Deciphering adaptation in the human genome

The case of CaSR in south-east Asian hunter-gatherers

Barbara Sinigaglia

DOCTORAL THESIS UPF / 2022

Thesis supervisor

Dr. Elena Bosch Fusté

DEPARTMENT OF MEDICINE AND LIFE SCIENCES

The single story creates stereotypes,
and the problem with stereotypes is not that they are untrue,
but that they are incomplete.
They make one story become the only story.

Chimamanda Adichie, *The danger of the single story*

Sugnu sempri alla finestra e viru a ranni civiltà
Ca ha statu, unni Turchi, Ebrei e Cristiani si stringeunu la manu
Tannu si pinsava ca "la diversità è ricchezza"
Tempi di biddizza e di puisia, d'amuri e di saggezza
Zoccu ha statu aieri, oggi forsi ca putissi riturnari
Si truvamu semi boni di chiantari

Carmen Consoli, *A finestra*

Acknowledgements

Llegando al final de este camino no puedo evitar de recordar el momento en el que lo empecé. Recuerdo exactamente la sensación de vértigo que sentí cuando llegó en mi correo la respuesta de que me aceptaban para hacer una estancia de prácticas en un laboratorio de Barcelona. De allí empezó esta aventura, en la que conocí personas maravillosas y aprendí cosas de valor inestimable que me permitieron crecer tanto profesionalmente como personalmente. Para ello tengo que agradecer tanto Elena como Francesc, que me dieron la oportunidad de entrar en la grande casa que es el IBE. Elena, no ha sido fácil llegar aquí, gracias por darme la oportunidad de trabajar en un proyecto tan ambicioso.

Pero en estos agradecimientos tengo que empezar desde el principio.

Ai miei genitori, che da sempre mi appoggiano, mi stimolano e che in questi anni sono stati costretti a fare i conti giornalmente con la testardaggine, le frustrazioni, sbalzi d'umore e preoccupazioni. Eppure, grazie a loro non mi sono mai sentita sola, perché ad accompagnarmi quotidianamente c'erano sempre tutte le frasi e gli insegnamenti che mi hanno tramandato, coscientemente o incoscientemente. A Mattia, simbolo esemplare di perseveranza ed ambizione. Ti ho visto superare tutte le difficoltà di questi anni difficoltà con una diligenza e un senso del dovere invidiabili. Ti ringrazio di essermi accanto con quei piccoli grandi gesti che riscaldano il cuore, tanto semplici quanto potenti, come una fetta di pane fresco con l'olio.

En estos años tantas personas me han enseñado muchísimas cosas, pero solo algunas dejaron una marca especial. Judit y Mónica, sois los pilares de Bioevo. En estos años me habéis acompañado en todo tipo de momentos, en las risas y en las no tan risas, tanto laboralmente como personalmente con la misma determinación e ironía. Siempre recordaré el momento en que os encontré sentadas en la plaza del PRBB con los cuchillos entre los dientes, preparadas para hacer todo el necesario para ayudarme cuando me encontré sin casa de un día para otro. Aunque

prácticamente no nos conocíamos, me transmitisteis protección y apoyo con una generosidad que nunca olvidaré. ¡Gracias!

Y los mejores los dejo para lo último. Simone, André, Carla, Ana, Neus, Rocío, Marco y Marcel. Los que lo han visto todo, viviendo conmigo cada paso y cada decisión. En muy breve tiempo os habéis convertido en la familia barcelonés que echaba de menos cuando volvía con mi familia italiana. Simone, che te devo di'? Sei stato come un fratello maggiore, sempre disponibile a darmi retta, tra una chiacchera e un consiglio, una battuta ed un caffè (o un calzone, una pizza, le orecchiette, il sushi [bonanima]) il tempo insieme passa sempre troppo velocemente. Mi hai accolta sin dal primo momento come se mi conoscessi da sempre, e piano piano sei diventato una delle persone imprescindibili. André, el hermanito juguetón. El que cuando está se nota porque es un volcán de energías y de chistes, es el que llena todas las conversaciones de lo mejor de la cultura popular. Hiciste tanto "ruido" en nuestro grupo que, aunque lleves ya un tiempo en fuera, nosotros seguimos aquí citando tus frases como fueran un legado. Marco y Marcel, no se en que idioma escribir: Marco te apañas que entiendes castellano, del revés no funcionaria. Vosotros llegasteis ya cuando el camino empezaba a ser cuesta arriba y yo andaba por el labo tipo un correcaminos. Desde el primer momento habéis entrado en nuestra rutina disfuncional añadiendo vuestro toque personal de humor que aligeraba los días.

Quiero agradecer a las mejores compañeras de viaje que podía encontrar. Carla, Ana, Neus y Rocío. No hay palabras suficientes para explicar lo importante que habéis sido a lo largo de estos años y el impacto positivo que habéis tenido en mi vida. Me enseñasteis qué es la determinación y la resiliencia, la aceptación y la amistad incondicional. Me acompañasteis en cada momento, haciendo que me cuestionara en cada paso, y luego buscando juntas las respuestas. Sois como el hogar durante una tormenta, cuando cierras la puerta detrás de ti y se apagan los ruidos externos. Solo en ese momento consigues reconocer cuanta lluvia te mojó, percibir las gotas que caen en tu piel, y apreciar el calorcito del interior. En los días de las peores tormentas supisteis estar

allí, nunca pretendiendo tener las soluciones para todo, sino acompañándome con un paraguas a buscar lo que había perdido por el camino.

Finalmente tengo que agradecer a toda la gente del IBE con las que en estos años he podido colaborar, intercambiar ideas y contestar preguntas científicas. Su aportación ha sido fundamental para el desarrollo de esta tesis doctoral.

Abstract

South-east Asian rainforest hunter-gatherers (SEA-RFHG) are characterized by specific morphological features including a particularly dark skin color (D), an extremely short stature (S), the development of woolly hair (W), and the presence of a localized fat accumulation in the hips called steatopygia (S) (DSWS phenotype). Using whole-genome sequences from different SEA-RFHG (Andamanese, Malaysians, and the Philippines) we characterized signatures of positive selection around the Calcium-Sensing Receptor (CaSR) gene and identified the R990G substitution (rs1042636) as a potential adaptive variant to experimentally follow up. At a systemic level, CaSR has a critical role in calcium homeostasis by regulating parathyroid hormone secretion. However, the gene is also expressed across a wide range of tissues being involved in many different biological functions. Thus, to investigate a putative adaptive phenotype at the organism level, we generated a knock-in mouse for the R990G substitution using the CRISPR-Cas9 technology and explored potential phenotypical differences related to CaSR functions between homozygote carriers of ancestral and derived alleles. We found that 990G homozygotes show, among other phenotypes, greater weight and fat accumulation, probably related to changes in the adipogenesis-lipolysis equilibrium. Such differential adipogenicity could facilitate an earlier sexual maturation, in agreement with the hypothesis that the DSWS phenotype could result from selection for an early onset of reproduction in a particularly hostile environment.

Riassunto

I cacciatori-raccoglitori della foresta pluviale nel Sud Est Asiatico presentano delle specifiche caratteristiche morfologiche che includono un colore della pelle particolarmente scuro (D), una statura estremamente ridotta (S), la presenza di capelli particolarmente ispidi (W) ed una accumulazione adiposa accentuata nella zona dei fianchi

definita steatopigia (S) (fenotipo DSWS). Attraverso l'uso di sequenze whole-genome appartenenti a diverse popolazioni del Sud Est Asiatico (Andamanesi, Malesi e Filippini), il nostro gruppo ha identificato segnali di selezione positiva in una regione genomica che include un recettore accoppiato a proteina G: il Calcium-Sensing Receptor (CaSR). All'interno del gene è stata rilevata la presenza di una sostituzione non-sinonima (R990G) associata all' rs1042636 che successivamente abbiamo scelto come variante potenzialmente adattativa da testare sperimentalmente. Il ruolo principale del CaSR a livello sistemico è quello di regolare l'omeostasi del calcio attraverso la modulazione della secrezione del paratormone (PTH). D'altro canto, essendo estremamente pleiotropico, il gene si esprime in molteplici tessuti in cui svolge differenti funzioni biologiche. Per queste ragioni abbiamo generato un modello di topo introducendo il cambio non sinonimo R990G attraverso le forbici molecolari CRISPR-Cas9 al fine di studiare le potenziali differenze ascrivibili alle molteplici funzioni del CaSR tra gli animali omozigoti per i due alleli (ancestrale e derivato) e dedurne un possibile fenotipo adattativo. I risultati ottenuti rivelano che gli animali omozigoti per l'allele 990G presentano, tra gli altri fenotipi, un incremento nell'accumulo di grasso, probabilmente dovuto ad un'accentuata adipogenesi ed alla riduzione della lipolisi. Queste caratteristiche potrebbero facilitare l'anticipazione della maturazione sessuale in accordo con l'ipotesi secondo cui il fenotipo DSWS nelle popolazioni di cacciatori-raccoglitori potrebbe essere stato generato in seguito ad una pressione selettiva che avvantaggiava la riproduzione anticipata in ambienti particolarmente ostili come la foresta pluviale.

Sinopsis

Los cazadores recolectores del Sudeste Asiático presentan varias características fenotípicas específicas como un color de piel particularmente oscuro (O), una estatura extremadamente baja (B), un pelo muy rizado (R) y una inusual acumulación de grasa en las caderas llamada esteatopigia (E) (fenotipo OBRE). A partir de la recopilación y análisis de secuencias genómicas completas de varias poblaciones humanas de las Islas Andamán, Malasia, y Filipinas, identificamos

señales de selección positiva alrededor del gen del receptor de detección de calcio (abreviado como CaSR en inglés)), y la sustitución no sinónima R990G como variante genética potencialmente adaptativa en estas poblaciones para validar experimentalmente. A nivel sistémico el CaSR actúa sobre la homeóstasis del calcio regulando la secreción de la hormona paratiroidea. Además, el gen se expresa en muchos tejidos y participa en diferentes funciones biológicas. Consecuentemente, para poder investigar el fenotipo adaptativo a nivel de organismo, generamos a un ratón knock-in por la sustitución R990G usando la tecnología del CRISPR-Cas9 y analizamos las posibles diferencias fenotípicas relacionadas con las funciones del CaSR entre los animales homocigotos por el alelo ancestral y derivado. Encontramos que los homocigotos 990G muestran, entre otros fenotipos, mayor peso y una mayor acumulación de grasa probablemente relacionada con cambios en el equilibrio adipogénesis-lipolisis. Estas características podrían facilitar una maduración sexual temprana, de acuerdo con la hipótesis que la baja estatura del fenotipo OBRE puede ser el resultado de la selección para un inicio temprano de la reproducción en un ambiente particularmente hostil.

Preface

Studying human diversity has interested researchers and explorers for centuries. Only in the last decades with the increasing amount of first genetic and later genomic data, scientists have been able to appreciate the high genetic similarity of all human populations (more than 99 % identity) and retrace human migrations through molecular fingerprints left in our genomes. Besides, results obtained from the Population Genetics field suggest that populations are melted into one another, their allele frequency variation is not discrete, but rather clinal and it is strongly affected by selective pressures experienced by populations in different environments.

In this thesis, we focused on the study of human adaptation in the tropical rainforest environment. Specifically, we analyzed tribal rainforest hunter-gatherer populations from South East Asia presenting certain morphological features including a particularly dark skin color (D), an extremely short stature (S), the development of woolly hair (W), and the presence of a marked fat accumulation localized in the hips called steatopygia (S) (DSWS phenotype). Analyzing whole-genome data from the Andamanese population, we retrieved signals of selection in the region encompassing the Calcium-sensing Receptor (CaSR). Here, the non-synonymous substitution R990G (rs1042636) was identified as one of the top 20 most differentiated non-synonymous variants between the Andamanese and mainland Indian populations showing signals of selection in the former group. Furthermore, we replicated the selection signal in the CaSR region in whole-genome data from Malaysians and Philippine populations also carrying the R990G substitution at high frequencies. Thus, we decided to study the phenotypic implication of the R990G substitution in a mouse model considering the following evidence in addition to the selection signals previously described: the R990G substitution was associated with a very high CADD score (18,42), indicating a strong functional impact; the substitution was also previously described as a gain of function, conferring the receptor higher sensitivity to extracellular Ca^{2+}; the

CASR gene is very pleiotropic, and it shows an active role in different biological pathways.
Results obtained from the mouse model showed a systemic unbalance in the lipid handling, being the homozygotes for the derived allele more prone to develop higher weight and BMI, higher fat percentage, higher number of small adipocytes in fat histology, and downregulation of lipolytic-related genes.

All the evidence gathered in this study suggests that the R990G substitution in rainforest hunter-gatherers from South East Asia favors an enhanced fat accumulation that could result adaptive in a difficult environment such as the tropical rainforest. In this line, the increased fat storage enhances energy storage and has been previously shown to result in a faster sexual maturation and early onset of menarche in girls, generating an anticipated closure of the growth plate in bones. Thus, we hypothesize that natural selection favored this mechanism in rainforest hunter-gatherer populations studied here generating their specific phenotype with a small stature and the development of the steatopygia.

When studying human adaptation in historically excluded populations, their social background and possible consequences of the research must be considered. Often, adaptation generates extreme phenotypes that have been historically used to racialize people and confine them to a life of social exclusion based on the assumption that the phenotype mirrored the psychological and moral qualities of individuals.

For these reasons, before moving to the body of this thesis, some considerations must be pointed out: neither clinal population differences nor any adaptative variants are enough to justify any kind of subdivision of the human race into biological discrete groups; any kind of deterministic association between genetics and psychological abilities or ethical values must be condemned, being the result of the resurrection of some old-fashioned nationalistic ideas; science and society are strictly interdependent, while on one hand, scientists are human beings socialized in a certain environment which could be

influencing their research, on the other hand, the study of human diversity from a scientific perspective has a strong impact on public opinion and politics, giving scientists the strong responsibility to design collaborative and respectful projects and disseminate consciously their scientific results.

Table of content

Abstract .... **i**

Preface .... **iv**

INTRODUCTION **1**

1. The human journey into South East Asia **3**

1.1. Genetic structure and demography of South East Asia 4
1.2. SEA Neolithic transition: in Archeological data 7
1.3. Human settling in the Sahul continent 11

2. Tribal rainforest hunter-gatherers of SEA, Papua New Guinea, and Australia **14**

2.1. Andamanese populations before and after the first contact with western society 14
2.2. Past and present of the hunter-gatherer tribes from the Malaysian peninsula 21
2.3. The Philippines hunter-gatherer tribes 22
2.4. Hunter-gatherers' settlement in Papua New Guinea 23
2.5. Australian Aborigines 25
2.6. Phenotype and ancestry: convergent evolution or phylogenetic relatedness? 28
2.7. Defining ourselves by creating the "otherness" 29

3. Calcium sensing receptor **33**

3.1. Molecular structure 33
3.2. The physiological function of CaSR 34
3.3. CASR is a pleiotropic gene 37
3.4. The CaSR R990G polymorphism: molecular phenotype, phenotypic associations 40
3.5. R990G frequency distribution in Asia and signals of selection 41

4. Detecting signals of selection and creating a mouse model **43**

4.1. Evolutionary forces and their impact on the genomes 43
4.2. Molecular fingerprints of positive selection 45
4.3. Finding adaptation in human populations 48
4.4. Gene editing history and applications 49
4.5. CRISPR-Cas9 50
4.6. CRISPR-Cas9 editing in mouse models: increasing the specificity. 52

OBJECTIVES **55**

RESULTS **58**

Exploring adaptive phenotypes for the calcium-sensing receptor polymorphism R990G **61**

Tables and figures 103

DISCUSSION **111**

5. Implication of CaSR in the systemic lipidic handling: **113**

5.1. The role of R990G substitution in adipocyte differentiation and lipolysis 113
5.2. The role of R990G substitution in liver 115
5.3. Connecting the dots: the R990G substitution and its systemic implication 116

6. Putative adaptive forces for an enhanced lipidic accumulation in populations with the DSWS phenotype **117**

6.1. Carrying a specific adaptive variant in a western-centric world 118

7. How can we talk about genetic differences between populations in a time of major historical changes? **120**

7.1. The never-ending story of the unknown's fear 120
7.2. Searching for non-self to find the self: an archaic instinct 121
7.3. The unexpected and underestimated impact of science on society 121
7.4. The responsibility of the correct scientific communication 123
7.5. Laying the foundation for a more conscious research 125

CONCLUSIONS **129**

REFERENCES **133**

ANNEX **119**

ANNEX: supplementary material 163

INTRODUCTION

1. The human journey into South East Asia

During the last 100 ky, since the Out of Africa, human populations have been spreading across the world occupying very different habitats in a relatively short period. The newly colonized environments acted as different selective pressures on the human genome generating, together with different demographic processes, the 10% variability present in nowadays populations (Fan et al., 2016; R. D. Hernandez et al., 2011; Rees et al., 2020). One of the harshest environments ever colonized by humans is the tropical rainforest, characterized by high temperature and humidity, as well as the presence of many pathogens and parasites. Furthermore, the rainforest was described as a challenging environment in terms of food accessibility, due to the scarcity of protein- and fat-rich fauna and the rare presence of carbohydrate-rich plants (Fan et al., 2016; Roberts & Petraglia, 2015). Despite these challenging conditions, several human groups settled in the tropical environment developing specific adaptations that allowed their survival (Migliano et al., 2007, 2013; Perry et al., 2014). In this thesis, I will examine the hunter-gatherer populations settled in tropical rainforests of South East Asia (SEA), illustrating their history and cultures as well as describing the population structure while focusing on the analysis of their genetic adaptation to the rainforest.

Before zooming in, I will here briefly delineate a summary of the SEA population structure that will be further detailed in the following sections. Archeological remains support the hypothesis that the first settlers of SEA arrived after the Out of Africa (OOA) around 70-50 kilo years ago (kya) following the south coastal route and colonizing a big landmass nowadays submerged called Sundaland (Aghakhanian et al., 2015). The majority of present-day populations that live in the area share a phenotypic affinity with East Asian populations, and additionally, archeological and linguistic data support the idea that the population diversity was strongly influenced by an expansion of farmers from the north (McColl et al., 2018). Additionally, the current SEA population also includes several hunter-gatherer minorities that live in the Andaman Islands, the Malaysian peninsula, the Philippine

archipelago, Papua New Guinea, and Australia. These minorities have been described to be descendants of the first settlers that arrived on the continent during the OOA (McColl et al., 2018). Traditionally, due to their distinct phenotype, their hunter-gatherer lifestyle, and culture, these tribal minorities were hypothesized to be monophyletic by anthropologists and archeologists (Endicott, 2013), but instead, a more complex scenario was later suggested by the genetic data (Endicott et al., 2003; Migliano et al., 2013). An in-depth description of the socio-cultural and phenotypic characteristics of these tribal minorities will follow in the next sections, here I will refer to them as "rainforest hunter-gatherers" (RFHG) as previously suggested by Fan et al. (2016) and I will describe the state of the art on their population history as inferred from archaeological data, historical records, linguistics as well as genetic structure and demography.

1.1. Genetic structure and demography of South East Asia

As previously mentioned, due to their phenotypic characteristics the SEA tribal RFHG were historically considered the descendants of an ancient African group of RFHG that peopled SEA, however, ancient mtDNA haplogroup characterization revealed instead an Asian origin for the RFHG from the Andaman Islands, Malaysia, and Philippines (Endicott et al., 2003; Reich et al., 2011).

Through the analysis of modern SEA genomes looking for an ancient Denisovan introgression on genotyping data, Reich et al. (2011) were able to describe the population movements during the first settling of SEA and Oceania. A very low signal of Denisovan introgression was found in the mainland SEA, in the modern RFHG Onge population native of the Andaman Islands, and in the Malaysian RFHG Jehai. Though, a signal of admixture was found in the Jehai group coming from SEA populations. On the contrary, the highest introgression signal was retrieved in Australia and New Guinea while this signal was diluted in the Philippines, South-eastern Indonesia, and Polynesia

(figure 1). Evidence of Denisovan introgression was also found in the tribal RFHG group of Mamanwa that live in the Philippines. These findings suggested that the Andaman RFHGs were the only group showing a very low signal of introgression from Denisovan and no admixture from SEA, thus resulting to be the closest group to the ancestral population.

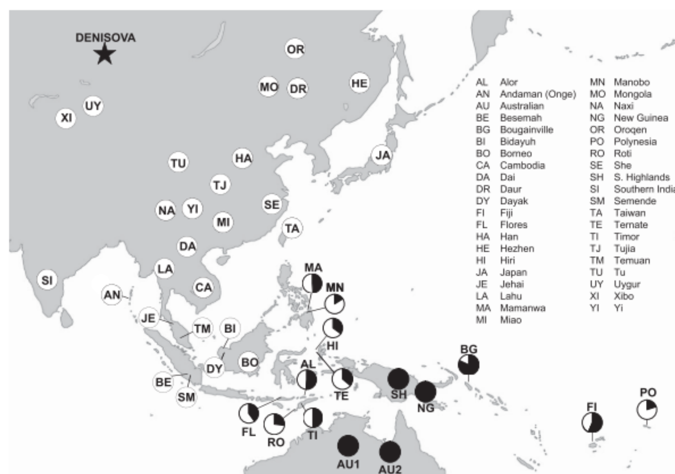


Figure 1: Denisovan Genetic Material as a Fraction of that in New Guineans from Reich et al. (2011)

In this line, whole-genome sequencing was performed on ancient DNA by McColl et al. (2018) and analyzed together with previously published modern samples from worldwide populations. Interestingly, the two oldest samples associated with the Hòabìnhian culture, one dating from 7,950 to 7,795 kya and the other from 4,415 to 4,160 kya, were found to be closely related to the Onge population, while the Jehai population was described to be an admixture between the oldest ancient samples and Neolithic-Bronze age individuals. Similar results were obtained in another study analyzing genotype data of the "Orang Asli" (OA) community (literally "Original People"), which are the tribal hunter-gatherers that constitute 0,6% of the Malaysian population (Aghakhanian et al., 2015). Here, the Treemix tree (figure 2 A) showed the two RFHG populations of the Andaman Islands (Onge and Jarawa) to be closely related and grouped in a deep clade. The rest of the Malaysian OA populations clustered together and were found to be

different from the rest of East Asia, supporting the idea that they are descendants of the Hòabìnhian hunter-gatherers that occupied the Sunda peninsula during the late Pleistocene. Further analyses on genome-wide data from the Philippine tribal RFHG populations of Agta, Aeta, Batak, and Mamanwa were performed by Jinam et al. (2017) together with genome-wide data from the Andamanese and the Malaysian OA population that self-identify as "Negrito". In the Neighbor-Joining tree (figure 2 B), all the hunter-gatherer populations included in the study appeared as basal to all the other SEA populations. Furthermore, the authors inferred the introgression of Denisova into the genome of the Philippine populations Aeta, Agta, and the admixture of the Philippine and Malaysian hunter-gatherer populations with the surrounding non-hunter-gatherer.

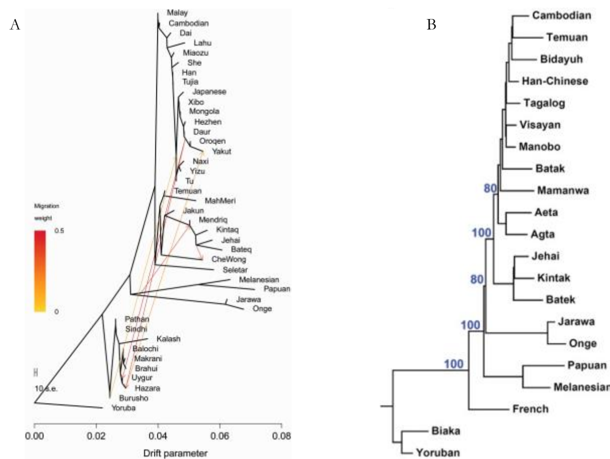


Figure 2: Phylogenetic relationships among SEA populations. A. Treemix tree from Aghakhanian et al. (2015). B. Neighbor-Joining Tree from Jinam et al. (2017)

Taken together, all the results reported here show a complex history for the peopling of SEA (figure 3) with the first settlers arriving from the southern coastal route (Aghakhanian et al., 2015) probably during the last glacial maximum, when the sea level was lower and the nowadays islands formed the Sundaland landmass (Chaubey & Endicott, 2013). The first wave of settlers reached the Andaman archipelago (where they became isolated) and spread through the Sundaland, where numerous nowadays Malaysian tribal RFHG groups are found to be carrying the

genetic signals of the ancestral Hòabìnhian hunter-gatherers (Aghakhanian et al., 2015). The ancestral settlers reached then the Philippines, Melanesia, and Papua where they met the indigenous Denisovan inhabitants of the area in Papua New Guinea and Australia (Aghakhanian et al., 2015; Jinam et al., 2017; Reich et al., 2011) receiving the introgression from the Denisovan population that Meyer et al. (2012) estimated to reach the 6% of the Papuan genomes.

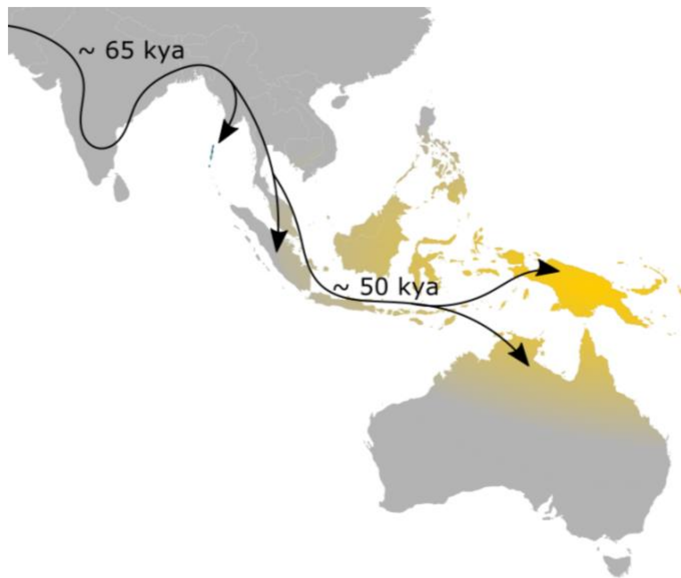


Figure 3: Settlement in SEA of the first hunter-gatherer groups. The black lines indicate the possible routes taken by the first settlers, whereas in yellow is represented the areal distribution of Denisovans.

1.2. SEA Neolithic transition: in Archeological data

Traditionally, archeological studies associated the Austronesian expansion into South East Asia with the spread of the Neolithic lifestyle in the area, leading to a cultural and linguistic replacement. Yet, Philip et al. (2017) combining archeological, linguistic, and genetic data identified for the Neolithic transition, on the one hand, an Eastern Migration Route (EMR), that was previously defined as the Out of Taiwan theory, and on the other hand, a Western Migration Route

(WMR) that involved people moving from China through Vietnam and Malaysia until Sumatra and Borneo. The group of people that entered SEA through the WMR have been associated with the cord-marked pottery which is a marker of the early Neolithic and was described to be related to the population movement that spread the Austroasiatic languages. No evidence of the corded-marked pottery was found in the EMR, where instead the archeological sites revealed the production of the red-slipped pottery that was associated with the Austronesian language expansion, as shown in figure 4 (Choin et al., 2021; Philip et al., 2017).

In summary, two waves of Neolithic expansion can be identified by analyzing the archeological data, both of them starting in south China. Following the WMR, A. Anderson (2005) described the expansion of the cord-marked pottery and the Austroasiatic languages that he defined as markers of the Neolithic expansion I. At the same time, the so-called Neolithic expansion II involved a population expansion following the EMR carrying the red-slippery pottery and the Austronesian languages from Taiwan towards the south, until reaching and replacing the Austroasiatic expansion occurred westwards.

