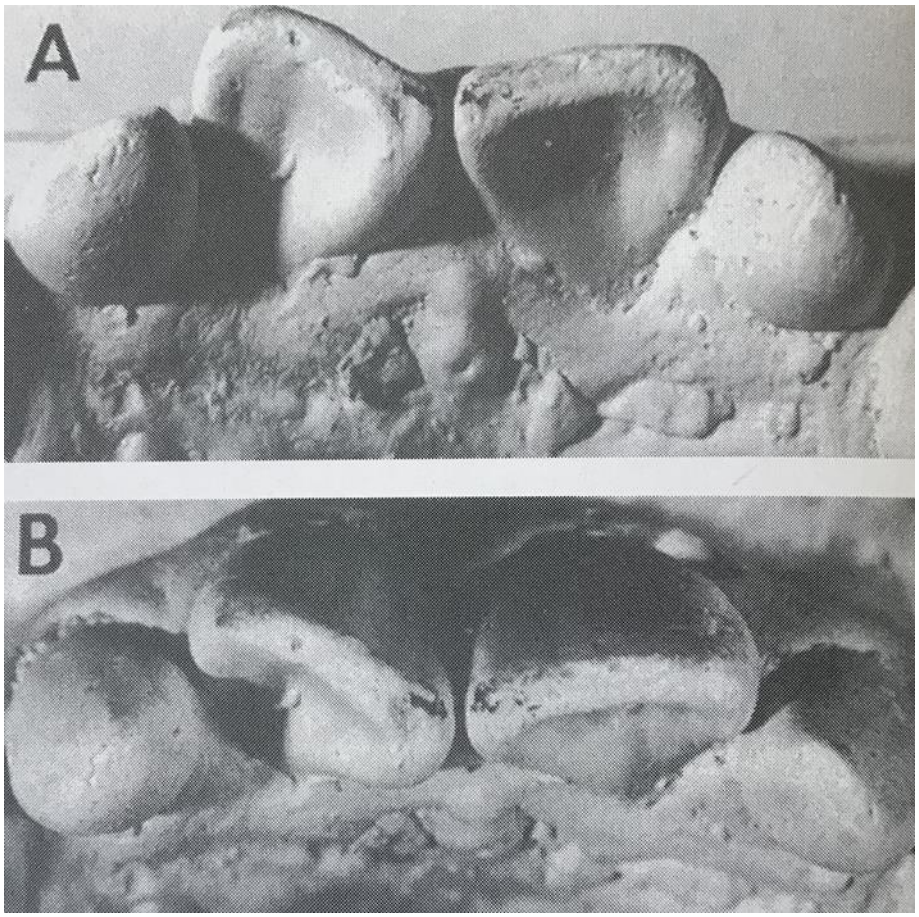


Figure 4.2. Winging upper central incisors shown from lingual (A) and occlusal (B) views. Figure from Scott and Turner (1997) [111].

Labial Convexity (LC-U1 and LC-U2) [111]

This trait is observed in the medial two-thirds of the upper incisors. From occlusal view, the labial surface goes from completely flat to showing a strong degree of convexity. Scoring: [111]

0. Labial surface is plane.

1. Labial surface shows trace convexity.

2. Labial surface shows weak convexity.

3. Labial surface shows moderate convexity.

4. Labial surface shows pronounced convexity.

Double-shovelling (DS-UI1, DS-UI2, DS-LI1 and DS-LI2)

Following Turner's (1991) scoring criteria this trait is observed in the upper and lower incisors, canine and premolar [372]. However, in Scott and Irish's (2017) scoring procedure, this trait is only evaluated in central and lateral upper incisors. In this thesis I followed Scott and Irish's 2017 procedure [374].

This trait refers to mesial and distal marginal ridges on the labial surface of the incisors (Figure 4.3). The scale of this trait goes from 0 to 6 [372].

0. *Absence*: Labial Surface is smooth.

1. *Faint:* In presence of strong contrasting light, the mesial and distal ridging can be observed. In some cases, just mesial ridges can be seen in this and stronger grades.

2. *Trace:* Ridge can be easily seen and palpated in comparison with faint expression but still slight.

3. *Slight:* Ridges distinct enough to be readily palpated.

4. *Moderate:* Ridging is pronounced along at least one half of crown height.

5. *Pronounced*: Ridging is very prominent and may occur from incisal edge to crown root junction.

6.*Very pronounced*: Extreme double shovel with well-develop ridges along both mesial and distal labial margins [374].

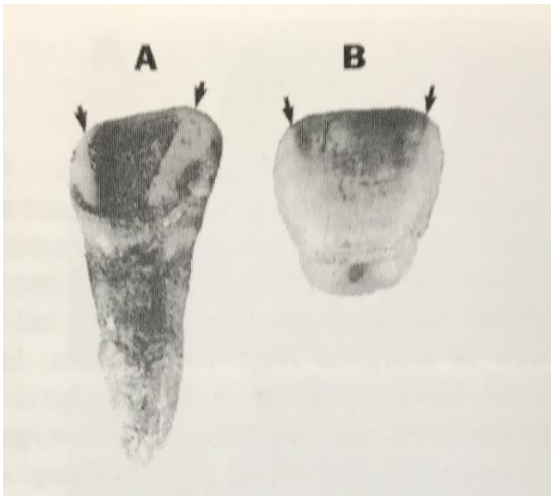


Figure 4.3. Double shoveling of upper central incisors. (A) Well-developed mesial and distal labial marginal ridges. (B) Weakly developed labial and marginal ridges. Figure from Scott and Turner (1997) [111].

Interruption Grooves (IG-UI1 and IG-UI2)

This feature is assessed in central and lateral upper incisors. Grooves can be observed crossing the cingulum, and often extend onto the root. They are different to most of the nonmetric dental traits. In some cases, the groove is restricted to one or more marginal ridges and it does not reach the root. These grooves can be present in different parts of the tooth crown and/or root, the morphogenesis is still not understood, and their expression has still to be classified in the same manner as other traits where presence varies from slight to pronounced (Figure 4.4). Classification:

0. Absence of grooves on lingual marginal ridges and basal cingula.

M. Groove on mesiolingual marginal ridge.

D. An interruption groove occurs on the distolingual border.

MD. Grooves in both mesio- and distolingual marginal ridges.

Med. A groove occur in the medial area of the cingulum.

Comments: In this thesis, this trait was only scored as absent or present, 0 = absent and 1 = present (any other category M, D, MD and Med) [111].

Figure 4.4. Interruption grooves of upper lateral incisors. Both upper lateral incisors express the trait. Figure from Scott and Turner (1997) [111].

Tuberculum Dentale (TD)

This trait is observed in upper incisors and canine. It corresponds to cingular projections on the lingual surface of the upper anterior teeth. This feature usually takes the form of ridges in the lingual surface, referred to as mediolingual ridges, or various degrees of expression of a cusp (Figure 4.5). The expression varies among the three anterior teeth, their score criteria are described separately [111]:

Upper central incisors (TD-UP1) [111]

TD is expressed in the form of ridges; they vary in size and number. The rank of expression goes from 0 to 6:

0: *No expression*. No ridges on basal eminence. Cingular region of the lingual surface is smooth.

1. *Faint ridging*. One slight ridge.

2. *Trace ridging*. One moderate ridge or two slight ridges.

3. *Strong ridging*. One moderate ridge and one slight ridge, or one pronounced ridge.

4. *Pronounced ridging*. Two moderate ridges, or one slight and one pronounced ridge.

5. One moderate ridge and one pronounced ridge.

6. Two pronounced ridges, or three slight and/or moderate ridges.

Upper lateral incisors (TD-UI2) [111]

These teeth vary in size, morphology, and form. The lowest manifestation of this feature on UI2 takes the shape of ridges, like those expressed in UI1, but the next level of expression seems more like tubercles. There is a thin line between a ridge and tubercle, a ridge is an elevation where the two sides are roughly parallel and as it advances blends into the lingual surface as it extends toward the occlusal margin. Tubercle is typically rounded and distinct (with or without a free apex) and does not gradually merge into the lingual surface. Score criteria:

0. Absence of ridge or tubercle formation.

1. Slight ridge.

2. Moderate ridge.

3. Small tubercle.

4. Moderate tubercle.

5. Large or double tubercle.

6. Welt on basal cingulum.

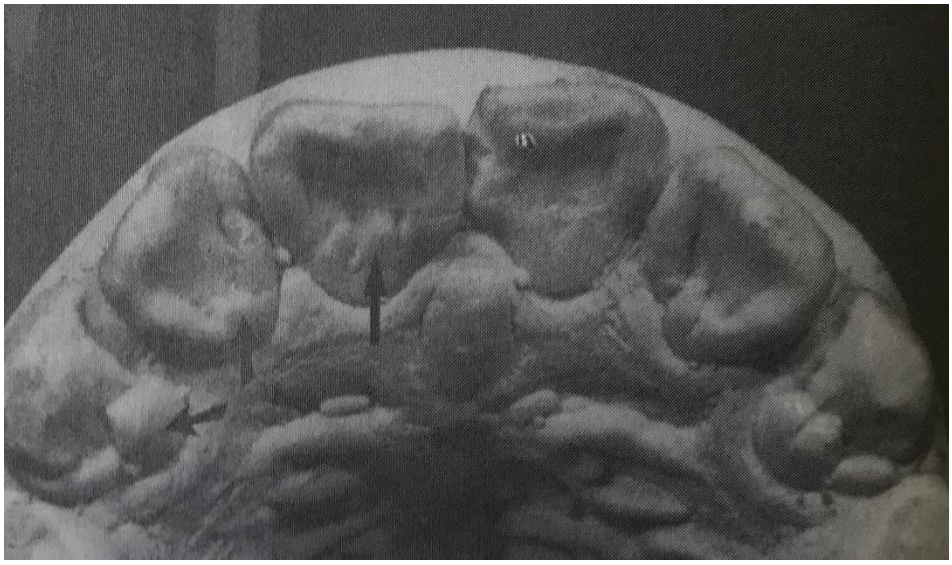


Figure 4.5. Tuberculum dentale in upper incisors and canines. The arrows show three levels of expression, from moderate to pronounced tuberculum projections on the right upper central incisor. Figure from Scott and Turner (1997) [111].

Peg-shaped Incisor (PS-UI2)

Upper lateral incisor. This tooth is very small and lacking the normal crown morphology, instead it has peg form [111]. Classification:

0. Normal sized incisor.

1. Incisor reduced in size, normal crown shape.

2. Peg-shaped incisor.

Midline Diastema (DIAS-U1) *

This corresponds to the space between the upper central incisors. Presence or absence is determined by a measurement taken between the upper central incisors midway, between the cervical part and the (unworn) incisal edge of the tooth [111]. Classification:

0. No diastema (space <0.5 mm)

1. Diastema (space >0.5 mm)

Canines:

Upper canine (TD-UC)

In upper canines, TD is mostly tubercles. In some cases, ridges can be found in upper canines, but it typically takes the form of a tubercle ranging from slightly to pronounced (Figure 4.5) [111].

0. Absence of ridge or tubercle formation.

1. Very slight tubercle, distinguished by a single groove.

2. Slight tubercle, outlined by two grooves.

3. Moderate tubercle.

4. Medium tubercle with no free apex.

5. Large tubercle with free apex.

6. Pronounced tubercle with free apex.

7. Hyper-pronounced tubercle with free apex.

Canine Mesial Ridge or Bushman Canine (MR-UC)

This trait can be observed in the upper canines. Frequently, canines may have large mesial ridges or tubercles, but when this two features merge and reach a point where the lingual sulcus is distal to the midline of the tooth, this can be considered as Bushman canine or Canine mesial ridge [111](Figure 4.6).

ASUDAS classifies the trait into four rank-scales grades of expression:

0. Mesial and distal lingual are the same size. Neither is attached to the tuberculum dentale if present.

1. Mesiolingual ridge is larger than the distolingual and is weakly attached to the tuberculum dentale.

2. Mesiolingual ridge is larger than the distolingual and is moderately attached to the tuberculum dentale.

3. Morris's type form. Mesiolingual ridge is much larger than the distolingual and is fully incorporated into the tuberculum dentale.

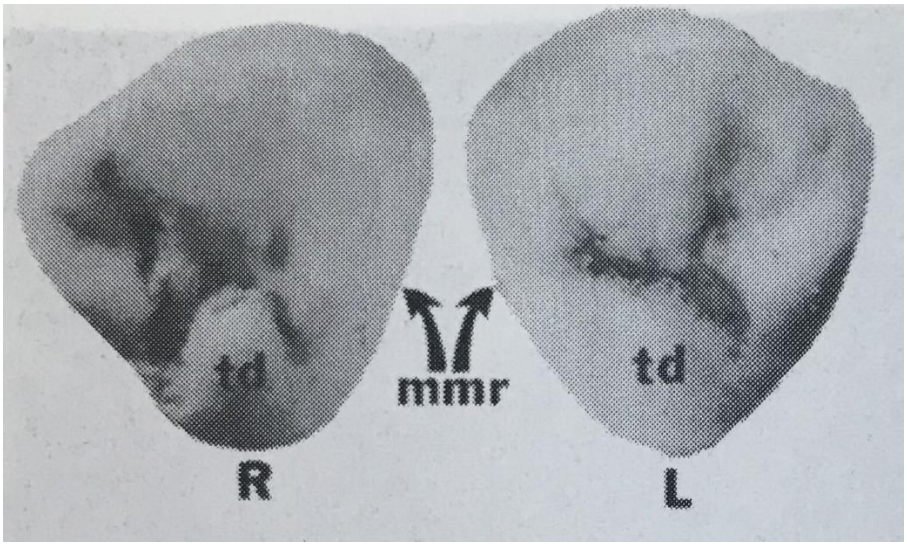


Figure 4.6. Mesial canine ridge or Bushman canine. The left canine (L) shows a mesial marginal ridge (mmr) and tuberculum dentale (td) that are not divided by a developmental groove. The right canine (R) shows a mesial marginal ridge and tuberculum dentale separated by a developmental groove. Figure from Scott and Turner (1997) [111].

Canine Distal Accessory Ridge (DAR-UC and DAR-LC)

Upper and lower canines show this feature on the lingual surface of the tooth between the apex of the tooth and the distal marginal ridge. It lies closer to the distal marginal ridge (Figure 4.7) [111]. Scoring:

0. Distal accessory ridge is absent.

1. Distal accessory ridge is very faint.

2. Distal accessory ridge is weakly developed.

3. Distal accessory ridge is moderately developed.

4. Distal accessory ridge is strongly developed.

5. Distal accessory ridge is very pronounced.

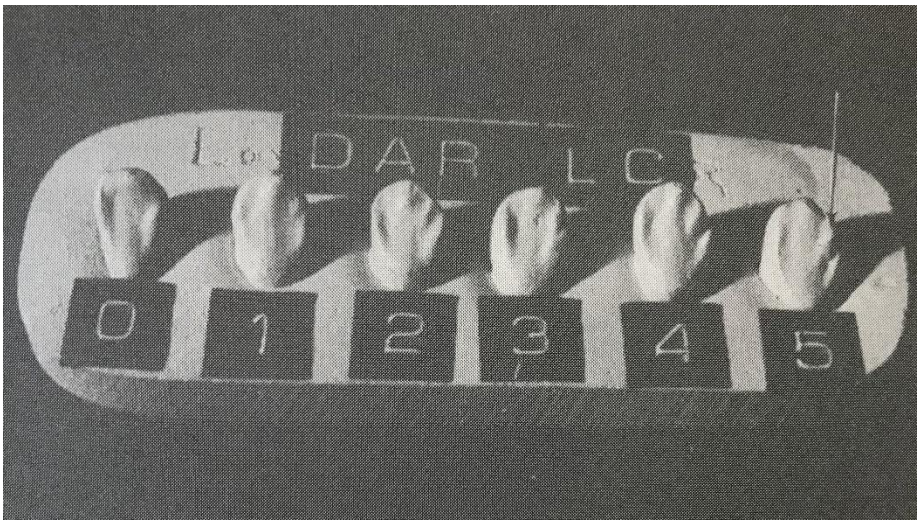


Figure 4.7. Canine distal accessory ridge. The plaque used to rank (0-5) lower canine distal accessory ridge. Number 5 is the highest level of expression; an arrow is showing the feature. Figure from Scott and Turner (1997) [111].

Congenital Absence (CA-UI2, CA-UP2 and CA-LP1)

This trait can be assessed in upper lateral and lower central incisors, upper and lower second premolars, and upper and lower third molars [111]. Scoring:

0. Tooth is present. Any degree of visible dental impaction is considered as present.

1. Tooth is congenitally absent.

Comment: In this thesis this feature was studied in upper lateral incisors, lower central incisors and lower first premolar.

Premolar:

Premolar Accessory Ridges (AR-UP1, AR-UP2)

This trait is observed on upper premolar 1 and 2, and lower premolar 1 and 2. This feature corresponds to the linear elevation between the buccal cusp ridge and the medial sulcus of maxillary first and second premolar (Figure 4.8) [111]. Scoring:

0. Absence of ridge.

T. Truncated ridge (is not continuous).

1. Trace (slight, but continuous ridge).

2. Small (thin continuous ridge).

3. Medium (moderately thick continuous ridge).

4. Pronounced (thick continuous ridge).

Comment: In this thesis, this trait was only scored in the upper premolar accessory ridges as absent=0 or present=1.

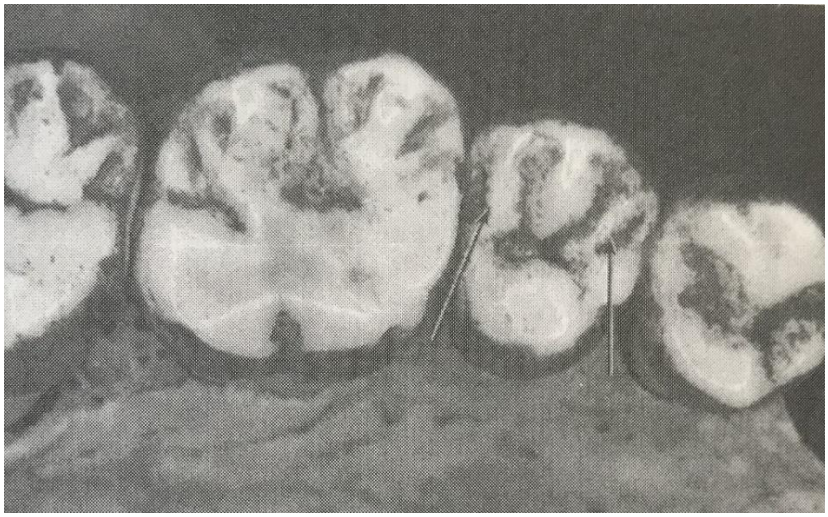


Figure 4.8. Premolar accessory ridges. The upper second premolar exhibits both mesial and distal accessory ridges. Figure from Scott and Turner (1997) [111].

Distosagittal Ridge or Uto-Aztecan Premolar (UTO-UP1)

This feature is seen in upper first premolar. A pronounced ridge from the apex of the buccal cusp extends to the distal occlusal edge at or near the sagittal sulcus. If straight lines are situated along the major axis of the buccal cusp and on the midline between

the two cusps, the angles of divergence range from 0 to 6 degrees (Morris, 1981) [375] (Figure 4.9). The Uto-Aztecan premolar is notorious, because this difference is greater than normal (35 to 45 degrees). There is also a mesial rotation of the buccal surface and a buccolingual expansion of the buccal cusp. [111] Classification:

0. Normal premolar shape.

1. Distosagittal ridge is present.

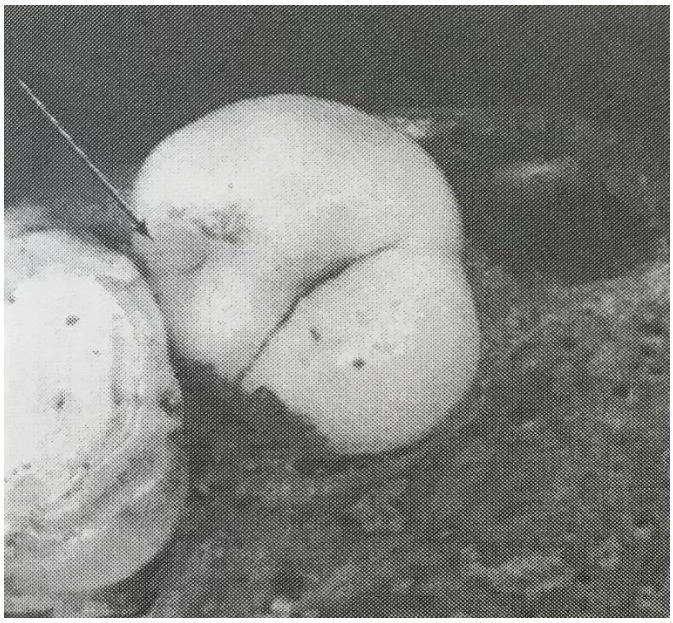


Figure 4.9. Distosagittal ridge (Uto-Aztecan premolars). Only present in upper first premolars. An arrow is showing the fossa adjacent to the marginal ridge. Figure from Scott and Turner (1997) [111].

Premolar Odontomes (ODO-UP1, ODO-UL1 and ODO-UL2)

This trait can be present in the central sulcus of upper and lower premolars. Odontomes have conical shape and they are the size of a pin (Figure 4.10) [111]. Scoring:

0. Absence.

1. Odontome is present.

Comment: In this thesis premolar odontome in the upper premolar two was not studied.

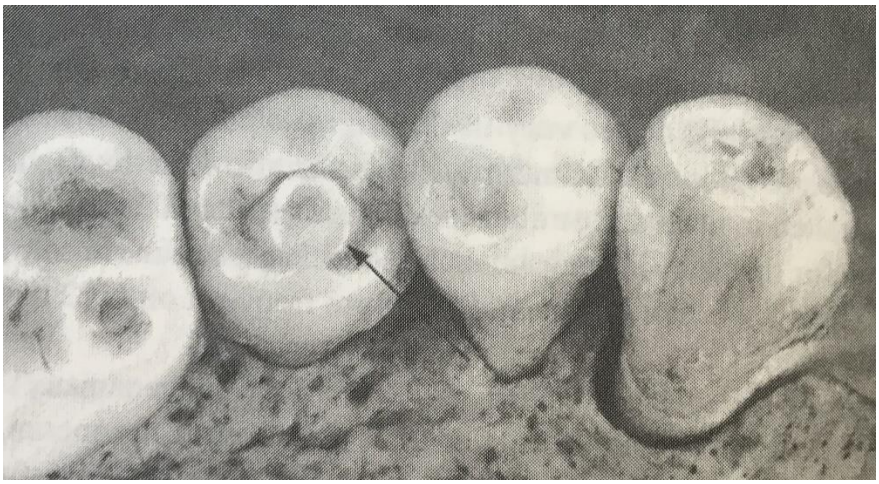


Figure 4.10. Premolar odontomes. Occlusal tubercles, or odontomes. On unworn teeth, odontomes are conical in form. In this case the arrow is showing the severe wear of the odontoma in lower second premolar. Figure from Scott and Turner (1997) [111].

Upper Premolar Mesial and Distal Accessory Cusps or Marginal tubercles (AMT-UP1 and AMT-UP2)

This trait is observed in upper premolars. Small accessory cusps are sometimes seen at the mesial or distal margin of the sagittal sulcus, which separates the buccal and lingual cusps [111]. Classification:

0. Accessory cusp absent.

1. Well-defined mesial accessory cusp with tangible cusp tip.

2. Well-defined distal accessory cusp with tangible cusp tip.

3. Well-defined mesial and distal accessory cusps with tangible cusp tips.

Tricusped premolars (3C-UP1 and 3C-UP2)

Upper premolars. Presence of 3 cusps in upper premolars [111]. Scoring:

0. Extra distal cusp (hypocone) is absent.

1. Hypocone is present. Its size is the same as regular lingual cusps.

Elongated upper and lower premolars (EP-LP1 and EP-LP2) (Edgar and Sciulli (2004)) [376].

This trait is observed in upper and lower premolars. It looks more rectangular than 'normal' premolars. This rectangular shape is due to primarily the mesiodistal length (Figure 4.11).

Classification:

0. Absent.

1. Present.

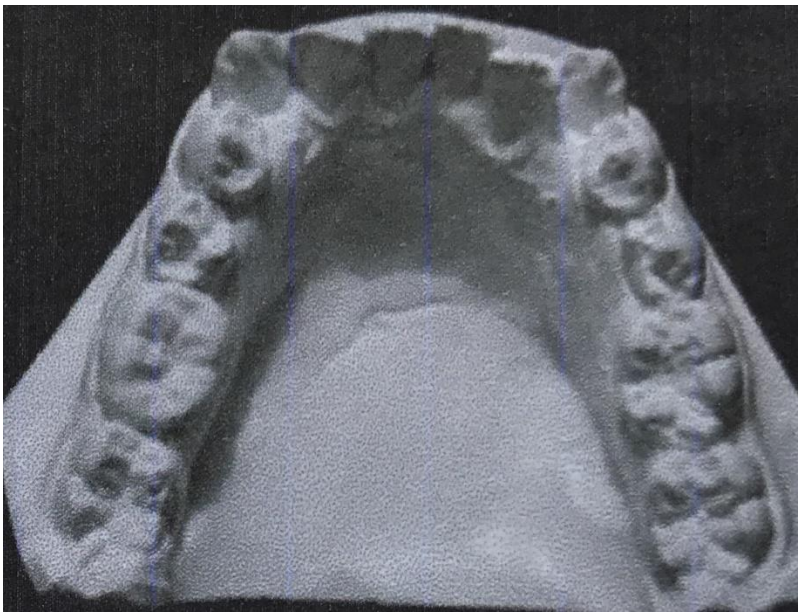


Figure 4.11. Elongated posterior premolars. The two posterior premolars express the feature. Figure from Edgar and Sciulli (2004) [376].

Molars:

Metacone (Met-UM1, Met-UM2 and Met-UM3)

This trait is seen in upper molars. Distobuccal cusp or cusp number 3 [111](Figure 4.12).
Classification:

0. Metacone absent.

1.There is a ridge at the metacone site but no free apex.

2. Metacone expressed as faint cuspule with a free apex.

3. A weak cusp is present.

3.5. An intermediate-sized cusp that falls between grades 3 and 4 (interpolation necessary)

4. Metacone is large.

5. Metacone is large, equal in size to a large UM1 hypocone.

Figure 4.12. Plaque used to score metacone expression. Figure from Scott and Irish (2017) [374].

3-Cusped upper second molar (3CUM2)

Upper molar two. It is the absence of the Hypocone [374]. Scoring:

0. Absence of the Hypocone.

1. Presence of the Hypocone.

Hypocone (HipUM1, HipUM2 and HipUM3)

Upper molars. It is the fourth cusp situated in the distolingual part of the tooth. Absence and severely reduced forms of this feature are more common on M2 (Figure 4.13) [111].Classification:

0. No hypocone.

1. Faint ridging.

2. Faint cuspule present.

3. Small cusp.

3.5. Moderate-sized cusp.

4. Large cusp.

5. Very large cusp.

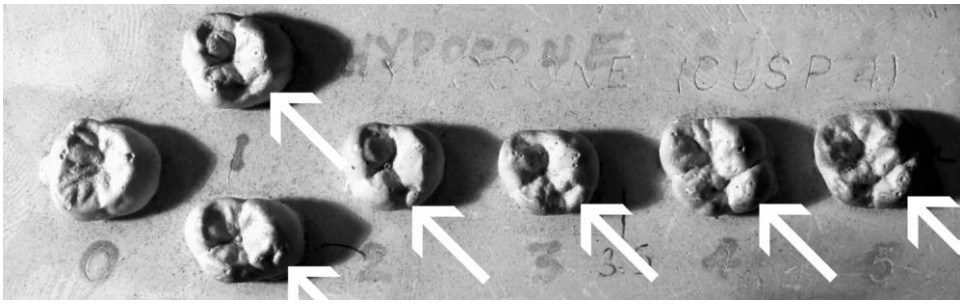


Figure 4.13. Plaque used to grade the expression of the Hypocone. Figure from Scott and Irish (2017) [374].

Cusp 5 (Metaconule) (C5UM1, C5UM2 and C5UM3)

This feature is observed in the upper molars. A fifth cusp shaped as a conule. May occasionally be present in the distal fovea between the hypocone and metacone of the upper molars [111](Figure 4.14). Scoring:

0. There is only a single distal groove present separating cusp 3 and 4.

1. Faint cuspule.

2. Trace cuspule.

3. Small cuspule.

4. Small cusp.

5. Medium-sized cusp.

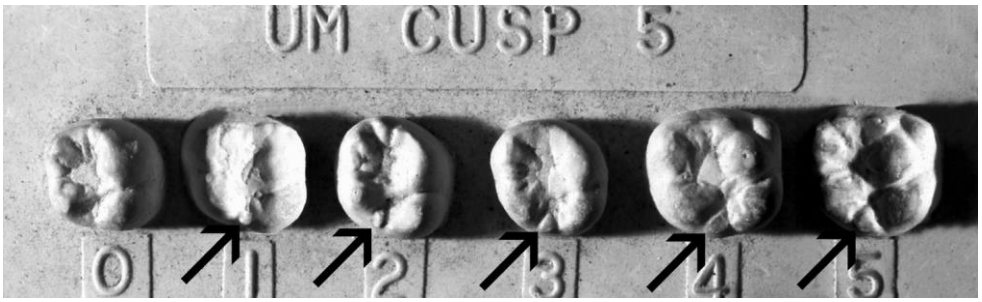


Figure 4.14. Plaque used to grade the expression of Cusp 5 (Metaconule). Figure from Scott and Irish (2017) [374].

Carabelli Tubercle (CarUM1, CarUM2 and CarUM3)

This trait can be seen in the upper molars. When it is present it is located in the lingual surface of the mesiolingual cusp (protocone or cusp 1) [111](Figure 4.15). Scoring:

0. Mesiolingual cusp 1 is smooth.

1. A vertical groove separates the protocone from the mesial marginal ridge complex; grade 1 expression occurs when there is a slight eminence that deflects distally from this groove.

2. A slight groove or eminence and takes the form of a pit.

3. A slightly Y-shaped depression.

4. A large Y-shaped depression.

5. A small tubercle without a free apex. The distal border of the cusp does not contact the lingual groove separating cusps 1 and 4.

6. Moderate tubercle with a free apex.

7. Pronounced tubercle with a free apex contacting the medial lingual groove is present.

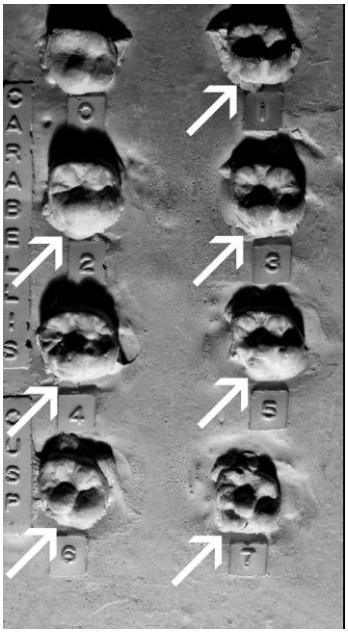


Figure 4.15. Plaque used to grade the expression of Carabelli's trait. Figure from Scott and Irish (2017) [374].

Parastyle (ParUM1, ParUM2 and ParUM3)

This feature is present in upper molars. It is sometimes related to Bolk´s paramolar tubercle, but they are different traits. Typically, the parastyle is expressed on the paracone or cusp 2. It ranges in size from a pit to a large free-standing tubercle. In some instances, it is expressed on cusp 3 (metacone) on any molar [111](Figure 4.16).

0. Buccal surfaces of cusps 2 and 3 are smooth.

1. A small pit near the buccal groove between cusps 2 and 3.

2. A small cusp but no free apex.

3. Medium cusp with free apex.

4. Large cusp with free apex.

5. Very large cusp with free apex that may extend onto the surfaces of both

cusps 2 and 3.

Comment: Grade 6. A free peg-shaped crown attached to the root of the third molar, it is very rare, and it is not shown on the plaque. In this thesis only grades from 0 to 5 were assessed.

Figure 4.16. Plaque used to grade the expression of Parastyle in upper molars. Figure from Scott and Irish (2017) [374].

Third Molars/peg-shaped/absent/reduced (CAUM3 and CALM3) Scott and Irish, 2017 [374].

Third molars, they are the most variable in terms of size, morphology, and number. As loss and reduction are elements of the same phenomenon, they are grouped together to form a single trait that involves pegged or reduced forms plus congenital absence. Scoring:

0. Third molar present and normal.

1. Third molar significantly reduced in size (half of normal size, with two or more cusps).

2. Third molar peg-shaped (only a single cusp evident) (Figure 4.17).

3. Third molar congenitally absent.

Comment: For population comparisons, the categories are often combined. In this thesis this trait was graded as present (grade 1) when third molar was reduced in size, peg-shaped and as absent or congenitally absent (grade 0) when there was no third molar.

Figure 4.17. Upper molar 3 showing a peg-shaped form. Figure from Scott and Irish (2017) [374].

Premolar Lingual Cusp Variation (LCV-LP1 and LCV-LP2)

This feature is present in lower premolars. Multiple lingual cusps take different forms on LP1 and LP2 and have separate classifications. There are 10 grades of classification, including absence and nine degrees of trait presence [111](Figure 4.18).

Scoring:

A. No lingual cusp. A ridge may be present, but reduced structure without a free tip.

0. One lingual cusp. Size and shape can vary but there is a tip present.

1. One or two lingual cusps.

2. Two lingual cusps. Mesial cusp is much larger than distal cusp.

3. Two lingual cusps. Mesial cusp is larger than distal cusp.

4. Two lingual cusps. Mesial and distal cusps are equal in size.

5. Two lingual cusps. Distal cusp is bigger than mesial cusp.

6. Two lingual cusps. Distal cusp is much larger than mesial cusp.

7. Two lingual cusps. Distal cusp is very much bigger than mesial cusp.

8. Three lingual cusps. Three of them are the same size.

9. Three lingual cusps. Mesial cusp is much larger than medial and/or distal cusp.

Figure 4.18. Plaque used to grade the expression of lower first premolar lingual cusp number. Figure from Scott and Irish (2017) [374].

Anterior Fovea (AF-LM1, AF-LM2 and AF-LM3) Scott and Irish (2107) [374].

It is observed in lower molars. It is located on the mesial aspect of the trigonid of the lower molars. There are three main elements that conformed this trait. Distinct essential ridges on the protoconid and metaconid that meet close to the centre of the trigonid, and a mesial marginal ridge that is expressed to varying degrees. When these three features join it produces a fovea, or depression, on the mesial section of the trigonid (Figure 4.19).

Scoring:

0. Anterior fovea is absent.

1. Trace, with slight development of mesial marginal ridge that connects the mesial aspects of cusps 1 and 2 producing a faint groove.

2. The essential ridges are better developed; thus, the groove is deeper than in grade 1.

3. The groove is longer than in grade 2.

4. Groove is very long, and mesial ridge is robust, producing a well-defined fovea.

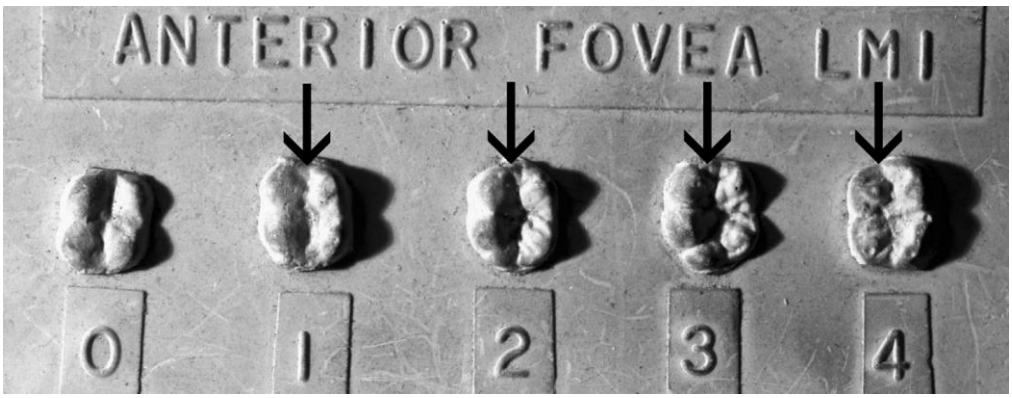


Figure 4.19. Plaque used to grade the expression of anterior fovea. Figure from Scott and Irish (2017) [374].

Lower Molar Groove Pattern (GP-LM1 and GP-LM2)

This trait is present in lower molars. Lower molars have five major cusps. (1) protoconid (mesiobuccal), (2) metaconid (mesiolingual), (3) hypoconid (distobuccal), (4) entoconid (distolingual), and (5) hypoconulid (distal). Classification of this trait is based on the contact between different cusps and the shape of these unions [111](Figure 4.20). Scoring:

Y. Cusps 2 and 3 are in contact.

+. Cusps 1, 2, 3 and 4 are in contact at the central sulcus.

X. Contact between cusps 1 and 4.

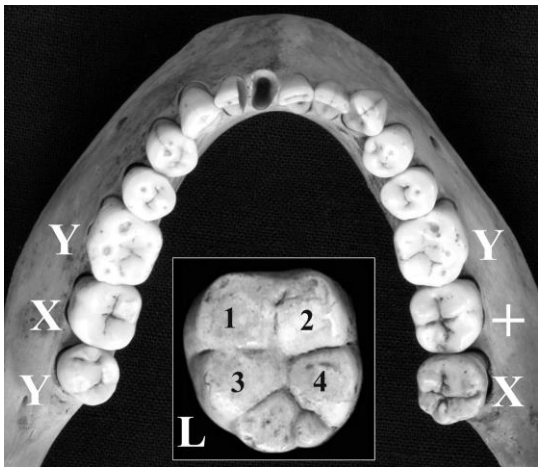


Figure 4.20. Lower molar groove pattern depends on the cusps contact. (L) shows five lowers molar cusps and their numbers). Both LM1s show Y patterns (2–3 contact) while left LM2 shows an X pattern (1–4 contact) and right LM2 shows a + pattern (1–2–3–4 contact). The abnormal left LM3 has a Y pattern, while its normal antimere has an X pattern. Figure from Scott and Irish (2017) [374].

Lower molar cusp number (CN-LM1, CN-LM2 and CN-LM3)

Lower molars. This feature depends entirely on the presence of cusp 5, or the hypoconulid [111](Figure 4.21). Scoring:

4. Cusps 1-4 (1, protoconid; 2, metaconid; 3, hypoconid; 4, entoconid) are present.

5. Cusp 5 (hypocoulid) is also present.

6. Cusp 6 (entoconulid) is also present.

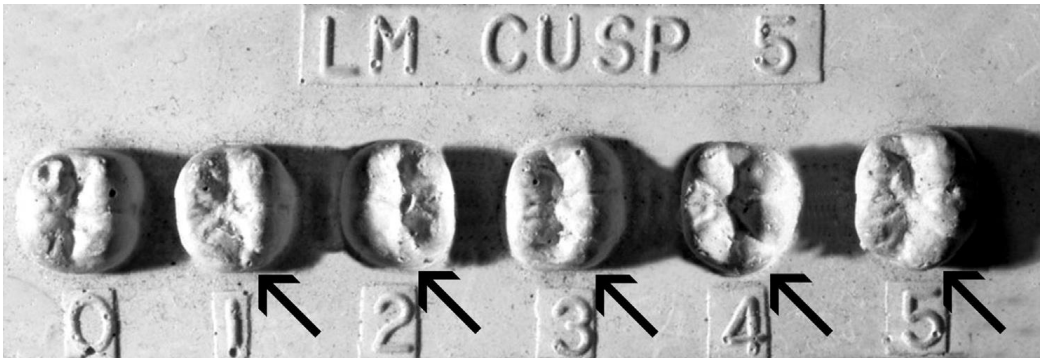


Figure 4.21. Plaque used to grade the expression of lower molar cusp number. This feature depends on the presence of Cusp 5. Figure from Scott and Irish (2017) [374].

Cusp 5 (Hypoconulid) (C5-LM1, C5-LM2 and C5-LM3)

This trait is observed in lower molars. It occurs on the distal occlusal surface of the tooth. The scoring system is based on the size only if cusp 6 is not present [111].

0. The molar has only 4 cusps. No hypoconulid.

1. Cusp 5 is present, but very little.

2. Cusp 5 is small.

3. Cusp 5 is medium-sized.

4. Cusp 5 is large.

5. Cusp 5 is very large.

Cusp 6 (Entoconulid) (C6-LM1, C6LM2 and C6LM3)

This feature can be seen in the lower molars. The entoconulid is expressed on the distal portion of the lower molars. Cusp 6 is associated with the entoconid (cusp 4). It is classified by size relative to cusp 5 [111](Figure 4.22). Scoring:

0. Cusp 6 is absent.

1. Cusp 6 is much smaller than cusp 5.

2. Cusp 6 is smaller than cusp 5.

3. Cusp 6 is equal in size to cusp 5.

4. Cusp 6 is larger than cusp 5.

5. Cusp 6 is much larger than cusp 5.

Figure 4.22. Plaque used to grade the expression of cusp 6 (entoconulid) of lower molars. Figure from Scott and Irish (2017) [374].

Cusp 7 (Metaconulid) (C7-LM1, C7-LM2 and C7-LM3)

It is observed in the lower molars. Cusp 7 or tuberculum intermedium is a wedge-shaped accessory cusp in the lingual groove, between cusps 2 and 4 [111](Figure 4.23). Scoring:

0. No cusp 7.

1. Faint wedge-shaped cusp between cusps 2 and 4.

1A. A faint cusp 7 without tip is present displaced on the lingual surface of cusp 2.

2. A small cusp 7.

3. A moderate cusp 7.

4. A large cusp 7.

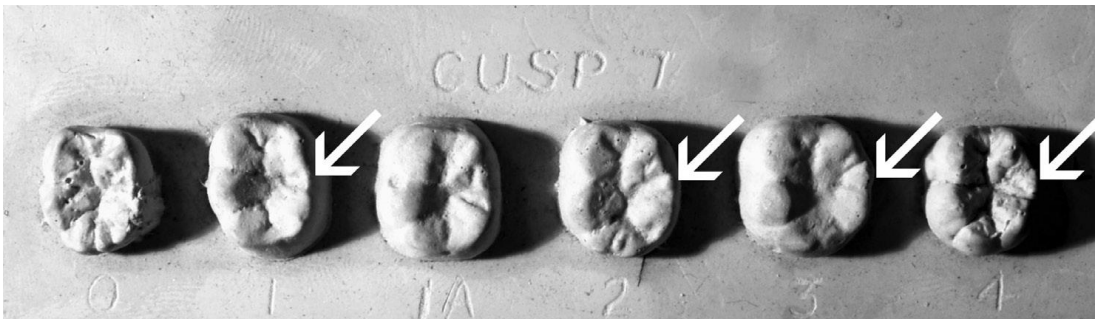


Figure 4.23. Plaque used to grade the expression of cusp 7 (metaconulid) of lower molars. Figure from Scott and Irish (2017) [374].

Deflecting Wrinkle (DW-LM1, DW-LM2 and DW-LM3) Scott and Irish (2017) [374].

Lower molars. The deflecting wrinkle is expressed on the occlusal surface of the mesiolingual cusp (metaconid). Basically, this trait is an unusual manifestation of the essential ridge of the metaconid. In most instances, this ridge runs a direct course from the cusp. But in some cases, the ridge changes course (or deflects) about halfway along its length before it terminates in the central sulcus (Figure 4.24). Scoring:

0. Deflecting wrinkle absent. Medial ridge of cusp 2 is straight.

1. Cusp 2 medial ridge is straight but with midpoint constriction.

2. Medial ridge deflects at halfway point toward central occlusal fossa but does not contact the cusp 4 (hypoconid).

3. Medial ridge shows strong deflection distally shaping an L-shaped ridge. And it does contact the hypoconid.

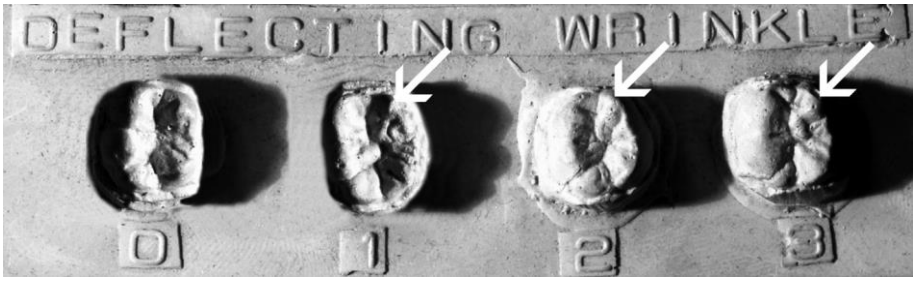


Figure 4.24. Plaque used to grade the expression of deflecting wrinkle of lower molars. Figure from Scott and Irish (2017) [374].

Distal Trigonid Crest (DTC-LM1, DTC-LM2 and DTC-LM3) [111]

This feature is observed in lower molars. Cusps 1 and 2 are joined by a ridge (Figure 4.25). Scoring:

0. Distal borders of protoconid and metaconid are not connected by a crest.

1. Distal trigonid crest present.

Figure 4.25. Distal trigonid crest; this crest is shaped when the distal accessory ridges of protoconid and metaconid meet at the central occlusal sulcus (can be continuous or discontinuous). Figure from Scott and Irish (2017) [374].

Protostylid (Protoconidal cingulum) (Prtost-LM1, Prtost-LM2 and Prtost-LM3)

This trait is present in lower molars. On the buccal surface of cusp 1. It is usually is associated with the buccal groove separating cusps 1 and 3 [111] (Figure 4.26). Scoring:

0. No expression of the trait.

1. A pit is present in the buccal grove.

2. Buccal groove is curved distally.

3. A faint secondary groove extends mesially from the buccal groove.

4. Secondary groove is slightly more pronounced.

5. Secondary groove is stronger and can be easily seen.

6. Secondary groove extends across most of the buccal surface of cusp 1. This is considered a weak or small cusp.

7. A cusp with a free apex occurs.

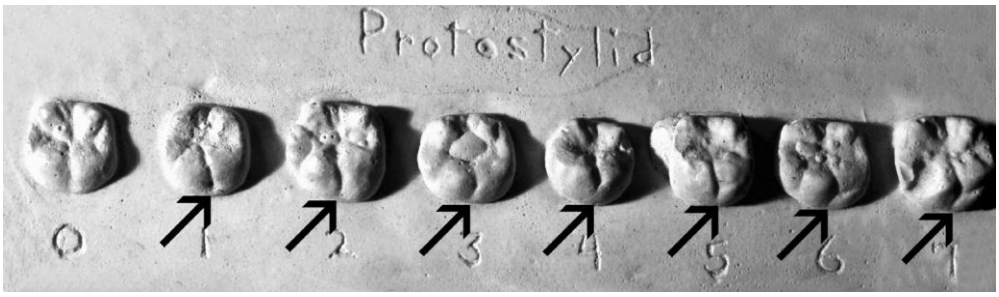


Figure 4.26. Plaque used to grade the expression of prototylid of lower molars. Figure from Scott and Irish (2017) [374].

4.1.3 Odontometric traits

Teeth and bones are the most resistant parts of the body, they disintegrate after death. Therefore, they are a good source of evidences for the evolutionary anthropologists, the archaeologists and the forensic scientists [377, 132].

Dental phenotypic data, both metric and the shape of the teeth has been widely used to assess the biological relationship among different populations (Section 1.3.3) [132] Based on some studies, in the last few decades the efforts have been focused on the non-metric dental traits [378] to characterize world-wide populations and establish population history, and origin of modern humans. Although, on the other hand dental measurements has often been used in the study of the evolution of the hominid and diversity of local and regional population groups [378]. In odontometrics, the most reported measurements are the maximum crown length (mesiodistal diameter) and the maximum crown wide (buccolingual diameter) [379]. Studies in current populations have classified these groups as microdontic, mesodontic and megadontic [380]. Despite odontometrics has been often used to differentiate among populations, there are

detractors that recognize that is useful among distant populations, but it does not work properly among closest populations [379].

The largest toothed population are native-Australians, their crowns are 30 to 35% larger than the crowns of Bushmen, Lapps, Iranians, etc., which are the populations with smallest teeth [379]. Hanihara et al. (2005) [378] measured teeth from 72 populations and seven geographic groups were assessed. They corroborated that the population with largest teeth are native Australians [379], follow by Melanesians, Micronesians, Sub-Saharan Africans and Native Americans. Intermediate size teeth were from East/southeast Asians and Polynesians and Jomon/Ainu and Western Eurasians have smaller teeth.

In this thesis I used odontometrics traits as quantitative phenotypes to perform the genome wide analysis due to metrics traits have a genetic component, based on the information presented above. Because if these measurements are used to differentiate across populations, they must have a genetic component that makes these phenotypes vary among world-wide populations and I attempted to find genomic regions associated to these traits.

4.2 Material and Methods

4.2.1 Study subjects

A sample of 501 volunteers from Medellín, Colombia were recruited and utilized to perform a genome-wide scan with categorical phenotypes. These people are part of the CANDELA Cohort [59]. Most of these individuals were students and staff from Antioquia University in Medellín. Adult subjects of both sexes aged between 18 and 40 years were invited to participate through media presentations, public lectures and personal contact with some volunteers (Table 4.2 and Figure 1.14). Then, a second GWAS with a bigger group of samples from the same Colombian Cohort (583 individuals) was performed with quantitative phenotypes (Table 4.3 and Figure 1.14).

		Total	Male	Female
Sample size		501	225	276
Percentage		100	44.91	55.09
Age (years)	Min	18	18	18
	Mean	23.03	23.57	22.60
	Max	40	40	39
	S.D.	4.94	5.12	4.76

Table 4.2. Features of the study sample used to perform the GWAS with ASUDAS data.

		Total	Male	Female
Sample size		564	252	312
Percentage		100	44.68	55.32
Age (years)	Min	18	18	18
	Mean	23.28	23.70	22.94
	Max	40	40	39
	S.D.	4.83	4.94	4.72

Table 4.3. Features of the study sample used to perform the GWAS with quantitative data.

Ethics approval was obtained from: Universidad de Antioquia (Colombia) and University College London (UK). All participants provided written informed consent.

Volunteers with antecedents of altered dental traits, several missing teeth, caries, apparent wear, maxillofacial surgery, orthognathic surgery, or any other massive craniofacial or dental surgery were excluded from this study.

Blood samples were collected by a certified phlebotomist and DNA extracted following standard laboratory procedures.

Intraoral digital photographs were obtained (by L.M.R.). The volunteer was seated in the dental chair, with an occipital support keeping the Frankfort plane (Figure 4.27 and Figure 4.28) parallel to the ground and with a labial retractor (PTJ Intl Co, Houdemont, France) located in the mouth. The operator used a Canon IE3 camera (12 megapixels) with a Canon 100mm macro and a Sigma ring flash. In the sagittal plane, the macro lens was focussed on the midline teeth (Figure 4.27). The transversal plane was parallel to the horizontal grid of the lens regarding the inter-canine plane. For arcade photographs, a metal mirror was used covering as much area as possible, while the lens remained focussed. The macro lens was used in manual mode and in a ratio of 2: 1 and 3: 1, which keeps the light, registration and focusing distance unchanged, which allows for comparison of images and distances among volunteers. To obtain the depth of field in the focus, the largest amplitude of the diaphragm was used. Subsequently, these photographs were used to obtain ordinal and quantitative phenotyping data.

Figure 4.27. Teeth photographs in different planes. Frontal, transversal and arcade photos covering as much of the area as possible. Images by L.M.R.

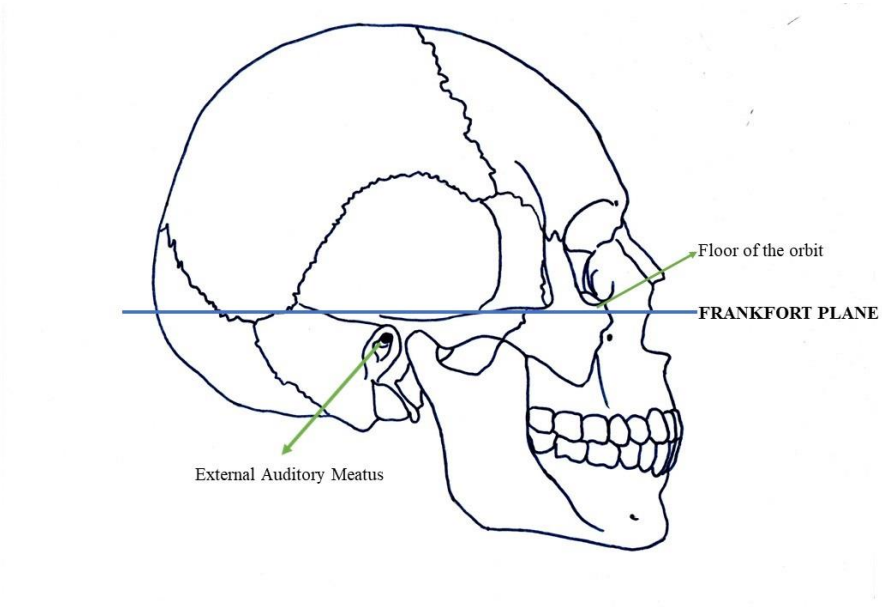


Figure 4.28. Frankfort plane. This plane passes through the floor of the orbit and the external auditory meatus [381].

Other phenotypes including height, weight, BMI, age and sex were also recorded for each participant [59].

Using these photographs, I performed several Genome Wide Association Studies between the genotypes and each phenotype. The first group of GWAS corresponds to ordinal traits (ASUDAS scoring, described in detail in Section 4.1.2.2 assessed in 501 individuals from the Colombian Cohort from CANDELA[59] (Table 4.2) and the second group of GWAS of quantitative traits (inciso-cervical, mesio-distal and bucco-lingual measurements of the teeth) performed in 564 people from the same cohort (Table 4.3).

4.2.2 Phenotyping

4.2.2.1 Ordinal traits (ASUDAS)

Intraoral digital photographs (Figure 4.27) were used to assess 86 dental traits following ASUDAS definitions [372] (Table 4.4, Figure 4.29, Section 4.1.2.2). The scoring was performed blindly with regards age, sex and genetic ancestry. When asymmetric expression was evident the score of the antimere with strongest expression was retained [382]. The scored traits are listed in Table 4.4.

Phenotypes	Acronym	Teeth Involved	Scale
Shovel Shape central upper incisors	**SSUI1**	Central Upper incisors	0─7
Double Shovel central upper incisors	**DSUI1**		0─6
Congenital absence lateral upper incisors	**CAUI2**		0─1
Winging central upper incisors	**WINUI1**		1─4
Labial curvature central upper incisors	**LCUI1**		0─4
Midline diastema upper central incisors	**DIASUI1**		0─1
Interruption grooves central upper incisors	**IGUI1**		0─1
Tuberculum dentale central upper incisors	**TDUI1**		0─6
Shovel Shape lateral upper incisors	**SSUI2**	Lateral Upper incisors	0-7
Double Shovel lateral upper incisors	**DSUI2**		0─6
Peg shape upper lateral incisors	**PSUI2**		0─2
Labial curvature lateral upper incisors	**LCUI2**		0─4
Interruption grooves lateral upper incisors	**IGUI2**		0─1
Tuberculum dentale lateral upper incisors	**TDUI2**		0─6
Tuberculum dentale upper canines	**TDUC**	Upper canines	0─6
Mesial ridge upper canine	**MRUC**		0─3
Distal accessory ridge of upper canine	**DARUC**		0─5
Accessory ridge upper premolar 1	**ARUP1**	Upper premolars	0─1
Accessory ridge upper premolar 2	**ARUP2**		0─1
Uto-Aztecan premolar 1	**UTOUP1**		0─1
Odontome upper premolar 1	**ODOUP1**		0─1
Accessory marginal tubercle upper premolar 1	**AMTUP1**		0─1
Accessory marginal tubercle upper premolar 2	**AMTUP2**		0─1
Congenital absence upper second premolar	**CAUP2**		0─1
Presence of 3 cusps in upper premolar 1	**3CUP1**		0─1
Presence of 3 cusps in upper premolar 2	**3CUP2**		0─1
Metacone upper molar 1	**MetUM1**	Upper molars	0─6
Metacone upper molar 2	**MetUM2**		0─6
Metacone upper molar 3	**MetUM3**		0─6
Presence of 3 cusps in upper molar 2	**3CUM2**		0─1
Hypocone upper molar 1	**HipUM1**		0─6
Hypocone upper molar 2	**HipUM2**		0─6
Hypocone upper molar 3	**HipUM3**		0─6
Cusp 5 upper molar 1	**C5UM1**		0─5
Cusp 5 upper molar 2	**C5UM2**		0─5
Cusp 5 upper molar 3	**C5UM3**		0─5
Carrabelli's tubercle upper molar 1	**CarUM1**		0─7
Carrabelli's tubercle upper molar 2	**CarUM2**		0─7
Carrabelli's tubercle upper molar 3	**CarUM3**		0─7
Parastyle upper molar 1	**ParUM1**		0─5
Parastyle upper molar 2	**ParUM2**		0─5
Parastyle upper molar 3	**ParUM3**		0─5
Upper molar 3, peg-shaped, absent, reduced	**CAUM3**		0─1

Phenotypes	Acronym	Teeth Involved	Scale
Shovel shape central lower incisor	**SSLI1**	Central lower incisors	0─3
Double shovel central lower inicisors	**DSLI1**		0─3
Congenital absence central lower incisors	**CALI1**		0─1
Shovel shape lateral lower incisor	**SSLI2**	Lateral lower incisors	0─3
Double shovel lateral lower inicisors	**DSLI2**		0─3
Accessory distal ridge of lower canine	**DARLC**	Lower canine	0─5
Lingual cusps variation central lower premolar	**LCVLP1**	Lower premolars	0─9
Lingual cusps variation lateral lower premolar	**LCVLP2**		0─9
Elongated premolar lower first premolar	**EPLP1**		0─1
Elongated premolar lower second premolar	**EPLP2**		0─1
Protoconid accesory ridge central lower premolar	**ARPrLP1**		0─1
Protoconid accesory ridge lateral lower premolar	**ARPrLP2**		0─1
Odontome lower first premolar	**ODOUL1**		0─1
Odontome lower second premolar	**ODOUL2**		0─1
Congenital absence lower first premolar	**CALP1**		0─1
Anterior fovea lower molar 1	**AFLM1**	Lower molars	0─4
Anterior fovea lower molar 2	**AFLM2**		0─4
Anterior fovea lower molar 3	**AFLM3**		0─4
Groove pattern lower molar 1	**GPLM1**		0─2
Groove pattern lower molar 2	**GPLM2**		0─2
Groove pattern lower molar 3	**GPLM3**		0─2
Cusp number lower molar 1	**CNLM1**		4─6
Cusp number lower molar 2	**CNLM2**		4─6
Cusp number lower molar 3	**CNLM3**		4─6
Cusp 5 lower molar 1	**C5LM1**		0─6
Cusp 5 lower molar 2	**C5LM2**		0─5
Cusp 5 lower molar 3	**C5LM3**		0─5
Cusp 6 lower molar 1	**C6LM1**		0─5
Cusp 6 lower molar 2	**C6LM2**		0─5
Cusp 6 lower molar 3	**C6LM3**		0─5
Cusp 7 lower molar 1	**C7LM1**		0─4
Cusp 7 lower molar 2	**C7LM2**		0─4
Cusp 7 lower molar 3	**C7LM3**		0─4
Deflecting wrinkle lower molar 1	**DWLM1**		0─3
Deflecting wrinkle lower molar 2	**DWLM2**		0─3
Deflecting wrinkle lower molar 3	**DWLM3**		0─3
Distal-Trigonid crest lower molar 1	**DTCLM1**		0─1
Distal-Trigonid crest lower molar 2	**DTCLM2**		0─1
Distal-Trigonid crest lower molar 3	**DTCLM3**		0─1
protostylid lower molar 1	**PrtostLM1**		0─7
protostylid lower molar 2	**PrtostLM2**		0─7
protostylid lower molar 3	**PrtostLM3**		0─7
Lower molar 3, peg-shaped, absent, reduced	**CALM3**		0─1

Table 4.4. ASUDAS traits assessed. The traits were scored following ASUDAS dental traits descriptions.

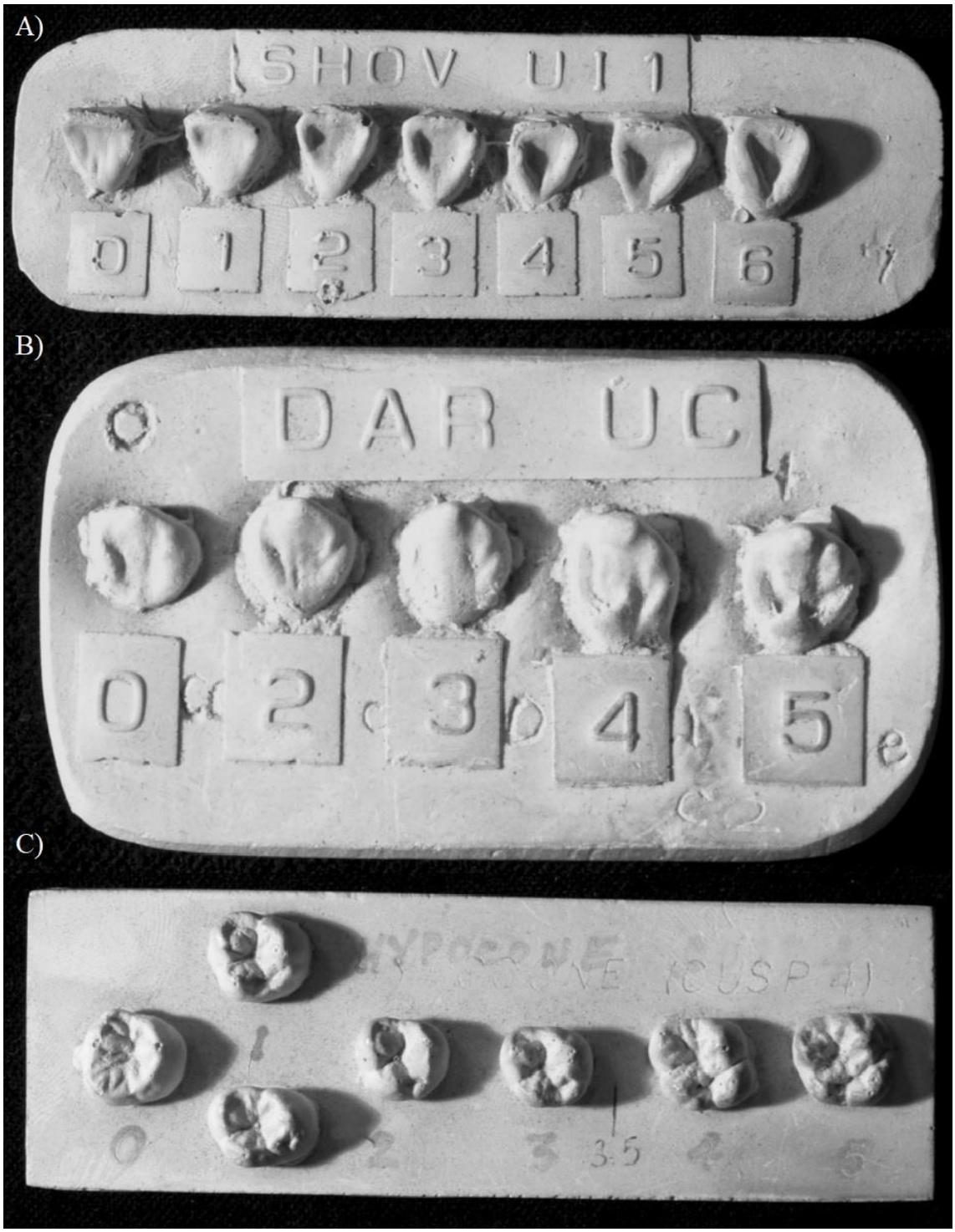


Figure 4.29. Examples ASUDAS scoring plaques. A) Central incisor shovel shape plaque. B) Upper canine distal accessory ridge plaque and C) Hypocone plaque. Figure from Scott and Turner (1997) [111].

4.2.2.1.1 Intra-observer rater reliability.

The intra-rater reliability rate is shown in Delgado, 2015 [135]. Photographs for all the volunteers were scored by the same rater (M.D.).

4.2.2.1.2 Quality control of phenotype data

The 86 ordinal phenotyped traits were filtered by missingness and the distribution of the data. The first filter was missingness, traits presenting more than 15% of missing data were removed. The next filter was the distribution of the categories within the traits, when one category was constant, this means that within a trait, all the individuals were phenotyped in the same category, rare with a threshold of <10% (or when more than 90% of the individuals were classified in one category and less than 10% of the remaining ones were put in the rest of the categories within a trait) and they were not normally distributed, the trait was removed (Table 4.5 and Appendix B). After the QC 46 traits were retained.

	All Data													
		Counts											**Filtering**	
Trait	**Type of Data**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**NA**	**Rarity**	**Missingness**
SSUI1	Ordinal	150	23	90	81	105	27	1	0			0		
SSUI2	Ordinal	141	18	94	88	104	25	4	0			3		
DSUI1	Ordinal	223	72	130	40	9	1	1				1		
DSUI2	Ordinal	249	73	113	26	9	2	0				5		
CAUI2	Binary	474	2									1	<10%	
PSUI2	Ordinal	456	16	3								2	<10%	
WINUI1	Ordinal		38	2	416	18						3		
LCUI1	Ordinal	445	22	9	0	0						1	<10%	
LCUI2	Ordinal	236	63	108	47	18						5		
DIASUI1	Binary	445	31									1	<10%	
IGUI1	Binary	367	107									3		
IGUI2	Binary	234	235									8		
TDUI1	Ordinal	283	5	69	89	19	4	3				5		
TDUI2	Ordinal	215	2	38	68	105	21	20				8		
TDUC	Ordinal	180	11	73	68	103	16	22				4		
MRUC	Ordinal	360	52	47	15	0						3		
DARUC	Ordinal	340	12	78	38	5	1					3		
ARUP1	Binary	297	135									45		
ARUP2	Binary	161	299									17		
UTOUP1	Binary	432	0									45	Constant	
ODOUP1	Binary	432	1									44	<10%	
AMTUP1	Binary	192	240									45		

Continue…

		All Data												
		Counts											Filtering	
Trait	**Type of Data**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**NA**	**Rarity**	**Missingness**
AMTUP2	Binary	196	264									17		
CAUP2	Binary	456	5									16	<10%	
3CUP1	Binary	431	1									45	<10%	
3CUP2	Binary	462	0									15	Constant	
MetUM1	Ordinal	0	0	0	15	224	237					1		
MetUM2	Ordinal	0	0	2	50	222	202					1		
MetUM3	Ordinal	1	1	7	60	32	38					338		>15%
3CUM2	Binary	337	137									3		
HipUM1	Ordinal	5	0	3	14	188	266					1		
HipUM2	Ordinal	119	9	92	80	81	93					3		
HipUM3	Ordinal	58	4	36	25	3	0					351		>15%
C5UM1	Ordinal	415	3	27	23	5	2					2		
C5UM2	Ordinal	415	0	26	21	10	4					1		
C5UM3	Ordinal	115	0	4	3	1	0					354	<10%	>15%
CarUM1	Ordinal	284	5	52	14	28	33	25	35			1		
CarUM2	Ordinal	430	2	17	7	6	5	4	5			1	<10%	
CarUM3	Ordinal	128	0	0	0	0	1	1	1			346	<10%	>15%
ParUM1	Ordinal	457	7	4	0	0	0					9	<10%	
ParUM2	Ordinal	455	2	2	2	1	7					8	<10%	
ParUM3	Ordinal	125	1	0	0	0	1					350	<10%	>15%
CAUM3	Binary	188	212									77		>15%
SSLI1	Ordinal	322	72	73	9	1						0		
SSLI2	Ordinal	328	70	67	11	1						0		
DSLI1	Ordinal	411	45	17	4							0		
DSLI2	Ordinal	424	32	18	3							0		
CALI1	Binary	476	1									0	<10%	
DARLC	Ordinal	368	23	56	18	7	1					4		
LCVLP1	Ordinal	86	96	82	60	19	46	22	21	6	8	31		
LCVLP2	Ordinal	18	62	113	99	46	35	10	43	36	9	6		
EPLP1	Binary	416	31									30	<10%	
EPLP2	Binary	460	11									6	<10%	
ARPrLP1	Binary	324	123									30		
ARPrLP2	Binary	269	201									7		
ODOUL1	Binary	445	1									31	<10%	
ODOUL2	Binary	470	0									7	Constant	
CALP1	Binary	451	0									26	Constant	
AFLM1	Ordinal	360	56	36	21	3						1		
AFLM2	Ordinal	232	102	75	60	6						2		
AFLM3	Ordinal	70	20	10	6	0						371		>15%
GPLM1	Ordinal	435	30	7								5	<10%	
GPLM2	Ordinal	155	260	57								5		
GPLM3	Ordinal	33	21	34								389		>15%

Continue…

	All Data													
		Counts											Filtering	
Trait	**Type of Data**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**NA**	**Rarity**	**Missingness**
CNLM1	Ordinal					33	340	81	20			3		
CNLM2	Ordinal					301	125	44	4			3		
CNLM3	Ordinal					39	39	7	2			390		>15%
C5LM1	Ordinal	35	1	20	125	260	33					3		
C5LM2	Ordinal	314	5	69	52	30	4					3		
C5LM3	Ordinal	39	1	10	19	11	4					393		>15%
C6LM1	Ordinal	409	1	57	4	2	0					4		
C6LM2	Ordinal	435	3	28	7	0	0					4	<10%	
C6LM3	Ordinal	76	0	4	2	0	0					395	<10%	>15%
C7LM1	Ordinal	413	2	19	15	18	6					4		
C7LM2	Ordinal	447	1	8	13	5						3	<10%	
C7LM3	Ordinal	79	0	2	1	2						393	<10%	>15%
DWLM1	Ordinal	409	21	30	13							4		
DWLM2	Ordinal	458	3	11	2							3	<10%	
DWLM3	Ordinal	91	1	2	2							381	<10%	>15%
DTCLM1	Binary	455	13									9	<10%	
DTCLM2	Binary	436	37									4	<10%	
DTCLM3	Binary	86	6									385	<10%	>15%
PrtostLM1	Ordinal	193	260	5	4	3	2	2	0			8		
PrtostLM2	Ordinal	333	126	4	3	2	2	0	1			6		
PrtostLM3	Ordinal	80	11	0	1	1	0	1	5			378		>15%
CALM3	Binary	181	203									93		>15%

Table 4.5. Summary of ASUDAS data, after quality control. Removed traits due to either Missingness >15% or distribution of the data (constant and rare traits < 10%). Fifty-five traits were left after the missingness filter. Finally, after the filter to check the distribution of traits, forty-six traits were left to perform the GWAS.

4.2.2.1.3 Regrouping ASUDAS dental traits

As mentioned above (Section 4.2.2.1), the 86 phenotypes were assessed following the ASUDAS scale and then filtered based on the distribution, missingness, etc.; 46 traits were retained for the genome-wide scans. Subsequently, following Scott and Irish (2017) [374] criteria and the group's physical anthropologist (M. Delgado) suggestions based on his expertise, some traits were regrouped. These new groups went through the same quality control (Section 4.2.2.1.2) as the original data. There were 3 types of regrouping criteria. In the first one, categories were regrouped differently to the original ASUDAS scale, i.e. Shovel Shape upper central incisor (SSUI1), commonly classified in a scale from 0 to 6, changed as follows, 0-1 → 0, 2-4 → 1 and 5-6 → 2, thus from 7 categories the number was lowered to 3, but still it is an ordinal trait (Table 4.6). After regrouping the ASUDAS trait, and the quality control was performed, 34 traits were left for further analyses.

	Regrouping						
						Filters	
Trait	**Suggested groups**					**Type of Data**	**Rarity**
SSUI1	0-1	2-3	4-6			Ordinal	
SSUI2	0-1	2-3	4-6			Ordinal	
DSUI1	0	1	2	3	4-5-6	Ordinal	
DSUI2	0	1	2	3	4-5-6	Ordinal	
WINUI1*	1-2	3-4				Binary	Constant
LCUI2	0	1-2	3-4			Ordinal	
TDUI1	0	1-2	3-4	5-6		Ordinal	
TDUI2	0	1-2	3-4	5-6		Ordinal	
TDUC	0	1-2	3-4	5-6		Ordinal	
MRUC	0	1-2	3			Ordinal	
DARUC	0	1-2	3-5			Ordinal	
MetUM1	0	1-2	3-4	5-6		Ordinal	
HipUM1	0	1-2	3-4	5-6		Ordinal	
HipUM2	0	1-2	3-4	5		Ordinal	
C5UM1	0	1-3	4-5			Ordinal	
C5UM2	0	1-3	4-5			Ordinal	
CarUM1	0	1-3	4-7			Ordinal	
CarUM2	0	1-3	4-7			Ordinal	
ParUM1	0	1-2				Binary	
ParUM2	0	1-2	3-5			Ordinal	
SSLI1	0	1	2	3-4		Ordinal	
SSLI2	0	1	2	3-4		Ordinal	
DSLI2	0	1-3				Binary	
DARLC	0	1-2	3-5			Ordinal	
LCVLP1	0-1	2-7	8-9			Ordinal	
LCVLP2	0-1	2-7	8-9			Ordinal	
AFLM1	0	1-2	3-4			Ordinal	
AFLM2	0	1-2	3-4			Ordinal	
C5LM1	0	1-3	4-5			Ordinal	
C5LM2	0	1-3	4-5			Ordinal	
C6LM1	0	1-2	3-4			Ordinal	
C7LM1	0	1-3	4-5			Ordinal	
DWLM1	0	1	2-3			Ordinal	
PrtostLM1	0	1	2-7			Ordinal	
PrtostLM2	0	1	2-7			Ordinal	

Table 4.6. Regrouping of ASUDAS traits based on frequency. After regrouping of the different categories of the ordinal phenotyping data (ASUDAS), based on Scott and Irish (2017) [374] and the Physical Anthropologist (M.D.) criteria, 35 new groups were suggested, but 1 trait was removed because it became constant (WINUI1*).

The second criterion was based on absence or presence of the trait in each tooth, this is called Breakpoint approach in Scott and Irish (2017) [374], for instance the same trait SSUI1, 0-1→ 0 (Absence) and 2-6 → 1 (Presence), originally the data was ordinal and it later became binary data (Table 4.7). After applying the breakpoint criterion and the quality control filters, 28 traits were left for further analyses.

	Breakpoint				
				Filter	
Trait	**Degrees**	**Presence(1)**	**Absence(0)**	**Type of Data**	**Rarity**
SSUI1	0–7	2–7	0–1	Binary	
SSUI2	0-7	2–7	0–1	Binary	
DSUI1	0–6	2–6	0–1	Binary	
DSUI2	0–6	2–6	0–1	Binary	
PSUI2	0–2	1–2	0	Binary	
WINUI1*	1–4	1	2–4	Binary	Rare
LCUI1	0–4	1–4	0	Binary	
LCUI2	0–4	1–4	0	Binary	
TDUI1	0–6	1–6	0	Binary	
TDUI2	0–6	1–6	0	Binary	
TDUC	0–6	1–6	0	Binary	
MRUC	0–3	1–3	0	Binary	
DARUC	0–5	2–5	0–1	Binary	
MetUM1*	0–5	2–5	0–1	Binary	Constant
MetUM2*	0–5	2–5	0–1	Binary	Constant
HipUM1*	0–5	2–5	0–1	Binary	Rare
HipUM2	0–5	2–5	0–1	Binary	
C5UM1	0–5	1–5	0	Binary	
C5UM2	0–5	1–5	0	Binary	
CarUM1	0–7	2–7	0–1	Binary	
CarUM2	0–7	2–7	0–1	Binary	
ParUM1	0–5	1–5	0	Binary	
ParUM2	0–5	1–5	0	Binary	
SSLI1	0–3	2–3	0–1	Binary	
SSLI2	0–3	2–3	0–1	Binary	
DSLI1*	0–3	2–3	0–1	Binary	Rare

Continue…

	Breakpoint				
				Filter	
Trait	**Degrees**	**Presence (1)**	**Absence (0)**	**Type of Data**	**Rarity**
DSLI2*	0─3	2─3	0─1	Binary	Rare
DARLC	0─5	2─5	0─1	Binary	
LCVLP1	0─9	2─9	0─1	Binary	
LCVLP2	0─9	2─9	0─1	Binary	
AFLM1	0─4	2─4	0─1	Binary	
AFLM2	0─4	2─4	0─1	Binary	
GPLM1	Y, ± ,X	Y	X, ±	Binary	
GPLM2	Y, ± ,X	Y	X, ±	Binary	
CNLM1	4─6	*6 or more	4─5	Binary	
CNLM2	4─6	5─6	4	Binary	
C5LM1*	0─5	1─5	0	Binary	Rare
C5LM2	0─5	1─5	0	Binary	
C6LM1	0─5	1─5	0	Binary	
C6LM2	0─5	1─5	0	Binary	
C7LM1	0─4	1─4	0	Binary	
C7LM2	0─4	1─4	0	Binary	
DWLM1*	0─3	2─3	0─1	Binary	Rare
DWLM2	0─3	2─3	0─1	Binary	
PrtostLM1	0─7	1─7	0	Binary	
PrtostLM2	0─7	1─7	0	Binary	

Table 4.7. Regrouping based on breakpoint criterion. All the traits were converted to binary traits after applying the filters. Eight traits were removed because they presented rare or constant frequency in one category. Finally, 28 new breakpoint groups were retained. * Traits removed after the filtering.

Finally, for the composite trait criterion, this consists of the absence or presence of a trait in any tooth. For instance, if shovelling is present in the lower central incisors, and not in the rest of the teeth where this trait should be expressed (upper central incisors and upper and lower lateral incisors), it is considered as the trait is present anyway (Table 4.8). After the composite criterion and the quality control filters were applied, 21 traits were retained for further analyses.

Composite Trait					
				Filter	
Trait	**Teeth**	**Absent(0)**	**Present(1)**	**Type of Data**	**Rarity**
Shoveling	UI1, UI2, LI1, LI2	0-1	2-6	Binary	
Double Shoveling	UI1, UI2, LI1, LI2	0-1	2-6	Binary	
Congenital Absence*	UI1, UI2, LI1, LI2	0		Binary	Rare
Labial Curvature	UI1,UI2	0	1-4	Binary	
Interruption Grooves	UI1,UI2	0	1	Binary	
Dental Tubercle	UI1,UI2,UC	0	1-6	Binary	
Distal Accesory Ridge	UC,LC	0-1	2-5	Binary	
Accesory Ridges	UP1,UP2,LP1,LP2	0	1	Binary	
Odontome	UP1,LP2	0	1	Binary	
Marginal Accesory Tubercle*	UP1,UP2	0	1	Binary	Rare
Three cusps	UP1,UP2,UM2	0	1	Binary	
Metacone*	UM1,UM2,UM3	0-1	2-5	Binary	Constant
Hypocone	UM1,UM2,UM3	0-1	2-5	Binary	
Metaconule	UM1,UM2,UM3	0	1-5	Binary	
Carabelli's tubercle	UM1,UM2,UM3	0-1	2-7	Binary	
Parastyle*	UM1,UM2,UM3	0	1-5	Binary	Rare
Peg-shaped reduced or absent	UM3,LM3	0	1	Binary	
Lingual cusp variation*	LP1, LP2	0-1	2-9	Binary	Rare
Elongated premolar	LP1, LP2	0	1	Binary	
Anterior foveae*	LM1,LM2, LM3	0-1	2-5	Binary	Rare
Groove pattern	LM1,LM2, LM3	+,X	Y	Binary	
Cusp number*	LM2,LM3	4	5+	Binary	Rare
Hypoconulid	LM1,LM2, LM3	0	1-5	Binary	
Entoconulid	LM1,LM2, LM3	0	1-5	Binary	
Metaconulid	LM1,LM2, LM3	0	1-4	Binary	
Deflecting Wrinkle	LM1,LM2, LM3	0-1	2-3	Binary	
Distal trigonid crest	LM1,LM2, LM3	0	1	Binary	
Protostylid	LM1,LM2, LM3	0	1-6	Binary	

Table 4.8. Regrouping based on absence or presence of the trait in one tooth (composite criterion). After applying the composite procedure, the number of traits decreased to 28 binary traits and after the filtering, 21 traits were retained.

* Removed trait due to the distribution of the trait or the trait presented some category frequency either rare or constant.

4.2.2.2 Quantitative phenotyping

In the same photos used for the ordinal phenotyping (Figure 4.27), three different distances were taken in incisors, canines and premolar teeth. Distance from the cervical margin to the incisal edge, distance from the mesial side to the distal side and the distance from the buccal face to the lingual face of the teeth. There are measurements

for 564 individuals, after filtering, 509 were left to run the GWAS. The quantitative measurements (distances) were assessed in incisors, canines and premolars (Figure 4.30 and Table 4.9).

4.2.2.2.1 Intra-observer rater reliability

Intraclass correlation coefficients (ICCs) [291] were used to evaluate the rater reliability for dental features measurements. This approach uses a two-way mixed effects ANOVA model, with two different scorings (from repeated scoring by one rater or from scores by two raters) as a fixed effect, and variation across subjects as a random effect. In this case it was calculated with repeated scoring by one rater. Scores from a set of photographs for 28 individuals (~ 5% of total sample size, from the same Colombian sample) were used for calculating ICCs for each dental measurement. The photographs were scored twice by one rater (L.M.R.), independently, two weeks apart. Photographs for all the volunteers were scored by the same rater (L.M.R.).

4.2.2.2.2 Quality control of phenotyping data

Summary statistics for the 60 measurements, corresponding to 3 measurements per tooth (20 teeth, upper and lower incisors, canines and premolars), were calculated to sum up the observations in order to detect if some traits have typos or any other mistake (Table B.1). Also, the distribution of each trait was checked (Figures B.67-B.115).

4.2.2.3 Imputation of dental traits

In order to replace missing values with substitute values in the ASUDAS and metric dental data to improve the further analyses, imputation of the missing data was applied to both. The methodology used was missForest (package in R), this method uses a trained random forest on the existing data to predict the missed values. It is applied on different types of data (continuous and/or categorical) [383]. The imputation was performed 5 times and an average of the 5 different predicted values was used, this process is called bagging, which is useful to reduce the variability in prediction and improve accuracy.

The variables and individuals retained had low missingness proportion. Thus, the overall effect of imputation was low - that is, if we use different algorithms, different parameters, etc. the final outcomes of any analysis applied to the imputed data will not

be hugely affected, but imputation is necessary because many methods such as PCA do not work on incomplete data.