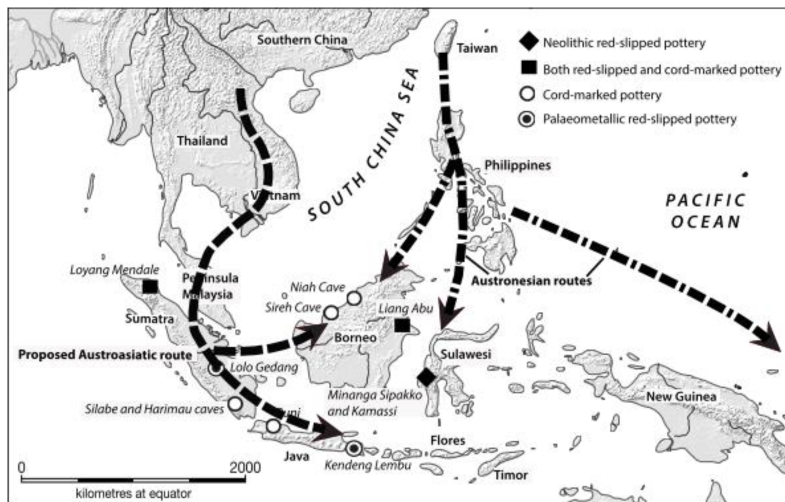


Figure 4: Proposed route of Austroasiatic and Austronesian migration into Indonesia. The figure shows the geographic distribution of sites that have produced red-slipped and cord-marked pottery from Philip et al. (2017)

SEA Neolithic transition: in linguistic and genetic data

Modern Austroasiatic languages are spoken nowadays by more than 65 million people, and most of them are characterized by many dialects without a written tradition. Culturally, the most important languages with a remarkable written record are Khmer, Mon, and Vietnamese, spoken in Vietnam and Cambodia (Diffloth, 2018), while the rest of the Austroasiatic speakers are scattered in remote mountains or are part of the Malayan tribal RFHGs as shown in figure 5 A (Benjamin, 1976; Toshihiro, 2009).

On the other hand, the Austronesian languages constitute the largest language family in the world and are spoken from Madagascar to Polynesia (R. A. Blust, 2018). Most of the Austronesian languages variability is contained within Taiwan, while the rest are Malayo-Polynesian speakers, as shown in figure 5 B. Genome-wide analysis of present-day populations from Austronesian-speaking Island South East Asia (ISEA) identified the presence of an Asian ancestry component associated with the modern Austroasiatic speakers. In this study, Lipson et al. (2014) explain their results hypothesizing the existence of an admixed Austroasiatic-Austronesian ancestral population in Vietnam and Malaysia that subsequently spread into the ISEA leading to the radiation of the Malayo-Polynesian Austronesian languages.

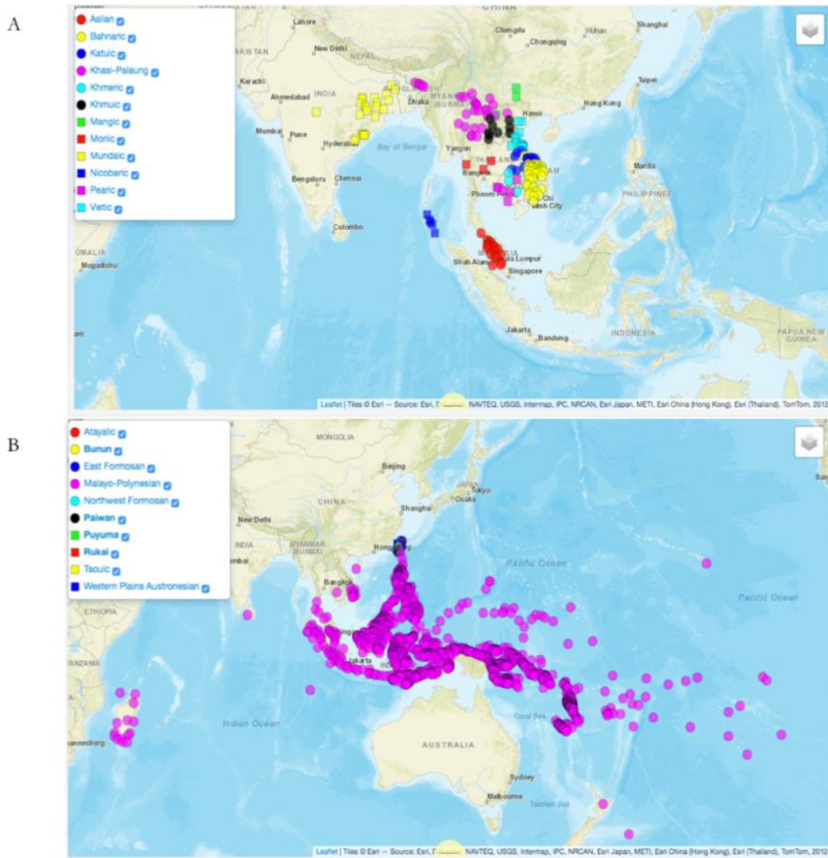


Figure 5: Modern linguistic distribution. A. Austroasiatic languages, B. Austronesian languages. from Glottolog 4.6 (Hammarström et al., 2022)

Considering all the evidence together, we can fairly say that the nowadays SEA tribal RFHG are the descendants of the first populations that arrived in Sundaland during the late Pleistocene in a single-wave migration from central Asia. Archeological findings revealed that these populations were of Hòabìnhian culture and spoke a language that is now lost. When the sea level raised, they remained isolated, and later on, they were forced to move to more remote areas (islands, remote mountains, and rainforests) by the subsequent Neolithic migrations from the mainland. The Sundaland people that settled in Malaysia acquired Austroasiatic languages by the Neolithic migration wave through the WMR, while the Sundaland people in the Philippines acquired Austronesian languages from the EMR (figure 6).

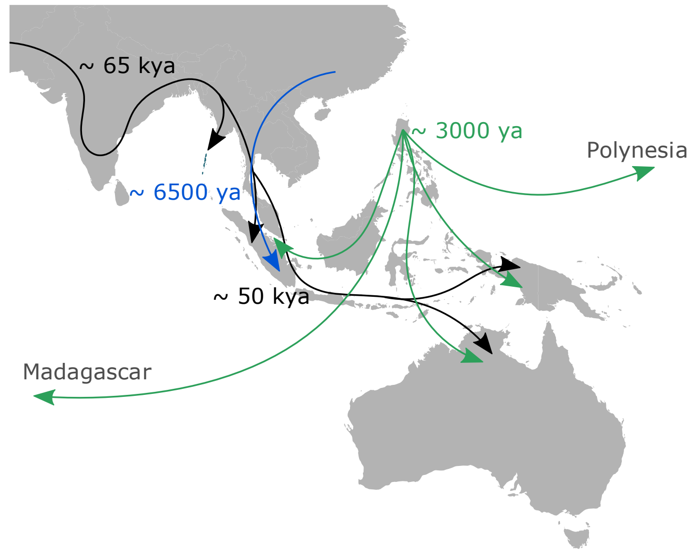


Figure 6: Plausible migration routes into SEA. Black arrows indicate the first settlement after the Out of Africa. The blue arrow indicates the WMR followed by the Austroasiatic expansion. Green arrows show the Austronesian expansion through the EMR.

1.3. Human settling in the Sahul continent

As previously described, during the late Pleistocene the South East Asia geomorphology was strongly characterized by lower sea level and the emergence of two continental landmasses: the Sunda Shelf and the Sahul continent. The latter included the nowadays Australia and New Guinea, and it was separated from Sundaland by a major ecological barrier of deep oceanic water containing small islands: the Wallacea (Brucato et al., 2021; O'Connell & Allen, 2010). Archeology and genetics agree on dating the colonization of the Sahul continent back to ~50 kya when the first wave of settlers from Sundaland crossed the Wallacea (Brucato et al., 2021). The routes followed by the first groups for Sahul's colonization are still under debate, but recent studies on both mitochondrial and autosomal genomes revealed that probably an ancestral population inhabiting Sundaland entered the Sahul continent in two waves, one settling the Northern Sahul (that now corresponds to New Guinea) and the other entering in the Southern Sahul (which

today is Australia) (Malaspinas et al., 2016; Pedro et al., 2020). In this line, Pedro et al. (2020) inferred that the ancestral Sunda population must have been already characterized by a strong population structure since the mitochondrial haplogroups variability of modern samples from Island South East Asia, Oceania, and Australia revealed that Northern and Southern Sahul hosted different lineages that rooted back to 50 kya during the first stages of the settlement. Additionally, the analysis of whole-genome sequencing data from Papua New Guinea, Indonesia, and Australia also supported the two-wave settlement hypothesis, revealing, however, a complex scenario for the peopling of the Sahul continent including settlement and gene flow while not totally excluding the possibility of the single wave migration (Figure 7) (Brucato et al., 2021).

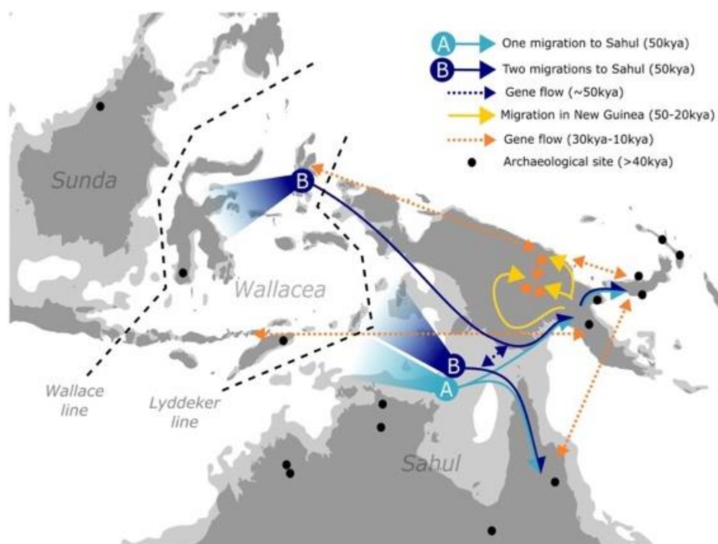


Figure 7: Genomic scenario of the human dynamics in north Sahul during the Upper Pleistocene. Light blue arrows represent the settlement of north Sahul following a scenario with one migration to Sahul (model A). The dark blue arrows represent the settlement of north Sahul following a scenario with two migrations to Sahul (model B). The dotted arrows represent gene flow, the dotted orange arrows represent gene flows after the initial settlement of Sahul. The black dots represent approximate locations of archaeological sites older than 40 ka, while the dashed black lines represent ecological lines in Wallacea. The dark gray areas represent current land area, with light gray areas representing the estimated land area around 50,000 years ago based on paleogeological reconstructions. From Brucato et al. (2021)

Furthermore, archeological data showed the presence of human settlements widespread in the entire Sahul dating back to 40-45 kya indicating a rapid population movement in the continent (O'Connell & Allen, 2010). In this line, Allen & O'Connell (2020) pointed out that the first colonization occurred in less than a millennium involving a group of fewer than 2,000 people that faced drastic sea-level falls and rapid changes in land accessibility. Hence, the authors suggest that "there is now consensus in the literature that all routes were possible and choosing the primacy of one over another seems only to depend on which variables are stressed." It is also important to bear in mind that present-day Papua New Guineans and Australians show the highest proportion of Denisovan introgression in their genomes as previously discussed (Reich et al., 2011 and figure 3), and recent analysis of whole-genome sequencing data corroborated these results postulating that the Denisovan introgression occurred before the split between the Aboriginal Australians and Papuans, hence before entering the Shul continent (Malaspinas et al., 2016).

2. Tribal rainforest hunter-gatherers of SEA, Papua New Guinea, and Australia

The tribal RFHGs of SEA comprise the Andamanese people, which can be further distinguished into the Onge, Jarawa, and the Great Andamanese, as well as several tribal RFHG groups from the Philippines, Malaysia, Papua New Guinea, and Australia. In the following sections, I will give some socio-cultural and historical insights regarding the RFHG groups to better understand their historical background and corresponding consequences on their lives in present days. Additionally, I will present a detailed description of the phenotypic characteristics shared by the aforementioned RFHG groups also giving some insights into how carrying such phenotypes could strongly impact the past and present hunter-gatherer's conception and their socio-political place in society.

2.1. Andamanese populations before and after the first contact with western society

The Andaman Islands are an extension of the Arakan mountains into the Indian Ocean and probably were part of a bigger landmass closer to the mainland during the last glacial maximum (R. Blust, 2013). Genetic studies on ancient and modern samples revealed that the first settlers of SEA arrived in the archipelago carrying the genetic signals of an ancestral population associated with the Hòabìnhian culture (Endicott et al., 2003; McColl et al., 2018). Mondal et al. (2016) described a shared ancestry between the Andaman population and the rest of the Asian and Pacific populations, supporting a single wave of migration from Africa for the peopling of SEA. The authors also found a signal of Neanderthal introgression in the Andamanese genomes, as well as a proportion of Denisovan introgression similar to the rest of Asian populations. In this work, the data also showed the presence of a reduction in the African ancestry in the Andamanese genomes which was interpreted as a signal of introgression from an extinct archaic population.

The first inhabitants of the Andaman Islands are considered to have become isolated after they arrived in the archipelago, allowing them to maintain the language originally spoken by the first settlers. They are considered a linguistic isolate speaking neither Austronesian nor Austroasiatic languages (Endicott, 2013). However, the present-day variability of the Andamanese languages is quite scarce due to the population decline experienced after the arrival of the English colonizers. Endicott et al. (2003), indeed, revealed that in the mid-19th century 13 different linguistic groups were known, but only 3 of them survived (figure 8).

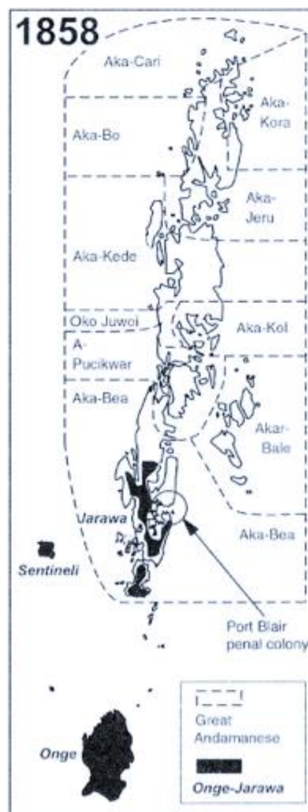


Figure 8: Original distribution of the communities in the Andaman Islands in 1858. From Endicott et al. (2003)

Despite the high linguistic heterogeneity, four main populations were described at the first contact: The Great Andamanese who occupied the North, the Middle, and part of the South Andaman archipelago, the Jarawa that lived in the South Andaman, the Onge inhabiting the Little Andaman island at the very south, and finally, the Sentinelese who lived in the Sentinel Island near the Jarawa's territory, with which they shared phenotypic and linguistic similarities (Endicott et al., 2003). Stock & Migliano (2009) reported that when the firsts English colonizers arrived in the archipelago in 1858, they made contact with the Great Andamanese population, which was the largest population numbering almost 6,000 individuals subdivided into different linguistic groups[1]. Soon after their arrival, a British penal colony was established in Port Blair in South Andaman (C. Anderson, 2018), and the Great Andamanese were encouraged to live in residential homes supervised by the convicts. Here a series of infectious diseases such as influenza, tuberculosis, syphilis, and measles started to spread among the populations, and those who became sick started to return to the rainforest. In 1871 the total number of convicts in Port Blair amounted to 8,643, while the Indigenous[2] Andaman populations were estimated to include a total of 2,847 individuals (C. Anderson, 2018). After a few decades, in 1900, the Great Andamanese population counted 600 individuals, and sixty years later the number dropped to 19 people (Stock & Migliano, 2009). On the other hand, the Onge and Jarawa populations always kept their distance from outsiders. It was estimated that in 1900 the Onge population was around 600 people, but during the first half of the 20th century, more than 70% of the population was wiped out by epidemics. Besides, the Jarawa population always refused any kind of contact with outsiders, rejecting every attempt to establish some relationship. They maintained their isolation until the 1970s (the

[1] It is worth noticing that the name was assigned by the colonizers just because they considered all the populations living in the Andaman archipelago as homogeneous, only on the bases of the physical appearance. Taking in consideration the high linguistic variability, most probably many different populations were grouped under the same name "Great Andamanese".

[2] The use of the capitalized letter for the word "Indigenous" identify a group of political and historical communities that suffered the consequences of colonialism, settler governments, displacement, and exploitation (Weeber, 2020)

Sentinelese until 1991), and their territory was officially recognized in 1932 (Stock & Migliano, 2009).

The Islanders' skepticism towards the newcomers should be interpreted always taking into consideration that the means for the establishment of the "friendly relationships" with the Andamanese included systematic kidnap of Indigenous peoples, a new dress-up with foreigners' clothes, days of captivity, and different kind of violence (from physical to psychological). Sometimes, the kidnapped people were brought to see the outside world hoping that they would learn about "The Civilization" and persuade the rest of the group to accept the outsiders. In reality, what often happened was either they did not survive the journey for the contraction of infectious diseases, or they would get ill and spread the disease on their return to the Islands (C. Anderson, 2018; Sen, 2009a; Venkateswar, 2004).

Great Andamanese culture and recent history

All the Indigenous populations of the Andaman Islands follow a hunter-gatherer lifestyle, having a diet based on fishing and hunting and recollecting roots and seeds from the rainforest. Their culture and religion are particularly entangled with the seasonality of the winds and the storms.

Bxown (1909) described some of the beliefs that constituted the Great Andamanese culture identifying the personification of the North-East monsoon (Biliku) and the South-West monsoon (Tarai). Biliku was also described to be the creator of the land and the sky as well as the keeper of the fire that created the sun in an impetus of rage, striking a firebrand in the sky after Andamanese's ancestors stole her the fire. In many legends, she would also manifest herself with thunderstorms and lightning when she would get angry for some disrespectful behavior of the Islanders: i.e., burning beeswax, or eating some roots and fruits during the wet season. The development of these legends is not trivial if we take into account that the honey season for the Andamanese

started at the end of the dry weather which would be followed by the beginning of the wet season, which often started with a strong thunderstorm. To recollect the beeswax, it was necessary to melt it down, hence it is understandable why the natives saw a cause-effect relationship between the melted wax and the following violent storms. However, Blust (2013) points out that this system of beliefs based on the punitive thunderstorms constitutes the "Thunder-complex", and it can be used to retrace the movements of the ancestral populations that peopled SEA since it is shared by all the tribal RFHG groups.

Today, the government of the South Andaman district (https://southandaman.nic.in) reports that the present-day Great Andaman population counts about 43 individuals that live in houses provided by the government and work at the coconut plantations. For clothes and food, they rely on the rations that the government distributes as well as on some sporadic hunting and fishing. Furthermore, in August 2020 four cases of Covid-19 were detected among people living in the tribe, probably contracting it during the movement between Port Blair and their island (Biswas, 2020). Fortunately, all the cases were promptly quarantined and recovered from the disease.

Onge's culture and recent history

Like the Great Andamanese, the Onge population has been described not to have a structured religion but rather as having a group of beliefs that are used to interpret nature (Venkateswar, 2004). They once occupied the entire Little Andaman Island where they had a hunter-gatherer lifestyle while later, they were displaced in Dugong Creek and South Bay of Little Andaman areas where the Indian government decided to confine them in 1976 to ensure the welfare of Indigenous Andaman communities (Gupta, 2017). Contrarily, the nomadic hunter-gathering lifestyle practices have nowadays been replaced by a more settled life, working in coconuts plantations and relying on the free rations of food provided by the government. As a consequence, the

lifestyle shift has had a strong impact on the Onge population generating a rapid cultural decline. As reported by Venkateswar (2004) in a long conversation with the Onge Bara Raju, the current Onge people lost part of their cultural identity: they consider themselves “less Onge” than the “old-time Onge” and for this reason, they neither can perform the traditional rituals, nor they are willing to learn. The displacement into the Andaman homes and the consequent cultural shift and identity loss coincided with the government’s development plans for the Little Andaman involving the clearing of the forest for agricultural purposes and the establishment of forest-based industries and villages. Venkateswar (2004) recognizes in this process a program of internal colonization in which the previous European colonizers are now substituted by the same Indian government. However, some of their ancestral knowledge survived thanks to the strong oral tradition passed along through the generations. A clear example of that was seen in 2004 after the strong earthquake that affected all of South East Asia. On that occasion, the Onge of South Bay were able to recognize the first signs of danger and rapidly escape into the forest before the tsunami inundated their land (Sandip, 2014).

Jarawa’s culture and recent history

As previously mentioned, the Jarawa population used to occupy the South Andaman forest, but they were forced to move when the first attempt to establish a new colonial territory occurred. They had no choice but to enter into the territorial boundaries of the Aka-Bea-da group (part of the Great Andamanese population) exacerbating the conflicts for the land (Venkateswar, 2004). The Jarawa have been described as the most controversial population of the Andaman archipelago, for their hostility and rejection toward the newcomers. The Victorian explorers described them either as skillful warriors or as violent beasts, in dependence on what the narration necessitated (Chatterjee, 2020). They were the only population to obtain official recognition of their territory by the government (Stock & Migliano, 2009), and their contact with outsiders has always been sporadic. Since

the Indian independence, the government has applied different policies regarding the relationships between the Indigenous people of the Andaman Islands and the settlers, trying to establish specific tribal areas. However, there has always been a small portion of the public opinion that believed in the assimilation of the Jarawa into today's mainstream population. Even if a certain distance with the tribe was maintained, the Jarawa's isolation was further compromised in the 1970s with the construction of a truck road that connected the North Andaman with the South Andaman. The road arrived into the Jarawa's territory, facilitating the interaction with the outsiders and the diffusion of alcoholism, sexual abuse, and infectious diseases among the tribal group. In 2007 the Indian government established a buffer zone around the reserve, for discouraging contact, but with scarce success (Talukdar, 2017). In fact, in 2012 public opinion was shocked by the diffusion of a video showing an encounter between a tourist and a tour guide with the Jarawa population in which the tourist offered food to the Jarawa's women in exchange for a dance. The video brought attention to the ethical implication of the practice of "human safaris" and raised some questions on how to balance the equilibrium between the Indigenous and non-indigenous populations (Liljeblad, 2014). The discussion on whether to include the Jarawa into the mainstream population or keep them isolated in the reserve is beyond the scope of this thesis, but it is important to acknowledge the conundrum that they are facing in dealing with modern society. Is modernization necessary for the Jarawa? And if not, is it right to isolate them into a small area of their own Land? And if we stand to let them choose whatever lifestyle they want, are we capable to leave them to exercise their culture on their own terms without trying to interfere or exploit them for economic, social, or political purposes?

2.2. Past and present of the hunter-gatherer tribes from the Malaysian peninsula

The tribal hunter-gatherer populations living in the Malaysian peninsula constitute an ethnic minority reaching a frequency of 0,6% of the population. They were assigned the name of "Orang Asli" (OA) meaning "original people" by the Malaysian government during the 1960s, including in this definition various Indigenous populations of Malaysia. Surprisingly, even if the Orang Asli people see that they were classified in a heterogeneous group, they actually tend to identify themselves more with the OA definition and not accordingly to their individual population (Toshihiro, 2009). As previously mentioned, the OA community speaks Austroasiatic languages (Endicott, 2013; Toshihiro, 2009) and is genetically close to the Asian populations that first settled in SEA (Aghakhanian et al., 2015; Migliano et al., 2013). Traditionally they occupied the Malaysian forest pursuing a hunter-gatherer lifestyle, but since the government adopted its integration policies they were relocated and encouraged to adopt a more settled lifestyle and work on the farms. However, the process of integration has not been successful on one hand because of the prejudice that the mainstream population still has towards the animist OA community, being considered pagans, and on the other hand because of the exclusion that the OA feels coming from the mainstream population (Toshihiro, 2009). Furthermore, Toshihiro (2009) describes the perception that the Malaysian have of the OA community as "primitive people inhabiting the forest with dangerous magical powers". In reality, the OA group can be divided into three main subgroups: the self-identified "Negrito" group that lives in the north of the peninsula as hunter-gatherers, the Senoi group that live in the central region and practice the slashing and burning cultivation, and the Proto-Malays that are actually farmers that also hunt and gather. R. Blust (2013) described the presence of the "thunder-complex" among the Orang Asli's beliefs, which can be seen as the cultural fingerprint of the ancestral hunter-gatherer population that spread into SEA in the late Pleistocene.

2.3. The Philippines hunter-gatherer tribes

The presence of humans in the Philippines dates back to the late Pleistocene and was detected in the rainforest of Luzon island (Roberts & Petraglia, 2015). As previously described (chapter 1.1), genetic data suggested that all the present-day hunter-gatherer populations from the Philippines and Malaysia fall in the basal branches of SEA populations confirming that they are related to the most ancestral inhabitants of the area (Jinam et al., 2017). The first contact of the western world with the Philippines people was in the 16th century when Magellan led a European expedition in service of Philip II of Spain (Borja, 2005). The first Spanish colonizers described a quite organized society with groups of farmers inhabiting the lowland that traded their products with hunter-gatherer tribes that lived in the rainforest. Archeological and linguistic evidence point to a long prehispanic history of mutual dependence between the inland tribal forest hunter-gatherers (that provided row materials, pelts, meats, and wild tuberous) and the coastal populations (that traded iron tools, weapons, pottery, and food). The establishment of this dynamic has had a strong impact on the tribal hunter-gatherer culture, that on one hand developed seasonal movements to the lowlands to collect wild resources to trade, and on the other hand acquired Austronesian languages related to the ones of their neighboring farmers (Junker, 2010). Furthermore, contact with the farmers has also changed the hunter-gatherer's diet, introducing rice as a fundamental element in their diet (Headland, 1991; Junker, 2010).

Nowadays the number of hunter-gatherer groups that live in the forest is decreasing. An ethnological study conducted by Headland in 1991 reported indeed that the increasing deforestation for cultivation purposes is shrinking the rainforest area and decreasing the number of wild animals that hunter-gatherers used to hunt. Therefore, hunting and fishing practices are decreasing while the diet is relying more on crops and rice coming from the outside. On top of that, recent studies revealed that the Philippine tribal RFHGs are undergoing a process of deculturation, caused by a traumatic cultural change in the 20th century linked to wild deforestation, and resulting in a feeling of cultural

subordination towards the mainstream culture. Hence, the younger generations are not interested in learning about their ancestral culture while the older ones seek to survive without preserving their traditional culture (Minter, 2010). However, since the early 1990s, the government adopted a more pluralistic and inclusive administration and installed the National Integrated Protected Area System which aimed to preserve the rainforest biodiversity allowing for the proclamation of protected areas under the direct management of the Indigenous tribes. As a consequence, the hunter-gatherer groups acquired the land right of the protected areas and could decide their path of development (Minter, 2010).

2.4. Hunter-gatherers' settlement in Papua New Guinea

Papua New Guinea (PNG) was rapidly settled soon after the human arrival in Sahul, from the nowadays south-east part of the region. From the analysis of whole-genome sequencing data, it was shown that PNG's settlement firstly affected the lowlands which were the most hospitable places during the earliest stages of colonization. Later on, the progression towards the Last Glacial Maximum contributed to a reduction in the resources' accessibility in the lowland, inducing the population to move into the highlands. Furthermore, gene flow was also detected between the lowlands and the highlands suggesting a semi-permanent settlement of the highlands during the first stages of the colonization, with continuous contact of the nomadic groups with the lowlanders. No gene flow was detected between 20 kya and 3 kya, suggesting that the contact between the two groups decreased in time, causing an increase in the genetic differentiation between them (Brucato et al., 2021). In this line, the present-day population differentiation between highlands and lowlands was described by Bergström et al. (2017) to have been developed during the last 20 ky, while the highlands population structure developed during the last 10 ky when the agricultural practices started to spread across the highlands independently from the arrival of an external population.

The agricultural invention in PNG

The cultural shift from the hunter-gatherer lifestyle to the implementation of farming practices in New Guinea has been at the center of the debate since it occurred in a very short time and with limited spatial distribution (Shaw et al., 2020). Indeed, when analyzing the evidence of the cultivation practices in New Guinean highlands, Denham (2004) did not associate the new habits as an invention of the Papuan hunter-gatherers, but rather as a result of the New Guinean people's interaction with the population of the West Island South East Asia that had already undergone to the Austronesian cultural shift. On the contrary, other authors suggested that by the time the Austronesian population arrived in PNG 3 kya, a complex cultural structure was already present on the island (Brucato et al., 2021). Archeological studies have indeed revealed the presence of tools for agricultural practices in the highlands dating back to at least 1,000 years earlier than the Austronesian migration into Papua New Guinea supporting the hypothesis that the introduction of agriculture in the highlands occurred without any contact with an external population (Shaw et al., 2020).