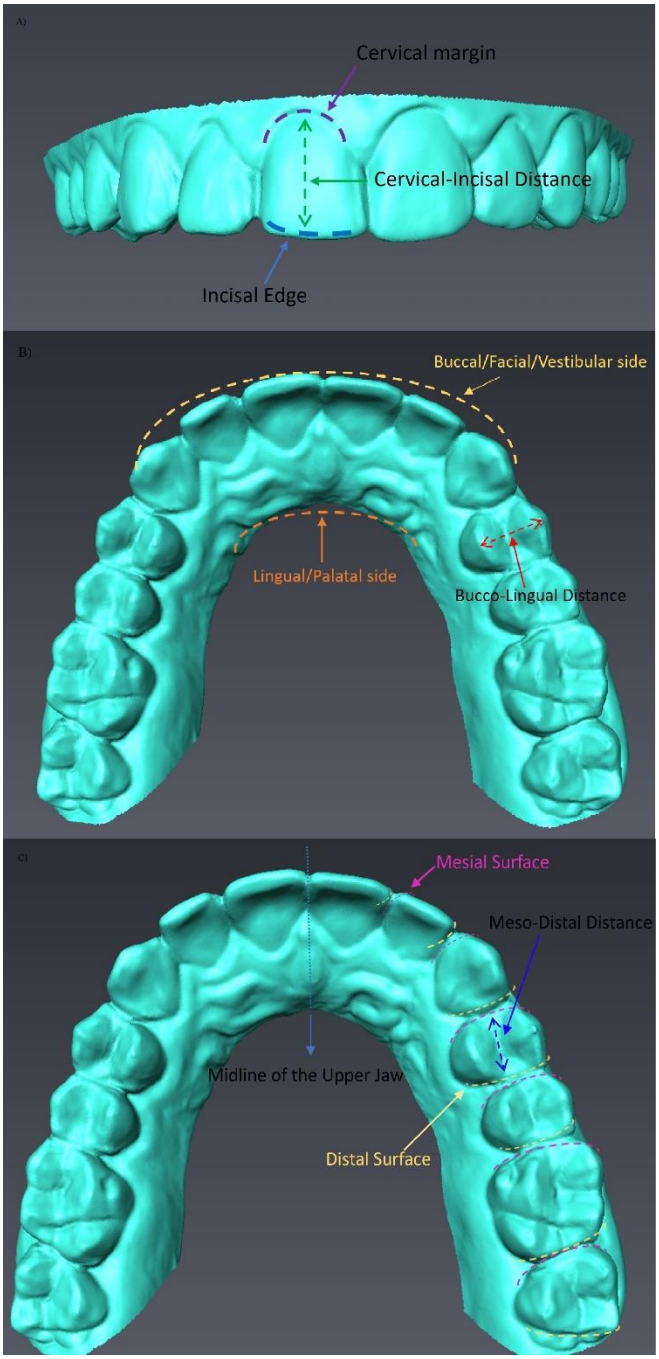


Figure 4.30. Measurements obtained in Incisors, Canines and Premolars.A) Distance between the cervical margin of a tooth (the area above the junction of the crown and the root of the tooth) and the incisal edge (which corresponds to the cutting edge of a tooth).B). Distance between the buccal part (the surface of a tooth that faces the buccal or labial mucosa of the vestibule of the mouth) and the lingual side (surface of a maxillary tooth, which is in direct contact with the tongue) of the mouth. C) Distance between the mesial surface (the part of a tooth that is next to the tooth in front of it or that is closest to the midline of the front of the jaw) and the distal part of a tooth (the surface of a tooth that is next to the tooth behind it or that is farthest from the middle of the front of the jaw).

Jaw	Measurement	Side	Tooth	Trait code used in the analyses
Upper	Incisal-Cervical Distance	Right	Central Incisor	IC_Up_Rt_I1
			Lateral Incisor	IC_Up_Rt_I2
			Canine	IC_Up_Rt_C
			First Premolar	IC_Up_Rt_P1
			Second Premolar	IC_Up_Rt_P2
		Left	Central Incisor	IC_Up_Lf_I1
			Lateral Incisor	IC_Up_Lf_I2
			Canine	IC_Up_Lf_C
			First Premolar	IC_Up_Lf_P1
			Second Premolar	IC_Up_Lf_P2
	Meso-Distal Distance	Right	Central Incisor	MD_Up_Rt_I1
			Lateral Incisor	MD_Up_Rt_I2
			Canine	MD_Up_Rt_C
			First Premolar	MD_Up_Rt_P1
			Second Premolar	MD_Up_Rt_P2
		Left	Central Incisor	MD_Up_Lf_I1
			Lateral Incisor	MD_Up_Lf_I2
			Canine	MD_Up_Lf_C
			First Premolar	MD_Up_Lf_P1
			Second Premolar	MD_Up_Lf_P2
	Bucco-Lingual Distance	Right	Central Incisor	BL_Up_Rt_I1
			Lateral Incisor	BL_Up_Rt_I2
			Canine	BL_Up_Rt_C
			First Premolar	BL_Up_Rt_P1
			Second Premolar	BL_Up_Rt_P2
		Left	Central Incisor	BL_Up_Lf_I1
			Lateral Incisor	BL_Up_Lf_I2
			Canine	BL_Up_Lf_C
			First Premolar	BL_Up_Lf_P1
			Second Premolar	BL_Up_Lf_P2

Continue…

Jaw	Measurement	Side	Tooth	Trait code used in the analyses
Lower	Incisal-Cervical Distance	Right	Central Incisor	IC_Lw_Rt_I1
			Lateral Incisor	IC_Lw_Rt_I2
			Canine	IC_Lw_Rt_C
			First Premolar	IC_Lw_Rt_P1
			Second Premolar	IC_Lw_Rt_P2
		Left	Central Incisor	IC_Lw_Lf_I1
			Lateral Incisor	IC_Lw_Lf_I2
			Canine	IC_Lw_Lf_C
			First Premolar	IC_Lw_Lf_P1
			Second Premolar	IC_Lw_Lf_P2
	Meso-Distal Distance	Right	Central Incisor	MD_Lw_Rt_I1
			Lateral Incisor	MD_Lw_Rt_I2
			Canine	MD_Lw_Rt_C
			First Premolar	MD_Lw_Rt_P1
			Second Premolar	MD_Lw_Rt_P2
		Left	Central Incisor	MD_Lw_Lf_I1
			Lateral Incisor	MD_Lw_Lf_I2
			Canine	MD_Lw_Lf_C
			First Premolar	MD_Lw_Lf_P1
			Second Premolar	MD_Lw_Lf_P2
	Bucco-Lingual Distance	Right	Central Incisor	BL_Lw_Rt_I1
			Lateral Incisor	BL_Lw_Rt_I2
			Canine	BL_Lw_Rt_C
			First Premolar	BL_Lw_Rt_P1
			Second Premolar	BL_Lw_Rt_P2
		Left	Central Incisor	BL_Lw_Lf_I1
			Lateral Incisor	BL_Lw_Lf_I2
			Canine	BL_Lw_Lf_C
			First Premolar	BL_Lw_Lf_P1
			Second Premolar	BL_Lw_Lf_P2

Table 4.9. Sixty quantitative dental traits assessed in the upper and lower jaw. The name used in the analyses was organized by jaw, type of measurement, side of the maxilla and type of tooth.

4.2.2.4 DNA genotyping and quality control

This is the same procedure described in Section 3.2.4, the only difference is that from the final cleaned genotyping file, only the samples used on these dental traits GWAS were retained (~564), to decrease the size of the files and increase the speed of the analysis.

4.2.2.5 SNP Genotype Imputation

This is the same procedure described in Section 3.2.5.

4.2.2.6 Statistical genetic analyses

4.2.2.6.1 Narrow-sense heritability (h2)

This is the same method described in Section 3.2.6.1.

4.2.2.6.2 Genome Wide Association Analyses

PLINK 1.9 [251] was used to perform both genome-wide association tests for all the phenotypes proposed above (no regrouping, regrouping, breakpoint and composite criteria applied to ASUDAS and metric dental phenotypes) using multiple linear regression with an additive genetic model incorporating age, sex and five genetic PCs as covariates. Association analyses were performed on the imputed data set with two approaches: using the best-guess imputed genotypes in PLINK[251] and using the IMPUTE2 [295] genotype probabilities in SNPTEST v2.5 [298]. Both were consistent with each other and with the results from the chip genotype data. For analysis of the X chromosome an inactivation model was used (male genotypes encoded as 0/2 and female genotypes as 0/1/2). The genetic PCs were obtained using PLINK 1.9 [251] from an LD-pruned dataset of 93,328 SNPs. They were selected by inspecting the proportion of variance explained and checking scree and PC scatter plots. Individual outliers were removed, and PCs re-calculated after each removal. The top PCs appear to be a good proxy for continental ancestry (Figure 4.31).

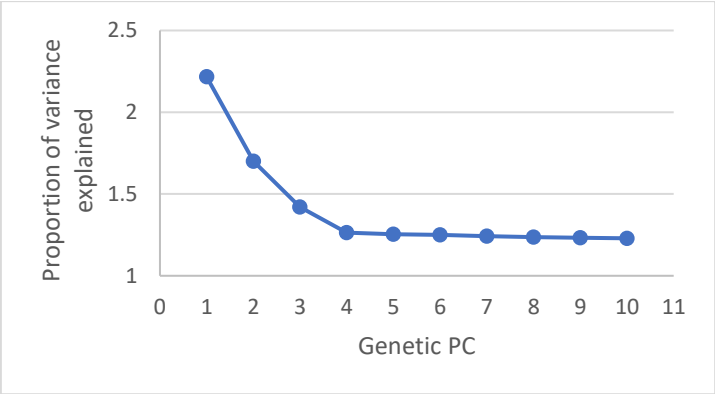


Figure 4.31. Selection of genetic Principal components to be used as covariables in the GWAS analysis. The proportion of the variance explained by each PC. PC 1 explains most of the variation in the sample.

Using these PCs the Q–Q plots (Figure 4.32) for all association tests showed no sign of inflation, the genomic control factor lambda being <1.02 in all cases (Figures B.116-B.176), thus confirming that we are appropriately accounting for population stratification [253] .

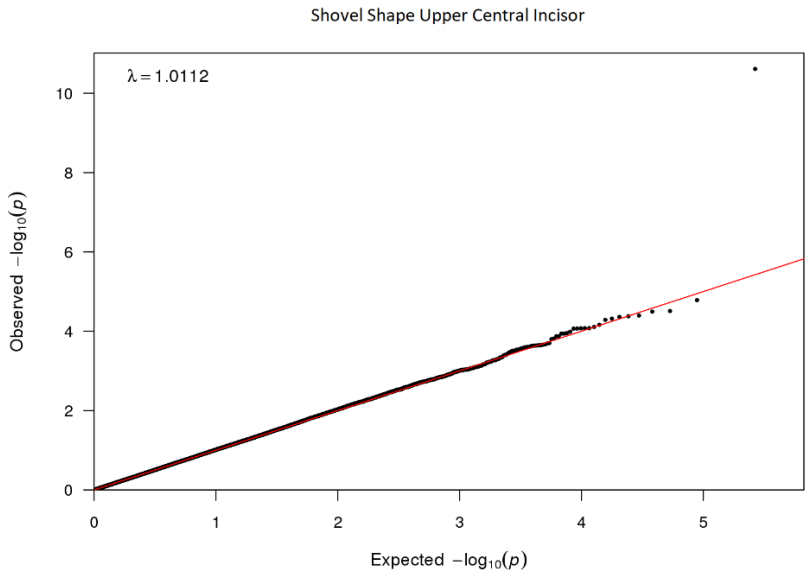


Figure 4.32. Q-Q plot for GWAS of shovel shaped upper incisor scored following the ASUDAS approach. The remaining Q-Q Plots are shown in Appendix B. This plot does not show signs of inflation between the expected and observed P-values. All the traits show a similar pattern, the genomic control factor lambda <1.02, demonstrating there is no population stratification. Similar plots were obtained for all phenotypes examined (Figures B.116-B.176).

4.2.2.6.3 Local Ancestry Inference

A large proportion of significant SNPs in both genome-wide scans results (ASUDAS and dental measurement traits) showed a high frequency in African populations (Appendix B, Table B3 and B4) and therefore rare in the Colombian sample (Figure 3.5). Thus, to check if the haplotypes of the Colombian individuals carrying this allele were indeed African in origin and not the result of an error in genotyping or other issue with the data, some of these SNPs were checked. From a previous local inference analysis performed with RFMix software some of the significant SNPs that were present were checked (rs3827760, rs1037804 and rs9899063). Not all the SNPs were checked because RFMix is very slow, one analysis could take weeks, therefore I just checked the SNPs that were available from this previous analysis.

A merged dataset prepared for the population genetic analysis of the CANDELA samples [59] was phased using SHAPEIT2 [294] to produce best-guess haplotypes for all samples. All samples, both admixed (CANDELA) and references, were used together in this step because the quality of phasing improves with a larger and more diverse dataset.

This phased dataset was sub-setted to prepare the input for RFMix v1 [366]. As RFMix is very slow and memory-intensive, only the Colombian CANDELA samples that were used in the GWAS analysis were retained among the admixed samples. The set of reference Native American samples was smallest in this data merge, only 348, when retaining samples with >95% Native ancestry, and including Native samples from across Latin America. The set of available European and African reference samples was much larger, due to availability of large cohorts such as the 1000 Genomes [293]. Therefore, following the recommendation of RFMix, only the most relevant (geographically proximate) populations were retained to reduce the sample sizes to approximately 350. Only South-West European samples (Spain, Portugal and Italy) and Central-West African samples (Yoruba, etc.) were used.

The phased haplotypes for this subset of samples were divided into chromosomes, and RFMix v1 [366] was run on each chromosome separately. The RFMix software processes windows of haplotypes across the entire dataset to initially produce an assignment of local ancestry for the admixed haplotypes by window, which it then refines through

iterations. The HLA region on chromosome 6 was more difficult for the software to process due to its high recombination rate, and therefore it encountered problems and could not produce an output within the allocated time. The software took several days, several weeks for larger chromosomes, to complete. Upon completion, for each chromosome the software produces local ancestry assignments for each haplotype, at each locus, it gives a most-probable ancestry assignment as 1, 2 or 3 corresponding to one of the three reference populations.

4.3 Results

4.3.1 Study Sample

The cohort used in this study correspond to a sub-sample of Colombian volunteers, which is part of CANDELA [59] (Table 4.2 and Table 4.3).

4.3.2 Ordinal traits (ASUDAS)

Based on the initial phenotyping, forty-six dental traits were assessed on an ordered categorical scale reflecting the distinctiveness of each trait using dental photographs of 501 individuals (Table 4.4). For most of the traits the highest frequency is in the lower categories (0 and 1). With some exceptions, LCVLP2, 3CUM2, and CNLM1, these traits present higher frequencies in upper categories of expression, these features characterize the Eurasian dental complex. Moreover, some traits that show high frequency in African populations (MRUC, DIASUI1 and C7LM1) presented low frequencies in this group (Figure 4.33).

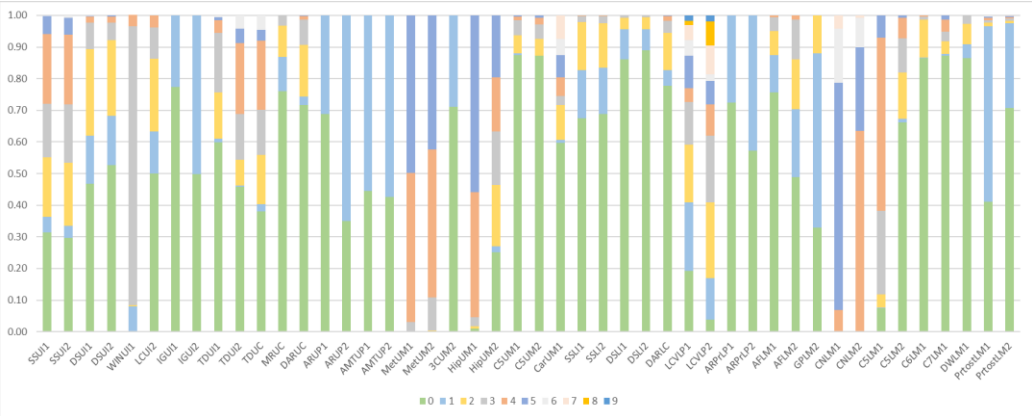


Figure 4.33. Frequency distribution of 46 ordinal dental traits, scored using the ASUDAS system, in the Colombian sample (CANDELA).

4.3.2.1 Rater reliability for ASUDAS scores

The intra-rater reliability rate is shown in Delgado, 2015 [135].

4.3.2.2 Correlations of initial 46 ASUDAS traits

Significant strongest positive correlations were observed between some ordinal phenotypes (using a Bonferroni-adjusted permutation *P* value threshold for significance of 1×10^{-3}, Table 4.10). From moderate to strong positive, and significant, correlations (r = 0.51) were observed between several of the traits, usually the same trait scored in different teeth.

Strongest correlation was observed between shovel shape and shoveling traits (highest correlation, r =0.92) in lateral and central incisors in the upper jaw (highest correlation, r = 0.93). Incisors and canines situated in the upper jaw showed weak to strong significant positive correlations with between shovel shape and double shoveling traits and Tuberculum dentale (TD) and mesial ridge canine (or Bushman canine) traits (from r = 0.17 to r = 0.52). Probably, this is due to both traits being present in the lingual surface of the tooth, and sometimes mesial tubercles merge and form the canine mesial ridge or Bushman canine. Something similar happens with the lower jaw, but the correlations are weaker (Table 4.10).

Cusp 5 on the second lower molar feature (C5LM2) showed a very strong, and positive, correlation with Cusp number of lower molars trait (CNLM) (r = 0.75). This is expected because CNLM depends entirely on the presence of Cusp 5 [374]. Cusp number in the first lower molar (CNLM1) also showed a strong, and positive, correlation with Cusp 6 and 7 in the first lower molar (r = 0.69) (Table 4.10).

The 3-Cusped upper second molar (3CUM2) trait presented a very strong, and negative, correlation with the Hypocone in the second upper molar, this trait correspond to the absence of the fourth cusp located in the distolingual part of the tooth, and its absence is more common in the second molar. Consequently, this negative correlation was totally expected (Table 4.10).

Groove pattern on the second lower molar (GPLM2) showed weak, and negative, correlations with most of the traits, and moderate negative correlations with Cusp number in the second lower molar (CNLM2) and Cusp 5 in the second lower molar (C5LM2) (r = 0.45). The scoring of this trait is based on the contact between different

cusps and the shape of the unions in the lower molars. Therefore, this outcome was expected (Table 4.10).

corr/p-value	SSUI1	SSUI2	DSUI1	DSUI2	WINUI1	LCUI2	TDUI1	TDUI2	TDUC	MRUC	DARUC	MetUM1
SSUI1		**8.6E-199**	**1.4E-50**	**3.4E-45**	1.8E-02	8.3E-02	**3.3E-22**	**4.9E-11**	**1.3E-07**	**1.9E-05**	8.6E-01	5.9E-01
SSUI2	0.92		**7.4E-43**	**1.1E-42**	1.2E-01	3.1E-02	**1.1E-23**	**4.9E-13**	**2.3E-11**	**4.2E-07**	2.8E-01	5.3E-01
DSUI1	0.61	0.57		**1.5E-161**	1.2E-01	5.6E-02	**9.7E-12**	**4.7E-08**	**1.5E-06**	1.0E-04	7.4E-01	6.8E-01
DSUI2	0.59	0.57	0.89		9.8E-02	2.0E-02	**6.8E-14**	**4.0E-08**	**2.9E-06**	**3.0E-07**	8.1E-01	9.7E-01
WINUI1	-0.11	-0.07	-0.07	-0.08		1.8E-01	3.1E-01	2.7E-01	3.5E-01	2.5E-01	6.5E-01	3.9E-01
LCUI2	0.08	0.10	0.09	0.11	-0.06		1.9E-02	5.1E-02	6.0E-04	3.1E-03	1.2E-01	5.3E-01
TDUI1	0.43	0.44	0.31	0.34	0.05	0.11		**2.2E-33**	**7.2E-20**	**3.8E-06**	5.5E-03	6.8E-01
TDUI2	0.30	0.33	0.25	0.25	-0.05	0.09	0.52		**2.2E-28**	2.0E-04	9.2E-02	2.1E-02
TDUC	0.24	0.30	0.22	0.21	-0.04	0.16	0.40	0.48		**1.5E-12**	4.5E-03	3.9E-01
MRUC	0.20	0.23	0.18	0.23	-0.05	0.14	0.21	0.17	0.32		4.4E-02	2.8E-01
DARUC	0.01	0.05	-0.02	0.01	-0.02	0.07	0.13	0.08	0.13	0.09		7.7E-01
MetUM1	0.02	0.03	-0.02	0.00	-0.04	0.03	0.02	0.11	0.04	-0.05	0.01	
MetUM2	0.02	0.03	-0.04	-0.03	-0.03	0.01	0.00	0.02	-0.04	-0.01	-0.02	0.19
HipUM1	0.00	0.01	-0.03	-0.02	-0.05	0.07	-0.04	-0.01	0.02	0.08	-0.01	0.14
HipUM2	0.12	0.11	0.08	0.08	-0.09	0.02	0.08	0.15	0.09	0.12	-0.02	0.05
C5UM1	0.04	0.04	0.02	0.05	0.05	0.02	0.04	0.04	0.06	0.08	0.12	0.05
C5UM2	-0.01	0.00	-0.03	0.02	-0.02	0.04	0.06	0.01	0.08	0.05	0.13	-0.03
CarUM1	-0.04	-0.02	0.02	0.01	0.00	0.01	0.05	0.14	0.14	0.09	0.07	0.09
SSLI1	0.47	0.45	0.38	0.40	-0.08	0.03	0.28	0.24	0.16	0.20	0.03	0.13
SSLI2	0.46	0.45	0.38	0.39	-0.06	0.05	0.31	0.25	0.19	0.19	0.04	0.13
DSLI1	0.30	0.25	0.33	0.32	-0.02	0.02	0.20	0.16	0.13	0.08	-0.02	0.08
DSLI2	0.27	0.23	0.35	0.36	-0.04	0.01	0.23	0.15	0.12	0.10	0.00	0.09
DARLC	0.03	0.07	0.01	0.04	0.04	0.15	0.09	0.09	0.18	0.11	0.21	0.02
LCVLP1	0.07	0.07	0.02	-0.02	-0.01	-0.03	0.13	0.11	0.17	-0.01	0.05	-0.04

corr/p-value	MetUM2	HipUM1	HipUM2	C5UM1	C5UM2	CarUM1	SSLI1	SSLI2	DSLI1	DSLI2	DARLC	LCVLP1
SSUI1	6.3E-01	9.7E-01	8.8E-03	4.3E-01	9.0E-01	4.3E-01	**5.9E-28**	**6.5E-26**	**2.1E-11**	**1.3E-09**	4.6E-01	1.7E-01
SSUI2	4.5E-01	7.6E-01	1.3E-02	3.6E-01	9.6E-01	6.6E-01	**5.2E-25**	**6.7E-25**	**2.1E-08**	**3.5E-07**	1.2E-01	1.6E-01
DSUI1	3.5E-01	5.3E-01	7.2E-02	6.3E-01	5.8E-01	7.2E-01	**3.4E-18**	**1.2E-17**	**2.3E-13**	**1.4E-15**	8.3E-01	6.7E-01
DSUI2	4.8E-01	6.1E-01	9.2E-02	2.4E-01	7.1E-01	8.2E-01	**4.0E-19**	**6.3E-19**	**7.1E-13**	**2.5E-16**	3.5E-01	7.2E-01
WINUI1	5.2E-01	2.7E-01	5.6E-02	2.9E-01	6.9E-01	9.6E-01	9.5E-02	1.6E-01	7.1E-01	3.3E-01	3.6E-01	8.9E-01
LCUI2	8.9E-01	1.3E-01	7.2E-01	7.4E-01	3.4E-01	8.8E-01	4.5E-01	2.8E-01	6.1E-01	8.4E-01	1.2E-03	5.0E-01
TDUI1	9.4E-01	3.6E-01	7.6E-02	4.3E-01	1.7E-01	2.5E-01	**6.4E-10**	**9.3E-12**	**1.4E-05**	**7.1E-07**	5.3E-02	7.5E-03
TDUI2	7.3E-01	8.8E-01	1.3E-03	3.6E-01	8.1E-01	1.7E-03	**2.4E-07**	**5.2E-08**	**3.5E-04**	1.2E-03	5.8E-02	2.1E-02
TDUC	3.9E-01	6.9E-01	5.2E-02	1.7E-01	1.0E-01	2.4E-03	5.6E-04	**4.2E-05**	5.8E-03	6.5E-03	9.7E-05	4.3E-04
MRUC	8.4E-01	9.2E-02	6.8E-03	6.8E-02	2.7E-01	5.9E-02	**1.2E-05**	**2.1E-05**	6.5E-02	3.3E-02	1.7E-02	7.8E-01
DARUC	6.5E-01	7.5E-01	6.5E-01	8.6E-03	4.8E-03	1.2E-01	5.5E-01	4.2E-01	6.2E-01	9.2E-01	**5.4E-06**	2.7E-01
MetUM1	**3.7E-05**	1.6E-03	3.1E-01	3.2E-01	4.8E-01	3.9E-02	3.8E-03	4.8E-03	7.1E-02	4.7E-02	6.0E-01	4.6E-01
MetUM2		5.0E-01	7.7E-03	3.0E-01	1.7E-01	2.0E-01	7.3E-02	1.1E-01	4.9E-01	6.9E-01	9.0E-01	1.8E-01
HipUM1	0.03		**3.0E-08**	4.9E-03	7.8E-02	**2.5E-07**	1.4E-01	1.5E-01	6.6E-01	3.6E-01	2.5E-01	4.9E-01
HipUM2	0.12	0.25		2.2E-02	1.8E-02	**8.7E-06**	1.3E-01	1.2E-01	5.5E-01	5.0E-01	1.8E-03	2.5E-01
C5UM1	0.05	0.13	0.11		**7.6E-15**	5.3E-02	2.8E-02	3.2E-02	9.3E-01	9.4E-01	3.7E-01	8.2E-01
C5UM2	0.06	0.08	0.11	0.35		9.0E-01	8.4E-01	9.5E-01	2.1E-01	4.0E-01	7.4E-01	2.2E-01
CarUM1	0.06	0.23	0.20	0.09	0.01		3.8E-01	5.7E-01	4.8E-01	4.1E-01	1.6E-01	3.4E-01
SSLI1	0.08	0.07	0.07	0.10	0.01	0.04		**4.2E-214**	**1.2E-33**	**1.1E-30**	2.8E-01	7.5E-01
SSLI2	0.07	0.07	0.07	0.10	0.00	0.03	0.93		**5.3E-34**	**6.0E-36**	1.9E-01	7.6E-01
DSLI1	-0.03	0.02	0.03	0.00	-0.06	0.03	0.51	0.52		**9.6E-146**	4.6E-01	5.9E-01
DSLI2	-0.02	0.04	0.03	0.00	-0.04	0.04	0.49	0.53	0.87		3.5E-01	9.2E-01
DARLC	-0.01	0.05	0.14	0.04	0.02	0.06	0.05	0.06	0.03	0.04		4.7E-01
LCVLP1	0.06	0.03	0.05	0.01	-0.06	0.05	0.01	0.01	0.03	0.00	-0.03	

Continue…

corr/p-value	LCVLP2	AFLM1	AFLM2	GPLM2	CNLM1	CNLM2	C5LM1	C5LM2	C6LM1	C7LM1	DWLM1
SSUI1	4.0E-01	3.2E-01	8.4E-01	9.7E-02	1.4E-01	5.4E-04	7.5E-01	2.2E-04	3.5E-01	3.8E-01	2.1E-01
SSUI2	4.3E-01	6.3E-01	5.0E-01	1.2E-01	1.4E-01	1.5E-03	5.7E-01	1.0E-03	3.4E-01	2.7E-01	6.9E-02
DSUI1	2.9E-02	9.0E-01	4.0E-01	3.1E-01	8.2E-01	1.2E-01	2.2E-01	2.1E-01	5.4E-01	7.2E-01	4.9E-01
DSUI2	7.3E-02	7.5E-01	1.5E-01	8.5E-02	1.7E-01	4.4E-02	4.9E-01	1.4E-01	2.4E-01	5.9E-01	1.6E-01
WINUI1	5.6E-01	4.4E-01	8.9E-01	8.6E-01	1.6E-01	5.9E-01	8.2E-01	3.4E-01	1.4E-01	3.3E-01	4.9E-01
LCUI2	6.4E-02	3.1E-03	**2.7E-06**	1.7E-02	8.0E-03	4.6E-01	9.5E-01	2.6E-01	1.0E-01	2.9E-03	1.9E-01
TDUI1	1.6E-02	2.3E-01	1.3E-01	3.2E-04	1.7E-01	3.3E-02	6.6E-01	2.5E-02	1.1E-01	6.8E-01	2.8E-01
TDUI2	8.2E-02	2.6E-01	2.1E-02	2.4E-02	3.3E-01	2.1E-01	2.6E-01	6.1E-02	9.5E-01	2.2E-01	6.4E-01
TDUC	2.0E-04	2.0E-03	9.4E-04	6.5E-02	5.8E-03	7.3E-01	6.1E-01	8.5E-01	8.2E-02	2.5E-02	7.0E-03
MRUC	1.2E-01	2.2E-02	5.0E-02	8.9E-02	1.5E-01	8.7E-02	4.5E-02	2.3E-01	8.7E-01	7.1E-01	1.2E-01
DARUC	6.4E-01	3.3E-03	2.5E-03	9.8E-01	3.7E-02	1.1E-02	8.8E-03	4.8E-02	3.7E-01	1.4E-01	5.2E-03
MetUM1	2.6E-01	3.1E-02	5.3E-03	5.4E-01	3.3E-02	3.8E-01	1.5E-02	4.6E-01	3.1E-01	2.4E-01	5.7E-02
MetUM2	4.4E-01	3.1E-01	4.8E-01	6.9E-02	3.7E-01	8.2E-01	2.8E-02	2.7E-01	8.1E-01	2.1E-01	8.3E-01
HipUM1	2.2E-01	3.7E-01	8.2E-01	3.8E-01	1.1E-04	2.5E-01	**1.2E-06**	4.4E-01	1.2E-01	1.6E-02	6.0E-01
HipUM2	1.8E-01	4.7E-01	7.0E-01	5.5E-01	6.5E-01	2.0E-01	2.2E-04	3.7E-03	6.6E-01	5.8E-02	7.4E-01
C5UM1	5.5E-01	3.5E-01	2.9E-01	3.7E-03	1.0E-01	2.6E-02	2.5E-04	2.8E-02	6.6E-01	9.6E-01	1.8E-02
C5UM2	3.3E-01	4.2E-01	5.7E-02	1.8E-02	5.1E-02	3.7E-03	8.9E-05	6.4E-04	6.7E-01	5.2E-01	9.6E-04
CarUM1	2.4E-02	4.1E-01	2.8E-01	1.5E-01	1.7E-01	3.5E-01	9.8E-02	7.2E-01	2.7E-02	7.1E-01	9.6E-01
SSLI1	2.1E-01	4.7E-01	6.6E-01	2.2E-01	4.9E-01	3.1E-02	3.1E-01	7.8E-03	3.4E-01	2.4E-01	7.6E-01
SSLI2	3.0E-01	8.9E-01	5.7E-01	1.7E-01	3.6E-01	2.2E-02	2.3E-01	3.1E-03	2.9E-01	3.1E-01	9.9E-01
DSLI1	2.7E-01	4.3E-01	6.8E-01	3.9E-01	6.8E-01	3.6E-01	1.3E-01	1.2E-01	2.9E-01	1.8E-01	6.2E-01
DSLI2	3.0E-01	8.4E-01	9.4E-01	5.4E-01	9.6E-01	6.0E-01	6.2E-01	2.5E-01	3.8E-01	1.7E-01	5.6E-01
DARLC	7.7E-01	8.7E-02	1.7E-02	5.7E-03	9.5E-03	7.2E-03	1.4E-02	2.5E-03	2.6E-01	5.7E-02	5.5E-01
LCVLP1	**4.2E-17**	3.5E-01	5.7E-01	1.2E-02	8.0E-01	2.0E-01	7.0E-01	4.6E-01	2.1E-01	5.3E-01	4.0E-01

Continue…

corr/p-value	PrtostLM1	PrtostLM2	IGUI1	IGUI2	ARUP1	ARUP2	AMTUP1	AMTUP2	3CUM2	ARPrLP1	ARPrLP2
SSUI1	2.4E-03	3.0E-01	**7.8E-06**	**2.5E-06**	5.0E-01	4.8E-03	4.7E-02	6.9E-01	3.5E-02	1.2E-02	1.3E-03
SSUI2	2.0E-03	1.4E-01	**7.4E-06**	**1.5E-08**	4.0E-02	3.8E-03	8.6E-02	4.2E-01	5.0E-02	2.8E-03	2.3E-03
DSUI1	4.7E-02	1.8E-01	1.7E-02	1.6E-04	4.6E-01	2.6E-01	2.6E-01	7.8E-01	2.5E-01	1.0E-01	2.0E-01
DSUI2	7.1E-02	8.4E-02	9.1E-04	**1.4E-05**	5.0E-01	2.7E-01	2.8E-02	5.5E-01	4.1E-01	6.3E-02	1.8E-01
WINUI1	8.7E-01	6.9E-01	3.1E-01	7.7E-01	1.7E-01	8.7E-01	5.0E-01	3.0E-01	6.7E-01	3.9E-01	2.3E-01
LCUI2	1.2E-01	1.0E+00	3.7E-01	6.0E-02	1.4E-02	1.1E-01	7.0E-01	2.0E-01	4.5E-01	1.5E-01	4.0E-01
TDUI1	2.0E-03	1.0E-03	**1.3E-19**	1.8E-04	2.4E-01	2.0E-01	1.3E-02	3.5E-01	1.4E-01	8.8E-03	7.9E-03
TDUI2	4.6E-02	1.9E-03	5.8E-05	1.5E-02	9.5E-02	1.5E-01	1.7E-01	7.6E-01	1.6E-02	1.6E-01	6.1E-01
TDUC	5.9E-02	1.5E-03	**9.6E-07**	**1.8E-05**	9.3E-02	4.3E-01	2.0E-01	1.1E-02	7.4E-02	1.3E-01	1.2E-01
MRUC	9.6E-02	2.9E-01	1.9E-02	2.0E-04	6.4E-01	3.6E-01	8.8E-02	2.9E-01	8.2E-03	7.8E-01	3.5E-01
DARUC	8.8E-01	2.8E-01	2.6E-04	4.8E-02	2.2E-02	2.2E-02	1.7E-01	3.2E-01	5.4E-01	1.6E-03	**2.6E-06**
MetUM1	6.6E-01	9.6E-01	2.8E-01	9.4E-01	2.8E-01	1.2E-01	3.6E-01	7.8E-01	2.1E-01	4.2E-01	1.6E-02
MetUM2	2.5E-01	6.5E-01	2.5E-01	3.6E-01	7.4E-01	9.8E-01	7.0E-01	7.2E-01	6.1E-02	9.8E-01	2.4E-01
HipUM1	1.4E-01	2.9E-01	5.5E-01	8.0E-01	2.6E-01	1.2E-01	3.7E-01	1.9E-02	**3.2E-06**	2.4E-01	7.4E-02
HipUM2	7.1E-02	3.2E-02	9.9E-01	5.9E-01	8.5E-01	8.1E-02	9.4E-01	9.8E-01	**4.1E-88**	9.9E-01	9.6E-01
C5UM1	9.2E-01	7.5E-01	3.0E-01	9.6E-01	1.9E-02	1.7E-02	6.7E-01	9.3E-01	2.3E-03	3.8E-01	7.8E-01
C5UM2	6.2E-01	4.2E-04	4.2E-02	4.1E-01	3.4E-03	2.7E-02	3.3E-01	7.8E-01	7.9E-03	1.3E-01	2.4E-01
CarUM1	4.6E-04	**2.4E-05**	3.3E-01	5.7E-02	1.3E-01	1.9E-01	7.7E-01	7.8E-01	1.1E-04	1.6E-01	1.0E-01
SSLI1	8.8E-03	2.5E-01	5.2E-05	**2.4E-06**	1.3E-01	1.1E-02	3.4E-02	4.7E-01	1.7E-01	3.5E-02	1.5E-03
SSLI2	2.2E-02	5.0E-01	1.0E-04	**1.4E-06**	9.2E-02	1.7E-02	8.7E-02	4.0E-01	2.5E-01	5.0E-02	2.7E-03
DSLI1	2.7E-04	5.5E-01	7.0E-02	9.1E-03	8.9E-01	1.1E-01	2.6E-02	5.7E-01	6.6E-01	7.9E-01	1.3E-01
DSLI2	6.8E-04	8.1E-01	5.1E-02	1.0E-02	6.6E-01	8.5E-02	2.4E-01	9.7E-01	8.8E-01	6.2E-01	4.0E-01
DARLC	8.4E-01	4.8E-01	6.4E-03	1.2E-02	3.7E-04	3.4E-03	7.7E-01	4.2E-01	2.2E-03	1.2E-02	1.0E-02
LCVLP1	6.0E-01	3.4E-01	6.1E-01	1.0E+00	4.1E-01	1.6E-01	8.4E-01	3.1E-01	1.2E-01	1.3E-01	8.1E-02

Continue…

corr/p-value	SSUI1	SSUI2	DSUI1	DSUI2	WINUI1	LCUI2	TDUI1	TDUI2	TDUC	MRUC	DARUC	MetUM1
LCVLP2	0.04	0.04	0.10	0.08	0.03	0.09	0.11	0.08	0.17	0.07	0.02	-0.05
AFLM1	-0.05	-0.02	0.01	0.01	0.04	0.14	0.06	0.05	0.14	0.11	0.13	0.10
AFLM2	0.01	0.03	0.04	0.07	-0.01	0.21	0.07	0.11	0.15	0.09	0.14	0.13
GPLM2	-0.08	-0.07	-0.05	-0.08	-0.01	-0.11	-0.17	-0.10	-0.09	-0.08	0.00	0.03
CNLM1	0.07	0.07	0.01	0.06	0.06	0.12	0.06	0.05	0.13	0.07	0.10	0.10
CNLM2	0.16	0.15	0.07	0.09	-0.03	-0.03	0.10	0.06	0.02	0.08	0.12	0.04
C5LM1	0.01	0.03	-0.06	-0.03	-0.01	0.00	-0.02	0.05	0.02	0.09	0.12	0.11
C5LM2	0.17	0.15	0.06	0.07	-0.04	-0.05	0.10	0.09	0.01	0.06	0.09	0.03
C6LM1	0.04	0.04	0.03	0.05	0.07	0.08	0.07	0.00	0.08	0.01	0.04	0.05
C7LM1	0.04	0.05	-0.02	0.02	0.05	0.14	0.02	0.06	0.10	0.02	0.07	0.05
DWLM1	0.06	0.08	0.03	0.06	0.03	0.06	0.05	0.02	0.12	0.07	0.13	0.09
PrtostLM1	0.14	0.14	0.09	0.08	-0.01	0.07	0.14	0.09	0.09	0.08	-0.01	-0.02
PrtostLM2	0.05	0.07	0.06	0.08	0.02	0.00	0.15	0.14	0.15	0.05	0.05	0.00
IGUI1	0.20	0.20	0.11	0.15	0.05	0.04	0.40	0.18	0.22	0.11	0.17	0.05
IGUI2	0.22	0.26	0.17	0.20	-0.01	0.09	0.17	0.11	0.20	0.17	0.09	0.00
ARUP1	0.03	0.10	0.04	0.03	0.07	0.12	0.06	0.08	0.08	0.02	0.11	0.05
ARUP2	0.13	0.14	0.05	0.05	0.01	0.07	0.06	0.07	0.04	0.04	0.11	0.07
AMTUP1	0.10	0.08	0.05	0.11	-0.03	0.02	0.12	0.07	0.06	0.08	0.07	0.04
AMTUP2	0.02	0.04	-0.01	0.03	-0.05	0.06	0.04	-0.01	0.12	0.05	0.05	0.01
3CUM2	-0.10	-0.09	-0.05	-0.04	0.02	0.04	-0.07	-0.11	-0.08	-0.12	-0.03	-0.06
ARPrLP1	0.12	0.14	0.08	0.09	0.04	0.07	0.12	0.07	0.07	0.01	0.15	0.04
ARPrLP2	0.15	0.14	0.06	0.06	0.06	0.04	0.12	-0.02	0.07	0.04	0.22	0.11

Continue…

corr/p-value	MetUM2	HipUM1	HipUM2	C5UM1	C5UM2	CarUM1	SSLI1	SSLI2	DSLI1	DSLI2	DARLC	LCVLP1
LCVLP2	0.03	0.01	0.02	0.05	0.09	0.05	0.02	0.03	-0.02	0.00	0.11	-0.03
AFLM1	-0.08	0.04	-0.03	-0.13	-0.11	0.07	-0.06	-0.06	-0.04	-0.03	-0.13	-0.12
AFLM2	-0.04	0.18	0.02	0.08	0.09	0.06	0.03	0.04	-0.02	0.00	0.12	0.01
GPLM2	0.01	0.05	0.06	0.10	0.13	-0.04	0.10	0.11	0.04	0.02	0.12	0.06
CNLM1	0.10	0.22	0.17	0.17	0.18	0.08	0.05	0.06	-0.07	-0.02	0.11	-0.02
CNLM2	0.05	0.04	0.13	0.10	0.16	-0.02	0.12	0.14	0.07	0.05	0.14	0.04
C5LM1	-0.01	0.07	0.02	0.02	0.02	0.10	0.04	0.05	0.05	0.04	0.05	0.06
C5LM2	-0.06	0.11	-0.09	0.00	0.03	-0.02	-0.05	-0.05	-0.06	-0.06	0.09	-0.03
C6LM1	0.01	0.02	0.02	0.11	0.15	0.00	0.01	0.00	0.02	0.03	0.03	-0.04
C7LM1	-0.05	0.07	0.08	0.00	0.02	0.16	0.12	0.11	0.17	0.16	-0.01	-0.02
DWLM1	0.02	0.05	0.10	0.01	0.16	0.19	0.05	0.03	0.03	0.01	0.03	0.05
PrtostLM1	0.05	0.03	0.00	0.05	0.09	0.04	0.18	0.18	0.08	0.09	0.13	0.02
PrtostLM2	-0.04	0.01	-0.02	0.00	0.04	0.09	0.22	0.22	0.12	0.12	0.12	0.00
IGUI1	0.02	0.05	0.01	0.11	0.14	0.07	0.07	0.08	-0.01	0.02	0.17	-0.04
IGUI2	0.00	0.07	0.08	0.11	0.10	0.06	0.12	0.11	0.08	0.08	0.14	0.07
ARUP1	-0.02	0.04	0.00	0.02	0.05	0.01	0.10	0.08	0.11	0.06	-0.01	-0.01
ARUP2	-0.02	0.11	0.00	0.00	0.01	0.01	0.03	0.04	-0.03	0.00	0.04	0.05
AMTUP1	-0.09	-0.21	-0.75	-0.14	-0.12	-0.18	-0.06	-0.05	-0.02	0.01	-0.14	-0.07
AMTUP2	0.00	0.06	0.00	0.04	0.07	0.07	0.10	0.09	0.01	-0.02	0.12	0.07
3CUM2	0.05	0.08	0.00	0.01	0.05	0.08	0.15	0.14	0.07	0.04	0.12	0.08
ARPrLP1	-0.03	0.02	0.03	0.00	-0.06	0.03	0.51	0.52		**9.6E-146**	4.6E-01	5.9E-01
ARPrLP2	-0.02	0.04	0.03	0.00	-0.04	0.04	0.49	0.53	0.87		3.5E-01	9.2E-01

Continue…

corr/p-value	LCVLP2	AFLM1	AFLM2	GPLM2	CNLM1	CNLM2	C5LM1	C5LM2	C6LM1	C7LM1	DWLM1
LCVLP2	0.06	0.51		4.8E-03	2.3E-02	5.4E-01	4.6E-02	3.9E-01	9.2E-01	1.2E-02	1.6E-03
AFLM1	-0.13	-0.10	-0.13		2.1E-02	**3.4E-22**	2.9E-02	**7.6E-27**	8.2E-02	2.4E-01	8.0E-01
AFLM2	0.08	0.15	0.10	-0.11		**7.1E-10**	**6.0E-15**	2.3E-03	**3.6E-68**	**2.2E-55**	1.4E-01
GPLM2	0.04	-0.01	0.03	-0.43	0.28		4.3E-04	**8.9E-87**	**4.1E-09**	4.7E-03	5.7E-01
CNLM1	0.00	0.04	0.09	-0.10	0.35	0.16		**5.2E-07**	3.7E-01	3.9E-01	4.8E-01
CNLM2	0.04	-0.03	0.04	-0.47	0.14	0.75	0.23		5.5E-03	6.7E-01	6.6E-01
C5LM1	0.08	0.12	0.00	-0.08	0.69	0.27	-0.04	0.13		6.3E-05	6.3E-02
C5LM2	0.03	0.14	0.12	-0.05	0.64	0.13	0.04	0.02	0.18		8.4E-01
C6LM1	0.00	0.14	0.15	-0.01	0.07	-0.03	0.03	-0.02	0.09	-0.01	
C7LM1	0.05	-0.08	0.03	-0.02	-0.03	0.01	-0.05	0.02	-0.01	-0.01	-0.01
DWLM1	0.07	-0.01	0.03	-0.09	0.02	0.16	0.08	0.19	0.05	-0.06	0.13
PrtostLM1	0.01	0.02	0.09	-0.02	0.08	0.02	0.10	0.05	0.03	0.05	0.01
PrtostLM2	-0.01	0.02	0.07	0.02	0.10	-0.04	0.06	0.00	0.02	0.08	0.02
IGUI1	0.06	0.12	0.25	-0.04	0.05	0.02	0.04	0.06	0.03	0.07	0.16
IGUI2	0.10	0.03	0.09	-0.06	0.10	0.07	0.05	0.07	0.05	0.09	0.14
ARUP1	0.03	0.02	0.22	-0.06	0.06	0.00	0.06	0.02	-0.05	0.10	-0.04
ARUP2	0.04	0.02	0.17	-0.05	0.01	-0.03	0.04	-0.04	-0.06	0.06	0.01
AMTUP1	-0.07	0.06	0.04	-0.04	-0.05	-0.07	-0.19	-0.14	-0.03	0.06	0.00
AMTUP2	0.13	0.05	0.10	-0.12	0.13	0.15	0.04	0.08	0.10	0.13	0.05
3CUM2	0.15	0.07	0.16	-0.08	0.05	0.08	0.07	0.10	-0.02	0.05	0.08

Continue…

corr/p-value	PrtostLM1	PrtostLM2	IGUI1	IGUI2	ARUP1	ARUP2	AMTUP1	AMTUP2	3CUM2	ARPrLP1	ARPrLP2
LCVLP2	5.5E-01	4.5E-01	4.7E-02	1.1E-01	**1.1E-07**	5.8E-02	**2.8E-06**	3.1E-04	3.3E-01	3.2E-02	6.8E-04
AFLM1	6.6E-01	4.1E-02	6.9E-01	5.9E-01	3.6E-01	2.0E-01	2.5E-01	3.1E-01	3.7E-01	1.3E-02	1.0E-01
AFLM2	4.9E-01	6.1E-01	7.3E-02	3.9E-02	3.0E-01	3.6E-02	2.3E-01	7.6E-01	3.1E-01	6.9E-03	2.9E-01
GPLM2	8.9E-01	5.5E-04	6.3E-01	4.2E-01	6.5E-01	1.2E-01	9.9E-01	5.1E-01	1.3E-01	1.4E-03	9.7E-02
CNLM1	2.6E-01	8.3E-02	2.3E-02	2.3E-01	4.3E-01	3.0E-01	2.3E-01	3.9E-01	**4.2E-05**	3.8E-01	1.2E-01
CNLM2	6.8E-01	**3.7E-05**	2.8E-01	9.5E-01	2.4E-01	1.1E-01	6.8E-01	3.9E-01	2.7E-03	1.0E-01	2.7E-02
C5LM1	7.9E-01	2.6E-01	4.5E-01	6.1E-01	5.6E-01	3.3E-01	2.8E-01	2.3E-01	5.4E-01	4.1E-02	7.2E-01
C5LM2	7.7E-01	2.0E-01	2.8E-01	6.8E-02	1.6E-01	5.9E-02	3.1E-02	2.2E-01	2.0E-01	7.6E-03	2.7E-01
C6LM1	8.0E-01	4.4E-03	8.8E-01	7.3E-01	1.1E-03	2.1E-03	3.6E-01	8.4E-01	9.7E-01	2.6E-01	1.0E-01
C7LM1		**1.3E-07**	8.1E-01	1.9E-02	9.3E-01	7.7E-01	**9.4E-04**	4.9E-01	1.4E-01	4.6E-01	5.8E-01
DWLM1	0.24		3.7E-02	1.4E-01	2.6E-01	6.1E-01	6.7E-01	7.8E-01	1.1E-01	5.8E-01	1.1E-01
PrtostLM1	-0.01	0.10		**2.5E-08**	5.9E-02	8.8E-01	1.9E-01	8.3E-01	9.4E-01	2.9E-02	1.3E-02
PrtostLM2	0.11	0.07	0.25		1.7E-01	1.0E-01	4.0E-01	3.3E-01	8.4E-01	8.2E-02	7.2E-01
IGUI1	0.00	0.05	0.09	0.07		**4.6E-21**	4.0E-01	3.5E-01	9.4E-01	**4.4E-06**	**2.2E-06**
IGUI2	0.01	0.02	0.01	0.08	0.44		2.0E-01	1.9E-01	2.1E-02	**6.9E-06**	**3.2E-09**
ARUP1	0.16	0.02	0.06	0.04	0.04	0.06		**2.0E-24**	8.2E-01	2.1E-02	1.5E-01
ARUP2	0.03	-0.01	0.01	0.05	0.05	0.06	0.47		5.9E-01	1.3E-01	2.0E-01
AMTUP1	-0.07	-0.07	0.00	-0.01	0.00	-0.11	0.01	0.03		3.9E-01	2.8E-01
AMTUP2	-0.04	0.03	0.10	0.08	0.22	0.21	0.11	0.07	-0.04		**1.5E-34**
3CUM2	0.03	0.07	0.12	0.02	0.23	0.27	0.07	0.06	-0.05	0.54	

Table 4.10. Simple correlation between 46 ordinal traits. Correlation values are presented in the lower left triangle, with corresponding permutation P values in the upper right triangle. Correlations with significant P values (<0.0005, Bonferroni-adjusted threshold) and their corresponding P values are highlighted in bold.

Several correlations between the 46 dental features and covariates were weak (Table 4.11a and b), but most of them significant (r values from 0.10 to 0.21, p-values $8.3x10^{-2}$ and $7x10^{-16}$, respectively). Negative correlations were observed between dental traits and age, sex and ancestry. Most of the traits related to cusps, ridges and foveae, which are structures susceptible to wear, showed a negative correlation with age (DSUI2, SSLI1, SSLI2, DSLI1, MetUM1, CNLM2, AFLM2, C6LM1, etc.). Some authors have suggested that despite the consumption of softer and processed foods, age has an important effect on dental wear [135]. All the traits observed in canines, showed negative correlation with sex. These teeth present the highest sexual dimorphism within teeth in humans, hominis and nonhuman primates [177]. Sixteen traits showed positive correlations with Native-American ancestry, and negative correlations with European ancestry (SSUI1, SSUI2, DSUI1, DSUI2, MetUM1, MetUM2, SLI1, SLI2, DSLI1, DSLI2, AFLM1, AFLM2, CNLM2, C5LM2, C6LM1 and 3CUM2). Finally, C5UM1, CNLM1, C6LM1 and C7LM1 presented low, and positive, correlations with African ancestry (Table 4.11a-Table 4.11b).

Covariate	Sex	Age	Africa	Europe	Native
SSUI1	0.01	-0.07	-0.01	-0.11	0.15
SSUI2	0.00	-0.08	0.01	-0.13	0.16
DSUI1	-0.03	-0.04	-0.08	-0.06	0.15
DSUI2	-0.02	-0.06	-0.05	-0.09	0.18
WINUI1	-0.02	0.03	-0.01	0.06	-0.07
LCUI2	-0.06	-0.02	0.05	-0.10	0.08
TDUI1	-0.14	0.05	0.03	-0.02	-0.01
TDUI2	-0.11	0.02	0.06	-0.06	0.02
TDUC	-0.18	0.02	0.11	-0.08	0.00
MRUC	-0.13	-0.02	0.03	-0.08	0.07
DARUC	-0.11	-0.16	0.02	-0.06	0.05
MetUM1	0.10	-0.08	0.07	-0.16	0.15
MetUM2	-0.01	-0.13	0.03	-0.14	0.16
HipUM1	-0.04	-0.14	0.08	-0.06	0.00
HipUM2	-0.06	-0.06	0.08	-0.02	-0.05
C5UM1	-0.02	-0.09	0.11	-0.09	0.00
C5UM2	-0.05	-0.04	0.13	-0.06	-0.06
CarUM1	-0.05	-0.04	-0.02	0.06	-0.06
SSLI1	0.04	-0.06	0.02	-0.17	0.21
SSLI2	0.04	-0.05	0.02	-0.17	0.21
DSLI1	0.01	0.05	0.02	-0.12	0.14
DSLI2	0.02	0.02	0.01	-0.12	0.15
DARLC	-0.13	-0.06	0.01	-0.02	0.01
LCVLP1	-0.01	-0.07	0.07	-0.02	-0.05
LCVLP2	0.01	-0.05	0.06	-0.03	-0.02
AFLM1	0.00	-0.09	0.07	-0.11	0.08
AFLM2	-0.02	-0.23	-0.02	-0.09	0.14
GPLM2	0.01	0.05	-0.03	0.06	-0.05
CNLM1	-0.06	-0.05	0.13	-0.15	0.07
CNLM2	-0.08	0.01	0.08	-0.15	0.13
C5LM1	-0.04	-0.10	0.07	-0.06	0.01
C5LM2	-0.04	-0.02	0.07	-0.14	0.12
C6LM1	-0.02	0.04	0.11	-0.15	0.09
C7LM1	-0.11	-0.08	0.10	-0.09	0.02
DWLM1	0.06	-0.10	0.06	-0.09	0.06
PrtostLM1	-0.06	0.18	0.00	-0.01	0.01
PrtostLM2	-0.09	0.01	0.05	-0.06	0.03
IGUI1	-0.11	-0.06	0.01	-0.04	0.05
IGUI2	-0.09	0.04	0.00	0.00	0.00
ARUP1	-0.03	-0.19	0.08	-0.10	0.05
ARUP2	-0.04	-0.17	0.05	-0.04	0.00
AMTUP1	-0.04	-0.14	0.01	-0.04	0.04
AMTUP2	-0.09	-0.15	0.06	-0.03	-0.02
3CUM2	0.07	0.07	-0.08	-0.02	0.10
ARPrLP1	0.00	-0.13	0.07	-0.04	-0.02
ARPrLP2	-0.03	-0.16	0.08	-0.08	0.03

Table 4.11a. Correlation between 46 ordinal dental traits and age, sex, and continental ancestry. Correlation values are presented in Table a), with corresponding *P* values in Table b). Correlations with significant *P* values (<0.0002, Bonferroni-adjusted threshold), are highlighted in bold. Anc.= Continental ancestry estimated from the genetic data. Sex coded as female=1, male=0.

p-value	Sex	Age	Africa	Europe	Native
SSUI1	1.3E-01	7.8E-01	**2.1E-02**	**7.6E-04**	**2.4E-16**
SSUI2	**7.6E-02**	8.6E-01	**5.8E-03**	**4.4E-04**	**1.7E-16**
DSUI1	3.5E-01	**9.0E-02**	2.2E-01	**7.3E-04**	**3.4E-07**
DSUI2	1.7E-01	2.6E-01	**4.2E-02**	**1.0E-04**	**2.4E-08**
WINUI1	4.8E-01	8.7E-01	2.0E-01	1.3E-01	5.1E-01
LCUI2	7.3E-01	3.0E-01	**3.8E-02**	**8.5E-02**	9.0E-01
TDUI1	2.5E-01	4.8E-01	6.6E-01	8.9E-01	**1.5E-02**
TDUI2	6.4E-01	1.9E-01	1.9E-01	6.8E-01	**2.1E-02**
TDUC	7.3E-01	**1.7E-02**	**7.6E-02**	9.7E-01	**9.0E-02**
MRUC	6.8E-01	4.6E-01	1.0E-01	1.5E-01	**7.7E-02**
DARUC	**7.0E-04**	6.3E-01	2.1E-01	2.4E-01	7.2E-01
MetUM1	**8.3E-02**	1.2E-01	**3.3E-04**	**1.4E-03**	2.2E-01
MetUM2	**4.7E-03**	4.8E-01	**1.7E-03**	**4.8E-04**	**7.0E-03**
HipUM1	**2.0E-03**	**8.5E-02**	2.0E-01	9.7E-01	5.4E-01
HipUM2	1.9E-01	**7.4E-02**	6.0E-01	2.7E-01	**2.2E-02**
C5UM1	**5.0E-02**	**1.5E-02**	**6.2E-02**	9.7E-01	7.5E-01
C5UM2	3.4E-01	**3.3E-03**	2.2E-01	1.9E-01	4.2E-01
CarUM1	4.4E-01	6.6E-01	2.1E-01	2.1E-01	2.1E-01
SSLI1	1.6E-01	6.1E-01	**1.3E-04**	**4.0E-06**	**7.0E-16**
SSLI2	2.4E-01	7.2E-01	**2.2E-04**	**4.3E-06**	**1.8E-14**
DSLI1	3.1E-01	6.0E-01	**6.4E-03**	**1.8E-03**	**2.8E-03**
DSLI2	6.7E-01	8.4E-01	**1.1E-02**	**1.4E-03**	**2.8E-03**
DARLC	1.6E-01	7.8E-01	6.9E-01	7.9E-01	4.9E-01
LCVLP1	1.7E-01	1.3E-01	7.3E-01	2.8E-01	5.8E-01
LCVLP2	2.6E-01	1.8E-01	4.8E-01	6.8E-01	7.5E-01
AFLM1	**5.2E-02**	1.5E-01	**2.0E-02**	**9.6E-02**	2.9E-01
AFLM2	**3.6E-07**	5.9E-01	**5.1E-02**	**2.0E-03**	4.4E-01
GPLM2	3.2E-01	5.6E-01	2.3E-01	3.0E-01	**6.5E-02**
CNLM1	2.6E-01	**4.5E-03**	**1.1E-03**	1.4E-01	4.6E-01
CNLM2	8.7E-01	**9.9E-02**	**9.5E-04**	**6.0E-03**	**1.8E-03**
C5LM1	**2.4E-02**	1.3E-01	1.7E-01	7.5E-01	4.2E-01
C5LM2	6.0E-01	1.4E-01	**2.5E-03**	**1.0E-02**	**1.4E-03**
C6LM1	4.1E-01	**1.9E-02**	**9.8E-04**	**4.3E-02**	1.9E-01
C7LM1	1.0E-01	**3.5E-02**	**4.9E-02**	6.1E-01	4.8E-01
DWLM1	**3.2E-02**	2.3E-01	**5.4E-02**	1.7E-01	7.5E-01
PrtostLM1	**8.8E-05**	9.5E-01	8.2E-01	8.1E-01	3.0E-01
PrtostLM2	8.7E-01	2.8E-01	1.9E-01	5.0E-01	9.8E-01
IGUI1	1.7E-01	8.8E-01	3.5E-01	2.7E-01	3.4E-01
IGUI2	3.6E-01	9.3E-01	9.6E-01	9.8E-01	**2.3E-02**
ARUP1	**7.5E-05**	1.1E-01	**4.0E-02**	2.7E-01	5.9E-01
ARUP2	**3.3E-04**	2.5E-01	4.2E-01	9.3E-01	4.4E-01
AMTUP1	**3.8E-03**	7.7E-01	4.2E-01	4.4E-01	4.7E-01
AMTUP2	**1.6E-03**	1.9E-01	4.7E-01	7.1E-01	8.2E-01
3CUM2	1.5E-01	1.0E-01	7.2E-01	**3.3E-02**	**5.8E-02**
ARPrLP1	**4.3E-03**	1.4E-01	3.9E-01	7.2E-01	9.1E-01
ARPrLP2	**3.9E-04**	**7.1E-02**	**6.8E-02**	5.4E-01	2.2E-01

Table 4.11b: Corresponding *P* values correlation between 46 ordinal dental traits and age, sex and continental ancestry.

4.3.2.3 Narrow-sense heritability (h2) of 46 ordinal traits

Based on a kinship matrix derived from the SNP data [60], the narrow-sense heritability for the initial 46 ordinal dental traits using GCTA [61] was estimated. Most of the values are either close to 1 or 0, meaning the algorithm convergence might have had problems given the small sample size, and therefore converged to one boundary or the other of the 0-1 interval, which is the set of legitimate values of heritability. Only a few values are intermediate (around 0.3 or 0.5). But the variance is huge - SD is around 40%. Thus, except those with high heritability values, all others are non-significant, which means that even a heritability of 50% can be null i.e. 0%. (Table 4.12). Heritability was not calculated for the regrouped ASUDAS traits or metric data due to the reasons explained above.

Trait	Heritability	S.E.	*P*-val
SSUI1	1	0.47	4.50E-03
SSUI2	0.92	0.45	8.60E-03
DSUI1	0.6	0.44	4.90E-02
DSUI2	0.65	0.42	3.10E-02
WINUI1	0.39	0.4	1.70E-01
LCUI2	0.12	0.32	3.40E-01
IGUI1	0	0.36	5.00E-01
IGUI2	0	0.4	5.00E-01
TDUI1	0	0.32	5.00E-01
TDUI2	0.82	0.48	3.70E-02
TDUC	0.48	0.39	3.80E-02
MRUC	0.75	0.48	6.90E-02
DARUC	0	0.3	5.00E-01
ARUP1	0.54	0.46	6.50E-02
ARUP2	0	0.29	5.00E-01
AMTUP1	0	0.4	5.00E-01
AMTUP2	0.01	0.24	4.70E-01
MetUM1	0.32	0.36	9.90E-02
MetUM2	0.93	0.45	4.10E-03
3CUM2	0.14	0.31	2.90E-01
HipUM1	0	0.18	5.00E-01
HipUM2	0.74	0.42	8.70E-03
C5UM1	1	0.44	3.90E-06
C5UM2	0.88	0.41	8.20E-04
CarUM1	0	0.36	5.00E-01
SSLI1	1	0.43	1.80E-04
SSLI2	1	0.41	5.70E-04
DSLI1	0.47	0.43	1.20E-01
DSLI2	0.27	0.4	2.50E-01
DARLC	0	0.41	5.00E-01
LCVLP1	0.32	0.41	1.80E-01
LCVLP2	0.29	0.41	2.10E-01

Table 4.12. Heritability estimates for the 46 ordinal dental traits examined using ASUDAS.

4.3.2.4 Genome Wide Association Analyses

I performed genome-wide association tests using multivariate linear regression, as implemented in PLINK [251] on not regrouped and regrouped (all criteria) ASUDAS traits, using an additive genetic model adjusting for: age, sex and the first five principal components (Figure 4.31) computed from the SNP data. The resulting statistics showed no evidence of residual population stratification for any of the traits (Figure 4.32). After quality control assessment, 412 samples and 640,094 SNPs were retained for the not imputed GWAS and 411 samples and 9,459,622 variants were retained for the imputed GWAS.

4.3.2.4.1 Genomic regions associated with ordinal dental traits

Eleven of the ordinal dental traits examined showed genome-wide significant association and Figure 4.34(P values < 5×10^{-8}) with SNPs in at least one genomic region (Table 4.13).

Chromosomal Region	Index SNP	Alleles[2]	Effect size	P value	Associated Trait	Candidate Gene[1]
1p13.1	rs10494193	T>C	1.54E+00	4.6E-08	Protostylid LM1	*CD101*
1q43	rs3851907	C>T	8.85E-01	2.3E-09	Cusp 7 LM1 *	*ACTN2*
2p12	rs9808165	G>T	-1.28E+00	4.0E-09	Hypocone UM1	*LOC101927907*
2q12.3	rs3827760	A>G	4.05E-01	2.2E-09	Shovel Shape LI1	*EDAR*
2q12.3	rs3827760	A>G	4.05E-01	2.2E-09	Shovel Shape LI1 *	*EDAR*
2q12.3	rs3827760	A>G	9.52E-01	2.4E-11	Shovel Shape UI1	*EDAR*
2q12.3	rs3827760	A>G	4.22E-01	4.3E-10	Shovel Shape UI1 *	*EDAR*
2q12.3	rs3827760	A>G	9.29E-01	5.4E-11	Shovel Shape UI2	*EDAR*
2q12.3	rs3827760	A>G	4.26E-01	8.4E-11	Shovel Shape UI2 *	*EDAR*
2q37.1	rs17862881	A>G	9.11E-01	4.4E-08	Double Shoveling LI1	***TRPM8***
2q37.1	rs17862881	A>G	1.04E+00	3.8E-11	Double Shoveling LI2	***TRPM8***
10q26.13	rs12570261	C>T	8.11E-01	4.6E-08	Protostylid LM1	***FGFR2***
10q26.2	rs1037804	T>A	1.02E+00	2.0E-13	Double Shoveling LI1	*ADAM12*
10q26.2	rs1037804	T>A	8.99E-01	1.6E-11	Double Shoveling LI2	*ADAM12*
12p11.21	rs16919218	C<T	1.35E+00	4.6E-08	Deflecting Wrinkle LM1	***KIAA1551/BICD1***
12q24.33	rs12299956	C>T	1.27E+00	1.3E-08	Protostylid LM2	*GPR133*
15q23	rs630603	C>T	1.13E+00	4.8E-13	Protostylid LM2	*TLE3*
17q11.2	rs9899063	G>A	9.62E-01	4.3E-08	Protostylid LM2	*NF1*
18q22.1	rs17072636	G>A	-4.49E-01	3.2E-08	Metacone UM2 *	*LOC284294*
22q12.1	rs16984020	G>A	8.64E-01	4.0E-08	Double Shoveling LI1	***LOC101929***

Table 4.13. Properties of index SNPs in chromosomal regions associated to the ASUDAS traits examined. [1] For intragenic SNPs, gene names are shown in bold.
[2] Derived alleles are shown after ancestral alleles.
* Regrouped traits (Section 4.2.2.1.3, Table 4.6).
MAF CAN: Minor Allele frequency in CANDELA Cohort
UI1: Upper central incisor
UI2: Upper lateral incisor
UM1: Upper molar 1
UM2: Upper molar 2
LI1: Lower central incisor
LI2: Lower lateral incisor
LM1: Lower molar 1
LM2: Lower molar 2

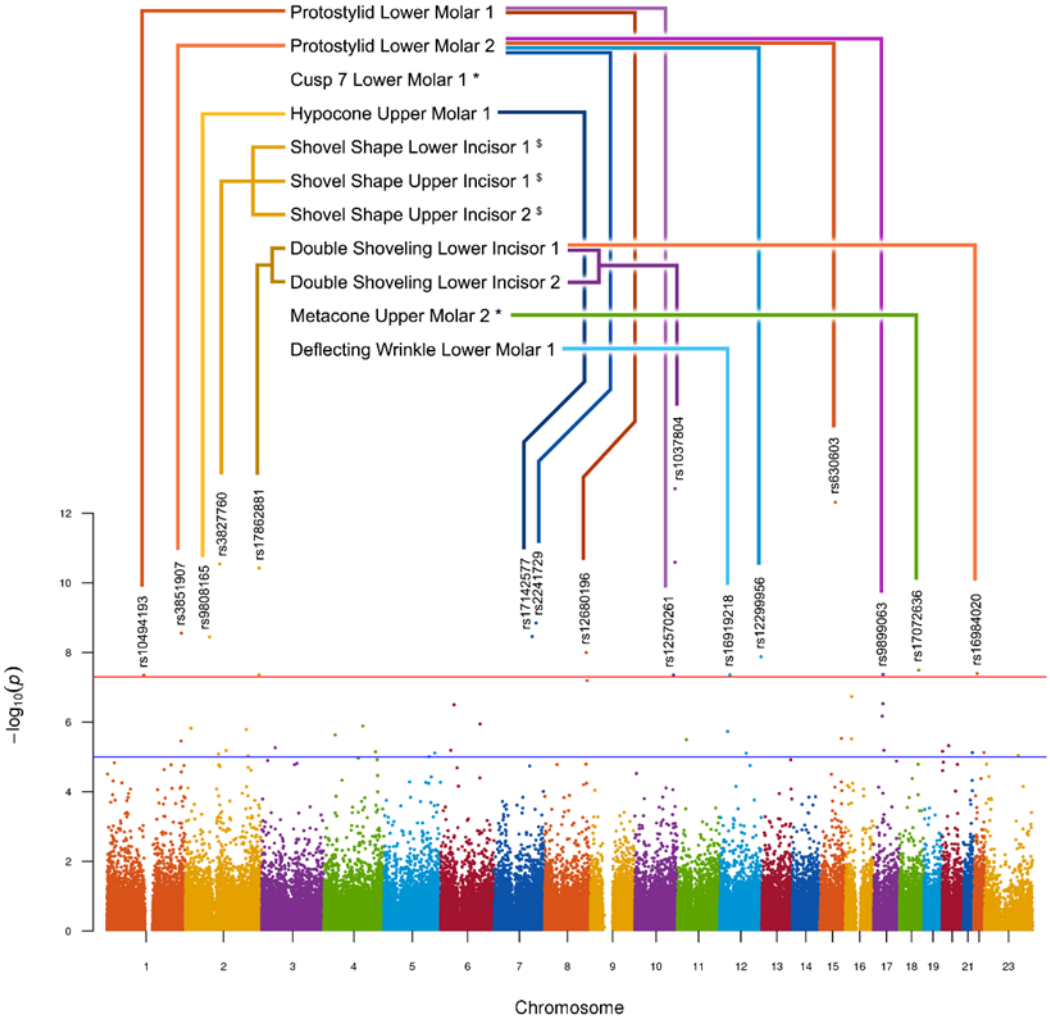


Figure 4.34. Manhattan plot summarizing the GWAS results for the ASUDAS dental traits examined in the Colombian CANDELA samples. These GWAS was performed with 46 traits phenotyped using the ASUDAS scoring method from 501 volunteers. A 'composite' of Manhattan plot shows the results across traits. All the SNPs with P values exceeding thresholds genome-wide suggestive (10^{-5}) are over the blue line and P values reaching the threshold genome-wide significant ($5x10^{-8}$) are above the red line.

The accessory cusp 7 of the first lower molar (C7LM1) regrouped (from 6 to 3 categories) showed association with the genomic variant, rs3851907, located in 1q43 chromosome region. It falls within the gene *ACTN2* genomic coordinates (Actinin, alpha 2) (Table 4.13, Figure 4.34 and Figure 4.35).

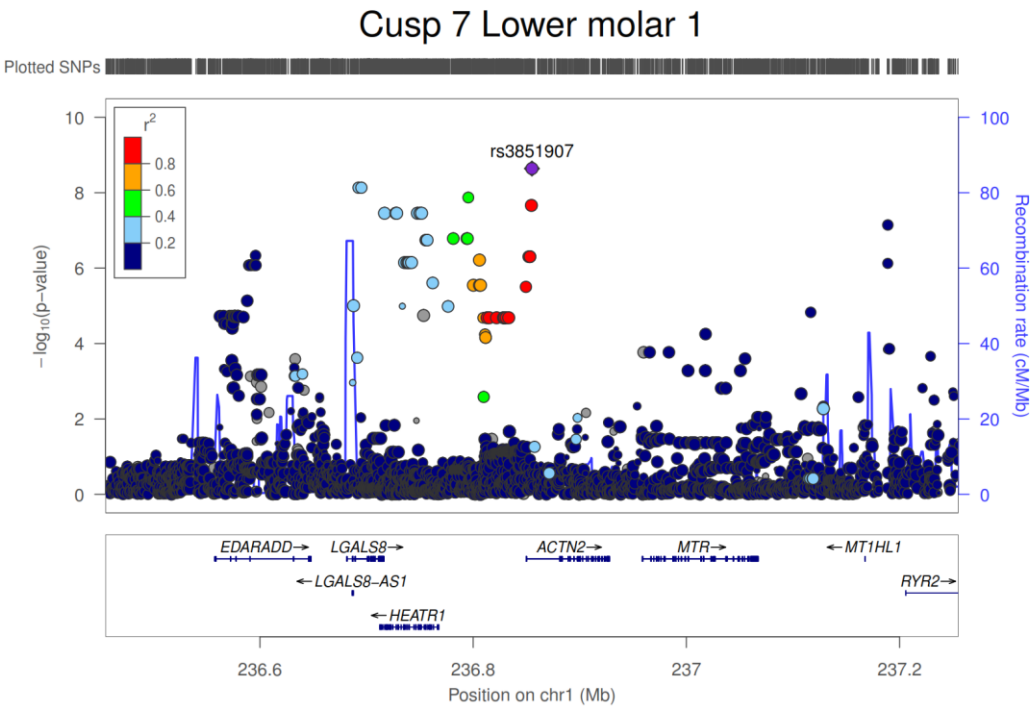


Figure 4.35. Regional association plot for SNPs in 1q43 and Cusp 7 first lower molar. This plot shows an area of 500 Mb around the index SNP (rs3851907). This plot was produced with Locus Zoom [310].

Double shoveling in the central lower incisor (DSLI1) showed association with 3 genetic markers, rs1037804, rs16984020 and rs17862881. The first one is in the chromosomic region 10q26.2 (Table 4.13, Figure 4.34 and Figure 4.36), corresponding to an intron on the gene *ADAM12*. The second SNP (rs16984020) associated with DSLI1 is in chromosome 22 (Table 4.13, Figure 4.34 and Figure 4.37) and it does not fall in a specific gene. It is around 200 Mb downstream from *LOC1929539*. Finally, the third marker that presented significant association with DSLI1 is rs17862881. This SNP is also located in an intergenic region (2q37.1) (Table 4.13, Figure 4.34 and Figure 4.38), but it is near several genes.

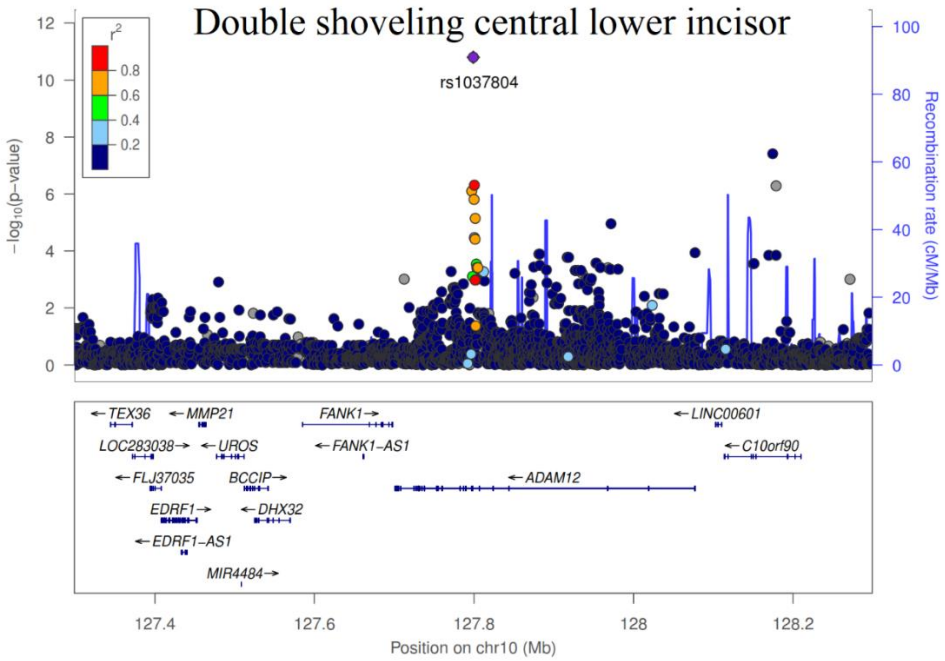


Figure 4.36. Regional association plot for SNPs in 10q26.2 and DSLI1. This plot shows an area of 500 Mb around the index SNP (rs1037804). This plot was produced with Locus Zoom [310].

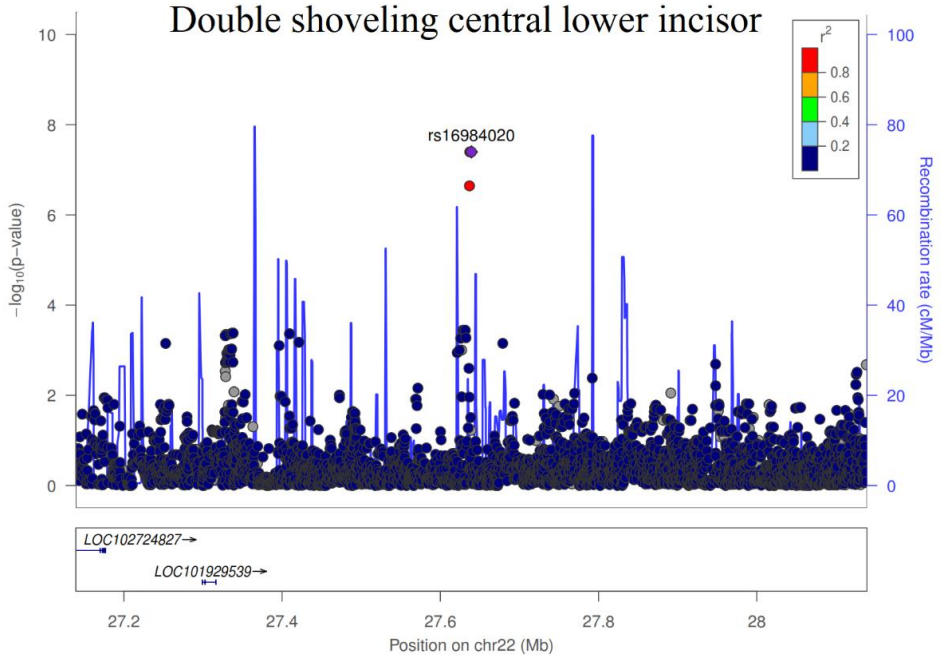


Figure 4.37. Regional association plot for SNPs in 22q12.1 and DSLI1. This plot shows an area of 500 Mb around the index SNP (rs16984020). This plot was produced with Locus Zoom [310].

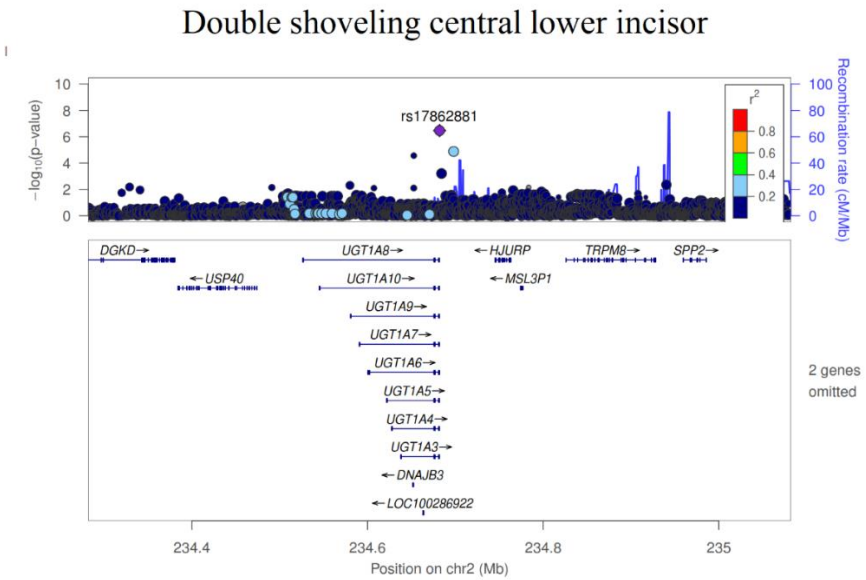


Figure 4.38. Regional association plot for SNPs in 2q37.1 and DSLI1. This plot shows an area of 500 Mb around the index SNP (rs17862881). This plot was produced with Locus Zoom [310].

Double shoveling on the lateral lower incisor (DSLI2) showed the same associations with SNPs rs17862881 and rs1037804 (Table 4.13, Figure 4.34, Figure 4.39 and Figure 4.40). This was expected since this is the same trait assessed in another tooth.

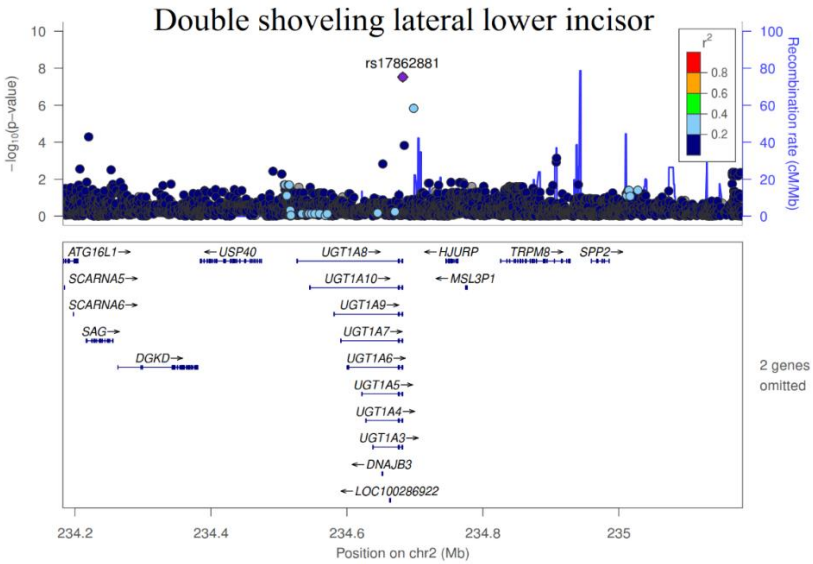


Figure 4.39. Regional association plot for SNPs in 2q37.1 and DSLI2. This plot shows an area of 500 Mb around the index SNP (rs17862881). This plot was produced with Locus Zoom [310].

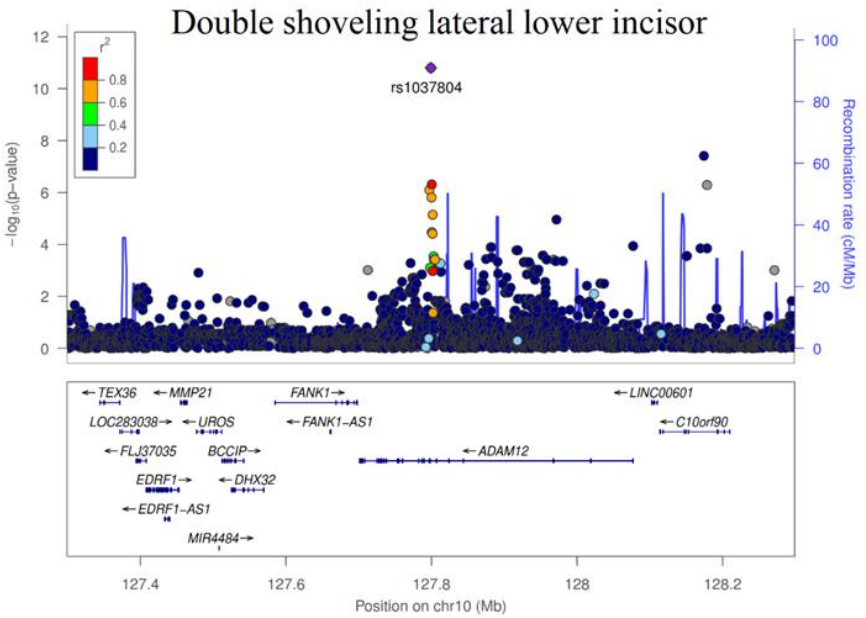


Figure 4.40. Regional association plot for SNPs in 10q26.2 and DSLI2. This plot shows an area of 500 Mb around the index SNP (rs1037804). This plot was produced with Locus Zoom [310].

Deflecting wrinkle of the first lower molar (DWLM1) showed association with SNPs in chromosome region 12p11.21 (Table 4.13, Figure 4.34 and Figure 4.41). The index marker was rs16919218. This genetic variant falls in an intergenic region near to *KIAA1551* and *BICD1* genes.

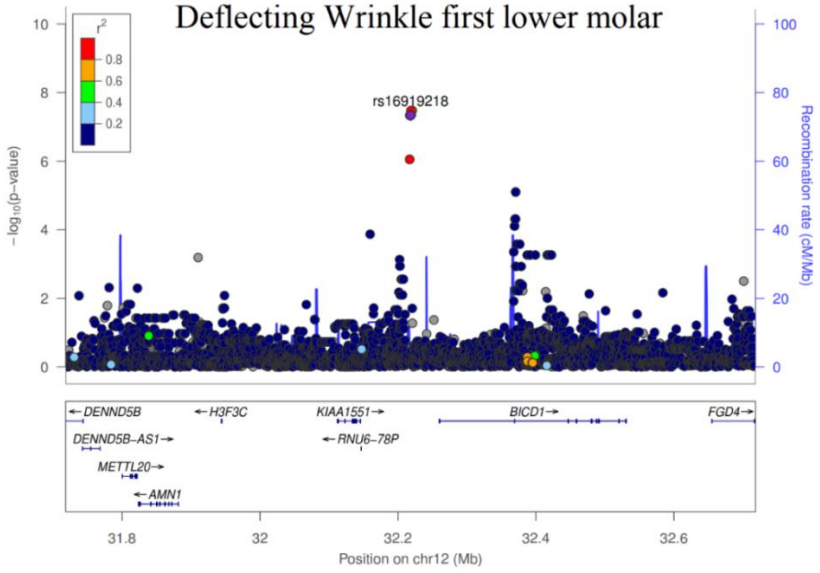


Figure 4.41. Regional association plot for SNPs in 12p11.21 and DWLM1. This plot shows an area of 500 Mb around the index SNP (rs16919218). This plot was produced with Locus Zoom [310].

Hypocone in the first upper molar (HypLM1) corresponds to the fourth cusp situated in the distolingual part of the molar. This feature showed a significant association with two genetic markers, rs9808165 and rs17142577 (Table 4.13 and Figure 4.34). The first SNP is in chromosome region 2p12 (Figure 4.42). The genomic variant rs17142577 is the second SNP associated with HypLM1, and it is located in the 7q31.31 region (Figure 4.43). It does not fall in a gene but is close to *KCND2*.

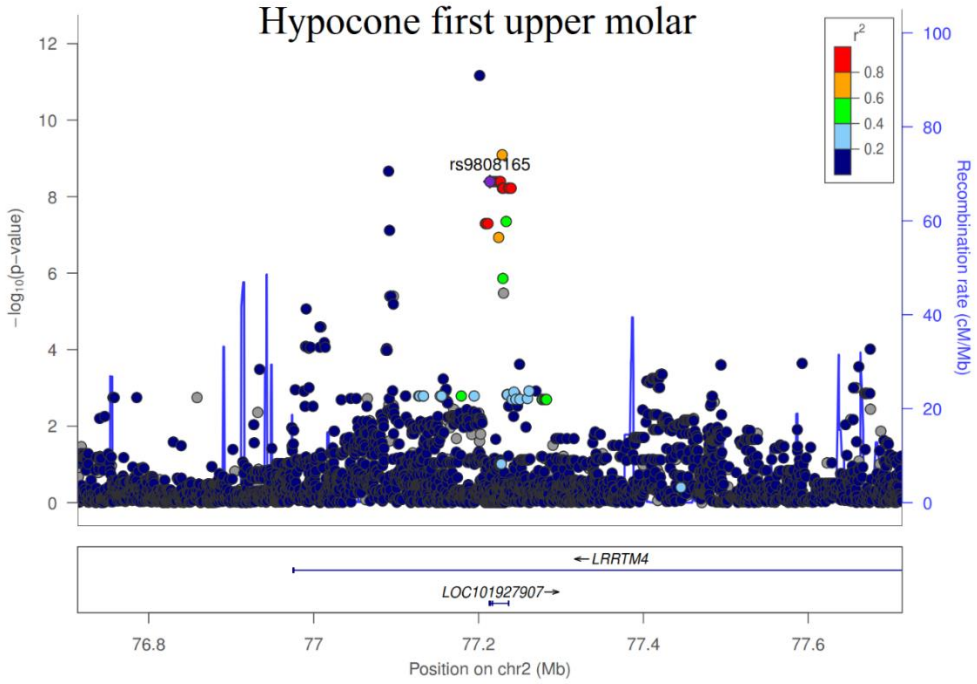


Figure 4.42. Regional association plot for SNPs in 2p12 and HypUM1. This plot shows an area of 500 Mb around the index SNP (rs9808165). This plot was produced with Locus Zoom [310].

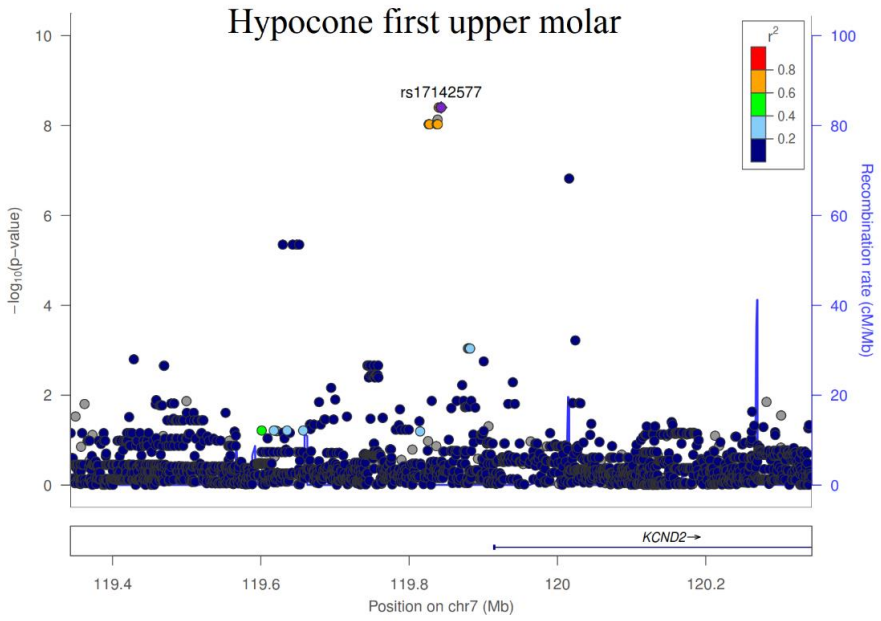


Figure 4.43. Regional association plot for SNPs in 7q31.31 and HypUM1. This plot shows an area of 500 Mb around the index SNP (rs17142577). This plot was produced with Locus Zoom [310].

The trait Metacone in the second upper molar (MetUM2) regrouped (from 7 to 4 categories) is the presence of cusp number 3 in the upper molars. It showed association with a SNP (rs17072636) situated in the 18q22.1 region and it falls in an intron in the uncharacterized gene *LOC284294* (Table 4.13, Figure 4.34 and Figure 4.44).

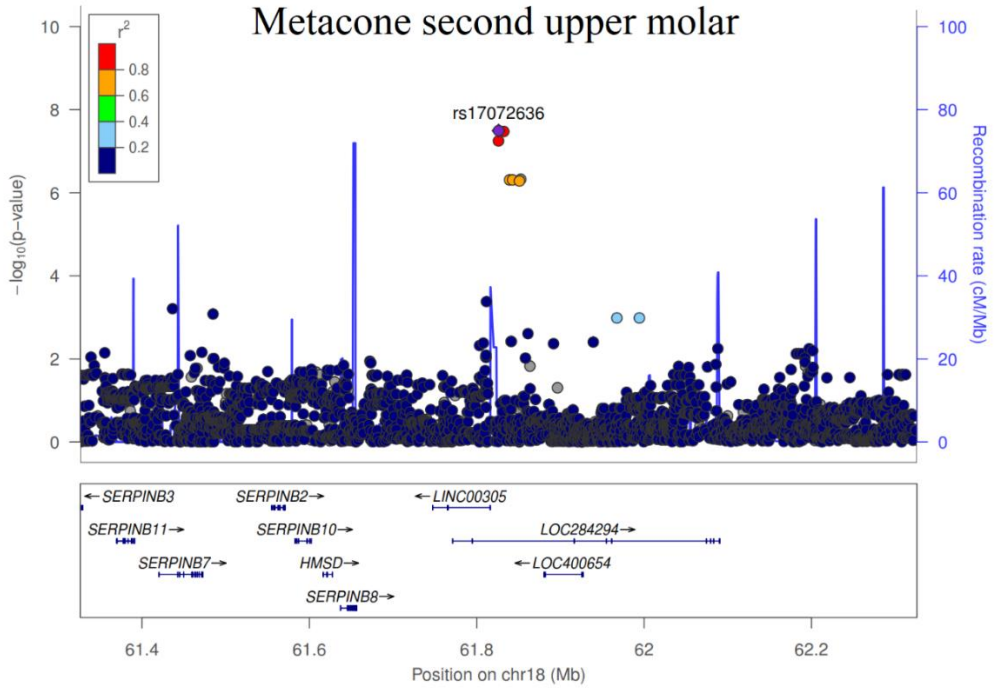


Figure 4.44. Regional association plot for SNPs in 18q22.1 and MetUM2. This plot shows an area of 500 Mb around the index SNP (rs17072636). This plot was produced with Locus Zoom [310].

Protostylid in the first lower molar (PrtstLM1), correspond to the appearance of a secondary groove associated to the buccal groove separating cusps 1 and 3 on the buccal surface of the lower molars. This feature has shown association with SNPs in three genomic regions (1p13.1, 8q24.22 and 10q26.13) (Table 4.13 and Figure 4.34). The first genomic variant (rs10494193) is located in *CD101* gene (Figure 4.45).

The second marker that showed association with this trait was rs12680196, it falls in an intergenic region (8q24.22) and it is 400Mb downstream from the gene *EFR3A* (Figure 4.46).

Finally, PrtostLM1 presented an association with a SNP (rs12570261), also located in an intergenic region in chromosome 10, band 26.13 (Figure 4.47). It is close to the gene *FGFR2* (~50Mb downstream).

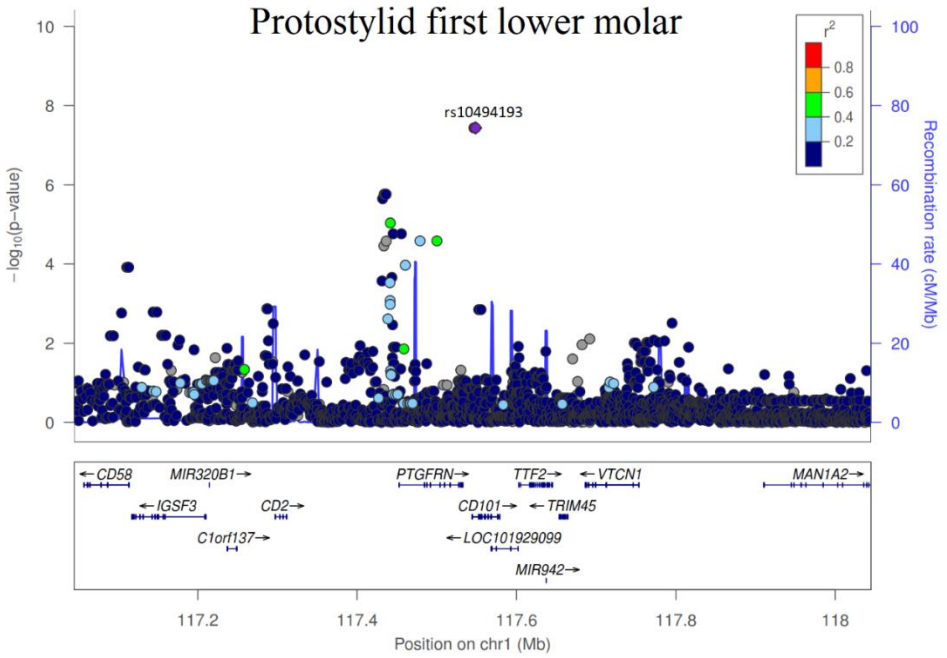


Figure 4.45. Regional association plot for SNPs in 1p13.1 and PrtstLM1. This plot shows an area of 500 Mb around the index SNP (rs10494193). This plot was produced with Locus Zoom [310].

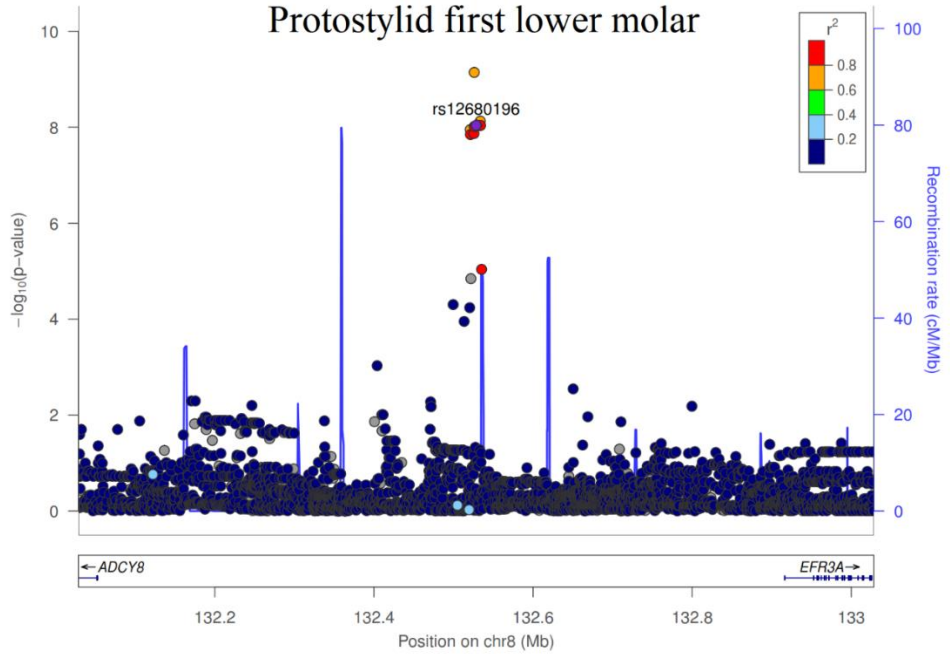


Figure 4.46. Regional association plot for SNPs in 8q21.22 and PrtstLM1. This plot shows an area of 500 Mb around the index SNP (rs12680196). This plot was produced with Locus Zoom [310].

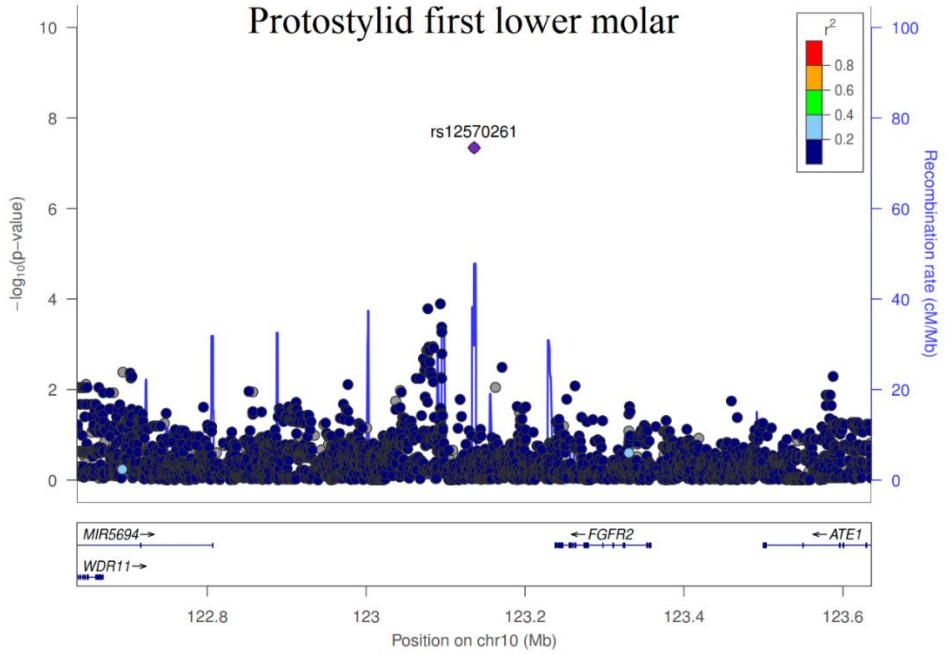


Figure 4.47. Regional association plot for SNPs in 10q26.13 and PrtstLM1. This plot shows an area of 500 Mb around the index SNP (rs12570261). This plot was produced with Locus Zoom [310].

Four genomic variants (rs2241729, rs12299956, rs630603 and rs9899063) located in four different chromosomes were found associated with Protostylid in the second lower molar (PrtstLM2) (Table 4.13 and Figure 4.34)

The first marker rs2241729 (7q32.3) (Figure 4.48) is situated in an intron in the gene *PLXNA4.*

The second genetic variant rs12299956, falls within the gene *GPR133* in 12q24.33 chromosome region (Figure 4.49).

Marker (rs630603) situated in the 15q23 chromosome region (Figure 4.50) is the third SNP associated with Protostylid in the second lower molar. This genomic variant falls in the *TLE3* gene.

Finally, the SNP rs9899063 which falls in the *NF1* gene showed an association with PrtstLM2. The *NF1* gene is in chromosome region 17q11.2 (Figure 4.51).

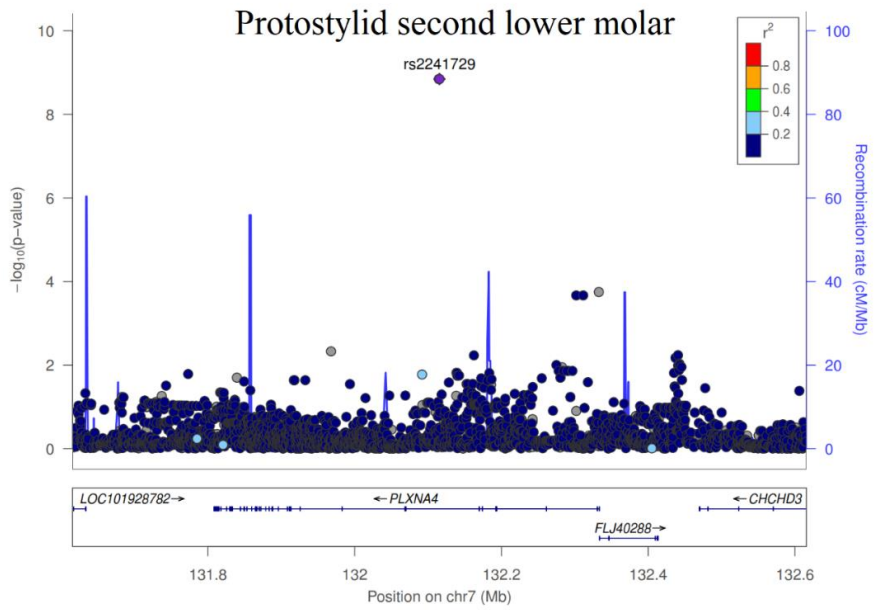


Figure 4.48. Regional association plot for SNPs in 7q32.3 and PrtstLM2. This plot shows an area of 500 Mb around the index SNP (rs2241729). This plot was produced with Locus Zoom [310].

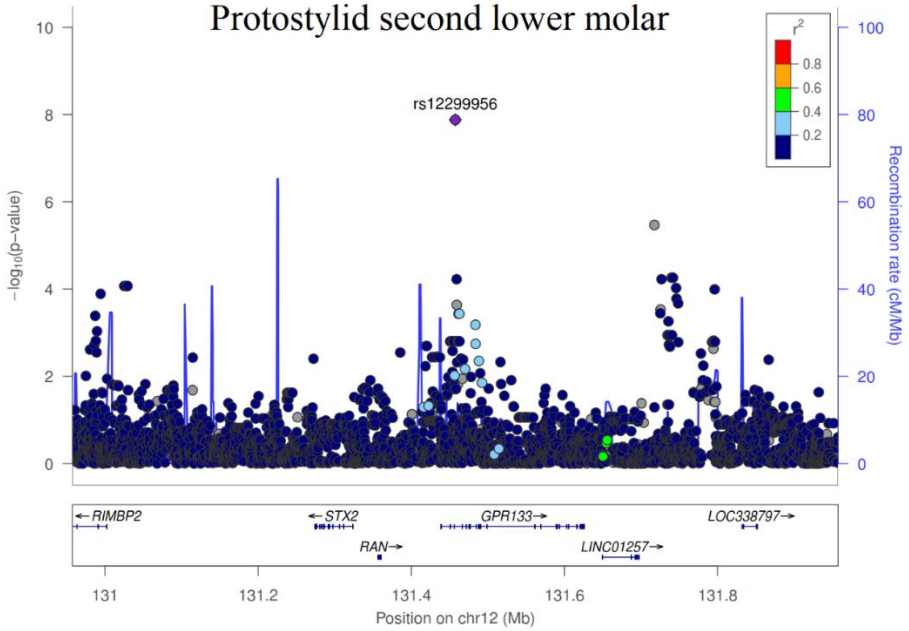


Figure 4.49. Regional association plot for SNPs in 12q24.33 and PrtstLM2. This plot shows an area of 500 Mb around the index SNP (rs12299956). This plot was produced with Locus Zoom [310].

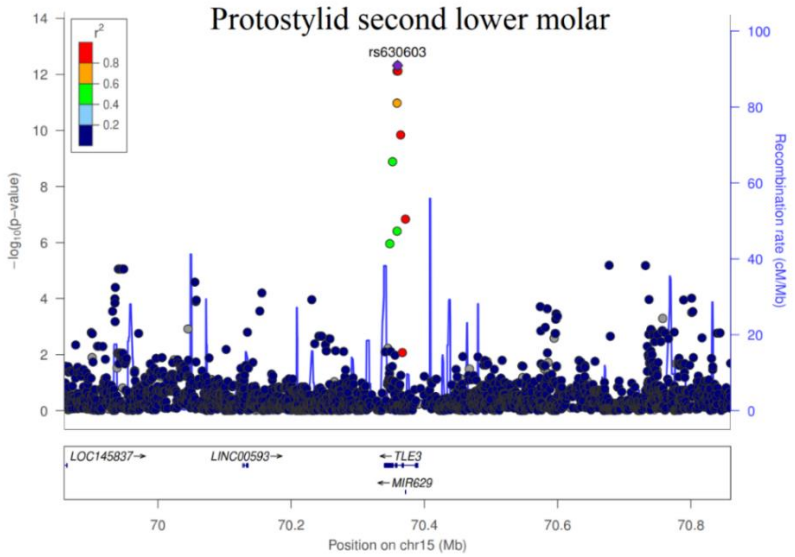


Figure 4.50. Regional association plot for SNPs in 15q23 and PrtstLM2. This plot shows an area of 500 Mb around the index SNP (rs630603). This plot was produced with Locus Zoom [310].

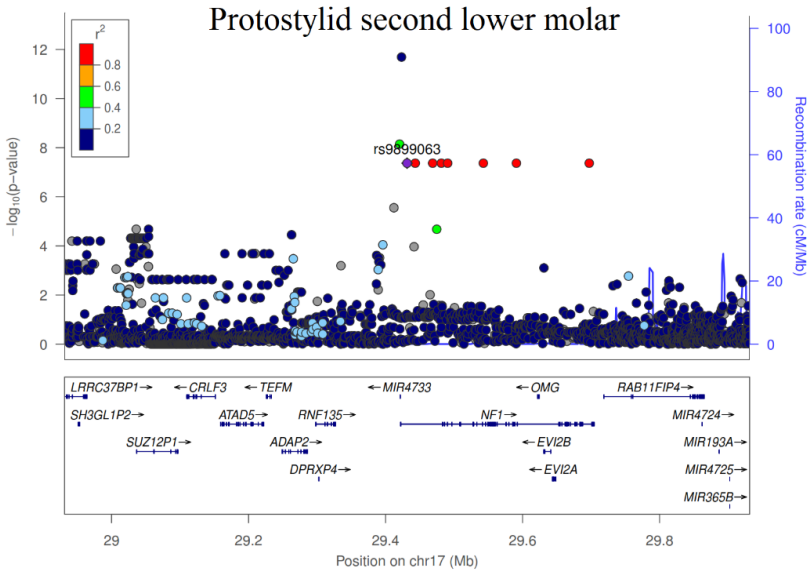


Figure 4.51. Regional association plot for SNPs in 17q11.2 and PrtstLM2. This plot shows an area of 500 Mb around the index SNP (rs9899063). This plot was produced with Locus Zoom [310].

The genomic variant rs3827760 falls within the *EDAR* gene (2q12.3) (Table 4.13 and Figure 4.34, Figure 4.52 - Figure 4.57), showed association with shovel shape in central and lateral upper incisors and in central lower incisors (SSUI1,SSUI1 regrouped, SSUI2, SSUI2 regrouped, SLI1 and SLI1 regrouped), both the normally categorized version (0-6) and regrouped version (0-3) of the phenotypes.

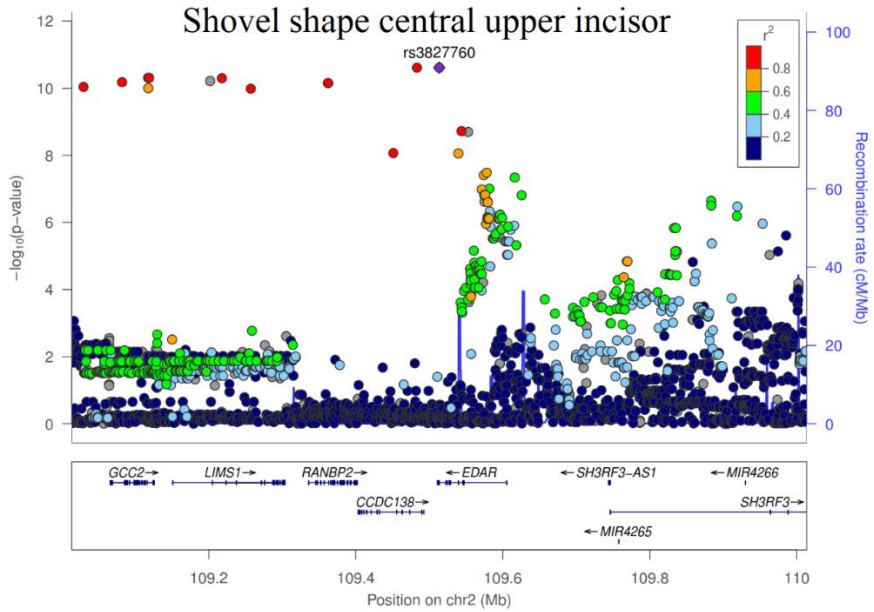


Figure 4.52. Regional association plot for SNPs in 2q12.3 and SSUI1. This plot shows an area of 500 Mb around the index SNP (rs3827760). This plot was produced with Locus Zoom [310].

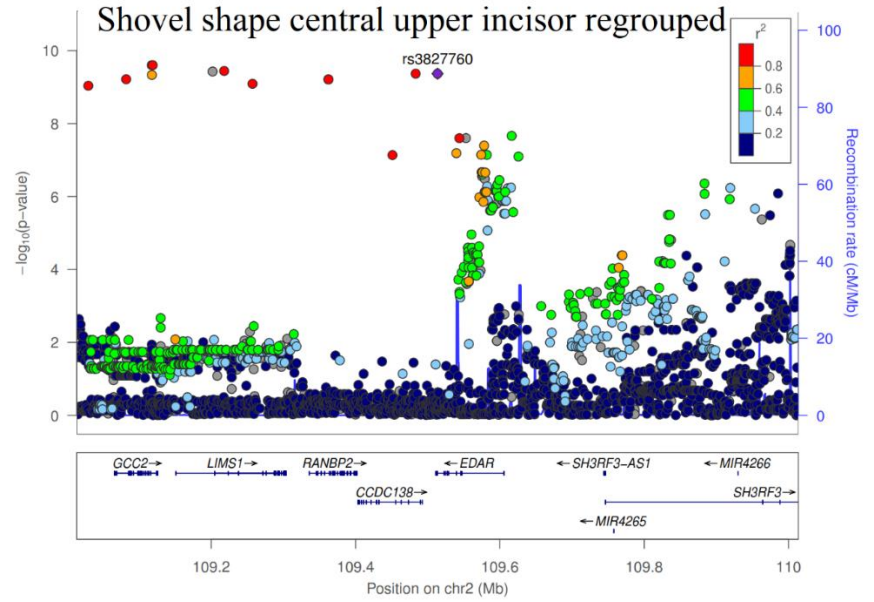


Figure 4.53. Regional association plot for SNPs in 2q12.3 and SSUI1 regrouped. This plot shows an area of 500 Mb around the index SNP (rs3827760). This plot was produced with Locus Zoom [310].

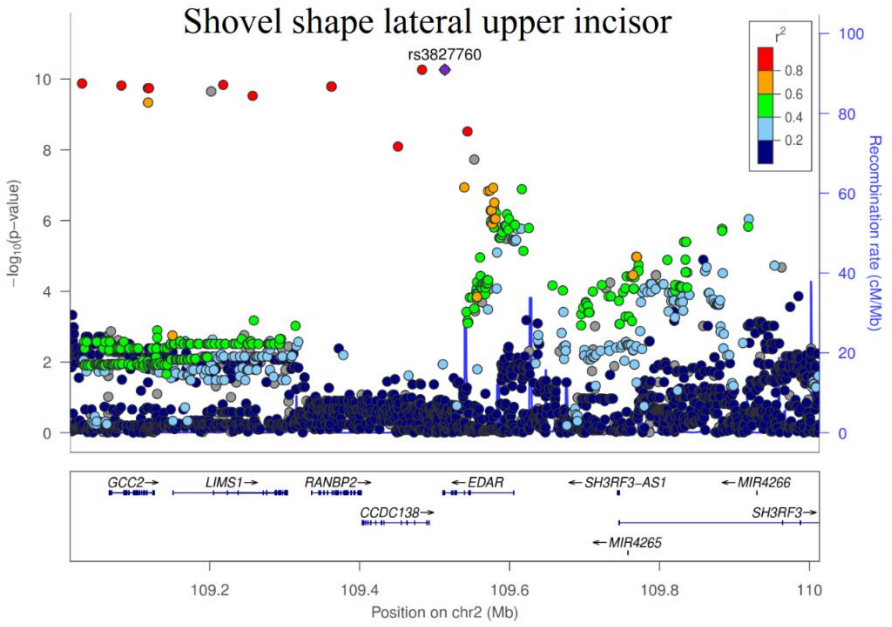


Figure 4.54. Regional association plot for SNPs in 2q12.3 and SSUI2. This plot shows an area of 500 Mb around the index SNP (rs3827760). This plot was produced with Locus Zoom [310].

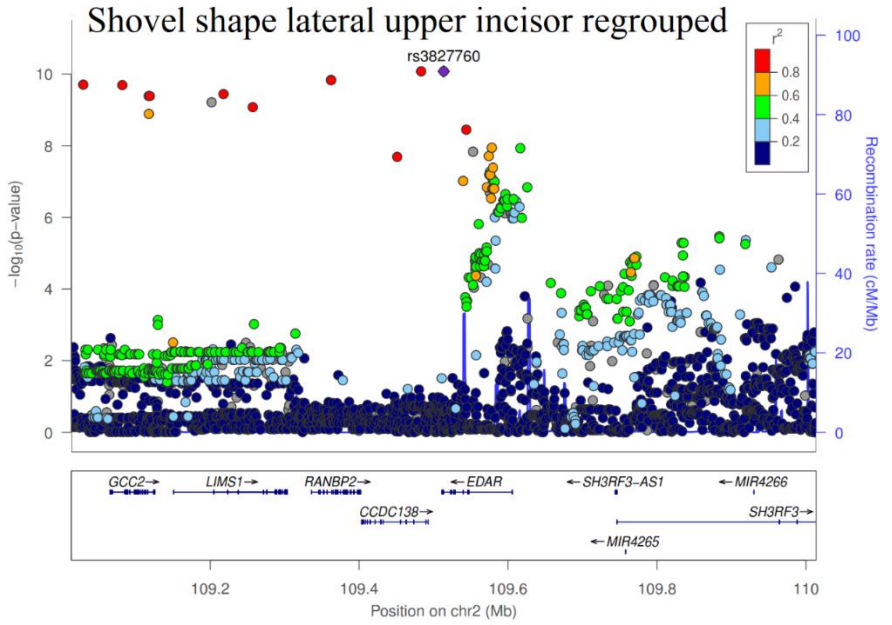


Figure 4.55. Regional association plot for SNPs in 2q12.3 and SSUI2 regrouped. This plot shows an area of 500 Mb around the index SNP (rs3827760). This plot was produced with Locus Zoom [310].

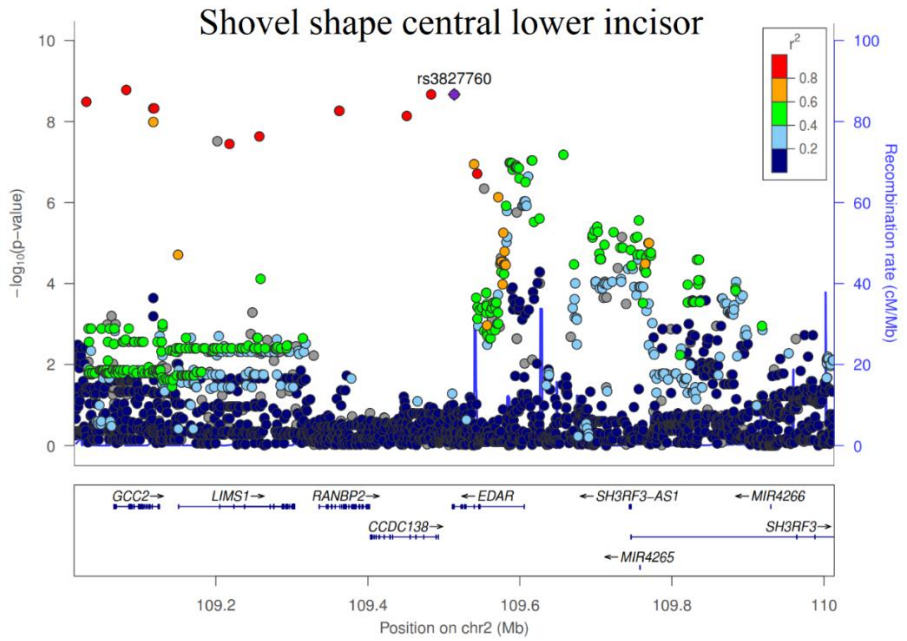


Figure 4.56. Regional association plot for SNPs in 2q12.3 and SSLI1. This plot shows an area of 500 Mb around the index SNP (rs3827760). This plot was produced with Locus Zoom [310].

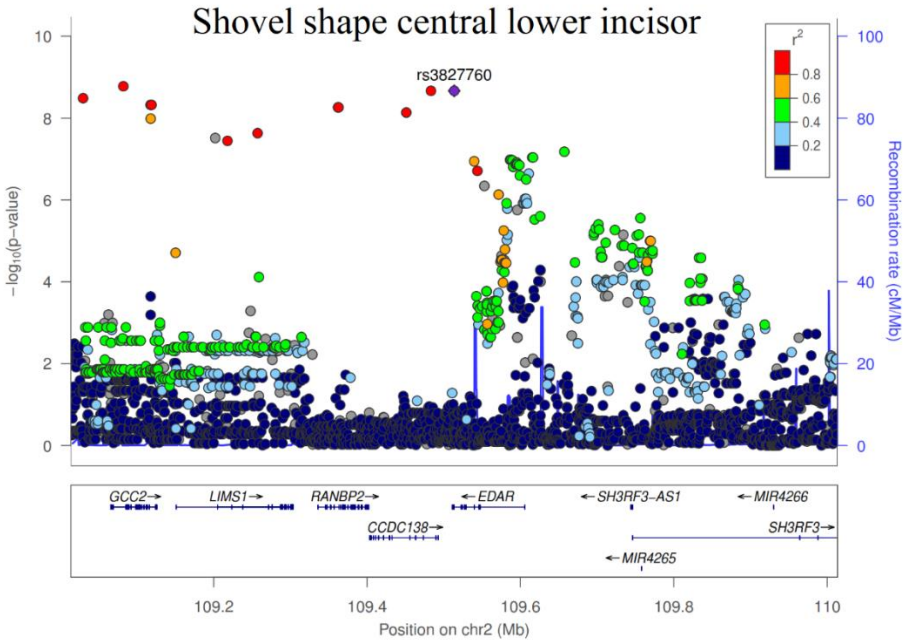


Figure 4.57. Regional association plot for SNPs in 2q12.3 and SSLI1 regrouped. This plot shows an area of 500 Mb around the index SNP (rs3827760). This plot was produced with Locus Zoom [310].

4.3.3 Quantitative dental features

Sixty distances were measured (Table 4.9 and Figure 4.30) using dental photographs (Figure 4.27) of 564 individuals.

4.3.3.1 Rater reliability

The rater reliability for metric dental traits scores were assessed by calculating intra-class correlation coefficients (ICC) [291]. They indicate high intra-rater reliability of the traits measured (Table 4.14). There are two traits with lower ICC (MD_Up_Rt_C and BL_Up_Lf_I2).

Jaw	Measurement	Side	Tooth	Trait code	ICC (L.M.R)
Upper	Incisal-Cervical Distance	Right	Central Incisor	IC_Up_Rt_I1	0.99
			Lateral Incisor	IC_Up_Rt_I2	0.96
			Canine	IC_Up_Rt_C	1.00
			First Premolar	IC_Up_Rt_P1	0.97
			Second Premolar	IC_Up_Rt_P2	0.98
		Left	Central Incisor	IC_Up_Lf_I1	0.99
			Lateral Incisor	IC_Up_Lf_I2	0.98
			Canine	IC_Up_Lf_C	0.98
			First Premolar	IC_Up_Lf_P1	0.99
			Second Premolar	IC_Up_Lf_P2	0.96
	Meso-Distal Distance	Right	Central Incisor	MD_Up_Rt_I1	0.96
			Lateral Incisor	MD_Up_Rt_I2	0.97
			Canine	MD_Up_Rt_C	0.76
			First Premolar	MD_Up_Rt_P1	0.99
			Second Premolar	MD_Up_Rt_P2	0.99
		Left	Central Incisor	MD_Up_Lf_I1	0.99
			Lateral Incisor	MD_Up_Lf_I2	0.94
			Canine	MD_Up_Lf_C	0.97
			First Premolar	MD_Up_Lf_P1	1.00
			Second Premolar	MD_Up_Lf_P2	0.99
	Bucco-Lingual Distance	Right	Central Incisor	BL_Up_Rt_I1	0.91
			Lateral Incisor	BL_Up_Rt_I2	0.91
			Canine	BL_Up_Rt_C	0.98
			First Premolar	BL_Up_Rt_P1	0.98
			Second Premolar	BL_Up_Rt_P2	0.98
		Left	Central Incisor	BL_Up_Lf_I1	0.88
			Lateral Incisor	BL_Up_Lf_I2	0.94
			Canine	BL_Up_Lf_C	0.98
			First Premolar	BL_Up_Lf_P1	0.99
			Second Premolar	BL_Up_Lf_P2	0.99

Jaw	Measurement	Side	Tooth	Trait code	ICC (L.M.R)
Lower	Incisal-Cervical Distance	Right	Central Incisor	IC_Lw_Rt_I1	0.95
			Lateral Incisor	IC_Lw_Rt_I2	0.99
			Canine	IC_Lw_Rt_C	1.00
			First Premolar	IC_Lw_Rt_P1	0.98
			Second Premolar	IC_Lw_Rt_P2	0.98
		Left	Central Incisor	IC_Lw_Lf_I1	0.96
			Lateral Incisor	IC_Lw_Lf_I2	1.00
			Canine	IC_Lw_Lf_C	1.00
			First Premolar	IC_Lw_Lf_P1	0.98
			Second Premolar	IC_Lw_Lf_P2	0.94
	Meso-Distal Distance	Right	Central Incisor	MD_Lw_Rt_I1	0.98
			Lateral Incisor	MD_Lw_Rt_I2	0.97
			Canine	MD_Lw_Rt_C	0.96
			First Premolar	MD_Lw_Rt_P1	0.98
			Second Premolar	MD_Lw_Rt_P2	0.98
		Left	Central Incisor	MD_Lw_Lf_I1	0.97
			Lateral Incisor	MD_Lw_Lf_I2	0.96
			Canine	MD_Lw_Lf_C	0.90
			First Premolar	MD_Lw_Lf_P1	0.98
			Second Premolar	MD_Lw_Lf_P2	0.98
	Bucco-Lingual Distance	Right	Central Incisor	BL_Lw_Rt_I1	0.85
			Lateral Incisor	BL_Lw_Rt_I2	0.92
			Canine	BL_Lw_Rt_C	0.99
			First Premolar	BL_Lw_Rt_P1	0.99
			Second Premolar	BL_Lw_Rt_P2	0.99
		Left	Central Incisor	BL_Lw_Lf_I1	0.97
			Lateral Incisor	BL_Lw_Lf_I2	0.80
			Canine	BL_Lw_Lf_C	0.99
			First Premolar	BL_Lw_Lf_P1	0.99
			Second Premolar	BL_Lw_Lf_P2	0.99

Table 4.14. Rater reliability for quantitative dental traits. There was 1 rater (L.M.R) that scored the same 28 people twice, 2 weeks apart.

4.3.3.2 Correlations

Overall, most of the traits showed from weak to very strong positive, and significant, correlations amongst each other (using a Bonferroni-adjusted permutation *P*-value threshold for significance of 1×10^{-4}, Table 4.15) (r values 0.10 to 0.93). The strongest

correlations were observed between inciso-cervical distances (length of the tooth) in the upper and lower jaw.

The highest correlation was seen between the inciso-cervical distances of the right and left central incisors in the upper jaw (r = 0.93) (IC_Up_Lf_I1 and IC_Up_Rt_I1). The lowest correlation was between the inciso-cervical distance of the upper left central incisor and the second upper premolar. Something similar was observed with the bucco-lingual and mesio-distal measurements. This is reflecting the difference in size of the different teeth as they move away from the midline line of the teeth.

Most of the correlations between bucco-lingual and mesio-distal distances were from moderate to strong positive correlations (r values from 0.30 to 0.64). Nevertheless, most of the correlations between these two measurements, bucco-lingual and meso-distal distances, with the length of the teeth (inciso-cervical measurements) were negligible or weak positive correlations (r values 0.0 to 0.16).

Trait	IC_Up_Rt_P2	IC_Up_Rt_P1	IC_Up_Rt_C	IC_Up_Rt_I2	IC_Up_Rt_I1	IC_Up_Lf_I1	IC_Up_Lf_I2	IC_Up_Lf_C	IC_Up_Lf_P1	IC_Up_Lf_P2
IC_Up_Rt_P2		**1.7E-110**	**1.5E-57**	**6.5E-50**	**5.7E-37**	**2.4E-32**	**6.5E-49**	**9.1E-52**	**2.6E-73**	**2.4E-126**
IC_Up_Rt_P1	0.81		**2.6E-56**	**1.1E-42**	**2.5E-32**	**8.8E-30**	**1.0E-46**	**1.9E-58**	**1.0E-121**	**8.5E-84**
IC_Up_Rt_C	0.61	0.64		**1.2E-98**	**3.8E-76**	**6.5E-70**	**2.1E-77**	**5.4E-167**	**4.5E-51**	**2.8E-50**
IC_Up_Rt_I2	0.58	0.58	0.74		**1.1E-97**	**1.4E-84**	**1.2E-140**	**5.7E-79**	**1.2E-42**	**1.8E-46**
IC_Up_Rt_I1	0.51	0.51	0.68	0.74		**1.2E-242**	**5.0E-84**	**9.8E-69**	**4.0E-32**	**6.2E-38**
IC_Up_Lf_I1	0.48	0.49	0.66	0.71	0.93		**7.4E-85**	**2.4E-69**	**8.5E-33**	**5.7E-34**
IC_Up_Lf_I2	0.57	0.60	0.68	0.83	0.70	0.71		**2.0E-87**	**1.1E-51**	**1.1E-46**
IC_Up_Lf_C	0.58	0.65	0.86	0.69	0.65	0.65	0.71		**9.4E-71**	**2.0E-55**
IC_Up_Lf_P1	0.71	0.83	0.62	0.58	0.51	0.51	0.62	0.70		**3.3E-80**
IC_Up_Lf_P2	0.81	0.75	0.58	0.56	0.51	0.49	0.56	0.60	0.74	
MD_Up_Lf_P2	0.30	0.27	0.27	0.26	0.30	0.28	0.24	0.25	0.25	0.34
MD_Up_Lf_P1	0.34	0.36	0.33	0.28	0.32	0.32	0.27	0.31	0.36	0.34
MD_Up_Lf_C	0.27	0.26	0.40	0.33	0.30	0.31	0.28	0.40	0.24	0.28
MD_Up_Lf_I2	0.20	0.19	0.29	0.34	0.26	0.27	0.35	0.30	0.17	0.23
MD_Up_Lf_I1	0.23	0.27	0.36	0.32	0.36	0.39	0.29	0.38	0.26	0.25
MD_Up_Rt_I1	0.24	0.29	0.38	0.34	0.38	0.40	0.30	0.37	0.27	0.26
MD_Up_Rt_I2	0.22	0.22	0.31	0.37	0.28	0.28	0.37	0.32	0.21	0.26
MD_Up_Rt_C	0.24	0.27	0.37	0.31	0.32	0.33	0.28	0.39	0.28	0.27
MD_Up_Rt_P1	0.35	0.39	0.35	0.31	0.33	0.33	0.29	0.31	0.36	0.32
MD_Up_Rt_P2	0.30	0.26	0.28	0.25	0.29	0.27	0.20	0.23	0.26	0.32
BL_Up_Lf_P2	0.30	0.34	0.35	0.31	0.35	0.36	0.29	0.35	0.33	0.38
BL_Up_Lf_P1	0.27	0.39	0.34	0.29	0.33	0.36	0.28	0.37	0.39	0.33
BL_Up_Lf_C	0.35	0.40	0.51	0.41	0.43	0.44	0.38	0.55	0.42	0.36
BL_Up_Lf_I2	0.30	0.34	0.45	0.44	0.42	0.40	0.41	0.45	0.34	0.32
BL_Up_Lf_I1	0.30	0.34	0.42	0.39	0.47	0.48	0.38	0.47	0.34	0.34
BL_Up_Rt_I1	0.30	0.34	0.42	0.37	0.44	0.46	0.38	0.46	0.33	0.33
BL_Up_Rt_I2	0.33	0.36	0.43	0.45	0.42	0.40	0.42	0.44	0.34	0.32
BL_Up_Rt_C	0.32	0.42	0.51	0.41	0.41	0.43	0.39	0.54	0.39	0.34
BL_Up_Rt_P1	0.24	0.38	0.34	0.31	0.34	0.37	0.31	0.37	0.37	0.31

Continue…

Trait	MD_Up_Lf_P2	MD_Up_Lf_P1	MD_Up_Lf_C	MD_Up_Lf_I2	MD_Up_Lf_I1	MD_Up_Rt_I1	MD_Up_Rt_I2	MD_Up_Rt_C	MD_Up_Rt_P1	MD_Up_Rt_P2
IC_Up_Rt_P2	**1.5E-12**	**7.1E-14**	**1.2E-10**	**2.1E-06**	**8.4E-08**	**7.9E-09**	**3.1E-07**	**7.3E-09**	**1.8E-14**	**4.1E-13**
IC_Up_Rt_P1	**3.9E-09**	**6.8E-16**	**6.3E-09**	**2.4E-05**	**1.9E-09**	**1.9E-10**	**9.1E-07**	**3.2E-09**	**1.1E-18**	**8.1E-09**
IC_Up_Rt_C	**1.0E-10**	**1.5E-13**	**2.1E-23**	**2.2E-12**	**4.2E-18**	**2.1E-20**	**6.8E-14**	**7.5E-20**	**6.4E-15**	**5.5E-11**
IC_Up_Rt_I2	**9.1E-10**	**5.4E-10**	**1.2E-15**	**8.9E-17**	**7.3E-15**	**1.8E-16**	**9.9E-20**	**8.7E-14**	**6.0E-12**	**6.4E-09**
IC_Up_Rt_I1	**3.6E-13**	**1.3E-12**	**3.0E-13**	**2.7E-10**	**3.4E-18**	**4.0E-21**	**9.6E-12**	**3.6E-15**	**1.2E-13**	**2.0E-12**
IC_Up_Lf_I1	**6.2E-11**	**1.1E-12**	**1.4E-13**	**1.3E-10**	**1.1E-21**	**9.5E-23**	**1.4E-11**	**6.0E-16**	**9.4E-14**	**2.8E-10**
IC_Up_Lf_I2	**2.2E-08**	**1.9E-09**	**1.2E-11**	**1.4E-17**	**2.2E-12**	**2.5E-13**	**3.0E-19**	**3.8E-11**	**9.2E-11**	**2.6E-06**
IC_Up_Lf_C	**4.3E-09**	**1.2E-11**	**3.6E-23**	**5.7E-13**	**2.5E-20**	**4.9E-20**	**1.5E-14**	**1.4E-21**	**9.2E-12**	**3.0E-08**
IC_Up_Lf_P1	**9.9E-08**	**4.3E-16**	**8.1E-08**	1.9E-04	**8.5E-09**	**3.7E-09**	**4.9E-06**	**8.4E-10**	**5.9E-16**	**9.6E-09**
IC_Up_Lf_P2	**7.4E-16**	**7.0E-14**	**2.4E-11**	**1.1E-07**	**6.7E-09**	**1.2E-09**	**1.0E-09**	**9.6E-11**	**2.2E-12**	**7.1E-14**
MD_Up_Lf_P2		**6.6E-69**	**6.5E-21**	**5.8E-13**	**6.2E-13**	**1.5E-14**	**2.2E-12**	**1.1E-23**	**6.1E-54**	**5.3E-139**
MD_Up_Lf_P1	0.70		**5.9E-30**	**4.5E-17**	**1.9E-22**	**5.3E-23**	**1.2E-19**	**1.1E-37**	**1.5E-125**	**4.8E-60**
MD_Up_Lf_C	0.39	0.49		**2.6E-31**	**1.6E-25**	**1.4E-25**	**5.1E-27**	**2.8E-138**	**7.2E-26**	**6.6E-24**
MD_Up_Lf_I2	0.30	0.37	0.46		**3.3E-44**	**2.9E-48**	**2.6E-141**	**2.9E-31**	**1.0E-21**	**3.0E-13**
MD_Up_Lf_I1	0.30	0.43	0.42	0.54		**8.3E-187**	**9.6E-41**	**2.3E-33**	**2.0E-26**	**6.3E-15**
MD_Up_Rt_I1	0.32	0.43	0.42	0.56	0.88		**2.6E-46**	**3.8E-32**	**1.7E-28**	**1.6E-16**
MD_Up_Rt_I2	0.30	0.40	0.43	0.83	0.53	0.55		**1.1E-30**	**8.5E-24**	**3.2E-17**
MD_Up_Rt_C	0.41	0.54	0.82	0.46	0.48	0.47	0.46		**3.5E-33**	**1.6E-28**
MD_Up_Rt_P1	0.64	0.84	0.46	0.42	0.46	0.48	0.44	0.51		**1.2E-55**
MD_Up_Rt_P2	0.83	0.66	0.41	0.31	0.33	0.34	0.35	0.45	0.64	
BL_Up_Lf_P2	0.50	0.58	0.48	0.45	0.48	0.48	0.44	0.51	0.61	0.51
BL_Up_Lf_P1	0.41	0.61	0.46	0.42	0.51	0.48	0.41	0.50	0.64	0.42
BL_Up_Lf_C	0.30	0.40	0.52	0.43	0.46	0.44	0.43	0.53	0.43	0.30
BL_Up_Lf_I2	0.25	0.34	0.43	0.49	0.40	0.41	0.52	0.42	0.37	0.27
BL_Up_Lf_I1	0.28	0.37	0.38	0.38	0.51	0.49	0.40	0.40	0.42	0.30
BL_Up_Rt_I1	0.27	0.37	0.41	0.38	0.50	0.48	0.41	0.42	0.42	0.29
BL_Up_Rt_I2	0.24	0.37	0.43	0.49	0.40	0.40	0.51	0.45	0.39	0.27
BL_Up_Rt_C	0.29	0.41	0.50	0.42	0.46	0.43	0.41	0.53	0.42	0.31
BL_Up_Rt_P1	0.42	0.61	0.44	0.45	0.50	0.48	0.46	0.49	0.67	0.44

Continue…

Trait	BL_Up_Lf_P2	BL_Up_Lf_P1	BL_Up_Lf_C	BL_Up_Lf_I2	BL_Up_Lf_I1	BL_Up_Rt_I1	BL_Up_Rt_I2	BL_Up_Rt_C	BL_Up_Rt_P1	BL_Up_Rt_P2
IC_Up_Rt_P2	**1.1E-12**	**2.3E-09**	**2.9E-17**	**1.7E-12**	**7.2E-13**	**1.3E-12**	**8.2E-15**	**1.5E-14**	**1.3E-07**	**2.0E-14**
IC_Up_Rt_P1	**3.7E-14**	**8.0E-19**	**4.0E-20**	**3.9E-14**	**2.2E-14**	**7.5E-14**	**5.0E-16**	**2.5E-21**	**1.3E-17**	**8.7E-13**
IC_Up_Rt_C	**2.8E-17**	**7.6E-14**	**9.0E-39**	**6.8E-29**	**5.7E-25**	**2.8E-25**	**3.7E-26**	**7.1E-38**	**2.2E-14**	**2.2E-14**
IC_Up_Rt_I2	**9.3E-14**	**1.9E-10**	**7.9E-24**	**1.0E-27**	**3.6E-21**	**3.2E-19**	**2.0E-28**	**3.5E-24**	**5.2E-12**	**3.4E-13**
IC_Up_Rt_I1	**9.2E-17**	**3.3E-13**	**1.1E-26**	**1.3E-24**	**3.4E-31**	**2.0E-28**	**2.7E-25**	**9.9E-24**	**1.2E-14**	**7.9E-16**
IC_Up_Lf_I1	**2.0E-18**	**1.4E-15**	**9.4E-28**	**1.3E-22**	**5.4E-33**	**2.1E-30**	**9.8E-23**	**1.4E-26**	**3.5E-17**	**3.8E-18**
IC_Up_Lf_I2	**8.1E-12**	**8.0E-10**	**6.5E-21**	**2.4E-24**	**9.2E-21**	**2.4E-20**	**1.6E-25**	**1.5E-21**	**6.4E-12**	**2.4E-12**
IC_Up_Lf_C	**1.6E-17**	**1.7E-16**	**6.8E-46**	**8.5E-29**	**2.3E-31**	**5.7E-31**	**4.0E-28**	**1.2E-43**	**9.8E-17**	**2.3E-15**
IC_Up_Lf_P1	**2.8E-13**	**9.5E-19**	**3.8E-21**	**7.0E-14**	**2.7E-14**	**3.5E-13**	**8.5E-14**	**9.1E-19**	**4.1E-17**	**5.5E-12**
IC_Up_Lf_P2	**2.8E-20**	**7.2E-13**	**4.7E-18**	**5.5E-14**	**4.8E-16**	**2.7E-15**	**3.1E-14**	**5.2E-16**	**6.8E-12**	**1.0E-20**
MD_Up_Lf_P2	**1.3E-36**	**3.9E-20**	**8.3E-13**	**2.3E-09**	**4.1E-11**	**6.7E-11**	**1.1E-08**	**2.7E-12**	**5.1E-21**	**5.4E-31**
MD_Up_Lf_P1	**9.8E-43**	**2.1E-49**	**6.3E-20**	**4.0E-14**	**5.1E-17**	**6.8E-17**	**2.1E-16**	**5.0E-21**	**6.9E-50**	**7.2E-34**
MD_Up_Lf_C	**3.7E-33**	**3.5E-26**	**9.7E-41**	**6.0E-27**	**3.9E-21**	**1.9E-24**	**5.2E-26**	**5.3E-37**	**3.4E-24**	**2.5E-28**
MD_Up_Lf_I2	**3.3E-29**	**5.4E-21**	**1.2E-26**	**5.6E-35**	**1.3E-20**	**9.2E-21**	**1.6E-34**	**8.6E-26**	**5.7E-25**	**1.2E-26**
MD_Up_Lf_I1	**5.5E-33**	**4.9E-33**	**1.2E-30**	**4.6E-23**	**3.2E-38**	**7.5E-36**	**5.5E-23**	**1.3E-30**	**4.3E-31**	**2.4E-29**
MD_Up_Rt_I1	**1.6E-32**	**2.9E-28**	**3.9E-28**	**4.3E-24**	**1.8E-34**	**2.3E-34**	**1.2E-22**	**5.7E-27**	**1.8E-28**	**8.0E-29**
MD_Up_Rt_I2	**1.6E-27**	**7.1E-21**	**2.2E-26**	**1.4E-39**	**5.9E-23**	**4.4E-24**	**4.6E-38**	**9.2E-24**	**1.3E-25**	**9.2E-30**
MD_Up_Rt_C	**1.3E-37**	**5.9E-31**	**5.4E-42**	**1.1E-25**	**1.2E-22**	**9.9E-25**	**1.9E-28**	**1.3E-41**	**7.7E-31**	**9.4E-35**
MD_Up_Rt_P1	**2.0E-49**	**1.2E-55**	**1.4E-22**	**1.9E-16**	**1.2E-21**	**1.3E-21**	**1.1E-18**	**2.6E-21**	**1.4E-64**	**2.8E-41**
MD_Up_Rt_P2	**5.5E-38**	**1.9E-21**	**3.2E-13**	**1.4E-10**	**7.4E-13**	**2.7E-12**	**3.4E-10**	**2.6E-13**	**2.7E-23**	**4.0E-35**
BL_Up_Lf_P2		**6.7E-114**	**8.1E-52**	**1.6E-31**	**6.8E-37**	**3.5E-37**	**4.6E-33**	**1.1E-50**	**2.1E-106**	**1.6E-194**
BL_Up_Lf_P1	0.82		**6.9E-58**	**6.4E-31**	**1.1E-38**	**6.0E-37**	**2.1E-30**	**1.4E-53**	**3.0E-175**	**6.7E-98**
BL_Up_Lf_C	0.58	0.65		**1.4E-67**	**1.7E-65**	**5.3E-71**	**3.0E-54**	**1.9E-143**	**1.4E-50**	**1.9E-48**
BL_Up_Lf_I2	0.47	0.50	0.65		**1.9E-67**	**9.5E-68**	**2.4E-161**	**1.5E-49**	**2.4E-30**	**4.0E-32**
BL_Up_Lf_I1	0.51	0.55	0.64	0.65		**1.9E-224**	**1.6E-67**	**7.5E-71**	**1.3E-39**	**2.5E-32**
BL_Up_Rt_I1	0.51	0.54	0.66	0.65	0.92		**7.1E-65**	**2.3E-68**	**1.2E-37**	**7.5E-32**
BL_Up_Rt_I2	0.48	0.50	0.59	0.86	0.65	0.64		**9.9E-54**	**9.3E-33**	**1.7E-31**
BL_Up_Rt_C	0.58	0.63	0.83	0.57	0.66	0.65	0.59		**2.4E-55**	**1.8E-47**
BL_Up_Rt_P1	0.80	0.90	0.61	0.49	0.55	0.54	0.51	0.64		**8.1E-104**

Continue…

Trait	IC_Lw_Rt_P2	IC_Lw_Rt_P1	IC_Lw_Rt_C	IC_Lw_Rt_I2	IC_Lw_Rt_I1	IC_Lw_Lf_I1	IC_Lw_Lf_I2	IC_Lw_Lf_C	IC_Lw_Lf_P1	IC_Lw_Lf_P2
IC_Up_Rt_P2	**2.4E-76**	**5.9E-57**	**6.1E-54**	**2.1E-43**	**2.0E-36**	**2.8E-33**	**5.7E-45**	**9.0E-59**	**1.9E-52**	**1.0E-53**
IC_Up_Rt_P1	**3.7E-70**	**4.2E-78**	**9.9E-51**	**2.7E-32**	**4.1E-29**	**2.6E-24**	**1.2E-37**	**6.1E-53**	**1.2E-71**	**2.3E-53**
IC_Up_Rt_C	**7.5E-53**	**3.9E-55**	**1.4E-89**	**8.1E-60**	**5.5E-46**	**5.3E-45**	**1.1E-61**	**4.2E-86**	**2.1E-46**	**1.1E-40**
IC_Up_Rt_I2	**9.6E-43**	**9.2E-36**	**3.6E-56**	**1.1E-58**	**2.9E-49**	**1.1E-45**	**1.8E-53**	**6.3E-60**	**9.2E-39**	**3.1E-43**
IC_Up_Rt_I1	**4.4E-37**	**8.4E-34**	**2.2E-47**	**1.4E-48**	**5.9E-50**	**9.8E-48**	**1.3E-50**	**4.3E-54**	**6.0E-37**	**1.9E-34**
IC_Up_Lf_I1	**4.3E-38**	**2.1E-37**	**1.5E-47**	**2.8E-49**	**3.9E-47**	**6.6E-50**	**2.5E-56**	**4.8E-59**	**4.2E-41**	**4.3E-36**
IC_Up_Lf_I2	**5.7E-42**	**1.0E-41**	**9.1E-51**	**9.5E-59**	**1.3E-55**	**7.9E-56**	**3.3E-67**	**1.3E-66**	**1.2E-46**	**4.3E-42**
IC_Up_Lf_C	**1.1E-56**	**6.9E-57**	**1.7E-82**	**3.7E-52**	**2.1E-47**	**6.4E-48**	**1.4E-61**	**7.5E-96**	**7.9E-61**	**9.0E-53**
IC_Up_Lf_P1	**6.6E-61**	**2.4E-76**	**1.4E-51**	**2.2E-34**	**1.4E-32**	**7.3E-30**	**1.6E-42**	**6.3E-59**	**1.2E-79**	**1.4E-61**
IC_Up_Lf_P2	**1.2E-71**	**6.4E-55**	**4.6E-48**	**1.2E-38**	**6.6E-34**	**6.1E-33**	**6.5E-40**	**8.3E-52**	**7.0E-60**	**3.3E-68**
MD_Up_Lf_P2	**1.4E-11**	**5.5E-08**	**1.3E-05**	**5.0E-06**	2.4E-04	**7.4E-05**	**3.3E-06**	**7.4E-06**	**6.6E-11**	**9.4E-11**
MD_Up_Lf_P1	**2.2E-09**	**8.9E-16**	**3.2E-11**	**1.2E-06**	**2.0E-05**	**1.1E-05**	**1.2E-08**	**3.1E-12**	**8.1E-19**	**2.8E-06**
MD_Up_Lf_C	**5.2E-10**	**5.1E-11**	**6.0E-17**	**3.0E-12**	**2.3E-07**	**4.1E-06**	**2.4E-10**	**1.1E-15**	**3.0E-09**	**2.6E-06**
MD_Up_Lf_I2	**5.6E-06**	**5.3E-08**	**4.9E-08**	**6.2E-06**	**1.1E-06**	**3.5E-07**	**1.3E-06**	**1.7E-09**	**3.0E-08**	**3.9E-05**
MD_Up_Lf_I1	**4.2E-08**	**4.5E-12**	**2.4E-15**	**8.4E-10**	**9.7E-10**	**2.5E-11**	**1.6E-10**	**5.9E-17**	**3.0E-12**	**2.2E-08**
MD_Up_Rt_I1	**8.8E-08**	**7.6E-11**	**4.7E-15**	**4.4E-12**	**6.7E-11**	**6.8E-12**	**2.9E-11**	**1.1E-14**	**3.4E-10**	**2.4E-08**
MD_Up_Rt_I2	**3.2E-06**	**2.0E-07**	**1.1E-07**	**2.0E-08**	**1.4E-08**	**1.3E-09**	**2.7E-07**	**9.9E-10**	**1.1E-08**	**2.8E-05**
MD_Up_Rt_C	**2.2E-10**	**1.7E-13**	**2.8E-13**	**3.1E-09**	**1.6E-06**	**5.9E-06**	**3.3E-09**	**5.8E-13**	**4.5E-11**	**4.4E-06**
MD_Up_Rt_P1	**7.2E-11**	**3.1E-16**	**1.1E-10**	**7.3E-08**	**1.8E-06**	**4.4E-06**	**7.0E-09**	**2.4E-11**	**1.2E-17**	**2.4E-06**
MD_Up_Rt_P2	**1.1E-10**	**3.3E-08**	**6.6E-06**	**6.0E-05**	2.1E-03	3.8E-04	**6.4E-05**	**6.0E-06**	**2.3E-08**	**5.9E-08**
BL_Up_Lf_P2	**3.0E-12**	**4.0E-16**	**7.5E-15**	**6.5E-11**	**1.4E-07**	**8.7E-09**	**1.3E-11**	**4.4E-14**	**7.6E-18**	**3.5E-11**
BL_Up_Lf_P1	**2.9E-09**	**1.9E-21**	**3.2E-16**	**3.2E-09**	**7.7E-08**	**3.2E-08**	**2.2E-12**	**1.7E-16**	**1.2E-22**	**2.1E-09**
BL_Up_Lf_C	**5.9E-22**	**2.9E-24**	**1.1E-31**	**4.1E-20**	**3.3E-16**	**4.4E-17**	**2.4E-23**	**1.5E-31**	**1.4E-25**	**1.1E-17**
BL_Up_Lf_I2	**1.8E-13**	**4.4E-17**	**4.4E-20**	**1.9E-22**	**4.4E-17**	**1.6E-17**	**3.0E-21**	**1.8E-20**	**9.0E-17**	**3.6E-11**
BL_Up_Lf_I1	**1.1E-13**	**1.2E-18**	**1.2E-22**	**4.6E-20**	**3.3E-17**	**3.0E-20**	**2.5E-21**	**5.7E-25**	**6.9E-20**	**1.6E-14**
BL_Up_Rt_I1	**2.9E-13**	**2.3E-16**	**6.7E-22**	**5.2E-20**	**6.2E-17**	**2.5E-18**	**1.3E-20**	**1.3E-24**	**1.3E-18**	**5.0E-14**
BL_Up_Rt_I2	**2.7E-13**	**1.7E-18**	**2.2E-22**	**1.9E-21**	**4.5E-17**	**1.0E-17**	**1.4E-21**	**9.3E-24**	**9.5E-15**	**2.6E-10**
BL_Up_Rt_C	**2.4E-18**	**1.1E-24**	**1.4E-33**	**1.6E-21**	**7.1E-16**	**1.7E-16**	**5.0E-23**	**7.6E-31**	**2.8E-26**	**1.6E-15**
BL_Up_Rt_P1	**1.4E-08**	**2.9E-19**	**7.8E-16**	**1.3E-09**	**1.9E-08**	**1.5E-08**	**2.6E-12**	**5.5E-17**	**4.5E-22**	**2.5E-09**

Continue…

Trait	MD_Lw_Lf_P2	MD_Lw_Lf_P1	MD_Lw_Lf_C	MD_Lw_Lf_I2	MD_Lw_Lf_I1	MD_Lw_Rt_I1	MD_Lw_Rt_I2	MD_Lw_Rt_C	MD_Lw_Rt_P1	MD_Lw_Rt_P2
IC_Up_Rt_P2	**4.3E-07**	**2.3E-06**	**9.6E-08**	5.5E-04	2.2E-03	**7.1E-05**	**8.3E-06**	**1.9E-09**	**5.1E-07**	1.4E-04
IC_Up_Rt_P1	**2.5E-09**	**3.7E-12**	**9.6E-09**	**3.1E-05**	**3.5E-05**	**1.2E-05**	**5.3E-07**	**3.2E-09**	**1.9E-10**	**3.3E-06**
IC_Up_Rt_C	**2.0E-08**	**3.7E-11**	**2.8E-17**	**8.4E-09**	**2.0E-09**	**1.1E-10**	**6.2E-10**	**1.9E-20**	**6.9E-11**	**7.6E-08**
IC_Up_Rt_I2	**1.6E-07**	**3.8E-10**	**2.2E-14**	**1.7E-08**	**2.5E-09**	**2.2E-10**	**5.9E-10**	**1.2E-14**	**2.3E-10**	**2.7E-07**
IC_Up_Rt_I1	**7.3E-09**	**4.2E-11**	**3.9E-13**	**1.5E-09**	**1.7E-11**	**1.3E-12**	**1.8E-12**	**4.2E-15**	**2.1E-10**	**1.6E-11**
IC_Up_Lf_I1	**8.5E-10**	**3.8E-13**	**1.2E-15**	**1.0E-11**	**2.8E-12**	**2.5E-12**	**5.0E-14**	**1.8E-17**	**4.6E-11**	**2.3E-12**
IC_Up_Lf_I2	**1.8E-07**	**1.9E-11**	**2.2E-10**	**1.7E-07**	**8.9E-07**	**7.5E-09**	**6.5E-10**	**2.1E-11**	**3.5E-09**	**1.1E-07**
IC_Up_Lf_C	**3.3E-11**	**2.5E-15**	**1.2E-16**	**4.9E-10**	**9.4E-10**	**1.4E-10**	**5.6E-13**	**9.8E-21**	**5.2E-12**	**2.9E-10**
IC_Up_Lf_P1	**2.2E-08**	**1.5E-14**	**1.5E-09**	**1.4E-05**	**5.2E-05**	**1.7E-05**	**1.7E-06**	**3.1E-09**	**2.3E-11**	**1.9E-05**
IC_Up_Lf_P2	**4.3E-07**	**2.3E-08**	**1.2E-09**	**8.8E-06**	8.7E-04	**4.3E-05**	**7.4E-07**	**1.9E-09**	**8.0E-08**	1.4E-04
MD_Up_Lf_P2	**5.1E-29**	**3.6E-14**	**8.7E-14**	**2.2E-11**	**6.9E-13**	**2.9E-11**	**6.6E-11**	**1.3E-08**	**1.6E-15**	**2.7E-31**
MD_Up_Lf_P1	**5.5E-28**	**8.3E-32**	**3.4E-20**	**6.3E-15**	**1.3E-15**	**1.1E-15**	**2.1E-18**	**1.2E-19**	**2.3E-35**	**7.5E-35**
MD_Up_Lf_C	**1.2E-17**	**1.1E-17**	**9.0E-29**	**3.8E-16**	**7.2E-14**	**8.5E-14**	**6.1E-20**	**8.4E-40**	**2.6E-19**	**2.5E-21**
MD_Up_Lf_I2	**1.4E-17**	**5.3E-14**	**5.9E-19**	**4.6E-28**	**1.2E-26**	**1.3E-29**	**7.4E-33**	**1.6E-17**	**7.8E-16**	**5.8E-21**
MD_Up_Lf_I1	**9.6E-19**	**7.1E-23**	**2.4E-23**	**8.8E-34**	**1.0E-43**	**9.1E-45**	**3.7E-44**	**3.6E-23**	**9.3E-23**	**7.4E-20**
MD_Up_Rt_I1	**2.5E-18**	**3.3E-19**	**1.9E-20**	**8.8E-34**	**4.3E-47**	**3.3E-47**	**1.3E-46**	**4.0E-22**	**2.3E-20**	**7.9E-18**
MD_Up_Rt_I2	**1.2E-16**	**1.3E-13**	**1.2E-15**	**6.3E-25**	**5.2E-27**	**1.6E-29**	**2.1E-31**	**1.8E-14**	**1.5E-15**	**1.2E-19**
MD_Up_Rt_C	**2.6E-21**	**9.0E-20**	**1.2E-32**	**3.8E-20**	**4.6E-19**	**1.6E-20**	**1.5E-21**	**1.7E-40**	**1.7E-21**	**1.4E-24**
MD_Up_Rt_P1	**1.5E-26**	**1.5E-31**	**2.6E-19**	**4.2E-16**	**6.6E-18**	**5.4E-18**	**3.1E-18**	**4.9E-15**	**3.1E-32**	**7.7E-33**
MD_Up_Rt_P2	**1.6E-28**	**9.0E-14**	**3.9E-13**	**7.2E-12**	**1.0E-13**	**7.2E-14**	**3.3E-12**	**3.2E-10**	**6.4E-15**	**5.8E-34**
BL_Up_Lf_P2	**7.5E-25**	**1.0E-25**	**4.2E-22**	**5.1E-18**	**8.8E-19**	**3.4E-17**	**2.0E-23**	**7.8E-21**	**2.3E-27**	**7.4E-33**
BL_Up_Lf_P1	**5.7E-20**	**3.8E-32**	**1.6E-18**	**3.5E-15**	**1.3E-14**	**1.9E-13**	**1.3E-18**	**2.7E-18**	**6.7E-32**	**7.9E-26**
BL_Up_Lf_C	**2.6E-17**	**6.4E-26**	**1.5E-25**	**1.2E-19**	**7.5E-17**	**7.8E-14**	**3.6E-20**	**5.0E-32**	**5.3E-25**	**4.1E-19**
BL_Up_Lf_I2	**1.1E-13**	**2.6E-14**	**5.5E-13**	**4.9E-14**	**3.1E-12**	**2.0E-13**	**1.8E-17**	**1.4E-15**	**3.7E-14**	**1.6E-15**
BL_Up_Lf_I1	**2.4E-14**	**1.9E-15**	**1.3E-16**	**3.9E-16**	**7.0E-15**	**2.1E-14**	**1.1E-18**	**2.1E-17**	**1.1E-13**	**1.2E-17**
BL_Up_Rt_I1	**3.3E-15**	**5.8E-17**	**1.5E-15**	**1.1E-16**	**2.9E-15**	**1.1E-14**	**5.1E-22**	**2.1E-20**	**1.9E-15**	**1.8E-18**
BL_Up_Rt_I2	**2.9E-13**	**1.6E-15**	**4.0E-14**	**1.8E-14**	**4.2E-13**	**2.2E-15**	**9.3E-17**	**3.6E-19**	**1.5E-15**	**2.5E-15**
BL_Up_Rt_C	**2.7E-17**	**5.3E-21**	**4.1E-25**	**3.4E-16**	**1.0E-14**	**6.0E-13**	**1.3E-17**	**7.9E-33**	**1.3E-20**	**1.2E-19**
BL_Up_Rt_P1	**5.5E-19**	**1.0E-29**	**1.1E-18**	**6.0E-14**	**1.3E-12**	**9.8E-13**	**5.1E-17**	**1.5E-17**	**1.6E-30**	**1.8E-26**

Continue…

Trait	BL_Lw_Lf_P2	BL_Lw_Lf_P1	BL_Lw_Lf_C	BL_Lw_Lf_I2	BL_Lw_Lf_I1	BL_Lw_Rt_I1	BL_Lw_Rt_I2	BL_Lw_Rt_C	BL_Lw_Rt_P1	BL_Lw_Rt_P2
IC_Up_Rt_P2	**5.7E-06**	**1.4E-08**	**8.6E-20**	**3.4E-10**	**9.8E-10**	**7.7E-09**	**2.2E-09**	**1.3E-18**	**1.3E-09**	**5.6E-05**
IC_Up_Rt_P1	**1.3E-09**	**8.3E-15**	**8.6E-22**	**2.0E-11**	**4.0E-11**	**1.1E-10**	**6.4E-11**	**2.9E-22**	**1.4E-13**	**3.2E-07**
IC_Up_Rt_C	**7.5E-13**	**1.5E-17**	**1.1E-34**	**1.2E-18**	**9.0E-17**	**1.8E-15**	**2.5E-15**	**4.1E-39**	**3.3E-15**	**2.3E-10**
IC_Up_Rt_I2	**5.3E-11**	**1.1E-17**	**2.1E-23**	**9.8E-16**	**1.6E-14**	**3.1E-14**	**3.8E-14**	**2.9E-24**	**9.3E-14**	**3.1E-10**
IC_Up_Rt_I1	**2.7E-13**	**2.0E-18**	**3.9E-21**	**9.7E-18**	**1.1E-16**	**6.3E-16**	**1.6E-14**	**1.3E-23**	**6.7E-15**	**1.9E-11**
IC_Up_Lf_I1	**1.1E-13**	**3.2E-18**	**1.9E-24**	**5.2E-20**	**4.7E-20**	**5.0E-19**	**2.7E-17**	**6.9E-26**	**1.1E-13**	**1.3E-11**
IC_Up_Lf_I2	**9.0E-10**	**4.5E-15**	**1.2E-19**	**8.9E-15**	**1.4E-13**	**1.2E-13**	**6.6E-14**	**1.7E-20**	**2.0E-12**	**4.7E-09**
IC_Up_Lf_C	**9.2E-16**	**1.2E-20**	**1.0E-36**	**1.1E-19**	**5.3E-19**	**3.9E-17**	**1.5E-17**	**1.3E-39**	**6.1E-19**	**3.3E-12**
IC_Up_Lf_P1	**1.6E-09**	**9.7E-14**	**5.1E-21**	**7.8E-11**	**2.3E-09**	**1.9E-10**	**5.9E-11**	**5.7E-21**	**8.3E-13**	**1.6E-07**
IC_Up_Lf_P2	**4.0E-10**	**3.0E-11**	**7.3E-23**	**2.5E-11**	**8.9E-10**	**1.1E-09**	**1.4E-10**	**5.5E-23**	**4.2E-12**	**1.6E-08**
MD_Up_Lf_P2	**1.8E-10**	**1.1E-11**	**4.2E-11**	**7.5E-08**	**6.3E-06**	**8.7E-07**	**2.9E-06**	**3.8E-10**	**1.0E-09**	**2.4E-13**
MD_Up_Lf_P1	**3.0E-16**	**3.0E-24**	**1.3E-18**	**3.5E-10**	**4.1E-09**	**1.2E-08**	**9.6E-09**	**1.0E-17**	**8.1E-23**	**9.4E-21**
MD_Up_Lf_C	**6.2E-18**	**2.6E-16**	**2.4E-29**	**6.2E-16**	**8.2E-12**	**7.8E-12**	**4.5E-16**	**6.5E-28**	**3.4E-17**	**3.5E-17**
MD_Up_Lf_I2	**4.2E-15**	**6.6E-15**	**7.1E-15**	**1.7E-15**	**6.9E-16**	**2.3E-16**	**1.3E-14**	**8.8E-13**	**2.4E-14**	**7.7E-17**
MD_Up_Lf_I1	**2.5E-21**	**3.7E-21**	**1.9E-18**	**5.8E-22**	**2.3E-26**	**2.3E-26**	**2.7E-21**	**9.0E-20**	**4.5E-24**	**1.3E-20**
MD_Up_Rt_I1	**4.4E-23**	**2.2E-19**	**6.9E-19**	**1.4E-21**	**3.3E-25**	**3.0E-26**	**1.7E-19**	**8.5E-21**	**4.2E-23**	**7.9E-22**
MD_Up_Rt_I2	**3.8E-14**	**7.9E-15**	**9.2E-14**	**1.7E-15**	**8.6E-17**	**1.6E-15**	**2.2E-14**	**5.8E-13**	**1.8E-13**	**9.6E-18**
MD_Up_Rt_C	**3.4E-20**	**2.9E-19**	**1.6E-31**	**3.8E-18**	**7.7E-16**	**6.8E-16**	**2.5E-19**	**5.3E-30**	**5.5E-21**	**8.7E-22**
MD_Up_Rt_P1	**1.6E-17**	**1.0E-23**	**3.0E-17**	**4.1E-11**	**2.3E-11**	**1.0E-11**	**9.2E-12**	**5.5E-16**	**3.3E-24**	**4.0E-23**
MD_Up_Rt_P2	**1.5E-11**	**5.8E-12**	**1.0E-10**	**1.6E-07**	**4.4E-06**	**1.0E-06**	**1.9E-06**	**8.6E-11**	**2.5E-10**	**1.5E-14**
BL_Up_Lf_P2	**2.6E-54**	**8.3E-49**	**1.3E-39**	**7.8E-31**	**1.2E-24**	**1.2E-24**	**9.7E-28**	**1.3E-34**	**3.3E-49**	**1.1E-68**
BL_Up_Lf_P1	**9.8E-40**	**1.3E-53**	**3.5E-39**	**4.3E-25**	**1.6E-21**	**1.1E-21**	**5.0E-24**	**8.8E-35**	**1.8E-51**	**1.3E-41**
BL_Up_Lf_C	**2.2E-34**	**2.7E-48**	**1.3E-71**	**8.9E-46**	**1.4E-42**	**6.5E-38**	**2.9E-41**	**6.4E-69**	**5.1E-48**	**1.6E-34**
BL_Up_Lf_I2	**5.5E-23**	**1.5E-25**	**7.6E-32**	**5.8E-37**	**1.0E-34**	**6.4E-34**	**6.6E-34**	**7.3E-32**	**5.7E-28**	**4.2E-23**
BL_Up_Lf_I1	**2.1E-28**	**9.0E-39**	**2.7E-39**	**3.0E-50**	**3.4E-55**	**6.7E-54**	**2.3E-45**	**6.9E-41**	**1.2E-37**	**2.0E-27**
BL_Up_Rt_I1	**1.3E-27**	**4.1E-36**	**8.1E-41**	**3.2E-49**	**2.5E-56**	**1.6E-55**	**1.1E-46**	**2.4E-42**	**7.6E-34**	**8.9E-28**
BL_Up_Rt_I2	**2.4E-21**	**4.1E-28**	**1.9E-30**	**6.1E-35**	**5.3E-35**	**1.0E-31**	**3.7E-33**	**4.4E-31**	**5.8E-28**	**6.9E-24**
BL_Up_Rt_C	**6.3E-34**	**3.5E-50**	**1.0E-71**	**2.5E-44**	**6.2E-40**	**5.0E-38**	**3.5E-40**	**5.0E-70**	**4.4E-50**	**1.6E-35**
BL_Up_Rt_P1	**2.6E-40**	**5.0E-57**	**6.9E-38**	**4.7E-27**	**3.1E-21**	**1.2E-22**	**2.2E-24**	**1.2E-34**	**9.0E-54**	**9.3E-46**