From a linguistic perspective, Bergström et al. (2017) revealed that the present New Guinea language diversity includes more than 850 different languages and that most of the language variability is caused by the presence of more than 20 unrelated autochthonous language families that were historically grouped under the name "Papuan", to differentiate them from the Austronesian languages also present in the island. As previously discussed (chapter 1.2), the Austronesian expansion carried across Island South East Asia a complex of cultural traits associated with the Neolithic transition that reshaped the genetics, culture, and language of South East Asia. However, in New Guinea, the Austronesian languages had less impact on the island's linguistic variability being nowadays spoken by small groups scattered along the coasts. Linguistics has also shown that the high diversity of the Papuan languages developed during the last 10 ky, during the population

expansion related to the development of the agricultural practices in the highlands (Pawley et al., 2005).

From the analysis of PNG mtDNA haplogroup distribution it has been shown that among groups of different ecozones (island, coast, mountain, and plain), language constituted the stronger barrier to the gene flow between the Austronesian speakers and the Papuan groups (Vilar et al., 2008). Moreover, when analyzing genome-wide data and calculating the population differentiation among Papuan speakers, the authors found that the genetic distance among PNG populations was much larger than the one between European and East Asian populations. On top of that, high F_{st} values were obtained when comparing highlanders *vs* lowlanders but also eastern *vs* western highlanders and southern vs northern lowlanders indicating that the differentiation was driven not only by geography but also by cultural-linguistics factors (Bergström et al., 2017).

Nowadays population distribution in PNG territory has been described to be heterogeneous, being influenced by the presence of some endemic diseases such as malaria, which is particularly frequent in the lowlands (Riley, 1983). A recent report from the *United Nations News* (UN News, 2022) detailed an increase in violence and abuses against the tribal PNG groups, reporting child killings, disappearances, torture, and enforce mass displacement. The socio-political situation of West Papua is particularly tense, with a strong militarization of the country and a growing people's discontent over the government policies of the land's exploitation that complicates the Indigenous people's survival ("Indonesia Attack: Gunmen Kill 24 Construction Workers in Papua," 2018).

2.5. Australian Aborigines

Archeological sites older than 40 ky are widespread in all of Australia, testifying the rapid movement of people that entered the continent from Sunda and migrated southward (O'Connell & Allen, 2010). Malaspinas et al. (2016) described the rapid population expansion by

analyzing whole-genome sequencing data in Aboriginal Australians identifying that populations from the south-western desert diverged from the northeastern regions around 31 kya. The authors hypothesized that Australia's colonization followed the coastal routes since the internal desert acted as a barrier to migration. This hypothesis was later corroborated by the creation of a migration model merging paleogeographic and hydrological data with physiological data from hunter-gatherers. The model suggested that people adopted a traveling strategy that included moving among Australia's water sources and choosing the most optimal route with higher visibility and being less calorically expensive. According to the authors, this strategy allowed people to spread all over the continent in a short period (Crabtree et al., 2021). Analysis of mitochondrial haplogroup variability revealed that the oldest lineages present in Australia (haplogroup P) are rooted back in the ancestral population settled in Sundaland before the Sahul colonization. Furthermore, unidirectional gene flow from northern Sahul (PNG) to southern Sahul (Australia) was detected during the last glacial maximum and through the Holocene (haplogroups M and Q), which stopped in the late Holocene (figure 9) (Pedro et al., 2020). The presence of a short-statured group of Aborigines in the northern Australian rainforest was described since the arrival of the first European colonizers and was soon interpreted as evidence of the primordial settlement closely related to the ancestral population that also settled in South East Asia. However, the mitochondrial haplogroup distribution revealed that the population shared many more haplogroups with Australian Aborigines than with South East Asia (McAllister et al., 2013).

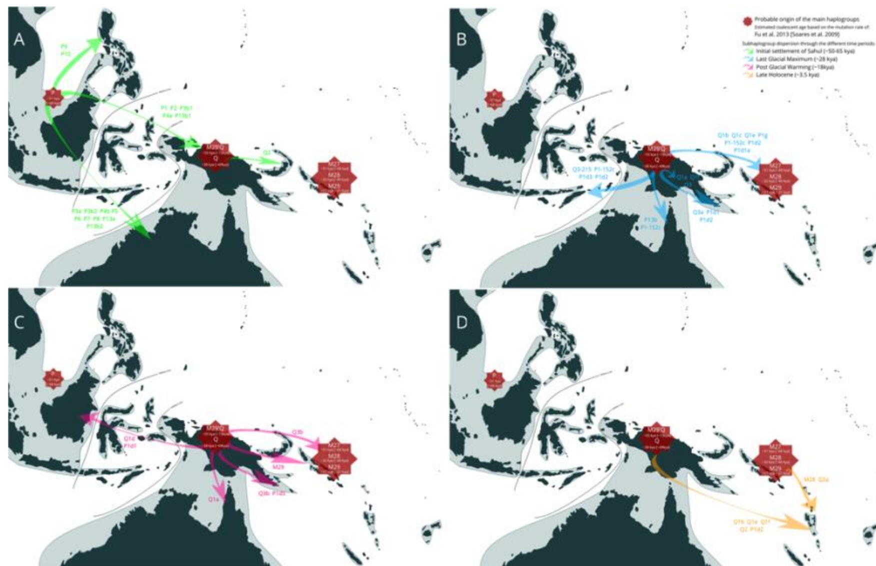


Figure 9: Proposed movements of maternal lineages P, M, and Q. Dark shading represents modern coastlines; light shading illustrates the extent of the Sunda continent at the Last Glacial Maximum. Red octagrams represent the probable approximate origins of haplogroups. Arrows represent probable migration paths during: a) the initial settlement of Sahul (~50 kya; green); b) the Last Glacial Maximum (~28 kya; blue); c) the postglacial warming period through to the Holocene (~18 kya; pink); and d) the Late Holocene (~3.5 kya; orange). Image from (Pedro et al., 2020)

The description of the Australian autochthonous populations recorded by the first European colonizers who arrived in Australia in the late 18th century drew a very diverse profile for the hunter-gatherer groups. Anthropological studies on seven different Australian regions where European contact took place described the population's lifestyle as being strongly influenced by the environment and the nature and location of the resources. Furthermore, in terms of linguistic and cultural traits, the different groups showed very specific regional identities as described by Keen (2003). In general, at the time of the early European contact, no clear signal of neolithization was described by the first colonizers, however, some signs of early domestication could be found (Bellwood & Hiscock, 2005). Nowadays it is difficult to appreciate the Aboriginal Australian ethnography pre-European contact since the strong impact of the European invasion on the autochthonous populations. Many Aboriginal groups saw a rapid population decline that, sometimes, reached extinction within only one year (White, 2011). Sadly, the causes of the extinction were not only the

spread of infectious diseases but also some systematic lawful persecution of the Aboriginal identity perpetuated by the early government (Saini, 2020).

2.6. Phenotype and ancestry: convergent evolution or phylogenetic relatedness?

As described in the previous chapter, the population structure of South East Asia, Papua New Guinea, and Australia traces its roots in the very first human migration that moved outside of Africa in the late Pleistocene and it was later reshaped by both the Austronesian migration and, in more recent times, by the European colonization. Scattered all over the area, living in remote islands, in the rainforest, or in isolated mountains, some populations maintained the ancestral hunter-gatherer lifestyle, showing different genetic histories with or without gene flow with the surrounding populations. These groups are the tribal RFHG of the Andaman Islands, Malaysia, Philippines, Papua New Guinea, and Australia, and, despite having different genetic histories, they all are characterized by very specific morphological features which include a particularly dark skin color (D), an extremely short stature (S), that could reach 150 cm in men and 145 cm in women, the development of woolly hair (W), and the presence of marked steatopygia (S) – fat accumulation specifically localized in the hips (Deng et al., 2022; Sen, 2009b; Stock & Migliano, 2009). From now on, in this thesis, the whole set of the above-mentioned characters will be referred to as DSWS phenotype.

The populations carrying the above-mentioned complete set of characters are not only present in SEA. It was, indeed, estimated that more than 900 k people in the central African rainforest carry the DSWS phenotype (Olivero et al., 2016). Perry et al. (2014) analyzed genome-wide data to understand the evolution of the DSWS phenotype in the African RFHG Batwa and Baka, finding specific signatures of polygenic adaptation for the DSWS phenotype in the Batwa not shared with the Baka and thus suggesting a convergent evolution for the DSWS phenotype in the two African RFHGs. As mentioned in chapter

1, the monophyletic hypothesis for a common origin of all the groups carrying the DSWS phenotype[3] living both in Africa and in South East Asia has been progressively replaced by the convergent evolution hypothesis (Endicott, 2013; Migliano et al., 2013). When analyzing, indeed, genome-wide data from Philippine and Papua New Guinean DSWS groups together with the populations from the HGDP-CEPH panel, Migliano et al. (2013) evidenced that African, Philippine and Papua New Guineans DSWS populations are genetically closer to their neighboring populations than to one another. All these findings support the aforementioned hypotheses of a single wave migration into Sunda and Sahul that firstly colonized the continent, and that subsequently some groups became isolated and adapted to the post-Holocene tropical environment developing the DSWS phenotype. But what are the implications of carrying such a specific set of characters?

2.7. Defining ourselves by creating the "otherness"

The distinctive appearance of the people carrying the DSWS phenotype has drawn the attention of the western world since the moment they were reached by the Colonial pioneers in the nineteenth century. This was the time of the greatest expeditions to the remote areas of the globe, from which the explorers took and brought back to Europe all kinds of food, objects, animals, and even people. Furthermore, the increasing interest in the hunter-gatherer lifestyle of the "new land's" Indigenous people magnified the sense of "otherness" towards them (Endicott, 2013). Indeed, the autochthonous inhabitants of the colonized areas were described to be least civilized, close to the "state of nature" and doomed to imminent extinction, hence, the tales about the wild nature of the "newly discovered lands" and their savage inhabitants shaped in the public opinion "The White Man's Burden"[4] justifying the colonial expeditions (Minter, 2010; Opubor, 2020). The rhetoric of the wild nature's seducing mystical forces, the paradisiac

[3] Hereinafter "DSWS peoples/groups/populations" for shortness.
[4] The "moral duty" felt by western colonizers to spread their "civilization" and to save other human groups from their "state of nature".

landscapes, and the Indigenous peoples' "state of nature" contributed to building in the western society a distorted image of the eastern colonies, that was later described as *Orientalism* by Edward Said (Donzé-magnier, 2017). Said refers to Orientalism as a "European invention […] which has helped to define Europe". It was political propaganda used to remark European superiority and rationalize colonial looting (Opubor, 2020). Many scholars, sailors, and minor aristocrats that self-fashioned themselves as anthropologists started to recollect anthropological and anthropometric data to define and categorize DSWS people under the name of "Negrito" (Manickham, 2009; Sen, 2009b). The term was created in the seventeenth century by the Spanish missionaries that arrived in the Philippines and encountered the autochthonous populations describing them as short and dark people. Subsequently, the term was used to visually categorize all the people that presented a specific aspect, regardless of their culture or geographic distribution (Manickham, 2009). At the same time, the deterministic association between the DSWS phenotype and their intelligence established a slow process of dehumanization that culminated in their infantilization, objectification, and hypersexualization legitimizing the explorers to treat them without any empathy or respect (Sen, 2009b). It is necessary here to acknowledge the story of Sarah Baartman, a Gonaqua Khoekhoe woman from today's Eastern Cape region in South Africa which in the nineteenth century attracted European curiosity towards her pronounced steatopygia typical of the Khoekhoe population (Scully & Crais, 2008).

Sarah Baartman, a life in a cage

The Gonaqua Khoekhoe population lived in south-east Africa as free pastoralists until the late 1700s, when the region became a Dutch colony. The real name of Sara Baartmann is unknown, we only know the name that was given to her by her slave master. She grew up in an ambiguous balance between slavery and freedom under the first Dutch and later British rule. As a Khoekhoe person, she could not be enslaved because the Dutch law forbade the slavery of the Indigenous people,

but, as a woman, she could not live as a "free black" but would always need a master protector. Like every other colonized woman, she could only engage in domestic work in the house of the master, being vulnerable to rape and sexual exploitation. However, the intimacy with the master's family was often used by enslaved women as an opportunity for manumission. She had indeed relatively more freedom than the rest of the slaves in her house (Scully & Crais, 2008). In fact, an entire court case was built afterward around the question of her free will and her freedom in choosing to leave her country and move to England. Willingly or not, since she was transferred to Cape Town, she experienced the European obsession towards the Khoekhoe women's body and started to be exhibited in the naval hospital to the military men, and later, in 1810, she was transferred to England where her master exhibited her to the public as a freakshow. As Scully & Crais (2008) report, the exhibition was not about Baartman herself, but rather selling the idea of the "Hottentot [5] Venus": an exotic being that incarnated what the audience expected from her. "Dressed in clothes supposedly appropriate to her race, forced to chant and dance again in ethnographic style, Sara Baartman came to enact a masquerade of culture and race much more acutely attuned to an English fantasy of her life than it was to the reality of her time in London or Cape Town" Scully & Crais (2008).

In time, the freakshows turned into the ethnographic display of an exotic woman that represented the European idea of the Indigenous people in their colonies (Scully & Crais, 2008). Furthermore, the erotic interest in Sara Baartman's steatopygia was accompanied by the belief that she belonged to a population considered the missing link with the apes; a "living fossil" that could be placed at the bases of all the other human groups (Sorgoni, 2003). To validate these theories, Georges Cuvier attempted to study her body while she was still alive, but even if she refused, he could do it anyway after her death by performing the

[5] Pejorative word invented by the Dutch to describe the pastoralists who spoke the difficult click languages. The world comes from huttentut: to stammer (Scully & Crais, 2008)

autopsy. Cuvier disposed of Baartman's body desiccating it and extracting her brain and genitalia that were exposed in the *Musée de l'Homme* until 2002 when they were returned to South Africa (Sorgoni, 2003).

Sara Baartman's life embodied what the colonized people suffered in their history as a consequence of the socio-political and economical equilibrium ruled by the wretched law of "the survival of the strongest". Their discrimination, historical marginalization, and exclusion from society were justified by the deterministic interpretation of their phenotype and lifestyle (Sorgoni, 2003). As previously mentioned, a similar fate occurred to the people carrying the DSWS phenotype in South East Asia, where the Eurocentric idea of "development" was unsuccessfully forced upon the hunter-gatherer populations, resulting in the belief that their "state of nature" limited their capability to grasp the complexity of an "evolved" culture, placing them in an "in-between" state: not quite animals, but not humans ether (C. Anderson, 2011; Saini, 2020; Sen, 2009b; Sorgoni, 2003).

In this thesis, I will talk about natural selection and human adaptation in nowadays people presenting the DSWS characteristics from South East Asia always bearing in mind that: just as they inherited from their ancestors the socio-political and economical exclusion as well as cultural discrimination that nowadays has translated into the loss of their own cultural identity (Venkateswar, 2004), in the same way, we also have inherited the responsibility of what was perpetrated by the oppressors, that alas, were our ancestors.

3. Calcium sensing receptor

3.1. Molecular structure

The calcium-sensing receptor (CaSR) is a G protein-coupled receptor (GPCR) involved in calcium homeostasis. In humans, the gene encoding CaSR (*CASR*) is localized in chromosome 3 and it is composed of 7 exons encoding for 1,085 amino acids (Magno et al., 2011). In the large N-terminal extracellular domain (ECD) are placed multiple calcium-binding sites (Geng et al., 2016), whereas seven alpha-helix transmembrane domains anchor the receptor to the plasma membrane (Hu & Spiegel, 2007). Finally, the intracellular C-terminal domain has not been yet described in a crystallographic structure, but it is known to be responsible for the G protein activation (Brown & Mc Leod, 2001; Conigrave & Ward, 2013). Like the rest of the GPCRs, the CaSR works in dimers that need to interact with each other to bind cations and transmit the signal into the cell (figure 10).

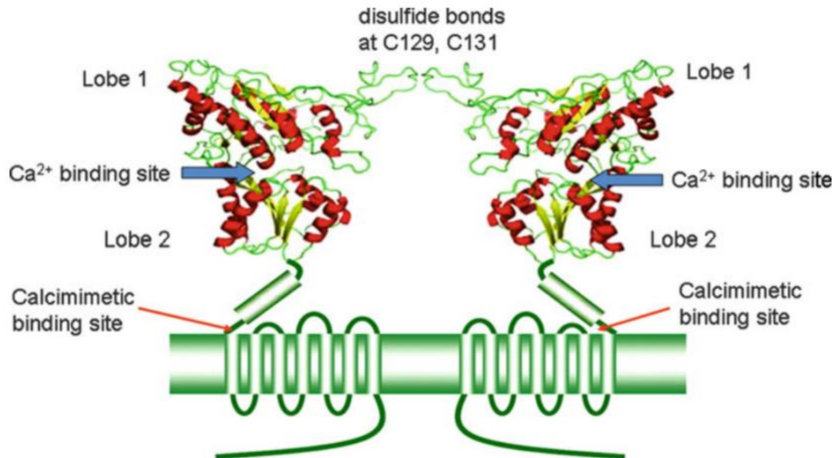


Figure 10: Schematic representation of the ECD of the CaSR based on the known structure of the ECDs of several mGluRs from Chakravarti et al. (2012)

3.2. The physiological function of CaSR

CaSR is mainly involved in calcium homeostasis through the direct regulation of the parathyroid hormone (PTH) secreted by the parathyroid glands (Magno et al., 2011). When calcium concentration in blood is low, PTH acts in the nephron's distal tubule to stimulate Ca^{2+} reabsorption, meanwhile, in the proximal tubule, it enhances the production of vitamin D's reactive form (1,25(OH)2D3). The synthesis of 1,25(OH)2D3 increases calcium reabsorption in the intestine, while in bone PTH induces osteoclasts activation and bone reabsorption and stimulates Ca^{2+} release into the system. The concerted interaction of PTH effects in both bones and kidneys, and the vitamin D absorption and activation in the kidney and the intestine generate the increase in serum calcium concentration. The serum Ca^{2+}increase is then sensed by CaSR in parathyroids, causing the inhibition of PTH secretion in a negative feedback loop as shown in figure 11 (Brown, 2013; Santa Maria et al., 2016).

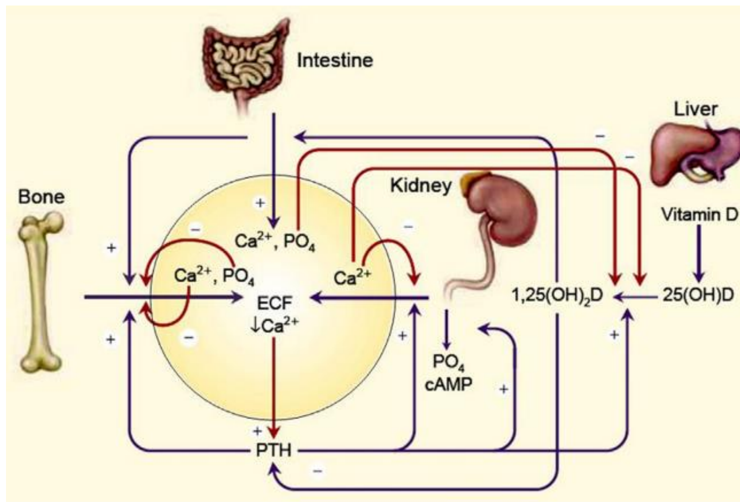


Figure 11: Schematic illustration of the system that maintains Ca^{2+} o homeostasis from Brown (2013).

Different polymorphisms of the *CASR* gene have been described as generating both lost and gain of function mutations (Brown & mac Leod, 2001). CaSR inactivating mutations lead to the development of primary hyperparathyroidism, familial hypocalciuric hypercalcemia (FHH), and, in the most extreme cases, neonatal severe hyperparathyroidism (NSHPT) (Frank-Raue et al., 2011; Fukumoto et al., 2001; Hendy et al., 2000). On the other hand, some activating mutations produce a variety of pathological conditions such as hypocalcemia, hypercalciuria, or hypoparathyroidism, while others are associated with kidney stone risk in human populations (Vezzoli et al., 2017). As shown in figure 12, numerous mutations were discovered all along the *CASR* gene being both inactivating and activating, whereas several polymorphisms were described to be involved in different human disorders regarding cations homeostasis in serum and urine (figure 13).

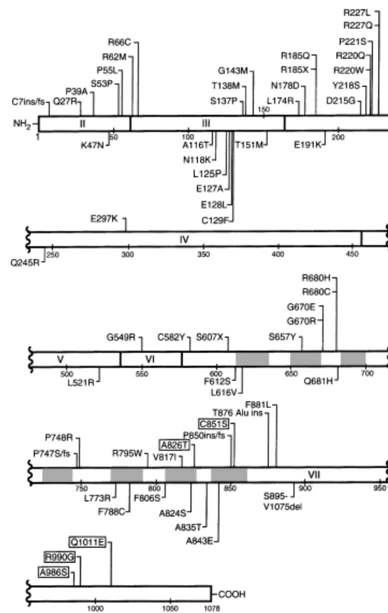


Figure 12: Locations of inactivating and activating mutations in the *CASR* gene. The locations of the inactivating mutations found in patients with FHH and /or NSHPT (as well as polymorphisms A826T, C8515, A986S, R990G, and Q1011E which are boxed) are shown above, and of activating mutations found in patients with ADH are shown below the bars representing the protein-coding exons from Hendy et al. (2000)

Single-Nucleotide Polymorphism	Position	Domain	Alleles	Minor Allele Frequency	Amino Acid Change	Disorder Associated With the Minor Allele
rs1801725	123,486,447	Exon 7	G>T	20%	986 A/S	Higher serum calcium
						Lower serum phosphate
rs1042636	122,284,922	Exon 7	A>G	5%	990R/G	Hypercalciuria, Higher calcium excretion, Calcium kidney stones
				50%[a]		
rs1801726	122,284,985	Exon 7	C>G	2%	1011 q/E	Higher serum calcium
rs17251221	123,475,937	Intron 4	A>G	19%	None	Higher serum calcium
						Higher serum magnesium
rs6776158	122,183,002	Promoter 1	A>G	26%	None	Calcium kidney stones
						Lower serum phosphate
rs1501899	123,390,018	Intron 1	G>a	26%	None	Calcium kidney stones
rs7652589	123,371,778	5′UTR	G>a	27%	None	Calcium kidney stones

[a]*Frequency in Asian people.*

Figure 13: Calcium-Sensing Receptor Gene Polymorphisms Involved in Human Disorders from Vezzoli et al. (2017)

3.3. CASR is a pleiotropic gene

CaSR expression was detected in a plethora of different tissues besides the parathyroid glands, such as bone, cartilage, brain, kidney, breast, skin, smooth cardiac muscle, pancreas, and intestine. Therefore, being expressed in such numerous tissues, the gene is involved in many different pathways besides its role in Ca^{2+} regulation in plasma (Diez-Fraile et al., 2013). In this thesis, we will describe CaSR involvement in the differentiation of the growth plate (GP) and bone remodeling (Santa Maria et al., 2016), in adipocytes differentiation (Y. H. He et al., 2012) as well as in epidermal differentiation (Tu et al., 2001, 2012).

CaSR in growth plate and bone

Several studies on the growth plate and bone formation highlighted the clear involvement of CaSR in chondrocytes maturation, and in balancing osteoblast maturation and osteoclast inhibition through the RANKL gene (Chang et al., 2010; Dvorak et al., 2007). Santa Maria et al. (2016) also described CaSR involvement in chondrocyte differentiation through the regulation of the Parathyroid Hormone 1 receptor (PTH1R) and the Parathyroid Hormone-related peptide (PTHrP). In the proliferating zone of the GP, chondrocytes are maintained in a proliferative state by PTHrP/PTH1r, until they spread away from the proliferative zone, where the PTHrP/PTH1r's concentration is low. Here, *CASR* expression starts to increase, enhancing chondrocyte differentiation and the closure of the GP (Matta & Mobasheri, 2014; Santa Maria et al., 2016). Both activating and inactivating mutations of CaSR in bone revealed that the gene plays a pivotal role in bone synthesis and remodeling. Bone tissue-specific deletion of CaSR in a mouse model showed a delay in the expression of osteogenic genes, and therefore, in osteoblasts differentiation (Chang et al., 2013). When inducing a double knock-out both for CaSR and 1alpa(OH)ase (the enzyme responsible for the hydroxylation, hence the activation, of VitD) abnormalities in the growth plate were detected, such as disorganization of the cellular progression and reduced mineral

deposition at the chondro-osseous junction (Richard et al., 2010). On the other hand, an activating mutation in Nuf mice caused ectopic calcifications in the eyes, tongue, blood vessels, and visceral smooth muscle (Hough et al., 2004). In the growth plate, CaSR accumulation was detected in maturating chondrocytes and increased in hypertrophic chondrocytes as well as in osteoblast in the chondro-osseous junction (Santa Maria et al., 2016). In the cancellous bone, transgenic mice with a constitutively active mutant CaSR (Act-CaSR) showed a reduction in bone mineral density and bone fraction, while having an increase in the trabecular spacing and a reduction in trabecular number and connectivity density, suggesting that Act-CaSR in osteoblast support osteoclastogenesis and bone reabsorption (Dvorak et al., 2007).

CaSR in adipose tissue

No evidence of CaSR expression in adipose tissue was present until Cifuentes et al. in 2005 isolated the protein and sequenced the gene in cultured adipocytes from human omental fat obtained from abdominal surgery. Later, studies carried out *in vivo* and *in vitro* highlighted the importance of CaSR for adipocyte proliferation, differentiation, and regulation of lipolysis (Bravo-Sagua et al., 2016). For example, the preadipose cell line LS14 was treated with three different CaSR activators, resulting in the elevation of the preadipocyte proliferation and the elevation of pro-inflammatory factors (figure 14 A) (Rocha et al., 2015). Villarroel et al. 2013 also treated the LS14 cell line with the calcimimetic cinacalcet for 10 days and discovered that CaSR activation elevated the expression of the adipogenic markers aP2, adiponectin, PPARγ, FASN, and GDP. Furthermore, the two CASR polymorphisms rs1801725 and rs7652589 were found to be associated with dyslipidemia in hemodialysis patients (Grzegorzewska et al., 2019). He et al. (2012) treated a confluent monolayer of SW872 preadipocytes with a calcium agonist for CaSR activation (GdCl3) and analyzed the triglyceride content as a measure of the adipocyte differentiation in three different conditions: control, GdCl3 treated cells and pretreating the cell culture with both GdCl3 and CaSR-siRNA. It was found that

the calcium-sensing receptor activation not only influenced the expression of genes related to lipid handling (LPL, aP2, C/EBPa) but also was directly linked with preadipocyte maturation through the activation of PPARγ, which is considered the master-gene for preadipocyte proliferation and differentiation (figure 14 B) (Lowell, 1999; Spiegelman & Flier, 1996). In mature adipocytes, the calcium-sensing receptor was described to play a role both in the lipogenic pathway through the regulation of genes involved in fatty acids synthesis, and in the lipolysis regulating the hormone-sensitive lipases and the adipose tissue triglycerides lipase (Villarroel et al., 2014).

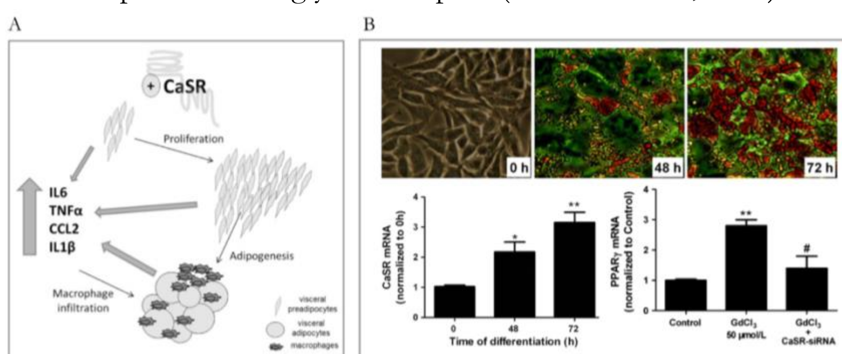


Figure 14: CaSR role in preadipocyte proliferation and differentiation. A. adapted from Rocha et al. (2015); B. on the top panels is shown a monolayer of SW872 preadipocytes treated with GdCl3 in 3 time points (0h, 48h, and 72h), in the bottom panels *CASR* expression increases directly with the GdCl3 treatment as well as PPARγ expression which in turn decreases when treating with CaSR-siRNA, adapted from He et al. (2012)

CaSR in epidermis

In keratinocytes two variants of the *CASR* gene are transcribed: the full-length CASR transcript and a spliced variant lacking the exon 5. The latter variant is fully cytoplasmatic and does not change its expression through keratinocyte differentiation. On the other hand, the full-length CaSR expression decreases in the terminally differentiated keratinocytes (Tu & Bikle, 2013). In human cultured keratinocytes, the inhibition of CaSR expression using an antisense cDNA revealed that the gene regulates the calcium-induced keratinocytes differentiation (Tu et al., 2001). Besides, knocking out the gene in a mouse model caused a

reduced expression of the differentiation genes for the epidermis and a gene expression delay of the permeability barrier establishment (Tu et al., 2012). On the other hand, transgenic mice injected with a construct carrying the basal CaSR under the control of the keratin promoter, showed a CaSR overexpression, a thicker back skin, and an acceleration of epidermal differentiation markers in 6-day-old mice (Turksen & Troy, 2003).

3.4. The CaSR R990G polymorphism: molecular phenotype, phenotypic associations

As previously mentioned, several polymorphisms in the *CASR* gene were described to be associated with different disorders (figure 13) (Vezzoli et al., 2017). Among them, the non-synonymous substitution R990G was described to be associated with Hypercalciuria, higher calcium excretion, and kidney stones. In exon 7 of the *CASR* gene, the nucleotide transition Adenine-to-Guanine generates an aminoacidic change from Arginine to Glycine at position 990 of the CaSR's COOH-terminal. A genome-wide meta-analysis on a cohort of Indian and European people detected the association of non-synonymous substitutions in the *CASR* gene with serum calcium concentration, and specifically, the rs1042636 (R990G) was found to be associated with low serum calcium concentration in linkage with the rs1801725 (Kapur et al., 2010). Guha et al. (2015) also analyzed the genetic variants in the *CASR* gene in a cohort of 200 people from eastern India affected by kidney stone risk and detected that calcium excretion was significantly higher in people carrying the 990G allele. Other studies also reported that European patients affected by primary hyperparathyroidism and carrying the G allele showed an increase in urinary calcium excretion and a low PTH (Corbetta et al., 2006). Consistent with these results, Vezzoli et al. (2007) highlighted that the 990G allele conferred the receptor a higher sensitivity to Ca^{2+} leading to a higher risk for the development of hypercalciuria, however, the subjects did not show any differences in serum PTH levels. Additionally, when HEK-293 cells were transfected with a 990G-CaSR vector they showed an increase in the CaSR response to the extracellular Ca^{2+} concentration (Vezzoli et

al., 2007), while the expression of CaSR's downstream genes was significantly increased in 990G-CaSR injected cells (Ranieri et al., 2013). These results together evidenced that the higher sensitivity to calcium was followed by an enhanced CaSR activity. Besides the role in calcium-sensing and homeostasis, the R990G substitution was also associated with hypertriglyceridemia in the Chinese population, where the G allele was significantly higher in hypertriglycemic subjects when compared with a non-hypertriglycemic group (Y. han He et al., 2014). Furthermore, when isolated adipocytes from human omental fat of patients without kidney disease were treated with cinacalcet, a CaSR activator, it was shown that cells carrying the G allele both in homozygosis and in heterozygosis showed an enhanced antilipolytic response when compared with cells carrying the A allele (Reyes et al., 2012).

3.5. R990G frequency distribution in Asia and signals of selection

To analyze the worldwide frequency distribution of a polymorphic variant of interest, it is possible to rely on samples from the 1000 genome project (Auton et al., 2015). The project constitutes the most comprehensive catalogue of human genetic variation based on whole genome sequencing data of more than 2,000 individuals from America, Africa, Europe, South Asia, and East Asia. The study includes common genetic variants with at least 1% frequency in at least 1,000 samples. When searching the rs1042636 in the catalogue, the populations from East Asia seem to carry the highest frequency of the derived allele (figure 15). Furthermore, Mondal et al. (2016) analyzed whole genome sequences of 60 individuals from mainland India together with sequences of 10 individuals from two populations from the Andaman Islands. The authors described a common origin of all the Asian and Pacific populations while also showing the presence of a small proportion of ancestry derived from an unknown ancient hominin in SEA populations. On top of that, signals of positive selection towards the Andamanese population were detected in several genomic regions with a significant excess of candidate genes for selection related to body

morphology and height. Notably, one of the candidate regions for positive selection included the *CASR* gene, where the non-synonymous substitution R990G was found among the top 20 most differentiated variants in high frequency among the Andamanese when compared with mainland Indians (personal communication from Mondal et al. (2016).

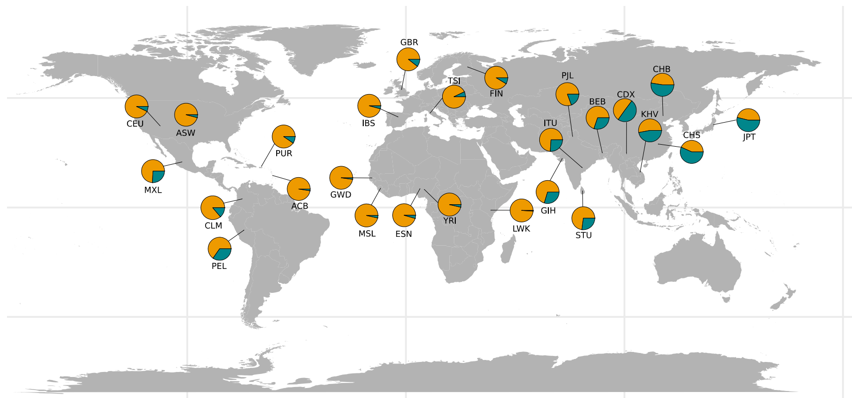


Figure 15: Allele frequency of the rs1042636 in 1000 genomes populations. In yellow is shown the ancestral allele frequency (A allele), while in green is the derived allele frequency (G allele).

4. Detecting signals of selection and creating a mouse model

4.1. Evolutionary forces and their impact on the genomes

The evolution of all living species is the result of a complex dynamic process driven by four main evolutionary forces that produce allele frequency changes through time and space: mutation, genetic drift, natural selection, and gene flow (Relethford & Harding, 2001). On a smaller scale, the evolutionary forces act primarily on populations, that define the smallest biological units that can evolve over time. Mutation introduces new alleles in populations, which are generated by random errors during DNA replication. Genetic drift is a stochastic mechanism of allele frequency fluctuation generated by the random sampling of alleles from one generation to the next, which can lead to the random fixation or extinction of existing polymorphic alleles (figure 18 a). Natural adaptive selection causes an increase in the allele frequency of specific mutations through the generations when they have a positive functional impact on survival and fitness (figure 18 b-d). Finally, gene flow introduces new variability into populations because of migration and reproduction with individuals from external groups with different genetic compositions. However, from an adaptive perspective, natural selection is the only evolutionary force that allows populations to better fit or adapt to their environments and changing conditions acting on the distribution of continuous phenotypical traits according to different modes (Gregory, 2008).

Modes of natural selection

Depending on the action of natural selection on the alleles upon selection, we can identify three main types of natural selection: positive, negative, and balancing selection. An allele under positive selection increases its frequency in the population when the resulting phenotype is advantageous for fitness (figure 18 b). In this scenario, all surrounding

mutations are dragged along with the positively selected variant in a mechanism known as genetic hitchhiking, causing a clear reduction in the haplotype[6] diversity (figure 19 b). Negative selection eliminates those alleles negatively affecting reproductive success, which can also lead to a reduction in genetic variability in the surrounding genomic region if the deleterious alleles are eliminated along with neutral alleles in linkage disequilibrium (LD)[7], a process known as background selection (figure 18 d). Finally, balancing selection favors multiple alleles at a given locus, driving them neither to fixation nor to extinction (figure 18 c) usually favoring the extreme phenotypes in a continuous trait phenotypic distribution and disadvantaging the intermediate phenotypes (Bamshad & Wooding, 2003; Rees et al., 2020).

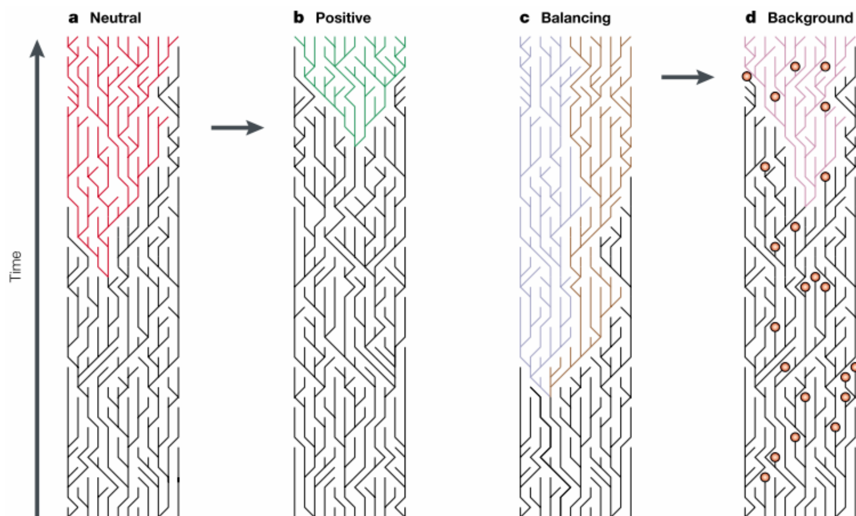


Figure 18: Effects of natural selection on gene genealogies and allele frequencies. Each panel (a–d) represents the complete genealogy for a population of 12 haploid individuals. Each line traces the ancestry of a lineage, and colored lines trace all descendants who have inherited an allele that is either neutral (a) or affected by natural selection (b–d) back to their common ancestor (that is, the coalescence of the genealogy). a | The genealogy of a neutral allele (red) as it drifts to fixation. b | The genealogy of an allele (green) that is driven to fixation more quickly after the onset of positive selection (arrow) compared with expectations under a neutral model. Note that the genealogy has a more recent coalescence. c | The genealogy of two alleles (blue and gold) under balancing selection, which are maintained in the

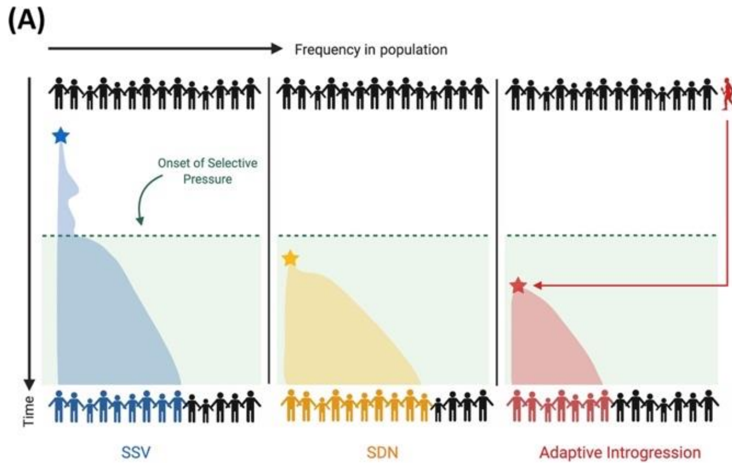
[6] Haplotype: combination of alleles in different loci inherited together.

[7] Linkage disequilibrium: non-random association of alleles at different loci

population. As a result, the genealogy of the two alleles has an older coalescence. d | The genealogy of an allele (purple) that drifts to fixation under the influence of background selection. Each circle represents the elimination of a deleterious (that is, lethal) mutation by background selection. The coalescence of the lineage is more recent than expected under a neutral model because a linked deleterious mutation caused the extinction of one lineage (arrow) more quickly than would be predicted (Bamshad & Wooding, 2003)

4.2. Molecular fingerprints of positive selection

Variants under positive selection can be *de novo* mutations (model of selection on a de novo mutation or SDN) that rapidly elevate their frequency generating a hard selective sweep. This rapid increase of the favored variant and surrounding linked variation causes an important initial decrease in the genetic diversity of the genomic region under selection, which will be followed by the progressive accumulation of new variants by mutation that will deviate the site frequency spectrum towards an excess of rare variants (Nielsen et al., 2007). Nevertheless, in humans, the most common scenario is that of the soft selective sweep generated by a new selective pressure acting on standing variation (model of selection on standing variation or SSV): mutations that before the selective sweep were neutrally fluctuating in frequency through generations in the population and that result adaptive after an environmental change. In this case, we expect different haplotypes carrying the positively selected allele, a narrower haplotype core region, and a lower occurrence of low-frequency variants (Hermisson & Pennings, 2017; Weigand & Leese, 2018). Another scenario of positive selection can take place when a beneficial variant is introduced through admixture with an external population, known as adaptive admixture or post-admixture selection when the allele was already under selection in the source population or not, respectively (Cuadros-Espinoza et al., 2022). Moreover, if the beneficial variant is introduced through introgression from archaic humans, the process is recognized as adaptive introgression (figure 19) (Rees et al., 2020).



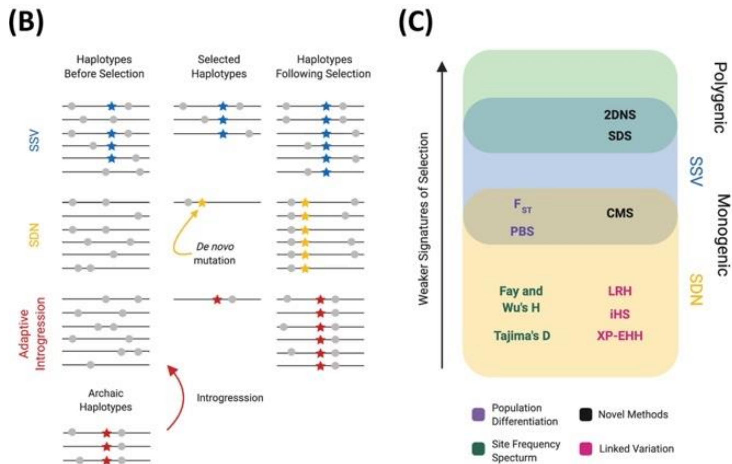


Figure 19: Characterization of Selection on Standing Variation, Selection, on de Novo Mutation, and Adaptive Introgression. (A) Cartoon showing the rise in frequency of an allele through a population according to its origin. The stars represent mutations (blue: mutation present in the population prior to selection; yellow: mutation occurring after the onset of selection; and red: mutation present in an archaic population which spreads through a receiving population following an admixture event). The frequency of each mutation in a population following a selection event is represented by the number of people icons of their respective colors under each panel. The red walking person icon in the top right represents an archaic human. (B) Cartoon depicting the haplotypes arising from selection on standing variation (SSV), selection on de novo mutation (SDN), and adaptive introgression. Stars represent the beneficial allele and grey circles represent neutral linked polymorphic alleles. (C) Some of the most common statistics used to identify local adaptation (not a comprehensive list). Abbreviations: iHS, integrated haplotype score; LRH, long-range haplotype; PBS, population branch statistic; SDS, Singleton density score; XP-EHH, cross-population extended haplotype homozygosity (Rees et al., 2020).

In any case, genetic adaptation leaves characteristic molecular footprints in the patterns of variation of genomic regions under positive selection that can be retrieved by different neutrality tests. Since both soft sweeps and hard sweeps generate a typical pattern of low diversity with an excess of low-frequency variants, different statistical tests can be used to detect the genomic regions whose genetic patterns of variation deviate from neutrality. For example, some tests aim to identify and quantify skews in the site frequency spectrum, including SFselect (Ronen et al., 2013a). Other statistical tests analyzing the allele frequency spectrum such as Tajima's D or Fay and Wu's H are often used to look for regions showing an excess of rare variants or of high frequency derived alleles, respectively (Fay & Wu, 2000; Sabeti et al., 2006a; Tajima, 1989). Tests that analyze allele frequency differences between populations like F_{ST} allow the detection of genomic adaptations occurring in a population under a specific selective pressure, using another genetically close population as a control to retrieve old selection events occurring at least since their corresponding divergence split (Akey, 2009; Pybus et al., 2014; Ronen et al., 2013b). Finally, a selective pressure acting on an advantageous variant generates an increase in its frequency sweeping along with it the surrounding variants in LD. If the selective event is recent, recombination will not have had time to break the physical linkage generating an unusually long region with variants inherited together (Biswas & Akey, 2006; Sabeti et al., 2006a; Weigand & Leese, 2018). For these reasons, methods based on linkage disequilibrium decay, such as the integrated Haplotype Score (iHS) and the cross-population Extended Haplotype Homozygosity (XP-EHH), are more useful to detect recent genetic adaptations. Finally, it is important to use different neutrality tests because each one of them focuses on different aspects of the selective sweep, allowing to identify, for example, both old selection events and more recent ones (Pybus et al., 2014).

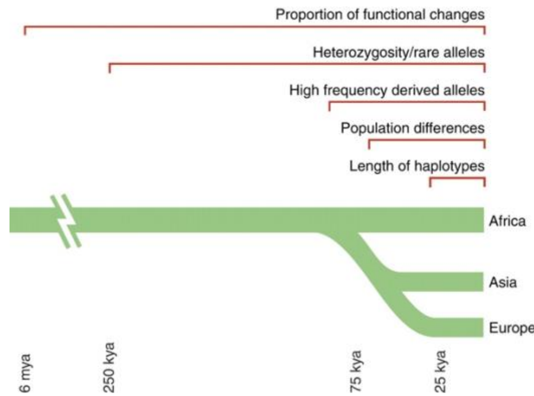


Figure 20: Time scales for the signatures of selection. The five signatures of selection persist over varying time scales. A rough estimate is shown of how long each is useful for detecting selection in humans. Image from (Sabeti et al., 2006a)

4.3. Finding adaptation in human populations

The first studies that aimed to discover genetic adaptations in humans applied a candidate gene approach, focusing on genes with a well-described function and having a precise evolutionary hypothesis to test. In other words, these studies relied on *a priori* assumptions regarding which phenotypes could be adaptive and which genes may be under selection. The candidate gene approach allowed to identify genes under selection for adaptive phenotypes related to the metabolism of specific dietary compounds after a dietary shift, the variation of skin pigmentation in response to solar radiation, and immunological adaptation in response to specific pathogens. However, strong limitations for these types of studies included on one hand having to hypothesize an *a priori* assumption that limited the analysis only to the predefined candidate genes, while, on the other hand, not taking into account the confounding effect of the demographic processes (Akey, 2009; Sabeti et al., 2006b; Szpak et al., 2019).

In the last decades, the increased availability of genomic data from different populations allowed researchers to perform genome-wide scans of selection that exponentially improved the sensitivity toward retrieving new genomic regions showing the typical footprints of positive selection. Nevertheless, the regions identified tend to be quite

large and include many different genes involved in disparate pathways (Akey, 2009; Grossman et al., 2010). For this reason, after retrieving the genomic regions displaying signals of positive selection, researchers often look for polymorphisms presenting extreme allele frequency differences between populations with and without the detected signatures of selection. Often, when a non-synonymous variant is identified inside a gene (or a regulatory variant close to a gene), researchers move to analyze their functional implications by using some prediction scores for functional impact (such as the CADD score) and performing genomic annotations of the selected variants in large-scale genotype/phenotype databases (such as the GWAS catalog (Buniello et al., 2019)) to infer each corresponding adaptive phenotype (Rentzsch et al., 2018; Roca-Umbert et al., 2022; Szpak et al., 2019; Walsh et al., 2020). Despite all the *in silico* predictions, however, to really understand the phenotypic outcome of a candidate allele, it must be tested either *in vitro* or *in vivo* to also include the background effect of the physiological environment (Vitti et al., 2013).

4.4. Gene editing history and applications

Gene editing is the most powerful tool of molecular biology to understand the function of the genes we are interested in. The classical approach for studying gene function is called reverse genetics and consists of modifying the DNA sequence and seeing the phenotypic outcome (Singh et al., 2016). The first evidence of *de novo* mutations induced by the treatment with radiations was observed by Muller and Auerbach in the early 20th century, but only in the late decades, the gene-editing had a real outbreak. The first experiments that attempted to modify the double helix used transposons to insert specific mutations, however, these methods induced genomic changes at random (Randhawa & Sengar, 2021). Afterward, homologous recombination (HR) was used in mice and yeast (Carroll, 2017), but the efficiency of the protocol was very low due to the low rate of recombination with each donor template and the high work investment in the selection of the corresponding suitable embryos (Randhawa & Sengar, 2021). In the

following years, the editing efficiency was improved by using endonucleases targeting specific loci with the introduction of a double-strand break (DSB). The damage induced the DNA repairing machinery of the cell to perform either the Non-Homologous End Joining (NHEJ) causing the introduction of an indel and therefore the loss of the gene function, or the Homologous Direct Repair (HR) with a donor vector introducing a precise insertion. The induced DSB followed by either the NHEJ or the HR is the core of the gene-editing technology subsequently developed: Zinc Finger Nucleases, TALENS, and CRISPR-Cas9 as shown in figure 16 (Randhawa & Sengar, 2021).

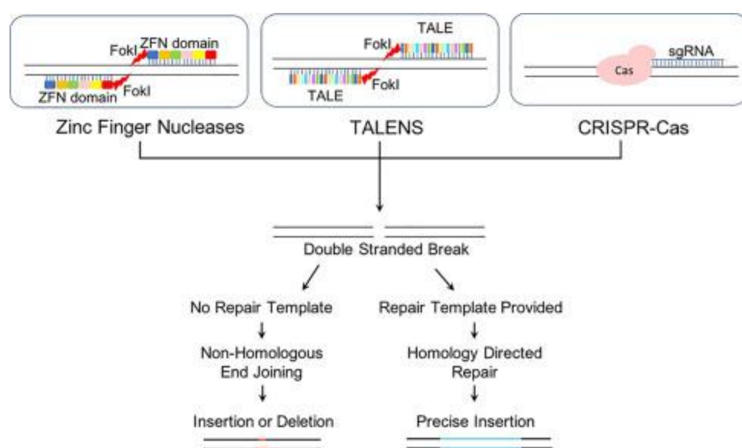


Figure 16: Genome editing using ZFN, TALEN, and CRISPR-Cas. Image from (Randhawa & Sengar, 2021).

4.5. CRISPR-Cas9

Since the late 1980s, the discovery of some short palindromic sequences in the genome of *E. coli* attracted scientific curiosity, and during the following years, different hypotheses were made about their possible involvement in RNA stability or replicon partitioning. It was only twenty years later that the first evidence of the spacer[8]'s similarity with the phage P1 was gathered and that the CRISPRs (clustered regularly

[8] Fragments of the viral DNA integrated into the bacterial genome and separated by short palindromic repeats that form harpin loops when expressed

interspaced short palindromic repeats) were proposed as part of the bacterial adaptive immune system. Finally, in 2011, the experimental evidence of the CRISPR functioning with the involvement of the Caspase 9 (Cas9) protein was obtained using bacterial clones selected for phage resistance that had acquired the CRISPR sequences into their genome only when Cas9 was present (Randhawa & Sengar, 2021). Since then, more data were obtained studying different Prokaryotes, detecting that the CRISPR elements were hypervariable across species, but all of them had some structural similarities and used different Caspases (Barrangou & Marraffini, 2014).

Molecular mechanism of CRISPR-Cas9

In Prokaryotes, the CRISPR-Cas complex is a ribonucleoprotein (RNP) composed of two RNAs (crRNA and tracrRNA) and an endonuclease that together target and digest the specific viral genome. The crRNA is a 20 nt RNA complementary to the target region[9], while the tracrRNA, in part complementary to the crRNA, folds into a hairpin-loop 3D structure that functions as a scaffold for the nuclease's anchoring. The recognition of the target region relies on one hand on the 20 bp complementarity of the crRNA, and on the other on the Caspase's PAM-interacting domain in the COOH-terminal. The PAM domain interacts with 3 nucleotides specific in each Caspase, and when correctly recognized confers more stability to the RNP inducing a conformational change that allows the opening of the double helix in the target region and the subsequent formation of the crRNA:DNA heteroduplex. The correct pairing of the 20 crRNA nt activates the two catalytic sites of the Caspase resulting in a double-strand break (figure 17) (Singh et al., 2016).

Since its discovery, researchers have been optimizing the CRISPR-Cas machinery to perform highly specific gene editing. Different studies reported that the use of a synthetic single-guide RNA (sgRNA) obtained by the fusion of the crRNA and the tracrRNA increased the

[9] The crRNA complementarity to the viral genome relies on the presence of the spacer sequences integrated in the bacterial DNA in a previous infection

editing efficiency, as well as the use of a shorter sgRNA with 17 bp of complementarity to the on-target. Furthermore, in dependence on the target region's genomic context, it was shown that different Caspases, being specific for different PAM sequences, could be used for increasing the specificity to the target sequence; the commonly used caspase is the *Streptococcus pyogenes* Cas9 with the PAM sequence NGG (Singh et al., 2016; Zhang et al., 2015). However, since the technology can tolerate some mismatches between the sgRNA and its target, the probability to detect some off-target activity is not so low, hence, for increasing the specificity and reducing the off-target effects, different strategies have been described (Ding et al., 2019; Hay et al., 2017; Zhang et al., 2015).

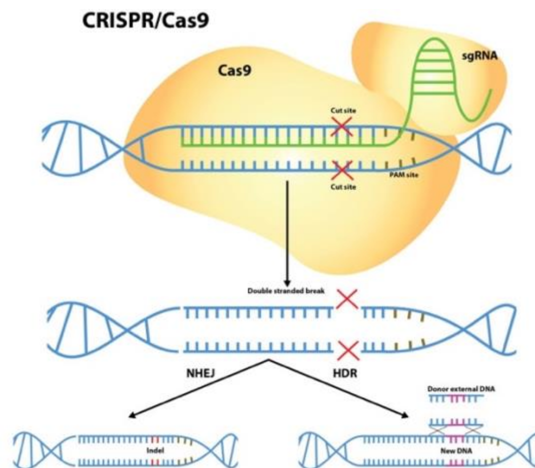


Figure 17: CRISPR-Cas9 mediated gene editing. Image from Cribbs & Perera (2017)

4.6. CRISPR-Cas9 editing in mouse models: increasing the specificity.

Traditional genetic engineering to generate mouse models involved homologous recombination to induce genetic modification to a target locus in embryonic stem cells (ES). Afterward, the edited ES were injected into a blastocyst and implanted into a foster mother. The newborns received some of the edited stem cells that could be passed on to the following generation through controlled breeding (Capecchi,

1989). This method was quite laborious on one hand for the slow efficiency of the homologous recombination using an external vector, and on the other for the ES injection that caused the loss of the majority of the embryos (Carroll, 2017; Chen et al., 2016). The implementation of the CRISPR-Cas9 technology in the process of generation of mouse models improved both the efficiency of the editing and the rapidity in the establishment of any mutant mouse line. Many different strategies were used to improve the efficiency of the sgRNAs and Cas9 delivery into the embryos, from the injection of viral vectors carrying the sgRNA and the Cas9 genes into harvested embryos to the microinjection in utero of the reagents for the CRISPR reaction followed by *in-situ* electroporation (Ehrke-Schulz et al., 2017; Gurumurthy et al., 2019; Nakagawa et al., 2015). However, the authors of two independent studies (Chen et al., 2016; Wang et al., 2016) increased the editing efficiency through the electroporation of the CRISPR-Cas9 ribonucleoprotein already assembled with the gRNA into mouse one-cell zygotes. It was shown that the efficiency for the RNP delivery into the zygotes reached 100%, the pick for the endonuclease's activity and the HDR-mediated knock-in occurred in the first genomic replication and their transient presence in the zygotes rapidly decreased, reducing the off-target mosaicism. Furthermore, it was shown that using an asymmetric vector specifically designed for pairing soon after the target strand cleavage increased both the frequency and the efficiency of the HDR (Hashimoto et al., 2016; Modzelewski et al., 2018; Richardson et al., 2016). In this study, to limit the off-target effect, we applied both the electroporation of RNP into one-cell zygotes and used an asymmetric vector for the knock-in experiment.