Continue…

Trait	IC_Up_Rt_P2	IC_Up_Rt_P1	IC_Up_Rt_C	IC_Up_Rt_I2	IC_Up_Rt_I1	IC_Up_Lf_I1	IC_Up_Lf_I2	IC_Up_Lf_C	IC_Up_Lf_P1	IC_Up_Lf_P2
BL_Up_Rt_P2	0.32	0.32	0.32	0.30	0.33	0.36	0.29	0.33	0.31	0.39
IC_Lw_Rt_P2	0.68	0.70	0.59	0.54	0.51	0.51	0.54	0.61	0.67	0.67
IC_Lw_Rt_P1	0.63	0.73	0.62	0.52	0.51	0.53	0.56	0.63	0.72	0.63
IC_Lw_Rt_C	0.59	0.61	0.72	0.60	0.56	0.56	0.58	0.69	0.62	0.57
IC_Lw_Rt_I2	0.54	0.51	0.62	0.61	0.57	0.57	0.61	0.58	0.52	0.52
IC_Lw_Rt_I1	0.50	0.48	0.55	0.57	0.57	0.56	0.60	0.56	0.51	0.49
IC_Lw_Lf_I1	0.48	0.44	0.55	0.55	0.56	0.57	0.60	0.56	0.49	0.48
IC_Lw_Lf_I2	0.55	0.54	0.62	0.59	0.57	0.60	0.65	0.62	0.57	0.52
IC_Lw_Lf_C	0.61	0.62	0.71	0.62	0.59	0.61	0.64	0.73	0.65	0.59
IC_Lw_Lf_P1	0.62	0.70	0.58	0.54	0.53	0.55	0.58	0.65	0.73	0.65
IC_Lw_Lf_P2	0.60	0.63	0.53	0.54	0.49	0.50	0.54	0.59	0.67	0.66
MD_Lw_Lf_P2	0.22	0.27	0.24	0.22	0.24	0.26	0.22	0.28	0.26	0.22
MD_Lw_Lf_P1	0.21	0.31	0.29	0.28	0.29	0.32	0.30	0.34	0.35	0.25
MD_Lw_Lf_C	0.22	0.26	0.35	0.32	0.30	0.33	0.26	0.34	0.27	0.26
MD_Lw_Lf_I2	0.15	0.19	0.24	0.24	0.25	0.28	0.22	0.26	0.20	0.19
MD_Lw_Lf_I1	0.13	0.19	0.25	0.25	0.28	0.29	0.21	0.25	0.18	0.14
MD_Lw_Rt_I1	0.17	0.20	0.27	0.26	0.29	0.29	0.24	0.27	0.20	0.17
MD_Lw_Rt_I2	0.19	0.23	0.26	0.26	0.29	0.31	0.26	0.30	0.22	0.21
MD_Lw_Rt_C	0.25	0.27	0.38	0.32	0.32	0.35	0.28	0.38	0.27	0.25
MD_Lw_Rt_P1	0.22	0.29	0.29	0.28	0.28	0.29	0.26	0.30	0.30	0.24
MD_Lw_Rt_P2	0.16	0.21	0.23	0.22	0.28	0.29	0.23	0.26	0.20	0.16
BL_Lw_Lf_P2	0.19	0.28	0.30	0.28	0.31	0.31	0.26	0.33	0.28	0.27
BL_Lw_Lf_P1	0.25	0.35	0.37	0.37	0.38	0.38	0.34	0.40	0.34	0.30
BL_Lw_Lf_C	0.37	0.42	0.49	0.40	0.38	0.41	0.37	0.50	0.41	0.40
BL_Lw_Lf_I2	0.26	0.30	0.36	0.33	0.35	0.38	0.32	0.37	0.29	0.28
BL_Lw_Lf_I1	0.26	0.30	0.34	0.32	0.34	0.38	0.31	0.36	0.27	0.26
BL_Lw_Rt_I1	0.24	0.29	0.33	0.32	0.33	0.37	0.31	0.35	0.29	0.26
BL_Lw_Rt_I2	0.25	0.29	0.33	0.31	0.32	0.35	0.31	0.35	0.30	0.27
BL_Lw_Rt_C	0.36	0.42	0.51	0.41	0.41	0.43	0.38	0.52	0.41	0.40
BL_Lw_Rt_P1	0.27	0.33	0.34	0.33	0.34	0.33	0.31	0.38	0.32	0.31
BL_Lw_Rt_P2	0.17	0.23	0.27	0.27	0.28	0.28	0.25	0.29	0.24	0.24

Continue…

Trait	MD_Up_Lf_P2	MD_Up_Lf_P1	MD_Up_Lf_C	MD_Up_Lf_I2	MD_Up_Lf_I1	MD_Up_Rt_I1	MD_Up_Rt_I2	MD_Up_Rt_C	MD_Up_Rt_P1	MD_Up_Rt_P2
BL_Up_Rt_P2	0.47	0.52	0.44	0.43	0.45	0.45	0.46	0.49	0.57	0.49
IC_Lw_Rt_P2	0.29	0.27	0.26	0.19	0.23	0.23	0.20	0.27	0.30	0.27
IC_Lw_Rt_P1	0.24	0.36	0.29	0.24	0.30	0.28	0.23	0.32	0.36	0.25
IC_Lw_Rt_C	0.19	0.30	0.34	0.23	0.33	0.32	0.22	0.30	0.29	0.19
IC_Lw_Rt_I2	0.19	0.22	0.29	0.19	0.26	0.29	0.23	0.25	0.24	0.17
IC_Lw_Rt_I1	0.16	0.19	0.22	0.20	0.25	0.27	0.24	0.20	0.22	0.13
IC_Lw_Lf_I1	0.17	0.20	0.19	0.21	0.28	0.28	0.25	0.19	0.21	0.15
IC_Lw_Lf_I2	0.20	0.26	0.26	0.20	0.27	0.28	0.22	0.25	0.26	0.17
IC_Lw_Lf_C	0.19	0.31	0.33	0.25	0.34	0.32	0.25	0.30	0.30	0.19
IC_Lw_Lf_P1	0.29	0.39	0.26	0.24	0.31	0.28	0.25	0.29	0.38	0.25
IC_Lw_Lf_P2	0.27	0.22	0.20	0.18	0.24	0.24	0.18	0.19	0.22	0.23
MD_Lw_Lf_P2	0.46	0.48	0.35	0.35	0.37	0.36	0.35	0.39	0.47	0.45
MD_Lw_Lf_P1	0.33	0.51	0.37	0.33	0.42	0.39	0.32	0.39	0.50	0.33
MD_Lw_Lf_C	0.31	0.40	0.44	0.36	0.40	0.38	0.33	0.47	0.39	0.30
MD_Lw_Lf_I2	0.28	0.35	0.33	0.44	0.48	0.48	0.42	0.37	0.36	0.29
MD_Lw_Lf_I1	0.30	0.35	0.31	0.43	0.54	0.55	0.43	0.36	0.38	0.31
MD_Lw_Rt_I1	0.28	0.36	0.31	0.45	0.54	0.56	0.45	0.38	0.38	0.31
MD_Lw_Rt_I2	0.27	0.39	0.37	0.47	0.54	0.55	0.47	0.39	0.38	0.29
MD_Lw_Rt_C	0.24	0.40	0.51	0.35	0.40	0.39	0.32	0.52	0.35	0.26
MD_Lw_Rt_P1	0.35	0.53	0.38	0.35	0.42	0.40	0.35	0.41	0.51	0.34
MD_Lw_Rt_P2	0.47	0.53	0.39	0.39	0.38	0.36	0.38	0.42	0.51	0.49
BL_Lw_Lf_P2	0.27	0.37	0.36	0.33	0.39	0.40	0.32	0.38	0.38	0.28
BL_Lw_Lf_P1	0.30	0.45	0.35	0.34	0.41	0.39	0.34	0.39	0.44	0.30
BL_Lw_Lf_C	0.28	0.39	0.45	0.32	0.36	0.36	0.31	0.46	0.37	0.27
BL_Lw_Lf_I2	0.23	0.28	0.33	0.33	0.39	0.39	0.33	0.36	0.30	0.22
BL_Lw_Lf_I1	0.19	0.27	0.28	0.33	0.43	0.42	0.34	0.33	0.30	0.19
BL_Lw_Rt_I1	0.21	0.26	0.28	0.34	0.43	0.43	0.33	0.33	0.31	0.21
BL_Lw_Rt_I2	0.20	0.26	0.33	0.32	0.39	0.37	0.32	0.37	0.31	0.20
BL_Lw_Rt_C	0.26	0.38	0.44	0.30	0.37	0.38	0.30	0.45	0.36	0.27
BL_Lw_Rt_P1	0.27	0.43	0.36	0.33	0.43	0.42	0.32	0.40	0.45	0.28
BL_Lw_Rt_P2	0.31	0.41	0.35	0.35	0.38	0.39	0.36	0.39	0.44	0.32

Continue…

Trait	BL_Up_Lf_P2	BL_Up_Lf_P1	BL_Up_Lf_C	BL_Up_Lf_I2	BL_Up_Lf_I1	BL_Up_Rt_I1	BL_Up_Rt_I2	BL_Up_Rt_C	BL_Up_Rt_P1	BL_Up_Rt_P2
BL_Up_Rt_P2	0.90	0.78	0.57	0.47	0.48	0.47	0.47	0.56	0.80	
IC_Lw_Rt_P2	0.29	0.27	0.39	0.31	0.31	0.31	0.31	0.36	0.26	0.29
IC_Lw_Rt_P1	0.36	0.42	0.43	0.36	0.38	0.36	0.38	0.43	0.40	0.35
IC_Lw_Rt_C	0.32	0.36	0.47	0.37	0.40	0.39	0.40	0.48	0.36	0.31
IC_Lw_Rt_I2	0.27	0.27	0.37	0.40	0.38	0.37	0.39	0.39	0.27	0.27
IC_Lw_Rt_I1	0.22	0.24	0.33	0.35	0.35	0.34	0.35	0.33	0.25	0.23
IC_Lw_Lf_I1	0.24	0.25	0.34	0.35	0.38	0.36	0.35	0.34	0.26	0.26
IC_Lw_Lf_I2	0.28	0.31	0.40	0.39	0.39	0.38	0.39	0.40	0.31	0.29
IC_Lw_Lf_C	0.31	0.37	0.46	0.38	0.42	0.41	0.41	0.46	0.37	0.33
IC_Lw_Lf_P1	0.37	0.43	0.44	0.36	0.39	0.38	0.34	0.45	0.42	0.37
IC_Lw_Lf_P2	0.28	0.27	0.35	0.28	0.32	0.32	0.27	0.33	0.27	0.27
MD_Lw_Lf_P2	0.42	0.41	0.35	0.31	0.32	0.33	0.31	0.35	0.40	0.40
MD_Lw_Lf_P1	0.45	0.51	0.45	0.33	0.35	0.36	0.35	0.40	0.49	0.41
MD_Lw_Lf_C	0.39	0.39	0.42	0.30	0.34	0.33	0.31	0.42	0.39	0.34
MD_Lw_Lf_I2	0.36	0.35	0.37	0.31	0.34	0.34	0.32	0.33	0.33	0.33
MD_Lw_Lf_I1	0.36	0.34	0.34	0.29	0.32	0.32	0.30	0.32	0.32	0.32
MD_Lw_Rt_I1	0.35	0.33	0.31	0.30	0.32	0.32	0.33	0.30	0.32	0.32
MD_Lw_Rt_I2	0.41	0.39	0.37	0.35	0.36	0.39	0.34	0.35	0.37	0.38
MD_Lw_Rt_C	0.38	0.38	0.47	0.33	0.35	0.38	0.37	0.47	0.38	0.35
MD_Lw_Rt_P1	0.46	0.51	0.44	0.33	0.32	0.35	0.35	0.40	0.50	0.42
MD_Lw_Rt_P2	0.48	0.46	0.37	0.33	0.36	0.36	0.33	0.37	0.47	0.46
BL_Lw_Lf_P2	0.60	0.56	0.49	0.40	0.45	0.44	0.39	0.49	0.56	0.57
BL_Lw_Lf_P1	0.60	0.63	0.59	0.44	0.54	0.52	0.47	0.60	0.65	0.57
BL_Lw_Lf_C	0.52	0.55	0.66	0.47	0.52	0.52	0.46	0.66	0.54	0.47
BL_Lw_Lf_I2	0.46	0.45	0.55	0.50	0.58	0.57	0.49	0.54	0.47	0.41
BL_Lw_Lf_I1	0.42	0.42	0.53	0.49	0.60	0.60	0.49	0.52	0.41	0.36
BL_Lw_Rt_I1	0.42	0.42	0.51	0.48	0.59	0.60	0.47	0.51	0.43	0.37
BL_Lw_Rt_I2	0.44	0.44	0.53	0.48	0.55	0.56	0.48	0.52	0.44	0.38
BL_Lw_Rt_C	0.49	0.52	0.65	0.47	0.53	0.53	0.46	0.65	0.52	0.44
BL_Lw_Rt_P1	0.60	0.62	0.59	0.46	0.53	0.51	0.47	0.60	0.63	0.57
BL_Lw_Rt_P2	0.66	0.57	0.49	0.41	0.44	0.44	0.41	0.50	0.59	0.62

Continue…

Trait	IC_Lw_Rt_P2	IC_Lw_Rt_P1	IC_Lw_Rt_C	IC_Lw_Rt_I2	IC_Lw_Rt_I1	IC_Lw_Lf_I1	IC_Lw_Lf_I2	IC_Lw_Lf_C	IC_Lw_Lf_P1	IC_Lw_Lf_P2
BL_Up_Rt_P2	**3.5E-12**	**5.0E-16**	**7.9E-14**	**8.1E-11**	**5.6E-08**	**1.0E-09**	**2.4E-12**	**1.4E-15**	**2.0E-17**	**1.7E-10**
IC_Lw_Rt_P2		**2.1E-96**	**3.2E-56**	**4.0E-41**	**3.2E-40**	**5.2E-36**	**6.0E-46**	**5.0E-63**	**9.9E-77**	**1.4E-131**
IC_Lw_Rt_P1	0.77		**7.3E-79**	**6.5E-48**	**7.8E-39**	**2.5E-38**	**2.4E-61**	**5.5E-80**	**4.3E-126**	**8.4E-68**
IC_Lw_Rt_C	0.60	0.71		**2.8E-94**	**3.7E-61**	**3.8E-57**	**1.9E-90**	**2.0E-171**	**1.9E-64**	**1.0E-49**
IC_Lw_Rt_I2	0.53	0.59	0.73		**1.3E-137**	**1.1E-125**	**6.1E-158**	**3.1E-86**	**4.3E-51**	**6.7E-40**
IC_Lw_Rt_I1	0.52	0.54	0.62	0.82		**1.3E-202**	**5.0E-126**	**1.4E-73**	**7.1E-45**	**2.1E-42**
IC_Lw_Lf_I1	0.50	0.53	0.60	0.80	0.90		**5.7E-142**	**1.9E-74**	**1.4E-40**	**8.1E-35**
IC_Lw_Lf_I2	0.55	0.65	0.72	0.85	0.80	0.82		**1.4E-120**	**2.3E-58**	**3.6E-43**
IC_Lw_Lf_C	0.63	0.71	0.86	0.70	0.66	0.67	0.79		**2.9E-83**	**4.9E-61**
IC_Lw_Lf_P1	0.71	0.82	0.66	0.60	0.57	0.55	0.64	0.72		**3.0E-93**
IC_Lw_Lf_P2	0.81	0.68	0.57	0.52	0.54	0.49	0.54	0.62	0.76	
MD_Lw_Lf_P2	0.29	0.33	0.26	0.23	0.24	0.20	0.25	0.28	0.34	0.25
MD_Lw_Lf_P1	0.32	0.42	0.35	0.31	0.30	0.28	0.32	0.37	0.42	0.28
MD_Lw_Lf_C	0.37	0.38	0.34	0.26	0.21	0.21	0.26	0.34	0.34	0.35
MD_Lw_Lf_I2	0.22	0.27	0.18	0.18	0.19	0.20	0.21	0.21	0.24	0.21
MD_Lw_Lf_I1	0.20	0.23	0.20	0.20	0.23	0.23	0.21	0.20	0.16	0.19
MD_Lw_Rt_I1	0.22	0.26	0.23	0.19	0.24	0.22	0.22	0.22	0.19	0.18
MD_Lw_Rt_I2	0.18	0.24	0.22	0.23	0.22	0.24	0.24	0.25	0.24	0.15
MD_Lw_Rt_C	0.34	0.36	0.37	0.26	0.24	0.22	0.29	0.37	0.32	0.30
MD_Lw_Rt_P1	0.34	0.40	0.33	0.29	0.29	0.26	0.30	0.33	0.37	0.28
MD_Lw_Rt_P2	0.28	0.28	0.25	0.24	0.21	0.21	0.24	0.26	0.27	0.20
BL_Lw_Lf_P2	0.24	0.35	0.30	0.25	0.20	0.21	0.26	0.29	0.35	0.23
BL_Lw_Lf_P1	0.34	0.41	0.38	0.32	0.31	0.28	0.32	0.37	0.40	0.34
BL_Lw_Lf_C	0.43	0.46	0.51	0.42	0.37	0.33	0.41	0.51	0.46	0.43
BL_Lw_Lf_I2	0.30	0.35	0.32	0.35	0.33	0.33	0.37	0.34	0.37	0.31
BL_Lw_Lf_I1	0.32	0.34	0.33	0.30	0.28	0.28	0.29	0.33	0.34	0.32
BL_Lw_Rt_I1	0.33	0.35	0.30	0.29	0.28	0.26	0.29	0.31	0.33	0.31
BL_Lw_Rt_I2	0.33	0.34	0.31	0.34	0.30	0.28	0.33	0.32	0.33	0.34
BL_Lw_Rt_C	0.42	0.45	0.52	0.43	0.38	0.33	0.41	0.49	0.46	0.43
BL_Lw_Rt_P1	0.27	0.39	0.38	0.30	0.31	0.27	0.31	0.37	0.40	0.26
BL_Lw_Rt_P2	0.23	0.31	0.28	0.25	0.18	0.20	0.24	0.29	0.33	0.22

Continue…

Trait	MD_Lw_Lf_P2	MD_Lw_Lf_P1	MD_Lw_Lf_C	MD_Lw_Lf_I2	MD_Lw_Lf_I1	MD_Lw_Rt_I1	MD_Lw_Rt_I2	MD_Lw_Rt_C	MD_Lw_Rt_P1	MD_Lw_Rt_P2
BL_Up_Rt_P2	**3.6E-22**	**8.9E-22**	**2.5E-16**	**2.4E-15**	**1.6E-14**	**1.5E-14**	**5.6E-21**	**1.6E-17**	**6.0E-23**	**2.3E-29**
IC_Lw_Rt_P2	**9.7E-12**	**2.5E-13**	**1.8E-19**	**1.6E-07**	**1.7E-06**	**2.3E-07**	**3.0E-05**	**1.4E-16**	**6.0E-15**	**2.0E-11**
IC_Lw_Rt_P1	**8.4E-14**	**1.7E-22**	**4.1E-19**	**9.1E-10**	**1.9E-07**	**1.8E-09**	**3.3E-08**	**2.3E-17**	**1.6E-20**	**2.3E-10**
IC_Lw_Rt_C	**3.5E-10**	**8.8E-16**	**3.6E-17**	**2.3E-05**	**8.8E-07**	**5.5E-08**	**2.2E-07**	**7.0E-20**	**2.3E-14**	**1.7E-09**
IC_Lw_Rt_I2	**3.0E-08**	**1.5E-12**	**4.3E-10**	**1.4E-05**	**2.7E-06**	**3.0E-06**	**1.6E-08**	**3.8E-10**	**3.0E-11**	**2.4E-08**
IC_Lw_Rt_I1	**2.5E-08**	**8.5E-12**	**3.1E-07**	**3.2E-06**	**2.6E-08**	**1.3E-08**	**1.1E-07**	**9.5E-09**	**6.2E-11**	**6.2E-07**
IC_Lw_Lf_I1	**3.6E-06**	**1.1E-10**	**7.9E-07**	**1.4E-06**	**3.3E-08**	**1.4E-07**	**4.6E-09**	**1.5E-07**	**5.0E-09**	**8.0E-07**
IC_Lw_Lf_I2	**2.9E-09**	**1.3E-13**	**5.9E-10**	**5.0E-07**	**2.5E-07**	**2.0E-07**	**5.0E-09**	**1.6E-12**	**3.5E-12**	**2.0E-08**
IC_Lw_Lf_C	**1.0E-11**	**6.7E-18**	**2.9E-17**	**7.5E-07**	**1.9E-06**	**1.6E-07**	**2.2E-09**	**3.3E-20**	**1.8E-14**	**6.0E-10**
IC_Lw_Lf_P1	**2.3E-14**	**3.8E-23**	**5.3E-15**	**5.1E-08**	2.1E-04	**1.6E-05**	**7.0E-08**	**1.9E-13**	**6.6E-18**	**1.3E-09**
IC_Lw_Lf_P2	**3.7E-09**	**3.3E-10**	**6.1E-17**	**6.1E-07**	**4.4E-06**	**1.7E-05**	3.2E-04	**3.0E-13**	**1.7E-10**	**1.6E-06**
MD_Lw_Lf_P2		**7.0E-73**	**9.9E-42**	**4.7E-32**	**2.3E-37**	**6.8E-35**	**8.3E-34**	**1.0E-37**	**1.0E-68**	**2.0E-141**
MD_Lw_Lf_P1	0.70		**3.3E-56**	**6.8E-33**	**1.3E-30**	**1.4E-30**	**3.6E-34**	**2.1E-47**	**1.5E-121**	**4.1E-66**
MD_Lw_Lf_C	0.53	0.63		**1.0E-60**	**1.9E-50**	**6.1E-49**	**1.6E-43**	**9.0E-117**	**6.5E-51**	**7.7E-38**
MD_Lw_Lf_I2	0.47	0.50	0.61		**1.1E-102**	**9.2E-102**	**3.0E-119**	**6.4E-49**	**2.1E-35**	**2.3E-26**
MD_Lw_Lf_I1	0.51	0.48	0.57	0.75		**1.2E-189**	**1.1E-94**	**2.7E-49**	**9.6E-35**	**2.2E-28**
MD_Lw_Rt_I1	0.49	0.48	0.56	0.74	0.88		**1.7E-94**	**6.1E-49**	**9.4E-35**	**1.0E-28**
MD_Lw_Rt_I2	0.48	0.51	0.53	0.78	0.73	0.73		**4.7E-56**	**6.2E-36**	**3.6E-34**
MD_Lw_Rt_C	0.51	0.58	0.78	0.56	0.56	0.56	0.60		**2.6E-57**	**1.8E-42**
MD_Lw_Rt_P1	0.68	0.82	0.60	0.51	0.51	0.51	0.52	0.63		**1.1E-67**
MD_Lw_Rt_P2	0.83	0.67	0.51	0.43	0.45	0.45	0.49	0.54	0.68	
BL_Lw_Lf_P2	0.46	0.50	0.39	0.35	0.35	0.35	0.37	0.42	0.50	0.43
BL_Lw_Lf_P1	0.43	0.56	0.53	0.43	0.41	0.40	0.37	0.50	0.55	0.44
BL_Lw_Lf_C	0.42	0.47	0.64	0.42	0.41	0.38	0.35	0.61	0.49	0.39
BL_Lw_Lf_I2	0.43	0.46	0.44	0.43	0.41	0.38	0.42	0.46	0.47	0.41
BL_Lw_Lf_I1	0.43	0.41	0.50	0.50	0.49	0.46	0.45	0.48	0.47	0.40
BL_Lw_Rt_I1	0.45	0.45	0.53	0.53	0.52	0.49	0.47	0.50	0.50	0.41
BL_Lw_Rt_I2	0.45	0.47	0.55	0.50	0.46	0.44	0.43	0.52	0.49	0.43
BL_Lw_Rt_C	0.40	0.47	0.62	0.41	0.42	0.40	0.38	0.61	0.49	0.37
BL_Lw_Rt_P1	0.40	0.55	0.46	0.40	0.38	0.39	0.39	0.46	0.56	0.40
BL_Lw_Rt_P2	0.44	0.48	0.37	0.33	0.35	0.35	0.36	0.42	0.49	0.48

Conitnue…

Trait	BL_Lw_Lf_P2	BL_Lw_Lf_P1	BL_Lw_Lf_C	BL_Lw_Lf_I2	BL_Lw_Lf_I1	BL_Lw_Rt_I1	BL_Lw_Rt_I2	BL_Lw_Rt_C	BL_Lw_Rt_P1	BL_Lw_Rt_P2
BL_Up_Rt_P2	**3.0E-49**	**3.3E-43**	**6.0E-32**	**2.0E-23**	**5.8E-18**	**5.2E-19**	**2.4E-20**	**5.4E-28**	**1.3E-43**	**2.2E-60**
IC_Lw_Rt_P2	**2.1E-08**	**1.9E-14**	**1.9E-26**	**3.4E-13**	**7.3E-15**	**4.4E-15**	**1.1E-15**	**6.8E-25**	**1.7E-09**	**3.7E-08**
IC_Lw_Rt_P1	**5.5E-16**	**2.2E-21**	**5.3E-28**	**3.6E-16**	**3.5E-15**	**3.2E-16**	**2.3E-15**	**5.9E-27**	**8.2E-20**	**1.8E-12**
IC_Lw_Rt_C	**6.0E-13**	**4.9E-19**	**7.4E-39**	**2.0E-14**	**1.4E-15**	**2.1E-13**	**6.7E-14**	**6.8E-40**	**6.0E-19**	**2.9E-11**
IC_Lw_Rt_I2	**1.4E-09**	**1.6E-13**	**8.4E-26**	**2.4E-17**	**3.1E-13**	**1.2E-12**	**7.2E-17**	**1.6E-26**	**4.0E-12**	**3.9E-09**
IC_Lw_Rt_I1	**1.6E-06**	**8.7E-13**	**2.0E-19**	**9.5E-16**	**3.5E-11**	**1.3E-11**	**7.2E-13**	**2.1E-20**	**2.3E-12**	**1.3E-05**
IC_Lw_Lf_I1	**9.1E-07**	**8.3E-11**	**4.9E-16**	**9.3E-16**	**1.3E-11**	**3.6E-10**	**1.6E-11**	**3.1E-16**	**7.0E-10**	**2.0E-06**
IC_Lw_Lf_I2	**1.0E-09**	**1.0E-13**	**1.5E-24**	**1.3E-19**	**2.2E-12**	**2.5E-12**	**2.5E-15**	**5.1E-24**	**1.6E-12**	**5.7E-09**
IC_Lw_Lf_C	**1.6E-12**	**4.9E-18**	**3.3E-38**	**6.0E-17**	**1.1E-15**	**4.0E-14**	**3.6E-15**	**4.5E-36**	**1.2E-17**	**6.3E-12**
IC_Lw_Lf_P1	**3.2E-15**	**1.3E-20**	**3.4E-28**	**5.0E-18**	**2.0E-15**	**1.4E-14**	**4.8E-14**	**4.8E-28**	**7.2E-21**	**6.4E-14**
IC_Lw_Lf_P2	**6.9E-08**	**1.7E-14**	**1.4E-26**	**2.8E-13**	**2.6E-14**	**4.7E-14**	**5.4E-16**	**2.6E-26**	**3.9E-09**	**1.7E-07**
MD_Lw_Lf_P2	**1.4E-30**	**5.1E-23**	**5.0E-25**	**1.5E-26**	**3.9E-26**	**1.3E-28**	**3.0E-29**	**5.8E-23**	**1.4E-20**	**5.0E-27**
MD_Lw_Lf_P1	**9.5E-33**	**7.5E-43**	**1.0E-28**	**5.1E-28**	**3.6E-22**	**2.0E-26**	**2.6E-29**	**1.7E-28**	**8.4E-41**	**1.6E-29**
MD_Lw_Lf_C	**1.2E-21**	**1.9E-38**	**4.1E-66**	**2.7E-28**	**9.7E-38**	**2.3E-41**	**9.8E-47**	**8.8E-62**	**1.3E-27**	**2.1E-19**
MD_Lw_Lf_I2	**1.7E-17**	**8.8E-24**	**1.9E-25**	**6.9E-27**	**1.2E-37**	**1.9E-41**	**1.9E-37**	**8.3E-25**	**2.4E-20**	**1.8E-15**
MD_Lw_Lf_I1	**2.9E-17**	**2.0E-22**	**6.3E-24**	**3.0E-24**	**7.1E-36**	**2.5E-40**	**6.5E-31**	**5.5E-26**	**2.1E-18**	**8.4E-18**
MD_Lw_Rt_I1	**1.0E-17**	**1.8E-20**	**9.7E-21**	**2.8E-21**	**1.4E-30**	**4.4E-36**	**1.3E-28**	**4.8E-23**	**1.2E-19**	**7.7E-18**
MD_Lw_Rt_I2	**1.4E-19**	**1.2E-17**	**2.6E-18**	**9.0E-26**	**7.6E-30**	**3.9E-32**	**2.7E-26**	**2.2E-20**	**8.7E-20**	**1.6E-18**
MD_Lw_Rt_C	**4.3E-25**	**5.8E-33**	**1.0E-59**	**8.5E-31**	**1.1E-33**	**2.9E-36**	**3.4E-40**	**8.3E-60**	**6.7E-28**	**1.5E-24**
MD_Lw_Rt_P1	**6.0E-32**	**2.5E-41**	**1.3E-31**	**1.6E-28**	**1.2E-28**	**1.2E-32**	**1.7E-31**	**2.4E-32**	**4.3E-43**	**4.0E-31**
MD_Lw_Rt_P2	**5.0E-26**	**2.9E-24**	**3.2E-21**	**2.1E-23**	**1.8E-22**	**2.4E-23**	**3.4E-26**	**3.7E-19**	**4.1E-20**	**2.7E-33**
BL_Lw_Lf_P2		**1.1E-85**	**2.5E-43**	**9.1E-38**	**7.0E-24**	**1.5E-29**	**2.9E-34**	**6.6E-39**	**6.6E-73**	**3.4E-113**
BL_Lw_Lf_P1	0.74		**5.2E-74**	**2.9E-48**	**2.2E-47**	**1.5E-50**	**1.7E-55**	**1.2E-69**	**1.7E-157**	**1.0E-67**
BL_Lw_Lf_C	0.54	0.69		**3.1E-69**	**2.2E-70**	**2.3E-65**	**1.3E-85**	**9.3E-198**	**2.9E-68**	**1.0E-36**
BL_Lw_Lf_I2	0.51	0.59	0.65		**2.3E-126**	**2.5E-126**	**2.8E-166**	**4.3E-69**	**3.3E-48**	**9.9E-36**
BL_Lw_Lf_I1	0.41	0.59	0.65	0.80		**5.1E-222**	**1.8E-143**	**1.9E-73**	**8.8E-41**	**1.4E-25**
BL_Lw_Rt_I1	0.46	0.60	0.64	0.80	0.91		**2.1E-153**	**2.1E-73**	**4.7E-44**	**5.6E-29**
BL_Lw_Rt_I2	0.49	0.62	0.70	0.86	0.83	0.84		**2.9E-90**	**3.1E-45**	**1.7E-31**
BL_Lw_Rt_C	0.52	0.68	0.89	0.65	0.67	0.67	0.72		**1.1E-58**	**1.1E-32**
BL_Lw_Rt_P1	0.70	0.87	0.67	0.59	0.55	0.57	0.57	0.64		**2.5E-77**
BL_Lw_Rt_P2	0.78	0.68	0.50	0.50	0.43	0.45	0.47	0.48	0.71	

Table 4.15. Simple correlation between 60 dental measurements. Correlation values are presented in the lower left triangle, with corresponding permutation P values in the upper right triangle. Correlations with significant P values (<0.0001, Bonferroni-adjusted threshold) and their corresponding P values are highlighted in bold.

Among the correlations between the dental measurements and the covariates, sex presented weak to moderate negative correlations with most of the traits (r = 0.42, p-value $2.76x10^{-6}$), although some were not significant (Table 4.16a and b).Women have smaller teeth than men [384].

Interestingly, the length of the teeth (inciso-cervical measurements) showed a weak correlation with age, especially canine and premolars (r values = 0.20 - 0.29), although they were not significant. This is expected due to the wear with the use of the teeth over the years [385].

Weak negative correlations were observed between the meso-distal and bucco-lingual distances, and European ancestry.

Trait	Sex	Age	Africa	Europe	Native
IC_Up_Rt_P2	-0.30	0.29	-0.11	0.06	0.03
IC_Up_Rt_P1	-0.37	0.19	-0.05	0.03	0.01
IC_Up_Rt_C	-0.42	0.20	-0.05	0.04	-0.01
IC_Up_Rt_I2	-0.30	0.18	-0.05	0.02	0.02
IC_Up_Rt_I1	-0.33	0.10	-0.04	0.05	-0.02
IC_Up_Lf_I1	-0.35	0.11	-0.04	0.07	-0.04
IC_Up_Lf_I2	-0.24	0.17	-0.05	0.03	0.01
IC_Up_Lf_C	-0.44	0.21	-0.04	0.03	0.00
IC_Up_Lf_P1	-0.36	0.20	-0.08	0.05	0.01
IC_Up_Lf_P2	-0.31	0.27	-0.07	0.03	0.04
MD_Up_Lf_P2	-0.16	0.07	0.02	-0.08	0.09
MD_Up_Lf_P1	-0.31	0.03	0.08	-0.10	0.05
MD_Up_Lf_C	-0.42	0.08	0.08	-0.14	0.11
MD_Up_Lf_I2	-0.18	0.02	0.05	-0.11	0.09
MD_Up_Lf_I1	-0.27	0.06	0.10	-0.06	-0.02
MD_Up_Rt_I1	-0.26	0.05	0.06	-0.06	0.02
MD_Up_Rt_I2	-0.16	0.03	0.09	-0.15	0.10
MD_Up_Rt_C	-0.41	0.09	0.06	-0.10	0.06
MD_Up_Rt_P1	-0.28	0.05	0.10	-0.09	0.02
MD_Up_Rt_P2	-0.22	0.07	0.03	-0.08	0.08
BL_Up_Lf_P2	-0.31	0.00	0.07	-0.10	0.06
BL_Up_Lf_P1	-0.35	-0.01	0.14	-0.16	0.07
BL_Up_Lf_C	-0.43	0.15	0.05	-0.07	0.04
BL_Up_Lf_I2	-0.30	0.07	0.01	-0.05	0.06
BL_Up_Lf_I1	-0.30	0.04	0.02	-0.02	0.01
BL_Up_Rt_I1	-0.31	0.07	0.01	-0.02	0.02
BL_Up_Rt_I2	-0.29	0.06	0.04	-0.05	0.03
BL_Up_Rt_C	-0.45	0.18	0.08	-0.08	0.03
BL_Up_Rt_P1	-0.34	0.01	0.12	-0.13	0.05
BL_Up_Rt_P2	-0.30	0.00	0.11	-0.12	0.05
IC_Lw_Rt_P2	-0.29	0.26	-0.11	0.12	-0.05

Trait	Sex	Age	Africa	Europe	Native
IC_Lw_Rt_P1	-0.38	0.21	-0.05	0.05	-0.02
IC_Lw_Rt_C	-0.39	0.19	-0.05	0.08	-0.05
IC_Lw_Rt_I2	-0.19	0.19	-0.08	0.05	0.02
IC_Lw_Rt_I1	-0.19	0.15	-0.10	0.08	-0.01
IC_Lw_Lf_I1	-0.17	0.14	-0.09	0.08	-0.01
IC_Lw_Lf_I2	-0.23	0.19	-0.09	0.09	-0.03
IC_Lw_Lf_C	-0.36	0.18	-0.05	0.07	-0.04
IC_Lw_Lf_P1	-0.33	0.19	-0.08	0.08	-0.03
IC_Lw_Lf_P2	-0.22	0.22	-0.13	0.12	-0.02
MD_Lw_Lf_P2	-0.23	0.02	0.10	-0.12	0.05
MD_Lw_Lf_P1	-0.26	0.03	0.08	-0.10	0.06
MD_Lw_Lf_C	-0.40	0.10	0.00	-0.06	0.08
MD_Lw_Lf_I2	-0.16	0.09	-0.01	-0.08	0.11
MD_Lw_Lf_I1	-0.16	0.04	0.00	-0.06	0.07
MD_Lw_Rt_I1	-0.14	0.04	0.02	-0.07	0.07
MD_Lw_Rt_I2	-0.21	0.05	0.02	-0.10	0.11
MD_Lw_Rt_C	-0.44	0.11	0.01	-0.07	0.08
MD_Lw_Rt_P1	-0.28	0.05	0.08	-0.08	0.02
MD_Lw_Rt_P2	-0.21	-0.01	0.13	-0.15	0.06
BL_Lw_Lf_P2	-0.28	-0.04	0.07	-0.03	-0.04
BL_Lw_Lf_P1	-0.34	0.05	0.12	-0.10	0.01
BL_Lw_Lf_C	-0.51	0.15	0.04	-0.05	0.02
BL_Lw_Lf_I2	-0.24	0.06	0.07	-0.07	0.02
BL_Lw_Lf_I1	-0.26	0.07	0.02	-0.02	0.01
BL_Lw_Rt_I1	-0.27	0.08	-0.01	0.01	-0.01
BL_Lw_Rt_I2	-0.27	0.13	0.01	-0.02	0.02
BL_Lw_Rt_C	-0.54	0.15	-0.01	-0.01	0.02
BL_Lw_Rt_P1	-0.34	0.01	0.11	-0.10	0.03
BL_Lw_Rt_P2	-0.22	-0.03	0.07	-0.02	-0.05

Table 4.16a. Correlation between quantitative dental traits and age, sex, and ancestry. Correlation values are presented in Table a), with corresponding *P* values in Table b). Correlations with significant *P* values (<0.001, Bonferroni-adjusted threshold), are highlighted in bold. Anc.= Continental ancestry estimated from the genetic data. Sex coded as female=1, male=0.

p-value	Sex	Age	Africa	Europe	Native
IC_Up_Rt_P2	**2.73E-12**	1.22E-02	1.49E-01	4.98E-01	8.40E-01
IC_Up_Rt_P1	**1.90E-05**	3.01E-01	4.88E-01	9.00E-01	9.76E-01
IC_Up_Rt_C	**2.76E-06**	2.67E-01	3.08E-01	8.33E-01	8.19E-01
IC_Up_Rt_I2	**2.67E-05**	2.96E-01	6.26E-01	6.61E-01	9.47E-01
IC_Up_Rt_I1	2.19E-02	3.47E-01	2.68E-01	6.12E-01	3.39E-01
IC_Up_Lf_I1	8.05E-03	3.13E-01	1.27E-01	3.21E-01	2.32E-01
IC_Up_Lf_I2	**3.56E-05**	2.80E-01	5.18E-01	7.97E-01	5.94E-01
IC_Up_Lf_C	**7.08E-07**	3.65E-01	5.28E-01	9.22E-01	2.65E-01
IC_Up_Lf_P1	**1.61E-05**	1.05E-01	2.65E-01	8.71E-01	5.21E-01
IC_Up_Lf_P2	**1.61E-10**	1.01E-01	5.36E-01	3.71E-01	8.93E-01
MD_Up_Lf_P2	8.50E-02	6.03E-01	5.43E-02	4.03E-02	1.84E-03
MD_Up_Lf_P1	4.71E-01	8.55E-02	4.01E-02	3.22E-01	4.30E-03
MD_Up_Lf_C	6.18E-02	4.91E-02	7.44E-04	1.37E-02	4.98E-02
MD_Up_Lf_I2	5.72E-01	2.09E-01	1.09E-02	3.59E-02	**4.83E-09**
MD_Up_Lf_I1	1.80E-01	1.81E-02	1.59E-01	5.66E-01	2.46E-04
MD_Up_Rt_I1	2.73E-01	1.56E-01	1.52E-01	6.60E-01	**2.28E-08**
MD_Up_Rt_I2	4.30E-01	3.78E-02	6.31E-04	1.69E-02	**1.70E-09**
MD_Up_Rt_C	4.18E-02	1.37E-01	2.61E-02	1.51E-01	3.41E-02
MD_Up_Rt_P1	2.81E-01	3.82E-02	4.83E-02	5.95E-01	3.51E-04
MD_Up_Rt_P2	9.91E-02	5.54E-01	8.16E-02	8.46E-02	1.46E-03
BL_Up_Lf_P2	9.35E-01	9.09E-02	1.88E-02	1.67E-01	1.71E-02
BL_Up_Lf_P1	9.08E-01	2.99E-03	5.56E-04	1.13E-01	1.05E-02
BL_Up_Lf_C	2.85E-04	2.14E-01	1.11E-01	3.97E-01	1.67E-02
BL_Up_Lf_I2	9.63E-02	8.53E-01	2.51E-01	1.78E-01	5.24E-02
BL_Up_Lf_I1	3.33E-01	5.85E-01	5.76E-01	8.58E-01	1.66E-01
BL_Up_Rt_I1	1.02E-01	7.75E-01	6.11E-01	7.02E-01	9.24E-02
BL_Up_Rt_I2	1.93E-01	3.87E-01	2.34E-01	4.94E-01	4.45E-02
BL_Up_Rt_C	**2.00E-05**	7.91E-02	5.94E-02	4.84E-01	3.43E-01
BL_Up_Rt_P1	8.60E-01	1.17E-02	7.38E-03	3.07E-01	2.40E-03
BL_Up_Rt_P2	9.18E-01	1.02E-02	5.15E-03	2.85E-01	4.27E-02
IC_Lw_Rt_P2	**6.15E-10**	1.27E-02	4.62E-03	2.31E-01	2.98E-01
IC_Lw_Rt_P1	**1.09E-06**	3.08E-01	2.79E-01	6.66E-01	2.99E-01
IC_Lw_Rt_C	**5.33E-06**	2.75E-01	8.03E-02	2.25E-01	4.84E-01
IC_Lw_Rt_I2	**3.26E-06**	6.69E-02	2.55E-01	7.03E-01	8.31E-01
IC_Lw_Rt_I1	3.13E-04	2.46E-02	5.00E-02	7.75E-01	5.82E-01
IC_Lw_Lf_I1	7.48E-04	2.90E-02	6.43E-02	8.38E-01	7.02E-01
IC_Lw_Lf_I2	**3.53E-06**	3.16E-02	2.83E-02	4.79E-01	7.94E-01
IC_Lw_Lf_C	**1.74E-05**	2.23E-01	1.23E-01	4.07E-01	7.54E-01
IC_Lw_Lf_P1	**1.26E-05**	8.94E-02	7.86E-02	5.14E-01	1.82E-01
IC_Lw_Lf_P2	**9.03E-08**	2.95E-03	7.63E-03	6.32E-01	7.25E-01
MD_Lw_Lf_P2	6.76E-01	1.60E-02	5.31E-03	2.21E-01	1.24E-02
MD_Lw_Lf_P1	5.32E-01	7.89E-02	2.18E-02	1.91E-01	2.03E-01
MD_Lw_Lf_C	1.20E-02	9.52E-01	1.55E-01	6.45E-02	1.50E-03

Continue…

p-value	Sex	Age	Africa	Europe	Native
MD_Lw_Lf_I2	3.69E-02	9.07E-01	7.66E-02	1.22E-02	**4.67E-10**
MD_Lw_Lf_I1	3.05E-01	9.66E-01	1.94E-01	8.92E-02	**1.86E-07**
MD_Lw_Rt_I1	3.26E-01	6.60E-01	1.02E-01	8.25E-02	**1.04E-07**
MD_Lw_Rt_I2	2.51E-01	6.01E-01	1.85E-02	8.84E-03	**3.74E-09**
MD_Lw_Rt_C	9.28E-03	7.77E-01	1.00E-01	5.63E-02	2.24E-02
MD_Lw_Rt_P1	2.71E-01	6.80E-02	9.54E-02	6.81E-01	2.42E-01
MD_Lw_Rt_P2	7.72E-01	2.64E-03	7.35E-04	1.67E-01	3.77E-02
BL_Lw_Lf_P2	3.18E-01	8.89E-02	4.74E-01	4.08E-01	5.68E-01
BL_Lw_Lf_P1	2.36E-01	6.31E-03	2.68E-02	8.17E-01	4.09E-01
BL_Lw_Lf_C	2.22E-04	3.80E-01	2.70E-01	5.72E-01	6.74E-01
BL_Lw_Lf_I2	1.32E-01	1.02E-01	1.18E-01	6.91E-01	7.36E-02
BL_Lw_Lf_I1	1.11E-01	6.74E-01	6.31E-01	8.36E-01	1.31E-01
BL_Lw_Rt_I1	7.39E-02	8.56E-01	7.39E-01	7.97E-01	8.94E-02
BL_Lw_Rt_I2	2.49E-03	8.59E-01	6.04E-01	6.10E-01	3.90E-01
BL_Lw_Rt_C	4.83E-04	8.37E-01	8.32E-01	6.24E-01	5.98E-01
BL_Lw_Rt_P1	7.52E-01	1.53E-02	2.35E-02	5.44E-01	3.18E-01
BL_Lw_Rt_P2	5.33E-01	8.49E-02	6.14E-01	2.50E-01	6.34E-01

Table 4.16b. Corresponding *P* values correlation between face traits and age, sex, and ancestry.

4.3.3.3 Genome Wide Association Analyses of dental measurements in the Colombian CANDELA sample

I performed a genome-wide association test using multivariate linear regression, as implemented in PLINK [251], using an additive genetic model adjusting for: age, sex and the first five principal components computed from the SNP data. The resulting statistics showed no evidence of residual population stratification for any of the traits (Figure 4.58). After quality control assessment, 509 samples and 640,094 SNPs were retained for the not imputed GWAS and 9,521,215 variants were retained for the imputed GWAS.

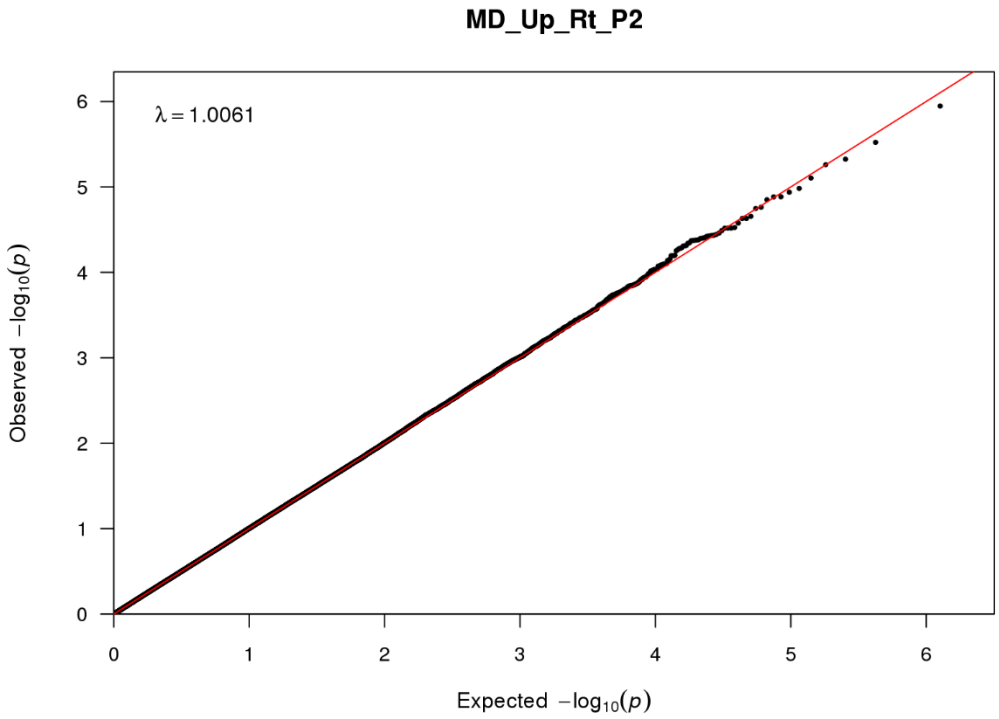


Figure 4.58. Q-Q plot for the GWAS of Mesio-distance of the right upper second premolar (Md_Up_Rt_P2). The remaining Q-Q Plots are shown in Appendix B Figures B.116-B.176. This plot does not show sign of inflation between the expected and observed *P*-values. All the traits show similar pattern, the genomic control factor lambda <1.02, demonstrating there is no population stratification.

4.3.3.3.1 Genomic regions associated to dental measurements

Seven dental measurements examined showed genome-wide significant association (*P* values < 5×10^{-8}) with SNPs in four genomic regions (Table 4.17 and Figure 4.59).

Chromosomal Region	Index SNP	Alleles	Effect size	P value	Associated Trait	Gene
1q42.2	rs16856377	A<G	-3.30E-01	4.5E-08	Bucco-Lingual Lw Lf Canine	***DISC2***
1q42.2	rs2883720	C>T	3.04E-01	8.7E-09	Incisal Cervical Lw Rt Premolar 1	***EXOC8***
2q12.3	rs3827760	A>G	2.08E-01	8.4E-09	Meso-Distal Lw Lf Lateral Incisor	*EDAR*
2q12.3	rs3827760	A>G	2.54E-01	1.2E-08	Meso-Distal Up Lf Lateral Incisor	*EDAR*
2q12.3	rs3827760	A>G	2.36E-01	9.7E-09	Meso-Distal Up Rt Central Incisor	*EDAR*
2q12.3	rs3827760	A>G	2.52E-01	6.5E-09	Meso-Distal Up Rt Lateral Incisor	*EDAR*
11q25	rs10894347	C>T	-2.54E-01	4.5E-08	Meso-Distal Lw Lf Premolar 1	***NTM***

Table 4.17. Properties of index SNPs in chromosomal regions showing genome-wide significant association with quantitative dental traits.

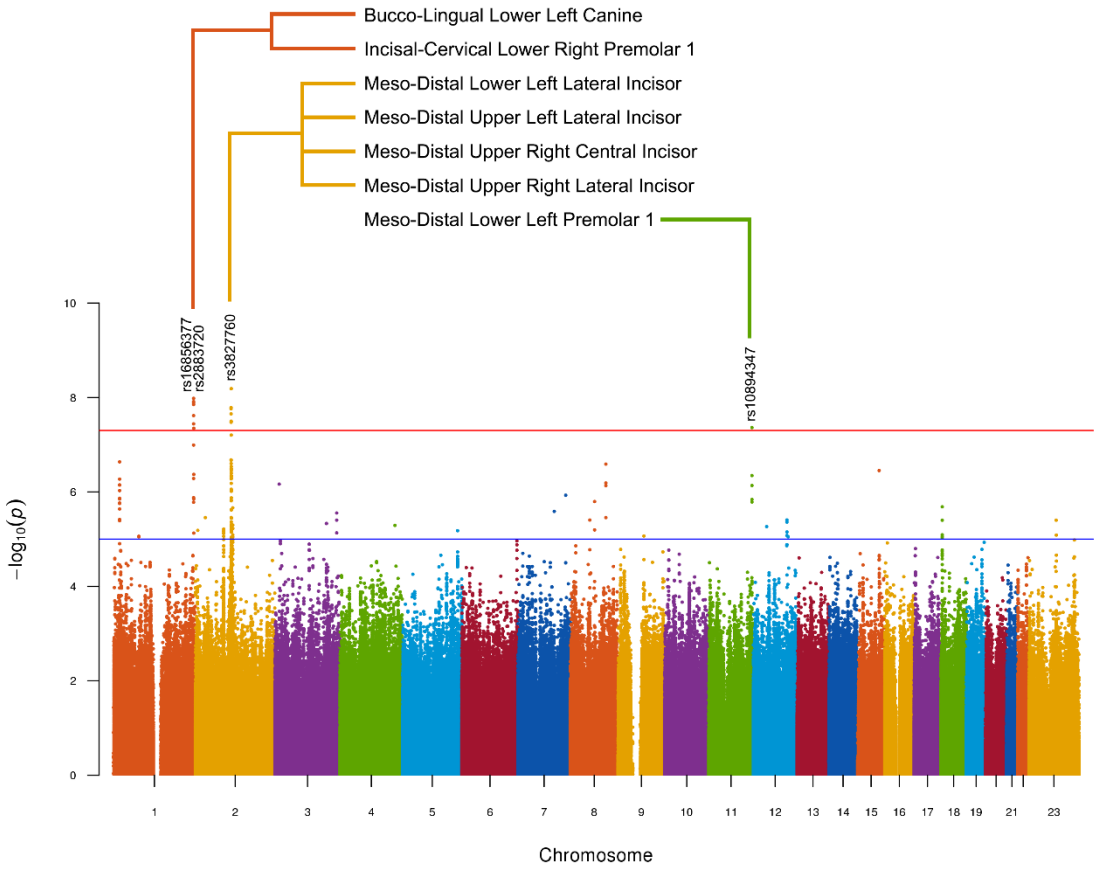


Figure 4.59. Summary of GWAS results for measurements of teeth in Colombian samples from CANDELA. These GWAS was performed with 3 different type of measurements of 20 teeth (8 incisors, 4 canines and 8 premolars) from 564 volunteers. A 'composite' of Manhattan plot shows the results across traits. All the SNPs with P values exceeding thresholds genome-wide suggestive (10^{-5}) are over the blue line and P values reaching the threshold genome-wide significant ($5x10^{-8}$) are above the red line.

Genetic variants in the chromosome region 1q42.2 (Table 4.17, Figure 4.59 and Figure 4.60) showed significant associations with the Bucco-Lingual distance of the left lower canine. The index SNP is rs16856377 and is located ~ 200Mb upstream of the *DISC2* gene.

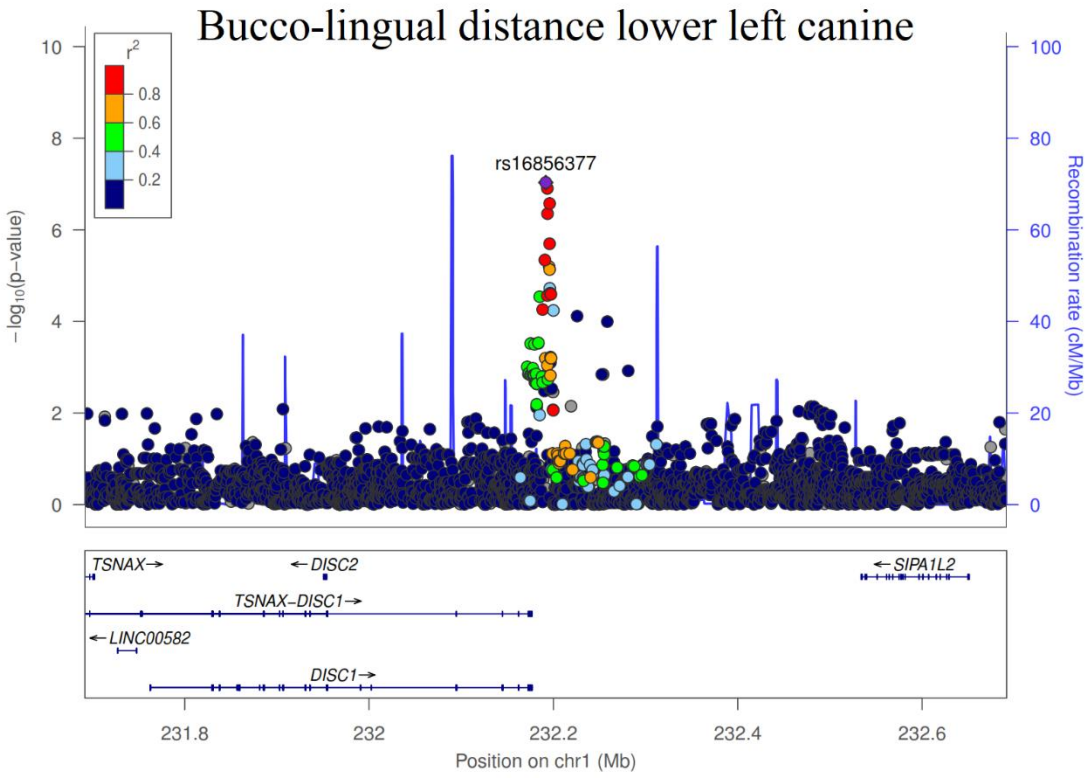


Figure 4.60. Regional association plot for SNPs in 1q42.2 and BL_Lw_Lf_C. This plot shows an area of 500 Mb around the marker (rs16856377). This plot was produced with Locus Zoom [310].

The Inciso-Cervical distance of the right lower first premolar showed a significant association with the SNP rs2883720. This marker falls in an intergenic area in the chromosome region 1q42.2 (Table 4.17, Figure 4.59 and Figure 4.61). It is surrounded by several genes, such as *TRIM67, TTC13, EXOC8, ARV1*, etc.

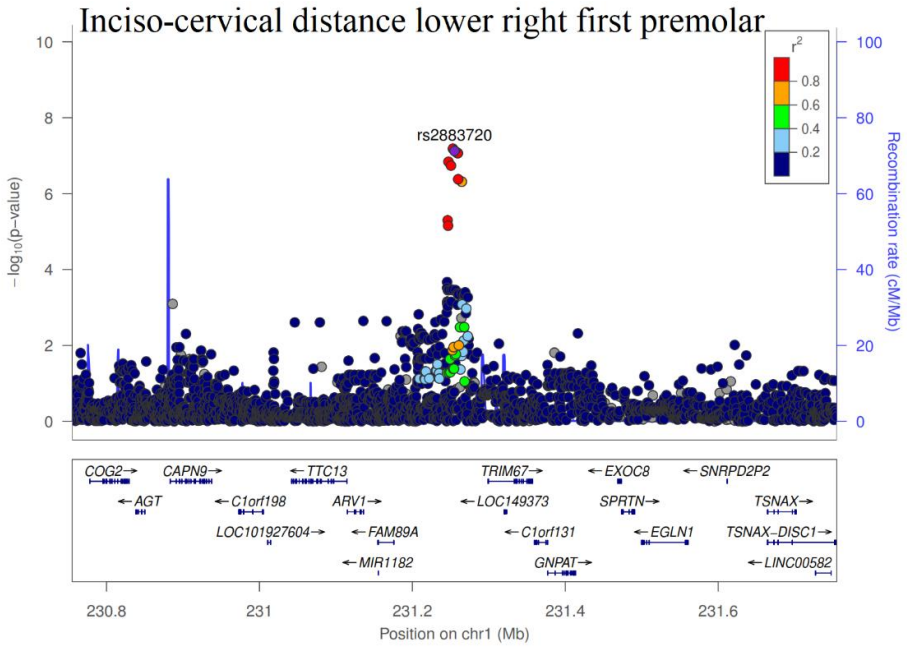


Figure 4.61. Regional association plot for SNPs in 1q42.2 and IC_Lw_Rt_P1. This plot shows an area of 500 Mb around the marker (rs2883720). This plot was produced with Locus Zoom [310].

Four meso-distal measurements from the incisors were associated with the genomic variant rs3827760 (*EDAR* gene), Meso-Distal distance in the lower left lateral incisor, Meso-Distal distance in the upper left lateral incisor, Meso-Distal distance in the upper right central incisor and Meso-Distal distance in the upper right lateral incisor (Table 4.17 and Figure 4.59, Figure 4.62 - Figure 4.65).

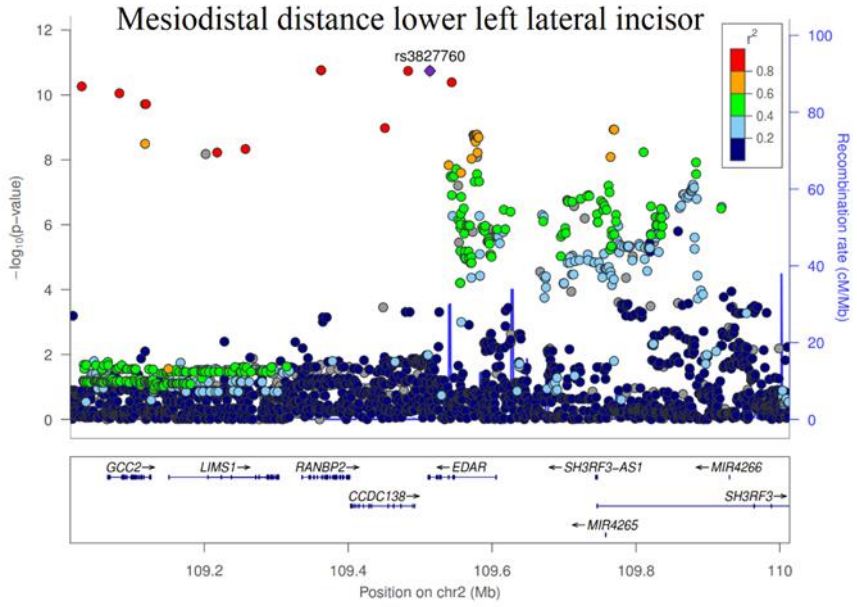


Figure 4.62. Regional association plot for SNPs in 2q12.3 and MD_Lw_Lf_I2. This plot shows an area of 500 Mb around the marker (rs3827760). This plot was produced with Locus Zoom [310].

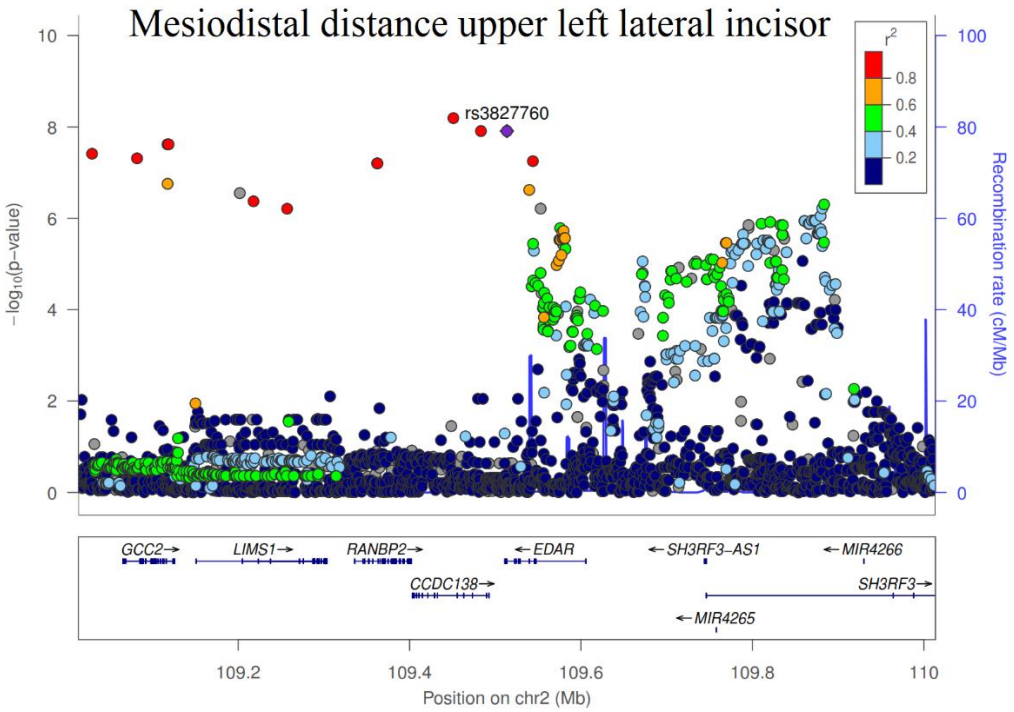


Figure 4.63. Regional association plot for SNPs in 2q12.3 and MD_Up_Lf_I2. This plot shows an area of 500 Mb around the marker (rs3827760). This plot was produced with Locus Zoom [310].

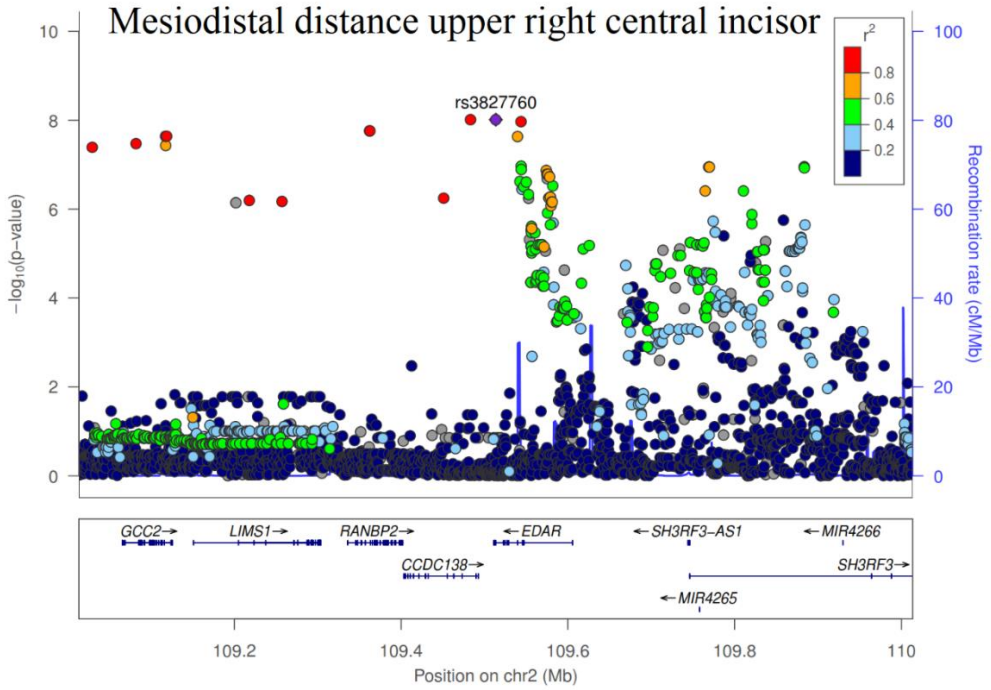


Figure 4.64. Regional association plot for SNPs in 2q12.3 and MD_Up_Rt_I1. This plot shows an area of 500 Mb around the marker (rs3827760). This plot was produced with Locus Zoom [310].

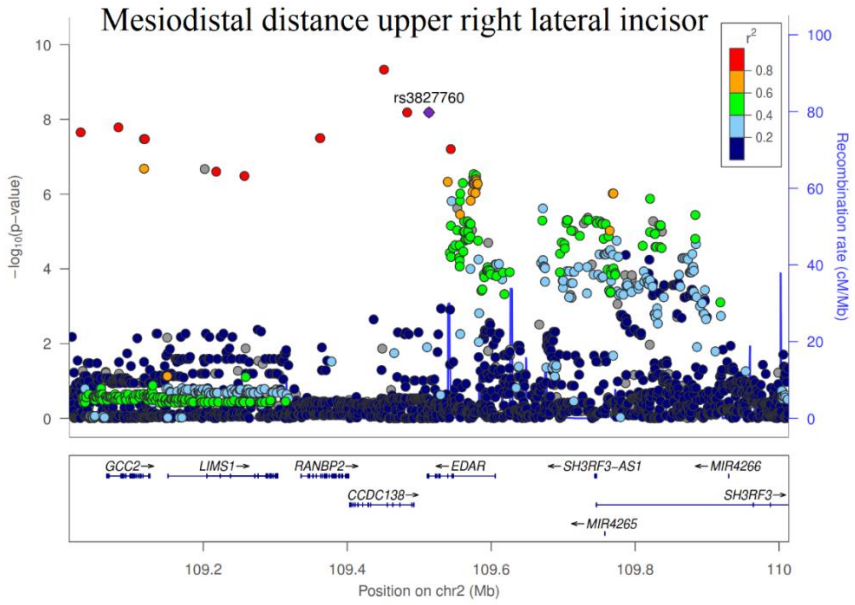


Figure 4.65. Regional association plot for SNPs in 2q12.3 and MD_Up_Rt_I2. This plot shows an area of 500 Mb around the marker (rs3827760). This plot was produced with Locus Zoom [310].

The Meso-Distal distance of the left lower first premolar showed association with a SNP (rs10894347) located in chromosome region 11q25 (Table 4.17, Figure 4.59 and Figure 4.66). This marker is between the *SNX19* and *NTM* genes.

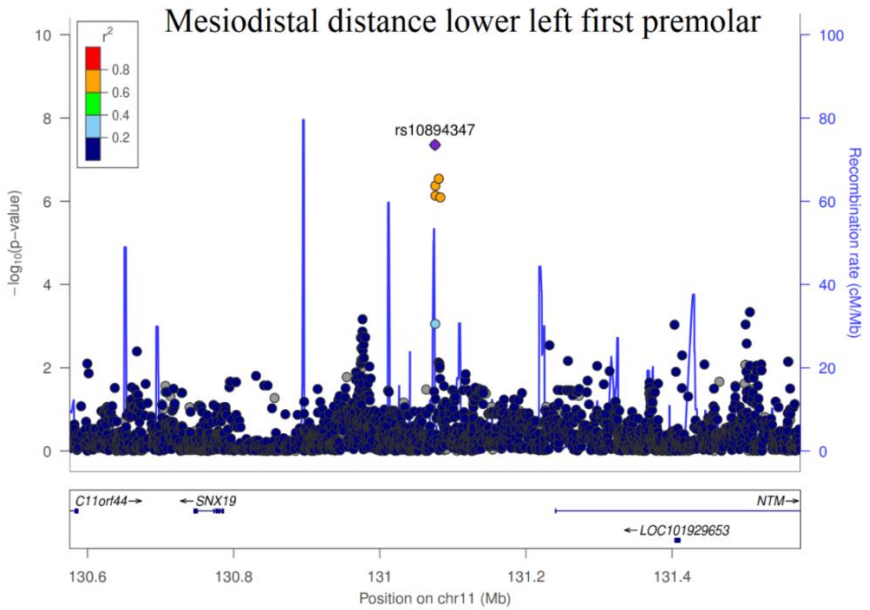


Figure 4.66. Regional association plot for SNPs in 2q12.3 and MD_Lw_Lt_P1. This plot shows an area of 500 Mb around the marker (rs10894347). This plot was produced with Locus Zoom [310].

4.3.4 Local Ancestry inference

A large proportion of SNPs in both genome-wide scans showed a high frequency in African populations (Appendix B, Table B.3 and B.4), and therefore rare in the Colombian sample, whose average African ancestry is low at around 9% (Figure 3.5). In order to check if the SNPs associated with some of the traits belong to African haplotypes in the volunteers carrying these markers, a local ancestry inference analysis was performed. From a previous local inference analysis performed with RFMix software [366] some of the significant SNPs that were present in this analysis were checked (rs3827760, rs1037804 and rs9899063). Not all the SNPs were checked because RFMix is very slow, one analysis could take weeks, therefore I just checked the SNPs that were available from this previous analysis (Table 4.18).

SNPs were marked as African-only if it was very rare in the other two continental populations (<2% in both Europeans and Native Americans) but not so rare in Africans (>2%). A SNP was rare in the Colombian sample studied if its allele frequency was <5%. These frequencies were calculated from the same set of samples used in the RFMix analysis.

For these SNPs, the minor allele in the Colombian population was also the African-specific allele. It was therefore checked in the RFMix results whether the haplotypes of the Colombian individuals carrying this allele is indeed African in origin. Otherwise, this could reflect an error in genotyping or other issues with the data.

To verify this for such a SNP, each haplotype in the Colombian data which carried the minor allele was filtered from the RFMix-produced haplotype and local ancestry files. The proportion of local ancestry assigned to Africa at this locus among these haplotypes (that carry the minor allele) was calculated. If this number is close to 100%, it implies nearly all such haplotypes around that locus indeed come from an African ancestor, and therefore carrying this rare Africa-specific allele is probably real instead of being a genotyping artefact. This in turn provides indirect validation that the observed association with this rare allele is probably real.

As a sanity check, the same process was applied to the African samples in this dataset, to see if for all such SNPs the proportion of assigned African ancestry was 100%, which indeed it was. This validates the process in general, but also lends some

reliability to each SNP as any SNP having low African ancestry proportion in this step would indicate errors in genotyping or the quality of data/model in that region.

In the end, for all such autosomal SNPs (except for those on chromosome 6), a table was prepared listing the allele frequencies in the Colombian and reference populations (Appendix B, Table B.3), as well as the assigned African ancestry proportion for the minor allele for the significant alleles that were present in the RFMix analysis . Out of three alleles checked, two of them were assigned to an African haplotype. The third one was rs3827760 (Table 4.18), this allele was used as a quality control because is well known to be present in East-Asian and Native-American populations.

		Minor allele frequency				Proportion African local ancestry from RFMix in Colombian CANDELA sample
Chr	SNP	Colombian	African	European	Native American	
2	rs3827760	0.25	0.00	0.00	0.96	0.00
10	rs1037804	0.02	0.22	0.00	0.00	0.89
17	rs9899063	0.02	0.27	0.01	0.00	1.00

Table 4.18. Results of local ancestry inference using RFMix in the Colombian CANDELA sample. The 3 SNPs that were previously analysed in RFMix [366]and were associated to some ordinal or quantitative dental trait.

4.4 Discussion

Genome-wide scans of ordinal and quantitative phenotypes were performed in a Colombian population, a subset of the CANDELA Cohort [59]. Genome-wide studies with the 46 traits scored using the ASUDAS method, based on the level of expression of the studied features. And then genome-wide scans were performed with three kind of different measurements (inciso-cervical, mesiodistal and bucco-lingual distances measured on incisors, canines and premolars). Both analyses, ordinal and quantitative phenotypes, showed new and previously reported genomic associations.

4.4.1 Candidate genes associated to ordinal traits

Twenty-three ordinal traits showed association with sixteen genomic variants (Table 4.13 and Figure 4.34). There were ten possible candidate genes, that have already been associated with some dental trait, in both development or disease of the teeth, whether by GWAS or other types of analyses (experimental analyses, bioinformatic prediction, etc.).

Cusp 7 LM1

The accessory cusp 7 of the first lower molar (C7LM1) regrouped (from 6 to 3 categories) showed association with the genomic variant, rs3851907, located in 1q43 chromosome region. It falls within the gene *ACTN2* (Actinin, alpha 2) (Table 4.13, Figure 4.34 and Figure 4.35). Interestingly, a SNP in this gene was found to be associated to dental caries in a GWAS in a European population [386]. The frequency of the associated allele is higher in African populations (Appendix Table B.3) and the trait C7LM1 is also more frequent in African populations as can be seen in Table 4.19.

Trait frequency	**WE**	**NE**	**NA**	**WA**	**SA**	**KH**	**CM**	**JO**	**RJ**	**NES**	**SS**	**AA**	**NWA**	**NSAI**
Cusp 7 LM1	0.05	0.05	0.09	0.44	0.27	0.26	0.08	0.03	0.06	0.06	0.10	0.09	0.07	0.09
Double Shoveling	0.04	0.05	0.09	0.03	0.02	0.00	0.29	0.01	0.20	0.33	0.15	0.35	0.57	0.71
Deflecting Wrinkle	0.05	0.16	0.08	0.17	0.18	0.17	0.16	0.05	0.15	0.40	0.17	0.30	0.37	0.38
Shoveling UI1	0.03	0.02	0.08	0.07	0.09	0.13	0.72	0.26	0.66	0.62	0.37	0.69	0.83	0.92

Table 4.19. Frequency of some associated traits in world-wide populations. WE: Western Europe, NE:Northern-Europe, NA:North-Africa, WA:West-Africa, SA: South-Africa, KH:Khoisan, CM:China-Mongolia, JO:Jomon, RJ:Recent-Japan, NES:Northeast-Siberia, SS: South Siberia, AA: American Arctic, NWA: Northwest North America, NSAI: N.&S. American Indian.[111].

Double shoveling LI1

Double shoveling in the central lower incisor (DSLI1) showed association with 3 genetic regions, the index SNPs were rs1037804, rs16984020 and rs17862881. The first one is in the chromosomic region 10q26.2 (Table 4.13, Figure 4.34 and Figure 4.36), corresponding to an intron on the gene *ADAM12*. Based on bioinformatic prediction, this marker falls in a promoter region [387]. Some authors using mouse models have demonstrated that the *ADAM12* gene, through a developmental pathway, is involved in regulating adipogenesis and myogenesis [388]. The second SNP (rs16984020) associated with DSLI1 is in chromosome 22 (Table 4.13, Figure 4.34 and Figure 4.37) and it does not fall in a specific gene. It is around 200 Mb downstream from *LOC1929539*, this gene is expressed in pharyngeal arches in murine embryos [389], and it is known that teeth are derived from the ectoderm of the first pharyngeal arch [390]. Finally, the third marker that presented significant association with DSLI1 is rs17862881. This SNP is also located in an intergenic region (2q37.1) (Table 4.13,

Figure 4.34 and Figure 4.38), but it is near several genes, which are part from a complex locus that encode different UDP-glucuronosyl transferases that work with different substrates. Another gene located close to this marker is *TRPM8* (~300 Mb). This gene belongs to the family of receptors called Transient receptor potential channels (TRP), which are involved with the sensory neurons innervating the dental pulp [391].

Double shoveling on the lateral lower incisor (DSLI2) showed the same associations with SNPs rs17862881 and rs1037804 (Table 4.13, Figure 4.34, Figure 4.39 and Figure 4.40). This was expected since this is the same trait assessed in another tooth.

Deflecting Wrinkle LM1

Deflecting wrinkle of the first lower molar (DWLM1) showed association with SNPs in chromosome region 12p11.21 (Table 4.13, Figure 4.34 and Figure 4.41). The index marker was rs16919218. This genetic variant falls in an intergenic region near the *KIAA1551* and *BICD1* genes. The first one is a protein coding gene that encodes a minor histocompatibility antigen and a tumor suppressor. The *BICD1* gene is involved in the apical localization in Drosophila blastoderm embryos [392]. This SNP has been associated with an increase in left ventricular mass in a genome-wide scan performed in a group of Caribbean-Hispanics [393]. Interestingly, the frequency of this trait in Africans is higher than in Europeans, although it is lower than in Asian populations (Table 4.19). This could be reflecting the genetic architecture of the trait because the MAF of this SNP is higher in Africans than in other population in the Colombian sample, maybe there is another marker affecting the trait in other populations.

Hypocone LM1

Hypocone in the first upper molar (HypLM1) showed a significant association with two genetic markers, rs9808165 and rs17142577 (Table 4.13, Figure 4.34, Figure 4.42 and 4.43). Contrary to the rest of the associations, the effect of both alleles is negative, this means that people carrying these alleles should have severely reduced forms or absence of this feature (Table 4.13). The first SNP is in chromosome region 2p12 (Figure 4.42). It falls in a genomic region containing only uncharacterized genomic elements, *LOC101927907* and 200 Mb upstream is *LRRTM4*. Interestingly a SNP located in the *LOC101927907* gene has been previously associated with missing teeth in a GWAS performed in a European population [386]. The genomic variant rs17142577

is the second SNP associated with HypLM1, and it is in the 7q31.31 region (Figure 4.43). It does not fall in a gene but is close to *KCND2.* This gene encodes the potassium voltage-gated channel D member 2 protein, and it is involved in regulating the heart rate. Bioinformatic tools have predicted that this gene belongs to a gene network that encodes channel interacting-proteins [394], where *KCND2* may be involved with *KLK4* [395], which is a gene that plays an important role in enamel formation [396].

Metacone UM2

The trait Metacone in the second upper molar (MetUM2) regrouped (from 7 to 4 categories) showed association with a SNP (rs17072636) situated in the 18q22.1 region and it falls in an intron in the uncharacterized gene *LOC284294 (*Table 4.13 and Figure 4.44).

Protostylid LM1

Protostylid in the first lower molar (PrtstLM1), correspond to the appearance of a secondary groove associated to the buccal groove separating cusps 1 and 3 on the buccal surface of the lower molars. This feature has shown association with SNPs in three genomic regions (1p13.1, 8q24.22 and 10q26.13) (Table 4.13 and Figure 4.34). The first genomic variant (rs10494193) is located in *CD101* gene, is also called *IGSF2* (Figure 4.45). It is involved in T-cells receptor/CD3-mediated T-cell activation [397].The second marker that showed association with this trait was rs12680196, it falls in an intergenic region (8q24.22) and it is 400Mb downstream from the gene *EFR3A* (Figure 4.46). SNPs on this gene have shown a suggestive association to dental cavities in the deciduous dentition in a group of children of European ancestry [398]. Finally, PrtostLM1 presented an association with a SNP (rs12570261), also located in an intergenic region in chromosome 10, band 26.13 (Figure 4.47). It is close to the gene Fibroblast Growth Factor Receptor 2 (*FGFR2)* (~50Mb downstream). The *FGFR2* gene belongs to a family of tyrosine kinase receptors genes. It has been implicated in the Crouzon syndrome [399], the main characteristic of this genetic disorder is craniosynostosis, but individuals with this syndrome also present underdeveloped jaw and dental problems [400].

Protostylid LM2

Four genomic variants (rs2241729, rs12299956, rs630603 and rs9899063) located in four different chromosomes were found associated with Protostylid in the second lower molar (PrtstLM2) (Table 4.13 and Figure 4.34). The first marker rs2241729 (7q32.3) (Figure 4.48) is situated in an intron in the gene *PLXNA4.* Plexin A4 belongs to the Plexin gene.family and encodes the receptor of Sema6A protein. They control lamina-specific neuronal stratification in the mouse retina [401]. Interestingly, another transmembrane semaphorin Sema3A has been proposed to participate in the innervation and development of the tooth [402]. The second genetic variant rs12299956, falls in the *GPR133* gene (Figure 4.49). The function of this gene in humans is to increase cellular cAMP formation via adenylate cyclase [403]. It has been proposed that the *GPR133* gene may participate in tooth loss, as cAMP signaling may be involved in periodontal homeostasis and immune response [386]. Marker rs630603, situated in the 15q23 chromosome region (Figure 4.50) is the third SNP associated to Protostylid in the second lower molar. This genomic variant falls in the *TLE3* gene. This gene is a member of the *Notch* signalling pathway, which controls important cellular interactions, but it has not been associated with dental traits before. Finally, SNP rs9899063 which falls in the *NF1* gene showed an association with PrtstLM2. The *NF1* gene is in chromosome region 17q11.2 (Figure 4.51). This gene encodes a protein called neurofibromin and it is produced in several types of cells (nerve cells, oligodendrocytes and Schwann cells) [404]. It is well known that different mutations in this gene cause Neurofibromatosis Type 1 (OMIM#162200) [405], [406].

Shovel shape UI1, UI1 regrouped, UI2, UI2 regrouped, LI1 and LI1 regrouped

The genomic variant rs3827760, which is a missense mutation situated in the *EDAR* gene (2q12.3) (Table 4.13, Figure 4.34, and Figure 4.52 - Figure 4.57), showed association with shovel shape in central and lateral upper incisors and in central lower incisors (SSUI1,SSUI1 regrouped, SSUI2, SSUI2 regrouped, SLI1 and SLI1 regrouped), both the normally categorized version (0-6) and regrouped version (0-3) of the phenotypes. This association has been previously found and described by several authors [284,407-409].