OBJECTIVES

The DSWS phenotype is probably the result of positive (adaptive) natural selection acting on human populations that live in tropical rainforests. However, little is known about the adaptive advantage conferred by the DSWS set of characters, and the selective pressures behind it. To understand the evolution of the phenotype, many studies scanned the genomes of DSWS populations in search of signatures of positive selection to subsequently identify the putative functional variants responsible for such phenotype. In one of the studies, whole genome sequencing data in the Andamanese were analyzed and classical signals of positive selection were found in a genomic region encompassing the SNP rs1042636, which causes the non-synonymous substitution R990G in the Calcium Sensing Receptor (CaSR). In this work, I first aim to characterize any putative adaptive phenotypic outcome of the R990G substitution in CaSR, and then I try to identify the selective pressure that drove the underlying adaptive event and detected selection signature.

To fulfill this purpose, we planned the following objectives:

i. Analyze the worldwide frequency distribution of rs1042636
ii. Analyze whether the signals of positive selection in the *CASR* gene originally described in the Andamanese replicate in populations from Malaysia, the Philippines, and Papua New Guinea, which also present the DSWS phenotype.
iii. Study the molecular function of the CaSR protein and understand in which pathways it plays an important physiological role.
iv. Generate a KI-mouse model carrying the R990G substitution to obtain a global view of all the possible tissues where the expression of 990G-CaSR could induce any phenotypic adaptive change.
v. Infer the putative adaptive phenotype of the 990G substitution at the organism level and investigate any relation with the development of the DSWS traits and the corresponding selective pressure behind them.

RESULTS

Exploring adaptive phenotypes for the calcium-sensing receptor polymorphism R990G

Barbara Sinigaglia[1], Jorge Escudero[1], Josep Moreno[3], Begoña Dobon[1], Sandra Acosta[1], Mayuk Mondal[2], Sandra Walsh[1], Gabriela Aguileta[1], Mònica Vallès[1], Stephen Forrow[4], Juan Martin-Caballero[3], Andrea Migliano[5], Jaume Bertranpetit[1], Francisco J Muñoz[6], Elena Bosch[1,7 *]

[1]Institut de Biologia Evolutiva (UPF-CSIC), Departament de Medicina i Ciències de la Vida, Universitat Pompeu Fabra, Parc de Recerca Biomèdica de Barcelona, Doctor Aiguader 88, 08003 Barcelona, Spain

[2] Institute of Genomics, University of Tartu, 51010 Tartu, Estonia

[3]PCB-PRBB Animal Facility Alliance, Parc de Recerca Biomèdica de Barcelona, Doctor Aiguader 88, 08003 Barcelona, Spain

[4]Mouse Mutant Core Facility, Institute for Research in Biomedicine (IRB), Baldiri Reixac 10, 08028 Barcelona, Spain

[5]Department of Anthropology, University of Zürich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

[6]Laboratory of Molecular Physiology, Departament de Medicina i Ciències de la Vida, Universitat Pompeu Fabra, Parc de Recerca Biomèdica de Barcelona, Doctor Aiguader 88, 08003 Barcelona, Spain

[7]Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), 43206 Reus, Spain

*Corresponding author

E-mail: elena.bosch@upf.edu

Manuscript in preparation

Abstract

Rainforest hunter-gatherers from South-East Asia are characterized by very specific morphological features including a particularly dark skin color (D), short stature (S), the development of woolly hair (W), and the presence of marked steatopygia (S) – fat accumulation specifically localized in the hips (DSWS phenotype). Using whole-genome sequencing data, we first characterized signatures of adaptive natural selection around the Calcium-Sensing Receptor (CaSR) gene in SE Asian rainforest human groups presenting the DSWS phenotype and identified the R990G substitution (rs1042636) as a putative adaptive variant to experimentally follow up. Although CaSR has a critical role in calcium homeostasis by directly regulating the parathyroid hormone (PTH) secretion, it is expressed across a wide range of tissues and has been described to be involved in many biological functions. Moreover, previous studies have characterized the R990G substitution as an activating polymorphism of CaSR associated with hypocalcemia. Thus, we generated a knock-in mouse for the R990G substitution and investigated organismal phenotypes that could have become adaptive in the rainforest hunter-gatherers from SE Asia. Interestingly, we found that mouse homozygotes for the derived allele show not only lower serum calcium concentration but greater fat accumulation and body weight, probably because of enhanced preadipocyte differentiation and lipolysis impairment resulting from the CaSR activation mediated by R990G. We hypothesize that such differential features in humans could facilitate an earlier sexual maturation, in agreement with the hypothesis that the short stature of the DSWS phenotype could result from selection for an early onset of reproduction in a particularly hostile environment.

Introduction

The Andamanese are the Indigenous people from the Andaman Islands, situated in the north-eastern Indian Ocean, where they have lived in genetic isolation following a hunter-gather lifestyle till the present times. From a cultural perspective, they are animists and rely on a system of beliefs strictly entangled with storm seasonality (Bxown, 1909; Stock & Migliano, 2009). Since the early colonial period, the Andamanese populations underwent a rapid population decline and in present times some groups are also experiencing a cultural shift toward a more sedentary lifestyle (Venkateswar, 2004). Like other hunter-gathers living in the tropical rainforests of Africa, Asia, and Australia, the Andamanese present characteristic phenotypic features including dark skin (D), very short stature (S), woolly hair (W), and sporadic steatopygia (S) (i.e., extreme development and accumulation of fat on the buttocks). From now on, this complete set of characters will be referred to as DSWS phenotype. Both a common origin and convergent adaptations to the environmental challenges found in their corresponding rainforest habitats and/or resulting from their lifestyle have been postulated to explain the shared DSWS phenotype within and across different continental hunter-gatherer human groups (Perry & Dominy, 2009; Verdu et al., 2009). Previously, the analysis of whole-genome sequences from 10 Onge and Jarawa individuals from the Andaman Islands and 60 mainland Indian individuals detected signatures of strong positive selection on genes related to human body size and recapitulated the Andamanese short stature when using a polygenic score based on height associated variants (Mondal et al., 2016). Among the top 20 most differentiated non-synonymous and stop gain SNPs between the Andamanese and mainland India on candidate genes for positive selection in the former group, we identified the R990G substitution (rs1042636) at the calcium-sensing receptor gene (*CASR*), with a CADD score value of 18.42, as a putative adaptive variant to experimentally follow up.

The calcium-sensing receptor (CaSR) has a critical role in calcium homeostasis by directly regulating the urinary calcium excretion and the

parathyroid hormone (PTH) secretion, which through its receptor (PTH1R) and by stimulating the production of 1,25-dihydroxy vitamin D acts on bone, kidney, and intestine to restore the appropriate circulating calcium levels in the body (Ho et al., 1995; Thakker et al., 2015). Moreover, the CaSR also participates in modulating bone formation and resorption as well as in the development of the bone growth plate (Chang et al., 2008; Goltzman & Hendy, 2015; Riccardi et al., 2013; Y. Wang et al., 2013). In particular, whereas early chondrocyte differentiation in the cartilaginous growth plate is activated via the PTH-related peptide (PTHrP)/PTH1R signaling, the insulin-like growth factor 1 (IGF1)/IGF1 receptor (IGF1R) and the extracellular Ca^{2+}/CaSR signaling pathways have been shown to mediate the terminal differentiation of the maturing and the hypertrophic chondrocytes as well as their subsequent transformation into the osteoblastic lineage (Santa Maria et al., 2016; Wang et al., 2013). However, the *CASR* gene is expressed across a wide range of tissues and has been involved in many other diverse biological functions including epidermal development, keratinocyte differentiation, hair follicle morphogenesis (Elsholz et al., 2014; Turksen & Troy, 2003), psoriasis (Zuo et al., 2015), human taste perception (Ohsu et al., 2010), as well as adipocyte differentiation and adipogenesis (Villarroel et al., 2013).

Interestingly, *in vitro* studies have already shown that the derived allele of the R990G substitution results in a gain-of-function of the CaSR that renders the receptor more sensitive to Ca^{2+} (Ranieri et al., 2013; Vezzoli et al., 2007). As expected for a CaSR activating variant, the 990G allele is associated with increased susceptibility to primary hypercalciuria (Vezzoli et al., 2007), decreased serum calcium (Kapur et al., 2010; Scillitani et al., 2004), and increased kidney stone disease risk (Guha et al., 2015). However, the pleiotropic nature of the *CASR* gene complicates any direct inference of a putative adaptive phenotype for such activating substitution in the Andamanese. Given the role of CaSR in promoting chondrocyte differentiation, we could hypothesize that the activation caused by the R990G polymorphism could induce an early closure of the growth plate. Likewise, any systemic hormonal or

mineral disturbance associated with R990G could also alter normal growth and lead to short stature. Furthermore, CaSR has an important role in epidermal development (Tu et al., 2012) and its overexpression in mice has been shown to accelerate differentiation of the epidermis, *in vivo* permeability barrier formation, and the early appearance of hair (Turksen & Troy, 2003). Although there is no direct experimental evidence between the R990G substitution and hair morphology or increased skin wound healing capabilities in humans, both phenotypes could be plausibly linked to known functions of CaSR and could have been selectively beneficial in a tropical rainforest habitat. Finally, CaSR is also expressed in adipocyte cells and its activation has been shown to promote inflammatory cytokine production, preadipocyte proliferation, and adipogenesis, while also inhibiting lipolysis (Bravo-Sagua et al., 2016; He et al., 2011; Villarroel et al., 2013). We could thus speculate whether the CaSR activation mediated by the 990G allele is associated with any differential adipose phenotype in the Andamanese, such as the steatopygia trait, which could have resulted from sexual selection or for facilitating greater fecundity and survival during famine periods.

Several general (Dong et al., 2015; Hannan et al., 2015; Ho et al., 1995; Hough et al., 2004) and tissue-specific (Chang et al., 2008, 2010; Toka et al., 2012) knock-out mice and knock-in mice for different inactivating and activating CaSR mutations have been obtained for the *CASR* gene and allowed not only to characterize many details of the CaSR function in humans but also to successfully recapitulate the disease phenotype of human illnesses such as familial hypocalciuria hypercalcaemia, severe neonatal hyperthyroidism and autosomal dominant hypocalcemia. Although we generally expect much less disturbing phenotypical effects in contrast to pathological conditions, knock-in mouse models have been proven a good strategy to experimentally validate the associated phenotype of adaptive alleles at the human *FOXP2* (Enard et al., 2009; Schreiweis et al., 2014) and *EDAR* (Kamberov et al., 2013) genes. Thus, the generation of a knock-in mouse for the R990G human substitution should be a promising approach to pinpoint any differential (and putative adaptive) phenotype across all those pleiotropic functions of the CaSR potentially shared between humans and mice.

Here, we first analyze genomic sequence data generated in the Andamanese and other SE Asian hunter-gatherer groups presenting the DSWS phenotype and test for signatures of recent positive selection across the *CASR* gene region by specifically interrogating the site frequency spectrum, and the patterns of linkage disequilibrium and population differentiation. Next, we generate a knock-in mouse for the human R990G substitution using the CRISPR-Cas9 genome editing technology and subsequently assess its phenotypic impact by exploring for potential differences between the homozygote carriers of the ancestral (R990) and derived (990G) alleles of this substitution in growth and wound healing processes as well as at adipose tissue, kidney, bone, and systemic level.

Results

Signals of positive selection and the R990G substitution

The *CASR* gene was identified by Mondal et al. (2016) as a strong candidate for positive selection in the Andamanese when applying a hierarchical boosting framework that combined eight neutrality tests to uncover the patterns of variation expected under a hard sweep scenario, while controlling for population demography (Mondal et al., 2016). Here, we recalculated three individual tests of positive selection (Tajima's D (Tajima, 1989), Fay and Wu's H (Fay & Wu, 2000), and XP-EHH (Sabeti et al., 2007) comparing Andamanese versus Yoruba) using the same whole genome sequencing data in the Andamanese and display them along 200 kb upstream and downstream the ~108 kb *CASR* gene region (Figure 1A-B, File S1). In agreement with previous results, departures from genome-wide empirical values were found in the *CASR* gene for the three selection tests. While significantly negative values of Tajima's D and Fay and Wu's H indicate an excess of low frequency polymorphisms and high-frequency derived SNPs relative to genome-wide expectations, respectively, the XP-EHH statistic points to unusually extended haplotype homozygosity on the Andamanese when compared to the Yoruba population. Thus, different footprints

of positive selection can be recognized in the patterns of genetic variation along the *CASR* gene region in the Andamanese.

When comparing the Jarawa and Onge Andamanese populations with six mainland Indian populations, the CaSR R990G substitution (rs1042636) was confirmed as the most highly differentiated functional variant in the *CASR* gene region (see Table S1). By exploring the worldwide frequencies of the derived 990G allele in populations from the 1000G project Phase 3 (Auton et al., 2015) and in an extended dataset of twelve additional populations from SE Asia (unpublished data; M Sikora and J Bertranpetit personal communication), we confirmed its highest frequencies in several hunter-gatherer groups presenting the DSWS phenotype (100% in Jarawa, 97% in Aeta, 85% in Agta, 83% in Kintaq, 78% in Jehai, and 75% in Onge and Mamanwa), while finding intermediate values in East Asia (52%) but much lower presence in other populations from South Asia (26%), Europe (7%) and Africa (3%; see Figure 1D). By contrast, the frequency of the CaSR 990G allele in populations from Papua New Guinea was clearly lower than that of the East Asian hunter-gatherer groups with the DSWS phenotype (ranging from 40% in Kosipe to 0% in Sepik). Furthermore, in accordance with the observed allele frequencies, when grouping these populations into three main SE Asian groups, we detected signatures of recent positive selection in Malaysia and the Philippines but not in Papua New Guinea (Figure 1C, Figure S1). Consequently, the CaSR R990G substitution can be hypothesized as a putative adaptive variant targeted by positive selection in several hunter-gatherer groups displaying the DSWS phenotype in SE Asia besides the Andamanese.

Generation of mice

To investigate the potential adaptive phenotype of the CaSR R990G substitution we generated a knock-in mouse model. Since the ancestral Arg allele (R990) is present in wildtype mice, we edited via CRISPR/Cas9 the corresponding orthologous nucleotide position of the mouse *Casr* gene to generate the R990G substitution in the encoded

protein (Figure S2). Briefly, embryos resulting from a cross between a B6CBAF1/J female and a C57B6/J male were harvested at stage E0.5, electroporated with the corresponding CRISPR/Cas9 reagents, and then transferred into 6-week-old CD1 recipient females. 17 offspring were obtained, and eight of them tested positive for the CaSR R990G substitution after PCR-RFLP analysis. Subsequent sequencing showed that five of the eight mosaic mice carried the expected sequence around the edited site (Table S2). Furthermore, all five animals transmitted the substitution to their offspring after breeding with CBA mice. Two independent lines were then generated from two mosaic mice (i.e., CaSR7 and CaSR15 in Table S2) by crossing the corresponding heterozygote descendants from the F1 generation. The F2 homozygotes obtained within each line, henceforth denoted as lines 7 and 15, were then used for the phenotyping analysis below. Both R990 and 990G homozygote littermates within each line were born at expected Mendelian ratios, appeared healthy, and presented normal longevity (data not shown).

Phenotyping screen in a knock-in mouse model

Body weight growth curves and Body Mass Index (BMI)

Body weight changes were recorded from weeks 3 to 20 and examined between lines (7/15), sexes (male/female), and genotypes (homozygotes R990/homozygotes 990G) weekly through the use of factorial ANOVAs, as well as by using a mixed-model ANOVA with repeated measures across weeks 4 to 12. When considering all mice together, sex was the variable that most influenced body weight, explaining 46.35 % of the total variation (average percentage across weeks 4 to 12), whereas genotype and line also displayed weekly significant differences in body weight and explained 8.17 % and 4.23 % of the variation, respectively (see weekly details in File S2). Notably, homozygote carriers for the 990G allele displayed significantly higher body weight than the homozygote carriers for the ancestral allele in the whole set of animals (repeated measures ANOVA, $p = 0.00003$) as well as when separating mice into the two generated lines (repeated

measures ANOVA, p = 0.00028 for line 7 and p = 0.031 for line 15; see weekly details in File S2). However, whereas the percentage of the body weight variation explained by the genotype was 13.46 % in line 7 (average percentage across weeks 4 to 12), in line 15 it explained only 4.74 % of the variation. Accordingly, when separating mice by sex and line, an ANOVA with repeated measures across weeks 4 to 12 showed significant differences in body weight between genotypes in both males (p = 0.022) and females (p = 0.005) in line 7 but only in males (p = 0.023) in line 15 (Figure 2A; Figure S3; File S2). Moreover, at the final point, homozygote carriers for the 990G allele displayed a clear tendency toward presenting higher BMI. In particular, we observed statistically significant differences between genotypes when analysing all animals together (Kruskal-Wallis test, p = 0.035), and when separating mice by sex (females Mann-Whitney U Test, p = 0.047; males Mann-Whitney U Test, p = 0.035) but only in the female group of line 7 when separating mice by line and sex (Mann-Whitney U Test, p = 0.021; Figure 2B; Figure S4). Since female homozygote carriers for the 990G allele in line 7 also presented significantly higher weight at the final point (Mann-Whitney U Test, p=0.033) and no statistically significant trend was observed between genotypes and mice crown-rump length (File S2), the observed differences on BMI can be mostly attributed to weight.

Blood and urine biochemistry

Blood and urine samples were taken at around 12-14 weeks of age to analyse 16 different metabolites and hormones mostly related to calcium homeostasis and lipid metabolism (File S3). In agreement with the CaSR activation produced by the R990G substitution (Ranieri et al., 2013; Vezzoli et al., 2007), when analysing all animals together, the homozygote carriers for the derived allele display a trend towards presenting lower calcium concentration in blood (factorial ANOVA; p = 0.060; Table 1). Serum calcium differences between genotypes were statistically significant only when considering males from both lines together (factorial ANOVA; p = 0.017), which also displayed significant differences for serum phosphorus (factorial ANOVA; p = 0.004).

Moreover, significant serum phosphorus differences between genotypes were also detected when only analysing males from line 15 (one-way ANOVA, $p = 0.038$) (see Table 1). No clear differential patterns between genotypes were found for urinary calcium and phosphorus excretion or the parathyroid hormone (PTH) and osteocalcin serum concentrations when testing a minimum of 5 animals per sex, genotype, and line (results not shown; File S3). However, we found consistent significant differences in the HDL and LDL cholesterol levels between the R990G CaSR genotypes (Table 1). In particular, the homozygote carriers for the derived allele presented significantly lower HDL and LDL blood concentrations when analysing all animals together (factorial ANOVA; $p = 0.020$ and $p = 0.001$ for HDL and LDL differences between genotypes, respectively) as well as when considering only those of line 15 (factorial ANOVA; $p = 0.016$ and $p = 0.002$ for HDL and LDL differences between genotypes, respectively), which also presented significantly lower total cholesterol (factorial ANOVA; $p = 0.019$). When further separating animals by line and sex, only line 15 displayed statistically significant differences in cholesterol levels: homozygote female carriers for the 990G derived allele presented lower total cholesterol (one-way ANOVA, $p = 0.043$), lower HDL (one-way ANOVA, $p = 0.009$) and lower LDL (one-way ANOVA, $p=0.031$) levels, while homozygote 990G male carriers presented only lower LDL levels (one-way ANOVA, $p = 0.022$).

Body composition and skeletal phenotype

Whole body fat and lean composition were determined with an EchoMRI instrument for 33 animals from line 7 at 14 weeks of age (File S4). Males and females significantly differed in their overall fat and lean content (Kruskal-Wallis test, $p = 0.021$ and $p < 0.001$, respectively), whereas when comparing body composition between genotypes we detected statistically significant differences only in females (Figure 2B; Figure S5). In particular, female homozygote carriers of the 990G CaSR allele presented a tendency towards higher overall fat content (Mann-Whitney U test, $p = 0.050$) and significantly lower lean percentage

(Mann-Whitney U test, p = 0.031) than the female homozygote carriers of the ancestral allele.
Since mutants of the *Casr* gene in mice have been associated with differential skeletal phenotypes (Chang et al., 2008; Goltzman & Hendy, 2015; Riccardi et al., 2013; Y. Wang et al., 2013), tibial cortical bone and trabecular bone at the proximal tibial metaphysis were trabecular ex-vivo by micro-computed tomography on 28 animals (including between 2-4 animals per sex, line, and genotype). No differences in bone micro-architecture were identified between the CaSR R990G genotypes as assessed by standard parameters such as mean density, bone volume, bone mineral content, bone mineral density, tissue mineral content, tissue mineral density, bone volume fraction, bone surface density, trabecular thickness, trabecular separation, and trabecular number (File S4; Figure S6).

Wound healing and epidermis histology

We also monitored wound healing and epidermis histology after performing 5 mm skin biopsy punches at 10 and 20 weeks of age. However, no major differences in epidermal thickness (results not shown) nor significant wound-healing rate differences or differential reepithelization patterns were detected between the genotypes of the CaSR R990G substitution (File S5; Figure S7).

Liver and fat histology

Liver biopsies from 92 animals between 20-26 weeks of age were categorized into a normal or a steatosis phenotype, depending on whether they presented normal to mild lipid droplets accumulation or moderate to severe fatty liver, respectively (Figure S8). Up to 41.3% of the animals displayed hepatic steatosis without significant differences across lines, sexes, and genotypes except for line 7 females, in which steatosis was more prevalent among the homozygote carriers of the 990G allele when compared to the homozygotes for the ancestral allele (Pearson's chi-squared test, $p<0.0033$; File S6). Subsequently, biopsies of visceral and subcutaneous fat from the same animals were used to

investigate differences in the adipocyte diameter and size distribution between genotypes. When analysing animals without hepatic steatosis, homozygote carriers of the 990G allele displayed a clear trend towards significantly smaller adipocyte diameters in both lines, sexes, and fat tissues as well as towards an excess of cells in the smallest size bins (Figure 3A-B; Files S7-8 and Figure S9). By contrast, animals with steatosis in the liver seem to show the opposite trend (data for both sexes and genotypes only available for line 15). In that case, homozygote male carriers of the 990G allele displayed significantly larger adipocyte diameters and excess of cells in the medium and largest size bins in both subcutaneous and visceral fat (Student's t-test $p < 0.0001$ and Chi-squared Test $p < 0.0001$). However, while homozygote 990G females with steatosis show significantly larger adipocyte diameters and excess of cells in the medium and largest size bins in subcutaneous fat (Two sample T-test, $p < 0.0001$; Chi-squared Test, $p < 0.0001$), they were found to present significantly smaller visceral adipocyte sizes (Two sample T-test, $p < 0.0001$; Chi-squared Test, $p < 0.0001$, File S8).

Expression of lipolytic and lipogenic genes

RNA was extracted from liver and fat (subcutaneous and visceral) biopsies at 20-26 weeks of age to analyze the relative expression of genes involved in lipogenesis, lipolysis, and inflammation (File S9, Figures S10-S11). In visceral fat from animals without hepatic steatosis, homozygote carriers of the 990G allele presented significantly lower mRNA expression of adipose triglyceride lipase (ATGL) (Student's t-test $p = 0.029$ in females and $p = 0.014$ in males) when compared with R990 homozygotes, whereas homozygote females for the 990G allele also showed significantly lower peroxisome proliferator activating receptor gamma (PPARγ) mRNA expression (Student's t-test $p = 0.036$, Figure 3). Moreover, even if not significant, we found a trend towards lower mRNA expression for the hormone-sensitive lipase (HSL) in 990G homozygote females (Student's t-test $p = 0.055$) and for the adiponectin in 990G homozygote males (Student's t-test $p = 0.088$) in visceral fat from animals without steatosis (File S9 and Figure S10). As

for the liver, 990G homozygotes displayed a clear tendency towards higher ATGL mRNA expression in both sexes (Student's t-test $p = 0.031$ in females and $p = 0.071$ in males when analyzing line 7 animals without steatosis). By contrast, animals with hepatic steatosis (results only available in line 15), show differential gene expression patterns between sexes when comparing genotypes: 990G homozygote males presented significantly lower mRNA expression of HSL and lipoprotein lipase (LPL) in visceral fat tissue (Student's t-test $p = 0.036$ and $p = 0.038$, respectively) as well as of fatty acid binding protein (FABP4/aP2) in the liver (Student's t-test $p = 0.035$) but 990G homozygote females presented significantly higher HSL mRNA expression when compared to homozygote wildtype animals (Student's t-test $p = 0.022$). Finally, no statistical differences in gene expression were found between genotypes in subcutaneous fat (File S9, Figures S10-S11), whereas CaSR gene expression was confirmed in all three tissues at the qPCR detection limit without differences between genotypes (results not shown).

BMI, weight, and height in the Agta

We next investigated whether the R990G polymorphism can be associated with differential trait values for BMI, weight, and height in 138 Agta individuals that have been genotyped with the Axiom Genome-Wide Human Origins array (Dobon et al., 2022; Migliano et al., 2013). After the corresponding QC filters and the imputation of the rs1042636 genotype using 81 SE Asian individuals sequenced with the 10X technology (see details in Materials and Methods), 68 Agta individuals were found to carry the GG genotype, 52 were heterozygotes GA and 11 were AA for rs1042636 (data not shown). Upon consideration of four different genetic models (i.e., additive, homozygote-only, recessive, and over-dominant), no differences in the means of the phenotype values for BMI, weight, and height were found between the genotypes in any of the two sexes (Kruskal-Wallis or t-test, Figure S12). Under the homozygote-only model, we also modeled the relationship between each phenotype and the rs1042636 genotype incorporating the effect of age and relatedness of the Agta individuals

while separating them by sex. For none of the phenotypes, the SNP-based association tests provided a better fit than the null model in either of the sexes. However, the phenotype values for BMI, weight, and height in the GG homozygotes at rs1042636 were higher than those in the AA homozygotes, even if the estimated effect size of the rs1042636 genotype was very small in all cases (≤ 0.201; data not shown) being weight the one with the higher effect size.

Discussion

After identifying the R990G substitution in the *CASR* gene as a putative adaptative variant in the Andamanese to follow-up, we first used genomic data to investigate whether other SE Asian hunter-gatherer groups presenting the DSWS phenotype shared the same accompanying signatures of recent positive selection and high prevalence of this nonsynonymous polymorphism. No unusual patterns of variation were found in Papuan New Guinea, but hunter-gatherer groups from Malaysia and the Philippines presented significant deviations from neutrality in the *CASR* flanking gene region and high frequencies for the associated G allele at rs1042636 (encoding for 990G). Thus, our results point to the action of positive selection acting on *CASR* in these hunter-gatherer populations. Next, we generated a knock-in mouse for the human R990G substitution and assessed its phenotypic impact by exploring for potential differences between the ancestral (R990) and derived (990G) homozygote sibling carriers resulting from several heterozygote crosses within two independent R990G KI mouse lines.

As expected for an activating CaSR mutation, the derived allele of the R990G (rs1042636) polymorphism is known to be significantly associated with lower serum calcium in humans (Kanai et al., 2018). In particular, the GWAS signal was described in the Japanese population (N=71,701) when after using a linear regression model under the assumption of an additive effect of the SNP dosages a beta value of -0.069 and a p-value of 1.23E-38 was originally reported for rs4678176 (Kanai et al., 2018), which is with high LD with rs1042636 (with r^2=1

in the Japanese, and ranging from 0.94 to 1 in all South Asian populations from the 1000G Project). Although only statistically significant when grouping male mice from the two lines (N=63), our humanized mouse model for the R990G substitution in CaSR showed a clear trend towards presenting lower serum calcium concentration, which, at least in males, was also accompanied by higher serum phosphorous. Assuming in mice the same effect size on serum calcium concentrations described in humans for the rs1042636 polymorphism, the chance for detecting an association with a p-value <0.05 was 84.06% in the whole mice group (N=120) and 52.37% in the female group (N=57). Overall, these results seem to recapitulate the expected hypocalcemia phenotype of the R990G substitution in mice.
The clearest differential phenotype found in our R990G KI mouse model was the higher weight of the derived allele homozygote carriers, which, at least in females, was accompanied by a tendency towards presenting higher BMI, higher fat as well as lower lean body content. Interestingly, calcium supplementation in the diet has been shown to elicit significant body weight loss and body fat content in mice and rats (Sun et al., 2012; Zhang et al., 2018). Similarly, supplementation of dietary calcium in humans has been suggested to reduce the body weight of children and adults (Heaney et al., 2002; P. Li et al., 2016) and the abdominal visceral adipose depots in overweight and obese adults (Rosenblum et al., 2012). On the contrary, rats fed with low calcium diets not only showed a greater visceral fat mass and CaSR expression in white adipose tissue but also lower serum fatty acids and glycerol concentrations than normal calcium-fed rats (He et al., 2011). Thus, the finding of greater body weight and fat content in 990G homozygotes that we have shown display a tendency towards presenting lower serum calcium when compared to homozygote R990 animals is consistent with previous studies.

Interestingly, the anti-obesity effects of dietary calcium have been suggested to result from the direct role of calcium in modulating fat metabolism (Zhang et al., 2019) with intracellular calcium having a key role in regulating lipid metabolism and triglyceride storage in adipocytes (Zemel, 2002). In agreement with these observations, distinct systemic

and molecular accompanying phenotypes related to lipidic handling and adipose tissue homeostasis were further observed between the homozygote carriers of the ancestral and the derived alleles of the R990G substitution in our KI mouse model. These included differential cholesterol, HDL, and LDL serum concentrations, as well as contrasting adipocyte sizes and expression of lipolytic genes (i.e., ATGL and PPARγ). In addition, we found that the presence of a fatty liver conditioned the direction of the effects of the R990G substitution on adipocyte size in visceral and subcutaneous fat as well as on specific gene expression patterns in visceral adipose tissue. Although under an excess of hepatic fat accumulation (steatosis) we observed some expected effects resulting from the activation of CaSR associated with obesity, we suggest that the putative adaptive role of the R990G substitution in hunter-gather human groups presenting the DSWS phenotype is probably determined in a physiological environment of no pathological accumulation of fatty acids.

Most of the lipidic and adipose features found associated with the 990G genotype in our study can be expected to result from either the higher preadipocyte proliferation or the antilipolytic effects widely recognized to happen upon activation of CaSR. On one hand, CaSR activation in visceral white adipose tissue has long been recognized to increase the proliferation of adipose progenitor cells and promote adipocyte differentiation (Bravo-Sagua et al., 2016). In human SW872 adipocytes, for example, the adipocyte differentiation and adipogenesis promoted after exposure to CaSR activators were suggested to result by upregulating PPARγ expression levels and that of subsequent downstream adipogenic genes (He et al., 2012). S imilarly, different CaSR agonists were found to increase preadipocyte proliferation through the ERK signaling pathway in human LS14 adipocytes (Rocha et al., 2015; Villarroel et al., 2013). Besides, the use of troglitazone (PPARγ agonist) in young obese rats has also been shown to result in the increase of small adipose cells and the decrease of large adipocytes without affecting the total mass of adipose tissue (Okuno et al., 1998). Therefore, we propose that the smaller adipocyte sizes observed in the 990G homozygote carriers presenting no hepatic steatosis could result from this early mitogenic effect of CaSR activation in preadipocytes.

Such an effect could lead to the maturation of a higher number of adipocytes, providing greater fat storage capacity and leading to the higher body weight and BMI observed in the 990G homozygotes. By contrast, 990G homozygote carriers with hepatic steatosis displayed a clear tendency towards greater visceral and subcutaneous adipocyte cell sizes and adipocyte hypertrophy. As in obesity, in these animals the maximum adipocyte capacity to store an excess of fatty acids as triglycerides in lipid droplets is probably being reached and, as a result, fatty acids start accumulating in the liver, among other non-adipose tissues. We did not detect, however, a clear pattern of differential expression of inflammatory or lipogenic genes associated with the R990G substitution at 23-26 weeks of age.

On the other hand, the stimulation of CaSR has been recognized to promote triglyceride accumulation and adipocyte hypertrophy (cell size increase) by coordinating not only the activation of adipogenesis but also the suppression of lipolysis (He et al., 2012). Particularly, in rats fed with low calcium diets, CaSR was suggested to affect fat accumulation via antilipolytic pathways (He et al., 2011). Furthermore, *in vitro* experiments demonstrated that the decreased expression of the two main lipolytic enzymes (i.e., HSL and ATGL) in SW872 adipocytes because of CaSR expression stimulation was attained by increasing the intracellular calcium and reducing the cAMP levels (He et al., 2011). In accordance with this antilipolytic effect of CaSR activation, our results clearly show that both male and female 990G homozygote carriers with normal liver present significantly lower ATGL mRNA expression in visceral fat. This result is also compatible with the observation that the human polymorphism rs1042636 causes a reduction of at least 20% of the lipolysis in omental adipose tissue (Reyes et al., 2012). In that case, not only the activation of CaSR with agonists, such as cinacalcet, in isolated human adipocytes has been shown to elicit an inhibitory effect of basal lipolysis but rs1042636 carriers of the G allele (encoding for 990G) were found to display higher sensitivity to the antilipolytic effect of cinacalcet (Reyes et al., 2012). Additionally, we also detected that the female homozygotes for the 990G allele presented significantly lower expression of PPARγ, which is in accordance with the lower mRNA

expression of lipolytic genes observed in that genotype. Indeed, activation of PPARγ with an agonist (rosiglitazone) has been shown to increase adipose tissue lipolysis in rats (Festuccia et al., 2006). Accordingly, a PPARγ-responsive element has been identified in the promoter sequence of the *Atgl* gene in mouse and rosiglitazone was found to increase *Atgl* expression (Nielsen et al., 2014).

As for mice with hepatic steatosis, we note that contrasting patterns were found between males (presenting lower expression of HSL, the second major lipolytic enzyme) and females (showing higher HSL expression levels) homozygotes for the R990G substitution, which could probably result from an additional regulatory role of estrogen influencing adipogenesis and lipolysis in females (Journé et al., 2004; Wohlers & Spangenburg, 2010). On the other hand, a metabolic link has been suggested between the lipolytic activity of adipocytes and the rate of cellular efflux to HDL (Verghese et al., 2007). Thus, the lower cholesterol levels observed in the homozygote carriers of the 990G allele could be reflecting the corresponding attenuation of lipolysis resulting from the lower visceral adipocyte expression of ATGL and HSL. Finally, 990G homozygote animals without hepatic steatosis showed a tendency towards higher expression levels of ATGL in the liver. Even if only significant in females, such a trend could help to compensate for the lipidic imbalances caused by R990G. Congruently with this, higher expression levels of ATGL and enhanced lipolysis rates in the liver have been previously described to prevent the development of steatosis (Li et al., 2021; Ong et al., 2011; Reid et al., 2008).

Different limitations of the present study must be acknowledged. First, despite the widespread use of the mouse as a fair experimental model to study biological processes and diseases in humans, our slightly different physiologies and anatomies might not only difficult the direct translation of experimental findings in mice into humans but also limit the mouse model applicability to accurately recapitulate human phenotypes. Secondly, although during the CRISPR-Cas9 editing experiment many strategies were applied to minimize the possible off-

target effect, such as RNP *in situ* electroporation directly into one-cell zygotes and the use of an asymmetric vector, the two lines deriving from two independent knock-ins displayed some phenotypical differences between them. No modifications were detected when sequencing the off-target loci predicted by the Benchling tool used for the two gRNAs design, nevertheless, rare off-target events could still occur (Atkins et al., 2021). Additionally, other background genetic differences could have been introduced within each line's initial crosses. Finally, we do not exclude mild phenotypic effects of the R990G substitution in bones, skin, or in other tissues that could not be retrieved in this study due to the sample size of the analyzed cohort. Similarly, we also knowledge that the sample size of Agta individuals with phenotypical information such as body weight is clearly limited to attain significance given the estimated effect size of the polymorphism.

While the 990G allele is found at intermediate frequencies in several populations from East Asia (i.e., 56.7% in CHS, 53.8% in JPT, and 52.4% in CHB), we note that the *CASR* gene region displayed signatures of positive selection only in some hunter-gatherer groups from SE Asia, where the polymorphism reaches its highest frequencies (i.e., 87.5% in Andamanese, 96.9% in the Philippine Aeta and 83.3% in the Malaysian Kintaq). In turn, whereas the physiological consequences we describe for this non-synonymous polymorphism are expected in all its carriers, these differences may have been selectively beneficial only under the environmental conditions, lifestyle, and genomic background of the SE Asian populations displaying the DSWS phenotype. Therefore, R990G cannot be inferred to provide a universal advantageous phenotype to its carriers. It cannot be used either to infer their ancestry. We propose that the differential adiposity profile and higher body weight associated with the CaSR R990G substitution could have been positively selected in hunter-gather human groups from SE Asia to facilitate fat storage, increasing survival in periods of food scarcity in their past. Indeed, adipocyte number is a major determinant for fat mass in human adults (Spalding et al., 2008) and the lower expression of lipolytic genes in visceral adipose tissue might further

favor adiposity preventing an unnecessary rapid mobilization of lipids under fasting or exercise conditions (Nielsen et al., 2014).

Moreover, we could further speculate that the distinct lipidic handling features and differential adipose tissue homeostasis caused by R990G could also contribute to some features of the DSWS phenotype such as the unusual short stature, by advancing puberty, or the development of steatopygia. Anthropometric studies show that fat percentage in the human body increases to facilitate puberty, especially in the central region of the female body, leading to an increase in leptin secretion, which regulates menarche inducing both the pubertal growth spurt and the growth plate fusion in the epiphysis (Lassek & Gaulin, 2007; Miranda et al., 2014; Siervogel et al., 2003). Such an adaptive scenario would also be consistent with a genome-wide analysis of age at menarche and bone mineral density conducted on 2,522 females from 414 pedigrees that reported suggestive linkage evidence with LOD score = 2.31 for the 3q13 region, containing the *CASR* gene (Pan et al., 2008) as well as with the results of Guo et al. (2008) that reported the 3q13 region significantly influencing menarche through interaction with genes on 22q13 (including *PPARG*, among others) (Dvornyk & Waqar-ul-Haq, 2012; Guo et al., 2006). Furthermore, the onset of sexual maturation was previously described to be a plastic mechanism with which humans react to stressful environments both in groups presenting the DSWS phenotype (Stock & Migliano, 2009) and in populations from central Europe and the Americas (Coall & Chisholm, 2003). Thus, overall, our results may support the hypothesis that the low life expectancy of the hunter-gather human groups presenting the DSWS phenotype has selected for early maturity and reproduction over physical growth to enhance fertility (Migliano et al., 2007, 2013).

Materials and Methods

Signals of selection

SNP genotypes from ten Andamanese and ten Indian Irula whole genome sequences generated at ~15x coverage (Mondal et al., 2016) and nine unrelated YRI whole genome sequences from the Complete Genomics dataset (Drmanac et al., 2010) were merged and curated by removing any SNP which had missing data for any individual and subsequently phased with SHAPEIT (Delaneau et al., 2012) using the 1000 G Project Phase 1 samples as reference. Tests of positive selection based on the site frequency spectrum (Tajima's D (Tajima, 1989) and Fay and Wu's H (Fay & Wu, 2000)) and on comparing the extended haplotype homozygosity profiles between populations (XP-EHH (Sabeti et al., 2007)) were then genome-wide calculated in the Andamanese as in Mondal et al. (2016). Tajima's D (Tajima, 1989) and Fay and Wu's H (Fay & Wu, 2000) were calculated using 3 kb windows, whereas XP-EHH (Sabeti et al., 2007) comparing Andamanese versus Yoruba was calculated by SNP.

To identify any putative adaptive variants in the *CASR* gene region, we then used ANNOVAR (Wang et al., 2010) to functionally annotate all SNPs presenting allele frequencies differences greater than 0.25 when comparing the Jarawa (JAR) and Onge (ONG) populations to six mainland Indian populations, including Uttar Pradesh Upper Caste Brahmins (UBR), Rajput (RAJ), Vellalar (VLR), Irula (ILA), Birhor (BIR) and Riang (RIA) (see Table S1). We also explored the allele frequencies for the R990G substitution and signals of positive selection along the *CASR* gene region in other south-eastern Asian populations for which phased whole genomes obtained with the 10X technology were available (unpublished data; M Sikora and J Bertranpetit personal communication). These included several DSWS groups from the Philippines (Aeta 2N=32, Batak 2N=24, Agta 2N=20, Mamanwa 2N=12) and Malaysia (Jehai 2N=18, Kintaq 2N=12, Temiar 2N=12) as well as different Indigenous groups from Papua New Guinea

(Koinambe 2N=24, Sepik 2N=4, Kosipe 2N=20, Papua New Guinea Highlands 2N=4, Australia 2N=4).

For all three south-east Asian groups, we applied a test of positive selection based on the site frequency spectrum (SF select (Ronen et al., 2013) available at https://github.com/rronen/Sfselect) following the same procedure as in Walsh et al. (2020), and computed the XP-EHH statistic (Sabeti et al., 2007) comparing the Philippines and the Malaysian DSWS groups to the Papua New Guinea group, with the rehh 2.0 program (Gautier et al., 2017). For each selection test and population group, z-scores across the *CASR* gene region were computed from the mean and standard deviation of the genome-wide raw values of each corresponding statistic (File S1).

Guide RNAs design and Cas9 assay for gRNA efficiency

To generate the CaSR R990G substitution, two guide RNAs (gRNAs) were designed using the Benchling design tool (www.benchling.com). In particular, gRNAs were selected to have the highest score for both the off-target and on-target sites and to have a cleavage site as close as possible to the targeted nucleotide. The resulting gRNAs sequences were as follows: 5'- CGGGAACTCCATGCGCCAGA -3' and 5'- ACGGGAACTCCATGCGCCAG -3'.
To test gRNA efficiency, a substrate DNA was first created by subcloning *Casr* exon 7 from mice into a Zero-Blunt cloning vector (Invitrogen) (see Figure S13) and a Cas9 cleavage assay was subsequently performed for each gRNA. Vectors were linearized and purified. To assay cleavage, a total of 300 nM of a 1:1 (vol/vol) mix of crRNA and tracrRNA was incubated for 5 min at 75ºC and subsequently allowed to cool to room temperature. 1ul of Cas9 nuclease (3ug/ul), 3 µl of NEB Cas9 buffer, and 11 µl of H2O were then added to the mix and incubated for 10 min at 25 ºC to allow the formation of the ribonucleoprotein (RNP) complex. 5 µl of substrate DNA (30 nM) was then added to the solution and incubated for 1h at 37 ºC. Cleavage products were subject to agarose gel electrophoresis, and the efficiency

of the gRNA was measured by the ability to efficiently cleave substrate DNA.

Electroporation of CRISPR RNP's and HDR template into zygotes

A homology-directed repair (HDR) template to generate the R990G substitution was designed according to guidelines (Richardson et al., 2016) and purchased from IDT (Integrated DNA Technologies). For gRNA annealing, 10ul (at 54 ng/μl) crRNA plus 15.6 ul (at 84 ng/μl) tracrRNA were incubated for 10 min at 75ºC followed by gradual cooling to room temp. Subsequently, 17.94 μl of the gRNA annealing reaction was incubated for 10 min at 25 ºC with 5.42 ul (at 125 ng/μl) of Cas9 nuclease, after which 32.5 ul (at 500 ng/μl) of ssDNA R990G template, plus 74.14 μl OPTIMEM were added. 40 μl of this solution (assembly reaction) were aliquoted and used for the electroporation.
4-week-old B6CBAF1/J female mice were superovulated with 5IU of PMSG and 5IUhCG and were subsequently mated with male C57B6/J stud mice. Zygotes were harvested at stage E0.5. Electroporation was then carried out with a NEPA21 electroporator (NEPAgene) using the following conditions: Assembly solution containing zygotes; 40ul. Poring pulse; 250V, length; 1.5mS, interval; 50mS, decay rate; 10, number of pulses; 4. Transfer pulse; 20V, length; 50mS, interval; 50mS, decay rate; 10, number of pulses; 5. Following electroporation, surviving zygotes were transferred into 6-weeks-old CD1 recipient females via oviduct transfer. All animal handling and experimental procedures conducted at the Mouse Mutant Core Facility (IRB) were approved by the Ethical Committee of the Parc Científic de Barcelona (PCB) (project 19-013-9170PRO-SPF).

DNA extraction and R990G genotyping

Genomic DNA was extracted from tail or ear biopsies of 2-week-old mice by a standard proteinase K and isopropanol-based method (File S10). Since the R990G substitution generates a consensus sequence for the BtgI restriction enzyme (Figure S2), we genotyped all newborn mice

by PCR-RFLP analysis. To do so, CaSR exon 7 genomic DNA was amplified using forward and reverse primers 5'-AACACCATCGAGGAGGTGC-3´ and 5'-CTCCACCGCTGATGACGAAG-3´, respectively. PCR reactions were performed in a volume of 50 µl containing 1x KOD-HS Buffer, dNTPs (200uM final), Mg^{2+} (1.5mM final), 2 µl of forward primer (400nM final), 2 µl of reverse primer (400nM final), 0.5 µl KOD Hot Start DNA polymerase (1 unit) and 1 µl of purified DNA (100-250 ng/µl). PCR conditions were 1 cycle of initial denaturation at 95 ºC for 3 min; 40 cycles of denaturation at 95 ºC for 30 sec, primer annealing at 60 ºC for 20 sec and extension at 70 ºC for 20 sec; final extension at 70 ºC for 5 min. After amplification, PCR products were run on agarose gels, purified with GeneJET Gel Extraction Kit (Thermo fisher scientific), and subsequently incubated for 90 min at 37ºC with the *BtgI* enzyme (BioLabs) following manufacturer's instructions.

Editing efficiency and quality

To evaluate the efficiency and quality of the editing, we cloned the amplified PCR products from each positive mouse into a pCR-Blunt vector (Invitrogen) following the manufacturer's conditions. A total of 24 colonies from each founder were analyzed. Plasmid DNA was amplified, and subject to RFLP analysis with BtgI. DNA from each colony was then extracted using PureLink™ Quick Plasmid Miniprep Kit (Thermo fisher scientific) and subsequently used for PCR-RFLP analysis as previously described to verify CRISPR editing efficiency. Up to 4 colonies from each confirmed positive mosaic were then Sanger sequenced to ensure the correct editing of the R990G substitution and discard any potential insertions or deletions introduced during the CRISPR Cas9 editing protocol.

Establishment of two lines and animal experimental procedures

Mosaic founder mice generated at the Mouse Mutant Core Facility (IRB) and harboring the corrected edited position were then bred with CBA mice (6 weeks) from Charles River Laboratories (France) to

generate founder lines at the Barcelona Biomedical Research Park (PRBB) animal facility. We selected one mosaic founder deriving from each separate editing experiment and gRNA to establish two independent CaSR R990G mice lines to be phenotyped in our study (Table S2). Genotyping was performed by PCR-RFLP analysis after extracting DNA from either tall or ear biopsies obtained just after weaning as described above. After establishing 17 crosses between heterozygote individuals in the CaSR7 F1, and 10 crosses between heterozygote individuals in the CaSR15 F1, a total of 125 F2 mice in line 7 and 86 F2 mice in line 15 were used for experimental procedures. Thus, all phenotyped animals belonged to the filial generation 2 (F2) to minimize the drift effect of inbreeding across several generations. Moreover, up to 7 and 11 potential off-target sites for each corresponding gRNAs, as predicted by the Benchling designing tool, were checked by standard Sanger sequencing in all founders for lines 7 (22 mice) and 15 (21 mice), respectively (aligned sequences are available in File S11). Off-target genes and primers for the amplification and sequencing are shown in Tables S3-4.

All generated mice for the phenotyping were kept under pathogen-free conditions in a controlled temperature and humidity environment with a 12:12 hr light: dark cycle and had free access to water and were fed ad libitum on a standard diet at the PRBB animal facility. All animal handling and experimental procedures for the phenotyping were conducted in accordance with the European Union Directive 2010/63/EU and the Spanish Legislation (Real Decreto 53/2013, BOE 34: 11370-11421) and were approved by the Ethical Committee of Animal Experimentation of the Barcelona Biomedical Research Park (project PML-18-0014) and the Animal Research Committee of the Department of Territory and Sustainability, Generalitat de Catalunya, Spain (project 10043).

Blood and urine biochemical analysis

Blood samples were collected from the retro-orbital vein at the 14th week of age and left to rest for 30 minutes. Serum was then separated

by centrifugation at 2000 rcf for 10 minutes at 4°C, aliquoted conveniently, and stored at -20ºC until further analysis. Serum samples were analyzed for calcium, magnesium, phosphorus, creatinine, fructosamine, triglycerides, NEFAs, total cholesterol, HDL, and LDL on an Automated clinical chemistry analyzer Olympus AU400. Serum PTH and osteocalcin concentrations were measured using the Mouse PTH 1-84 ELISA Kit (Immutopics, Cat.# 60-2305) and the Mouse Osteocalcin EIA Kit (Alfa Aesar, Catalog No. J64239), respectively, following the manufacturers' protocols.
Urine samples were collected from mice housed in metabolic cages at around 13 weeks of age. After 48 hours, twenty-four-hour urine was collected from each animal and rapidly stored at –20°C until further analysis. Urinary Ca, inorganic phosphate (Pi), and creatinine were measured on an Automated clinical chemistry analyzer Olympus AU400. All blood and urine biochemical analyses were determined at the Servei de Bioquímica Clínica Veterinària, Universitat Autònoma de Barcelona (Cerdanyola del Vallés, Barcelona).

Wound healing experiments and skin histology

Skin biopsies were performed after depilation using 5 mm disposable punches on anesthetized mice (line 7). Wound healing was evaluated at a macroscopic level and monitored by taking digital photographs at different time points after the lesion was performed at either 10 (6 mice) or 20 weeks (12 mice). At the same time points, the larger and minor diameters of the skin lesions were also manually measured by using a caliper. Wound areas were determined by applying the formula: (large diameter/2) x (minor diameter/2) x π, whereas the wound-healing rate was then estimated as ((area of original wound – area of actual wound)/area of original wound) x 100. Skin biopsies at 20 weeks of age were incubated in paraformaldehyde 4% overnight, paraffine embedded to be cut in 5μm thickness, and then stained with hematoxylin and eosin. Differences in epidermal thickness between genotypes were evaluated using 5 measures per sample and region (derma or epidermis) (Figures S6 and S7, File S5).

Body weight growth curves and Body Mass Index

Body weight changes were recorded from weeks 3 to 20 and examined weekly between lines, sexes, and genotypes through the use of factorial ANOVAs. Since several experimental procedures (such as blood extraction, use of metabolic cages for urine collection as well as transport and anesthetics for the micro-CT analysis performed outside the PRBB animal facility facilities) could be affecting unevenly numbers of individuals across the studied groups, we also applied a mixed-model ANOVA with repeated measures along weeks 4 to 12 to analyze the overall effect of the examined variables before performing any of the aforementioned experimental procedures. Body weight and crown-rump length were recorded at the final point (between weeks 20 and 26) for 128 animals. The Body Mass Index (BMI) was calculated as the ratio between body weight (g) and square crown-rump length (mm). Weight, crown-rump length, and BMI among experimental groups were subsequently normalized by age.

Body composition analysis and micro-computed tomography

Body composition analyses were performed at 14 weeks of age at the Animal Facility of the Parc Científic de Barcelona (PCB) within the PCB-PRBB Animal Facility Alliance. Fat and lean body mass of non-anesthetized live mice were determined using the Echo-MRI Analyzer system (Echo Medical Systems, Houston, TX, USA). For ex-vivo micro-CT analysis, femur and tibia bones were isolated at the final point (25 weeks of age), fixed in PFA 4% overnight at 4ºC, washed in PBS, and then stored in PBS at 4ºC. The bone micro-architecture of the tibial cortical bone and the trabecular bone at the proximal tibial metaphysis were assessed using a Perkin Elmer Quantum Fx instrument at the Preclinical Imaging Platform of the Lab Animal Service of the Vall d'Hebron Institut de Recerca (VHIR).

Histopathological analyses

Skin biopsies and liver, kidney, and thyroid gland samples obtained at the terminal point were fixed in 10% buffered formaldehyde, dehydrated in increasing concentrations of ethanol, and embedded in paraffin wax. Tissue samples were subsequently sectioned and stained with hematoxylin and eosin for pathological evaluation at the Departamento de Medicina y Cirugia Animal, Facultad de Veterinaria, Universidad Complutense de Madrid.

Dissection and histology of liver and fat tissue samples

Adult animals were sacrificed from the 5th up to the 6th month of their life to retrieve the medial and left liver lobes, inguinal subcutaneous white fat pads, and visceral white fat pads. One-half of each dissected tissue was incubated in RNAlaterTM stabilization solution (Invitrogen) according to the manufacturer's instruction to preserve the RNA of each specimen for gene expression quantification. The other half was incubated for 12 h at 4ºC in paraformaldehyde (4% in PBS), then, paraffine inclusion and hematoxylin and eosin stains were performed by the Histology Unit of the Centre for Genomic Regulation in Barcelona. 12 μm thick slides were produced for each tissue and observed using a Zeiss AxioImager Z1 microscope (Apotome) (Carl Zeiss MicroImaging).

Liver biopsies were categorized into a normal or a steatosis phenotype according to the degree of lipid accumulation in sections of the tissue stained with hematoxylin and eosin using as references figure 1 in Xu et al. (2012) and figure 4 in Kristiansen et al., 2016). In particular, we classified those liver sections presenting normal and mildly phenotype with lipid droplets as "normal", and those presenting moderate to severe fatty liver as "steatosis" (see Figure S9). For each animal, the phenotypic assignation was performed genotype blinded three times in three independent days. A Pearson's chi-squared test was used to determine whether there is a statistically significant difference between

the expected frequencies and the observed frequencies between the two steatosis phenotype categories and the R990G genotypes.
The open-access image analysis software ImageJ was used to calculate cell diameters and areas in both subcutaneous and visceral fat stains from five to seven animals per each sex and genotype. Five to seven images per animal were analyzed to retrieve most of the tissue's variability. A T-test was used to evaluate differences between genotypes within each line and sex. The frequency distribution of fat cell areas was calculated using bins from 0 to 15,000 μm^2 in 500 increments as reported in Parlee et al. (2014). A Pearson's chi-squared test was used to determine differences between genotypes across all bins after collapsing those with expected frequencies <5 (raw data and statistical analyses are available in File S7). Analyses were performed separating animals with normal liver and hepatic steatosis.

RNA extraction and quantitative PCR

Total RNA was extracted from liver and adipose tissue specimens following the standard protocol of the Higher purity tissue total RNA purification kit (Canvax, Cordoba, Spain). RNA quality and quantity were analyzed using Nanodrop spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). DNase treatment was performed on 300 ng of total RNA for highly expressed genes, and on 1000 ng of total RNA for genes with low expression using DNase I – RNase Free standard protocol (Thermo Fisher Scientific). The Transcription First strand cDNA synthesis kit (Roche Diagnostics) was used to perform the reverse transcription reaction. Quantitative PCR reactions (qPCR) were performed using the iTaq Universal SYBR Green Supermix (Bio-Rad Laboratories). Amplifications were carried out with an initial denaturation step at 95 °C for 3 min, followed by 44 cycles of 95 °C for 10 s and 57 °C for 1 min. Each qPCR reaction was followed by a melting curve analysis to verify specificity.
The expression of nine genes related to inflammation and lipid metabolism was quantified in liver and fat samples using the actin housekeeping gene as internal control (see primer details in Table S5). The relative gene expression for ATGL, HSL, PPARγ, PPARalpha,

FAS, FABP4/aP2, Tnf-alpha, Adiponectin, and LPL was quantified using the $2^{-\Delta\Delta Ct}$ method and given as copies of mRNA per 1,000 copies of actin mRNA. Statistical significance between genotypes was analyzed by the two-sample t-test and ANOVA.