Certainly, the EDAR V370A missense mutation, which is a gold standard marker of selective sweep in an East-Asian population. This marker demonstrates how a

beneficial allele is selected and then brought to other geographic regions, in this case America. Also shows how the genetic architecture is shaped across different populations. This recently introduced allele might have greater size effect because it was saved from the filtering of evolution and hence, is more likely to detect the associations with different phenotypes [249]. The *EDAR* gene is known to have several pleiotropic effects, sweat gland density [410], hair shape [211], chin protrusion [287], among other. All these features can be observed in Latin-American populations. Therefore, this association between shovel shape traits and rs3827760, was expected. This marker was also found associated with chin protrusion in the quantitative facial GWAS described in Section 3.3.8.

4.4.2 Candidate genes associated to dental measurements

The GWAS with dental measurements detected seven associations with four genomic regions (Table 4.17).

Bucco-lingual distance Lower Left Canine

Genetic variants in the chromosome region 1q42.2 (Table 4.17, Figure 4.59 and Figure 4.60) showed significant associations with the Bucco-Lingual distance of the left lower canine. The index SNP is rs16856377 and is located ~ 200Mb upstream of the *DISC2* gene. Mutations in this gene has been associated to schizophrenia [411]. The gene *SIPA1L2* is ~300Mb upstream, this gene encodes a protein called Signal-induced proliferation-associated 1-like protein 2 [412].

Inciso-cervical distance right lower first premolar

The Inciso-Cervical distance of the right lower first premolar showed a significant association with the SNP rs2883720. This marker falls in an intergenic area in the chromosome region 1q42.2 (Table 4.17, Figure 4.59 and Figure 4.61). It is surrounded by several genes, such as *TRIM67, TTC13, EXOC8, ARV1*, etc. The first two have not been characterized. Mutations in *EXOC8* have been related to Joubert syndrome [413] and mutations in *ARV1* to Epileptic encephalopathy. None of them has been associated to tooth traits before.

Mesiodistal distance lower left lateral incisor, upper left lateral incisor, upper right central incisor and upper right lateral incisor

Four meso-distal measurements from the incisors were associated with the genomic variant rs3827760 (*EDAR* gene), Meso-Distal distance in the lower left lateral incisor, Meso-Distal distance in the upper left lateral incisor, Meso-Distal distance in the upper right central incisor and Meso-Distal distance in the upper right lateral incisor (Table 4.17, Figure 4.59 and Figure 4.62 - Figure 4.65). As mentioned before, this SNP has been previously associated with shovel shape incisors [284,408,409]. All these distances correspond to the width of the teeth, the presence of marginal ridges affect the width of the tooth. This is also reflected in the correlations between the quantitative and ordinal traits, where shovel shape traits showed weak positive correlations with meso-distal measurements in the incisors (Appendices Table B.2a and B.2b).

The genome-wide scans performed with different dental measurements as phenotypes, also showed some associations, but none of them were previously associated with dental phenotypes, except for rs3827760, which was associated to mesiodistal distances in upper and lower incisors. The marginal ridges in the lingual surface of incisors define the presence or absence of the shovel shape, not just at the deepest point in the middle of the lingual fossa [414] and these ridges affect the maximum distance from the mesial to the distal surface of the teeth. The allele effect size (~$2.54E^{-01}$) is positive, thus, individuals carrying this allele will have wider incisors (Table 4.17).

4.4.3 Limitations of the sample

Local ancestry assignment

Despite the possible candidate genes, an issue to be noted is that many associated SNPs showed a higher frequency in African populations (Appendix B, Table B.3 and B.4). Nevertheless, the Colombian sample used to perform the GWAS presented a low African ancestry (~9%). In order to assess if this SNPs were Africans, from a previous analysis I checked if the haplotypes where these SNPs belong in the individuals carrying them were Africans (local ancestry assignment), to double check and discard genotyping errors, etc. The local ancestry assignment is more robust than single-SNP based decisions, as it aggregates haplotype information across a broad region for many individuals to decide for one SNP. When local ancestry assignments are averaged

across the two haplotypes for an individual and across the whole genome, the ancestry proportions are very similar to those obtained by other methods such as ADMIXTURE [415] that are well-established for ancestry estimation, thus demonstrating the accuracy of RFMix [366]. The results were congruent and the SNPs with high frequency in Africans were assigned to African haplotypes, clearing possible errors in the genotyping or the quality control, etc.

Population structure

Therefore, this led me to think of other options that affect the performance of GWAS, such as population structure [416]. Accordingly, very stringent quality control was carried out to remove outliers, and PCs were recalculated after each removal, and Q-Q plots for all associations showed no sign of inflation (lambda <1.02), confirming no presence of population structure.

Number of samples

The success of genome-wide association studies depends on several variables, such as the number of samples, frequency of causal variants, the proportion of trait variance explained by a marker among other things depending on the type of GWAS[6]. In this case to clarify how important the number of samples is for a quantitative trait GWAS I present an example modified from Schmid et al. (2019) [417], in this study they replied to the question about what sample size can a genetic variant be significantly associated to a trait? For instance, for a SNP that explains 1% of variance, the power at n=2000 is only 15%, but at n=5000 is 95%, and n=500 the power of detection is merely 1%. Among other reasons, this could be the main cause of some associations found here are not totally reliable and further analyses are needed.

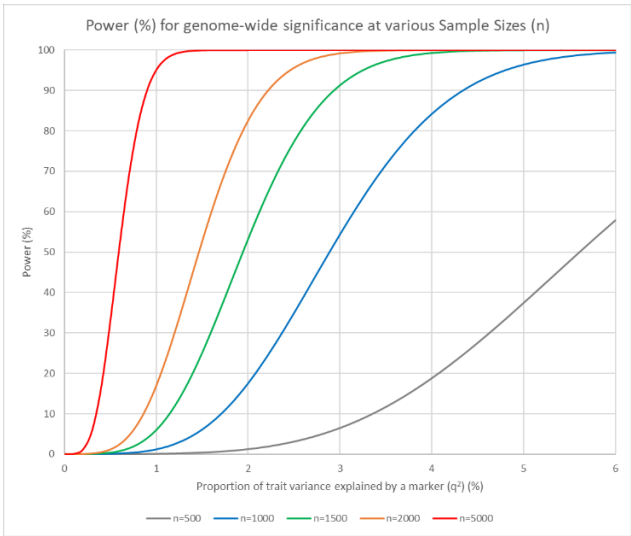


Figure 4.67. Power (%) for genome-wide significance at various sample sizes (n). Estimated power for various samples sizes, the x-axis represents the proportion of trait variance (q2) explained by a marker and y-axis denotes the estimated power (in percentage %).

Heritability calculation is also affected by the number of samples, this can be observed in the results obtained for heritability of ordinal traits (Section 4.3.2.3). Most values are either close to 1 or 0, meaning the algorithm convergence might have had problems given the small sample size [61], and therefore converged to one boundary or the other of the 0-1 interval. Only a few values are intermediate (i.e.0.3 or 0.5). But the variance is huge - SD is around 40%. Thus, except those with high heritability values, all others are non-significant, which means that even a heritability of 50% can be null i.e. 0%.

4.4.4 Future work

Finally, despite the sample size being very small and more samples and analyses are needed to improve these results, this exploratory research has contributed to draft the guidelines for further analysis, increasing the number of samples and the variability of the ancestry. Currently, our group is obtaining more samples in Colombia and Chile, to improve on the results obtained here.

Chapter 5: Conclusions

In this thesis I have performed several genome-wide association scans in a Latin-American population. As a result, I discovered new genomic variants related to facial and dental features.

In chapter 1 I have explored how facial and dental features are involved in different disciplines. I have also described the current knowledge about the evolution of the face and teeth. I have presented a summary on the genetics of the face and teeth. I have also described the genetic history of the settlement of America, focusing in Central and South America and how this affects the study of the different phenotypes. I have presented a summary about GWAS, highlighting advantages and limitations. And finally, I have described the studied samples. All these to support the main aim of this thesis, which is to identify genetic loci associated with human facial and dental traits in a Latin American population.

In chapter 2 I have described and explained the principal methods used on this thesis. The theory behind the genome-wide association studies and I have briefly mentioned about the history of geometric morphometry and the ASUDAS scoring method, to contextualize and explain the methods.

In chapter 3 I have performed a genome-wide scan seeking genetic variants associated with ordinal facial traits. I also did some follow-up analyses to endorse the associations found. I performed a second GWAS, this time with quantitative traits to prove if the associations obtained with ordinal phenotypes were real. I also performed a functional analysis of the effect of *EDAR* on chin protrusion.

In chapter 4 I have performed two genome-wide association studies with ordinal and quantitative dental traits. I have discovered some new possible genomic variants associated with dental traits and I also replicated the previous association between shovel shape incisors and the SNP rs3827760 (*EDAR* gene). I have also highlighted some limitations of the results obtained due to the number of samples being very small.

5.1 Impact of phenotyping strategies on GWAS

An important part of the genome-wide study is the assessed phenotype. The more the resolution is captured by the phenotype the more is the probability to detect associations [418]. Generally, the effect size of the genomic variants on phenotypes follows a gradient. The largest effect for those phenotypes directly affected by genes (endophenotypes), such as proteins. This is due to the competition that exist between genetic and non-genetic factors that contribute to the phenotype [249]. Thus, it would be expected that in phenotypes highly affected by non-genetic factors it will be harder to find genomic variants with a larger effect. This is reflecting how important is to obtain as much information as possible from the phenotype.

In this regard, in this thesis I presented two kinds of phenotyping methods, ordinal and quantitative phenotypes. Both methods were able to detect new genomic variants associated with facial and dental features, but with the quantitative method, especially morphometric geometry, the levels of significance obtained were higher and more associations were detected. In the case of ordinal phenotypes, the categories are assigned based on different criteria and not necessarily reflecting the level of variation within one category to another, thus important information is missed. In this thesis I use photogrammetry to obtain coordinates, distances and angles from 2D photographs, although this method is better than the categorical method. It depends on the expertise of the software operator to add the landmarks on the photos, additionally the photos may have different resolution, or the face is tilted, etc. and usually the number of landmarks is not enough to capture the whole phenotype variation. In order to improve this, several groups over the years have developed several phenotyping methods [419, 168] using MRI scans or even AI.

This thesis has offered a spectrum of different types of phenotypes and not only have they been useful for detecting new genomic variants, but also as way to compare between different phenotypes.

5.2 Future and significance of Genome Wide Association Studies

Over the years, countless genome-wide scans have been successful in finding genomic variants associated with complex disease or traits. But there are still several challenges to overcome to increase the significance of these findings. In this section I briefly attempt to mention some limitations, advantages and challenges of GWAS.

Genome-wide studies are already a part of an analysis pipeline because they offer an unbiased assessment of a large number of common alleles in complex phenotypes [418]. Although, it is expensive and it requires a big number of samples, it also gives the opportunity to narrow down the possible associations, and subsequently apply follow-up analyses, i.e. sequencing a specific region, which is cheaper than sequencing the whole genome. Furthermore, the amount of information that is generated by GWAS is priceless. Much effort should be done to increase the amount of information shared among different groups.

During the last year several scientific publications have discussed the issue about the diversity of the samples used to perform GWAS [239,420]. Around 73% of the populations analysed are from Europe or Asia. This difference has a huge impact in pharmacogenetics and pharmacogenomics because a drug can be useful in one population, but not in another population because of inter-individual variation dosage. Another issue is the lack of replication across different populations, due to differences in linkage disequilibrium across different ethnic groups. Markers in LD with the risk variant, may not be in LD in some populations [239].

An advantage of GWAS is that it can be used to identify individuals at high risk to present certain diseases, thus the disease can be detected, prevented and the treatment improved upon [420]. On the other hand, if the GWAS are performed in specific groups this advantage will not reach all the population.

Another benefit of Genome-wide scans is that they can help in the discovery of novel biological mechanisms that have never been studied before. And the follow-up of that finding may represent the discovery of new biological mechanisms [420].

The data produced by GWAS are utilized for several applications. For instance, reconstruction of population history, fine-scale estimation of location of birth, estimation of SNP heritability, polygenic risk scores, among others [420].

Many limitations have arisen since the beginning of the GWAS era. Nevertheless, most of the limitations can be overcome. For instance, the use of larger samples, increase in the diversity of the populations, advances in technology, the use of whole genome sequence (WGS) data instead of SNPs and this will allow to improve the risk prediction, detection of population stratification, identify more causal variants, explain more the missing heritability, etc.

Several authors have mentioned that whole genome sequence is the future of GWAS as prices decrease [421, 422]. Meanwhile, GWAS based in SNP genotyping arrays will keep on increasing the knowledge of the genetics variants behind complex phenotypes [420].

This thesis represents a contribution not only due to the new genetic variants associated with physical phenotypes, but also because of the different kind of phenotypes used, as well as the GWAS being performed in a population that was underrepresented in previous genome-wide scans, a group of Latin-American individuals.

5.3 The importance of the findings on this thesis in the morphology and development of the teeth and craniofacial features

Most of the genes or genetics regions implicated in morphology and development of craniofacial and dental features have been detected by studying familial cases, single cases of individuals presenting craniofacial and/or dental abnormalities or using animal models first, and then stablishing the relationship with 'normal craniofacial and dental features' and analysing how these changes affect the gene pathways acting over these traits and proposing how these genes are implicated in the normal development and morphology of the traits. In this case, I started searching for genes associated to normal facial and dental features. I found some genes previously associated to teeth or craniofacial anomalies, as well as genes that are involved in normal development of facial and dental traits.

For example, Hypohidrotic ectodermal dysplasia is a genetic condition that affects skin, hair, nails, teeth and sweat glands [423, 424, 425]. The most common phenotype is HED/EDA 1 (OMIM #305100), 1 out of 17,000 people present this X-linked hypohidrotic disorder. They show reduce capability to sweat, abnormal shape or lack of some teeth, sparse hair [426], some cases may show dysmorphic features, such as prognathism, forehead bumps, depressed nasal bridge and protruded forehead [427]. People with mutations in any of the genes belonging to the ectodysplasin A pathway (*EDA-EDAR* and *EDARADD*) present hypohidrotic ectodermal dysplasia [324].

The jaws are developed from the maxillary and mandibular process of pharyngeal arch 1 [22]. It is also known that the gene *Edn1*, which is expressed in ectodermal epithelium

of the mandibular process communicates via endothelin receptor A (encoded by *Edar* gene) [13, 23].

I detected an allele rs3827760 that falls in *EDAR* gene associated to chin protrusion in the face and shovel shape in the incisors. This allele, as mentioned before, was selected in an Asian population and later it was brought to America by the first settlers of the continent. It has a pleiotropic effect because the gene is affecting several phenotypes at the same time. However, this gene is expressed in the ectodermal epithelium and most of the phenotypes that derive from this layer during embryology development are shaped by the activity of this pathway (*EDA-EDAR* and *EDARADD*) where this gene is crucial. Thus, this pathway is very important in the embryonic ectodermal development and controls an early stage of embryology development and any mutation occurred at this point causes disturbance in the interaction between cells situated on the surface and the underlaying mesenchyme, leading to alterations in any stage of the skin appendages development.

Another interesting finding was the association between *SUPT3H/RUNX2* genomic region and nose bridge breadth. It is well known that *Runx2* is very important in the development of the bones, because *Runx2* is the first transcription factor required for the osteoblast maturation. In this genomic region there are two primers (P1 and P2), it has been proposed that they may control each other by producing two isoforms of *Runx2* (*Runx2-I* and *Runx2-II*) [428]. Mutations in *RUNX2* that cause loss of function have been associated to cleidocranial dysplasia, a delayed closure in cranial sutures [429]. Interestingly, *Runx2* transcription factor is necessary for the proliferation of osteoblasts by regulating *Fgfr2* and *Fgfr3* expression [430], and *FGFR2* has been associated with cleft lip and/or palate. Hence, the association I have found between *SUPT3H/RUNX2* region and nose bridge breadth suggest that this gene might be regulating not just the closure of sutures, but also the wide of this area of the nose by stimulating the proliferation and maturation of osteoblasts.

An allele (rs12570261) near by the gene *FGFR2* showed association with the protostylid in the first lower molar, this structure corresponds to the appearance of a secondary groove associated to the buccal groove separating cusps 1 and 3 on the buccal surface of the lower molars. During the tooth development, *Fgfr2* gene participates in the morphogenesis stage, when the primary enamel knot begins its

formation [106]. It is also known that individuals with mutations in this gene present Apert Syndrome, which is characterized by craniosynostosis, midface hypoplasia, pseudo cleft-palate [431], among other characteristics. Therefore, the association found in this study may be correct because I found this genetic marker that could be influencing the formation and normal development of teeth in humans.

The set of changes that allow the embryonic layers (ectoderm, mesoderm and endoderm) to transform into the different organs that make up an organism is a complex physiological process. An intricate array of signalling molecules such as FGFs, bone morphogenetic proteins (BMPs), Wnt, and Hedgehog (Hh) families are known to regulate the formation, differentiation, and maintenance of the tooth [432]and face during the development and throughout adulthood, and several of the genes that we have found through this study are implicated somehow or other in these processes.

Although we did not find all the genes or genomic regions detected by different methods and/or studies involved in the morphology and development of the face and teeth, this does not mean these genomic markers or genes are not associated to some of the processes shaping facial and dental features. This could be due to the genetic architecture of the studied population or the power of detection was not enough, etc. Perhaps because I was analysing dental and facial traits in adults, therefore their faces and teeth were totally developed, I only could detect signals from markers acting in the last stage of developing and not genomic regions expressing in other stages. Certainly, this is a question and discussion that I still have regarding this study and, it requires further analysis.

5.4 Is it feasible to predict the phenotype of a person based on their genetic information?

Prediction of phenotypes by Forensic Sciences has been very ambitious since the beginning and with the advances in DNA technology the Forensic DNA developing (FDP) was presented as the solution, for some traits such as hair and eye colour, FDP model prediction works, but it is not the same regarding other complex traits, such as facial and dental features.

Genome-wide associations scans were developed as a statistical analysis capable to detect a massive number of genetic markers associated to different complex

phenotypes (illnesses, physical features, etc.). It was thought that with this great amount of information, the prediction of certain features would be possible, but more than 10 years have passed since the beginning of GWAS and the accurate prediction of certain traits does not seem close.

In this thesis I presented the percentage of phenotypic variance explained by the associated alleles to different complex traits (facial and dental features), but none of these values surpassed the 2% of variance. This means that any of the detected associations explains more than 2% of the difference among one person to another for a particular trait. Then, Narrow-sense heritability (additive genetic variance) estimates were calculated on ordinal and quantitative facial phenotypes, and for some quantitative facial traits it was relatively high (between 0.45 and 0.90, Table 3.22), as expected for continuous variables. But this only explains a fraction of the total phenotypic variance.

The issue of missing heritability has been widely discussed, the fact that GWAS have identified a countless number of genetic variants showing association with complex traits. Nevertheless, these genomic variants only explain a modest proportion of the heritability of a trait. One possible explanation is that SNPs with small effects do not reach the significant threshold [420]. Consequently, the prediction of the phenotype of a complex trait based on the genetic information of a person or forensic DNA phenotyping (FDP), at this moment, seems very difficult. There are still many factors to consider in order to predict accurately complex phenotypes based only on DNA data, and these factors are not entirely related to genetics, also the environment plays a crucial role in shaping complex phenotypes.

References

1. Lacruz, R.S. *et al.* The evolutionary history of the human face. *Nature ecology & evolution*, 1 (2019).

2. Helms, J.A., Cordero, D. & Tapadia, M.D. New insights into craniofacial morphogenesis. *Development* **132**, 851-861 (2005).

3. Crouch, D.J. *et al.* Genetics of the human face: Identification of large-effect single gene variants. *Proceedings of the National Academy of Sciences* **115**, E676-E685 (2018).

4. Bouchard, T.J. Twins reared together and apart: What they tell us about human diversity. in *Individuality and determinism* 147-184 (Springer, 1984).

5. Lippert, C. *et al.* Identification of individuals by trait prediction using whole-genome sequencing data. *Proceedings of the National Academy of Sciences* **114**, 10166-10171 (2017).

6. Visscher, P.M. *et al.* 10 years of GWAS discovery: biology, function, and translation. *The American Journal of Human Genetics* **101**, 5-22 (2017).

7. Liu, F. *et al.* A genome-wide association study identifies five loci influencing facial morphology in Europeans. *PLoS Genet* **8**, e1002932 (2012).

8. Paternoster, L. *et al.* Genome-wide association study of three-dimensional facial morphology identifies a variant in PAX3 associated with nasion position. *Am J Hum Genet* **90**, 478-85 (2012).

9. Pillas, D. *et al.* Genome-wide association study reveals multiple loci associated with primary tooth development during infancy. *PLoS genetics* **6**, e1000856 (2010).

10. Geller, F. *et al.* Genome-wide association study identifies four loci associated with eruption of permanent teeth. *PLoS genetics* **7**, e1002275 (2011).

11. Li, Y., Willer, C., Sanna, S. & Abecasis, G. Genotype imputation. *Annual review of genomics and human genetics* **10**, 387-406 (2009).

12. Schoenwolf, G.C., Bleyl, S.B., Brauer, P.R. & Francis-West, P.H. *Larsen's human embryology E-book*, (Elsevier Health Sciences, 2014).

13. Twigg, S.R. & Wilkie, A.O. New insights into craniofacial malformations. *Human molecular genetics* **24**, R50-R59 (2015).

14. Lana-Elola, E., Rice, R., Grigoriadis, A.E. & Rice, D.P. Cell fate specification during calvarial bone and suture development. *Developmental biology* **311**, 335-346 (2007).

15. Deckelbaum, R.A. *et al.* Regulation of cranial morphogenesis and cell fate at the neural crest-mesoderm boundary by engrailed 1. *Development* **139**, 1346-1358 (2012).

16. Zhao, H. *et al.* The suture provides a niche for mesenchymal stem cells of craniofacial bones. *Nature cell biology* **17**, 386 (2015).

17. Minoux, M. & Rijli, F.M. Molecular mechanisms of cranial neural crest cell migration and patterning in craniofacial development. *Development* **137**, 2605-2621 (2010).

18. Leslie, E.J. *et al.* Comparative analysis of IRF6 variants in families with Van der Woude syndrome and popliteal pterygium syndrome using public whole-exome databases. *Genetics in Medicine* **15**, 338 (2013).

19. Kurosaka, H., Iulianella, A., Williams, T. & Trainor, P.A. Disrupting hedgehog and WNT signaling interactions promotes cleft lip pathogenesis. *The Journal of clinical investigation* **124**, 1660-1671 (2014).

20. Milunsky, J.M. *et al.* TFAP2A mutations result in branchio-oculo-facial syndrome. *The American Journal of Human Genetics* **82**, 1171-1177 (2008).

21. Iwata, J.-i. *et al.* Smad4-Irf6 genetic interaction and TGFβ-mediated IRF6 signaling cascade are crucial for palatal fusion in mice. *Development* **140**, 1220-1230 (2013).

22. Couly, G., Creuzet, S., Bennaceur, S., Vincent, C. & Le Douarin, N.M. Interactions between Hox-negative cephalic neural crest cells and the foregut endoderm in patterning the facial skeleton in the vertebrate head. *Development* **129**, 1061-1073 (2002).

23. Sato, T. *et al.* An endothelin-1 switch specifies maxillomandibular identity. *Proceedings of the National Academy of Sciences* **105**, 18806-18811 (2008).

24. Netter, F.H. *Atlas of Human Anatomy E-Book*, (Elsevier Health Sciences, 2012).

25. Marur, T., Tuna, Y. & Demirci, S. Facial anatomy. *Clinics in dermatology* **32**, 14-23 (2014).

26. Siemionow, M.Z. & Sonmez, E. Face as an organ: the functional anatomy of the face. in *The Know-How of Face Transplantation* 3-10 (Springer, 2011).

27. Dirkmaat, D. *A companion to forensic anthropology*, (John Wiley & Sons, 2014).

28. Martínez-Abadías, N. *et al.* Heritability of human cranial dimensions: comparing the evolvability of different cranial regions. *Journal of Anatomy* **214**, 19-35 (2009).

29. Dale, J., Lank, D.B. & Reeve, H.K. Signaling individual identity versus quality: a model and case studies with ruffs, queleas, and house finches. *The American Naturalist* **158**, 75-86 (2001).

30. Tibbetts, E.A. & Dale, J. Individual recognition: it is good to be different. *Trends in Ecology & Evolution* **22**, 529-537 (2007).

31. Du, S. & Martinez, A.M. Compound facial expressions of emotion: from basic research to clinical applications. *Dialogues in clinical neuroscience* **17**, 443 (2015).

32. Corruccini, R. The interaction between neurocranial and facial shape in hominid evolution. *Homo* **26**, 136-140 (1975).

33. Lieberman, D.E. Sphenoid shortening and the evolution of modern human cranial shape. *Nature* **393**, 158 (1998).

34. Lieberman, D.E., McBratney, B.M. & Krovitz, G. The evolution and development of cranial form in Homo sapiens. *Proceedings of the National Academy of Sciences* **99**, 1134-1139 (2002).

35. Navarrete, A., van Schaik, C.P. & Isler, K. Energetics and the evolution of human brain size. *Nature* **480**, 91 (2011).

36. Conroy, G.C., Jolly, C.J., Cramer, D. & Kalb, J.E. Newly discovered fossil hominid skull from the Afar depression, Ethiopia. *Nature* **276**, 67 (1978).

37. Liu, W., Zhang, Y. & Wu, X. Middle Pleistocene human cranium from Tangshan (Nanjing), Southeast China: a new reconstruction and comparisons with Homo erectus from Eurasia and Africa. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists* **127**, 253-262 (2005).

38. Stringer, C. & Galway-Witham, J. Palaeoanthropology: On the origin of our species. *Nature* **546**, 212 (2017).

39. White, T.D. *et al.* Pleistocene homo sapiens from middle awash, ethiopia. *Nature* **423**, 742 (2003).

40. Ferring, R. *et al.* Earliest human occupations at Dmanisi (Georgian Caucasus) dated to 1.85–1.78 Ma. *Proceedings of the National Academy of Sciences* **108**, 10432-10436 (2011).

41. Harvati, K.H. *Handbook of Paleoanthropology*, (Springer, 2014).

42. Dean, M.C. Another look at the nose and the functional significance of the face and nasal mucous membrane for cooling the brain in fossil hominids. *Journal of Human Evolution* **17**, 715-718 (1988).

43. Stewart, J.R. & Stringer, C.B. Human evolution out of Africa: the role of refugia and climate change. *Science* **335**, 1317-1321 (2012).

44. von Cramon-Taubadel, N. Evolutionary insights into global patterns of human cranial diversity: population history, climatic and dietary effects. *J. Anthropol. Sci* **92**(2014).

45. Hubbe, M., Hanihara, T. & Harvati, K. Climate signatures in the morphological differentiation of worldwide modern human populations. *The Anatomical*

Record: Advances in Integrative Anatomy and Evolutionary Biology **292**, 1720-1733 (2009).

46. Evteev, A., Cardini, A.L., Morozova, I. & O'Higgins, P. Extreme climate, rather than population history, explains mid-facial morphology of Northern Asians. *American journal of physical anthropology* **153**, 449-462 (2014).

47. Reyes-Centeno, H., Ghirotto, S. & Harvati, K. Genomic validation of the differential preservation of population history in modern human cranial anatomy. *American journal of physical anthropology* **162**, 170-179 (2017).

48. Noback, M.L., Harvati, K. & Spoor, F. Climate-related variation of the human nasal cavity. *American journal of physical anthropology* **145**, 599-614 (2011).

49. Cieri, R.L. *et al.* Craniofacial feminization, social tolerance, and the origins of behavioral modernity. *Current Anthropology* **55**, 000-000 (2014).

50. Godinho, R.M., Spikins, P. & O'Higgins, P. Supraorbital morphology and social dynamics in human evolution. *Nature ecology & evolution* **2**, 956 (2018).

51. Weinberg, S.M., Parsons, T.E., Marazita, M.L. & Maher, B.S. Heritability of Face Shape in Twins: A Preliminary Study using 3D Stereophotogrammetry and Geometric Morphometrics. *Dent 3000* **1**(2013).

52. Hughes, B.O. Heredity as a factor in cranial and facial development. (1942).

53. Kraus, B.S., Wise, W.J. & Frei, R.H. Heredity and the craniofacial complex. *American journal of orthodontics* **45**, 172-217 (1959).

54. Moore, G.R. & Hughes, B.O. Familial factors in diagnosis, treatment, and prognosis of dentofacial disturbances. *American Journal of Orthodontics and Oral Surgery* **28**, 603-639 (1942).

55. Cole, J.B. & Spritz, R.A. The Genetics of Facial Morphology. *eLS*, 1-9 (2001).

56. Byard, P., Lewis, A., Ohtsuki, F., Siervogel, R. & Roche, A. Sibling correlations for cranial measurements from serial radiographs. *Journal of craniofacial genetics and developmental biology* **4**, 265-269 (1984).

57. Kohn, L. The role of genetics in craniofacial morphology and growth. *Annual Review of Anthropology* **20**, 261-278 (1991).

58. Adhikari, K. *et al.* A genome-wide association study identifies multiple loci for variation in human ear morphology. *Nature communications* **6**, 7500 (2015).

59. Ruiz-Linares, A. *et al.* Admixture in Latin America: geographic structure, phenotypic diversity and self-perception of ancestry based on 7,342 individuals. *PLoS genetics* **10**, e1004572 (2014).

60. Speed, D., Hemani, G., Johnson, M.R. & Balding, D.J. Improved heritability estimation from genome-wide SNPs. *The American Journal of Human Genetics* **91**, 1011-1021 (2012).

61. Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M. GCTA: a tool for genome-wide complex trait analysis. *The American Journal of Human Genetics* **88**, 76-82 (2011).

62. Carson, E.A. Maximum likelihood estimation of human craniometric heritabilities. *American Journal of Physical Anthropology: The Official*

Publication of the American Association of Physical Anthropologists **131**, 169-180 (2006).

63. Naini, F.B. & Moss, J.P. Three-dimensional assessment of the relative contribution of genetics and environment to various facial parameters with the twin method. *American Journal of Orthodontics and Dentofacial Orthopedics* **126**, 655-665 (2004).

64. Chacón-Duque, J.-C. *et al.* Latin Americans show wide-spread Converso ancestry and imprint of local Native ancestry on physical appearance. *Nature communications* **9**(2018).

65. Carels, C. *et al.* A quantitative genetic study of cephalometric variables in twins. *Clinical orthodontics and research* **4**, 130-140 (2001).

66. Cole, J.B. *et al.* Human facial shape and size heritability and genetic correlations. *Genetics* **205**, 967-978 (2017).

67. Jelenkovic, A., Poveda, A., Susanne, C. & Rebato, E. Common genetic and environmental factors among craniofacial traits in Belgian nuclear families: comparing skeletal and soft-tissue related phenotypes. *Homo* **61**, 191-203 (2010).

68. Richmond, S., Howe, L.J., Lewis, S., Stergiakouli, E. & Zhurov, A. Facial Genetics: A Brief Overview. *Frontiers in genetics* **9**(2018).

69. Lange, S., Shield, K., Koren, G., Rehm, J. & Popova, S. A comparison of the prevalence of prenatal alcohol exposure obtained via maternal self-reports versus meconium testing: a systematic literature review and meta-analysis. *BMC pregnancy and childbirth* **14**, 127 (2014).

70. Jaillet, J., Robert-Gnansia, E., Till, M., Vinciguerra, C. & Edery, P. Biliary lithiasis in early pregnancy and abnormal development of facial and distal limb bones (Binder syndrome): a possible role for vitamin K deficiency. *Birth Defects Research Part A: Clinical and Molecular Teratology* **73**, 188-193 (2005).

71. Wickstrom, R. Effects of nicotine during pregnancy: human and experimental evidence. *Current neuropharmacology* **5**, 213-222 (2007).

72. Mirghani, H., Osman, N., Dhanasekaran, S., Elbiss, H.M. & Bekdache, G. Transplacental transfer of 2-naphthol in human placenta. *Toxicology reports* **2**, 957-960 (2015).

73. Hoyme, H.E. *et al.* Updated clinical guidelines for diagnosing fetal alcohol spectrum disorders. *Pediatrics* **138**(2016).

74. Beaty, T. *et al.* Confirming genes influencing risk to cleft lip with/without cleft palate in a case–parent trio study. *Human genetics* **132**, 771-781 (2013).

75. Buretić-Tomljanović, A., Giacometti, J., Ostojić, S. & Kapović, M. Sex-specific differences of craniofacial traits in Croatia: the impact of environment in a small geographic area. *Annals of human biology* **34**, 296-314 (2007).

76. McKusick, V. Online Mendelian Inheritance in Man, OMIM (TM). McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University (Baltimore, MD) and National Center for Biotechnology Information, National Library of Medicine (Bethesda, MD). *URL http://www. ncbi. nlm. nih. gov/omim* (2000).

77. Jones, K.L., Jones, M.C. & Del Campo, M. *SPEC-Smith's Recognizable Patterns of Human Malformation, 12-Month Access, eBook: Expert Consult-Online and Print*, (Elsevier Health Sciences, 2013).

78. Szabo-Rogers, H.L., Smithers, L.E., Yakob, W. & Liu, K.J. New directions in craniofacial morphogenesis. *Developmental biology* **341**, 84-94 (2010).

79. Gorlin, R.J., Cohen Jr, M.M. & Hennekam, R.C. *Syndromes of the head and neck*, (Oxford university press, 2001).

80. Johnson, D., Iseki, S., Wilkie, A. & Morriss-Kay, G. Expression patterns of Twist and Fgfr1,-2 and-3 in the developing mouse coronal suture suggest a key role for twist in suture initiation and biogenesis. *Mechanisms of development* **91**, 341-345 (2000).

81. Bialek, P. *et al.* A twist code determines the onset of osteoblast differentiation. *Developmental cell* **6**, 423-435 (2004).

82. El Ghouzzi, V. TWIST mutations disrupting the b-HLH domain are specific to Saether-Chotzen syndrome. *Am J Hum Genet Suppl* **61**, 332 (1997).

83. Sharma, V.P. *et al.* Mutations in TCF12, encoding a basic helix-loop-helix partner of TWIST1, are a frequent cause of coronal craniosynostosis. *Nature genetics* **45**, 304 (2013).

84. Buchanan, E.P., Xue, A.S. & Hollier Jr, L.H. Craniofacial syndromes. *Plastic and reconstructive surgery* **134**, 128e-153e (2014).

85. Azoury, S.C., Reddy, S., Shukla, V. & Deng, C.-X. Fibroblast growth factor receptor 2 (FGFR2) mutation related syndromic craniosynostosis. *International journal of biological sciences* **13**, 1479 (2017).

86. Dixon, M.J., Marazita, M.L., Beaty, T.H. & Murray, J.C. Cleft lip and palate: understanding genetic and environmental influences. *Nature Reviews Genetics* **12**, 167 (2011).

87. Deshpande, A.S. & Goudy, S.L. Cellular and molecular mechanisms of cleft palate development. *Laryngoscope Investigative Otolaryngology* **4**, 160-164 (2019).

88. Jones, M.C. Etiology of facial clefts: prospective evaluation of 428 patients. *The Cleft palate journal* **25**, 16-20 (1988).

89. Venkatesh, R. Syndromes and anomalies associated with cleft. *Indian journal of plastic surgery: official publication of the Association of Plastic Surgeons of India* **42**, S51 (2009).

90. Kati, F.A. Cleft lip and palate: review article. *World Journal of Pharmaceutical Research* **4**, 155-163 (2018).

91. Vanderas, A.P. Incidence of cleft lip, cleft palate, and cleft lip and palate among races: a review. *The Cleft palate journal* **24**, 216-225 (1987).

92. Dixon, J. *et al.* Tcof1 Treacle is required for neural crest cell formation and proliferation deficiencies that cause craniofacial abnormalities. *PNAS* **103**, 13403-13408 (2006).

93. Ludwig, K.U. *et al.* Genome-wide meta-analyses of nonsyndromic cleft lip with or without cleft palate identify six new risk loci. *Nature genetics* **44**, 968 (2012).

94. Robin, N.H., Moran, R.T. & Ala-Kokko, L. Stickler syndrome. in *GeneReviews®[Internet]* (University of Washington, Seattle, 2017).

95. Adhikari, K. *et al.* A genome-wide association scan implicates DCHS2, RUNX2, GLI3, PAX1 and EDAR in human facial variation. **7**, 11616 (2016).

96. Zalc, A., Rattenbach, R., Auradé, F., Cadot, B. & Relaix, F. Pax3 and Pax7 play essential safeguard functions against environmental stress-induced birth defects. *Developmental cell* **33**, 56-66 (2015).

97. Ludwig, K. *et al.* Strong association of variants around FOXE1 and orofacial clefting. *Journal of Dental Research* **93**, 376-381 (2014).

98. Al Chawa, T. *et al.* Nonsyndromic cleft lip with or without cleft palate: Increased burden of rare variants within Gremlin-1, a component of the bone morphogenetic protein 4 pathway. *Birth Defects Research Part A: Clinical and Molecular Teratology* **100**, 493-498 (2014).

99. Bureau, A. *et al.* Whole exome sequencing of distant relatives in multiplex families implicates rare variants in candidate genes for oral clefts. *Genetics* **197**, 1039-1044 (2014).

100. Butali, A. *et al.* Rare functional variants in genome–wide association identified candidate genes for nonsyndromic clefts in the African population. *American Journal of Medical Genetics Part A* **164**, 2567-2571 (2014).

101. Butali, A. *et al.* Replication of Genome Wide Association Identified Candidate Genes Confirm the Role of Common and Rare Variants in PAX 7 and VAX 1 in the Etiology of Nonsyndromic CL (P). *American Journal of Medical Genetics Part A* **161**, 965-972 (2013).

102. Ludwig, K.U. *et al.* Evaluating eight newly identified susceptibility loci for nonsyndromic cleft lip with or without cleft palate in a Mesoamerican population. *Birth Defects Research Part A: Clinical and Molecular Teratology* **100**, 43-47 (2014).

103. Chen, Q. *et al.* Joint testing of genotypic and gene-environment interaction identified novel association for BMP4 with non-syndromic CL/P in an Asian

population using data from an International Cleft Consortium. *PloS one* **9**, e109038 (2014).

104. Uslu, V.V. *et al.* Long-range enhancers regulating Myc expression are required for normal facial morphogenesis. *Nature genetics* **46**, 753 (2014).

105. Gómez-Robles, A. Dental evolutionary rates and its implications for the Neanderthal–modern human divergence. *Science advances* **5**, eaaw1268 (2019).

106. Bei, M. Molecular genetics of tooth development. *Current opinion in genetics & development* **19**, 504-510 (2009).

107. Vaahtokari, A., Åberg, T., Jernvall, J., Keränen, S. & Thesleff, I. The enamel knot as a signaling center in the developing mouse tooth. *Mechanisms of development* **54**, 39-43 (1996).

108. Salazar-Ciudad, I. Tooth patterning and evolution. *Current opinion in genetics & development* **22**, 585-592 (2012).

109. Wang, X.-P. *et al.* An integrated gene regulatory network controls stem cell proliferation in teeth. *PLoS biology* **5**, e159 (2007).

110. Butler, P.M. The ontogeny of molar pattern. *Biological Reviews* **31**, 30-69 (1956).

111. Scott, G.R. & Turner, C.G. *Anthropology of modern human teeth*, (Cambridge University Press Cambridge, 1997).

112. Türp, J., Brace, C. & Alt, K. Richard Owen and the comparative anatomy of teeth. (1997).

113. Weiss, K., Stock, D. & Zhao, Z. Dynamic interactions and the evolutionary genetics of dental patterning. *Critical Reviews in Oral Biology & Medicine* **9**, 369-398 (1998).

114. Darwin, C. *On the origin of species, 1859*, (Routledge, 2004).

115. Bateson, W. *Materials for the study of variation: treated with especial regard to discontinuity in the origin of species*, (Macmillan, 1894).

116. Gregory, W.K. A Half Century of Trituberculy the Cope-Osborn Theory of Dental Evolution with a Revised Summary of Molar Evolution from Fish to Man. *Proceedings of the American Philosophical Society* **73**, 169-317 (1934).

117. Mario, F. & Ricardo, G. Anatomía odontológica funcional y aplicada. *El Ateneo* (1994).

118. García Sívoli, C.E. *Estudio diacrónico de los rasgos dentales en poblaciones del Mediterráneo occidental Mallorca y Cataluña*, (Universitat Autònoma de Barcelona, 2009).

119. Hlusko, L.J. Integrating the genotype and phenotype in hominid paleontology. *Proceedings of the National Academy of Sciences* **101**, 2653-2657 (2004).

120. Donoghue, P.C. & Rücklin, M. The ins and outs of the evolutionary origin of teeth. *Evolution & development* **18**, 19-30 (2016).

121. Suwa, G. *et al.* The Ardipithecus ramidus skull and its implications for hominid origins. *Science* **326**, 68-68e7 (2009).

122. Brunet, M. *et al.* A new hominid from the Upper Miocene of Chad, Central Africa. *Nature* **418**, 145 (2002).

123. Hardt, T. *Handbook of Paleoanthropology: Vol I: Principles, Methods and Approaches Vol II: Primate Evolution and Human Origins Vol III: Phylogeny of Hominids*, (Springer Science & Business Media, 2007).

124. Sponheimer, M. *et al.* Hominins, sedges, and termites: new carbon isotope data from the Sterkfontein valley and Kruger National Park. *Journal of Human Evolution* **48**, 301-312 (2005).

125. Emes, Y., Aybar, B. & Yalcin, S. On the evolution of human jaws and teeth: A review. *Bulletin of the International association for paleodontology* **5**, 37-47 (2011).

126. Wood, B. & Lieberman, D.E. Craniodental variation in Paranthropus boisei: a developmental and functional perspective. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists* **116**, 13-25 (2001).

127. Schoetensack, O. Der Unterkiefer des Homo heidelbergensis aus den Sanden von Mauer bei Heidelberg. Ein Beitrag zur Paläontologie des Menschen. *Molecular and General Genetics MGG* **1**, 408-410 (1908).

128. Scott, G.R. & Irish, J.D. *Anthropological perspectives on tooth morphology: genetics, evolution, variation*, (Cambridge University Press, 2013).

129. Scott, G.R. *et al.* Sinodonty, Sundadonty, and the Beringian Standstill model: Issues of timing and migrations into the New World. *Quaternary international* **466**, 233-246 (2018).

130. Rathmann, H., Saltini Semerari, G. & Harvati, K. Evidence for Migration Influx into the Ancient Greek Colony of Metaponto: A Population Genetics Approach Using Dental Nonmetric Traits. *International Journal of Osteoarchaeology* **27**, 453-464 (2017).

131. Harvati, K. *et al.* A human deciduous molar from the Middle Stone Age (Howiesons Poort) of Klipdrift Shelter, South Africa. *Journal of human evolution* **82**, 190-196 (2015).

132. Rathmann, H. *et al.* Reconstructing human population history from dental phenotypes. *Scientific reports* **7**, 12495 (2017).

133. Harris, E.F. *Anthropologic and genetic aspects of the dental morphology of Solomon Islanders, Melanesia*, (Arizona State University, 1977).

134. Hubbard, A.R., Guatelli-Steinberg, D. & Irish, J.D. Do nuclear DNA and dental nonmetric data produce similar reconstructions of regional population history? An example from modern coastal K enya. *American journal of physical anthropology* **157**, 295-304 (2015).

135. Delgado, M. *et al.* Variation in dental morphology and inference of continental ancestry in admixed Latin Americans. *American journal of physical anthropology* **168**, 438-447 (2019).

136. Rubicz, R., Melton, P. & Crawford, M. H. In Anthropological Genetics: Theory, Methods, and Applications

in *In Anthropological Genetics: Theory, Methods, and Applications* (ed. Crawford, M.H.) 141-186 (Cambridge University Press, Cambridge, 2007).

137. Mikkola, M.L. Genetic basis of skin appendage development. in *Seminars in cell & developmental biology* Vol. 18 225-236 (Elsevier, 2007).

138. Fleischmannova, J., Matalova, E., Tucker, A.S. & Sharpe, P.T. Mouse models of tooth abnormalities. *European journal of oral sciences* **116**, 1-10 (2008).

139. Bei, M. Molecular genetics of ameloblast cell lineage. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* **312**, 437-444 (2009).

140. Catón, J. & Tucker, A.S. Current knowledge of tooth development: patterning and mineralization of the murine dentition. *Journal of Anatomy* **214**, 502-515 (2009).

141. Tummers, M. & Thesleff, I. The importance of signal pathway modulation in all aspects of tooth development. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* **312**, 309-319 (2009).

142. Kondo, S. & Miura, T. Reaction-diffusion model as a framework for understanding biological pattern formation. *science* **329**, 1616-1620 (2010).

143. Thesleff, I. Epithelial-mesenchymal signalling regulating tooth morphogenesis. *Journal of cell science* **116**, 1647-1648 (2003).

144. Dassule, H.R. & McMahon, A.P. Analysis of epithelial–mesenchymal interactions in the initial morphogenesis of the mammalian tooth. *Developmental biology* **202**, 215-227 (1998).

145. Kapadia, H., Mues, G. & D'souza, R. Genes affecting tooth morphogenesis. *Orthodontics & craniofacial research* **10**, 237-244 (2007).

146. Miletich, I. & Sharpe, P.T. Normal and abnormal dental development. *Human molecular genetics* **12**, R69-R73 (2003).

147. Haworth, K.E., Healy, C., Morgan, P. & Sharpe, P.T. Regionalisation of early head ectoderm is regulated by endoderm and prepatterns the orofacial epithelium. *Development* **131**, 4797-4806 (2004).

148. Ghergie, M., Cocîrla, E., Lupan, I., Kelemen, B.S. & Popescu, O. Genes and dental disorders. *Clujul Medical* **86**, 196 (2013).

149. Tucker, A.S., Matthews, K.L. & Sharpe, P.T. Transformation of tooth type induced by inhibition of BMP signaling. *Science* **282**, 1136-1138 (1998).

150. Tucker, A. & Sharpe, P. The cutting-edge of mammalian development; how the embryo makes teeth. *Nature Reviews Genetics* **5**, 499 (2004).

151. Klein, M., Nieminen, P., Lammi, L., Niebuhr, E. & Kreiborg, S. Novel mutation of the initiation codon of PAX9 causes oligodontia. *Journal of dental research* **84**, 43-47 (2005).

152. Vastardis, H., Karimbux, N., Guthua, S.W., Seidman, J. & Seidman, C.E. A human MSX1 homeodomain missense mutation causes selective tooth agenesis. *Nature genetics* **13**, 417 (1996).

153. Stockton, D.W., Das, P., Goldenberg, M., D'Souza, R.N. & Patel, P.I. Mutation of PAX9 is associated with oligodontia. *Nature genetics* **24**, 18 (2000).

154. Lammi, L. *et al.* Mutations in AXIN2 cause familial tooth agenesis and predispose to colorectal cancer. *The American Journal of Human Genetics* **74**, 1043-1050 (2004).

155. Tao, R. *et al.* A novel missense mutation of the EDA gene in a Mongolian family with congenital hypodontia. *Journal of human genetics* **51**, 498 (2006).

156. Kayser, M. & Schneider, P.M. DNA-based prediction of human externally visible characteristics in forensics: motivations, scientific challenges, and ethical considerations. *Forensic Science International: Genetics* **3**, 154-161 (2009).

157. Kayser, M. & De Knijff, P. Improving human forensics through advances in genetics, genomics and molecular biology. *Nature Reviews Genetics* **12**, 179 (2011).

158. Chaitanya, L. *et al.* The HIrisPlex-S system for eye, hair and skin colour prediction from DNA: Introduction and forensic developmental validation. *Forensic Science International: Genetics* **35**, 123-135 (2018).

159. Draus-Barini, J. *et al.* Bona fide colour: DNA prediction of human eye and hair colour from ancient and contemporary skeletal remains. *Investigative genetics* **4**, 3 (2013).

160. King, T.E. *et al.* Identification of the remains of King Richard III. *Nature communications* **5**, 5631 (2014).

161. Olalde, I. *et al.* Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature* **507**, 225 (2014).

162. Kanetsky, P.A. *et al.* A polymorphism in the agouti signaling protein gene is associated with human pigmentation. *The American Journal of Human Genetics* **70**, 770-775 (2002).

163. Stokowski, R.P. *et al.* A genomewide association study of skin pigmentation in a South Asian population. *The American Journal of Human Genetics* **81**, 1119-1132 (2007).

164. Duffy, D.L. *et al.* A three–single-nucleotide polymorphism haplotype in intron 1 of OCA2 explains most human eye-color variation. *The American Journal of Human Genetics* **80**, 241-252 (2007).

165. Sturm, R.A. *et al.* A single SNP in an evolutionary conserved region within intron 86 of the HERC2 gene determines human blue-brown eye color. *The American Journal of Human Genetics* **82**, 424-431 (2008).

166. Walsh, S. *et al.* IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. *Forensic Science International: Genetics* **5**, 170-180 (2011).

167. Walsh, S. *et al.* Developmental validation of the IrisPlex system: determination of blue and brown iris colour for forensic intelligence. *Forensic Science International: Genetics* **5**, 464-471 (2011).

168. Claes, P. *et al.* Modeling 3D facial shape from DNA. *PLoS genetics* **10**, e1004224 (2014).

169. Weedon, M.N. *et al.* Genome-wide association analysis identifies 20 loci that influence adult height. *Nature genetics* **40**, 575 (2008).

170. Biedermann, A., Bozza, S. & Taroni, F. Prediction in forensic science: a critical examination of common understandings. *Frontiers in psychology* **6**, 737 (2015).

171. Kayser, M. Forensic DNA phenotyping: predicting human appearance from crime scene material for investigative purposes. *Forensic Science International: Genetics* **18**, 33-48 (2015).

172. Divakar, K. Forensic odontology: The new dimension in dental analysis. *International journal of biomedical science: IJBS* **13**, 1 (2017).

173. Adams, C., Carabott, R. & Evans, S. *Forensic Odontology: an essential guide*, (John Wiley & Sons, 2013).

174. Kotrashetti, V.S., Hollikatti, K., Mallapur, M., Hallikeremath, S.R. & Kale, A.D. Determination of palatal rugae patterns among two ethnic populations of India by logistic regression analysis. *Journal of forensic and legal medicine* **18**, 360-365 (2011).

175. Ohtani, M. *et al.* Indication and limitations of using palatal rugae for personal identification in edentulous cases. *Forensic science international* **176**, 178-182 (2008).

176. Krishan, K., Kanchan, T. & Garg, A.K. Dental evidence in forensic identification–An overview, methodology and present status. *The open dentistry journal* **9**, 250 (2015).

177. Plavcan, J.M. Sexual size dimorphism, canine dimorphism, and male-male competition in primates. *Human Nature* **23**, 45-67 (2012).

178. Reich, D. *et al.* Reconstructing native American population history. *Nature* **488**, 370 (2012).

179. Pitulko, V.V. *et al.* The Yana RHS site: humans in the Arctic before the last glacial maximum. *Science* **303**, 52-56 (2004).

180. Tamm, E. *et al.* Beringian standstill and spread of Native American founders. *PloS one* **2**, e829 (2007).

181. Raghavan, M. *et al.* Genomic evidence for the Pleistocene and recent population history of Native Americans. *Science* **349**, aab3884 (2015).

182. Pinotti, T. *et al.* Y Chromosome Sequences Reveal a Short Beringian Standstill, Rapid Expansion, and early Population structure of Native American Founders. *Current Biology* **29**, 149-157. e3 (2019).

183. Dillehay, T.D. Probing deeper into first American studies. *Proceedings of the National Academy of Sciences* **106**, 971-978 (2009).

184. Adhikari, K., Chacón-Duque, J.C., Mendoza-Revilla, J., Fuentes-Guajardo, M. & Ruiz-Linares, A. The genetic diversity of the Americas. *Annual review of genomics and human genetics* **18**, 277-296 (2017).

185. Rasmussen, M. *et al.* The genome of a Late Pleistocene human from a Clovis burial site in western Montana. *Nature* **506**, 225 (2014).

186. Moreno-Mayar, J.V. *et al.* Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans. *Nature* **553**, 203 (2018).

187. Moreno-Mayar, J.V. *et al.* Early human dispersals within the Americas. *Science* **362**, eaav2621 (2018).

188. Skoglund, P. & Reich, D. A genomic view of the peopling of the Americas. *Current opinion in genetics & development* **41**, 27-35 (2016).

189. Rothhammer, F. & Dillehay, T.D. The late Pleistocene colonization of South America: an interdisciplinary perspective. *Annals of human genetics* **73**, 540-549 (2009).

190. Denevan, W.M. *The native population of the Americas in 1492*, (Univ of Wisconsin Press, 1992).

191. Bellwood, P.S. First farmers: the origins of agricultural societies. (2005).

192. Kent, R.B. *Latin America: regions and people*, (Guilford Publications, 2016).

193. Salzano, F.M. & Bortolini, M.C. *The evolution and genetics of Latin American populations*, (Cambridge University Press, 2005).

194. Montenegro, R.A. & Stephens, C. Indigenous health in Latin America and the Caribbean. *The Lancet* **367**, 1859-1869 (2006).

195. Ongaro, L. *et al.* The genomic impact of European colonization of the Americas. *bioRxiv*, 676437 (2019).

196. Browning, S.R. *et al.* Ancestry-specific recent effective population size in the Americas. *PLoS genetics* **14**, e1007385 (2018).

197. Curtin, P.D. *The Atlantic slave trade: a census*, (Univ of Wisconsin Press, 1972).

198. Ruiz-Linares, A. How genes have illuminated the history of early Americans and Latino Americans. *Cold Spring Harbor perspectives in biology* **7**, a008557 (2015).

199. Carvajal-Carmona, L.G. *et al.* Strong Amerind/white sex bias and a possible Sephardic contribution among the founders of a population in northwest Colombia. *The American Journal of Human Genetics* **67**, 1287-1295 (2000).

200. Alves-Silva, J. *et al.* The ancestry of Brazilian mtDNA lineages. *The American Journal of Human Genetics* **67**, 444-461 (2000).

201. Green, L.D., Derr, J.N. & Knight, A. mtDNA affinities of the peoples of North-Central Mexico. *The American Journal of Human Genetics* **66**, 989-998 (2000).

202. Rothhammer, F., Puddu, G. & Fuentes-Guajardo, M. CAN MITOCHONDRIAL DNA PROVIDE INFORMATION ON THE ETHNOGENESIS OF CHILEAN NATIVE POPULATIONS? *STRUCTURE* **26**, 635-+ (2018).

203. Adhikari, K., Mendoza-Revilla, J., Chacón-Duque, J.C., Fuentes-Guajardo, M. & Ruiz-Linares, A. Admixture in Latin America. *Current opinion in genetics & development* **41**, 106-114 (2016).

204. Wang, S. *et al.* Geographic patterns of genome admixture in Latin American Mestizos. *PLoS genetics* **4**, e1000037 (2008).

205. Kehdy, F.S. *et al.* Origin and dynamics of admixture in Brazilians and its effect on the pattern of deleterious mutations. *Proceedings of the National Academy of Sciences* **112**, 8696-8701 (2015).

206. Homburger, J.R. *et al.* Genomic insights into the ancestry and demographic history of South America. *PLoS genetics* **11**, e1005602 (2015).

207. Conomos, M.P. *et al.* Genetic diversity and association studies in US Hispanic/Latino populations: applications in the Hispanic Community Health Study/Study of Latinos. *The American Journal of Human Genetics* **98**, 165-184 (2016).

208. Moreno-Estrada, A. *et al.* The genetics of Mexico recapitulates Native American substructure and affects biomedical traits. *Science* **344**, 1280-1285 (2014).

209. Eyheramendy, S., Martinez, F.I., Manevy, F., Vial, C. & Repetto, G.M. Genetic structure characterization of Chileans reflects historical immigration patterns. *Nature communications* **6**, 6472 (2015).

210. Bedoya, G. *et al.* Admixture dynamics in Hispanics: a shift in the nuclear genetic ancestry of a South American population isolate. *Proceedings of the National Academy of Sciences* **103**, 7234-7239 (2006).

211. Adhikari, K. *et al.* A genome-wide association scan in admixed Latin Americans identifies loci influencing facial and scalp hair features. **7**, 10815 (2016).

212. Hordes, S.M. *To the end of the earth: a history of the crypto-Jews of New Mexico*, (Columbia University Press, 2008).

213. Velez, C. *et al.* The impact of Converso Jews on the genomes of modern Latin Americans. *Human genetics* **131**, 251-263 (2012).

214. Richardson, D. & Eltis, D. *Atlas of the transatlantic slave trade*, (Yale University Press, 2015).

215. Barry, B. *Senegambia and the Atlantic slave trade*, (Cambridge University Press, 1998).

216. Edwards, E. Slavery in Argentina. (2014).

217. Goetz, L.H., Uribe-Bruce, L., Quarless, D., Libiger, O. & Schork, N.J. Admixture and clinical phenotypic variation. *Human heredity* **77**, 73-86 (2014).

218. Pino-Yanes, M. *et al.* Genetic ancestry influences asthma susceptibility and lung function among Latinos. *Journal of Allergy and Clinical Immunology* **135**, 228-235 (2015).

219. Bermejo, J.L. *et al.* Subtypes of Native American ancestry and leading causes of death: Mapuche ancestry-specific associations with gallbladder cancer risk in Chile. *PLoS genetics* **13**, e1006756 (2017).

220. Mountain, J.L. & Risch, N. Assessing genetic contributions to phenotypic differences among'racial'and'ethnic'groups. *Nature genetics* **36**, S48 (2004).

221. Tishkoff, S.A. & Verrelli, B.C. Patterns of human genetic diversity: implications for human evolutionary history and disease. *Annual review of genomics and human genetics* **4**, 293-340 (2003).

222. Eichstaedt, C.A. *et al.* The Andean adaptive toolkit to counteract high altitude maladaptation: genome-wide and phenotypic analysis of the Collas. *PloS one* **9**, e93314 (2014).

223. Mallick, S. *et al.* The Simons genome diversity project: 300 genomes from 142 diverse populations. *Nature* **538**, 201 (2016).

224. Relethford, J.H. Apportionment of global human genetic diversity based on craniometrics and skin color. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists* **118**, 393-398 (2002).

225. Relethford, J.H. Race and global patterns of phenotypic variation. *American journal of physical anthropology* **139**, 16-22 (2009).

226. Visscher, P.M. *et al.* 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am J Hum Genet* **101**, 5-22 (2017).

227. Visscher, P.M., Brown, M.A., McCarthy, M.I. & Yang, J. Five years of GWAS discovery. *Am J Hum Genet* **90**, 7-24 (2012).

228. Boyle, E.A., Li, Y.I. & Pritchard, J.K. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* **169**, 1177-1186 (2017).

229. Eichler, E.E. *et al.* Missing heritability and strategies for finding the underlying causes of complex disease. *Nat Rev Genet* **11**, 446-50 (2010).

230. Manolio, T.A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461**, 747-53 (2009).

231. Shi, H., Kichaev, G. & Pasaniuc, B. Contrasting the Genetic Architecture of 30 Complex Traits from Summary Association Data. *Am J Hum Genet* **99**, 139-53 (2016).

232. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* **42**, 565-9 (2010).

233. Botstein, D. & Risch, N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat Genet* **33 Suppl**, 228-37 (2003).

234. Li, Y.I. *et al.* RNA splicing is a primary link between genetic variation and disease. *Science* **352**, 600-4 (2016).

235. Reich, D.E. *et al.* Linkage disequilibrium in the human genome. *Nature* **411**, 199-204 (2001).

236. Sudmant, P.H. *et al.* An integrated map of structural variation in 2,504 human genomes. *Nature* **526**, 75-81 (2015).

237. Reich, D., Price, A.L. & Patterson, N. Principal component analysis of genetic data. *Nat Genet* **40**, 491-2 (2008).

238. Popejoy, A.B. & Fullerton, S.M. Genomics is failing on diversity. *Nature News* **538**, 161 (2016).

239. Sirugo, G., Williams, S.M. & Tishkoff, S.A. The Missing Diversity in Human Genetic Studies. *Cell* **177**, 26-31 (2019).

240. Stewart, C. & Pepper, M.S. Cystic fibrosis in the African diaspora. *Annals of the American Thoracic Society* **14**, 1-7 (2017).

241. Tishkoff, S.A. *et al.* The genetic structure and history of Africans and African Americans. *science* **324**, 1035-1044 (2009).

242. Kilpeläinen, T.O. *et al.* Multi-ancestry study of blood lipid levels identifies four loci interacting with physical activity. *Nature communications* **10**, 376 (2019).

243. Welter, D. *et al.* The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res* **42**, D1001-6 (2014).

244. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661-78 (2007).

245. Lee, M.K. *et al.* Genome-wide association study of facial morphology reveals novel associations with FREM1 and PARK2. **12**, e0176566 (2017).

246. Peng, Q. *et al.* EDARV370A associated facial characteristics in Uyghur population revealing further pleiotropic effects. *Hum Genet* **135**, 99-108 (2016).

247. He, H. *et al.* Correlation between facial morphology and gene polymorphisms in the Uygur youth population. *Oncotarget* **8**, 28750-28757 (2017).

248. Botstein, D. & Risch, N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nature genetics* **33**, 228 (2003).

249. Marian, A.J. Molecular genetic studies of complex phenotypes. *Translational Research* **159**, 64-79 (2012).

250. Teo, Y.Y. *et al.* A genotype calling algorithm for the Illumina BeadArray platform. *Bioinformatics* **23**, 2741-2746 (2007).

251. Chang, C.C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).

252. Anderson, C.A. *et al.* Data quality control in genetic case-control association studies. *Nature protocols* **5**, 1564 (2010).

253. Patterson, N., Price, A.L. & Reich, D. Population structure and eigenanalysis. *PLoS genetics* **2**, e190 (2006).

254. Adams, D.C., Rohlf, F.J. & Slice, D.E. A field comes of age: geometric morphometrics in the 21st century. *Hystrix* **24**, 7 (2013).

255. Zelditch, M.L., Swiderski, D.L. & Sheets, H.D. *Geometric morphometrics for biologists: a primer*, (Academic Press, 2012).

256. Adams, D.C., Rohlf, F.J. & Slice, D.E. Geometric morphometrics: ten years of progress following the 'revolution'. *Italian Journal of Zoology* **71**, 5-16 (2004).

257. Remagnino, P., Mayo, S., Wilkin, P., Cope, J. & Kirkup, D. *Computational Botany: Methods for Automated Species Identification*, (Springer, 2016).

258. Bookstein, F.L. Biometrics, biomathematics and the morphometric synthesis. *Bulletin of mathematical biology* **58**, 313-365 (1996).

259. Rohlf, F.J. & Marcus, L.F. A revolution morphometrics. *Trends in Ecology & Evolution* **8**, 129-132 (1993).

260. Bookstein, F.L. Foundations of morphometrics. *Annual Review of Ecology and Systematics* **13**, 451-470 (1982).

261. Kendall, D.G. The diffusion of shape. *Advances in applied probability* **9**, 428-430 (1977).

262. Bookstein, F.L. Size and shape spaces for landmark data in two dimensions. *Statistical science*, 181-222 (1986).

263. Wagner, G.P. The biological homology concept. *Annual Review of Ecology and Systematics* **20**, 51-69 (1989).

264. Klingenberg, C.P. & Monteiro, L.R. Distances and directions in multidimensional shape spaces: implications for morphometric applications. *Systematic Biology* **54**, 678-688 (2005).

265. Bookstein, F.L. *Morphometric tools for landmark data: geometry and biology*, (Cambridge University Press, 1997).

266. Rohlf, F.J. Morphometrics. *Annual Review of ecology and Systematics* **21**, 299-316 (1990).

267. Klingenberg, C.P. Evolution and development of shape: integrating quantitative approaches. *Nature reviews. Genetics* **11**, 623 (2010).

268. Mitteroecker, P., Gunz, P., Windhager, S. & Schaefer, K. A brief review of shape, form, and allometry in geometric morphometrics, with applications to human facial morphology. *Hystrix, the Italian Journal of Mammalogy* **24**, 59-66 (2013).

269. Klingenberg, C.P. Cranial integration and modularity: insights into evolution and development from morphometric data. *Hystrix, the Italian Journal of Mammalogy* **24**, 43-58 (2013).

270. Slice, D.E. Geometric morphometrics. *Annu. Rev. Anthropol.* **36**, 261-281 (2007).

271. Mitteroecker, P., Windhager, S., Muller, G.B. & Schaefer, K. The morphometrics of "masculinity" in human faces. *PLoS One* **10**, e0118374 (2015).

272. Sanchez-Pages, S., Rodriguez-Ruiz, C. & Turiegano, E. Facial masculinity: how the choice of measurement method enables to detect its influence on behaviour. *PLoS One* **9**, e112157 (2014).

273. Windhager, S., Schaefer, K. & Fink, B. Geometric morphometrics of male facial shape in relation to physical strength and perceived attractiveness, dominance, and masculinity. *Am J Hum Biol* **23**, 805-14 (2011).

274. Quinto-Sanchez, M. *et al.* Facial asymmetry and genetic ancestry in Latin American admixed populations. *Am J Phys Anthropol* **157**, 58-70 (2015).

275. Claes, P. *et al.* Genome-wide mapping of global-to-local genetic effects on human facial shape. *Nature genetics* **50**, 414 (2018).

276. Lee, M.K. *et al.* Genome-wide association study of facial morphology reveals novel associations with FREM1 and PARK2. *PLoS One* **12**, e0176566 (2017).

277. Pickrell, J.K. *et al.* Detection and interpretation of shared genetic influences on 42 human traits. *Nature genetics* **48**, 709 (2016).

278. Shaffer, J.R. *et al.* Genome-wide association study reveals multiple loci influencing normal human facial morphology. *PLoS genetics* **12**, e1006149 (2016).

279. Adhikari, K. *et al.* A genome-wide association scan implicates DCHS2, RUNX2, GLI3, PAX1 and EDAR in human facial variation. *Nature communications* **7**(2016).

280. Qiao, L. *et al.* Genome-wide Variants of Eurasian Facial Shape Differentiation and a prospective model of DNA based Face Prediction. *Journal of Genetics and Genomics* **45**, 419-432 (2018).

281. Cha, S. *et al.* Identification of five novel genetic loci related to facial morphology by genome-wide association studies. *BMC genomics* **19**, 481 (2018).

282. Cole, J.B. *et al.* Genomewide association study of African children identifies association of SCHIP1 and PDE8A with facial size and shape. *PLoS genetics* **12**, e1006174 (2016).

283. Wu, W. *et al.* Whole-exome sequencing identified four loci influencing craniofacial morphology in northern Han Chinese. *Human genetics*, 1-11 (2019).

284. Kimura, R. *et al.* A common variation in EDAR is a genetic determinant of shovel-shaped incisors. *The American Journal of Human Genetics* **85**, 528-535 (2009).

285. Wu, S. *et al.* Genome-wide scans reveal variants at EDAR predominantly affecting hair straightness in Han Chinese and Uyghur populations. *Human genetics* **135**, 1279-1286 (2016).

286. Li, Y. *et al.* The effect of EDARV370A on facial and ear morphologies in Uyghur population. *Yi chuan= Hereditas* **40**, 1024-1032 (2018).

287. Peng, Q. *et al.* EDARV370A associated facial characteristics in Uyghur population revealing further pleiotropic effects. *Human genetics* **135**, 99-108 (2016).

288. Li, Y. *et al.* EDAR, LYPLAL1, PRDM16, PAX3, DKK1, TNFSF12, CACNA2D3, and SUPT3H gene variants influence facial morphology in a Eurasian population. *Human genetics*, 1-9 (2019).

289. Ritz-Timme, S. *et al.* A new atlas for the evaluation of facial features: advantages, limits, and applicability. *International journal of legal medicine* **125**, 301-306 (2011).

290. MATLAB, M. & Release, N.N.T. Natick. *Massachusetts: The MathWorks Inc* (2013).

291. Shrout, P.E. & Fleiss, J.L. Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin* **86**, 420 (1979).

292. Klingenberg, C.P. MorphoJ: an integrated software package for geometric morphometrics. *Molecular ecology resources* **11**, 353-357 (2011).

293. Consortium, G.P. A global reference for human genetic variation. *Nature* **526**, 68 (2015).

294. O'Connell, J. *et al.* A general approach for haplotype phasing across the full spectrum of relatedness. *PLoS genetics* **10**, e1004234 (2014).

295. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G.R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nature genetics* **44**, 955 (2012).

296. Alexander, D.H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome research* **19**, 1655-1664 (2009).

297. Consortium, I.H. The international HapMap project. *Nature* **426**, 789 (2003).

298. Marchini, J. & Howie, B. Genotype imputation for genome-wide association studies. *Nature reviews. Genetics* **11**, 499 (2010).

299. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-2191 (2010).

300. Monreal, A.W. *et al.* Mutations in the human homologue of mouse dl cause autosomal recessive and dominant hypohidrotic ectodermal dysplasia. *Nature genetics* **22**, 366-369 (1999).

301. Mou, C. *et al.* Enhanced ectodysplasin-A receptor (EDAR) signaling alters multiple fiber characteristics to produce the East Asian hair form. *Human mutation* **29**, 1405-1411 (2008).

302. Rohlf, F. tpsDig, Digitize Landmarks and Outlines, version 2.05 (software). *Stony Brook, NY: Department of Ecology and Evolution, State University of New York at Stony Brook* (2006).

303. Rohlf, F.J. Morphometrics at SUNY Stony Brook. Vol. 2016 (2016).

304. Carson, E. Maximum likelihood estimation of human craniometric heritabilities. *American Journal of Physical Anthropology* **131**, 169-180 (2006).

305. Van der Laan, M.J. *Multiple Testing Procedures with Applications to Genomics. Springer Series in Statistics*, (Springer, 2008).

306. O'Reilly, P.F. *et al.* MultiPhen: joint model of multiple phenotypes can increase discovery in GWAS. *PloS one* **7**, e34861 (2012).

307. Efron, B. Linear Statistical Inference and its Applications. (JSTOR, 1967).

308. Zhou, X. & Stephens, M. Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nature methods* **11**, 407-409 (2014).

309. Stephens, M. A unified framework for association analysis with multiple related phenotypes. *PloS one* **8**, e65245 (2013).

310. Pruim, R.J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336-2337 (2010).

311. Boehringer, S. *et al.* Genetic determination of human facial morphology: links between cleft-lips and normal variation. *European Journal of Human Genetics* **19**, 1192 (2011).

312. Beaty, T. *et al.* Confirming genes influencing risk to cleft lip with/without cleft palate in a case–parent trio study. *Human genetics* **132**, 771-781 (2013).

313. Beaty, T.H. *et al.* A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. *Nature genetics* **42**, 525-529 (2010).

314. Birnbaum, S. *et al.* Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. *Nature genetics* **41**, 473-477 (2009).

315. Grant, S.F. *et al.* A genome-wide association study identifies a locus for nonsyndromic cleft lip with or without cleft palate on 8q24. *The Journal of pediatrics* **155**, 909-913 (2009).

316. Ludwig, K.U. *et al.* Genome-wide meta-analyses of nonsyndromic cleft lip with or without cleft palate identify six new risk loci. *Nature genetics* **44**, 968-971 (2012).

317. Mangold, E. *et al.* Genome-wide association study identifies two susceptibility loci for nonsyndromic cleft lip with or without cleft palate. *Nature genetics* **42**, 24-26 (2010).

318. Rahimov, F. *et al.* Disruption of an AP-2α binding site in an IRF6 enhancer is associated with cleft lip. *Nature genetics* **40**, 1341-1347 (2008).

319. Sun, Y. *et al.* Genome-wide association study identifies a new susceptibility locus for cleft lip with or without a cleft palate. *Nature communications* **6**, 6414 (2015).

320. Zucchero, T.M. *et al.* Interferon regulatory factor 6 (IRF6) gene variants and the risk of isolated cleft lip or palate. *New England Journal of Medicine* **351**, 769-780 (2004).

321. Sadier, A., Viriot, L., Pantalacci, S. & Laudet, V. The ectodysplasin pathway: from diseases to adaptations. *Trends Genet* **30**, 24-31 (2014).

322. Kamberov, Y.G. *et al.* Modeling recent human evolution in mice by expression of a selected EDAR variant. *Cell* **152**, 691-702 (2013).

323. Goodwin, A.F. *et al.* Craniofacial morphometric analysis of individuals with X-linked hypohidrotic ectodermal dysplasia. *Mol Genet Genomic Med* **2**, 422-9 (2014).

324. Mikkola, M.L. Molecular aspects of hypohidrotic ectodermal dysplasia. *Am J Med Genet A* **149a**, 2031-6 (2009).

325. Headon, D.J. *et al.* Gene defect in ectodermal dysplasia implicates a death domain adapter in development. *Nature* **414**, 913-6 (2001).

326. Clauss, F. *et al.* Dento-craniofacial phenotypes and underlying molecular mechanisms in hypohidrotic ectodermal dysplasia (HED): a review. *J Dent Res* **87**, 1089-99 (2008).

327. Lesot, H., Clauss, F., Maniere, M.C. & Schmittbuhl, M. Consequences of X-linked hypohidrotic ectodermal dysplasia for the human jaw bone. *Front Oral Biol* **13**, 93-9 (2009).

328. Cluzeau, C. *et al.* The EDAR370A allele attenuates the severity of hypohidrotic ectodermal dysplasia caused by EDA gene mutation. *Br J Dermatol* **166**, 678-81 (2012).

329. Mou, C. *et al.* Enhanced ectodysplasin-A receptor (EDAR) signaling alters multiple fiber characteristics to produce the East Asian hair form. *Hum Mutat* **29**, 1405-11 (2008).

330. Sabeti, P.C. *et al.* Genome-wide detection and characterization of positive selection in human populations. *Nature* **449**, 913-8 (2007).

331. Barrett, J.C., Fry, B., Maller, J. & Daly, M.J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263-265 (2004).

332. Consortium, E.P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57 (2012).

333. Romanoski, C.E., Glass, C.K., Stunnenberg, H.G., Wilson, L. & Almouzni, G. Epigenomics: Roadmap for regulation. *Nature* **518**, 314-6 (2015).

334. Rojas, A. *et al.* Epigenetic control of the bone-master Runx2 gene during osteoblast-lineage commitment by the histone demethylase JARID1B/KDM5B. *Journal of biological chemistry* **290**, 28329-28342 (2015).

335. Drissi, H. *et al.* Transcriptional autoregulation of the bone related CBFA 1/RUNX 2 gene. *Journal of cellular physiology* **184**, 341-350 (2000).

336. Zambotti, A., Makhluf, H., Shen, J. & Ducy, P. Characterization of an osteoblast-specific enhancer element in the CBFA1 gene. *Journal of Biological Chemistry* **277**, 41497-41506 (2002).

337. Lee, M.-H. *et al.* Dlx5 specifically regulates Runx2 type II expression by binding to homeodomain-response elements in the Runx2 distal promoter. *Journal of Biological Chemistry* **280**, 35579-35587 (2005).

338. Napierala, D. *et al.* Mutations and promoter SNPs in RUNX2, a transcriptional regulator of bone formation. *Molecular genetics and metabolism* **86**, 257-268 (2005).

339. Tai, P.W. *et al.* Epigenetic landscape during osteoblastogenesis defines a differentiation-dependent Runx2 promoter region. *Gene* **550**, 1-9 (2014).

340. Naruse, I., Ueta, E., Sumino, Y., Ogawa, M. & Ishikiriyama, S. Birth defects caused by mutations in human GLI3 and mouse Gli3 genes. *Congenit Anom (Kyoto)* **50**, 1-7 (2010).

341. Rice, D.P. *et al.* Gli3 Xt− J/Xt− J mice exhibit lambdoid suture craniosynostosis which results from altered osteoprogenitor proliferation and differentiation. *Human molecular genetics* **19**, 3457-3467 (2010).

342. Buttitta, L., Mo, R., Hui, C.-C. & Fan, C.-M. Interplays of Gli2 and Gli3 and their requirement in mediating Shh-dependent sclerotome induction. *Development* **130**, 6233-6243 (2003).

343. Le Pabic, P., Ng, C. & Schilling, T.F. Fat-Dachsous signaling coordinates cartilage differentiation and polarity during craniofacial development. *PLoS genetics* **10**, e1004726 (2014).

344. Nakajima, H. *et al.* Wnt modulators, SFRP-1, and SFRP-2 are expressed in osteoblasts and differentially regulate hematopoietic stem cells. *Biochemical and biophysical research communications* **390**, 65-70 (2009).

345. Kurosaka, H., Iulianella, A., Williams, T. & Trainor, P.A. Disrupting hedgehog and WNT signaling interactions promotes cleft lip pathogenesis. *The Journal of clinical investigation* **124**, 1660 (2014).

346. Fujita, T. *et al.* Runx2 induces osteoblast and chondrocyte differentiation and enhances their migration by coupling with PI3K-Akt signaling. *J Cell Biol* **166**, 85-95 (2004).

347. Yoshida, C.A. *et al.* Runx2 and Runx3 are essential for chondrocyte maturation, and Runx2 regulates limb growth through induction of Indian hedgehog. *Genes & development* **18**, 952-963 (2004).

348. Sears, K., Goswami, A., Flynn, J. & Niswander, L. The correlated evolution of Runx2 tandem repeats, transcriptional activity, and facial length in carnivora. *Evolution & development* **9**, 555-565 (2007).

349. Visel, A. *et al.* ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* **457**, 854-8 (2009).

350. Pan, A., Chang, L., Nguyen, A. & James, A.W. A review of hedgehog signaling in cranial bone development. *Frontiers in physiology* **4**(2013).

351. Chong, H.J. *et al.* Signaling by SHH rescues facial defects following blockade in the brain. *Developmental Dynamics* **241**, 247-256 (2012).

352. Marcucio, R.S., Cordero, D.R., Hu, D. & Helms, J.A. Molecular interactions coordinating the development of the forebrain and face. *Developmental biology* **284**, 48-61 (2005).

353. Kundaje, A. *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317 (2015).

354. Vortkamp, A., Gessler, M. & Grzeschik, K.-H. GLI3 zinc-finger gene interrupted by translocations in Greig syndrome families. (2008).

355. Hui, C.-c. & Joyner, A.L. A mouse model of Greig cephalo–polysyndactyly syndrome: the extra–toesJ mutation contains an intragenic deletion of the Gli3 gene. *Nature genetics* **3**, 241-246 (1993).

356. Maynard, T.M., Jain, M.D., Balmer, C.W. & LaMantia, A.-S. High-resolution mapping of the Gli3 mutation Extra-toes J reveals a 51.5-kb deletion. *Mammalian genome* **13**, 58-61 (2002).

357. Veistinen, L. *et al.* Loss-of-Function of Gli3 in Mice Causes Abnormal Frontal Bone Morphology and Premature Synostosis of the Interfrontal Suture. *Front Physiol* **3**, 121 (2012).

358. Hurst, J.A. *et al.* Metopic and sagittal synostosis in Greig cephalopolysyndactyly syndrome: five cases with intragenic mutations or complete deletions of GLI3. *European journal of human genetics* **19**, 757 (2011).

359. Attanasio, C. *et al.* Fine tuning of craniofacial morphology by distant-acting enhancers. *Science* **342**, 1241006 (2013).

360. Takimoto, A., Mohri, H., Kokubu, C., Hiraki, Y. & Shukunami, C. Pax1 acts as a negative regulator of chondrocyte maturation. *Experimental cell research* **319**, 3128-3139 (2013).

361. Wagner, G.P. The developmental genetics of homology. *Nature reviews. Genetics* **8**, 473 (2007).

362. Pohl, E. *et al.* A hypofunctional PAX1 mutation causes autosomal recessively inherited otofaciocervical syndrome. *Human genetics* **132**, 1311-1320 (2013).

363. Rosenberg, N.A. *et al.* Genome-wide association studies in diverse populations. *Nature Reviews Genetics* **11**, 356 (2010).

364. Bustamante, C.D. & Burchard, E.G. De la Vega FM. Genomics for the world. *Nature* **475**, 163-5 (2011).

365. Asimit, J.L., Hatzikotoulas, K., McCarthy, M., Morris, A.P. & Zeggini, E. Trans-ethnic study design approaches for fine-mapping. *European journal of human genetics* **24**, 1330 (2016).

366. Maples, B.K., Gravel, S., Kenny, E.E. & Bustamante, C.D. RFMix: a discriminative modeling approach for rapid and robust local-ancestry inference. *The American Journal of Human Genetics* **93**, 278-288 (2013).

367. Fatemifar, G. *et al.* Genome-wide association study of primary tooth eruption identifies pleiotropic loci associated with height and craniofacial distances. *Human Molecular Genetics* **22**, 3807-3817 (2013).

368. Morrison, J. *et al.* Genome-wide association study of dental caries in the Hispanic Communities Health Study/Study of Latinos (HCHS/SOL). *Human molecular genetics* **25**, 807-816 (2015).

369. Shaffer, J. *et al.* GWAS of dental caries patterns in the permanent dentition. *Journal of dental research* **92**, 38-44 (2013).

370. Wang, X. *et al.* Genome-wide association scan of dental caries in the permanent dentition. *BMC oral health* **12**, 57 (2012).

371. Le Goff, C. *et al.* Regulation of procollagen amino-propeptide processing during mouse embryogenesis by specialization of homologous ADAMTS proteases: insights on collagen biosynthesis and dermatosparaxis. *Development* **133**, 1587-1596 (2006).

372. Turner, C.I. Scoring produces for key morphological traits of the permanent dentition: The Arizona State University dental anthropology system. *Advances in dental anthropology*, 13-31 (1991).

373. Hrdlička, A. Shovel-shaped teeth. *American Journal of Physical Anthropology* **3**, 429-465 (1920).

374. Scott, G.R. & Irish, J.D. *Human tooth crown and root morphology*, (Cambridge University Press, 2017).

375. Morris, D.H. Maxillary first premolar angular differences between North American Indians and non-North American Indians. *American journal of physical anthropology* **54**, 431-433 (1981).

376. Edgar, H.J. & Sciulli, P.W. Elongated mandibular premolar: a new morphological variant. *Dent Anthropol* **17**, 24-27 (2004).

377. Kieser, J.A. & Julius, K. *Human adult odontometrics: the study of variation in adult tooth size*, (Cambridge University Press, 1990).

378. Hanihara, T. & Ishida, H. Metric dental variation of major human populations. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists* **128**, 287-298 (2005).

379. Scott, G.R. & Turner, C.G. Dental anthropology. *Annual review of Anthropology* **17**, 99-126 (1988).

380. Harris, E.F. Ethnic differences in the apportionment of tooth size. *Advances in dental anthropology*, 121-142 (1991).

381. Chan, C.A. A review of the clinical significance of the occlusal plane: its variation and effect on head posture. *International College of Craniomandibular Orthopedics (ICCMO) Anthology* **8**, 1-63 (2007).

382. Scott, G.R. j L\. v—.. Population Variation of Carabelli's Trait. *Human biology* **52**, 63-78 (1980).

383. Stekhoven, D.J. Using the missForest Package. *R package*, 1-11 (2011).

384. Adeyemi, T. & Isiekwe, M. Comparing permanent tooth sizes (mesio-distal) of males and females in a Nigerian population. *West African journal of medicine* **22**, 219-221 (2003).