Imputation and association analysis in the Agta population

A total of 138 Agta samples genotyped with the Axiom Genome-Wide Human Origins array (~626,000 SNPs) and for which several anthropological traits have been compiled (Dobon et al., 2022; Migliano et al., 2013) were phased using SHAPEIT2 version v2 (R900) with the duoHMM method to improve the phasing by integrating the known pedigree information. SNPs with missing data were removed and window size was set to 5Mb for phasing. We filtered the dataset to keep only bi-allelic autosomal SNPs with Minor Allele Frequency (MAF) > 5% and without missing data with PLINK 1.9. The dataset was pruned for LD using --indep-pairwise 50 5 0.2, and we performed a principal component analysis (PCA) with EIGENSOFT (version 7.2.1) to identify ancestry outliers and exclude them from the analysis. We also calculated heterozygosity per sample with PLINK and samples with overall increased/decreased heterozygosity rates (±3 SD from the mean of the population) were removed. A total of 131 Agta samples remained after all QC filters (data not shown). Sex and age of the samples were used as covariates. A kinship matrix calculated by KING using identical by descent segment inference was included as random effects. The reference panel for imputation were 81 south-eastern Asian samples with the DSWS phenotype previously sequenced with the 10X technology (unpublished data; M Sikora and J Bertranpetit personal communication) and phased with SHAPEIT. We imputed the whole chromosome 3 of the SNP-array using IMPUTE5 (v1.1.5) with default parameters. The imputed alleles for rs1042636 are G/A (major/minor), with an IMPUTE info score of 0.92. Because we used south-eastern Asian populations as a reference for imputation, the allele G (derived) is annotated here as the major allele, contrary to the global databases where A (ancestral) is the major. For each of the fitted phenotypes and genetic models (additive model: 0 = GG, 1 = GA, 2 = AA; homozygote

only model: 0 = GG, 1 = AA; recessive model: 0 = GG, 1 = GA/ AA; over-dominant model: 0 = GG/AA, 1 = GA), we tested whether there is a difference in means between the genotypes (Kruskal-Wallis or t-test). Per each sex and under the homozygote-only model, a linear mixed effect modeling for quantitative trait loci mapping as available at the Rpackage lme4qtl was used to analyze the relationship between phenotype and genotype incorporating the effect of age and relatedness. The corresponding effect sizes of the genotype were standardized for all phenotypes.

Acknowledgments

We thank the PRBB Animal Facility of the Parc de Recerca Biomèdica de Barcelona and Yolanda Saco from the Servei de Bioquímica Clínica Veterinària at the Universitat Autònoma de Barcelona for their technical help. We also especially thank Patricia Robledo for providing advice and lending us metabolic cages. This work was supported by Agencia Estatal de Investigación (AEI), Fondo Europeo de Desarrollo Regional (FEDER) and Ministerio de Ciencia e Innovación (MICINN) with project grants BFU2016-77961-P and PID2019-110933GB-I00/AEI/10.13039/501100011033; by the Unidad de Excelencia María de Maeztu funded by Ministerio de Ciencia e Innovación and Agencia Estatal de Investigación (DOI: 10.13039/501100011033, Ref: CEX2018-000792-M); and by Direcció General de Recerca, Generalitat de Catalunya (2017SGR00702). BS was supported with an FPI-MINECO PhD fellowship (BES-

References

Atkins, A., Chung, C.-H., Allen, A. G., Dampier, W., Gurrola, T. E., Sariyer, I. K., Nonnemacher, M. R., & Wigdahl, B. (2021). Off-Target Analysis in Gene Editing and Applications for Clinical Translation of CRISPR/Cas9 in HIV-1 Therapy. *Frontiers in Genome Editing*, *3*. https://doi.org/10.3389/fgeed.2021.673022

Auton, A., Abecasis, G. R., Altshuler, D. M., Durbin, R. M., Abecasis, G. R., Bentley, D. R., Chakravarti, A., Clark, A. G., Donnelly, P., Eichler, E. E., Flicek, P., Gabriel, S. B., Gibbs, R. A., Green, E. D., Hurles, M. E., Knoppers, B. M., Korbel, J. O., Lander, E. S., Lee, C., … Abecasis, G. R. (2015). A global reference for human genetic variation. *Nature*, *526*(7571), 68–74. https://doi.org/10.1038/nature15393

Bravo-Sagua, R., Mattar, P., Díaz, X., Lavandero, S., & Cifuentes, M. (2016). Calcium sensing receptor as a novel mediator of adipose tissue dysfunction: Mechanisms and potential clinical implications. *Frontiers in Physiology*, *7*(SEP), 1–7. https://doi.org/10.3389/fphys.2016.00395

Bxown, A. K. (1909). The Religion of the Andaman Islanders. *Folklore*, *20*(3), 257–371. https://doi.org/10.1080/0015587X.1909.9719883

Chang, W., Dvorak, M., & Shoback, D. (2010). Assessing Constitutive Activity of Extracellular Calcium-Sensing Receptors In Vitro and in Bone. *Methods in Enzymology*, 253–266. https://doi.org/10.1016/B978-0-12-381298-8.00013-7

Chang, W., Tu, C., Chen, T. H., Bikle, D., & Shoback, D. (2008). The extracellular calcium-sensing receptor (CaSR) is a critical modulator of skeletal development. *Science Signaling*, *1*(35), 1–14. https://doi.org/10.1126/scisignal.1159945

Coall, D. A., & Chisholm, J. S. (2003). Evolutionary perspectives on pregnancy: maternal age at menarche and infant birth weight. *Social Science & Medicine*, *57*(10), 1771–1781. https://doi.org/10.1016/S0277-9536(03)00022-4

Delaneau, O., Marchini, J., & Zagury, J. F. (2012). A linear complexity phasing method for thousands of genomes. *Nature Methods*, *9*(2), 179–181. https://doi.org/10.1038/nmeth.1785

Dobon, B., Musciotto, F., Mira, A., Greenacre, M., Schlaepfer, R., Aguileta, G., Astete, L. H., Latora, V., Battiston, F., Vinicius, L., Migliano, A. B., & Bertranpetit, J. (2022). The making of the oral microbiome in Agta hunter-gatherers. *BioRxiv 2022.05.03.490437; Doi: Https://Doi.Org/10.1101/2022.05.03.490437*.

Dong, B., Endo, I., Ohnishi, Y., Kondo, T., Hasegawa, T., Amizuka, N., Kiyonari, H., Shioi, G., Abe, M., Fukumoto, S., & Matsumoto, T. (2015). Calcilytic Ameliorates Abnormalities of Mutant Calcium-Sensing Receptor (CaSR) Knock-In Mice Mimicking Autosomal Dominant Hypocalcemia (ADH). *Journal of Bone and Mineral Research*, *30*(11), 1980–1993. https://doi.org/10.1002/jbmr.2551

Drmanac, R., Sparks, A. B., Callow, M. J., Halpern, A. L., Burns, N. L., Kermani, B. G., Carnevali, P., Nazarenko, I., Nilsen, G. B., Yeung, G., Dahl, F., Fernandez, A., Staker, B., Pant, K. P., Baccash, J., Borcherding, A. P., Brownley, A., Cedeno, R., Chen, L., … Reid, C. A. (2010). Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science*, *327*, 78–81. https://doi.org/10.1126/science.1181498

Dvornyk, V., & Waqar-ul-Haq. (2012). Genetics of age at menarche: A systematic review. *Human Reproduction Update*, *18*(2), 198–210. https://doi.org/10.1093/humupd/dmr050

Elsholz, F., Harteneck, C., Muller, W., & Friedland, K. (2014). Calcium - a central regulator of keratinocyte differentiation in health and disease. *European Journal of Dermatology*, *24*(6), 650–661. https://doi.org/10.1684/ejd.2014.2452

Enard, W., Gehre, S., Hammerschmidt, K., Hölter, S. M., Blass, T., Somel, M., Brückner, M. K., Schreiweis, C., Winter, C., Sohr, R., Becker, L., Wiebe, V., Nickel, B., Giger, T., Müller, U., Groszer, M., Adler, T., Aguilar, A., Bolle, I., … Pääbo, S. (2009). A Humanized Version of Foxp2 Affects Cortico-Basal Ganglia Circuits in Mice. *Cell*, *137*(5), 961–971. https://doi.org/10.1016/j.cell.2009.03.041

Fay, J. C., & Wu, C.-I. (2000). Hitchhiking Under Positive Darwinian Selection. *Genetics*, *155*(3), 1405–1413. https://doi.org/10.1093/genetics/155.3.1405

Festuccia, W. T., Laplante, M., Berthiaume, M., Gélinas, Y., & Deshaies, Y. (2006). PPARγ agonism increases rat adipose tissue lipolysis, expression of glyceride lipases, and the response of lipolysis to hormonal control. *Diabetologia*, *49*(10), 2427–2436. https://doi.org/10.1007/s00125-006-0336-y

Gautier, M., Klassmann, A., & Vitalis, R. (2017). rehh 2.0: a reimplementation of the R package <scp>rehh</scp> to detect positive selection from haplotype structure. *Molecular Ecology Resources*, *17*(1), 78–90. https://doi.org/10.1111/1755-0998.12634

Goltzman, D., & Hendy, G. N. (2015). The calcium-sensing receptor in bone-mechanistic and therapeutic insights. *Nature Reviews Endocrinology*, *11*(5), 298–307. https://doi.org/10.1038/nrendo.2015.30

Guha, M., Bankura, B., Ghosh, S., Pattanayak, A. K., Ghosh, S., Pal, D. K., Puri, A., Kundu, A. K., & Das, M. (2015). Polymorphisms in CaSR and CLDN14 Genes Associated with Increased Risk of Kidney Stone Disease in Patients from the Eastern Part of India. *PLOS ONE*, *10*(6), e0130790. https://doi.org/10.1371/journal.pone.0130790

Guo, Y., Shen, H., Xiao, P., Xiong, D.-H., Yang, T.-L., Guo, Y.-F., Long, J.-R., Recker, R. R., & Deng, H.-W. (2006). Genomewide Linkage Scan for Quantitative Trait Loci Underlying Variation in Age at Menarche. *The Journal of Clinical Endocrinology & Metabolism*, *91*(3), 1009–1014. https://doi.org/10.1210/jc.2005-2179

Hannan, F. M., Walls, G. v., Babinsky, V. N., Nesbit, M. A., Kallay, E., Hough, T. A., Fraser, W. D., Cox, R. D., Hu, J., Spiegel, A. M., & Thakker, R. v. (2015). The Calcilytic Agent NPS 2143 Rectifies Hypocalcemia in a Mouse Model With an Activating Calcium-Sensing Receptor (CaSR) Mutation: Relevance to Autosomal Dominant Hypocalcemia Type 1 (ADH1). *Endocrinology*, *156*(9), 3114–3121. https://doi.org/10.1210/en.2015-1269

He, Y.-H., He, Y., Liao, X.-L., Niu, Y.-C., Wang, G., Zhao, C., Wang, L., Tian, M.-J., Li, Y., & Sun, C.-H. (2012). The calcium-sensing receptor promotes adipocyte differentiation and adipogenesis through PPARγ pathway. *Molecular and Cellular Biochemistry*, *361*(1–2), 321–328. https://doi.org/10.1007/s11010-011-1118-5

He, Y.-H., Song, Y., Liao, X.-L., Wang, L., Li, G., Alima, Li, Y., & Sun, C.-H. (2011). The Calcium-Sensing Receptor Affects Fat Accumulation via Effects on Antilipolytic Pathways in Adipose Tissue of Rats Fed Low-Calcium Diets. *The Journal of Nutrition*, *141*(11), 1938–1946. https://doi.org/10.3945/jn.111.141762

Heaney, R. P., Davies, K. M., & Barger-Lux, M. J. (2002). Calcium and Weight: Clinical Studies. *Journal of the American College of Nutrition*, *21*(2), 152S-155S. https://doi.org/10.1080/07315724.2002.10719213

Ho, C., Conner, D. A., Pollak, M. R., Ladd, D. J., Kifor, O., Warren, H. B., Brown, E. M., Seidman, J. G., & Seidman, C. E. (1995). A mouse model of human familial hypocalciuric hypercalcemia and neonatal severe hyperparathyroidism. *Nature Genetics*, *11*(4), 389–394. https://doi.org/10.1038/ng1295-389

Hough, T. A., Bogani, D., Cheeseman, M. T., Favor, J., Nesbit, M. A., Thakker, R. v., & Lyon, M. F. (2004). Activating calcium-sensing receptor mutation in the mouse is associated with cataracts and ectopic calcification. *Proceedings of the National Academy of Sciences*, *101*(37), 13566–13571. https://doi.org/10.1073/pnas.0405516101

Journé, F., Dumon, J.-C., Kheddoumi, N., Fox, J., Laïos, I., Leclercq, G., & Body, J.-J. (2004). Extracellular calcium downregulates estrogen receptor alpha and increases its transcriptional activity through calcium-sensing receptor in breast cancer cells. *Bone*, *35*(2), 479–488. https://doi.org/10.1016/j.bone.2004.03.021

Kamberov, Y. G., Wang, S., Tan, J., Gerbault, P., Wark, A., Tan, L., Yang, Y., Li, S., Tang, K., Chen, H., Powell, A., Itan, Y., Fuller, D., Lohmueller, J., Mao, J., Schachar, A., Paymer, M., Hostetter, E., Byrne, E., … Sabeti, P. C. (2013). Modeling recent human evolution in mice by expression of a selected EDAR variant. *Cell*, *152*(4), 791–702. https://doi.org/10.1016/j.cell.2013.01.016

Kanai, M., Akiyama, M., Takahashi, A., Matoba, N., Momozawa, Y., Ikeda, M., Iwata, N., Ikegawa, S., Hirata, M., Matsuda, K., Kubo, M., Okada, Y., & Kamatani, Y. (2018). Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nature Genetics*, *50*(3), 390–400. https://doi.org/10.1038/s41588-018-0047-6

Kapur, K., Johnson, T., Beckmann, N. D., Sehmi, J., Tanaka, T., Kutalik, Z., Styrkarsdottir, U., Zhang, W., Marek, D., Gudbjartsson, D. F., Milaneschi, Y., Holm, H., DiIorio, A., Waterworth, D., Li, Y., Singleton, A. B., Bjornsdottir, U. S., Sigurdsson, G., Hernandez, D. G., … Bergmann, S. (2010). Genome-Wide Meta-Analysis for Serum Calcium Identifies Significantly Associated SNPs near the Calcium-Sensing Receptor (CASR) Gene. *PLoS Genetics*, *6*(7), e1001035. https://doi.org/10.1371/journal.pgen.1001035

Kristiansen, M. N. B., Veidal, S. S., Rigbolt, K. T. G., Tølbøl, K. S., Roth, J. D., Jelsing, J., Vrang, N., & Feigh, M. (2016). Obese diet-induced mouse models of nonalcoholic steatohepatitis-tracking disease by liver biopsy. *World Journal of Hepatology*, *8*(16), 673. https://doi.org/10.4254/wjh.v8.i16.673

Lassek, W. D., & Gaulin, S. J. C. (2007). Menarche is related to fat distribution. *American Journal of Physical Anthropology*, *133*(4), 1147–1151. https://doi.org/10.1002/ajpa.20644

Li, P., Fan, C., Lu, Y., & Qi, K. (2016). Effects of calcium supplementation on body weight: a meta-analysis. *The American Journal of Clinical Nutrition*, *104*(5), 1263–1273. https://doi.org/10.3945/ajcn.116.136242

Li, T., Guo, W., & Zhou, Z. (2021). Adipose Triglyceride Lipase in Hepatic Physiology and Pathophysiology. *Biomolecules*, *12*(1), 57. https://doi.org/10.3390/biom12010057

Migliano, A. B., Romero, I. G., Metspalu, M., Leavesley, M., Pagani, L., Antao, T., Huang, D.-W., Sherman, B. T., Siddle, K., Scholes, C., Hudjashov, G., Kaitokai, E., Babalu, A., Belatti, M., Cagan, A., Hopkinshaw, B., Shaw, C., Nelis, M., Metspalu, E., … Kivisild, T. (2013). Evolution of the Pygmy Phenotype: Evidence of Positive Selection from Genome-wide Scans in African, Asian, and Melanesian Pygmies. *Human Biology*, *85*(1–3), 251–284. https://doi.org/10.3378/027.085.0313

Migliano, A. B., Vinicius, L., & Lahr, M. M. (2007). Life history trade-offs explain the evolution of human pygmies. *Proceedings of the National Academy of Sciences*, *104*(51), 20216–20219. https://doi.org/10.1073/pnas.0708024105

Miranda, V. P. N., de Faria, F. R., de Faria, E. R., & Priore, S. E. (2014). Somatic maturation and body composition in female healthy adolescents with or without adjustment for body fat. *Revista Paulista de Pediatria : Orgao Oficial Da Sociedade de Pediatria de Sao Paulo*, *32*(1), 78–84. https://doi.org/10.1590/s0103-05822014000100013

Mondal, M., Casals, F., Xu, T., Dall'Olio, G. M., Pybus, M., Netea, M. G., Comas, D., Laayouni, H., Li, Q., Majumder, P. P., & Bertranpetit, J. (2016). Genomic analysis of Andamanese provides insights into ancient human migration into Asia and adaptation. *Nature Genetics*, *48*(9), 1066–1070. https://doi.org/10.1038/ng.3621

Nielsen, T. S., Jessen, N., Jørgensen, J. O. L., Møller, N., & Lund, S. (2014). Dissecting adipose tissue lipolysis: molecular regulation and implications for metabolic disease. *Journal of Molecular Endocrinology*, *52*(3), R199–R222. https://doi.org/10.1530/JME-13-0277

Ohsu, T., Amino, Y., Nagasaki, H., Yamanaka, T., Takeshita, S., Hatanaka, T., Maruyama, Y., Miyamura, N., & Eto, Y. (2010). Involvement of the calcium-sensing receptor in human taste perception. *Journal of Biological Chemistry*, *285*(2), 1016–1022. https://doi.org/10.1074/jbc.M109.029165

Okuno, A., Tamemoto, H., Tobe, K., Ueki, K., Mori, Y., Iwamoto, K., Umesono, K., Akanuma, Y., Fujiwara, T., Horikoshi, H., Yazaki, Y., & Kadowaki, T. (1998). Troglitazone increases the number of small adipocytes without the change of white adipose tissue mass in obese Zucker rats. *Journal of Clinical Investigation*, *101*(6), 1354–1361. https://doi.org/10.1172/JCI1235

Ong, K. T., Mashek, M. T., Bu, S. Y., Greenberg, A. S., & Mashek, D. G. (2011). Adipose triglyceride lipase is a major hepatic lipase that regulates triacylglycerol turnover and fatty acid signaling and

partitioning. *Hepatology*, *53*(1), 116–126. https://doi.org/10.1002/hep.24006
Pan, F., Xiao, P., Guo, Y., Liu, Y.-J., Deng, H.-Y., Recker, R. R., & Deng, H.-W. (2008). Chromosomal regions 22q13 and 3p25 may harbor quantitative trait loci influencing both age at menarche and bone mineral density. *Human Genetics*, *123*(4), 419–427. https://doi.org/10.1007/s00439-008-0490-z
Parlee, S. D., Lentz, S. I., Mori, H., & MacDougald, O. A. (2014). *Quantifying Size and Number of Adipocytes in Adipose Tissue* (pp. 93–122). https://doi.org/10.1016/B978-0-12-411619-1.00006-9
Perry, G. H., & Dominy, N. J. (2009). Evolution of the human pygmy phenotype. *Trends in Ecology and Evolution*, *24*(4), 218–225. https://doi.org/10.1016/j.tree.2008.11.008
Ranieri, M., Tamma, G., di Mise, A., Vezzoli, G., Soldati, L., Svelto, M., & Valenti, G. (2013). Excessive signal transduction of gain-of-function variants of the calcium-sensing receptor (CaSR) are associated with increased ER to cytosol calcium gradient. *PLoS ONE*, *8*(11). https://doi.org/10.1371/journal.pone.0079113
Reid, B. N., Ables, G. P., Otlivanchik, O. A., Schoiswohl, G., Zechner, R., Blaner, W. S., Goldberg, I. J., Schwabe, R. F., Chua, S. C., & Huang, L.-S. (2008). Hepatic Overexpression of Hormone-sensitive Lipase and Adipose Triglyceride Lipase Promotes Fatty Acid Oxidation, Stimulates Direct Release of Free Fatty Acids, and Ameliorates Steatosis. *Journal of Biological Chemistry*, *283*(19), 13087–13099. https://doi.org/10.1074/jbc.M800533200
Reyes, M., Rothe, H. M., Mattar, P., Shapiro, W. B., & Cifuentes, M. (2012). Antilipolytic effect of calcimimetics depends on the allelic variant of calcium-sensing receptor gene polymorphism rs1042636 (Arg990Gly). *European Journal of Human Genetics*, *20*(4), 480–482. https://doi.org/10.1038/ejhg.2011.221
Riccardi, D., Brennan, S. C., & Chang, W. (2013). The extracellular calcium-sensing receptor, CaSR, in fetal development. *Best Practice & Research Clinical Endocrinology & Metabolism*, *27*(3), 443–453. https://doi.org/10.1016/j.beem.2013.02.010
Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L., & Corn, J. E. (2016). Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nature Biotechnology*, *34*(3), 339–344. https://doi.org/10.1038/nbt.3481
Rocha, G., Villalobos, E., Fuentes, C., Villarroel, P., Reyes, M., Díaz, X., Mattar, P., & Cifuentes, M. (2015). Preadipocyte proliferation is elevated by calcium sensing receptor activation. *Molecular and Cellular*

Endocrinology, *412*, 251–256. https://doi.org/10.1016/j.mce.2015.05.011

Ronen, R., Udpa, N., Halperin, E., & Bafna, V. (2013). Learning Natural Selection from the Site Frequency Spectrum. *Genetics*, *195*(1), 181–193. https://doi.org/10.1534/genetics.113.152587

Rosenblum, J. L., Castro, V. M., Moore, C. E., & Kaplan, L. M. (2012). Calcium and vitamin D supplementation is associated with decreased abdominal visceral adipose tissue in overweight and obese adults. *The American Journal of Clinical Nutrition*, *95*(1), 101–108. https://doi.org/10.3945/ajcn.111.019489

Sabeti, P. C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., Xie, X., Byrne, E. H., McCarroll, S. A., Gaudet, R., Schaffner, S. F., Lander, E. S., Frazer, K. A., Ballinger, D. G., Cox, D. R., Hinds, D. A., Stuve, L. L., Gibbs, R. A., Belmont, J. W., … Stewart, J. (2007). Genome-wide detection and characterization of positive selection in human populations. *Nature*, *449*, 913–918. https://doi.org/10.1038/nature06250

Santa Maria, C., Cheng, Z., Li, A., Wang, J., Shoback, D., Tu, C. L., & Chang, W. (2016). Interplay between CaSR and PTH1R signaling in skeletal development and osteoanabolism. *Seminars in Cell and Developmental Biology*, *49*, 11–23. https://doi.org/10.1016/j.semcdb.2015.12.004

Schreiweis, C., Bornschein, U., Burguière, E., Kerimoglu, C., Schreiter, S., Dannemann, M., Goyal, S., Rea, E., French, C. A., Puliyadi, R., Groszer, M., Fisher, S. E., Mundry, R., Winter, C., Hevers, W., Pääbo, S., Enard, W., & Graybiel, A. M. (2014). Humanized Foxp2 accelerates learning by enhancing transitions from declarative to procedural performance. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1414542111

Scillitani, A., Guarnieri, V., de Geronimo, S., Muscarella, L. A., Battista, C., D'Agruma, L., Bertoldo, F., Florio, C., Minisola, S., Hendy, G. N., & Cole, D. E. C. (2004). Blood ionized calcium is associated with clustered polymorphisms in the carboxyl-terminal tail of the calcium-sensing receptor. *Journal of Clinical Endocrinology and Metabolism*, *89*, 5634–5638. https://doi.org/10.1210/jc.2004-0129

Siervogel, R. M., Demerath, E. W., Schubert, C., Remsberg, K. E., Chumlea, W. C., Sun, S., Czerwinski, S. A., & Towne, B. (2003). Puberty and Body Composition. *Hormone Research in Paediatrics*, *60*(Suppl. 1), 36–45. https://doi.org/10.1159/000071224

Spalding, K. L., Arner, E., Westermark, P. O., Bernard, S., Buchholz, B. A., Bergmann, O., Blomqvist, L., Hoffstedt, J., Näslund, E., Britton, T.,

Concha, H., Hassan, M., Rydén, M., Frisén, J., & Arner, P. (2008). Dynamics of fat cell turnover in humans. *Nature*, *453*(7196), 783–787. https://doi.org/10.1038/nature06902

Stock, J. T., & Migliano, A. B. (2009). Stature, Mortality, and Life History among Indigenous Populations of the Andaman Islands, 1871–1986. *Current Anthropology*, *50*(5), 713–725. https://doi.org/10.1086/605429

Sun, C., Wang, L., Yan, J., & Liu, S. (2012). Calcium ameliorates obesity induced by high-fat diet and its potential correlation with p38 MAPK pathway. *Molecular Biology Reports*, *39*(2), 1755–1763. https://doi.org/10.1007/s11033-011-0916-x

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*(3), 585–595. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1203831&tool=pmcentrez&rendertype=abstract

Thakker, R. v., Richard Bringhurst, F., & Jüppner, H. (2015). Regulation of Calcium Homeostasis and Genetic Disorders that Affect Calcium Metabolism**Chapter titles shaded in green indicate chapters dedicated predominantly to pediatric endocrinology content. In *Endocrinology: Adult and Pediatric*. https://doi.org/10.1016/b978-0-323-18907-1.00061-5

Toka, H. R., Al-Romaih, K., Koshy, J. M., DiBartolo, S., Kos, C. H., Quinn, S. J., Curhan, G. C., Mount, D. B., Brown, E. M., & Pollak, M. R. (2012). Deficiency of the Calcium-Sensing Receptor in the Kidney Causes Parathyroid Hormone–Independent Hypocalciuria. *Journal of the American Society of Nephrology*, *23*(11), 1879–1890. https://doi.org/10.1681/ASN.2012030323

Tu, C. L., Crumrine, D. A., Man, M. Q., Chang, W., Elalieh, H., You, M., Elias, P. M., & Bikle, D. D. (2012). Ablation of the calcium-sensing receptor in keratinocytes impairs epidermal differentiation and barrier function. *Journal of Investigative Dermatology*, *132*(10), 2350–2359. https://doi.org/10.1038/jid.2012.159

Turksen, K., & Troy, T. C. (2003). Overexpression of the calcium sensing receptor accelerates epidermal differentiation and permeability barrier formation in vivo. *Mechanisms of Development*, *120*(6), 733–744. https://doi.org/10.1016/S0925-4773(03)00045-5

Venkateswar, S. (2004). *Development and Ethnocide: Colonial Practices in the Andaman Islands* (Christian Erni & Sille Stidsen, Eds.). IWGIA.