385. Lovejoy, C.O. Dental wear in the Libben population: its functional pattern and role in the determination of adult skeletal age at death. *American journal of physical anthropology* **68**, 47-56 (1985).

386. Everman, A. University of Pittsburgh (2018).

387. Yates, A. *et al.* Ensembl 2016. *Nucleic acids research* **44**, D710-D716 (2015).

388. Kurisaki, T. *et al.* Phenotypic analysis of Meltrin α (ADAM12)-deficient mice: involvement of Meltrin α in adipogenesis and myogenesis. *Molecular and cellular biology* **23**, 55-61 (2003).

389. Visel, A., Minovitsky, S., Dubchak, I. & Pennacchio, L.A. VISTA Enhancer Browser—a database of tissue-specific human enhancers. *Nucleic acids research* **35**, D88-D92 (2006).

390. Thesleff, I., Vaahtokari, A., Kettunen, P. & Aberg, T. Eplthelial-mesenchymal signaling during tooth development. *Connective tissue research* **32**, 9-15 (1995).

391. Michot, B., Lee, C.S. & Gibbs, J.L. TRPM8 and TRPA1 do not contribute to dental pulp sensitivity to cold. *Scientific reports* **8**, 13198 (2018).

392. Bullock, S.L. & Ish-Horowicz, D. Conserved signals and machinery for RNA transport in Drosophila oogenesis and embryogenesis. *Nature* **414**, 611 (2001).

393. Della-Morte, D. *et al.* A follow-up study for left ventricular mass on chromosome 12p11 identifies potential candidate genes. *BMC medical genetics* **12**, 100 (2011).

394. Szklarczyk, D. *et al.* STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic acids research* **43**, D447-D452 (2014).

395. Huttlin, E.L. *et al.* Architecture of the human interactome defines protein communities and disease networks. *Nature* **545**, 505 (2017).

396. Hart, T. & Hart, P. Genetic studies of craniofacial anomalies: clinical implications and applications. *Orthodontics & craniofacial research* **12**, 212-220 (2009).

397. Ruegg, C.L. *et al.* V7, a novel leukocyte surface protein that participates in T cell activation. II. Molecular cloning and characterization of the V7 gene. *The Journal of Immunology* **154**, 4434-4443 (1995).

398. Zeng, Z. *et al.* Genome-wide association study of primary dentition pit-and-fissure and smooth surface caries. *Caries research* **48**, 330-338 (2014).

399. Reardon, W. *et al.* Mutations in the fibroblast growth factor receptor 2 gene cause Crouzon syndrome. *Nature genetics* **8**, 98 (1994).

400. Glaser, R.L. *et al.* Paternal origin of FGFR2 mutations in sporadic cases of Crouzon syndrome and Pfeiffer syndrome. *The American Journal of Human Genetics* **66**, 768-777 (2000).

401. Matsuoka, R.L. *et al.* Transmembrane semaphorin signalling controls laminar stratification in the mammalian retina. *Nature* **470**, 259 (2011).

402. Shrestha, A. Semaphorin 3A and Class 4 Semaphorins in Tooth Innervation and Development. (2014).

403. Bohnekamp, J. & Schöneberg, T. Cell adhesion receptor GPR133 couples to Gs protein. *Journal of Biological Chemistry* **286**, 41912-41916 (2011).

404. Trovo-Marqui, A. & Tajara, E. Neurofibromin: a general outlook. *Clinical genetics* **70**, 1-13 (2006).

405. Wallace, M.R. *et al.* A de novo Alu insertion results in neurofibromatosis type 1. *Nature* **353**, 864 (1991).

406. Eisenbarth, I., Beyer, K., Krone, W. & Assum, G. Toward a survey of somatic mutation of the NF1 gene in benign neurofibromas of patients with neurofibromatosis type 1. *The American Journal of Human Genetics* **66**, 393-401 (2000).

407. Park, J.-H. *et al.* Effects of an Asian-specific nonsynonymous EDAR variant on multiple dental traits. *Journal of human genetics* **57**, 508 (2012).

408. Tan, J. *et al.* Characteristics of dental morphology in the Xinjiang Uyghurs and correlation with the EDARV370A variant. *Science China Life Sciences* **57**, 510-518 (2014).

409. Kimura, R. *et al.* Common polymorphisms in WNT10A affect tooth morphology as well as hair shape. *Human molecular genetics* **24**, 2673-2680 (2015).

410. Hlusko, L.J. *et al.* Environmental selection during the last ice age on the mother-to-infant transmission of vitamin D and fatty acids through breast milk. *Proceedings of the National Academy of Sciences* **115**, E4426-E4432 (2018).

411. Millar, J.K. *et al.* Disruption of two novel genes by a translocation co-segregating with schizophrenia. *Human molecular genetics* **9**, 1415-1423 (2000).

412. Ota, T. *et al.* Complete sequencing and characterization of 21,243 full-length human cDNAs. *Nature genetics* **36**, 40 (2004).

413. Dixon-Salazar, T.J. *et al.* Exome sequencing can improve diagnosis and alter patient management. *Science translational medicine* **4**, 138ra78-138ra78 (2012).

414. Carayon, D. *et al.* A geometric morphometric approach to the study of variation of shovel-shaped incisors. *American journal of physical anthropology* **168**, 229-241 (2019).

415. Alexander, D.H., Shringarpure, S.S., Novembre, J. & Lange, K. Admixture 1.3 Software Manual. (2015).

416. Martin, L.S. & Eskin, E. Population Structure in Genetic Studies: Confounding Factors and Mixed Models. *bioRxiv*, 092106 (2016).

417. Schmid, A.B. *et al.* Genetic components of human pain sensitivity: a protocol for a genome-wide association study of experimental pain in healthy volunteers. *BMJ open* **9**, e025530 (2019).

418. MacRae, C.A. & Vasan, R.S. Next-generation genome-wide association studies: time to focus on phenotype? (Am Heart Assoc, 2011).

419. Sero, D. *et al.* Facial recognition from DNA using face-to-DNA classifiers. *Nature communications* **10**, 2557 (2019).

420. Tam, V. *et al.* Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics*, 1 (2019).

421. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285 (2016).

422. Steinthorsdottir, V. *et al.* Identification of low-frequency and rare sequence variants associated with elevated or reduced risk of type 2 diabetes. *Nature genetics* **46**, 294 (2014).

423. Salinas, C.F., Jorgenson, R.J., Wright, J.T., DiGiovanna, J.J. & Fete, M.D. 2008 international conference on ectodermal dysplasias classification:

Conference report. *American Journal of Medical Genetics Part A* **149**, 1958-1969 (2009).

424. Visinoni, A.F., Lisboa-Costa, T., Pagnan, N.A. & Chautard-Freire-Maia, E.A. Ectodermal dysplasias: clinical and molecular review. *American Journal of Medical Genetics Part A* **149**, 1980-2002 (2009).

425. Reyes-Reali, J. *et al.* Hypohidrotic ectodermal dysplasia: clinical and molecular review. *International journal of dermatology* **57**, 965-972 (2018).

426. Trzeciak, W.H. & Koczorowski, R. Molecular basis of hypohidrotic ectodermal dysplasia: an update. *Journal of applied genetics* **57**, 51-61 (2016).

427. Pozo-Molina, G. *et al.* Novel missense mutation in the EDA 1 gene identified in a family with hypohidrotic ectodermal dysplasia. *International journal of dermatology* **54**, 790-794 (2015).

428. Zhang, S. *et al.* Dose-dependent effects of Runx2 on bone development. *Journal of Bone and Mineral Research* **24**, 1889-1904 (2009).

429. Molin, A. *et al.* Patients with isolated oligo/hypodontia caused by RUNX2 duplication. *American journal of medical genetics Part A* **167**, 1386-1390 (2015).

430. Kawane, T. *et al.* Runx2 is required for the proliferation of osteoblast progenitors and induces proliferation by regulating Fgfr2 and Fgfr3. *Scientific reports* **8**, 13551 (2018).

431. Villarreal Becerra, E.N. *et al.* General and oral aspects in Apert syndrome: report of a case. *Clinical Medical Reviews and Case Reports, 2015, vol. 2, num. 3, p. 1-5* (2015).

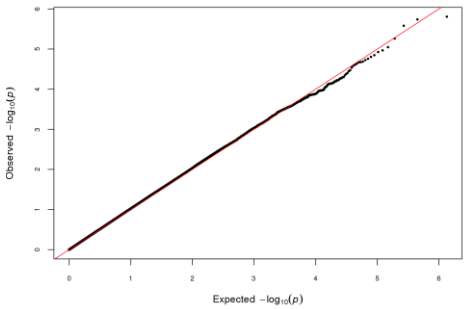
432. Du, W., Du, W. & Yu, H. The role of fibroblast growth factors in tooth development and incisor renewal. *Stem Cells International* **2018**(2018).

Appendices

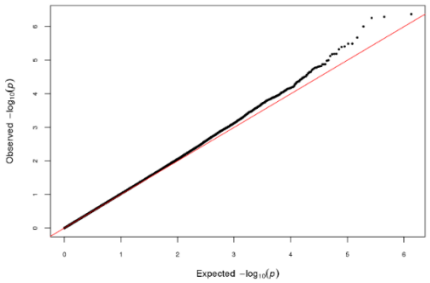
Appendix A to Chapter 3: Genome-Wide Association Studies of Human Facial traits

Figures A.1-A.13. Q-Q plots for genome-wide association tests of the facial features studied in the CANDELA sample.

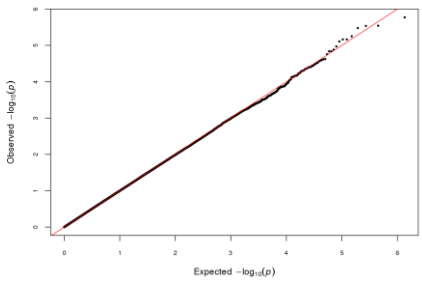
A.1. Forehead Profile



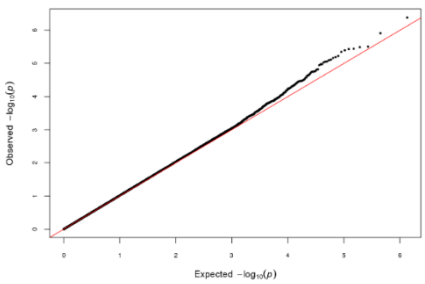
A.2. Brow Ridge Protrusion



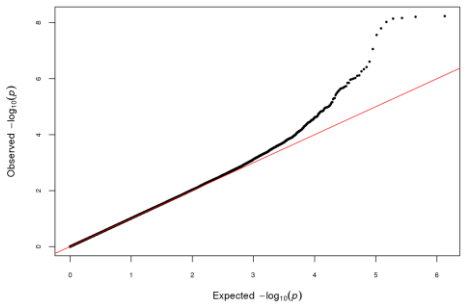
A.3. Cheekbone Protrusion



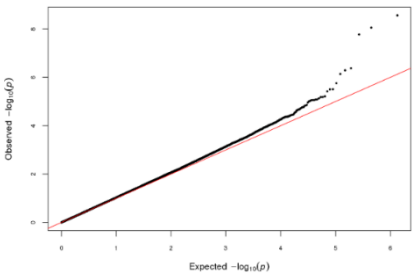
A.4. Nasal Root Breadth



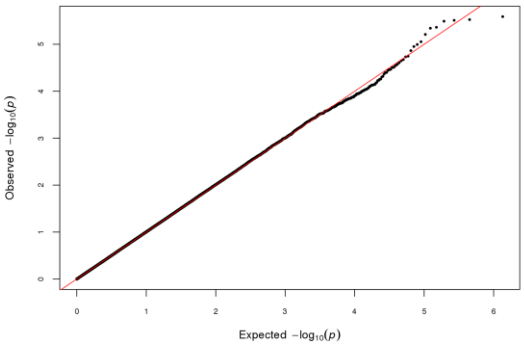
A.5. Nose Bridge Breadth



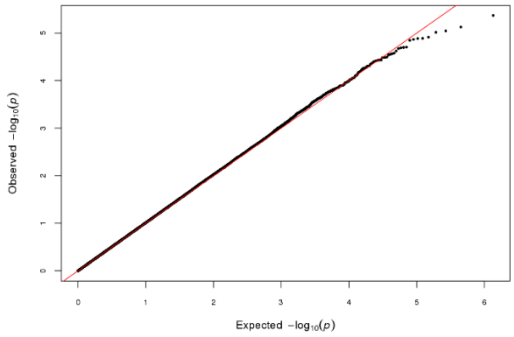
A.6. Nose Wing Breadth



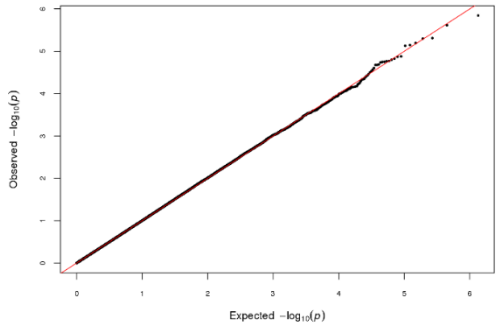
A.7. Nose Protrusion



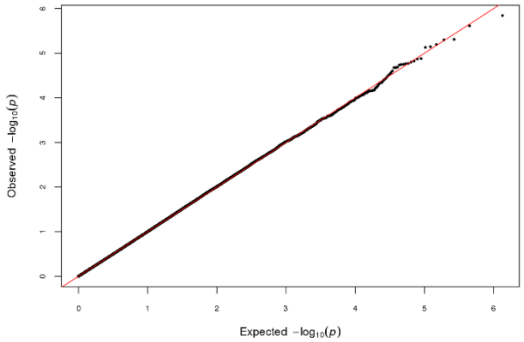
A.8. Nose Tip Shape



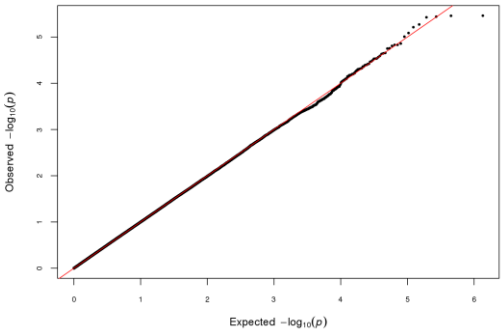
A.9. Columella Inclination



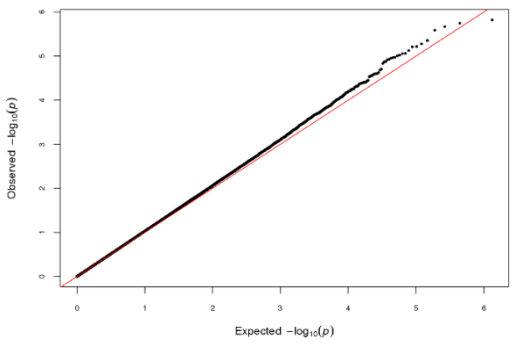
A.10. Upper Lip Thickness



A.11. Lower Lip Thickness



A.12. Chin Shape



A.13. Chin Protrusion

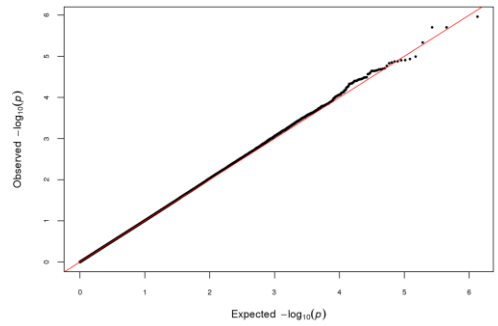


Table A.1 Genomic inflation factor for categorical traits:

Trait	Genomic Inflation Factor λ
Forehead profile	1.012
Brow ridge protrusion	1.002
Cheekbone protrusion	1.014
Nasal root breadth	1.013
Nose bridge breadth	1.008
Nose wing breadth	1.009
Nose profile	1.007
Nose protrusion	1.014
Nose tip shape	1.004
Columella inclination	1.015
Upper lip thickness	1.001
Lower lip thickness	1.002
Chin shape	1.016
Chin protrusion	1.014

Tables A.2a and A.2b. Fraction of trait variance explained by the index SNPs associated to ordinal and quantitative facial features. For all traits we calculated the fraction of trait variance explained by the covariates and by all index SNPs. A similar regression model to that used in the GWAS was used to estimate the R^2 of the model (representing the fraction of the variance of the trait explained by all regressors). The following models were applied:Model 1: Trait ~ age + sex + BMI + PC1 … PC5

Model 2: Trait ~ age + sex + BMI + PC1 … PC5 + index SNP

Model 3: Trait ~ age + sex + BMI + PC1 … PC5 + all index SNPs

Model 2 was applied separately for each index SNP.

From models 2 and 3, the fraction of trait variance explained by the covariates (i.e. Model 1) was subtracted to get the additional contribution of the SNP(s).

SNPs showing genome-wide significant associations to a trait are highlighted in bold.

Table A.2a. Ordinal traits

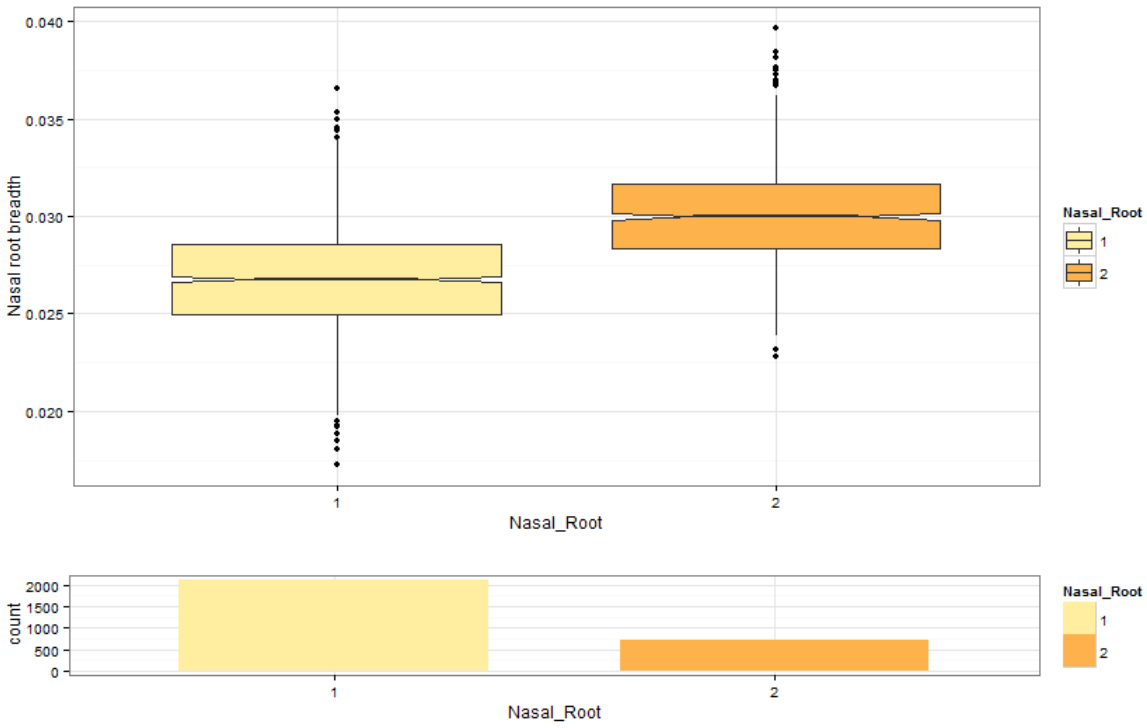
	R^2 explained by (%)								
Trait	Covariates	rs3827760	rs7559271	rs2045323	rs12644248	rs1852985	rs17640804	rs927833	All SNPs
Forehead profile	23.60	0.00	0.01	0.14	0.12	0.02	0.05	0.04	0.29
Brow ridge protrusion	39.81	0.01	0.02	0.02	0.22	0.24	0.00	0.00	0.46
Cheekbone protrusion	8.75	0.03	0.00	0.01	0.01	0.02	0.08	0.02	0.09
Nasal root breadth	5.84	0.07	0.17	0.02	0.02	0.21	0.02	0.00	0.45
Nose bridge breadth	7.08	0.01	0.01	0.00	0.09	**0.71**	0.06	0.01	0.92
Nose wing breadth	9.61	0.00	0.08	0.01	0.14	0.01	**0.62**	**0.66**	1.53
Nose profile	4.77	0.00	0.01	0.14	0.11	0.02	0.04	0.00	0.26
Nose protrusion	1.93	0.00	0.10	0.01	0.06	0.03	0.00	0.05	0.27
Nose tip shape	11.57	0.03	0.02	0.01	0.00	0.01	0.01	0.03	0.11
Columella inclination	5.29	0.00	0.07	0.44	**0.49**	0.09	0.03	0.00	0.86
Upper lip thickness	12.09	0.02	0.14	0.09	0.22	0.01	0.01	0.03	0.51
Lower lip thickness	9.61	0.08	0.09	0.01	0.19	0.04	0.00	0.01	0.41
Chin shape	2.77	0.00	0.03	0.03	0.09	0.02	0.01	0.03	0.14
Chin protrusion	4.83	0.15	0.01	0.03	0.03	0.02	0.03	0.09	0.36

Table A.2.b. Quantitative traits

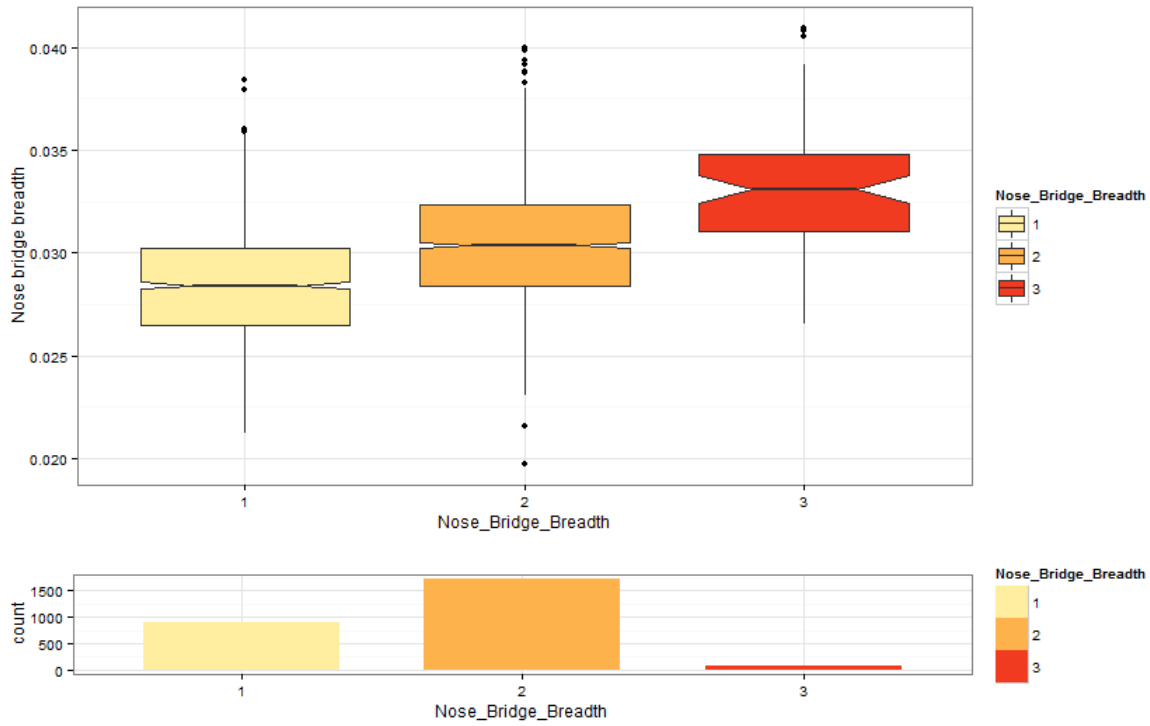
	R^2 explained by (%)								
Trait	Covariates	rs3827760	rs7559271	rs2045323	rs12644248	rs1852985	rs17640804	rs927833	All SNPs
Nasal root breadth	19.31	0.16	0.03	0.14	0.39	0.17	0.01	0.12	0.94
Nose bridge breadth	13.84	0.00	0.10	0.09	0.15	**1.18**	0.26	0.24	1.80
Nose wing breadth	15.27	0.03	0.13	0.05	0.17	0.10	**1.15**	**0.57**	2.28
Nose protrusion	17.72	0.00	0.28	**0.95**	0.29	0.01	0.01	0.07	1.32
Nose tip angle	10.15	0.01	0.07	**1.08**	0.52	0.05	0.02	0.02	1.37
Columella inclination	6.06	0.01	0.14	**0.63**	**0.44**	0.13	0.00	0.01	1.13
Upper lip thickness	16.75	0.09	0.10	0.09	0.39	0.08	0.02	0.00	0.32
Lower lip thickness	13.66	0.22	0.06	0.18	0.22	0.07	0.00	0.07	0.35
Chin protrusion	11.35	**1.32**	0.11	0.12	0.07	0.09	0.03	0.02	1.09
Nasion position	30.21	0.01	**1.33**	0.22	0.15	0.09	0.01	0.04	1.15

Figures A.14 - A.21. Boxplots showing the correlation between ordinal and quantitative facial features.

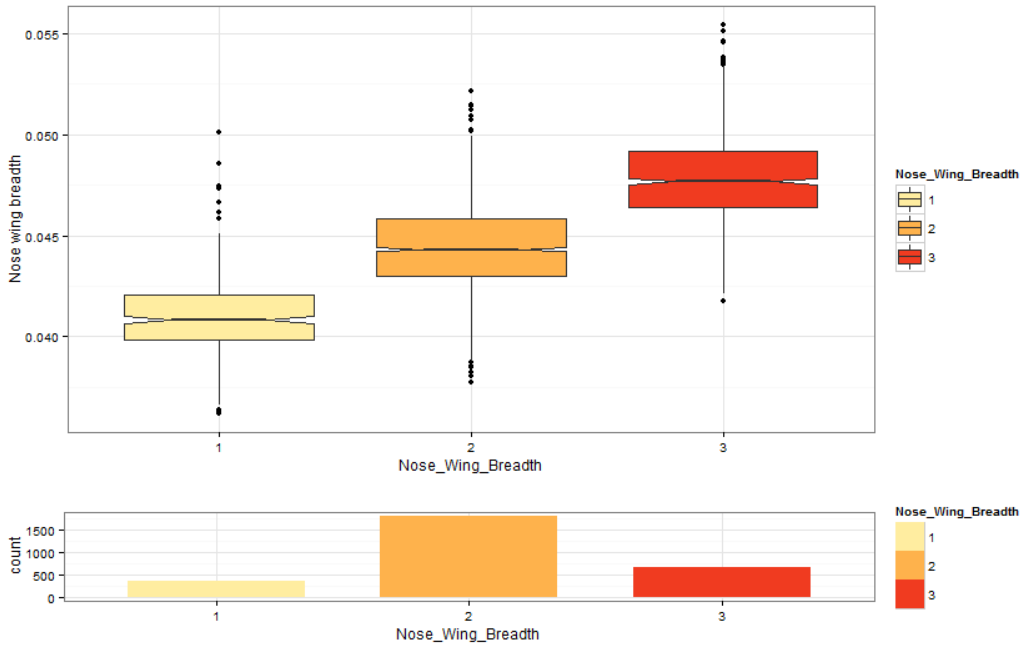
A.14. Nasal Root



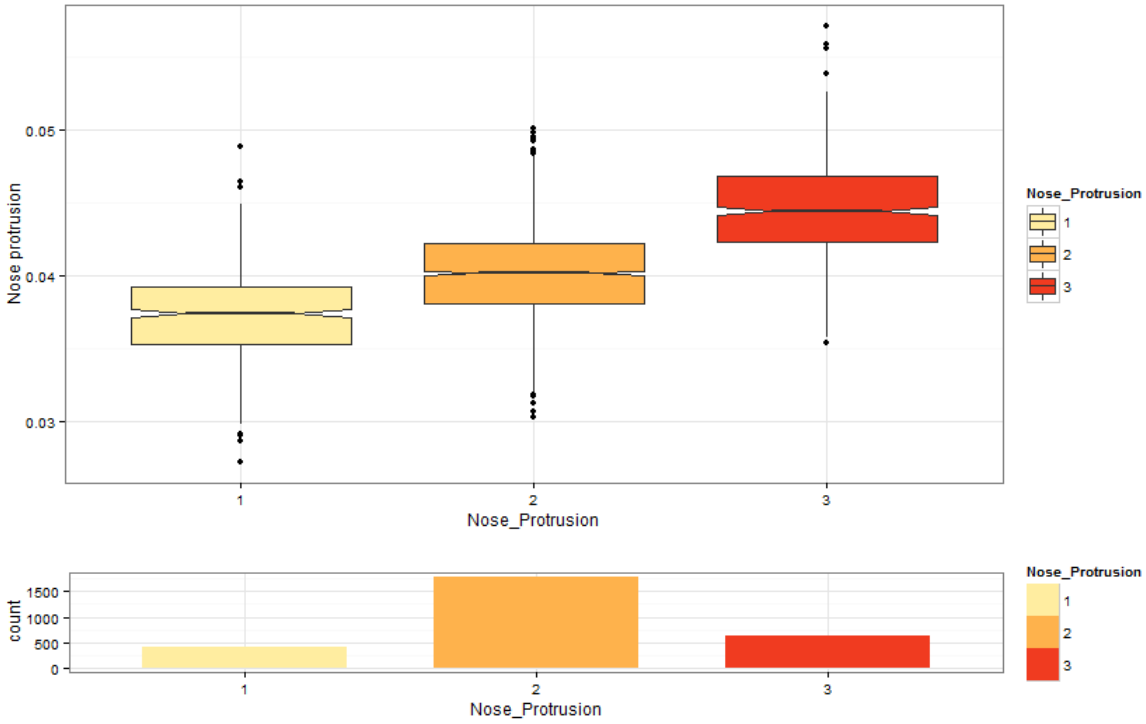
A.15. Nose Bridge Breadth



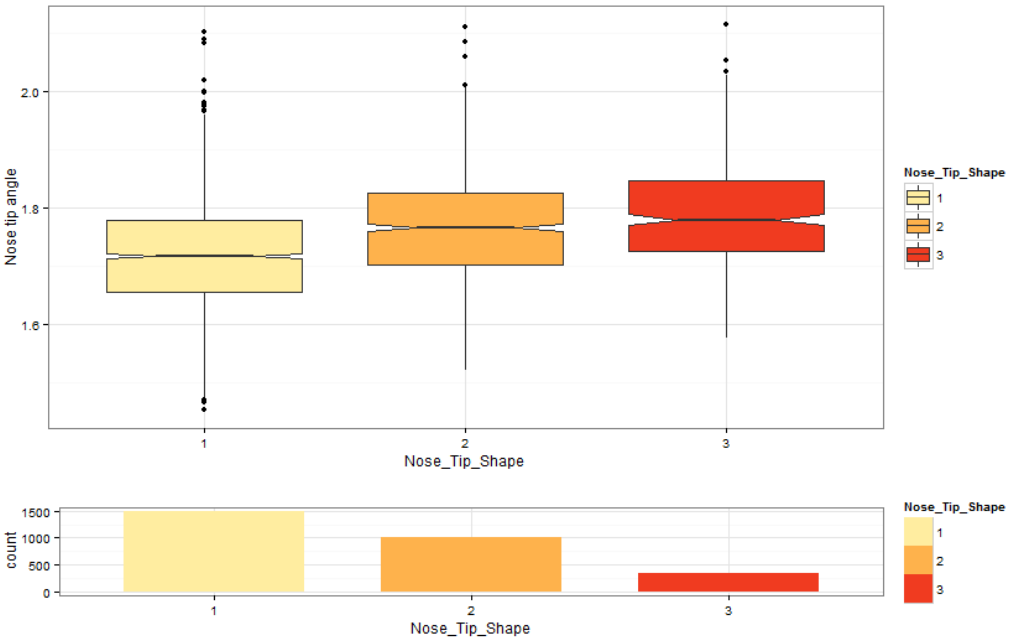
A.16. Nose Wing Breadth



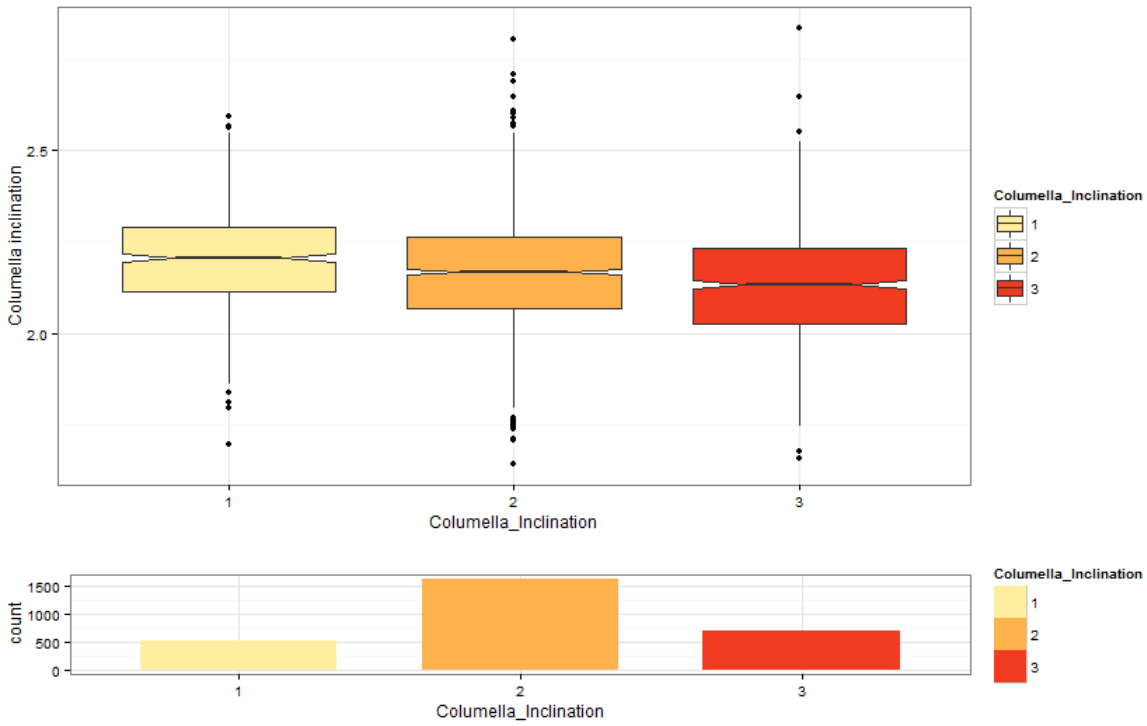
A.17. Nose Protrusion



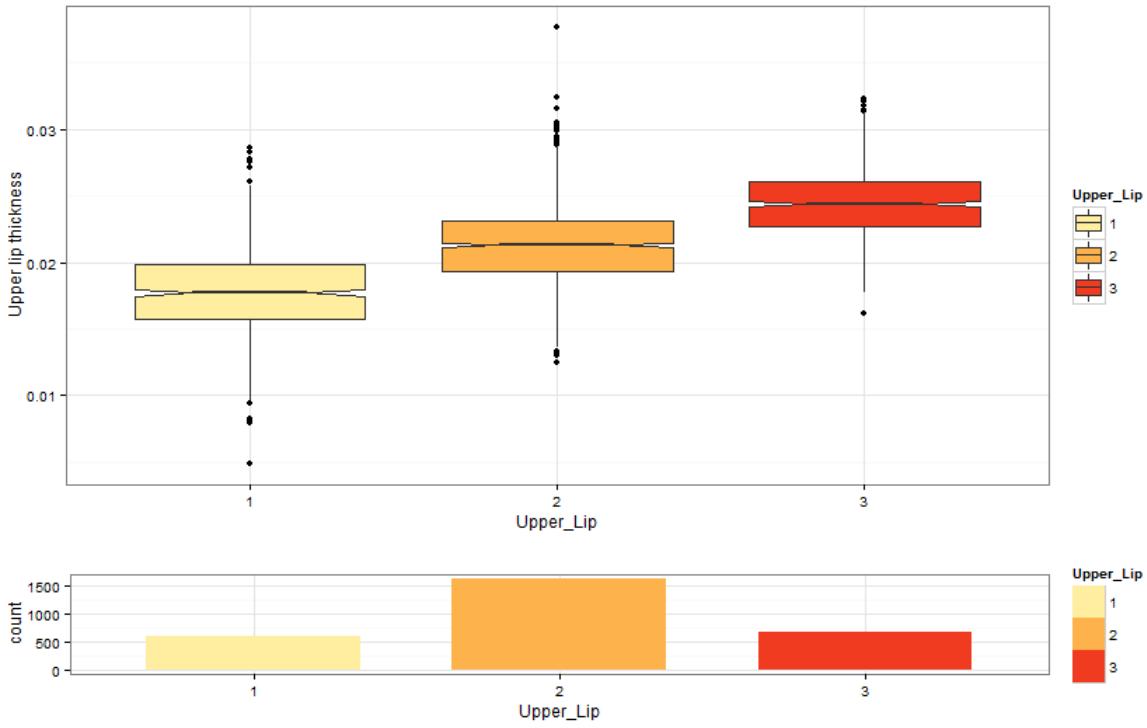
A.18. Nose Tip Angle



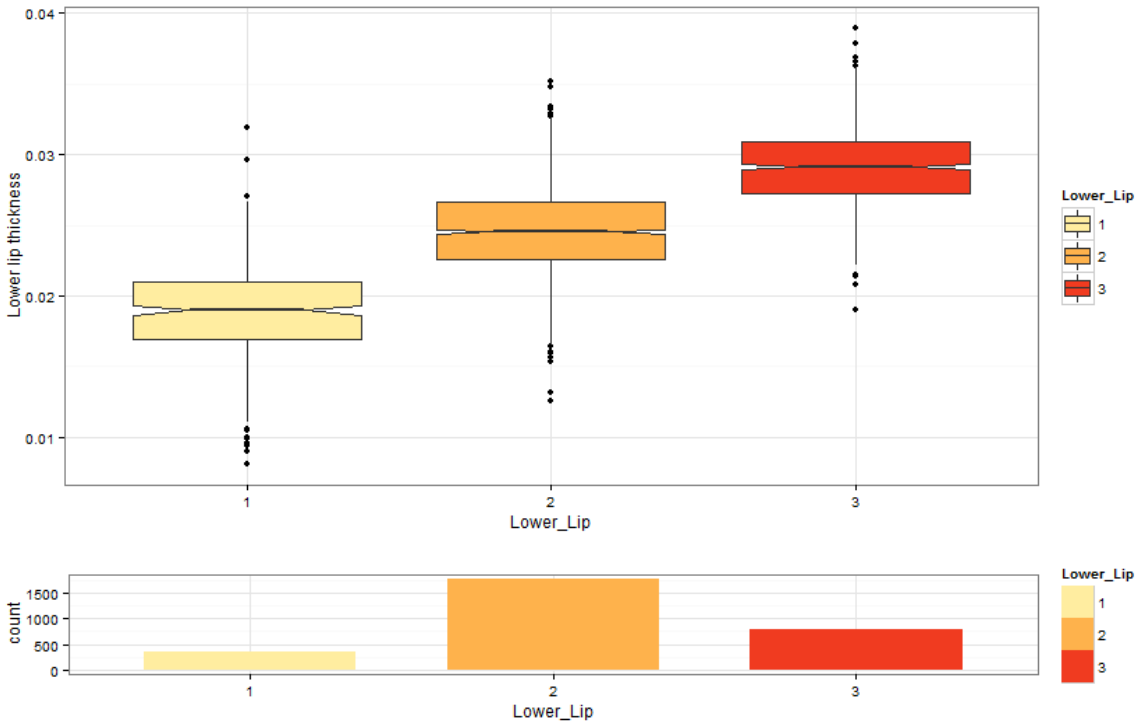
A.19. Columella Inclination



A.20. Upper Lip Thickness



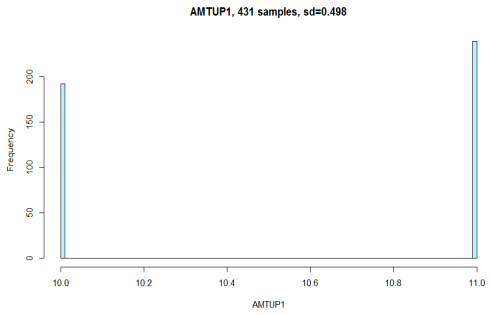
A.21. Lower Lip Thickness



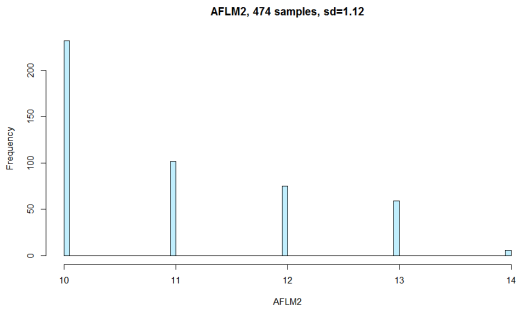
Appendix B to Chapter 4: Genome-Wide Association Studies of Human Dental traits

Figures B.1-B.66. Histograms of 66 out of 86 initial traits showing the distribution of ordinal dental features categorized with ASUDAS scoring method. Traits where all the individuals were scored in the same category are not present here.

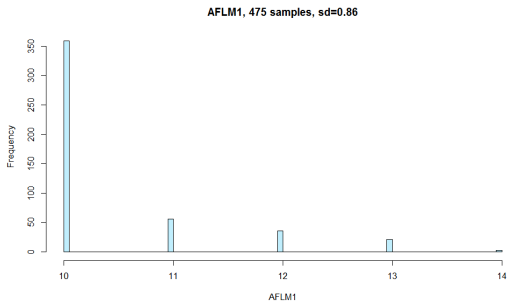
B.1



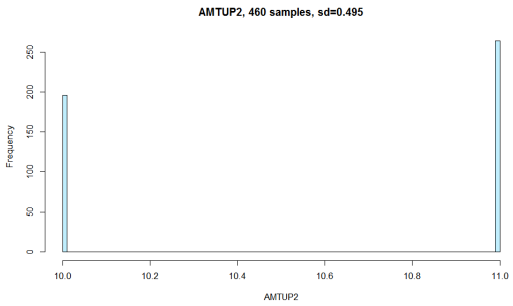
B.2



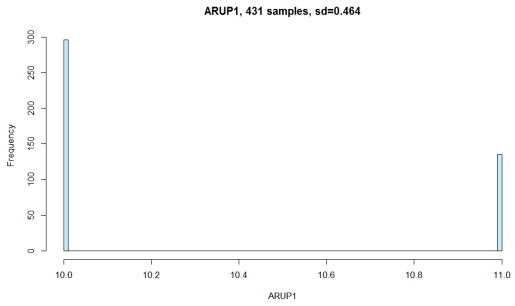
B.3



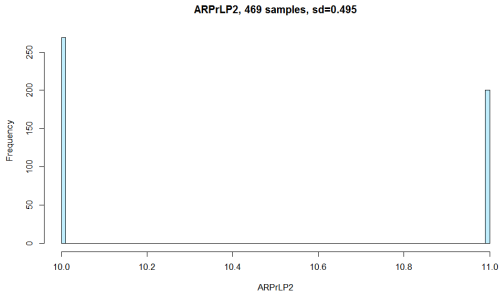
B.4



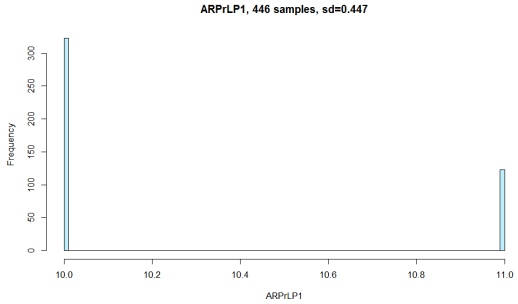
B.5



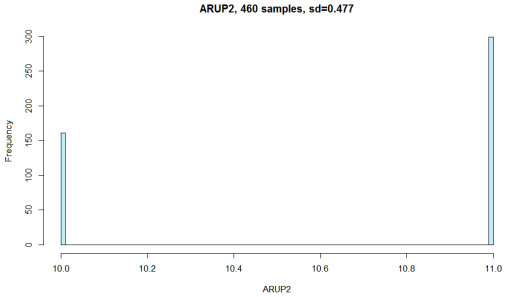
B.6



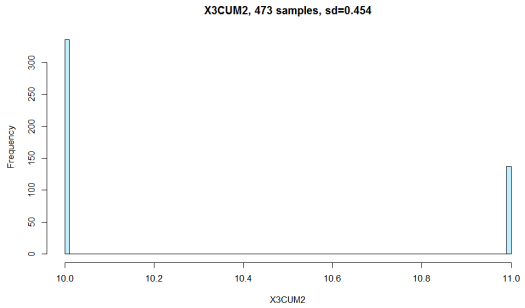
B.7



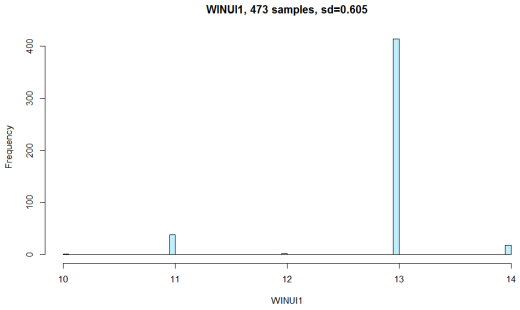
B.8



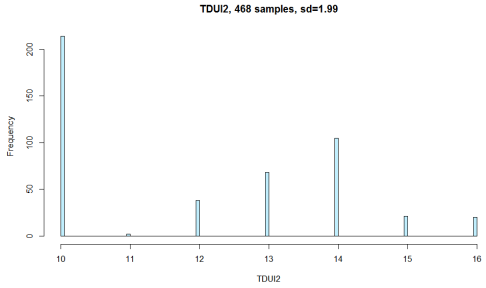
B.9



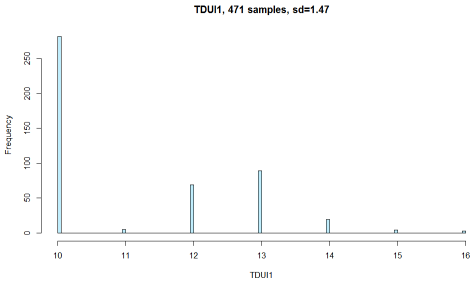
B.10



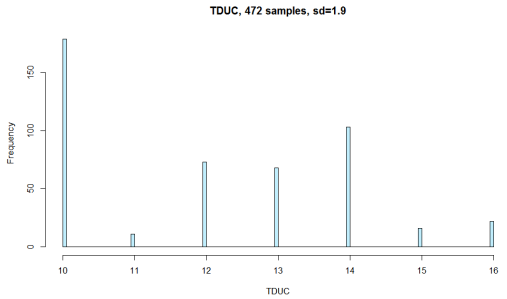
B.11



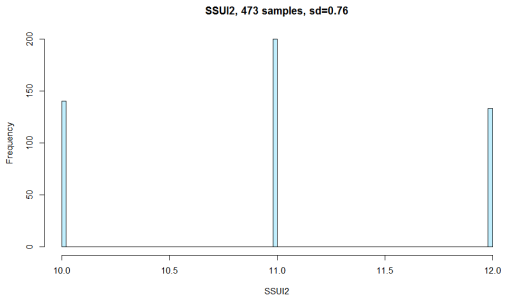
B.12



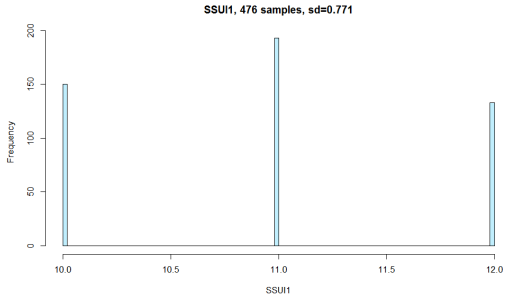
B.13



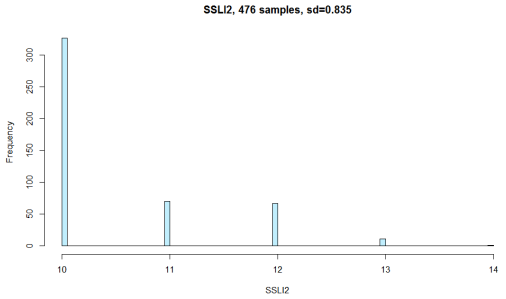
B.14



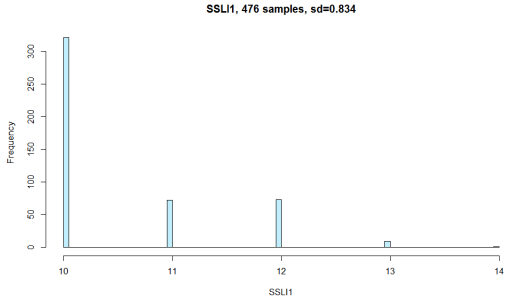
B.15



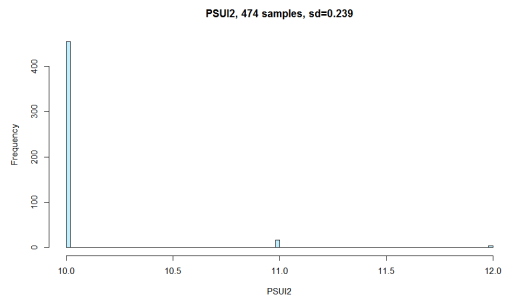
B.16



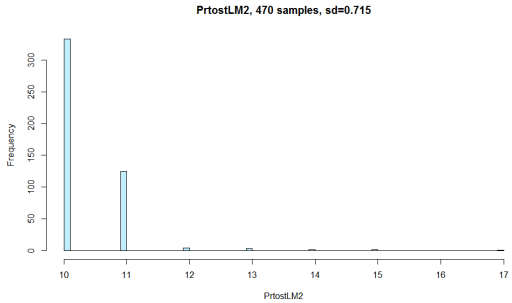
B.17



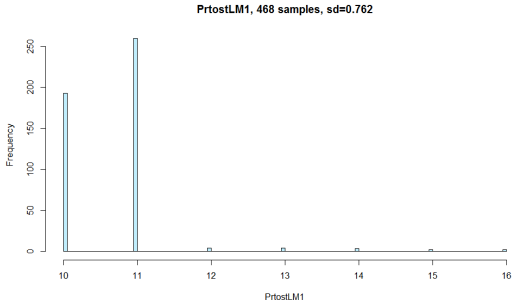
B.18



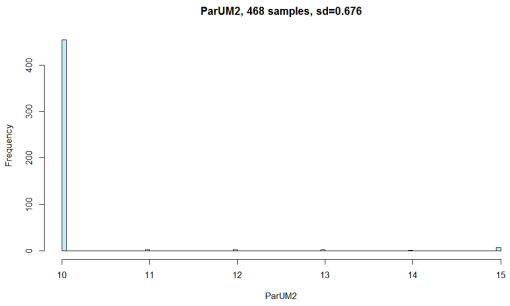
B.19



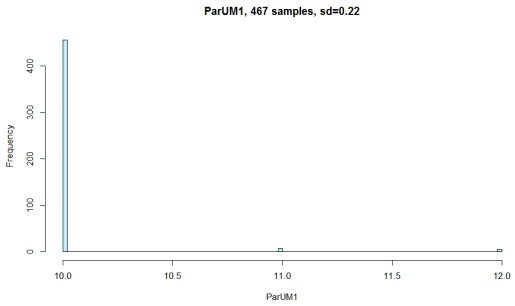
B.20



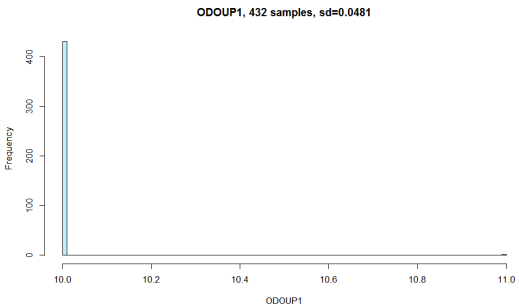
B.21



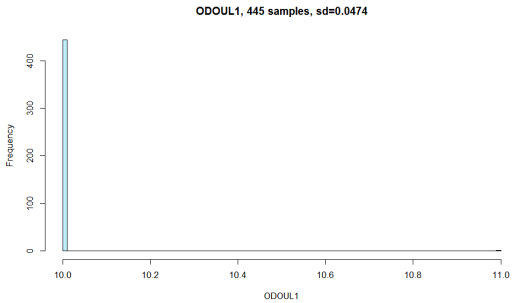
B.22



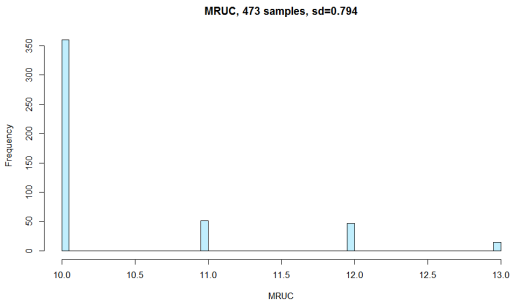
B.23



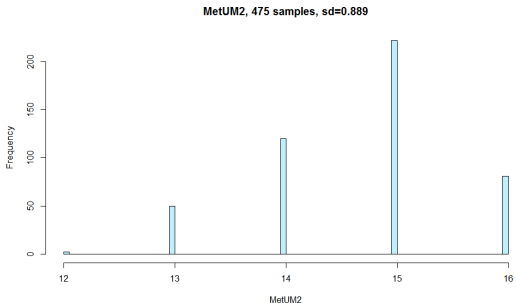
B.24



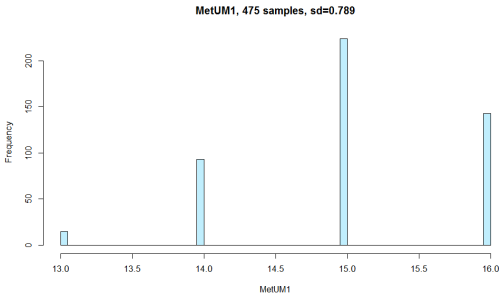
B.25



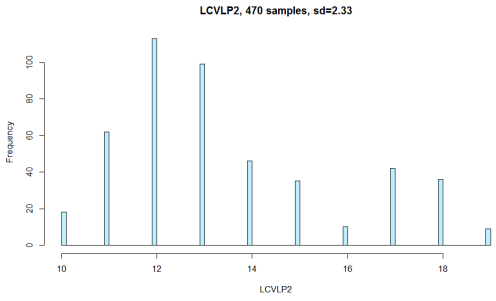
B.26



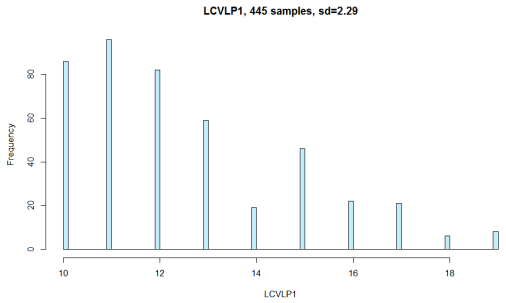
B.27



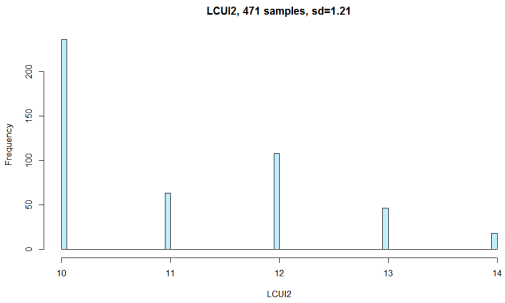
B.28



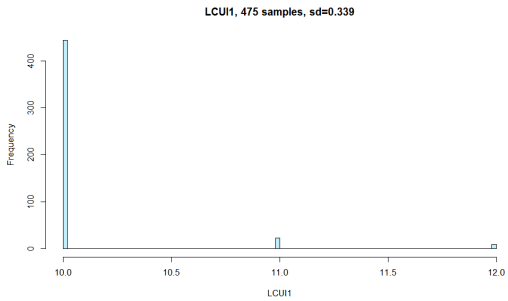
B.29



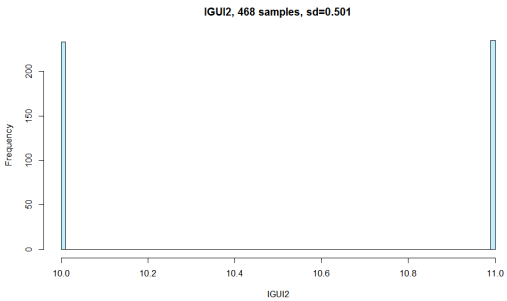
B.30



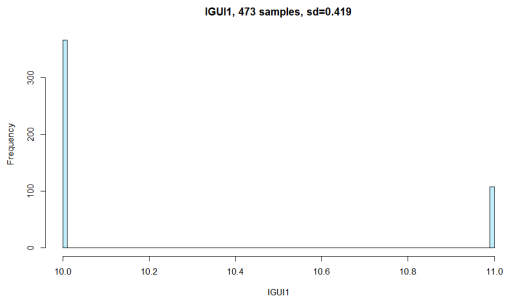
B.31



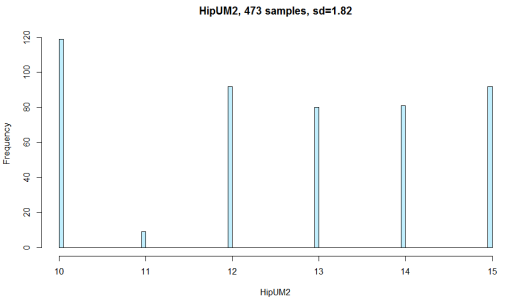
B.32



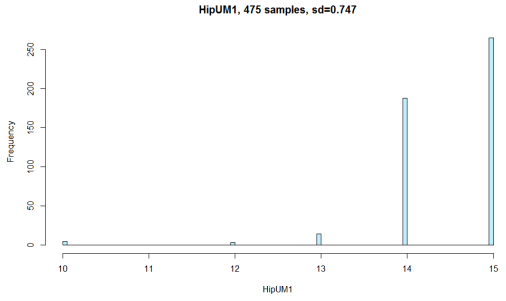
B.33



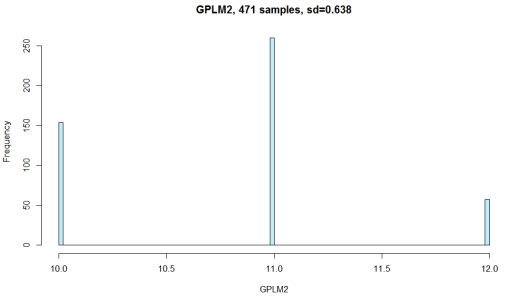
B.34



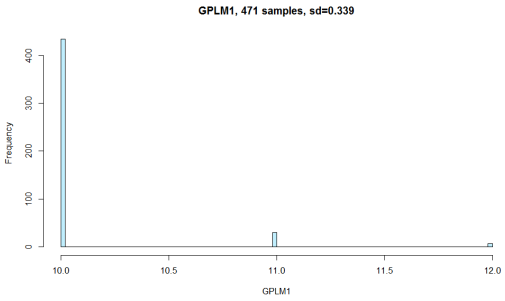
B.35



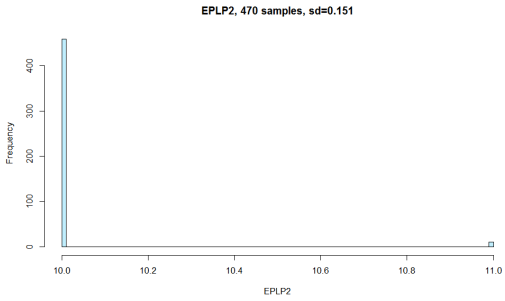
B.36



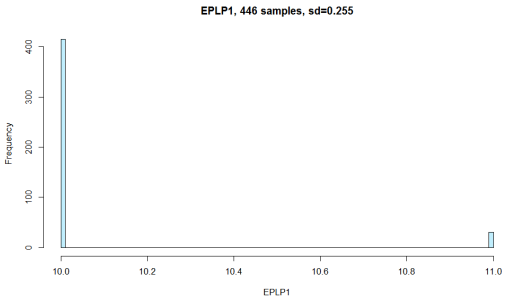
B.37



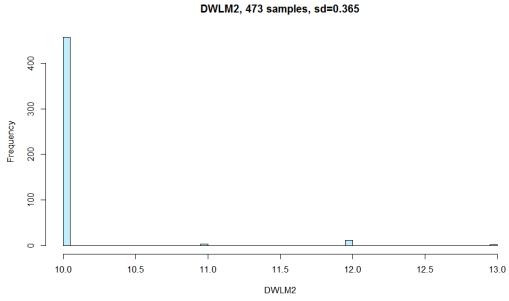
B.38



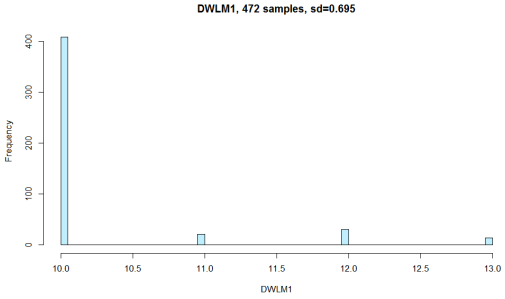
B.39



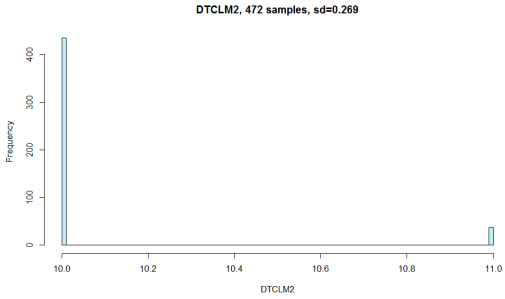
B.40



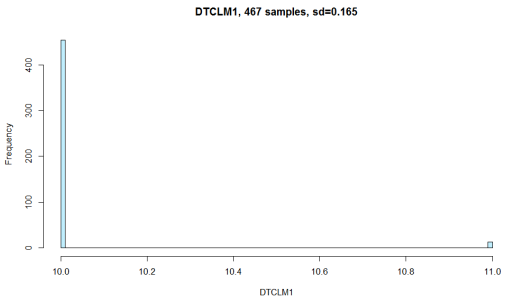
B.41



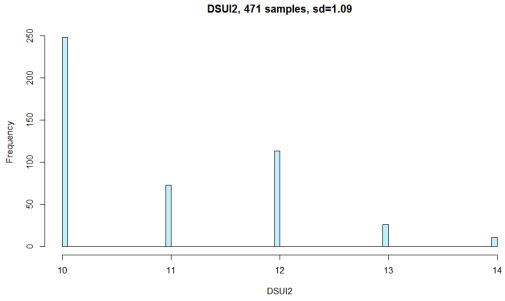
B.42



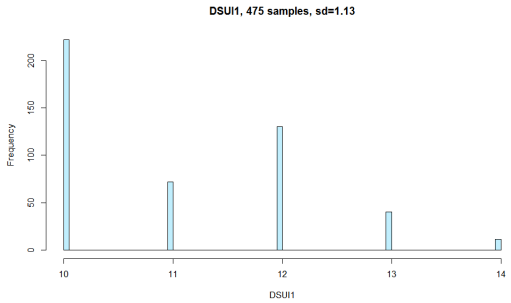
B.43



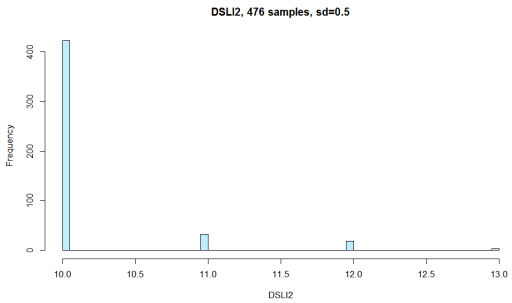
B.44



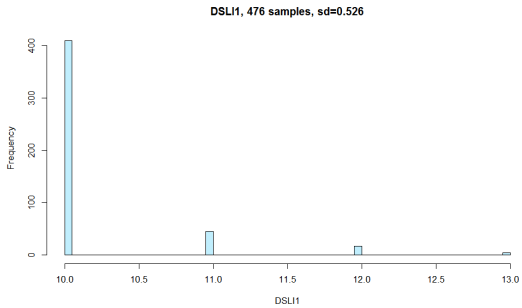
B.45



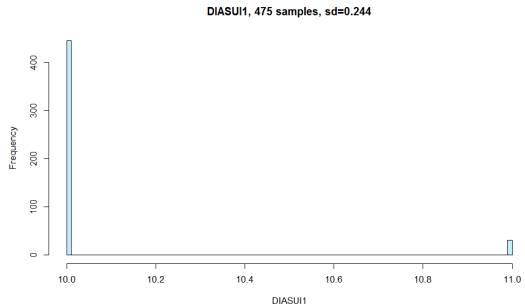
B.46



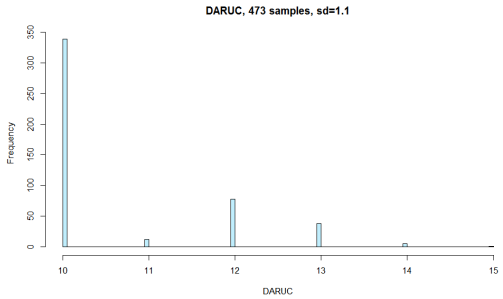
B.47



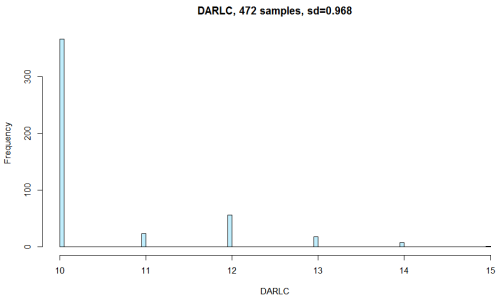
B.48



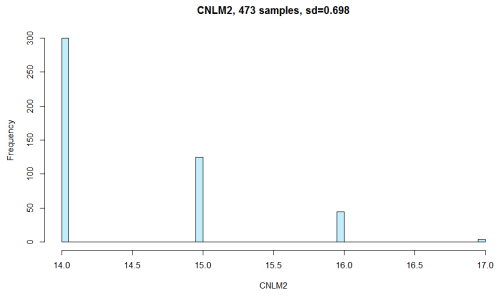
B.49



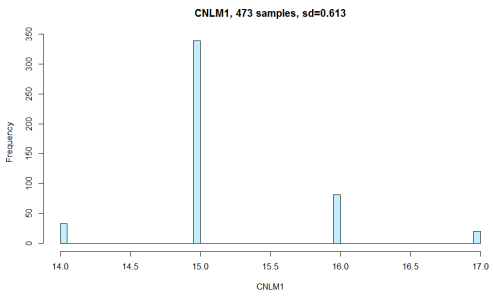
B.50



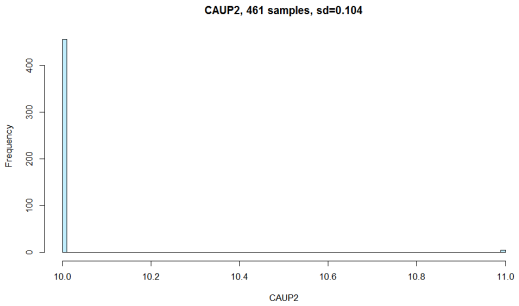
B.51



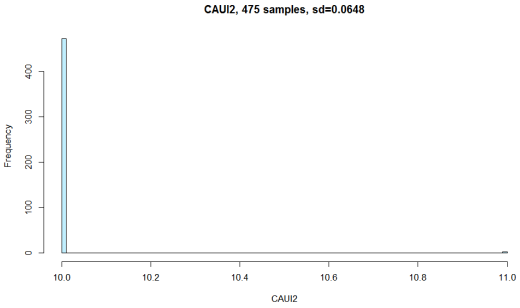
B.52



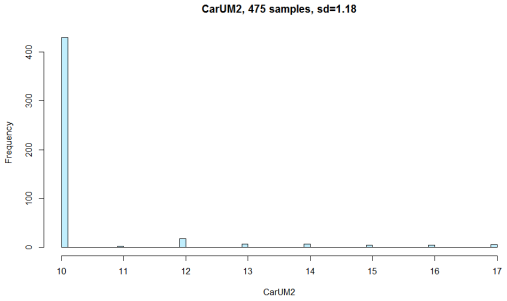
B.53



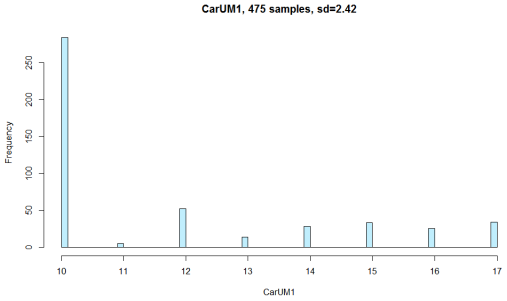
B.54



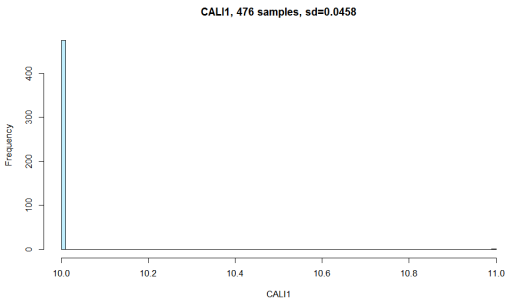
B.55



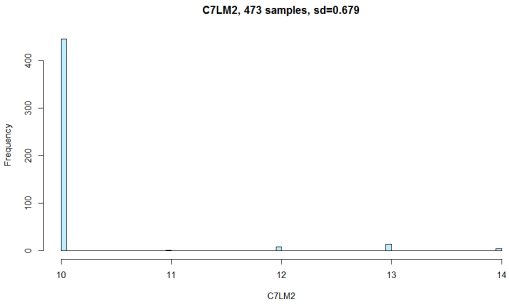
B.56



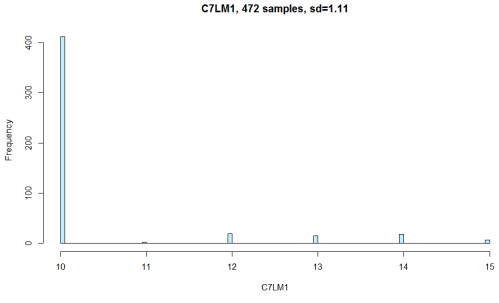
B.57



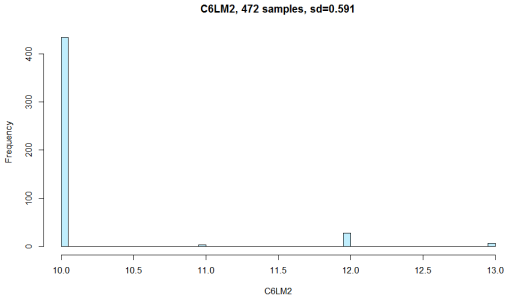
B.58



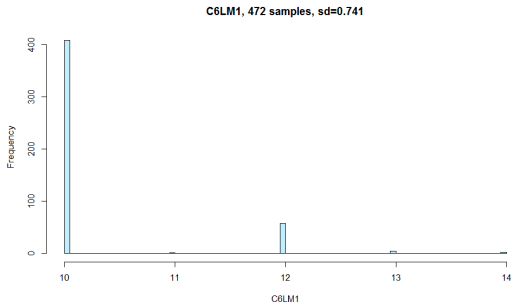
B.59



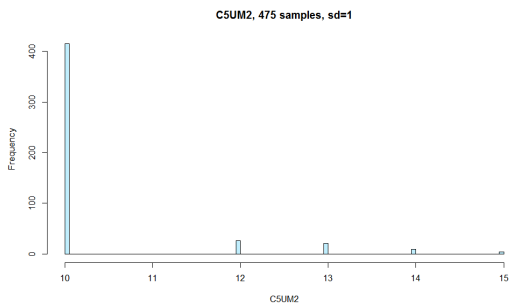
B.60



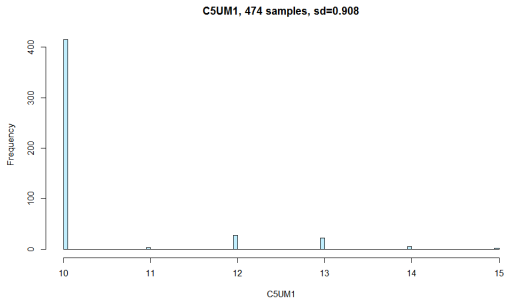
B.61



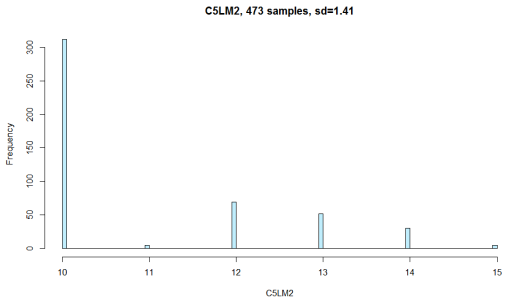
B.62



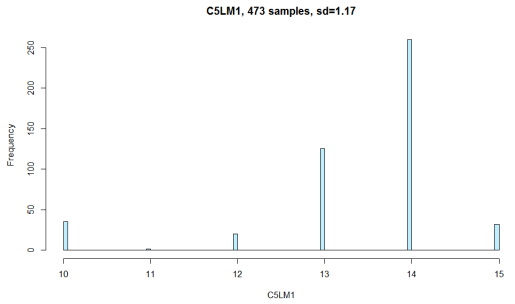
B.63



B.64



B.65



B.66

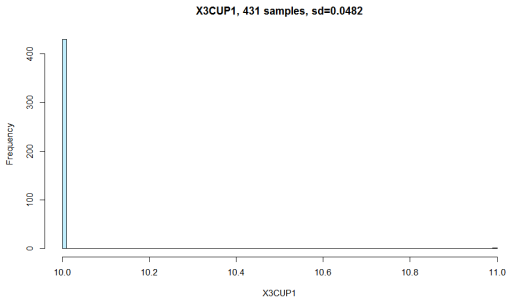
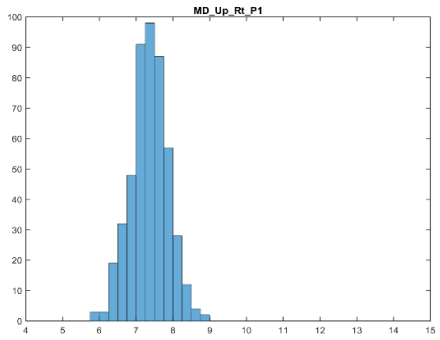


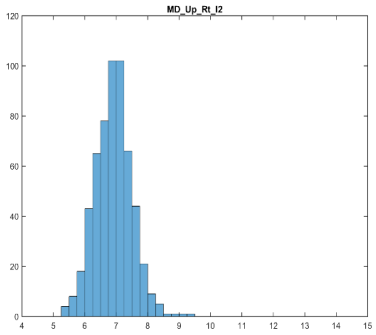
Table B.1. Summary Statistics of dental measurements data.

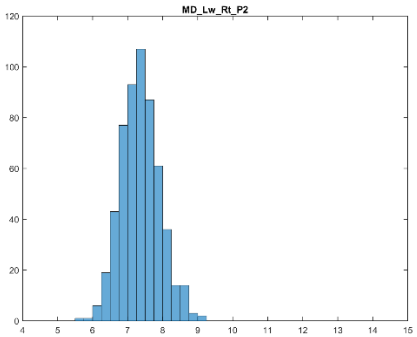
name	min	mean	median	max	std	upper_range	lower_range	max_range
IC_Up_Rt_P2	4.21	6.86	6.86	12.49	0.93	5.63	2.64	5.63
IC_Up_Rt_P1	5.31	7.91	7.91	11.88	0.90	3.97	2.60	3.97
IC_Up_Rt_C	7.07	9.87	9.82	13.31	1.09	3.48	2.76	3.48
IC_Up_Rt_I2	6.05	8.78	8.78	12.46	0.99	3.68	2.73	3.68
IC_Up_Rt_I1	7.56	10.20	10.12	13.58	0.95	3.46	2.56	3.46
IC_Up_Lf_I1	7.54	10.22	10.20	12.91	0.96	2.71	2.67	2.71
IC_Up_Lf_I2	6.38	8.94	8.90	12.55	0.98	3.65	2.52	3.65
IC_Up_Lf_C	6.68	9.95	9.94	13.18	1.05	3.24	3.26	3.26
IC_Up_Lf_P1	5.23	7.97	7.99	11.79	0.85	3.81	2.75	3.81
IC_Up_Lf_P2	4.38	6.83	6.82	10.23	0.88	3.41	2.44	3.41
MD_Up_Lf_P2	5.13	6.77	6.76	10.36	0.61	3.59	1.64	3.59
MD_Up_Lf_P1	5.88	7.36	7.34	10.76	0.53	3.42	1.46	3.42
MD_Up_Lf_C	6.15	7.82	7.81	9.56	0.55	1.75	1.66	1.75
MD_Up_Lf_I2	4.97	6.84	6.80	8.87	0.60	2.07	1.83	2.07
MD_Up_Lf_I1	6.92	8.75	8.75	10.56	0.56	1.81	1.84	1.84
MD_Up_Rt_I1	6.96	8.80	8.79	10.95	0.56	2.16	1.83	2.16
MD_Up_Rt_I2	4.43	6.91	6.90	9.33	0.59	2.43	2.47	2.47
MD_Up_Rt_C	6.41	7.88	7.87	9.82	0.54	1.96	1.46	1.96
MD_Up_Rt_P1	5.26	7.34	7.34	9.80	0.52	2.46	2.08	2.46
MD_Up_Rt_P2	5.18	6.85	6.87	9.83	0.59	2.96	1.69	2.96
BL_Up_Lf_P2	7.41	9.58	9.55	11.77	0.65	2.22	2.14	2.22
BL_Up_Lf_P1	7.70	9.57	9.53	11.51	0.62	1.97	1.83	1.97
BL_Up_Lf_C	6.76	8.50	8.49	10.37	0.62	1.88	1.74	1.88
BL_Up_Lf_I2	4.69	6.89	6.88	9.44	0.61	2.56	2.19	2.56
BL_Up_Lf_I1	5.85	7.62	7.61	9.54	0.56	1.93	1.76	1.93
BL_Up_Rt_I1	5.71	7.61	7.58	10.02	0.58	2.44	1.87	2.44
BL_Up_Rt_I2	4.35	6.82	6.83	9.60	0.62	2.77	2.49	2.77
BL_Up_Rt_C	6.46	8.52	8.51	10.38	0.64	1.87	2.05	2.05
BL_Up_Rt_P1	7.75	9.63	9.62	11.52	0.62	1.90	1.87	1.90

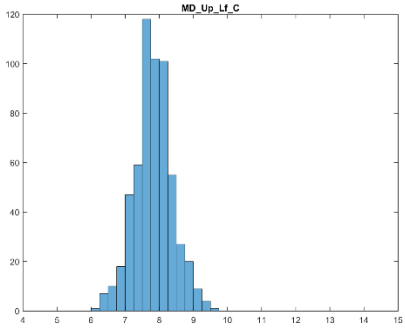
name	min	mean	median	max	std	upper_range	lower_range	max_range
IC_Lw_Rt_P2	4.74	7.15	7.12	11.60	0.86	4.48	2.38	4.48
IC_Lw_Rt_P1	5.30	8.23	8.24	10.91	0.85	2.67	2.94	2.94
IC_Lw_Rt_C	6.75	10.02	10.00	14.00	1.23	4.00	3.25	4.00
IC_Lw_Rt_I2	6.34	8.82	8.81	12.23	0.95	3.42	2.47	3.42
IC_Lw_Rt_I1	6.08	8.66	8.65	12.22	0.92	3.57	2.57	3.57
IC_Lw_Lf_I1	6.37	8.73	8.75	12.32	0.94	3.58	2.38	3.58
IC_Lw_Lf_I2	6.46	8.93	8.91	11.79	0.94	2.88	2.44	2.88
IC_Lw_Lf_C	7.17	10.14	10.04	13.76	1.21	3.73	2.87	3.73
IC_Lw_Lf_P1	5.76	8.32	8.31	11.48	0.83	3.17	2.55	3.17
IC_Lw_Lf_P2	4.36	7.26	7.21	12.82	0.89	5.62	2.85	5.62
MD_Lw_Lf_P2	5.95	7.43	7.42	11.62	0.59	4.21	1.46	4.21
MD_Lw_Lf_P1	5.92	7.31	7.32	11.44	0.55	4.13	1.39	4.13
MD_Lw_Lf_C	5.67	6.94	6.92	13.27	0.60	6.34	1.25	6.34
MD_Lw_Lf_I2	4.83	6.11	6.08	9.70	0.48	3.63	1.24	3.63
MD_Lw_Lf_I1	4.19	5.64	5.60	8.99	0.44	3.39	1.41	3.39
MD_Lw_Rt_I1	4.57	5.65	5.62	8.60	0.43	2.97	1.05	2.97
MD_Lw_Rt_I2	5.00	6.14	6.09	8.57	0.43	2.49	1.09	2.49
MD_Lw_Rt_C	5.04	6.97	6.96	10.96	0.58	4.00	1.92	4.00
MD_Lw_Rt_P1	5.54	7.32	7.33	10.97	0.56	3.64	1.79	3.64
MD_Lw_Rt_P2	5.72	7.37	7.36	11.39	0.57	4.03	1.63	4.03
BL_Lw_Lf_P2	7.12	8.82	8.80	11.40	0.60	2.60	1.68	2.60
BL_Lw_Lf_P1	6.65	8.12	8.11	12.05	0.59	3.94	1.46	3.94
BL_Lw_Lf_C	6.28	7.94	7.91	13.96	0.69	6.05	1.63	6.05
BL_Lw_Lf_I2	5.36	6.69	6.66	9.03	0.48	2.37	1.30	2.37
BL_Lw_Lf_I1	5.11	6.40	6.35	11.37	0.57	5.02	1.24	5.02
BL_Lw_Rt_I1	5.08	6.39	6.34	11.00	0.55	4.66	1.26	4.66
BL_Lw_Rt_I2	5.35	6.73	6.70	11.55	0.52	4.84	1.35	4.84
BL_Lw_Rt_C	6.45	7.97	7.95	13.63	0.69	5.68	1.49	5.68
BL_Lw_Rt_P1	6.63	8.15	8.13	11.71	0.58	3.59	1.49	3.59
BL_Lw_Rt_P2	6.95	8.84	8.86	10.93	0.57	2.07	1.91	2.07

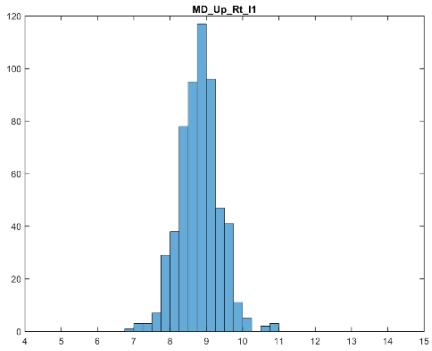
Figures B.67-B.115. Histograms showing the distribution of measurements of dental features.

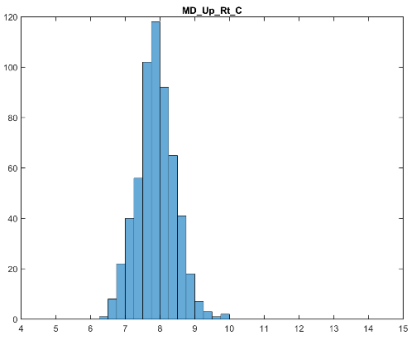


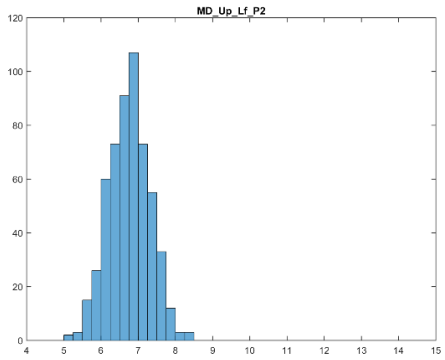


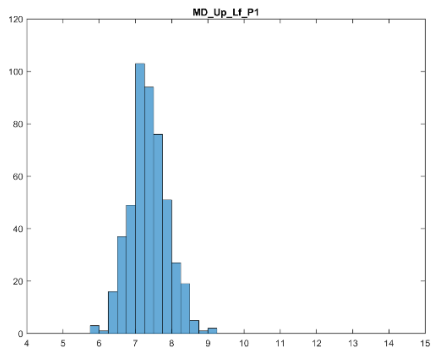


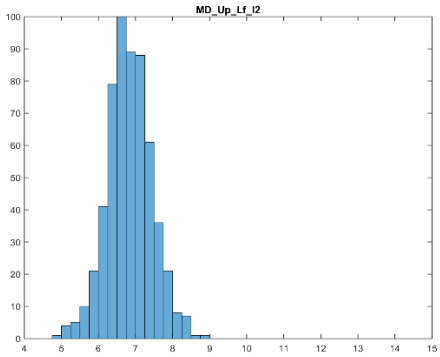


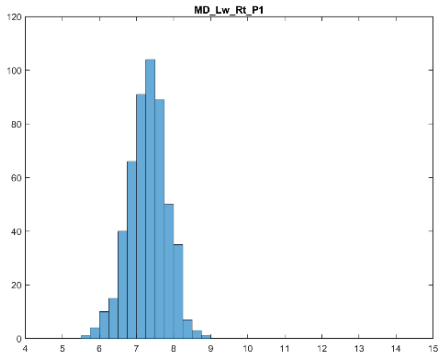


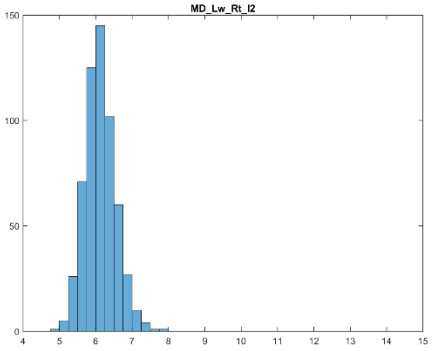


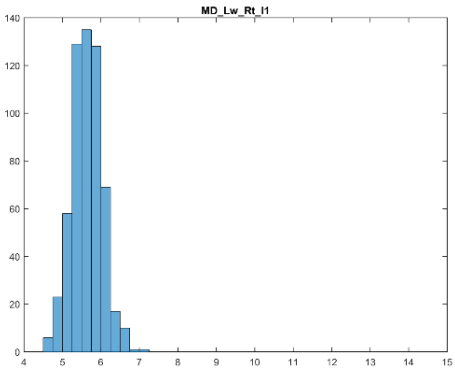


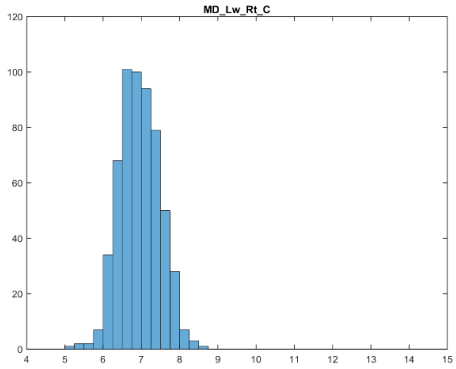


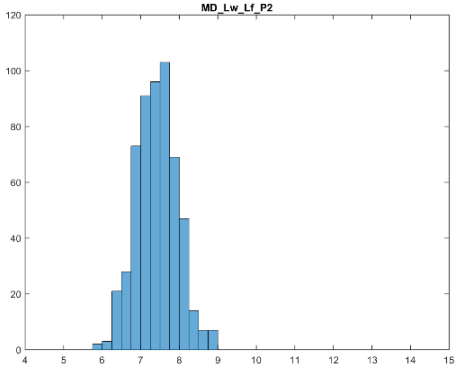

MD_Up_Lf_I2

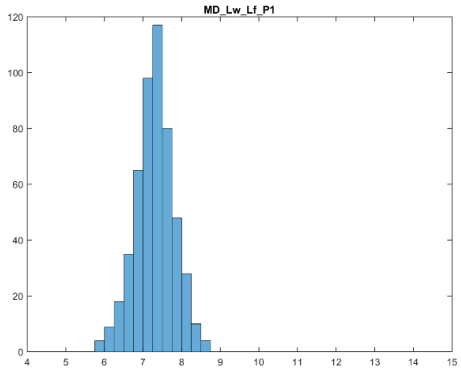

MD_Lw_Rt_P1

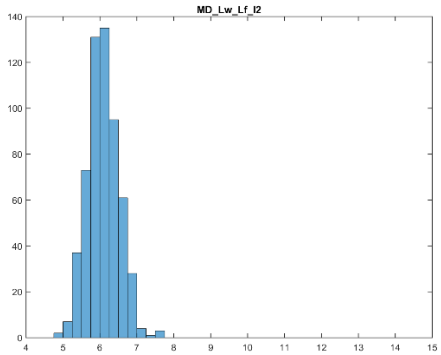

MD_Lw_Rt_I2

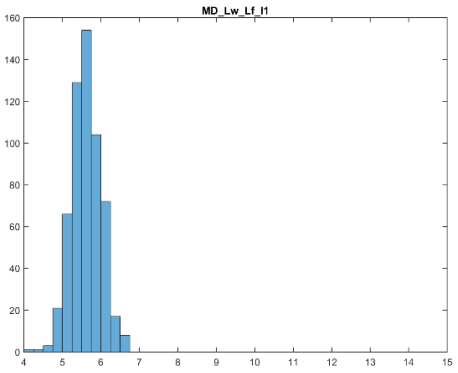

MD_Lw_Rt_I1

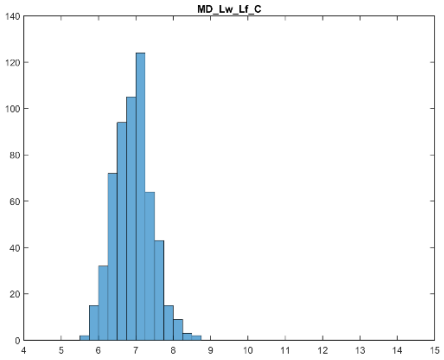

MD_Lw_Rt_C

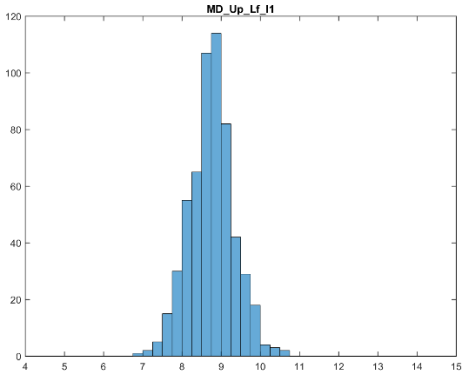

MD_Lw_Lf_P2

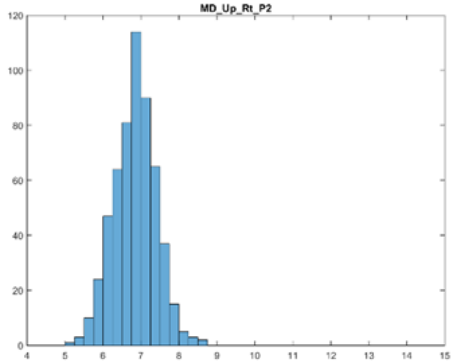

MD_Lw_Lf_P1

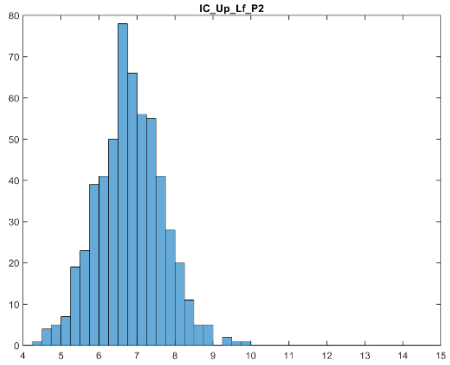

MD_Lw_Lf_I2

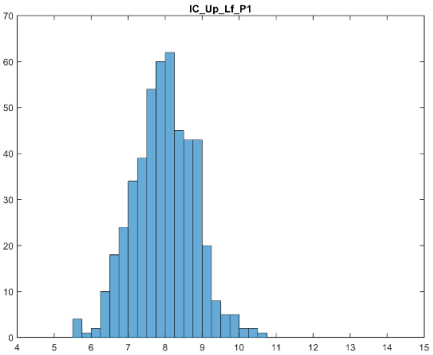

MD_Lw_Lf_I1

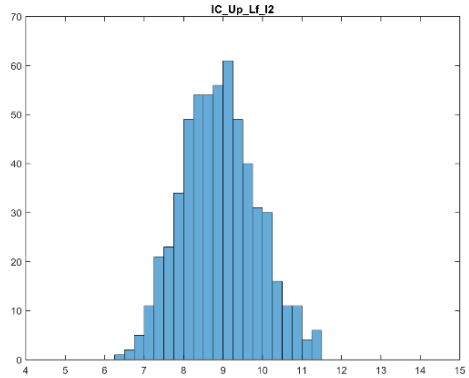

MD_Lw_Lf_C

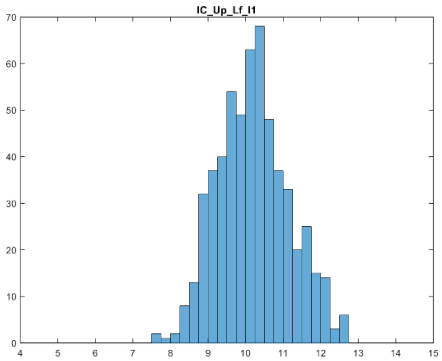

MD_Up_Lf_I1

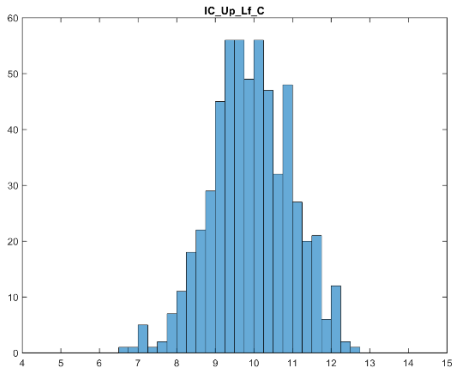

MD_Up_Rt_P2

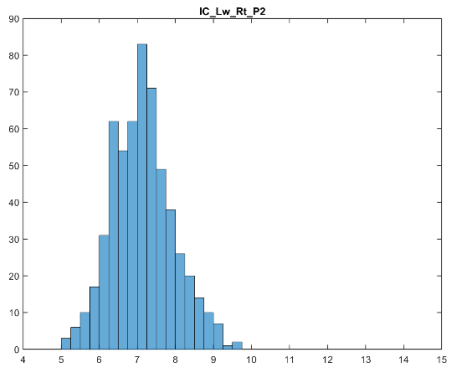

IC_Up_Lf_P2

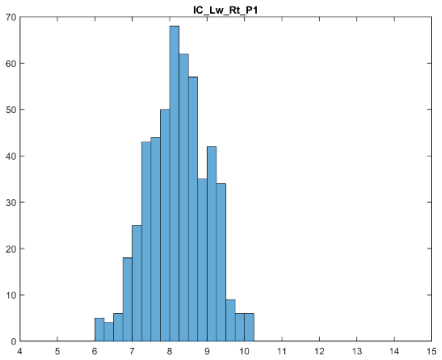

IC_Up_Lf_P1

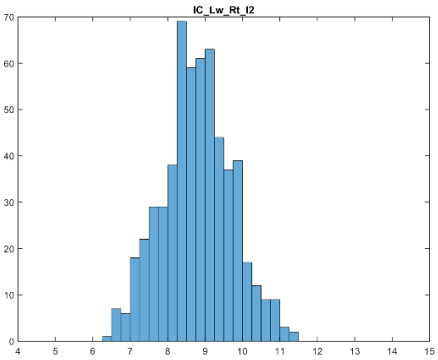

IC_Up_Lf_I2

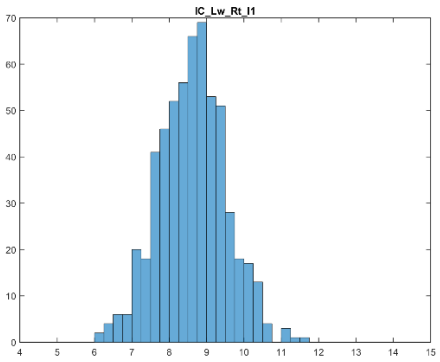

IC_Up_Lf_I1

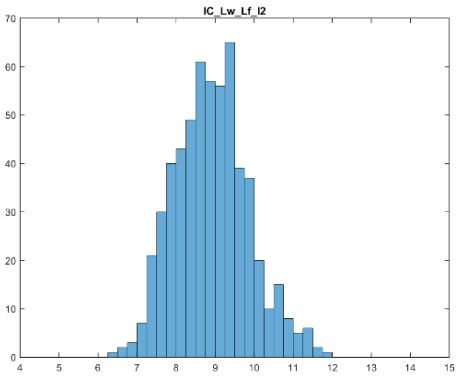

IC_Up_Lf_C

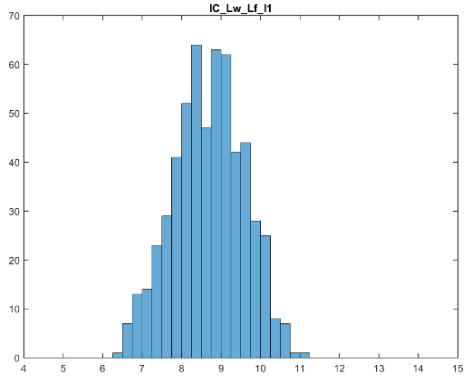

IC_Lw_Rt_P2

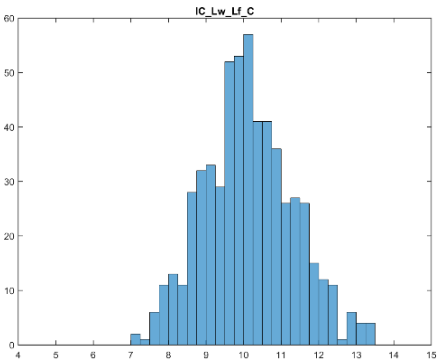

IC_Lw_Rt_P1

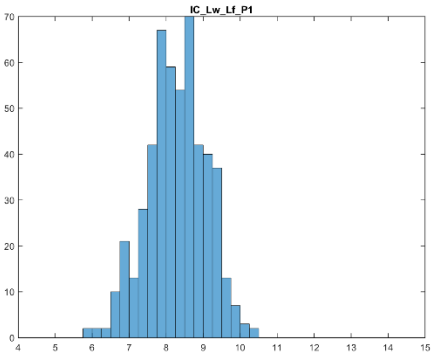

IC_Lw_Rt_I2

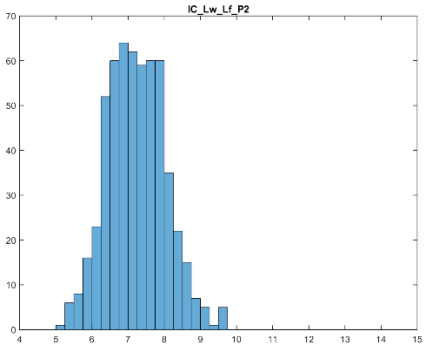

IC_Lw_Rt_I1

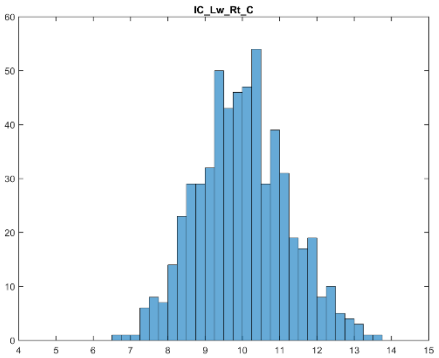

IC_Lw_Lf_I2

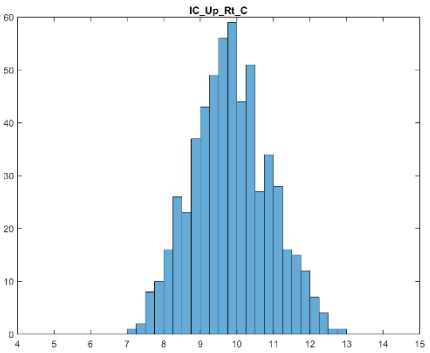

IC_Lw_Lf_I1

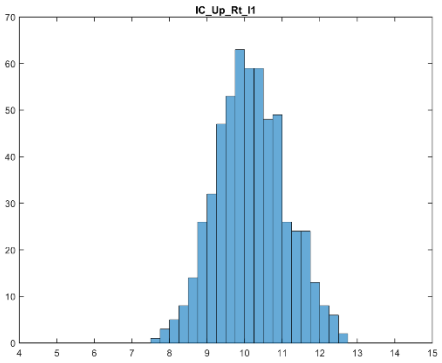

IC_Lw_Lf_C

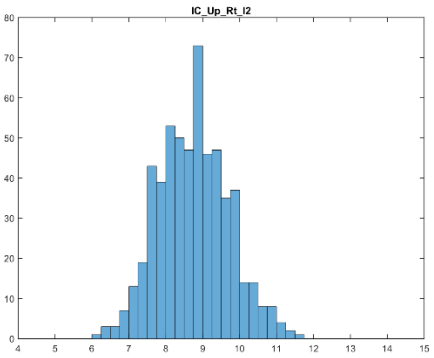

IC_Lw_Lf_P1

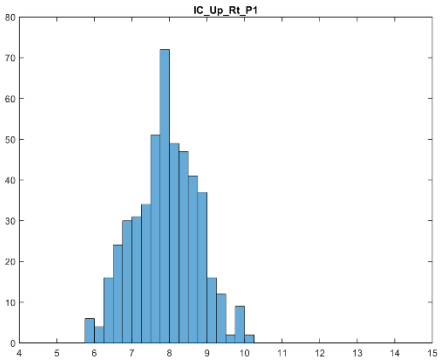

IC_Lw_Lf_P2

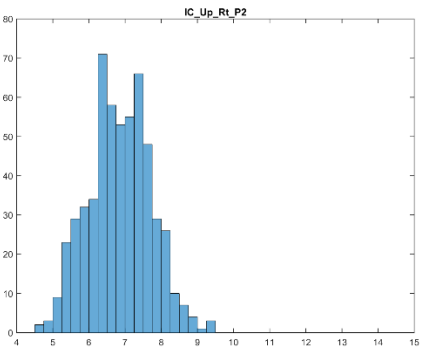

IC_Lw_Rt_C

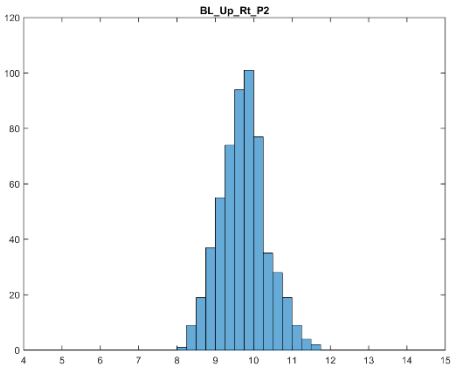

IC_Up_Rt_C

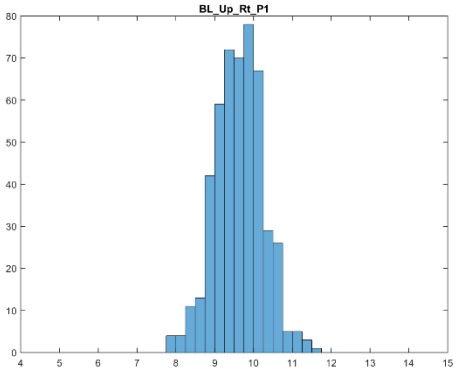

IC_Up_Rt_I1

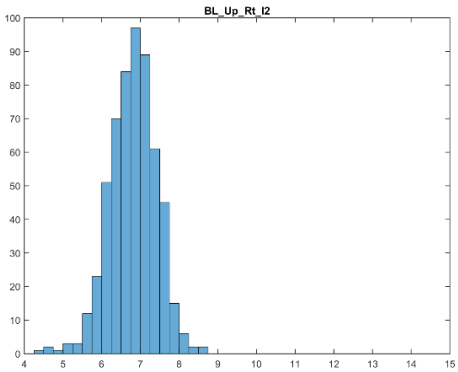

IC_Up_Rt_I2

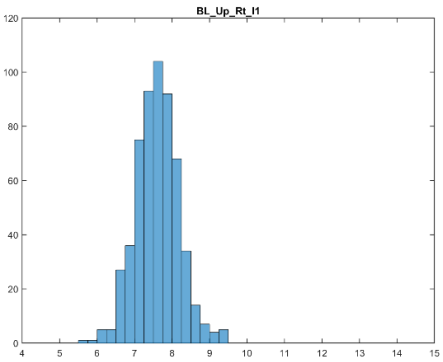

IC_Up_Rt_P1

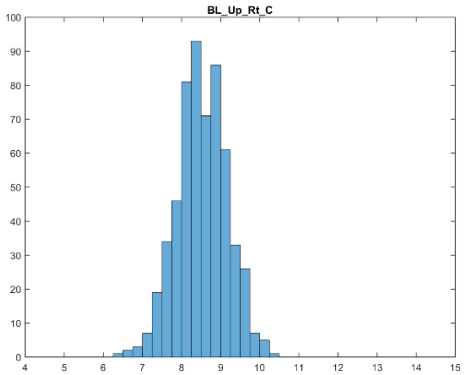

IC_Up_Rt_P2

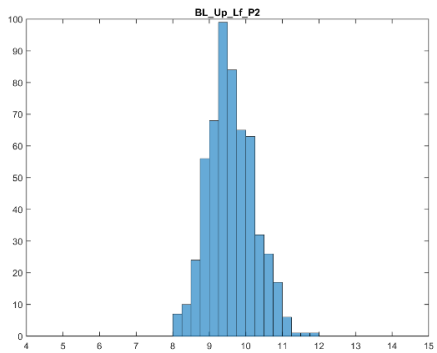

BL_Up_Rt_P2

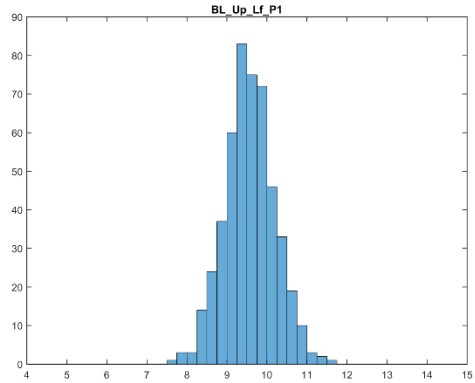

BL_Up_Rt_P1

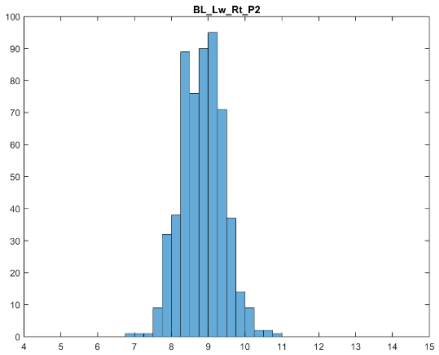

BL_Up_Rt_I2

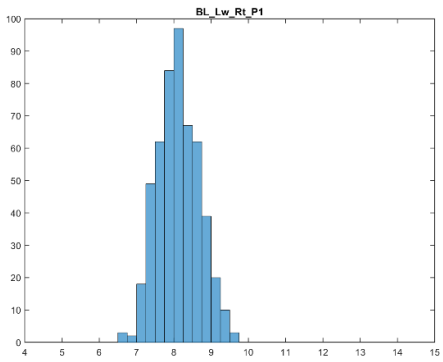

BL_Up_Rt_I1

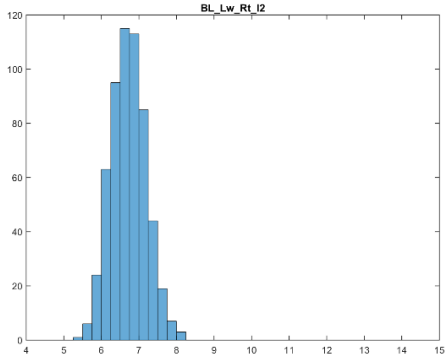

BL_Up_Rt_C

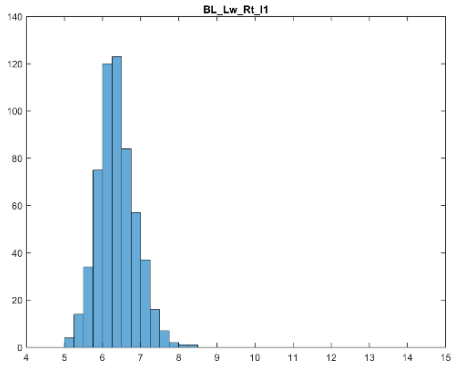

BL_Up_Lf_P2

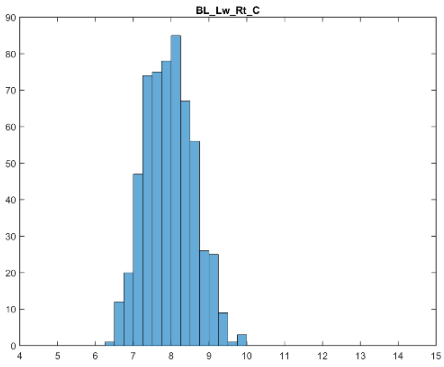

BL_Up_Lf_P1

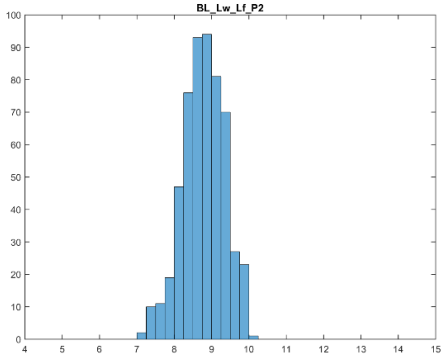

BL_Lw_Rt_P2

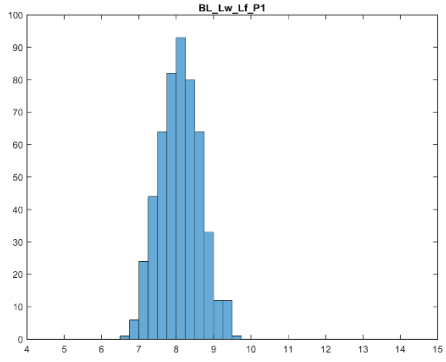

BL_Lw_Rt_P1

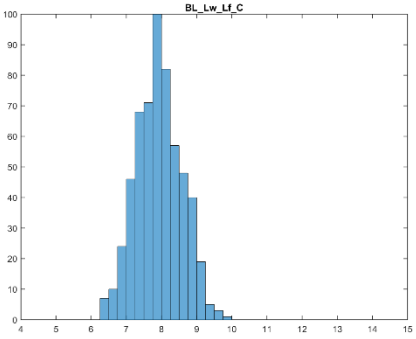

BL_Lw_Rt_I2

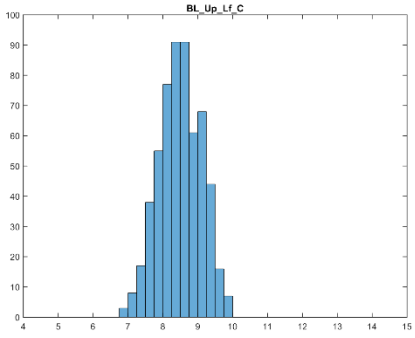

BL_Lw_Rt_I1

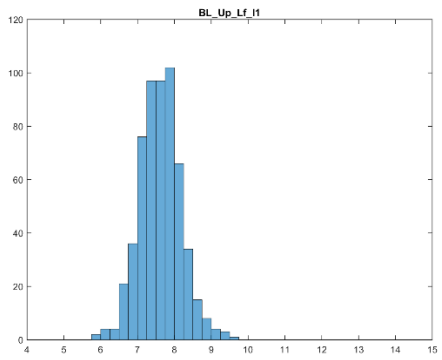

BL_Lw_Rt_C

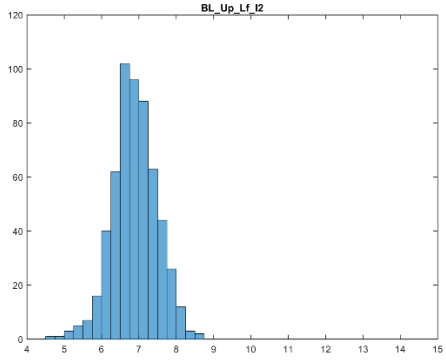

BL_Lw_Lf_P2

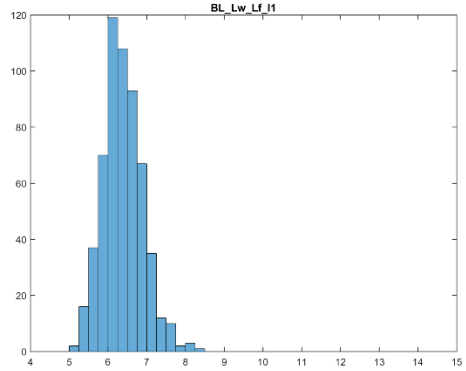

BL_Lw_Lf_P1

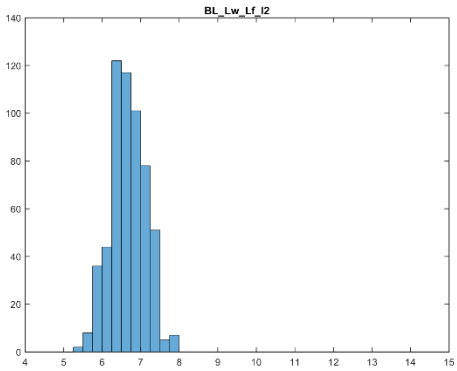

BL_Lw_Lf_C


BL_Up_Lf_C

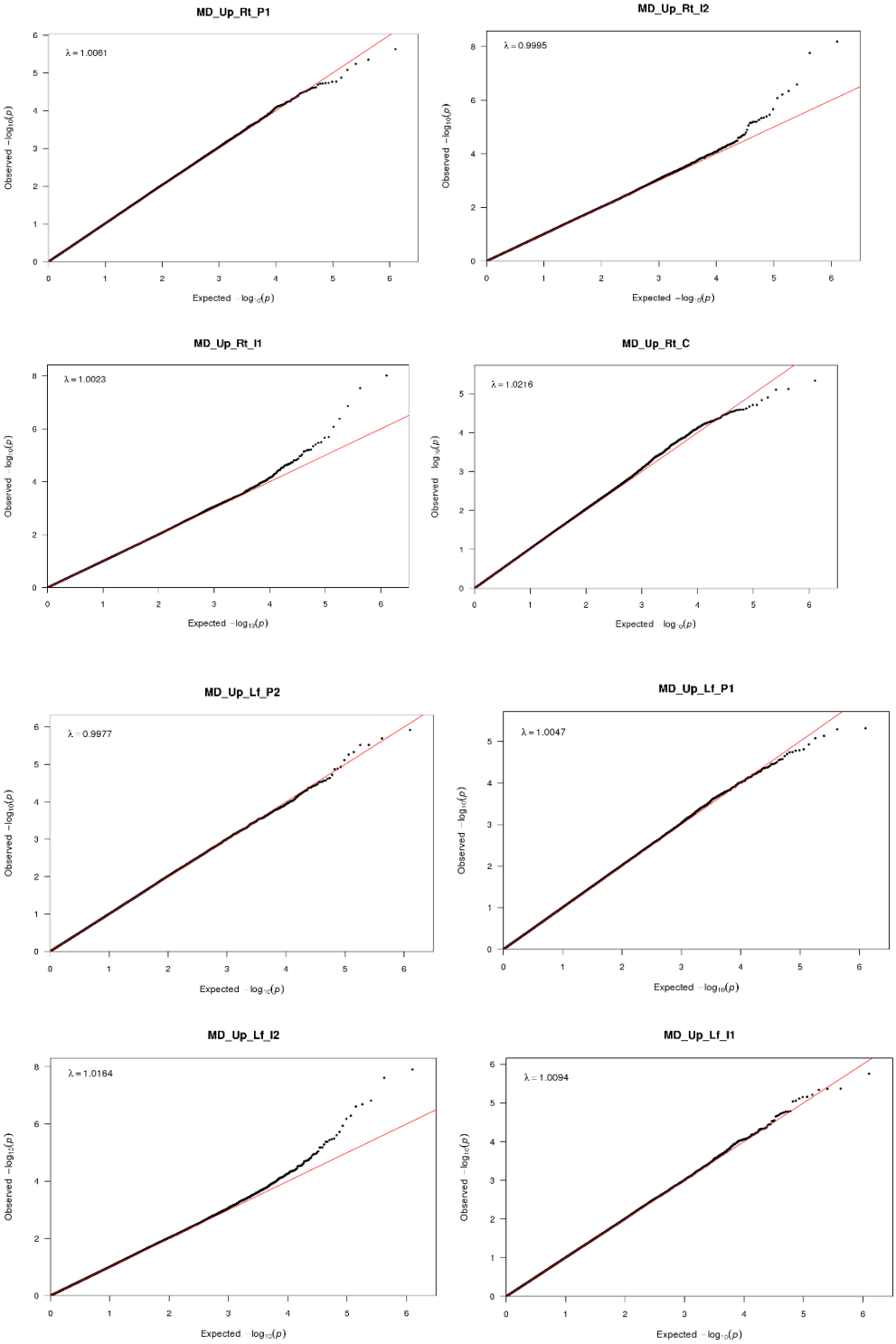

BL_Up_Lf_l1
120
100
80
60
40
20
0
4 5 6 7 8 9 10 11 12 13 14 15

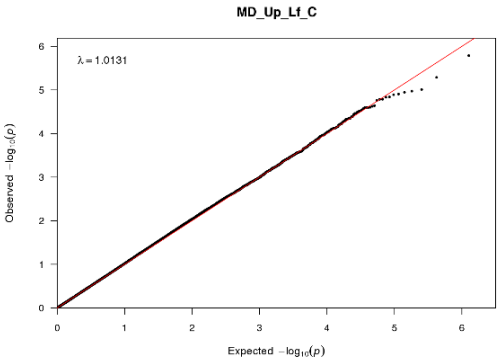

BL_Up_Lf_l2
120
100
80
60
40
20
0
4 5 6 7 8 9 10 11 12 13 14 15

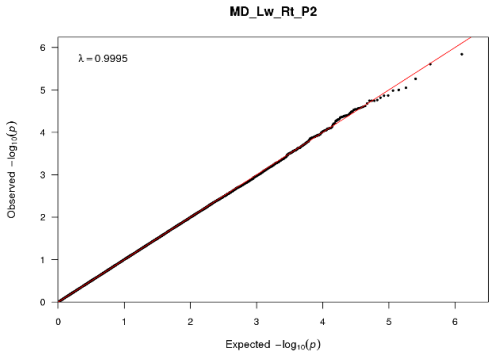

BL_Lw_Lf_l1
120
100
80
60
40
20
0
4 5 6 7 8 9 10 11 12 13 14 15

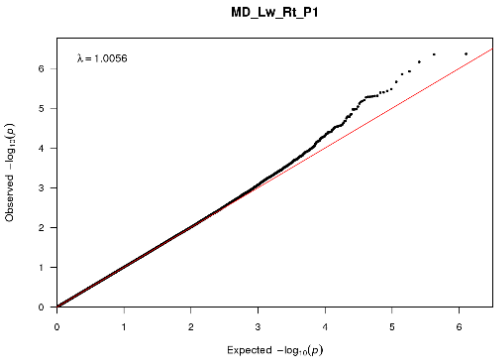

BL_Lw_Lf_l2
140
120
100
80
60
40
20
0
4 5 6 7 8 9 10 11 12 13 14 15

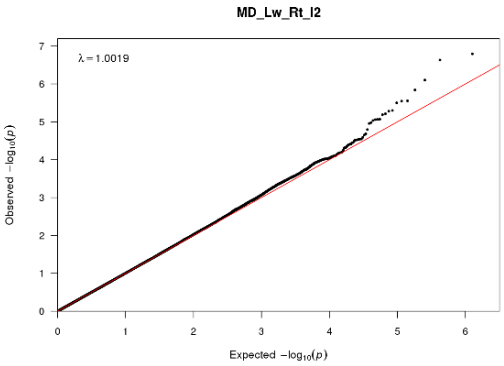
Figures B.116-B.176. Q–Q plots for all association tests (60 measurements) showed no sign of inflation, the genomic control factor lambda being <1.02 in all cases.

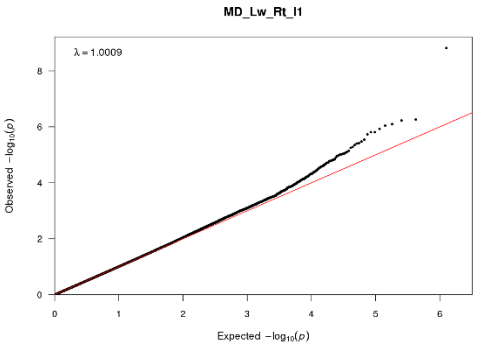


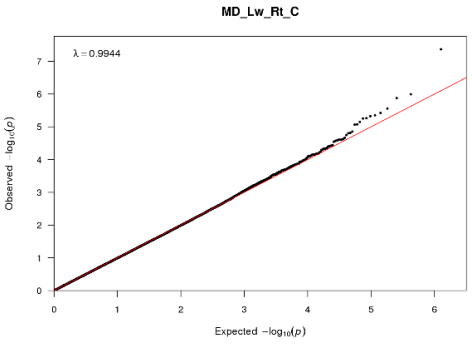

MD_Up_Lf_C
λ = 1.0131
Observed −log10(p)
Expected −log10(p)

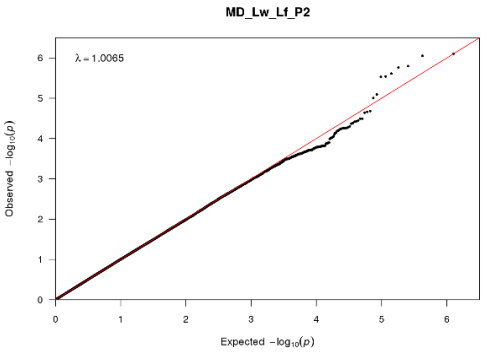

MD_Lw_Rt_P2
λ = 0.9995
Observed −log10(p)
Expected −log10(p)

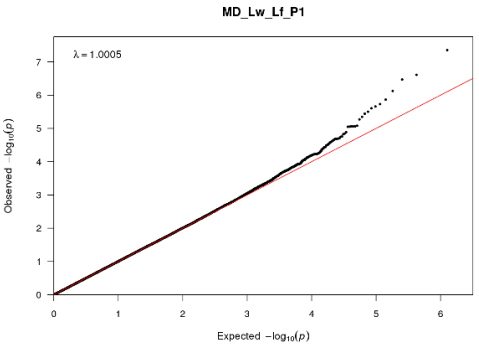

MD_Lw_Rt_P1
λ = 1.0056
Observed −log10(p)
Expected −log10(p)

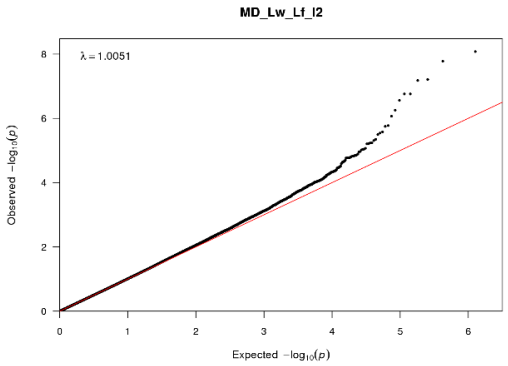

MD_Lw_Rt_I2
λ = 1.0019
Observed −log10(p)
Expected −log10(p)

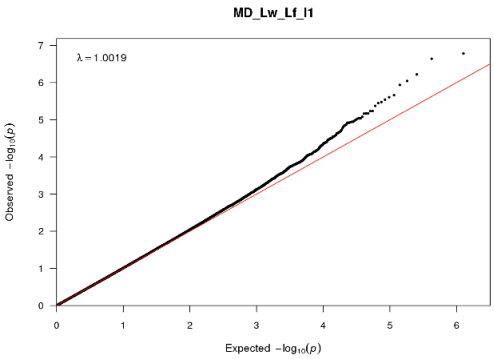

MD_Lw_Rt_I1
λ = 1.0009
Observed −log10(p)
Expected −log10(p)

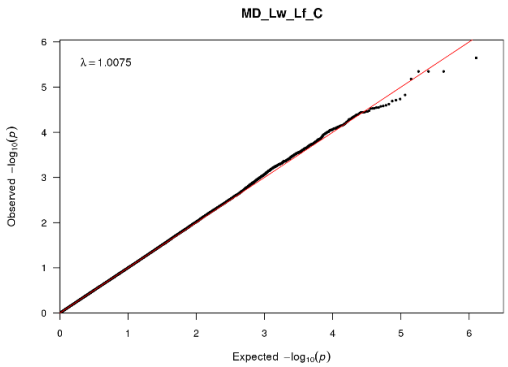

MD_Lw_Rt_C
λ = 0.9944
Observed −log10(p)
Expected −log10(p)

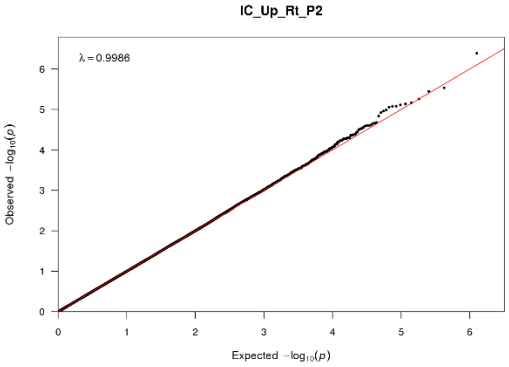

MD_Lw_Lf_P2
λ = 1.0065
Observed −log10(p)
Expected −log10(p)

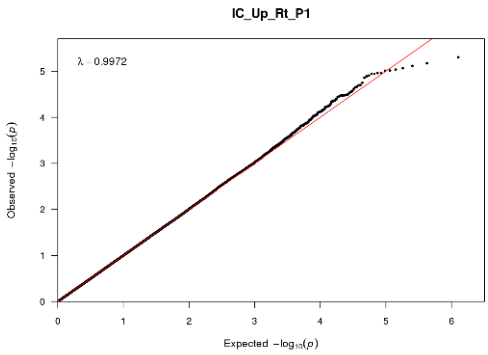

MD_Lw_Lf_P1
λ = 1.0005
Observed −log10(p)
Expected −log10(p)

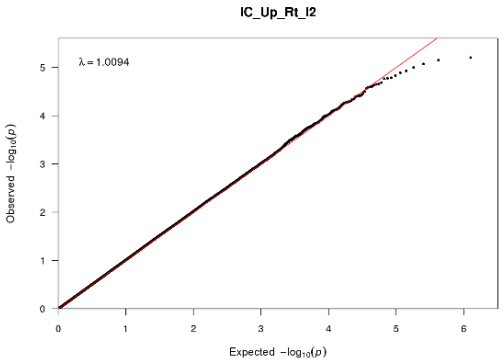

MD_Lw_Lf_I2
λ = 1.0051
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

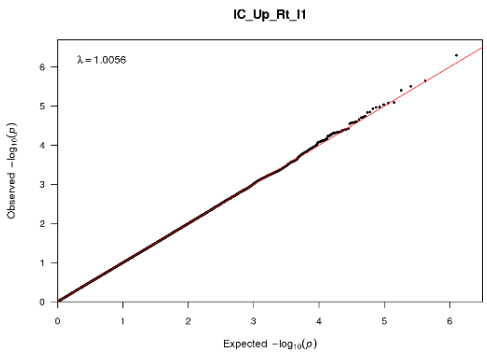

MD_Lw_Lf_I1
λ = 1.0019
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

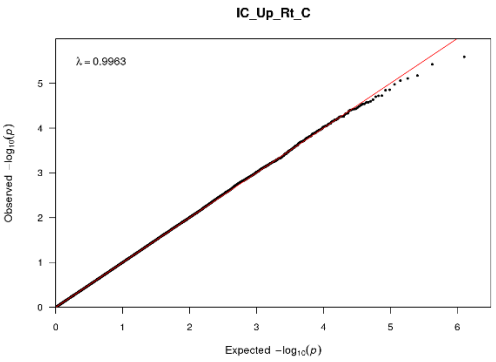

MD_Lw_Lf_C
λ = 1.0075
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

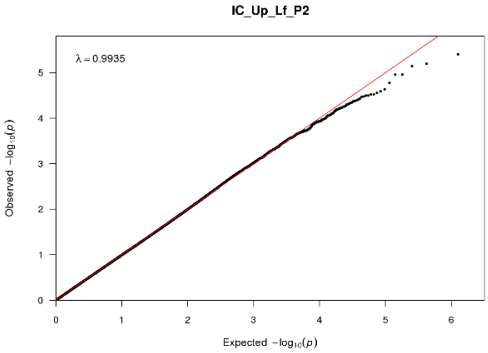

IC_Up_Rt_P2
λ = 0.9986
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

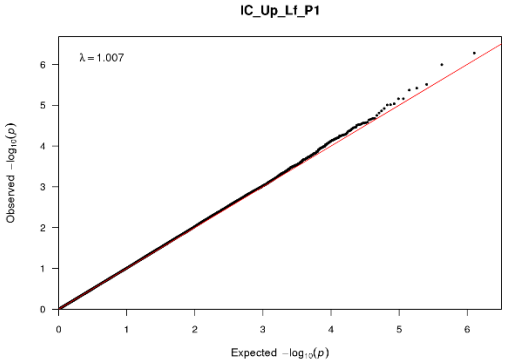

IC_Up_Rt_P1
λ = 0.9972
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

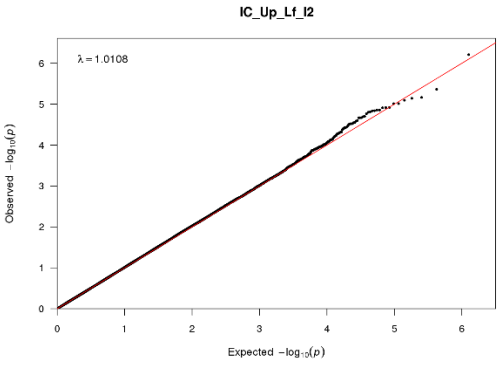

IC_Up_Rt_I2
λ = 1.0094
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

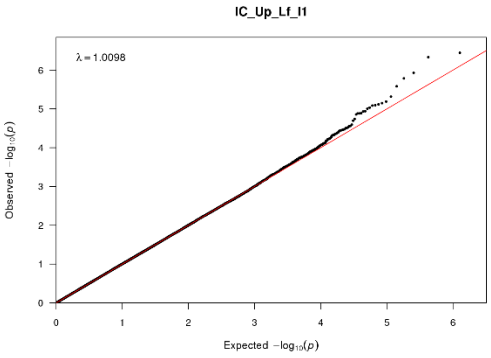

IC_Up_Rt_I1
λ = 1.0056
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

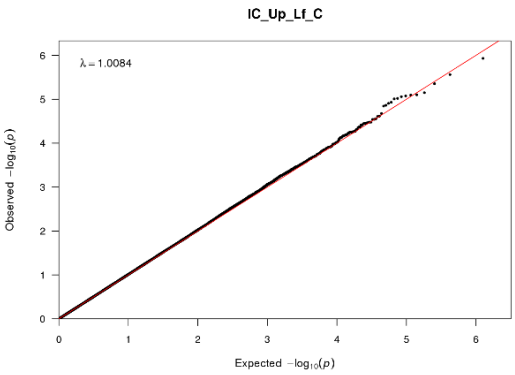

IC_Up_Rt_C
λ = 0.9963
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

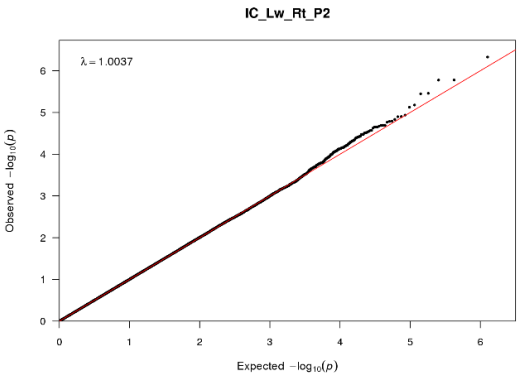

IC_Up_Lf_P2
λ = 0.9935
Observed −log10(p)
Expected −log10(p)

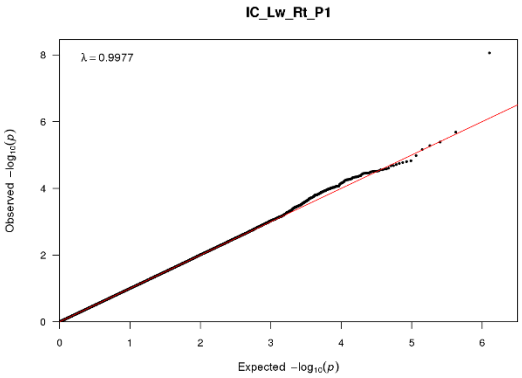

IC_Up_Lf_P1
λ = 1.007
Observed −log10(p)
Expected −log10(p)

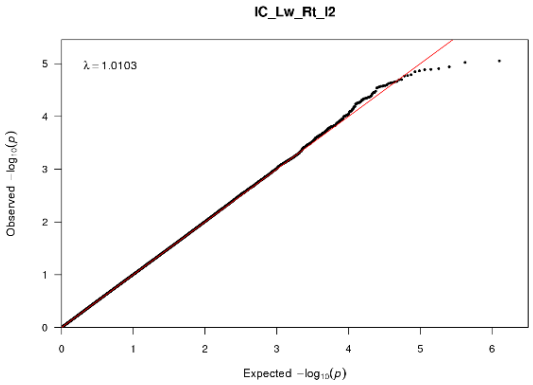

IC_Up_Lf_I2
λ = 1.0108
Observed −log10(p)
Expected −log10(p)

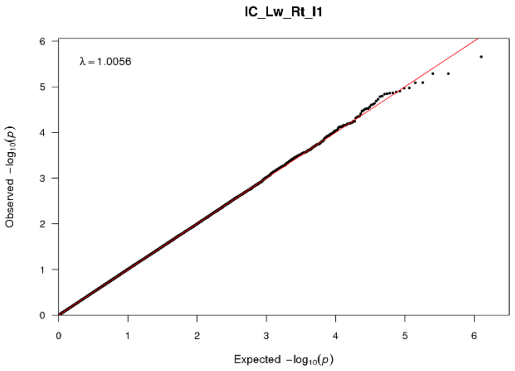

IC_Up_Lf_I1
λ = 1.0098
Observed −log10(p)
Expected −log10(p)

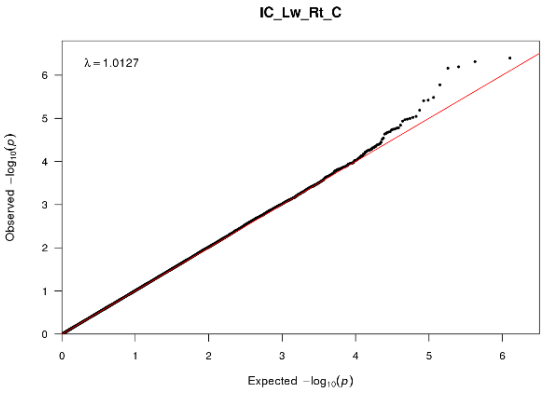

IC_Up_Lf_C
λ = 1.0084
Observed −log10(p)
Expected −log10(p)

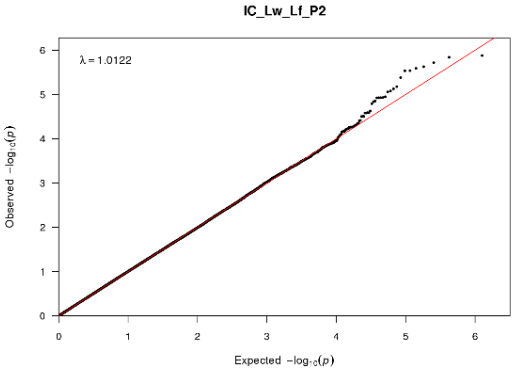

IC_Lw_Rt_P2
λ = 1.0037
Observed −log10(p)
Expected −log10(p)

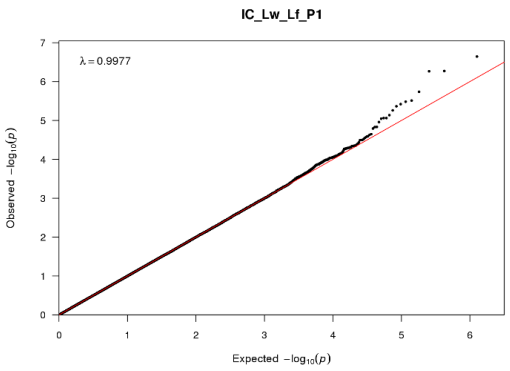

IC_Lw_Rt_P1
λ = 0.9977
Observed −log10(p)
Expected −log10(p)

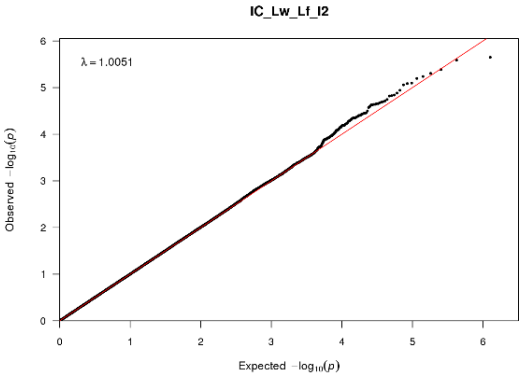

IC_Lw_Rt_I2
λ = 1.0103
Observed −log10(p)
Expected −log10(p)

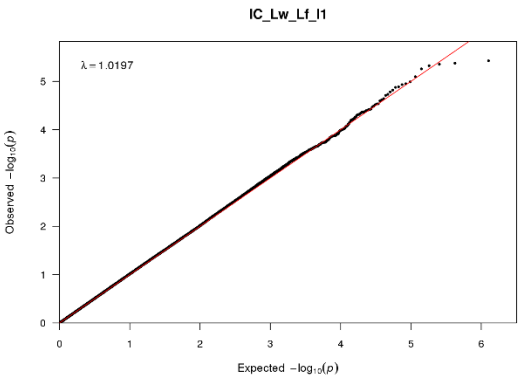

IC_Lw_Rt_I1
λ = 1.0056
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

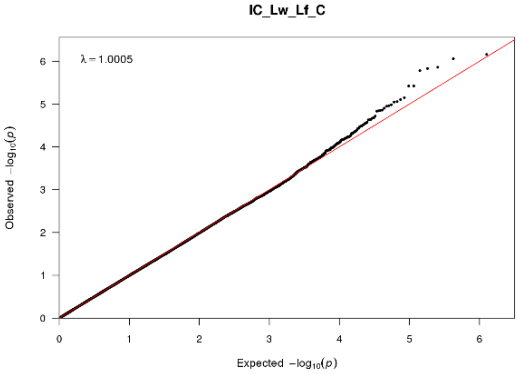

IC_Lw_Rt_C
λ = 1.0127
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

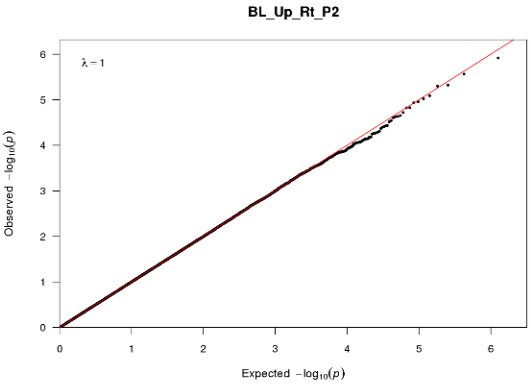

IC_Lw_Lf_P2
λ = 1.0122
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

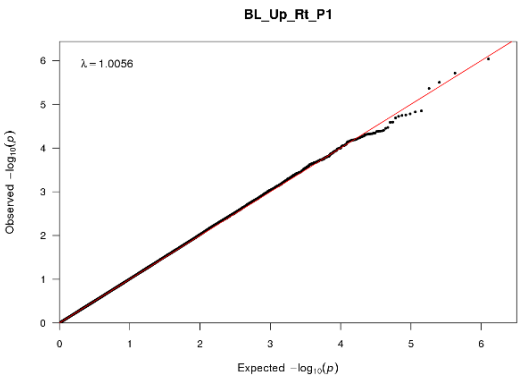

IC_Lw_Lf_P1
λ = 0.9977
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

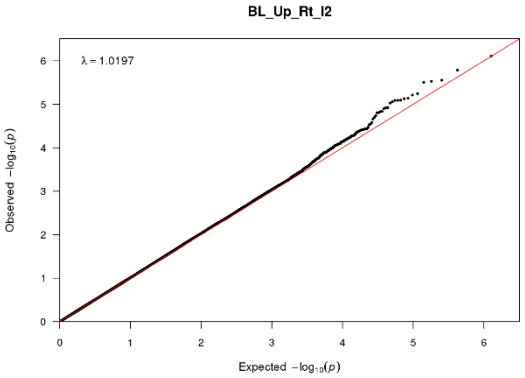

IC_Lw_Lf_I2
λ = 1.0051
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

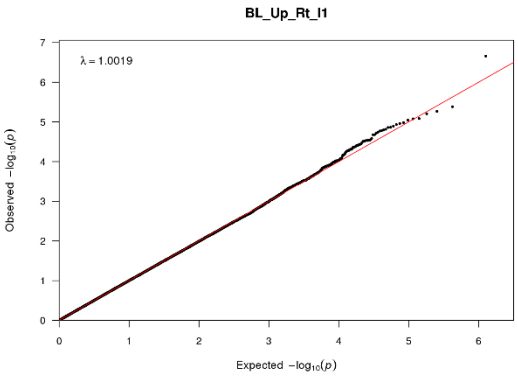

IC_Lw_Lf_I1
λ = 1.0197
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

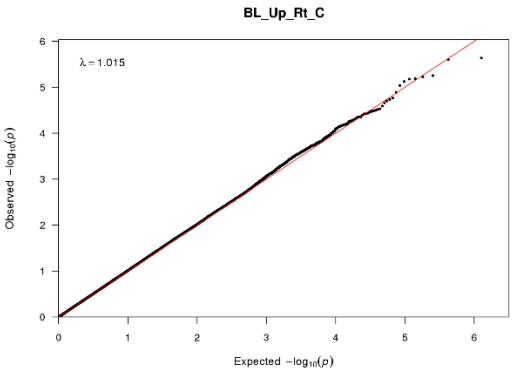

IC_Lw_Lf_C
λ = 1.0005
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

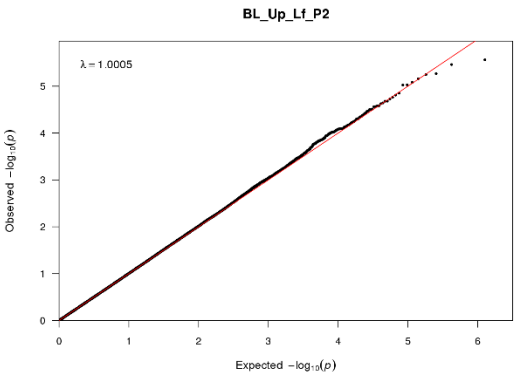

BL_Up_Rt_P2
λ = 1
Observed $-\log_{10}(p)$
Expected $-\log_{10}(p)$

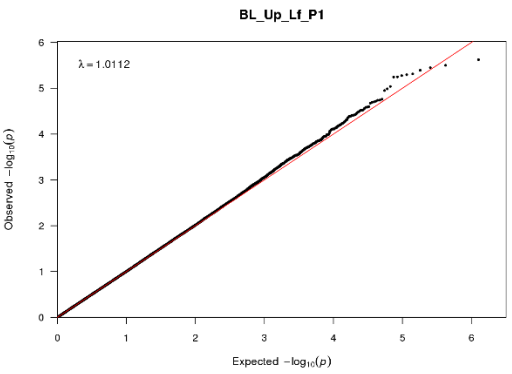

BL_Up_Rt_P1
λ = 1.0056
Observed −log10(p)
Expected −log10(p)

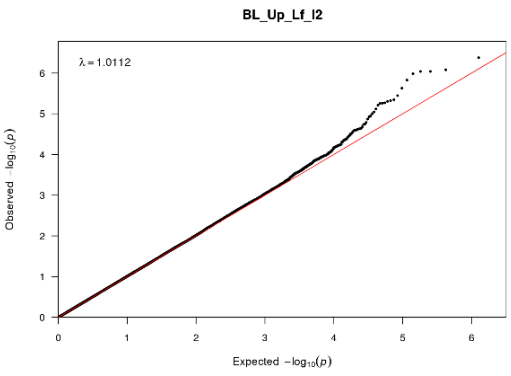

BL_Up_Rt_I2
λ = 1.0197
Observed −log10(p)
Expected −log10(p)

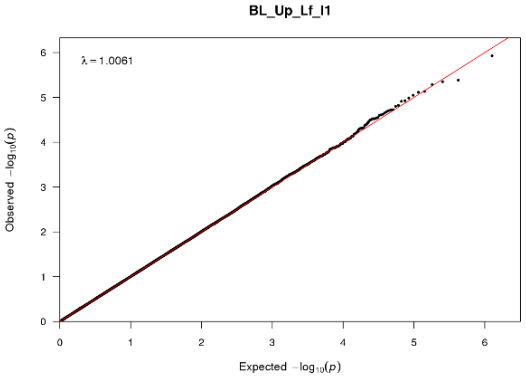

BL_Up_Rt_I1
λ = 1.0019
Observed −log10(p)
Expected −log10(p)

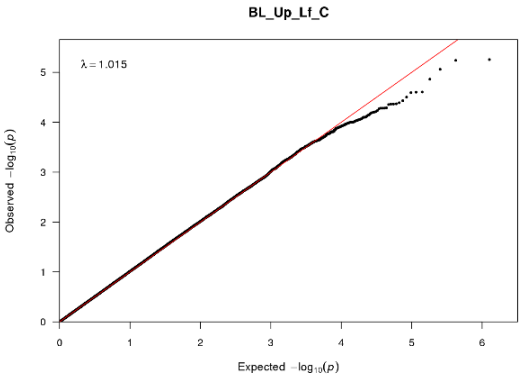

BL_Up_Rt_C
λ = 1.015
Observed −log10(p)
Expected −log10(p)

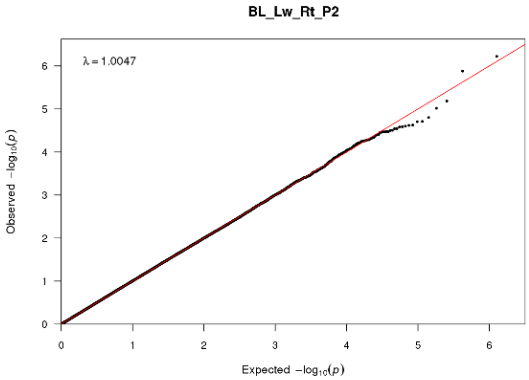

BL_Up_Lf_P2
λ = 1.0005
Observed −log10(p)
Expected −log10(p)

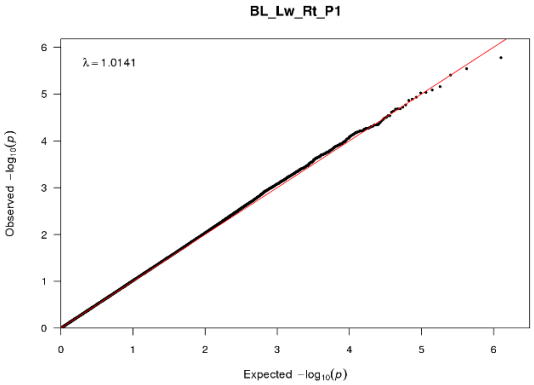

BL_Up_Lf_P1
λ = 1.0112
Observed −log10(p)
Expected −log10(p)

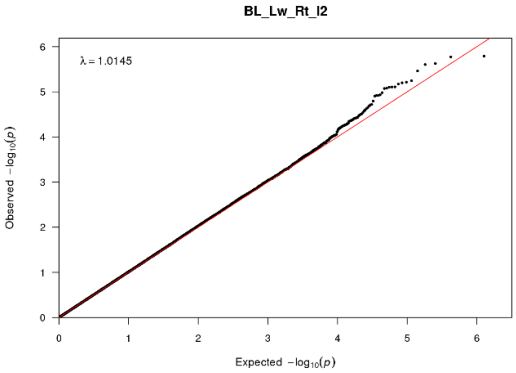

BL_Up_Lf_I2
λ = 1.0112
Observed −log10(p)
Expected −log10(p)

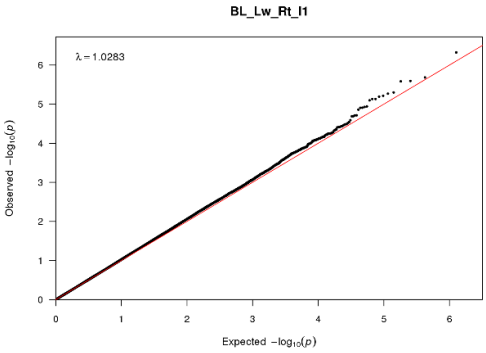

BL_Up_Lf_I1
λ = 1.0061
Observed −log10(p)
Expected −log10(p)

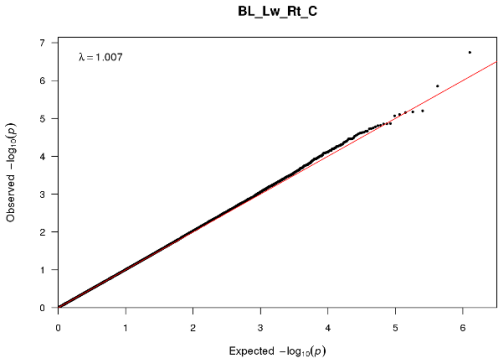

BL_Up_Lf_C
λ = 1.015
Observed −log10(p)
Expected −log10(p)

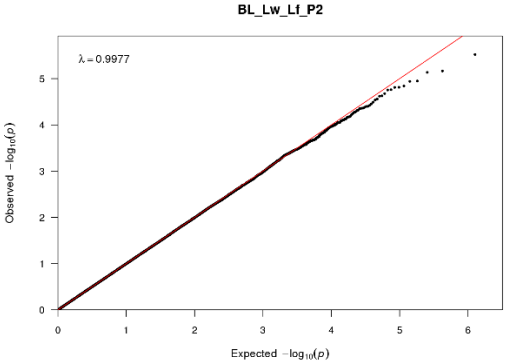

BL_Lw_Rt_P2
λ = 1.0047
Observed −log10(p)
Expected −log10(p)

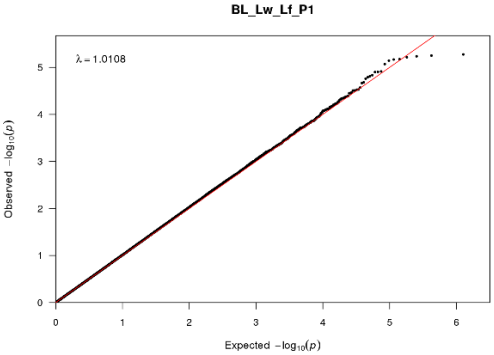

BL_Lw_Rt_P1
λ = 1.0141
Observed −log10(p)
Expected −log10(p)

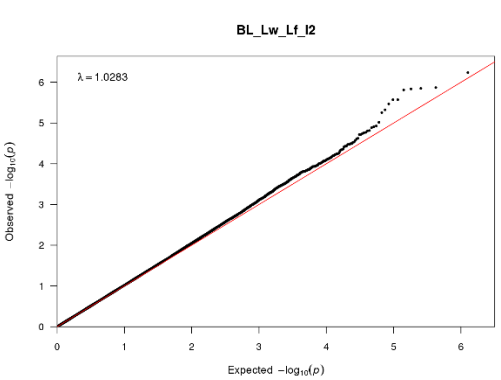

BL_Lw_Rt_I2
λ = 1.0145
Observed −log10(p)
Expected −log10(p)

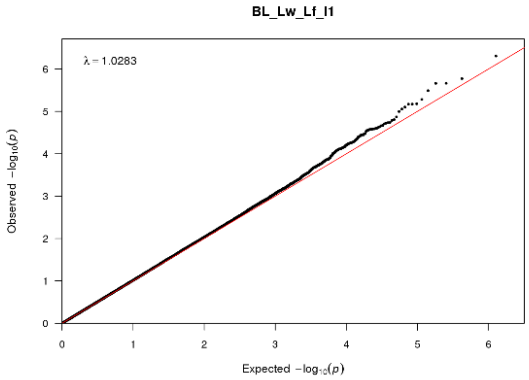

BL_Lw_Rt_I1
λ = 1.0283
Observed −log10(p)
Expected −log10(p)

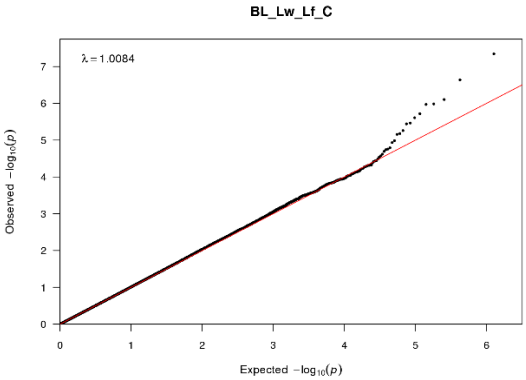

BL_Lw_Rt_C
λ = 1.007
Observed −log10(p)
Expected −log10(p)


BL_Lw_Lf_P2
λ = 0.9977
Observed −log10(p)
Expected −log10(p)


BL_Lw_Lf_P1
λ = 1.0108
Observed −log10(p)
Expected −log10(p)


BL_Lw_Lf_I2
λ = 1.0283
Observed −log10(p)
Expected −log10(p)


BL_Lw_Lf_I1
λ = 1.0283
Observed −log10(p)
Expected −log10(p)


BL_Lw_Lf_C
λ = 1.0084
Observed −log10(p)
Expected −log10(p)

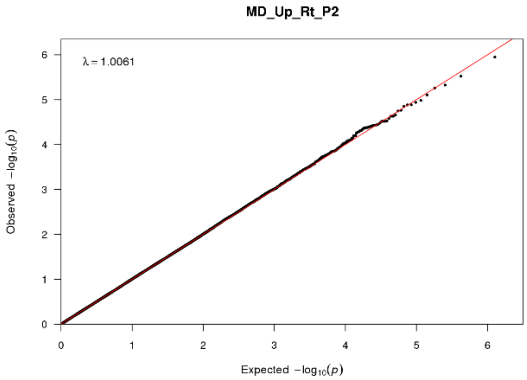

MD_Up_Rt_P2
λ = 1.0061
Observed −log10(p)
Expected −log10(p)

Table B.2a. Correlations between 46 ordinal dental traits (ASUDAS) and 60 dental measurements. Correlation values are presented in Table B.2a, with corresponding *P* values in Table B.2b). Correlations with significant *P* values (<0.00002, Bonferroni-adjusted threshold), are highlighted in bold.

	SSUI1	SSUI2	DSUI1	DSUI2	WINUI1	LCUI2	TDUI1	TDUI2	TDUC	MRUC	DARUC	MetUM1	MetUM2	HipUM1	HipUM2
MD_Up_Rt_P2	8.89	9.33	12.16	6.92	2.20	-3.02	11.96	10.93	11.80	5.15	7.90	1.30	-0.02	-0.99	10.87
BL_Up_Lf_P2	-2.61	-1.05	-2.67	-0.90	8.18	1.27	8.96	9.13	16.43	0.77	9.69	-2.71	-1.39	0.93	4.71
BL_Up_Lf_P1	-0.66	-0.08	-0.47	1.38	6.68	3.89	9.25	5.55	11.62	2.54	10.35	-3.89	-5.71	1.15	4.11
BL_Up_Lf_C	0.57	5.65	-1.23	-1.31	12.83	-1.36	11.93	13.40	27.49	8.48	7.33	-3.77	3.00	-4.57	6.58
BL_Up_Lf_I2	6.04	10.28	8.38	4.57	6.09	1.86	16.70	19.56	20.00	4.64	6.38	0.93	-2.94	1.60	4.40
BL_Up_Lf_I1	7.78	12.46	4.26	3.46	16.39	3.12	23.30	13.92	18.70	9.13	7.69	0.28	-0.26	1.51	6.44
BL_Up_Rt_I1	7.00	12.36	5.05	4.17	12.43	1.01	21.85	15.89	22.81	10.77	6.16	1.06	-2.07	0.20	8.87
BL_Up_Rt_I2	3.44	7.20	7.97	5.85	7.16	4.34	17.88	19.77	22.34	2.86	5.79	3.66	-1.04	2.43	5.48
BL_Up_Rt_C	1.74	3.76	-1.45	-0.43	9.83	3.38	11.89	13.47	27.45	7.32	5.03	-4.83	2.17	-5.02	6.28
BL_Up_Rt_P1	-1.57	-1.28	-3.27	-0.20	3.55	0.65	5.69	4.20	11.41	0.15	13.19	-5.15	-2.81	3.17	5.21
BL_Up_Rt_P2	-1.88	-0.39	-3.96	-1.64	6.34	-0.70	7.57	11.59	18.29	1.90	8.40	-0.66	-0.77	3.35	2.99
IC_Lw_Rt_P2	1.00	-0.01	0.51	-2.81	6.15	1.69	5.57	8.71	13.70	3.79	2.50	-7.41	-9.77	-3.37	-1.81
IC_Lw_Rt_P1	-2.74	-2.13	-2.24	-2.89	6.79	5.93	7.10	7.59	18.26	5.42	8.00	-2.14	-10.29	0.69	-6.10
IC_Lw_Rt_C	-6.28	-5.11	-4.09	-5.65	11.18	6.85	5.65	6.13	12.47	2.69	4.85	-5.28	-3.25	-6.55	-1.22
IC_Lw_Rt_I2	-0.65	-0.88	0.49	0.53	7.67	7.93	9.18	9.82	12.63	2.90	6.85	0.33	-6.10	-1.16	1.59
IC_Lw_Rt_I1	-3.30	-2.16	-1.96	0.00	3.86	9.04	7.96	8.91	7.38	4.06	6.08	0.61	0.75	-2.68	0.32
IC_Lw_Lf_I1	-2.79	-2.34	0.79	1.85	6.37	7.10	9.35	8.83	6.77	1.80	4.79	2.60	-0.59	-3.69	-0.01
IC_Lw_Lf_I2	-0.60	0.48	-0.16	-1.39	4.43	8.60	10.09	12.75	17.26	5.70	6.28	0.34	-3.13	-0.22	2.41
IC_Lw_Lf_C	-4.62	-2.56	-2.46	-1.99	5.83	8.13	7.58	4.75	12.64	5.37	4.62	-5.41	-2.00	-2.33	0.53
IC_Lw_Lf_P1	-4.88	-5.30	-3.23	-2.29	7.57	3.81	11.47	6.16	16.84	6.96	11.77	-1.52	-9.75	1.84	-6.29

Continue…

	C5UM1	C5UM2	CarUM1	SSLI1	SSLI2	DSLI1	DSLI2	DARLC	LCVLP1	LCVLP2	AFLM1	AFLM2	GPLM2	CNLM1	CNLM2	C5LM1
MD_Up_Rt_P2	2.81	4.97	10.26	11.48	7.80	4.68	1.67	1.91	1.83	4.21	2.15	-7.97	1.73	-1.17	6.51	-5.09
BL_Up_Lf_P2	9.09	5.86	16.54	4.85	3.72	5.63	4.45	8.86	-0.18	6.76	-2.58	-1.07	2.17	1.96	3.90	-6.75
BL_Up_Lf_P1	11.21	7.31	15.99	4.77	4.68	7.35	3.26	2.71	2.02	3.17	-5.43	-3.65	-0.18	2.38	-0.73	-3.72
BL_Up_Lf_C	3.08	-0.15	15.53	3.62	5.13	5.18	1.71	7.25	5.98	4.73	-3.33	-5.72	-1.09	1.41	6.08	-5.09
BL_Up_Lf_I2	9.09	3.69	13.64	10.43	8.62	4.19	-1.26	3.15	3.27	8.59	-10.22	-1.79	-3.60	1.79	7.81	-4.53
BL_Up_Lf_I1	6.51	-2.94	18.70	4.46	5.12	6.22	0.77	3.53	9.80	12.39	-7.11	-6.20	-3.57	-0.25	4.81	-5.04
BL_Up_Rt_I1	6.90	-2.75	18.52	5.55	6.37	5.37	-0.96	3.46	9.41	13.03	-7.52	-6.23	-4.12	1.03	3.76	-6.80
BL_Up_Rt_I2	10.25	3.24	16.37	11.90	11.07	4.09	0.65	10.25	1.70	9.92	-6.11	0.70	-7.04	1.09	8.64	-0.08
BL_Up_Rt_C	4.91	-1.14	12.58	0.57	3.96	6.41	2.52	8.34	3.17	3.82	-2.69	-7.96	-0.79	4.84	3.58	-5.00
BL_Up_Rt_P1	11.82	4.06	15.89	6.50	5.70	2.46	-1.42	2.16	0.68	3.04	-6.80	-6.02	-0.23	3.36	-0.74	0.00
BL_Up_Rt_P2	11.47	8.52	18.69	5.74	4.40	7.81	2.90	6.79	1.61	6.03	-3.60	-2.84	0.64	3.50	5.50	-0.74
IC_Lw_Rt_P2	-8.15	-7.63	14.95	-5.14	-3.35	1.63	-2.57	4.16	0.58	1.10	-9.44	-8.57	1.95	-4.02	-5.95	-8.02
IC_Lw_Rt_P1	0.46	-5.73	14.98	-3.09	-0.72	3.47	1.32	7.25	1.71	0.84	-1.35	-7.15	-6.16	3.59	0.47	-3.41
IC_Lw_Rt_C	-0.29	-3.34	7.09	-3.89	-1.24	1.74	0.05	8.82	-1.71	6.09	-8.67	-12.16	-3.26	1.79	-1.78	-5.19
IC_Lw_Rt_I2	2.87	-2.41	6.89	5.40	7.28	3.79	1.32	5.34	-2.17	2.42	-9.91	-14.15	-0.49	-0.45	-4.94	-4.33
IC_Lw_Rt_I1	-1.56	-8.70	9.27	4.70	6.09	5.43	2.13	4.36	-2.52	-1.30	-7.86	-10.23	-2.00	-2.74	-7.33	-7.34
IC_Lw_Lf_I1	-3.18	-9.93	8.90	3.43	3.65	5.22	2.93	5.11	-5.98	-0.89	-9.35	-12.53	-3.49	-2.66	-2.65	-4.84
IC_Lw_Lf_I2	-0.82	-5.68	10.51	3.58	5.64	1.54	-1.19	4.88	-3.13	4.66	-10.29	-12.92	0.09	-4.13	-5.48	-5.66
IC_Lw_Lf_C	1.28	-3.18	10.71	-1.08	2.06	4.46	1.54	6.52	-2.44	6.03	-8.22	-9.28	-3.03	-0.77	-4.88	-2.94
IC_Lw_Lf_P1	4.20	-6.55	9.55	-1.69	1.09	4.63	3.04	1.53	-1.56	-2.14	-2.54	-3.09	1.31	-1.95	-8.71	-7.53

Continue…

	C5LM2	C6LM1	C7LM1	DWLM1	PrtostLM1	PrtostLM2	IGUI1	IGUI2	ARUP1	ARUP2	AMTUP1	AMTUP2	3CUM2	ARPrLP1	ARPrLP2
IC_Up_Rt_P2	3.15	-0.05	-1.88	-5.73	10.61	13.27	7.17	7.43	-4.91	-6.88	1.80	-6.58	6.26	1.10	6.86
IC_Up_Rt_P1	5.90	-2.75	-0.82	-8.01	4.72	17.15	9.60	10.42	-2.46	-6.01	7.31	1.14	5.35	-0.88	5.21
IC_Up_Rt_C	1.99	6.58	2.41	-4.84	8.00	12.24	9.30	10.67	0.11	8.75	-2.57	-0.53	-4.38	9.41	10.37
IC_Up_Rt_I2	3.58	6.22	3.73	-3.12	4.67	14.71	8.55	3.64	0.66	3.46	-0.98	-3.67	-0.95	12.03	14.44
IC_Up_Rt_I1	-1.28	3.41	3.87	1.02	4.78	11.15	10.78	10.16	6.43	14.60	-0.27	0.74	-3.55	11.56	15.95
IC_Up_Lf_I1	0.90	5.63	3.50	-1.59	5.01	10.02	12.33	10.45	4.70	12.10	-0.55	-1.97	-3.58	9.21	12.01
IC_Up_Lf_I2	2.82	1.13	-0.20	-2.84	6.63	9.48	12.08	8.27	5.43	7.74	2.35	-0.65	0.73	6.54	10.42
IC_Up_Lf_C	-0.99	1.73	2.71	-6.59	8.55	9.92	11.62	8.57	1.75	5.54	-0.55	-0.14	-0.53	5.43	9.80
IC_Up_Lf_P1	3.72	-1.55	-0.28	-8.54	9.38	14.63	8.36	10.62	-0.54	-2.01	6.17	1.11	5.46	-2.70	7.00
IC_Up_Lf_P2	5.99	-2.45	-0.59	-9.33	7.30	13.12	6.22	6.15	-0.01	0.48	0.14	-2.36	4.01	-0.34	11.65
MD_Up_Lf_P2	9.15	0.58	-7.73	2.70	9.10	11.41	-1.01	2.70	12.09	13.08	-1.94	-4.53	-17.24	9.76	15.00
MD_Up_Lf_P1	7.62	5.78	-1.17	-2.81	12.10	9.28	4.29	1.47	10.19	15.54	3.95	3.25	-10.91	2.28	8.93
MD_Up_Lf_C	7.86	-0.64	3.47	-8.38	11.20	14.03	4.12	3.68	13.41	11.79	4.38	-4.37	-10.14	7.69	5.35
MD_Up_Lf_I2	19.45	8.11	3.12	2.54	10.65	10.57	7.17	11.95	10.53	3.99	7.02	4.77	-8.74	13.98	10.22
MD_Up_Lf_I1	12.93	13.61	3.29	-5.35	14.77	15.11	4.44	8.03	11.12	8.02	7.73	6.01	-9.67	13.79	13.60
MD_Up_Rt_I1	11.92	7.75	4.05	-6.82	16.20	12.93	7.64	11.67	9.55	9.03	7.49	8.21	-11.15	12.71	11.78
MD_Up_Rt_I2	17.28	2.30	0.66	-1.71	11.49	9.08	7.68	9.37	7.79	5.51	9.53	1.29	-12.59	13.06	13.99
MD_Up_Rt_C	10.29	3.08	4.19	-5.02	7.47	13.35	9.89	2.39	10.32	15.08	1.64	-2.19	-14.85	11.03	12.41
MD_Up_Rt_P1	6.66	1.53	-0.24	-2.57	11.27	7.07	1.11	-0.05	6.59	14.02	5.54	8.76	-10.02	4.47	9.70

Continue…

	SSUI1	SSUI2	DSUI1	DSUI2	WINUI1	LCUI2	TDUI1	TDUI2	TDUC	MRUC	DARUC	MetUM1	MetUM2	HipUM1	HipUM2
MD_Up_Rt_P2	8.89	9.33	12.16	6.92	2.20	-3.02	11.96	10.93	11.80	5.15	7.90	1.30	-0.02	-0.99	10.87
BL_Up_Lf_P2	-2.61	-1.05	-2.67	-0.90	8.18	1.27	8.96	9.13	16.43	0.77	9.69	-2.71	-1.39	0.93	4.71
BL_Up_Lf_P1	-0.66	-0.08	-0.47	1.38	6.68	3.89	9.25	5.55	11.62	2.54	10.35	-3.89	-5.71	1.15	4.11
BL_Up_Lf_C	0.57	5.65	-1.23	-1.31	12.83	-1.36	11.93	13.40	27.49	8.48	7.33	-3.77	3.00	-4.57	6.58
BL_Up_Lf_I2	6.04	10.28	8.38	4.57	6.09	1.86	16.70	19.56	20.00	4.64	6.38	0.93	-2.94	1.60	4.40
BL_Up_Lf_I1	7.78	12.46	4.26	3.46	16.39	3.12	23.30	13.92	18.70	9.13	7.69	0.28	-0.26	1.51	6.44
BL_Up_Rt_I1	7.00	12.36	5.05	4.17	12.43	1.01	21.85	15.89	22.81	10.77	6.16	1.06	-2.07	0.20	8.87
BL_Up_Rt_I2	3.44	7.20	7.97	5.85	7.16	4.34	17.88	19.77	22.34	2.86	5.79	3.66	-1.04	2.43	5.48
BL_Up_Rt_C	1.74	3.76	-1.45	-0.43	9.83	3.38	11.89	13.47	27.45	7.32	5.03	-4.83	2.17	-5.02	6.28
BL_Up_Rt_P1	-1.57	-1.28	-3.27	-0.20	3.55	0.65	5.69	4.20	11.41	0.15	13.19	-5.15	-2.81	3.17	5.21
BL_Up_Rt_P2	-1.88	-0.39	-3.96	-1.64	6.34	-0.70	7.57	11.59	18.29	1.90	8.40	-0.66	-0.77	3.35	2.99
IC_Lw_Rt_P2	1.00	-0.01	0.51	-2.81	6.15	1.69	5.57	8.71	13.70	3.79	2.50	-7.41	-9.77	-3.37	-1.81
IC_Lw_Rt_P1	-2.74	-2.13	-2.24	-2.89	6.79	5.93	7.10	7.59	18.26	5.42	8.00	-2.14	-10.29	0.69	-6.10
IC_Lw_Rt_C	-6.28	-5.11	-4.09	-5.65	11.18	6.85	5.65	6.13	12.47	2.69	4.85	-5.28	-3.25	-6.55	-1.22
IC_Lw_Rt_I2	-0.65	-0.88	0.49	0.53	7.67	7.93	9.18	9.82	12.63	2.90	6.85	0.33	-6.10	-1.16	1.59
IC_Lw_Rt_I1	-3.30	-2.16	-1.96	0.00	3.86	9.04	7.96	8.91	7.38	4.06	6.08	0.61	0.75	-2.68	0.32
IC_Lw_Lf_I1	-2.79	-2.34	0.79	1.85	6.37	7.10	9.35	8.83	6.77	1.80	4.79	2.60	-0.59	-3.69	-0.01
IC_Lw_Lf_I2	-0.60	0.48	-0.16	-1.39	4.43	8.60	10.09	12.75	17.26	5.70	6.28	0.34	-3.13	-0.22	2.41
IC_Lw_Lf_C	-4.62	-2.56	-2.46	-1.99	5.83	8.13	7.58	4.75	12.64	5.37	4.62	-5.41	-2.00	-2.33	0.53
IC_Lw_Lf_P1	-4.88	-5.30	-3.23	-2.29	7.57	3.81	11.47	6.16	16.84	6.96	11.77	-1.52	-9.75	1.84	-6.29

Continue…

	C5UM1	C5UM2	CarUM1	SSLI1	SSLI2	DSLI1	DSLI2	DARLC	LCVLP1	LCVLP2	AFLM1	AFLM2	GPLM2	CNLM1	CNLM2	C5LM1
MD_Up_Rt_P2	2.81	4.97	10.26	11.48	7.80	4.68	1.67	1.91	1.83	4.21	2.15	-7.97	1.73	-1.17	6.51	-5.09
BL_Up_Lf_P2	9.09	5.86	16.54	4.85	3.72	5.63	4.45	8.86	-0.18	6.76	-2.58	-1.07	2.17	1.96	3.90	-6.75
BL_Up_Lf_P1	11.21	7.31	15.99	4.77	4.68	7.35	3.26	2.71	2.02	3.17	-5.43	-3.65	-0.18	2.38	-0.73	-3.72
BL_Up_Lf_C	3.08	-0.15	15.53	3.62	5.13	5.18	1.71	7.25	5.98	4.73	-3.33	-5.72	-1.09	1.41	6.08	-5.09
BL_Up_Lf_I2	9.09	3.69	13.64	10.43	8.62	4.19	-1.26	3.15	3.27	8.59	-10.22	-1.79	-3.60	1.79	7.81	-4.53
BL_Up_Lf_I1	6.51	-2.94	18.70	4.46	5.12	6.22	0.77	3.53	9.80	12.39	-7.11	-6.20	-3.57	-0.25	4.81	-5.04
BL_Up_Rt_I1	6.90	-2.75	18.52	5.55	6.37	5.37	-0.96	3.46	9.41	13.03	-7.52	-6.23	-4.12	1.03	3.76	-6.80
BL_Up_Rt_I2	10.25	3.24	16.37	11.90	11.07	4.09	0.65	10.25	1.70	9.92	-6.11	0.70	-7.04	1.09	8.64	-0.08
BL_Up_Rt_C	4.91	-1.14	12.58	0.57	3.96	6.41	2.52	8.34	3.17	3.82	-2.69	-7.96	-0.79	4.84	3.58	-5.00
BL_Up_Rt_P1	11.82	4.06	15.89	6.50	5.70	2.46	-1.42	2.16	0.68	3.04	-6.80	-6.02	-0.23	3.36	-0.74	0.00
BL_Up_Rt_P2	11.47	8.52	18.69	5.74	4.40	7.81	2.90	6.79	1.61	6.03	-3.60	-2.84	0.64	3.50	5.50	-0.74
IC_Lw_Rt_P2	-8.15	-7.63	14.95	-5.14	-3.35	1.63	-2.57	4.16	0.58	1.10	-9.44	-8.57	1.95	-4.02	-5.95	-8.02
IC_Lw_Rt_P1	0.46	-5.73	14.98	-3.09	-0.72	3.47	1.32	7.25	1.71	0.84	-1.35	-7.15	-6.16	3.59	0.47	-3.41
IC_Lw_Rt_C	-0.29	-3.34	7.09	-3.89	-1.24	1.74	0.05	8.82	-1.71	6.09	-8.67	-12.16	-3.26	1.79	-1.78	-5.19
IC_Lw_Rt_I2	2.87	-2.41	6.89	5.40	7.28	3.79	1.32	5.34	-2.17	2.42	-9.91	-14.15	-0.49	-0.45	-4.94	-4.33
IC_Lw_Rt_I1	-1.56	-8.70	9.27	4.70	6.09	5.43	2.13	4.36	-2.52	-1.30	-7.86	-10.23	-2.00	-2.74	-7.33	-7.34
IC_Lw_Lf_I1	-3.18	-9.93	8.90	3.43	3.65	5.22	2.93	5.11	-5.98	-0.89	-9.35	-12.53	-3.49	-2.66	-2.65	-4.84
IC_Lw_Lf_I2	-0.82	-5.68	10.51	3.58	5.64	1.54	-1.19	4.88	-3.13	4.66	-10.29	-12.92	0.09	-4.13	-5.48	-5.66
IC_Lw_Lf_C	1.28	-3.18	10.71	-1.08	2.06	4.46	1.54	6.52	-2.44	6.03	-8.22	-9.28	-3.03	-0.77	-4.88	-2.94
IC_Lw_Lf_P1	4.20	-6.55	9.55	-1.69	1.09	4.63	3.04	1.53	-1.56	-2.14	-2.54	-3.09	1.31	-1.95	-8.71	-7.53

Continue…

	C5LM2	C6LM1	C7LM1	DWLM1	PrtostLM1	PrtostLM2	IGUI1	IGUI2	ARUP1	ARUP2	AMTUP1	AMTUP2	3CUM2	ARPrLP1	ARPrLP2
MD_Up_Rt_P2	8.69	1.33	-4.91	1.63	11.54	9.15	-0.38	0.74	2.96	13.55	-1.83	0.71	-14.43	6.78	13.38
BL_Up_Lf_P2	4.31	8.14	1.78	1.14	14.07	12.66	4.18	-2.11	16.57	14.48	5.37	10.44	-8.91	10.69	12.92
BL_Up_Lf_P1	1.22	7.98	0.03	-0.10	8.47	12.00	7.66	0.52	13.73	9.46	10.45	11.59	-7.92	10.08	9.76
BL_Up_Lf_C	5.32	6.13	1.23	-7.79	14.38	12.16	8.01	10.50	5.30	3.18	1.99	2.88	-8.30	2.35	6.35
BL_Up_Lf_I2	8.49	3.02	5.68	-3.37	11.77	6.90	7.38	7.74	11.64	9.61	8.52	4.48	-6.25	10.64	8.95
BL_Up_Lf_I1	4.46	3.63	2.73	-4.72	10.97	5.36	15.05	8.30	10.92	6.62	5.92	6.87	-10.14	10.61	10.80
BL_Up_Rt_I1	4.35	3.30	5.64	-5.13	12.67	6.06	9.83	8.74	7.62	8.00	8.63	6.49	-9.25	10.97	10.30
BL_Up_Rt_I2	10.89	2.53	4.67	1.42	8.85	8.85	11.71	9.07	13.82	10.64	3.89	4.08	-7.03	14.14	15.12
BL_Up_Rt_C	1.46	9.70	5.63	-5.08	11.50	10.77	8.88	7.37	7.97	4.07	1.17	6.68	-6.99	4.38	3.76
BL_Up_Rt_P1	0.10	6.36	0.20	-3.80	7.41	7.91	5.82	0.37	12.03	9.44	5.42	12.69	-6.90	9.82	10.54
BL_Up_Rt_P2	4.20	7.05	1.75	-0.43	14.71	13.61	7.27	-1.58	15.59	14.48	10.09	8.55	-10.23	9.78	10.48
IC_Lw_Rt_P2	-0.22	-1.59	0.71	-9.09	10.73	10.04	3.46	6.51	-0.57	-1.45	-2.94	0.79	0.72	-4.75	4.90
IC_Lw_Rt_P1	0.71	3.14	6.15	-3.96	9.12	10.93	3.18	7.95	-0.11	1.88	5.04	5.67	5.65	0.37	6.34
IC_Lw_Rt_C	-0.39	2.54	9.38	-5.03	5.14	5.87	6.69	5.49	1.34	7.16	2.35	1.01	2.91	5.88	12.46
IC_Lw_Rt_I2	-3.64	-0.50	3.15	-8.32	4.22	8.09	7.62	6.56	-0.67	6.42	-2.11	-2.21	-1.55	7.07	14.10
IC_Lw_Rt_I1	-3.77	-0.14	4.80	-8.49	3.13	8.61	5.99	5.30	-3.84	2.62	-2.46	-0.91	1.52	1.28	7.44
IC_Lw_Lf_I1	1.47	-0.80	0.57	#####	7.08	6.26	3.20	2.18	-4.53	5.26	-1.29	1.46	-0.52	-0.39	7.86
IC_Lw_Lf_I2	-1.61	-3.21	1.10	-8.82	6.66	8.50	6.46	6.96	-0.50	6.34	0.25	2.47	-0.83	1.94	10.92
IC_Lw_Lf_C	0.09	0.89	2.13	-4.34	10.59	11.15	5.94	8.61	1.04	5.43	1.75	4.50	1.71	1.98	12.50
IC_Lw_Lf_P1	-5.03	-5.01	4.35	-2.15	10.23	9.90	1.50	5.08	4.52	1.11	4.57	1.36	5.50	3.98	8.34

Continue…

	SSUI1	SSUI2	DSUI1	DSUI2	WINUI1	LCUI2	TDUI1	TDUI2	TDUC	MRUC	DARUC	MetUM1	MetUM2	HipUM1	HipUM2
IC_Lw_Lf_P2	-2.37	-1.27	-2.80	-4.44	10.75	-0.24	6.62	4.85	8.31	3.45	2.35	-3.38	-6.00	-3.94	-2.97
MD_Lw_Lf_P2	14.27	14.25	8.59	7.64	2.26	0.92	10.05	8.39	11.66	9.42	14.39	0.03	0.31	2.59	8.68
MD_Lw_Lf_P1	8.69	8.28	2.67	3.77	3.39	5.69	7.04	11.42	15.21	4.91	15.70	0.91	-2.45	0.63	8.16
MD_Lw_Lf_C	6.27	9.48	5.20	5.40	3.15	4.11	11.98	9.17	14.67	8.59	3.00	-3.91	4.56	4.57	7.58
MD_Lw_Lf_I2	22.12	25.08	14.71	16.64	-1.55	8.60	23.78	19.90	21.28	7.85	8.44	10.18	6.44	3.53	12.29
MD_Lw_Lf_I1	16.07	17.62	8.74	12.25	4.95	5.90	21.52	15.30	18.98	5.28	7.90	1.25	7.28	1.03	8.88
MD_Lw_Rt_I1	20.05	21.75	15.41	17.54	4.73	7.65	22.78	17.31	22.22	5.86	10.00	2.45	6.13	3.35	10.13
MD_Lw_Rt_I2	23.19	25.17	14.08	17.03	3.97	3.16	24.64	22.38	21.52	11.62	8.36	7.04	0.89	4.02	10.78
MD_Lw_Rt_C	4.33	7.23	4.42	5.57	3.66	4.07	11.10	10.85	14.36	8.30	3.85	-2.42	2.34	2.83	9.13
MD_Lw_Rt_P1	3.22	5.16	0.02	-0.37	3.65	8.88	8.10	10.14	11.89	2.27	15.78	-2.44	-2.29	2.17	9.34
MD_Lw_Rt_P2	11.60	11.59	8.10	6.10	4.21	-3.87	6.39	10.59	10.49	6.68	17.06	0.18	-3.00	4.40	13.05
BL_Lw_Lf_P2	2.98	3.48	-3.52	-1.62	8.32	-0.76	13.21	8.09	11.01	3.67	11.32	1.44	1.82	5.40	5.27
BL_Lw_Lf_P1	-0.84	0.15	-7.91	-6.45	12.32	-0.93	7.91	7.77	12.67	-2.40	7.65	-1.55	3.51	2.97	8.55
BL_Lw_Lf_C	-1.58	1.59	-3.93	-2.40	9.79	3.25	10.07	8.10	15.11	0.35	6.82	-4.26	5.91	1.60	5.21
BL_Lw_Lf_I2	9.11	10.48	3.02	5.90	12.20	4.09	17.83	14.88	20.96	7.62	8.35	2.59	-0.53	4.58	7.56
BL_Lw_Lf_I1	10.10	11.24	4.94	6.85	12.67	1.97	17.60	11.21	15.69	6.35	4.66	-2.90	2.52	1.21	8.87
BL_Lw_Rt_I1	8.67	11.16	4.35	7.17	11.92	6.33	16.43	13.34	15.41	7.94	5.65	-3.74	0.66	4.31	9.28
BL_Lw_Rt_I2	4.07	5.80	3.81	5.81	8.17	1.82	15.14	12.56	17.13	8.74	7.14	-5.01	-0.87	4.27	7.00
BL_Lw_Rt_C	-2.44	0.67	-2.97	-0.49	9.71	1.94	10.09	9.14	17.10	4.60	7.30	-8.82	3.38	0.42	4.79
BL_Lw_Rt_P1	-4.39	-2.34	-7.65	-3.93	12.61	-0.28	9.53	9.71	14.83	-4.75	10.04	0.87	3.00	3.35	7.63
BL_Lw_Rt_P2	-2.82	-2.21	-4.98	-1.68	9.22	-2.80	7.96	10.47	14.43	-2.72	12.63	-3.52	0.82	4.47	7.40

Continue…

	C5UM1	C5UM2	CarUM1	SSLI1	SSLI2	DSLI1	DSLI2	DARLC	LCVLP1	LCVLP2	AFLM1	AFLM2	GPLM2	CNLM1	CNLM2	C5LM1
IC_Lw_Lf_P2	-4.27	-4.89	13.94	-0.99	1.39	3.46	-0.89	4.45	-4.00	0.24	-9.70	-11.11	-1.05	0.26	-4.53	-6.84
MD_Lw_Lf_P2	10.86	5.30	18.20	7.44	6.43	3.59	-0.33	-1.08	0.13	10.92	3.99	-4.78	-7.36	18.46	14.34	1.30
MD_Lw_Lf_P1	10.45	-1.57	16.84	8.98	8.24	9.00	5.82	7.42	-1.00	3.08	2.34	-1.56	0.59	12.26	1.96	4.16
MD_Lw_Lf_C	1.40	-2.48	15.81	6.44	7.45	11.80	7.75	15.25	-5.98	2.30	-6.02	-6.28	-4.22	11.90	8.42	-5.79
MD_Lw_Lf_I2	2.31	1.10	17.10	11.27	12.03	5.44	5.53	12.81	-0.14	7.50	-1.11	1.35	-7.12	13.95	8.65	-4.82
MD_Lw_Lf_I1	4.39	1.94	17.21	10.00	9.05	4.10	1.14	11.65	-2.32	7.23	5.45	3.97	-9.17	15.33	5.68	-2.68
MD_Lw_Rt_I1	3.12	3.01	17.06	12.61	11.77	7.22	3.79	10.56	0.40	8.09	1.86	2.28	-9.11	14.20	6.16	-3.33
MD_Lw_Rt_I2	5.10	-0.88	12.65	17.05	16.74	7.05	6.26	6.10	-0.45	7.94	-0.09	2.38	-3.12	7.82	7.55	-5.39
MD_Lw_Rt_C	2.98	0.73	18.29	7.88	10.30	9.16	6.75	13.95	-3.85	3.12	-3.99	-8.14	-0.35	10.56	4.07	-3.32
MD_Lw_Rt_P1	2.34	0.50	20.12	5.68	4.81	7.63	4.10	8.11	-4.11	5.87	1.99	-2.52	0.90	10.79	5.34	-1.74
MD_Lw_Rt_P2	8.07	6.11	18.17	7.48	5.38	2.31	-2.62	5.35	3.62	14.91	-0.87	-6.38	-3.10	17.86	12.66	3.73
BL_Lw_Lf_P2	4.39	4.61	17.97	7.16	6.54	0.07	0.96	8.87	8.15	16.20	3.59	-2.26	-2.90	7.79	7.14	-2.55
BL_Lw_Lf_P1	5.48	5.87	16.87	0.63	1.09	1.05	0.34	9.18	9.83	9.76	0.34	-1.35	0.07	10.52	5.00	-1.28
BL_Lw_Lf_C	2.35	1.29	16.00	0.07	1.05	5.81	2.89	12.17	-1.11	-0.93	-5.85	-6.39	-3.26	5.60	6.51	-2.97
BL_Lw_Lf_I2	5.51	-2.03	18.55	0.55	2.87	6.26	4.61	8.16	4.15	9.15	1.21	-4.03	-9.52	10.59	5.68	-3.48
BL_Lw_Lf_I1	2.02	-1.17	22.46	2.62	3.11	7.43	5.65	8.13	-0.09	6.81	0.55	-2.83	-10.68	14.50	7.45	-0.85
BL_Lw_Rt_I1	1.51	-1.05	22.92	0.32	2.04	5.08	1.79	9.85	1.68	7.24	3.66	-0.67	-10.28	14.81	4.16	-3.13
BL_Lw_Rt_I2	5.20	-0.93	20.25	0.83	2.21	5.39	3.69	12.86	1.61	6.13	1.77	-5.31	-10.74	14.42	6.78	-5.32
BL_Lw_Rt_C	1.57	-3.39	14.45	-0.58	1.52	5.87	2.62	10.43	1.60	-0.39	-3.46	-7.73	-2.75	7.91	4.23	-4.60
BL_Lw_Rt_P1	5.56	9.75	15.27	-2.29	-1.04	1.05	-0.39	5.31	8.38	10.45	1.42	-0.46	0.10	6.84	3.66	-2.05
BL_Lw_Rt_P2	8.11	7.12	16.28	4.28	4.44	1.42	1.30	8.94	8.07	18.10	1.48	-0.54	-2.42	9.73	3.06	1.34

Continue…

	C5LM2	C6LM1	C7LM1	DWLM1	PrtostLM1	PrtostLM2	IGUI1	IGUI2	ARUP1	ARUP2	AMTUP1	AMTUP2	3CUM2	ARPrLP1	ARPrLP2
IC_Lw_Lf_P2	0.87	3.21	2.45	-9.09	8.97	7.89	3.19	8.81	-2.41	-4.62	-3.82	-5.95	5.37	2.87	5.79
MD_Lw_Lf_P2	14.48	15.23	14.96	-1.02	5.29	15.54	6.00	6.70	2.54	10.02	2.89	2.93	-11.10	8.49	9.10
MD_Lw_Lf_P1	3.17	5.46	14.10	-0.71	5.56	15.14	4.45	1.31	0.05	8.25	5.09	5.95	-6.89	4.03	5.63
MD_Lw_Lf_C	2.80	21.76	7.85	-3.08	12.64	9.55	8.74	10.22	1.25	4.56	-0.01	0.14	-5.39	4.82	4.51
MD_Lw_Lf_I2	3.17	20.43	7.63	-3.65	16.07	10.35	9.32	15.58	5.14	2.86	5.74	6.81	-10.72	10.88	5.87
MD_Lw_Lf_I1	5.13	18.97	8.04	-6.48	13.01	11.81	10.03	12.94	-1.93	2.69	9.21	9.98	-11.16	8.15	9.53
MD_Lw_Rt_I1	4.72	18.97	8.33	-6.10	10.70	11.99	10.14	12.30	0.98	3.83	8.75	6.82	-10.85	9.15	11.90
MD_Lw_Rt_I2	5.50	9.81	4.38	-7.85	15.63	11.13	7.38	15.53	-2.04	3.87	15.58	7.36	-9.64	7.37	7.60
MD_Lw_Rt_C	-0.47	16.08	5.92	-4.57	11.88	12.33	6.88	11.84	-2.39	2.77	3.44	3.31	-3.65	0.18	2.01
MD_Lw_Rt_P1	3.58	12.60	10.01	-4.82	7.37	13.81	6.90	5.67	2.71	6.24	5.08	5.09	-7.04	4.13	5.20
MD_Lw_Rt_P2	9.47	13.27	12.01	-0.19	5.55	13.38	10.05	3.26	6.68	14.75	-1.20	3.02	-12.96	6.54	8.95
BL_Lw_Lf_P2	5.47	10.63	9.95	-6.38	8.21	18.85	3.76	2.71	3.96	9.86	10.07	6.45	-7.45	11.44	12.51
BL_Lw_Lf_P1	2.36	16.17	7.88	-2.85	4.69	11.10	5.89	1.53	4.05	6.43	2.77	5.27	-7.95	8.81	9.40
BL_Lw_Lf_C	4.35	17.30	2.59	-3.28	9.09	9.62	7.63	6.97	-0.21	7.21	0.52	3.24	-5.56	3.05	5.43
BL_Lw_Lf_I2	5.87	15.78	3.24	-3.39	8.05	13.19	3.77	6.86	-1.64	2.25	8.82	6.99	-8.44	4.44	8.90
BL_Lw_Lf_I1	5.00	23.39	3.04	-4.03	14.55	8.54	8.82	6.93	-2.50	1.39	6.08	2.97	-10.42	3.47	7.87
BL_Lw_Rt_I1	3.42	23.61	6.52	-3.81	11.94	8.67	6.32	11.04	-1.31	-0.05	6.89	4.65	-12.26	7.37	8.63
BL_Lw_Rt_I2	2.98	24.29	5.85	-7.02	8.71	5.02	6.74	7.35	-4.18	-2.12	6.53	3.33	-7.72	3.14	3.23
BL_Lw_Rt_C	-2.14	18.21	4.19	-7.93	7.22	9.43	7.25	6.61	-4.74	3.63	0.90	0.46	-5.30	2.25	2.30
BL_Lw_Rt_P1	2.04	11.79	9.34	-1.44	3.64	14.71	3.47	-3.39	9.53	6.77	5.15	6.04	-7.82	10.40	9.91
BL_Lw_Rt_P2	2.06	7.41	12.77	-4.29	7.65	16.01	5.01	-3.21	8.71	13.21	6.47	7.56	-8.90	7.48	12.21

Table B.2b. Corresponding *P* values correlation between 46 ordinal dental traits (ASUDAS) and 60 dental measurements.

p-value	SSUI1	SSUI2	DSUI1	DSUI2	WINUI1	LCUI2	TDUI1	TDUI2	TDUC	MRUC	DARUC	MetUM1	MetUM2	HipUM1	HipUM2
IC_Up_Rt_P2	6.E-01	7.E-01	6.E-01	8.E-01	5.E-01	2.E-01	4.E-03	6.E-02	1.E-01	3.E-01	4.E-01	7.E-01	2.E-01	8.E-01	2.E-01
IC_Up_Rt_P1	1.E+00	9.E-01	8.E-01	7.E-01	8.E-01	1.E-01	9.E-03	4.E-02	8.E-03	2.E-01	9.E-02	2.E-01	2.E-02	7.E-01	7.E-01
IC_Up_Rt_C	7.E-01	7.E-01	6.E-01	8.E-01	2.E-01	1.E-01	3.E-02	2.E-02	7.E-05	1.E-01	6.E-01	1.E-01	5.E-01	9.E-01	3.E-01
IC_Up_Rt_I2	6.E-01	7.E-01	1.E+00	5.E-01	2.E-01	2.E-01	2.E-01	1.E-01	5.E-03	1.E-01	9.E-01	6.E-01	4.E-01	5.E-01	4.E-01
IC_Up_Rt_I1	2.E-01	2.E-01	8.E-02	3.E-01	1.E-01	3.E-01	5.E-03	1.E-02	2.E-03	9.E-02	1.E-01	8.E-01	3.E-01	6.E-01	1.E-01
IC_Up_Lf_I1	1.E-01	1.E-01	9.E-02	3.E-01	7.E-02	4.E-01	4.E-03	1.E-02	2.E-03	8.E-02	6.E-02	8.E-01	3.E-01	7.E-01	1.E-01
IC_Up_Lf_I2	4.E-01	7.E-01	9.E-01	8.E-01	5.E-01	3.E-01	3.E-02	1.E-02	3.E-03	2.E-01	9.E-01	6.E-01	6.E-01	9.E-01	6.E-01
IC_Up_Lf_C	6.E-01	4.E-01	5.E-01	1.E+00	1.E-01	6.E-01	4.E-03	4.E-02	1.E-04	6.E-02	6.E-01	3.E-01	8.E-01	7.E-01	7.E-01
IC_Up_Lf_P1	7.E-01	7.E-01	6.E-01	7.E-01	1.E+00	4.E-01	6.E-03	5.E-02	1.E-02	2.E-01	4.E-02	6.E-01	2.E-01	8.E-01	7.E-01
IC_Up_Lf_P2	7.E-01	8.E-01	3.E-01	6.E-01	6.E-01	4.E-01	6.E-03	5.E-02	3.E-01	2.E-01	4.E-02	3.E-01	3.E-02	6.E-01	4.E-01
MD_Up_Lf_P2	2.E-01	8.E-02	3.E-01	9.E-01	1.E-01	4.E-01	1.E-02	3.E-02	3.E-02	3.E-01	5.E-02	8.E-01	8.E-01	9.E-01	2.E-02
MD_Up_Lf_P1	3.E-01	4.E-01	7.E-01	9.E-01	6.E-01	4.E-01	7.E-02	5.E-02	3.E-02	6.E-01	**1.E-06**	5.E-01	5.E-01	4.E-01	8.E-02
MD_Up_Lf_C	7.E-01	7.E-01	5.E-01	6.E-01	4.E-01	3.E-01	2.E-02	3.E-02	2.E-04	2.E-03	3.E-02	4.E-01	7.E-01	8.E-01	1.E-02
MD_Up_Lf_I2	6.E-03	1.E-03	8.E-02	4.E-03	1.E+00	8.E-01	**2.E-05**	**5.E-06**	8.E-05	6.E-02	8.E-02	4.E-01	8.E-01	8.E-01	7.E-02
MD_Up_Lf_I1	3.E-02	2.E-03	9.E-01	6.E-01	1.E-01	3.E-01	**2.E-07**	8.E-05	**6.E-08**	2.E-01	2.E-01	4.E-01	2.E-01	9.E-01	2.E-02
MD_Up_Rt_I1	6.E-04	**1.E-05**	6.E-01	3.E-01	4.E-02	5.E-01	**3.E-08**	4.E-05	**4.E-08**	9.E-02	5.E-02	8.E-01	2.E-01	1.E+00	1.E-02
MD_Up_Rt_I2	2.E-02	7.E-03	6.E-02	2.E-02	2.E-01	4.E-01	1.E-04	**1.E-05**	**2.E-05**	6.E-02	8.E-02	9.E-01	6.E-01	8.E-01	4.E-02
MD_Up_Rt_C	6.E-01	3.E-01	4.E-01	3.E-01	9.E-01	6.E-01	1.E-03	2.E-02	3.E-05	5.E-04	2.E-03	5.E-01	5.E-01	5.E-01	3.E-03
MD_Up_Rt_P1	3.E-01	3.E-01	6.E-01	4.E-01	9.E-01	4.E-01	7.E-02	1.E-01	3.E-03	4.E-01	6.E-05	7.E-01	1.E+00	5.E-01	1.E-01

Continue…

p-value	C5UM1	C5UM2	CarUM1	SSLI1	SSLI2	DSLI1	DSLI2	DARLC	LCVLP1	LCVLP2	AFLM1	AFLM2	GPLM2	CNLM1	CNLM2	C5LM1
IC_Up_Rt_P2	5.E-01	7.E-01	8.E-04	4.E-01	4.E-01	2.E-01	6.E-01	2.E-01	7.E-01	4.E-01	4.E-02	3.E-02	5.E-01	3.E-01	9.E-01	3.E-02
IC_Up_Rt_P1	7.E-01	8.E-01	2.E-02	8.E-01	8.E-01	6.E-01	9.E-01	5.E-01	4.E-01	6.E-01	1.E-01	2.E-01	1.E-01	5.E-01	9.E-01	3.E-01
IC_Up_Rt_C	8.E-01	5.E-01	6.E-02	9.E-01	7.E-01	6.E-01	8.E-01	4.E-01	4.E-01	1.E-01	5.E-02	3.E-02	4.E-01	6.E-01	8.E-01	2.E-01
IC_Up_Rt_I2	4.E-01	9.E-01	2.E-02	3.E-01	5.E-01	8.E-01	9.E-01	1.E-01	9.E-01	2.E-01	9.E-02	3.E-01	9.E-01	6.E-01	5.E-01	2.E-01
IC_Up_Rt_I1	9.E-01	3.E-01	2.E-03	4.E-01	7.E-01	8.E-01	9.E-01	2.E-01	6.E-02	4.E-02	2.E-01	8.E-01	8.E-01	8.E-01	7.E-01	2.E-01
IC_Up_Lf_I1	7.E-01	2.E-01	2.E-04	4.E-01	5.E-01	3.E-01	6.E-01	2.E-01	7.E-02	9.E-02	4.E-02	3.E-01	1.E+00	5.E-01	8.E-01	3.E-01
IC_Up_Lf_I2	5.E-01	2.E-01	1.E-02	7.E-01	6.E-01	6.E-01	7.E-01	3.E-01	9.E-01	2.E-01	7.E-03	2.E-01	3.E-01	9.E-01	8.E-01	1.E-01
IC_Up_Lf_C	9.E-01	4.E-02	6.E-02	6.E-01	5.E-01	6.E-01	1.E+00	5.E-01	3.E-01	2.E-01	5.E-02	2.E-01	6.E-01	8.E-01	9.E-01	2.E-01
IC_Up_Lf_P1	7.E-01	5.E-01	1.E-02	1.E+00	9.E-01	4.E-01	9.E-01	1.E-01	7.E-01	4.E-01	1.E-01	5.E-01	5.E-01	7.E-01	9.E-01	5.E-01
IC_Up_Lf_P2	6.E-01	4.E-01	8.E-02	6.E-01	5.E-01	1.E-01	4.E-01	9.E-01	8.E-01	4.E-01	7.E-02	6.E-02	9.E-01	2.E-01	7.E-01	1.E-02
MD_Up_Lf_P2	7.E-01	7.E-01	5.E-02	3.E-02	2.E-01	6.E-01	8.E-01	1.E+00	3.E-01	5.E-01	8.E-01	4.E-02	9.E-01	3.E-01	2.E-01	4.E-01
MD_Up_Lf_P1	1.E-01	2.E-02	5.E-02	2.E-02	8.E-02	1.E-01	5.E-01	1.E-01	6.E-01	5.E-01	4.E-01	8.E-01	4.E-01	7.E-01	1.E-01	9.E-01
MD_Up_Lf_C	1.E-01	2.E-01	7.E-02	7.E-02	4.E-02	1.E-02	1.E-01	2.E-03	5.E-01	7.E-01	7.E-01	5.E-01	8.E-01	7.E-01	1.E-01	6.E-01
MD_Up_Lf_I2	4.E-02	3.E-01	9.E-03	1.E-04	1.E-04	5.E-03	3.E-02	2.E-02	5.E-01	2.E-01	7.E-01	5.E-01	8.E-04	1.E-01	**1.E-06**	8.E-01
MD_Up_Lf_I1	1.E-01	1.E-01	**5.E-07**	4.E-02	2.E-02	1.E-01	4.E-01	3.E-03	3.E-01	7.E-02	5.E-01	1.E+00	7.E-03	1.E-01	3.E-02	3.E-01
MD_Up_Rt_I1	2.E-02	4.E-02	**2.E-06**	5.E-03	3.E-03	2.E-01	4.E-01	5.E-03	5.E-01	1.E-01	5.E-01	1.E+00	3.E-03	2.E-01	3.E-02	8.E-01
MD_Up_Rt_I2	2.E-03	2.E-01	1.E-02	4.E-05	1.E-04	5.E-03	5.E-02	3.E-02	3.E-01	4.E-01	9.E-01	5.E-01	8.E-03	4.E-01	2.E-04	4.E-01
MD_Up_Rt_C	6.E-02	4.E-01	4.E-02	2.E-02	1.E-02	2.E-02	1.E-01	**1.E-06**	1.E+00	7.E-01	4.E-01	4.E-01	4.E-01	7.E-01	3.E-02	8.E-01
MD_Up_Rt_P1	9.E-03	1.E-02	7.E-02	3.E-02	6.E-02	3.E-01	5.E-01	3.E-01	8.E-01	7.E-01	6.E-01	5.E-01	4.E-01	7.E-01	2.E-01	4.E-01

Continue…

p-value	C5LM2	C6LM1	C7LM1	DWLM1	PrtostLM1	PrtostLM2	IGUI1	IGUI2	ARUP1	ARUP2	AMTUP1	AMTUP2	3CUM2	ARPrLP1	ARPrLP2
IC_Up_Rt_P2	5.E-01	1.E+00	7.E-01	2.E-01	2.E-02	5.E-03	1.E-01	1.E-01	3.E-01	1.E-01	7.E-01	2.E-01	2.E-01	8.E-01	1.E-01
IC_Up_Rt_P1	2.E-01	6.E-01	9.E-01	1.E-01	3.E-01	6.E-04	5.E-02	4.E-02	6.E-01	2.E-01	1.E-01	8.E-01	3.E-01	9.E-01	3.E-01
IC_Up_Rt_C	7.E-01	2.E-01	6.E-01	3.E-01	9.E-02	9.E-03	5.E-02	2.E-02	1.E+00	7.E-02	6.E-01	9.E-01	3.E-01	5.E-02	3.E-02
IC_Up_Rt_I2	4.E-01	2.E-01	4.E-01	5.E-01	3.E-01	2.E-03	7.E-02	4.E-01	9.E-01	5.E-01	8.E-01	4.E-01	8.E-01	1.E-02	2.E-03
IC_Up_Rt_I1	8.E-01	5.E-01	4.E-01	8.E-01	3.E-01	2.E-02	2.E-02	3.E-02	2.E-01	2.E-03	1.E+00	9.E-01	4.E-01	2.E-02	6.E-04
IC_Up_Lf_I1	8.E-01	2.E-01	5.E-01	7.E-01	3.E-01	3.E-02	8.E-03	3.E-02	3.E-01	1.E-02	9.E-01	7.E-01	4.E-01	6.E-02	1.E-02
IC_Up_Lf_I2	5.E-01	8.E-01	1.E+00	5.E-01	2.E-01	4.E-02	1.E-02	8.E-02	3.E-01	1.E-01	6.E-01	9.E-01	9.E-01	2.E-01	3.E-02
IC_Up_Lf_C	8.E-01	7.E-01	6.E-01	2.E-01	7.E-02	3.E-02	1.E-02	7.E-02	7.E-01	2.E-01	9.E-01	1.E+00	9.E-01	3.E-01	4.E-02
IC_Up_Lf_P1	5.E-01	8.E-01	1.E+00	9.E-02	6.E-02	3.E-03	1.E-01	3.E-02	9.E-01	7.E-01	2.E-01	8.E-01	3.E-01	6.E-01	2.E-01
IC_Up_Lf_P2	2.E-01	6.E-01	9.E-01	5.E-02	1.E-01	5.E-03	2.E-01	2.E-01	1.E+00	9.E-01	1.E+00	6.E-01	4.E-01	9.E-01	1.E-02
MD_Up_Lf_P2	5.E-02	9.E-01	1.E-01	6.E-01	6.E-02	2.E-02	8.E-01	6.E-01	1.E-02	6.E-03	7.E-01	3.E-01	2.E-04	5.E-02	1.E-03
MD_Up_Lf_P1	1.E-01	2.E-01	8.E-01	6.E-01	2.E-02	6.E-02	4.E-01	8.E-01	4.E-02	2.E-03	4.E-01	5.E-01	3.E-02	6.E-01	7.E-02
MD_Up_Lf_C	9.E-02	9.E-01	5.E-01	7.E-02	2.E-02	3.E-03	4.E-01	4.E-01	6.E-03	1.E-02	4.E-01	4.E-01	3.E-02	1.E-01	3.E-01
MD_Up_Lf_I2	3.E-05	8.E-02	5.E-01	6.E-01	2.E-02	2.E-02	1.E-01	1.E-02	3.E-02	4.E-01	2.E-01	3.E-01	6.E-02	4.E-03	3.E-02
MD_Up_Lf_I1	6.E-03	4.E-03	5.E-01	3.E-01	2.E-03	1.E-03	3.E-01	9.E-02	2.E-02	9.E-02	1.E-01	2.E-01	4.E-02	4.E-03	4.E-03
MD_Up_Rt_I1	1.E-02	1.E-01	4.E-01	1.E-01	5.E-04	6.E-03	1.E-01	1.E-02	5.E-02	6.E-02	1.E-01	8.E-02	2.E-02	8.E-03	1.E-02
MD_Up_Rt_I2	2.E-04	6.E-01	9.E-01	7.E-01	1.E-02	5.E-02	1.E-01	5.E-02	1.E-01	2.E-01	5.E-02	8.E-01	7.E-03	7.E-03	3.E-03
MD_Up_Rt_C	3.E-02	5.E-01	4.E-01	3.E-01	1.E-01	4.E-03	3.E-02	6.E-01	3.E-02	1.E-03	7.E-01	6.E-01	1.E-03	2.E-02	8.E-03
MD_Up_Rt_P1	2.E-01	8.E-01	1.E+00	6.E-01	2.E-02	2.E-01	8.E-01	1.E+00	2.E-01	5.E-03	3.E-01	8.E-02	4.E-02	4.E-01	5.E-02

Continue…

p-value	SSUI1	SSUI2	DSUI1	DSUI2	WINUI1	LCUI2	TDUI1	TDUI2	TDUC	MRUC	DARUC	MetUM1	MetUM2	HipUM1	HipUM2
MD_Up_Rt_P2	6.E-02	5.E-02	1.E-02	1.E-01	6.E-01	5.E-01	1.E-02	2.E-02	1.E-02	3.E-01	9.E-02	8.E-01	1.E+00	8.E-01	2.E-02
BL_Up_Lf_P2	6.E-01	8.E-01	6.E-01	8.E-01	8.E-02	8.E-01	6.E-02	5.E-02	5.E-04	9.E-01	4.E-02	6.E-01	8.E-01	8.E-01	3.E-01
BL_Up_Lf_P1	9.E-01	1.E+00	9.E-01	8.E-01	2.E-01	4.E-01	6.E-02	3.E-01	2.E-02	6.E-01	4.E-02	4.E-01	3.E-01	8.E-01	4.E-01
BL_Up_Lf_C	9.E-01	2.E-01	8.E-01	8.E-01	6.E-03	8.E-01	1.E-02	4.E-03	**2.E-09**	7.E-02	1.E-01	4.E-01	5.E-01	3.E-01	2.E-01
BL_Up_Lf_I2	2.E-01	3.E-02	7.E-02	3.E-01	2.E-01	7.E-01	3.E-04	3.E-05	**2.E-05**	3.E-01	2.E-01	8.E-01	5.E-01	7.E-01	3.E-01
BL_Up_Lf_I1	1.E-01	8.E-03	4.E-01	5.E-01	4.E-04	5.E-01	**5.E-07**	3.E-03	6.E-05	5.E-02	1.E-01	1.E+00	1.E+00	7.E-01	2.E-01
BL_Up_Rt_I1	1.E-01	8.E-03	3.E-01	4.E-01	8.E-03	8.E-01	**3.E-06**	7.E-04	**9.E-07**	2.E-02	2.E-01	8.E-01	7.E-01	1.E+00	6.E-02
BL_Up_Rt_I2	5.E-01	1.E-01	9.E-02	2.E-01	1.E-01	4.E-01	1.E-04	2.E-05	**1.E-06**	5.E-01	2.E-01	4.E-01	8.E-01	6.E-01	2.E-01
BL_Up_Rt_C	7.E-01	4.E-01	8.E-01	9.E-01	4.E-02	5.E-01	1.E-02	4.E-03	**2.E-09**	1.E-01	3.E-01	3.E-01	6.E-01	3.E-01	2.E-01
BL_Up_Rt_P1	8.E-01	8.E-01	5.E-01	1.E+00	5.E-01	9.E-01	3.E-01	4.E-01	2.E-02	1.E+00	8.E-03	3.E-01	6.E-01	5.E-01	3.E-01
BL_Up_Rt_P2	7.E-01	9.E-01	4.E-01	7.E-01	2.E-01	9.E-01	1.E-01	1.E-02	9.E-05	7.E-01	7.E-02	9.E-01	9.E-01	5.E-01	5.E-01
IC_Lw_Rt_P2	8.E-01	1.E+00	9.E-01	6.E-01	2.E-01	7.E-01	2.E-01	6.E-02	3.E-03	4.E-01	6.E-01	1.E-01	4.E-02	5.E-01	7.E-01
IC_Lw_Rt_P1	6.E-01	7.E-01	6.E-01	6.E-01	2.E-01	2.E-01	1.E-01	1.E-01	2.E-04	3.E-01	1.E-01	7.E-01	3.E-02	9.E-01	2.E-01
IC_Lw_Rt_C	2.E-01	3.E-01	4.E-01	2.E-01	2.E-02	1.E-01	2.E-01	2.E-01	7.E-03	6.E-01	3.E-01	3.E-01	5.E-01	2.E-01	8.E-01
IC_Lw_Rt_I2	9.E-01	9.E-01	9.E-01	9.E-01	1.E-01	9.E-02	5.E-02	4.E-02	7.E-03	5.E-01	1.E-01	9.E-01	2.E-01	8.E-01	7.E-01
IC_Lw_Rt_I1	5.E-01	6.E-01	7.E-01	1.E+00	4.E-01	5.E-02	9.E-02	6.E-02	1.E-01	4.E-01	2.E-01	9.E-01	9.E-01	6.E-01	9.E-01
IC_Lw_Lf_I1	5.E-01	6.E-01	9.E-01	7.E-01	2.E-01	1.E-01	5.E-02	6.E-02	1.E-01	7.E-01	3.E-01	6.E-01	9.E-01	4.E-01	1.E+00
IC_Lw_Lf_I2	9.E-01	9.E-01	1.E+00	8.E-01	3.E-01	7.E-02	3.E-02	6.E-03	2.E-04	2.E-01	2.E-01	9.E-01	5.E-01	1.E+00	6.E-01
IC_Lw_Lf_C	3.E-01	6.E-01	6.E-01	7.E-01	2.E-01	8.E-02	1.E-01	3.E-01	7.E-03	2.E-01	3.E-01	2.E-01	7.E-01	6.E-01	9.E-01
IC_Lw_Lf_P1	3.E-01	3.E-01	5.E-01	6.E-01	1.E-01	4.E-01	2.E-02	2.E-01	5.E-04	2.E-01	2.E-02	8.E-01	4.E-02	7.E-01	2.E-01

Continue…

p-value	C5UM1	C5UM2	CarUM1	SSLI1	SSLI2	DSLI1	DSLI2	DARLC	LCVLP1	LCVLP2	AFLM1	AFLM2	GPLM2	CNLM1	CNLM2	C5LM1
MD_Up_Rt_P2	6.E-01	3.E-01	3.E-02	1.E-02	1.E-01	3.E-01	7.E-01	7.E-01	7.E-01	4.E-01	6.E-01	9.E-02	7.E-01	8.E-01	2.E-01	3.E-01
BL_Up_Lf_P2	5.E-02	2.E-01	4.E-04	3.E-01	4.E-01	2.E-01	3.E-01	6.E-02	1.E+00	2.E-01	6.E-01	8.E-01	6.E-01	7.E-01	4.E-01	2.E-01
BL_Up_Lf_P1	2.E-02	1.E-01	1.E-03	3.E-01	3.E-01	1.E-01	5.E-01	6.E-01	7.E-01	5.E-01	3.E-01	5.E-01	1.E+00	6.E-01	9.E-01	5.E-01
BL_Up_Lf_C	5.E-01	1.E+00	8.E-04	4.E-01	3.E-01	3.E-01	7.E-01	1.E-01	2.E-01	3.E-01	5.E-01	2.E-01	8.E-01	8.E-01	2.E-01	3.E-01
BL_Up_Lf_I2	5.E-02	4.E-01	3.E-03	3.E-02	6.E-02	4.E-01	8.E-01	5.E-01	5.E-01	7.E-02	3.E-02	7.E-01	4.E-01	7.E-01	1.E-01	3.E-01
BL_Up_Lf_I1	2.E-01	5.E-01	6.E-05	3.E-01	3.E-01	2.E-01	9.E-01	5.E-01	4.E-02	8.E-03	1.E-01	2.E-01	4.E-01	1.E+00	3.E-01	3.E-01
BL_Up_Rt_I1	1.E-01	6.E-01	7.E-05	2.E-01	2.E-01	3.E-01	8.E-01	5.E-01	5.E-02	5.E-03	1.E-01	2.E-01	4.E-01	8.E-01	4.E-01	1.E-01
BL_Up_Rt_I2	3.E-02	5.E-01	4.E-04	1.E-02	2.E-02	4.E-01	9.E-01	3.E-02	7.E-01	3.E-02	2.E-01	9.E-01	1.E-01	8.E-01	7.E-02	1.E+00
BL_Up_Rt_C	3.E-01	8.E-01	7.E-03	9.E-01	4.E-01	2.E-01	6.E-01	7.E-02	5.E-01	4.E-01	6.E-01	9.E-02	9.E-01	3.E-01	4.E-01	3.E-01
BL_Up_Rt_P1	2.E-02	4.E-01	1.E-03	2.E-01	3.E-01	6.E-01	8.E-01	7.E-01	9.E-01	5.E-01	2.E-01	2.E-01	1.E+00	5.E-01	9.E-01	1.E+00
BL_Up_Rt_P2	1.E-02	7.E-02	6.E-05	2.E-01	3.E-01	1.E-01	5.E-01	1.E-01	7.E-01	2.E-01	4.E-01	5.E-01	9.E-01	5.E-01	2.E-01	9.E-01
IC_Lw_Rt_P2	8.E-02	1.E-01	1.E-03	3.E-01	5.E-01	7.E-01	6.E-01	4.E-01	9.E-01	8.E-01	4.E-02	7.E-02	7.E-01	4.E-01	2.E-01	9.E-02
IC_Lw_Rt_P1	9.E-01	2.E-01	2.E-03	5.E-01	9.E-01	5.E-01	8.E-01	1.E-01	7.E-01	9.E-01	8.E-01	1.E-01	2.E-01	5.E-01	9.E-01	5.E-01
IC_Lw_Rt_C	1.E+00	5.E-01	1.E-01	4.E-01	8.E-01	7.E-01	1.E+00	6.E-02	7.E-01	2.E-01	6.E-02	9.E-03	5.E-01	7.E-01	7.E-01	3.E-01
IC_Lw_Rt_I2	5.E-01	6.E-01	1.E-01	2.E-01	1.E-01	4.E-01	8.E-01	3.E-01	7.E-01	6.E-01	3.E-02	2.E-03	9.E-01	9.E-01	3.E-01	4.E-01
IC_Lw_Rt_I1	7.E-01	6.E-02	5.E-02	3.E-01	2.E-01	2.E-01	6.E-01	4.E-01	6.E-01	8.E-01	9.E-02	3.E-02	7.E-01	6.E-01	1.E-01	1.E-01
IC_Lw_Lf_I1	5.E-01	3.E-02	6.E-02	5.E-01	4.E-01	3.E-01	5.E-01	3.E-01	2.E-01	8.E-01	4.E-02	7.E-03	5.E-01	6.E-01	6.E-01	3.E-01
IC_Lw_Lf_I2	9.E-01	2.E-01	2.E-02	4.E-01	2.E-01	7.E-01	8.E-01	3.E-01	5.E-01	3.E-01	3.E-02	5.E-03	1.E+00	4.E-01	2.E-01	2.E-01
IC_Lw_Lf_C	8.E-01	5.E-01	2.E-02	8.E-01	7.E-01	3.E-01	7.E-01	2.E-01	6.E-01	2.E-01	8.E-02	5.E-02	5.E-01	9.E-01	3.E-01	5.E-01
IC_Lw_Lf_P1	4.E-01	2.E-01	5.E-02	7.E-01	8.E-01	3.E-01	5.E-01	8.E-01	7.E-01	7.E-01	6.E-01	5.E-01	8.E-01	7.E-01	7.E-02	1.E-01

Continue…

p-value	C5LM2	C6LM1	C7LM1	DWLM1	PrtostLM1	PrtostLM2	IGUI1	IGUI2	ARUP1	ARUP2	AMTUP1	AMTUP2	3CUM2	ARPrLP1	ARPrLP2
MD_Up_Rt_P2	7.E-02	8.E-01	3.E-01	7.E-01	1.E-02	5.E-02	9.E-01	9.E-01	5.E-01	5.E-03	7.E-01	9.E-01	2.E-03	2.E-01	5.E-03
BL_Up_Lf_P2	4.E-01	8.E-02	7.E-01	8.E-01	3.E-03	7.E-03	4.E-01	7.E-01	7.E-04	2.E-03	3.E-01	3.E-02	6.E-02	3.E-02	6.E-03
BL_Up_Lf_P1	8.E-01	1.E-01	1.E+00	1.E+00	9.E-02	2.E-02	1.E-01	9.E-01	6.E-03	6.E-02	4.E-02	2.E-02	1.E-01	4.E-02	5.E-02
BL_Up_Lf_C	3.E-01	2.E-01	8.E-01	1.E-01	2.E-03	9.E-03	9.E-02	2.E-02	3.E-01	5.E-01	7.E-01	5.E-01	7.E-02	6.E-01	2.E-01
BL_Up_Lf_I2	7.E-02	5.E-01	2.E-01	5.E-01	1.E-02	1.E-01	1.E-01	1.E-01	2.E-02	4.E-02	8.E-02	3.E-01	2.E-01	3.E-02	6.E-02
BL_Up_Lf_I1	3.E-01	4.E-01	6.E-01	3.E-01	2.E-02	3.E-01	1.E-03	8.E-02	3.E-02	2.E-01	2.E-01	1.E-01	3.E-02	3.E-02	2.E-02
BL_Up_Rt_I1	4.E-01	5.E-01	2.E-01	3.E-01	7.E-03	2.E-01	4.E-02	6.E-02	1.E-01	9.E-02	8.E-02	2.E-01	5.E-02	2.E-02	3.E-02
BL_Up_Rt_I2	2.E-02	6.E-01	3.E-01	8.E-01	6.E-02	6.E-02	1.E-02	5.E-02	5.E-03	3.E-02	4.E-01	4.E-01	1.E-01	3.E-03	1.E-03
BL_Up_Rt_C	8.E-01	4.E-02	2.E-01	3.E-01	1.E-02	2.E-02	6.E-02	1.E-01	1.E-01	4.E-01	8.E-01	2.E-01	1.E-01	4.E-01	4.E-01
BL_Up_Rt_P1	1.E+00	2.E-01	1.E+00	4.E-01	1.E-01	1.E-01	2.E-01	9.E-01	2.E-02	6.E-02	3.E-01	1.E-02	2.E-01	5.E-02	3.E-02
BL_Up_Rt_P2	4.E-01	1.E-01	7.E-01	9.E-01	2.E-03	4.E-03	1.E-01	7.E-01	1.E-03	2.E-03	4.E-02	7.E-02	3.E-02	4.E-02	3.E-02
IC_Lw_Rt_P2	1.E+00	7.E-01	9.E-01	5.E-02	2.E-02	3.E-02	5.E-01	2.E-01	9.E-01	8.E-01	5.E-01	9.E-01	9.E-01	3.E-01	3.E-01
IC_Lw_Rt_P1	9.E-01	5.E-01	2.E-01	4.E-01	6.E-02	2.E-02	5.E-01	1.E-01	1.E+00	7.E-01	3.E-01	2.E-01	2.E-01	9.E-01	2.E-01
IC_Lw_Rt_C	9.E-01	6.E-01	4.E-02	3.E-01	3.E-01	2.E-01	2.E-01	2.E-01	8.E-01	1.E-01	6.E-01	8.E-01	5.E-01	2.E-01	8.E-03
IC_Lw_Rt_I2	4.E-01	9.E-01	5.E-01	8.E-02	4.E-01	8.E-02	1.E-01	2.E-01	9.E-01	2.E-01	7.E-01	6.E-01	7.E-01	1.E-01	3.E-03
IC_Lw_Rt_I1	4.E-01	1.E+00	3.E-01	7.E-02	5.E-01	7.E-02	2.E-01	3.E-01	4.E-01	6.E-01	6.E-01	8.E-01	7.E-01	8.E-01	1.E-01
IC_Lw_Lf_I1	8.E-01	9.E-01	9.E-01	2.E-02	1.E-01	2.E-01	5.E-01	6.E-01	4.E-01	3.E-01	8.E-01	8.E-01	9.E-01	9.E-01	9.E-02
IC_Lw_Lf_I2	7.E-01	5.E-01	8.E-01	6.E-02	2.E-01	7.E-02	2.E-01	1.E-01	9.E-01	2.E-01	1.E+00	6.E-01	9.E-01	7.E-01	2.E-02
IC_Lw_Lf_C	1.E+00	8.E-01	6.E-01	4.E-01	2.E-02	2.E-02	2.E-01	7.E-02	8.E-01	3.E-01	7.E-01	3.E-01	7.E-01	7.E-01	7.E-03
IC_Lw_Lf_P1	3.E-01	3.E-01	4.E-01	7.E-01	4.E-02	4.E-02	8.E-01	3.E-01	4.E-01	8.E-01	4.E-01	8.E-01	3.E-01	4.E-01	9.E-02

Continue…

p-value	SSUI1	SSUI2	DSUI1	DSUI2	WINUI1	LCUI2	TDUI1	TDUI2	TDUC	MRUC	DARUC	MetUM1	MetUM2	HipUM1	HipUM2
IC_Lw_Lf_P2	6.E-01	8.E-01	6.E-01	3.E-01	2.E-02	1.E+00	2.E-01	3.E-01	8.E-02	5.E-01	6.E-01	5.E-01	2.E-01	4.E-01	5.E-01
MD_Lw_Lf_P2	2.E-03	2.E-03	7.E-02	1.E-01	6.E-01	8.E-01	3.E-02	8.E-02	1.E-02	5.E-02	2.E-03	1.E+00	9.E-01	6.E-01	7.E-02
MD_Lw_Lf_P1	7.E-02	9.E-02	6.E-01	4.E-01	5.E-01	2.E-01	2.E-01	2.E-02	2.E-03	3.E-01	1.E-03	9.E-01	6.E-01	9.E-01	9.E-02
MD_Lw_Lf_C	2.E-01	4.E-02	3.E-01	2.E-01	5.E-01	4.E-01	1.E-02	5.E-02	2.E-03	6.E-02	5.E-01	4.E-01	3.E-01	3.E-01	1.E-01
MD_Lw_Lf_I2	**1.E-06**	**4.E-08**	1.E-03	3.E-04	7.E-01	6.E-02	**2.E-07**	**2.E-05**	**4.E-06**	9.E-02	7.E-02	3.E-02	2.E-01	4.E-01	8.E-03
MD_Lw_Lf_I1	5.E-04	1.E-04	6.E-02	8.E-03	3.E-01	2.E-01	**3.E-06**	1.E-03	4.E-05	3.E-01	9.E-02	8.E-01	1.E-01	8.E-01	6.E-02
MD_Lw_Rt_I1	**1.E-05**	**2.E-06**	8.E-04	1.E-04	3.E-01	1.E-01	**7.E-07**	2.E-04	**1.E-06**	2.E-01	3.E-02	6.E-01	2.E-01	5.E-01	3.E-02
MD_Lw_Rt_I2	**4.E-07**	**4.E-08**	2.E-03	2.E-04	4.E-01	5.E-01	**8.E-08**	**1.E-06**	**3.E-06**	1.E-02	7.E-02	1.E-01	8.E-01	4.E-01	2.E-02
MD_Lw_Rt_C	4.E-01	1.E-01	3.E-01	2.E-01	4.E-01	4.E-01	2.E-02	2.E-02	2.E-03	7.E-02	4.E-01	6.E-01	6.E-01	5.E-01	5.E-02
MD_Lw_Rt_P1	5.E-01	3.E-01	1.E+00	9.E-01	5.E-01	7.E-02	1.E-01	4.E-02	1.E-02	6.E-01	1.E-03	6.E-01	6.E-01	7.E-01	5.E-02
MD_Lw_Rt_P2	1.E-02	1.E-02	8.E-02	2.E-01	4.E-01	4.E-01	2.E-01	2.E-02	3.E-02	2.E-01	3.E-04	1.E+00	5.E-01	3.E-01	5.E-03
BL_Lw_Lf_P2	5.E-01	5.E-01	5.E-01	7.E-01	8.E-02	9.E-01	5.E-03	9.E-02	2.E-02	4.E-01	2.E-02	8.E-01	7.E-01	2.E-01	3.E-01
BL_Lw_Lf_P1	9.E-01	1.E+00	1.E-01	2.E-01	1.E-02	9.E-01	1.E-01	1.E-01	9.E-03	6.E-01	1.E-01	8.E-01	5.E-01	5.E-01	8.E-02
BL_Lw_Lf_C	7.E-01	7.E-01	4.E-01	6.E-01	3.E-02	5.E-01	3.E-02	8.E-02	1.E-03	9.E-01	1.E-01	4.E-01	2.E-01	7.E-01	3.E-01
BL_Lw_Lf_I2	5.E-02	2.E-02	5.E-01	2.E-01	9.E-03	4.E-01	1.E-04	1.E-03	**6.E-06**	1.E-01	7.E-02	6.E-01	9.E-01	3.E-01	1.E-01
BL_Lw_Lf_I1	3.E-02	2.E-02	3.E-01	1.E-01	7.E-03	7.E-01	2.E-04	2.E-02	8.E-04	2.E-01	3.E-01	5.E-01	6.E-01	8.E-01	6.E-02
BL_Lw_Rt_I1	6.E-02	2.E-02	4.E-01	1.E-01	1.E-02	2.E-01	4.E-04	4.E-03	1.E-03	9.E-02	2.E-01	4.E-01	9.E-01	4.E-01	5.E-02
BL_Lw_Rt_I2	4.E-01	2.E-01	4.E-01	2.E-01	8.E-02	7.E-01	1.E-03	7.E-03	2.E-04	6.E-02	1.E-01	3.E-01	9.E-01	4.E-01	1.E-01
BL_Lw_Rt_C	6.E-01	9.E-01	5.E-01	9.E-01	4.E-02	7.E-01	3.E-02	5.E-02	2.E-04	3.E-01	1.E-01	6.E-02	5.E-01	9.E-01	3.E-01
BL_Lw_Rt_P1	4.E-01	6.E-01	1.E-01	4.E-01	1.E-02	1.E+00	5.E-02	5.E-02	2.E-03	3.E-01	4.E-02	9.E-01	5.E-01	5.E-01	1.E-01
BL_Lw_Rt_P2	5.E-01	6.E-01	3.E-01	7.E-01	5.E-02	6.E-01	9.E-02	3.E-02	2.E-03	6.E-01	7.E-03	5.E-01	9.E-01	3.E-01	1.E-01

Continue…

p-value	C5UM1	C5UM2	CarUM1	SSLI1	SSLI2	DSLI1	DSLI2	DARLC	LCVLP1	LCVLP2	AFLM1	AFLM2	GPLM2	CNLM1	CNLM2	C5LM1
IC_Lw_Lf_P2	4.E-01	3.E-01	3.E-03	8.E-01	8.E-01	5.E-01	8.E-01	3.E-01	4.E-01	1.E+00	4.E-02	2.E-02	8.E-01	1.E+00	3.E-01	1.E-01
MD_Lw_Lf_P2	2.E-02	3.E-01	1.E-04	1.E-01	2.E-01	4.E-01	9.E-01	8.E-01	1.E+00	2.E-02	4.E-01	3.E-01	1.E-01	8.E-05	2.E-03	8.E-01
MD_Lw_Lf_P1	3.E-02	7.E-01	5.E-04	6.E-02	9.E-02	6.E-02	2.E-01	1.E-01	8.E-01	5.E-01	6.E-01	7.E-01	9.E-01	1.E-02	7.E-01	4.E-01
MD_Lw_Lf_C	8.E-01	6.E-01	6.E-04	2.E-01	1.E-01	1.E-02	9.E-02	1.E-03	2.E-01	6.E-01	2.E-01	2.E-01	4.E-01	1.E-02	7.E-02	2.E-01
MD_Lw_Lf_I2	6.E-01	8.E-01	2.E-04	1.E-02	9.E-03	2.E-01	2.E-01	6.E-03	1.E+00	1.E-01	8.E-01	8.E-01	1.E-01	3.E-03	6.E-02	3.E-01
MD_Lw_Lf_I1	3.E-01	7.E-01	2.E-04	3.E-02	5.E-02	4.E-01	8.E-01	1.E-02	6.E-01	1.E-01	2.E-01	4.E-01	5.E-02	9.E-04	2.E-01	6.E-01
MD_Lw_Rt_I1	5.E-01	5.E-01	2.E-04	6.E-03	1.E-02	1.E-01	4.E-01	2.E-02	9.E-01	8.E-02	7.E-01	6.E-01	5.E-02	2.E-03	2.E-01	5.E-01
MD_Lw_Rt_I2	3.E-01	9.E-01	6.E-03	2.E-04	3.E-04	1.E-01	2.E-01	2.E-01	9.E-01	9.E-02	1.E+00	6.E-01	5.E-01	9.E-02	1.E-01	2.E-01
MD_Lw_Rt_C	5.E-01	9.E-01	7.E-05	9.E-02	3.E-02	5.E-02	1.E-01	3.E-03	4.E-01	5.E-01	4.E-01	8.E-02	9.E-01	2.E-02	4.E-01	5.E-01
MD_Lw_Rt_P1	6.E-01	9.E-01	3.E-05	2.E-01	3.E-01	1.E-01	4.E-01	1.E-01	4.E-01	2.E-01	7.E-01	6.E-01	9.E-01	3.E-02	3.E-01	7.E-01
MD_Lw_Rt_P2	9.E-02	2.E-01	1.E-04	1.E-01	3.E-01	6.E-01	6.E-01	3.E-01	5.E-01	1.E-03	9.E-01	2.E-01	5.E-01	1.E-04	7.E-03	4.E-01
BL_Lw_Lf_P2	4.E-01	3.E-01	1.E-04	1.E-01	2.E-01	1.E+00	8.E-01	6.E-02	9.E-02	5.E-04	4.E-01	6.E-01	5.E-01	1.E-01	1.E-01	6.E-01
BL_Lw_Lf_P1	3.E-01	2.E-01	5.E-04	9.E-01	8.E-01	8.E-01	9.E-01	6.E-02	4.E-02	5.E-02	9.E-01	8.E-01	1.E+00	3.E-02	3.E-01	8.E-01
BL_Lw_Lf_C	6.E-01	8.E-01	5.E-04	1.E+00	8.E-01	2.E-01	5.E-01	9.E-03	8.E-01	8.E-01	2.E-01	2.E-01	5.E-01	2.E-01	2.E-01	5.E-01
BL_Lw_Lf_I2	2.E-01	7.E-01	6.E-05	9.E-01	5.E-01	2.E-01	3.E-01	8.E-02	4.E-01	5.E-02	8.E-01	4.E-01	4.E-02	2.E-02	2.E-01	5.E-01
BL_Lw_Lf_I1	7.E-01	8.E-01	**1.E-06**	6.E-01	5.E-01	1.E-01	2.E-01	8.E-02	1.E+00	1.E-01	9.E-01	5.E-01	2.E-02	2.E-03	1.E-01	9.E-01
BL_Lw_Rt_I1	7.E-01	8.E-01	**7.E-07**	9.E-01	7.E-01	3.E-01	7.E-01	4.E-02	7.E-01	1.E-01	4.E-01	9.E-01	3.E-02	1.E-03	4.E-01	5.E-01
BL_Lw_Rt_I2	3.E-01	8.E-01	**1.E-05**	9.E-01	6.E-01	2.E-01	4.E-01	6.E-03	7.E-01	2.E-01	7.E-01	3.E-01	2.E-02	2.E-03	1.E-01	3.E-01
BL_Lw_Rt_C	7.E-01	5.E-01	2.E-03	9.E-01	7.E-01	2.E-01	6.E-01	3.E-02	7.E-01	9.E-01	5.E-01	1.E-01	6.E-01	9.E-02	4.E-01	3.E-01
BL_Lw_Rt_P1	3.E-01	4.E-02	2.E-03	6.E-01	8.E-01	8.E-01	9.E-01	3.E-01	9.E-02	3.E-02	8.E-01	9.E-01	1.E+00	2.E-01	5.E-01	7.E-01
BL_Lw_Rt_P2	8.E-02	1.E-01	5.E-04	4.E-01	3.E-01	8.E-01	8.E-01	6.E-02	1.E-01	1.E-04	8.E-01	9.E-01	6.E-01	4.E-02	5.E-01	8.E-01

Continue…

p-value	C5LM2	C6LM1	C7LM1	DWLM1	PrtostLM1	PrtostLM2	IGUI1	IGUI2	ARUP1	ARUP2	AMTUP1	AMTUP2	3CUM2	ARPrLP1	ARPrLP2
IC_Lw_Lf_P2	9.E-01	5.E-01	6.E-01	5.E-02	6.E-02	9.E-02	5.E-01	6.E-02	6.E-01	3.E-01	4.E-01	2.E-01	3.E-01	6.E-01	2.E-01
MD_Lw_Lf_P2	2.E-03	1.E-03	1.E-03	8.E-01	3.E-01	1.E-03	2.E-01	2.E-01	6.E-01	4.E-02	6.E-01	5.E-01	2.E-02	8.E-02	5.E-02
MD_Lw_Lf_P1	5.E-01	3.E-01	4.E-03	9.E-01	3.E-01	2.E-03	4.E-01	8.E-01	1.E+00	9.E-02	3.E-01	2.E-01	2.E-01	4.E-01	2.E-01
MD_Lw_Lf_C	5.E-01	**2.E-06**	9.E-02	5.E-01	7.E-03	4.E-02	6.E-02	3.E-02	8.E-01	3.E-01	1.E+00	1.E+00	2.E-01	3.E-01	3.E-01
MD_Lw_Lf_I2	5.E-01	**9.E-06**	1.E-01	4.E-01	5.E-04	3.E-02	4.E-02	8.E-04	3.E-01	5.E-01	2.E-01	1.E-01	2.E-02	2.E-02	2.E-01
MD_Lw_Lf_I1	3.E-01	4.E-05	8.E-02	2.E-01	5.E-03	1.E-02	3.E-02	5.E-03	7.E-01	6.E-01	6.E-02	3.E-02	2.E-02	9.E-02	4.E-02
MD_Lw_Rt_I1	3.E-01	4.E-05	7.E-02	2.E-01	2.E-02	1.E-02	3.E-02	8.E-03	8.E-01	4.E-01	7.E-02	1.E-01	2.E-02	6.E-02	1.E-02
MD_Lw_Rt_I2	2.E-01	4.E-02	3.E-01	9.E-02	8.E-04	2.E-02	1.E-01	8.E-04	7.E-01	4.E-01	1.E-03	1.E-01	4.E-02	1.E-01	1.E-01
MD_Lw_Rt_C	9.E-01	5.E-04	2.E-01	3.E-01	1.E-02	8.E-03	1.E-01	1.E-02	6.E-01	6.E-01	5.E-01	5.E-01	4.E-01	1.E+00	7.E-01
MD_Lw_Rt_P1	5.E-01	1.E-02	4.E-02	3.E-01	1.E-01	5.E-03	2.E-01	2.E-01	6.E-01	2.E-01	3.E-01	3.E-01	1.E-01	4.E-01	3.E-01
MD_Lw_Rt_P2	4.E-02	5.E-03	1.E-02	1.E+00	2.E-01	4.E-03	3.E-02	5.E-01	2.E-01	2.E-03	8.E-01	5.E-01	6.E-03	2.E-01	6.E-02
BL_Lw_Lf_P2	2.E-01	2.E-02	3.E-02	2.E-01	8.E-02	6.E-05	4.E-01	6.E-01	4.E-01	4.E-02	4.E-02	2.E-01	1.E-01	2.E-02	8.E-03
BL_Lw_Lf_P1	6.E-01	9.E-04	1.E-01	6.E-01	3.E-01	2.E-02	2.E-01	8.E-01	4.E-01	2.E-01	6.E-01	3.E-01	1.E-01	7.E-02	5.E-02
BL_Lw_Lf_C	3.E-01	2.E-04	6.E-01	5.E-01	5.E-02	4.E-02	1.E-01	1.E-01	1.E+00	1.E-01	9.E-01	5.E-01	2.E-01	5.E-01	2.E-01
BL_Lw_Lf_I2	2.E-01	7.E-04	5.E-01	5.E-01	9.E-02	5.E-03	4.E-01	1.E-01	7.E-01	6.E-01	7.E-02	1.E-01	7.E-02	4.E-01	6.E-02
BL_Lw_Lf_I1	3.E-01	**4.E-07**	5.E-01	4.E-01	2.E-03	7.E-02	6.E-02	1.E-01	6.E-01	8.E-01	2.E-01	5.E-01	3.E-02	5.E-01	9.E-02
BL_Lw_Rt_I1	5.E-01	**3.E-07**	2.E-01	4.E-01	1.E-02	6.E-02	2.E-01	2.E-02	8.E-01	1.E+00	2.E-01	3.E-01	9.E-03	1.E-01	7.E-02
BL_Lw_Rt_I2	5.E-01	**1.E-07**	2.E-01	1.E-01	6.E-02	3.E-01	1.E-01	1.E-01	4.E-01	7.E-01	2.E-01	5.E-01	1.E-01	5.E-01	5.E-01
BL_Lw_Rt_C	6.E-01	8.E-05	4.E-01	9.E-02	1.E-01	4.E-02	1.E-01	2.E-01	3.E-01	4.E-01	9.E-01	9.E-01	3.E-01	6.E-01	6.E-01
BL_Lw_Rt_P1	7.E-01	2.E-02	6.E-02	8.E-01	5.E-01	3.E-03	5.E-01	5.E-01	6.E-02	2.E-01	3.E-01	2.E-01	1.E-01	3.E-02	4.E-02
BL_Lw_Rt_P2	7.E-01	1.E-01	6.E-03	4.E-01	1.E-01	6.E-04	3.E-01	5.E-01	8.E-02	5.E-03	2.E-01	1.E-01	6.E-02	1.E-01	9.E-03

Table B.3. Allele frequencies of index SNPs that showed genome-wide significant association to ordinal dental traits (ASUDAS method) in different populations across the World [387].

Associated Trait	Index	MAF	ALL	AFR	ACB	ASW	ESN	GWD	LWK	MSL	YRI	AMR	CLM	MXL	PEL	PUR	EAS	CDX	CHB	CHS	JPT	KHV	EUR	CEU	FIN	GBR	IBS	TSI
	SNP	CAN																										
Cusp 7 LM1 *	rs3851907	0.02	0.05	0.16	0.15	0.07	0.12	0.17	0.16	0.19	0.19	0.02	0.03	0.00	0.01	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
Double Shoveling LI1	rs1037804	0.02	0.09	0.21	0.21	0.16	0.25	0.19	0.19	0.20	0.24	0.03	0.03	0.02	0.01	0.05	0.14	0.24	0.14	0.13	0.07	0.12	0.00	0.00	0.01	0.01	0.00	0.01
Double Shoveling LI1	rs16984020	0.02	0.05	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.03	0.01	0.00	0.15	0.22	0.14	0.17	0.06	0.15	0.03	0.04	0.01	0.03	0.01	0.04
Double Shoveling LI1	rs17862881	0.01	0.01	0.03	0.05	0.04	0.02	0.03	0.01	0.02	0.04	0.01	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.03	0.01	0.01	0.02	0.01
Double Shoveling LI2	rs17862881	0.01	0.01	0.03	0.05	0.04	0.02	0.03	0.01	0.02	0.04	0.01	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.03	0.01	0.01	0.02	0.01
Double Shoveling LI2	rs1037804	0.02	0.09	0.21	0.21	0.16	0.25	0.19	0.19	0.20	0.24	0.03	0.03	0.02	0.01	0.05	0.14	0.24	0.14	0.13	0.07	0.12	0.00	0.00	0.01	0.01	0.00	0.01
Deflecting Wrinkle LM1	rs16919218	0.02	0.07	0.23	0.26	0.22	0.23	0.25	0.23	0.24	0.19	0.01	0.01	0.01	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
Hypocone UM1	rs9808165	0.03	0.07	0.26	0.19	0.16	0.24	0.29	0.34	0.32	0.26	0.02	0.02	0.02	0.01	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.03	0.01
Hypocone UM1	rs17142577	0.02	0.07	0.23	0.22	0.16	0.20	0.27	0.26	0.21	0.25	0.02	0.03	0.00	0.02	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
Metacone UM2 *	rs17072636	0.02	0.05	0.05	0.06	0.07	0.06	0.02	0.09	0.02	0.04	0.03	0.01	0.02	0.07	0.01	0.07	0.07	0.04	0.07	0.07	0.08	0.02	0.01	0.04	0.01	0.03	0.01
Protostylid LM1	rs10494193	0.02	0.07	0.22	0.23	0.18	0.25	0.18	0.24	0.23	0.22	0.02	0.02	0.02	0.01	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Protostylid LM1	rs12680196	0.03	0.13	0.10	0.12	0.01	0.10	0.15	0.13	0.06	0.07	0.04	0.04	0.06	0.02	0.04	0.33	0.28	0.41	0.34	0.35	0.25	0.06	0.04	0.08	0.10	0.05	0.03
Protostylid LM1	rs12570261	0.04	0.09	0.16	0.13	0.08	0.22	0.17	0.14	0.14	0.19	0.03	0.02	0.02	0.02	0.04	0.07	0.07	0.11	0.07	0.04	0.08	0.05	0.06	0.03	0.08	0.04	0.05
Protostylid LM2	rs2241729	0.02	0.14	0.23	0.25	0.22	0.27	0.24	0.15	0.16	0.29	0.02	0.01	0.02	0.02	0.03	0.30	0.33	0.27	0.32	0.23	0.35	0.01	0.01	0.04	0.01	0.01	0.00
Protostylid LM2	rs12299956	0.01	0.03	0.12	0.15	0.11	0.14	0.13	0.08	0.14	0.11	0.01	0.02	0.00	0.00	0.03	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Protostylid LM2	rs630603	0.02	0.07	0.25	0.23	0.23	0.32	0.18	0.23	0.28	0.27	0.01	0.02	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Protostylid LM2	rs9899063	0.03	0.07	0.25	0.25	0.21	0.26	0.32	0.20	0.24	0.26	0.02	0.03	0.01	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01
Shovel Shape LI1	rs3827760	0.25	0.24	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.39	0.24	0.48	0.76	0.17	0.87	0.90	0.94	0.91	0.80	0.82	0.01	0.00	0.06	0.00	0.00	0.00
Shovel Shape LI1 *	rs3827760	0.25	0.24	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.39	0.24	0.48	0.76	0.17	0.87	0.90	0.94	0.91	0.80	0.82	0.01	0.00	0.06	0.00	0.00	0.00
Shovel Shape UI1	rs3827760	0.25	0.24	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.39	0.24	0.48	0.76	0.17	0.87	0.90	0.94	0.91	0.80	0.82	0.01	0.00	0.06	0.00	0.00	0.00
Shovel Shape UI1 *	rs3827760	0.25	0.24	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.39	0.24	0.48	0.76	0.17	0.87	0.90	0.94	0.91	0.80	0.82	0.01	0.00	0.06	0.00	0.00	0.00
Shovel Shape UI2	rs3827760	0.25	0.24	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.39	0.24	0.48	0.76	0.17	0.87	0.90	0.94	0.91	0.80	0.82	0.01	0.00	0.06	0.00	0.00	0.00
Shovel Shape UI2 *	rs3827760	0.25	0.24	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.39	0.24	0.48	0.76	0.17	0.87	0.90	0.94	0.91	0.80	0.82	0.01	0.00	0.06	0.00	0.00	0.00

* Regrouped traits (Section 4.2.2.1.3, Table 4.6). Acronyms are presented in the section Acronyms, page 29.

Table B.4. Allele frequencies of index SNPs that showed genome-wide significant association to quantitative dental traits in different populations across the globe [387].

Associated Trait	Index SNP	MAF CAN	ALL	AFR	ACB	ASW	ESN	GWD	LWK	MSL	YRI	AMR	CLM	MXL	PEL	PUR	EAS	CDX	CHB	CHS	JPT	KHV	EUR	CEU	FIN	GBR	IBS	TSI
Bucco-Lingual Lw Lf Canine	rs16856377	0.09	0.18	0.43	0.34	0.24	0.39	0.46	0.52	0.57	0.44	0.10	0.08	0.09	0.11	0.13	0.19	0.11	0.24	0.20	0.21	0.19	0.03	0.04	0.02	0.01	0.02	0.03
Incisal Cervical Lw Rt Premolar 1	rs2883720	0.39	0.40	0.29	0.28	0.21	0.31	0.24	0.39	0.28	0.30	0.43	0.39	0.52	0.47	0.39	0.54	0.39	0.60	0.63	0.60	0.48	0.35	0.32	0.34	0.34	0.35	0.38
Meso-Distal Lw Lf Lateral Incisor	rs3827760	0.25	0.24	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.39	0.24	0.48	0.76	0.17	0.87	0.90	0.94	0.91	0.80	0.82	0.01	0.00	0.06	0.00	0.00	0.00
Meso-Distal Up Lf Lateral Incisor	rs3827760	0.25	0.24	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.39	0.24	0.48	0.76	0.17	0.87	0.90	0.94	0.91	0.80	0.82	0.01	0.00	0.06	0.00	0.00	0.00
Meso-Distal Up Rt Central Incisor	rs3827760	0.25	0.24	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.39	0.24	0.48	0.76	0.17	0.87	0.90	0.94	0.91	0.80	0.82	0.01	0.00	0.06	0.00	0.00	0.00
Meso-Distal Up Rt Lateral Incisor	rs3827760	0.25	0.24	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.39	0.24	0.48	0.76	0.17	0.87	0.90	0.94	0.91	0.80	0.82	0.01	0.00	0.06	0.00	0.00	0.00
Meso-Distal Lw Lf Premolar 1	rs10894347	0.15	0.18	0.05	0.03	0.02	0.04	0.06	0.08	0.04	0.05	0.18	0.19	0.16	0.19	0.17	0.30	0.36	0.24	0.31	0.23	0.36	0.14	0.14	0.09	0.14	0.17	0.16