Verdu, P., Austerlitz, F., Estoup, A., Vitalis, R., Georges, M., Théry, S., Froment, A., le Bomin, S., Gessain, A., Hombert, J. M., van der Veen, L., Quintana-Murci, L., Bahuchet, S., & Heyer, E. (2009). Origins and

Genetic Diversity of Pygmy. *Current Biology*, *19*(4), 312–318. https://doi.org/10.1016/j.cub.2008.12.049

Verghese, P. B., Arrese, E. L., & Soulages, J. L. (2007). Stimulation of lipolysis enhances the rate of cholesterol efflux to HDL in adipocytes. *Molecular and Cellular Biochemistry*, *302*(1–2), 241–248. https://doi.org/10.1007/s11010-007-9447-0

Vezzoli, G., Terranegra, A., Arcidiacono, T., Biasion, R., Coviello, D., Syren, M. L., Paloschi, V., Giannini, S., Mignogna, G., Rubinacci, A., Ferraretto, A., Cusi, D., Bianchi, G., & Soldati, L. (2007). R990G polymorphism of calcium-sensing receptor does produce a gain-of-function and predispose to primary hypercalciuria. *Kidney International*, *71*, 1155–1162. https://doi.org/10.1038/sj.ki.5002156

Villarroel, P., Reyes, M., Fuentes, C., Segovia, M. P., Tobar, N., Villalobos, E., Martínez, J., Hugo, E., Ben-Jonathan, N., & Cifuentes, M. (2013). Adipogenic effect of calcium sensing receptor activation. *Molecular and Cellular Biochemistry*, *384*(1–2), 139–145. https://doi.org/10.1007/s11010-013-1791-7

Walsh, S., Pagani, L., Xue, Y., Laayouni, H., Tyler-Smith, C., & Bertranpetit, J. (2020). Positive selection in admixed populations from Ethiopia. *BMC Genetics*, *21*(S1), 108. https://doi.org/10.1186/s12863-020-00908-5

Wang, K., Li, M., & Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research*, *38*(16), e164. https://doi.org/10.1093/nar/gkq603

Wang, Y., Bikle, D. D., & Chang, W. (2013). Autocrine and Paracrine Actions of IGF-I Signaling in Skeletal Development. *Bone Research*, *1*(3), 249–259. https://doi.org/10.4248/BR201303003

Wohlers, L. M., & Spangenburg, E. E. (2010). 17β-estradiol supplementation attenuates ovariectomy-induced increases in ATGL signaling and reduced perilipin expression in visceral adipose tissue. *Journal of Cellular Biochemistry*, n/a-n/a. https://doi.org/10.1002/jcb.22553

Xu, J., Kulkarni, S. R., Li, L., & Slitt, A. L. (2012). UDP-Glucuronosyltransferase Expression in Mouse Liver Is Increased in Obesity- and Fasting-Induced Steatosis. *Drug Metabolism and Disposition*, *40*(2), 259–266. https://doi.org/10.1124/dmd.111.039925

Zemel, M. B. (2002). Regulation of Adiposity and Obesity Risk By Dietary Calcium: Mechanisms and Implications. *Journal of the American College of Nutrition*, *21*(2), 146S-151S. https://doi.org/10.1080/07315724.2002.10719212

Zhang, F., Ye, J., Meng, Y., Ai, W., Su, H., Zheng, J., Liu, F., Zhu, X., Wang, L., Gao, P., Shu, G., Jiang, Q., & Wang, S. (2018). Calcium Supplementation Enhanced Adipogenesis and Improved Glucose Homeostasis Through Activation of Camkii and PI3K/Akt Signaling Pathway in Porcine Bone Marrow Mesenchymal Stem Cells (pBMSCs) and Mice Fed High Fat Diet (HFD). *Cellular Physiology and Biochemistry*, *51*(1), 154–172. https://doi.org/10.1159/000495171

Zhang, F., Ye, J., Zhu, X., Wang, L., Gao, P., Shu, G., Jiang, Q., & Wang, S. (2019). Anti-Obesity Effects of Dietary Calcium: The Evidence and Possible Mechanisms. *International Journal of Molecular Sciences*, *20*(12), 3072. https://doi.org/10.3390/ijms20123072

Zuo, X., Sun, L., Yin, X., Gao, J., Sheng, Y., Xu, J., Zhang, J., He, C., Qiu, Y., Wen, G., Tian, H., Zheng, X., Liu, S., Wang, W., Li, W., Cheng, Y., Liu, L., Chang, Y., Wang, Z., … Zhang, X. (2015). Whole-exome SNP array identifies 15 new susceptibility loci for psoriasis. *Nature Communications*, *6*, 1–7. https://doi.org/10.1038/ncomms7793

Tables and figures

Table 1. Differential serum biochemistries between CaSR R990G genotypes. **104**

Figure 1. Signals of positive selection in the *CASR* gene region and worldwide frequencies for the R990G substitution. **105**

Figure 2. Body weight growth curves, body mass index and body composition for CaSR R990G knock-in mice (line 7). **107**

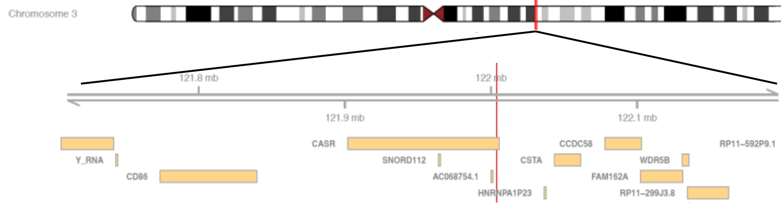
Figure 3. Differential adipose tissue features between CaSR R990G genotypes. **108**

Supplementary Material **Annex**

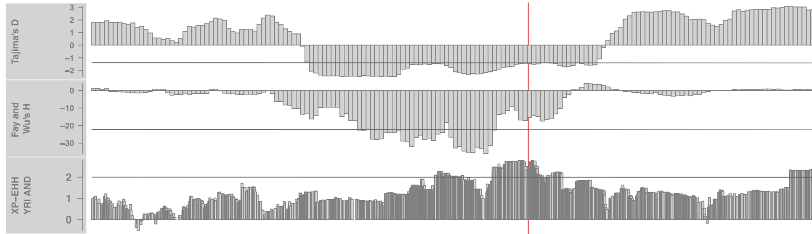
Table 1. Differential serum biochemistries between CaSR R990G genotypes. For each group and metabolite, mean concentrations per genotype are given with the standard deviation. Significance between genotypes was evaluated by means of a factorial ANOVA (considering genotype, sex, and line when applicable) or a one-way ANOVA when just evaluating differences among genotypes within a given sex in a line. HO, homozygote carriers of the derived allele (990G); WT, homozygote carriers of the ancestral allele (R990).

	Both sexes			Females			Males		
Both lines	**HO (N = 64)**	**WT (N = 56)**		**HO (N = 33)**	**WT (N = 24)**		**HO (N = 31)**	**WT (N = 32)**	
Calcium (mg/dL)	8.914 ± 0.268	9.015 ± 0.261	p = 0.060	8.950 ± 0.237	8.955 ± 0.214	p = 0.942	8.876 ± 0.297	9.060 ± 0.286	**p = 0.017**
Phosphorus (mg/dL)	6.407 ± 1.333	6.015 ±1.374	p = 0.141	6.169 ± 1.276	6.428 ± 1.551	p = 0.542	6.660 ± 1.366	5.705 ± 1.155	**p = 0.004**
Cholesterol (mg/dL)	99.034 ± 20.054	105.027 ± 20.379	p = 0.147	84.652 ± 11.603	92.829 ± 17.625	p = 0.051	114.345 ± 15.204	114.175 ± 17.459	p = 0.906
HDL (mmol/L)	1.736 ± 0.425	1.921 ± 0.363	**p = 0.020**	1.489 ± 0.234	1.736 ± 0.352	**p = 0.003**	2.000 ± 0.426	2.060 ± 0.309	p = 0.575
LDL (mmol/L)	0.307 ± 0.055	0.346 ± 0.076	**p = 0.001**	0.293 ± 0.052	0.337 ± 0.084	**p = 0.018**	0.323 ± 0.056	0.354 ± 0.070	**p = 0.031**
Line 7	**HO (N = 39)**	**WT (N = 30)**		**HO (N = 21)**	**WT (N = 12)**		**HO (N = 18)**	**WT (N = 18)**	
Calcium (mg/dL)	8.917 ± 0.231	9.011 ± 0.260	p = 0.162	8.947 ± 8.940	8.940 ± 0.193	p = 0.916	8.882 ± 0.292	9.058 ± 0.292	p = 0.080
Phosphorus (mg/dL)	6.335 ± 1.211	6.079 ± 1.065	p = 0.457	6.112 ± 1.128	6.474 ± 1.014	p = 0.371	6.589 ± 1.286	5.815 ± 1.042	p = 0.055
Cholesterol (mg/dL)	97.703 ± 21.891	101.067 ± 21.152	p = 0.955	83.857 ± 13.335	89.258 ± 20.170	p = 0.361	113.856 ± 18.679	108.939 ± 18.329	p = 0.431
HDL (mmol/L)	1.675 ± 0.455	1.844 ± 0.391	p = 0.234	1.462 ± 0.241	1.598 ± 0.318	p = 0.178	1.924 ± 0.523	2.009 ± 0.351	p = 0.571
LDL (mmol/L)	0.309 ± 0.059	0.323 ± 0.057	p = 0.405	0.295 ± 0.061	0.317 ± 0.066	p = 0.353	0.325 ± 0.052	0.328 ± 0.052	p = 0.889
Line 15	**HO (N = 25)**	**WT (N = 26)**		**HO (N = 12)**	**WT (N = 12)**		**HO (N = 13)**	**WT (N = 14)**	
Calcium (mg/dL)	8.910 ± 0.324	9.021 ± 0.267	p = 0.214	8.955 ± 0.340	8.971 ± 0.240	p = 0.896	8.869 ± 0.316	9.064 ± 0.289	p = 0.106
Phosphorus (mg/dL)	6.519 ± 1.523	5.940 ± 1.682	p = 0.238	6.262 ± 1.552	6.381 ± 1.999	p = 0.872	6.757 ± 1.517	5.563 ± 1.313	**p = 0.038**
Cholesterol (mg/dL)	101.112 ± 17.017	109.596 ± 18.819	**p = 0.019**	86.042 ± 8.064	96.400 ± 14.658	**p = 0.043**	115.023 ± 9.099	120.907 ± 14.169	p = 0.215
HDL (mmol/L)	1.831 ± 0.360	2.010 ± 0.313	**p = 0.016**	1.534 ± 0.224	1.874 ± 0.342	**p = 0.009**	2.105 ± 0.210	2.126 ± 0.241	p = 0.811
LDL (mmol/L)	0.305 ± 0.051	0.373 ± 0.087	**p = 0.002**	0.288 ± 0.031	0.357 ± 0.098	**p = 0.031**	0.321 ± 0.062	0.387 ± 0.078	**p = 0.022**

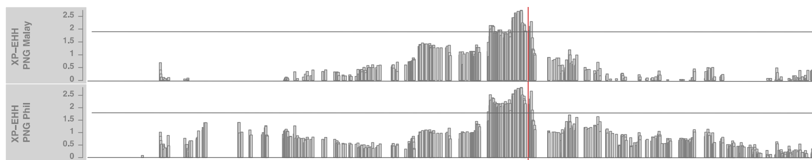
A



B



C



D

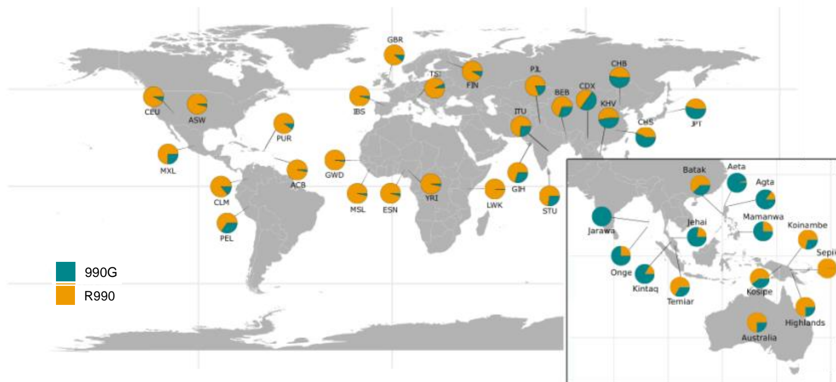


Figure 1. Signals of positive selection in the *CASR* gene region and worldwide frequencies for the R990G substitution. (A) On the top, the location of the *CASR* gene region on chromosome 3. Below, zoom-in with Ensembl gene tracks comprising 200 kb upstream and downstream of the *CASR* gene. (B) Tests of positive selection in the Andamanese (Tajima's D, Fay, and Wu's H, and XP-EHH comparing Andamanese versus Yoruba). (C) Tests of positive selection in Malaysian and Philippine hunter-gather groups presenting the DSWS phenotype (XP-EHH comparing each south-east Asian population group versus Papua New Guinea). For each population and selection statistic, the horizontal bar indicates the top or bottom 5th percentile of their respective genome-wide distribution (z-scores are available in File S1). The position of the CaSR R990G polymorphism is represented by the red vertical line across the panels. (D) Worldwide allele frequencies of the CaSR R990G polymorphism. Each pie graph represents the allele frequency in one population. Populations names from the 1000G project Phase 3 are abbreviated as follows: CHB, Han Chinese in Beijing; JPT, Japanese in Tokyo; CHS, Southern Han Chinese; CDX, Chinese Dai in Xishuangbanna; KHV, Kinh in Ho Chi Minh City; CEU, Utah Residents (CEPH) with Northern and Western European Ancestry; TSI, Toscani in Italia; FIN, Finnish in Finland; GBR, British in England and Scotland; IBS, Iberian Population in Spain; YRI, Yoruba in Ibadan, Nigeria; LWK, Luhya in Webuye, Kenya; GWD, Gambian in Western Divisions in the Gambia; MSL, Mende in Sierra Leone; ESN, Esan in Nigeria; ASW, Americans of African Ancestry in SW USA; ACB, African Caribbeans in Barbados; MXL, Mexican Ancestry from Los Angeles USA; PUR, Puerto Ricans from Puerto Rico; CLM, Colombians from Medellin, Colombia; PEL, Peruvians from Lima, Peru; GIH, Gujarati Indian from Houston, Texas; PJL, Punjabi from Lahore, Pakistan, BEB, Bengali from Bangladesh; STU, Sri Lankan Tamil from the UK; and ITU, Indian Telugu from the UK.

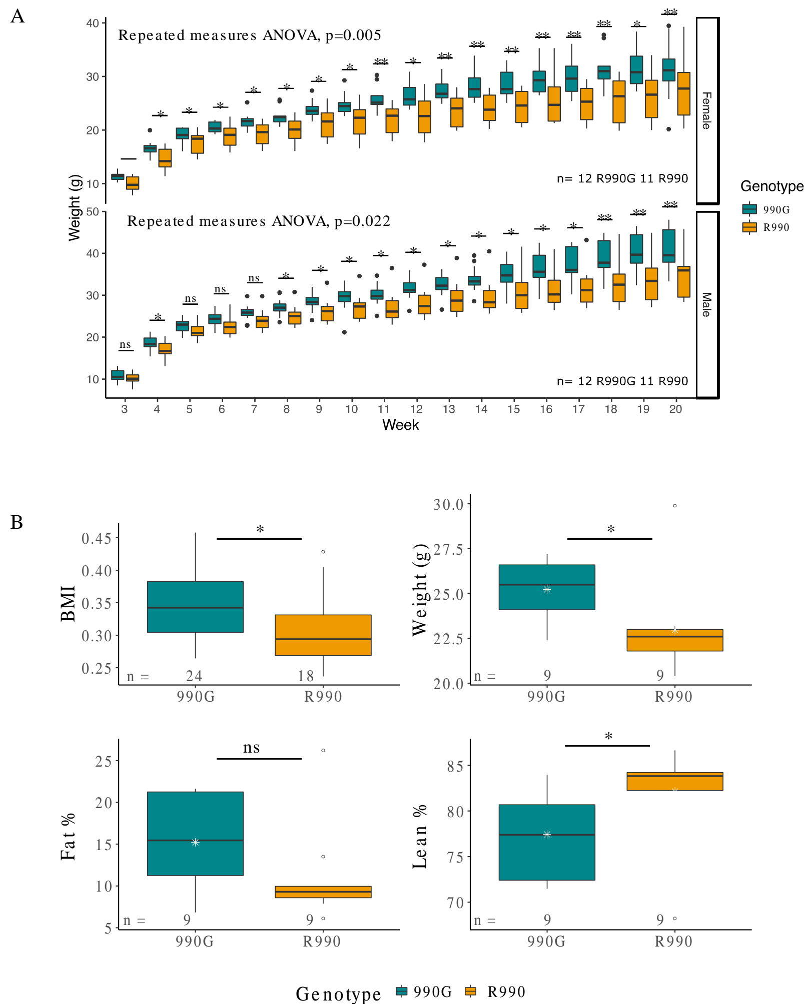
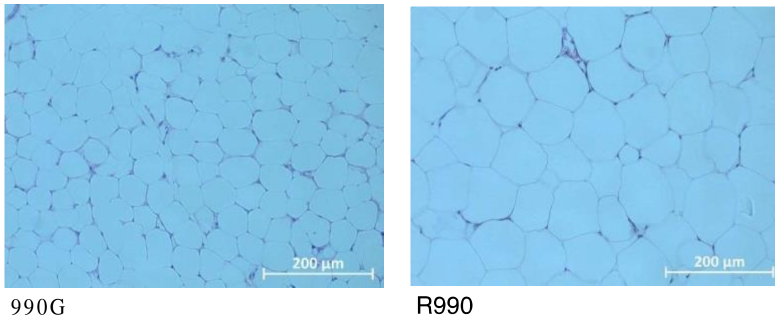
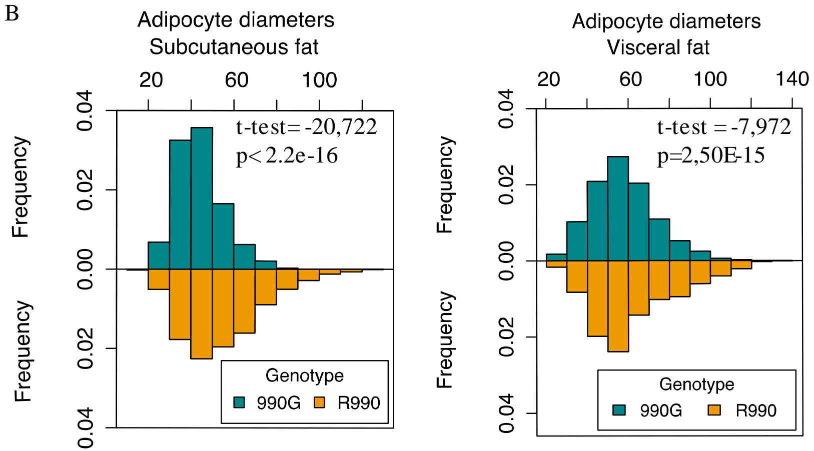


Figure 2. Body weight growth curves, body mass index, and body composition for CaSR R990G knock-in mice (line 7). (A) Boxplots showing the differences in mice body weight from 3 to 20 weeks of age between the homozygote carriers of the ancestral and derived alleles of the CaSR R990G polymorphism. *$p < 0.05$; **$p < 0.005$; ns, not significance (one-way ANOVA). (B) Body Mass Index (BMI) at the final point (20-26 weeks of age), and weight and body weight composition differences between genotypes in females as determined by EchoMRI analysis on week 14 of age. *$p < 0.05$; ns, not significance (Mann-Whitney U test). For all boxplots, the line in the middle of each box represents the median for each group examined, and the edges of the box represent the first and third quartiles. Filled circles represent outliers (values greater than 1.5 and less than 3 interquartile ranges from the edge of the box). Detailed significance levels for all tests are available in Files S2 and S3.

A



B



C

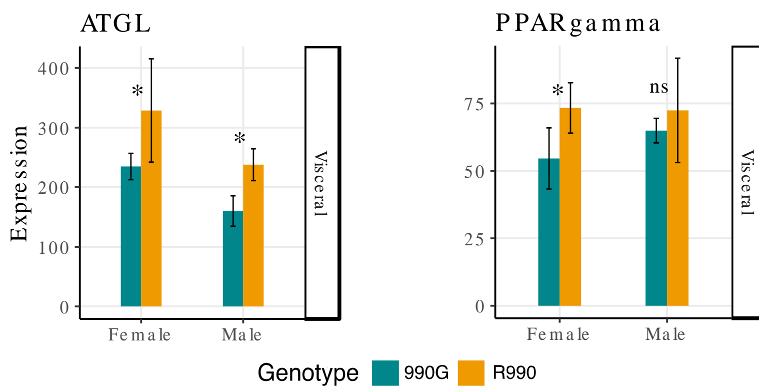


Figure 3. **Differential adipose tissue features between CaSR R990G genotypes**. (A) Visceral fat histology differences between genotypes in hematoxylin and eosin staining. On the right, the histology of a 990G female homozygote (22 weeks). On the left, the histology of a 990R female homozygote (23 weeks). (B) Adipocyte diameters in subcutaneous and visceral fat. Differences between genotypes were tested through a t-test (N= 11 females from line 7, including five 990G homozygotes and six R990 homozygotes, respectively). (C) Differential ATGL and PPARγ gene expression between genotypes in visceral fat. *$p < 0.05$; ns, non-significant. Relative expression is represented per 1,000 copies of Actin.

DISCUSSION

5. Implication of CaSR in the systemic lipidic handling:

5.1. The role of R990G substitution in adipocyte differentiation and lipolysis

Bravo-Sagua et al. (2016) reported the Calcium sensing receptor to be involved in regulating white adipose tissue lipid handling. Furthermore, He et al. (2011) demonstrated that rats fed with a low [Ca2^{+}] diet showed a significant fat accumulation caused by the enhanced expression of the CaSR gene, while lower protein levels of the hormone-sensitive lipase (HSL) and of the adipose triglyceride lipase (ATGL) were detected, supporting the evidence for an antilipolytic effect resulting from the calcium-sensing receptor activation. Additionally, the human polymorphism rs1042636 causing the R990G substitution in CaSR was also described to be associated with a reduction of 20% of the lipolysis in omental adipose tissue (Reyes et al., 2012). Our results confirm that the CaSR activation mediated by the 990G allele is associated with an impairment of lipid metabolism in white adipose tissue. We also detected downregulation of PPARγ in 990G females which is known to be involved in the regulation of lipogenesis in mature adipocytes (Kersten, 2001). Furthermore, Joosen et al. (2006), revealed that PPARγ primarily regulates FABP4/aP2, whose expression was shown to be a function of PPARγ activity. In agreement with Joosen et al. (2006), our results showed that 990G animals had the tendency to express less FABP4/aP2 mRNA than their R990 littermates, mirroring the downregulation of the PPARγ. Moreover, FABP4/aP2 expression was previously described to be important for HSL and ATGL expression in mature adipocytes through fatty acid ligation (Dou et al., 2020; Krušinová & Pelikánová, 2008; H. L. Li et al., 2021). Mice knocked out for adipose FABP4/aP2 showed, in fact, a significant reduction of HSL and ATGL, an increase in fat mass depots, and an enhanced LPL expression, indicating that FABP deficient mice had lower lipolysis but higher lipogenesis (Hertzel et al., 2006). In line with the literature, our results showed a lower expression of ATGL and HSL genes, without, however, any evidence for an

increase in the expression of lipogenic genes (LPL and FAS). In turn, it has also been shown that the interaction between FABP4/aP2 and free fatty acids (FFA) is required for the relocation of the complex FABP/FFA to the nucleus as well as for the activation of PPARγ itself (Krušinová & Pelikánová, 2008). In other words, FABP4/aP2 in adipose tissue is involved in the regulation of the lipolysis on one hand by modulating downstream genes (HSL and ATGL) and on the other hand regulating its own transcription through PPARγ activation in a feedback loop (H. L. Li et al., 2021).

PPARγ was also described to be downregulated by estrogen treatment in human abdominal adipose tissue (Lundholm et al., 2008). Jeong & Yoon (2011) treated mice with an activator of PPARγ (troglitazone TZN) together with estrogen, showing a reduction in body weight and adipocyte size. Moreover, a reduction in the expression of PPARγ target genes was detected. In our results, the fact that the reduction of PPARγ expression reached significance only in females could be related to the synergic activity of the estrogen with the CaSR. As a consequence of the PPARγ downregulation together with the CaSR activation, a lower expression of downstream lipolysis-related genes (FABP4/aP2, ATGL, and HSL) is generated in mature adipocytes. However, the sole lipolysis reduction in 990G mature adipocytes does not fully explain the higher weight and the smaller adipocyte diameters we found in 990G animals. To understand the mechanism behind this phenotype CaSR's role in the adipocytes' maturation must be considered.

As reported in chapter 3.3, He et al. (2012) showed that the activation of CaSR in adipose cell cultures upregulated PPARγ expression. Furthermore, it was shown that CaSR activation induced by its allosteric modulator CdCl3 enhanced the preadipocytes LS14's proliferation (Rocha et al., 2015). PPARγ is in fact considered to be the master activator of adipogenesis in preadipocytes where it promotes the expression of a large gene group that induces the cellular differentiation towards the adipocyte phenotype (Tang & Lane, 2012). In the study conducted by Okuno et al. (1998), the use of TZN in obese rats resulted in the increase of small adipose cells and the decrease of large

adipocytes without affecting the total mass of adipose tissue. Besides, the activation of PPARγ using pioglitazone in wild-type mice resulted in a significant decrease in adipocyte sizes and an increase in adipocyte number, as a consequence of adipocyte differentiation (Kubota et al., 1999).

Gene expression analysis by quantitative PCR in visceral fat from animals developing steatosis revealed a different expression profile from those animals not developing steatosis. The hormone-sensitive lipase was upregulated in 990G females (p-value <0,05), while 990G males showed downregulation of both HSL and LPL (p-value < 0,05). Since these differences were observed only in animals showing an impaired fat accumulation[10], we could hypothesize that the presence of an hyperlipidic environment could upregulate the lipolytic response in female visceral fat. However, the antilipolytic effect of CaSR activation in visceral fat was confirmed in males (HSL downregulation in males p-value<0,05). One reasonable explanation for detecting differences between sexes could be seen in the role of estrogen in regulating female adipogenesis and lipolysis (Journé et al., 2004; Wohlers & Spangenburg, 2010).

5.2. The role of R990G substitution in liver

When analyzing gene expression in non-steatotic livers we detected higher expression of ATGL in 990G females (p<0,05) and the same tendency in males, while no expression of TNFalpha or IL6 was found. The ATGL overexpression was, in fact, previously described to prevent the development of steatosis and hepatic inflammation, increase PPARalpha expression and promote fatty acid oxidation (T. Li et al., 2022; Ong et al., 2011; Reid et al., 2008). At the same time, when comparing 990G females having a normal liver with 990G females with steatosis in the CaSR7 line, we found a significant downregulation in ATGL expression in the steatosis animals (t = 5.78, df = 4.03, p-value = >0.01). On the other hand, no differences in the liver ATGL

[10] Both genotypes of CaSR15 females showed high weight, while only 990G females showed lower FFA concentration in blood.

expression were found between 990G and R990 homozygotes in CaSR15 animals with steatosis. In this cohort, instead, we found a significant decrease in FABP4/aP2 expression in 990G males when compared with their R990 homozygote littermates while in females we did not find any difference. As previously reported, FABP4/aP2 hepatic expression levels are sex and age-specific, suggesting the existence of some hormone-related compensatory mechanism (Atshaves et al., 2010).

5.3. Connecting the dots: the R990G substitution and its systemic implication

Considering all our results together, we can argue that the CaSR activation in 990G young animals' visceral fat enhances the preadipocyte proliferation probably through PPARγ upregulation, leading to the maturation of a higher number of small adipocytes as we detected in the histology's morphometric analysis. As a result, 990G homozygotes developed higher weight and BMI. However, in mature adipocytes from adult animals, CaSR activation confirmed its antilipolytic role in downregulating lipolysis-related genes. Furthermore, 990G adult females presented, in the liver histology, a higher tendency to develop steatosis, suggesting an enhanced fat uptake. On the other hand, animals without steatosis seemed to compensate for the lipidic imbalance in the liver with enhanced lipolysis. All these results together suggest the presence of a generalized imbalance in the lipidic handling with the tendency towards increased fat storage in different tissues: in visceral fat through the enhanced cell proliferation and in liver developing steatosis.

6. Putative adaptive forces for an enhanced lipidic accumulation in populations with the DSWS phenotype

Lipid accumulation, reproduction, and lifespan have been described to be causally linked both in different animal models and in humans through a reproductive-endocrine signaling axis (Hansen et al., 2013). During puberty the human body undergoes major morphological changes in body fat and lean, the growth rate is accelerated and there is the closure of the epiphysis' growth plate (Miranda et al., 2014). Anthropometric analysis in humans, in fact, revealed that fat percentage increases during puberty leading to an increase in leptin secretion. Leptin is a hormone primarily secreted by the adipose tissue that regulates the energy balance, and during the 6 months preceding menarche, it increases its levels by 28% as a consequence of the enhanced fat accumulation (Lassek & Gaulin, 2007; Siervogel et al., 2003). Specifically, Lassek & Gaulin (2007) found that the gluteo-femoral fat adipose tissue produced more leptin than upper body fat and that the age of menarche was inversely correlated to the leptin levels. In other words, during puberty, girls have an enhanced fat accumulation in the central region of the body, an increase in leptin production that regulates the menarche inducing the pubertal growth spurt and the growth plate fusion in the epiphysis.

Diachronic cross-sectional studies on anthropometrical data collected during the first colonial period in the Andaman Islands revealed a complex relationship between mortality, reproduction, and body height (Stock & Migliano, 2009). In this study, it was shown that during the most difficult periods of higher mortality and epidemics, the individuals' height was significantly reduced both in men and women. Furthermore, when analyzing the stature, growth, and fitness of the Philippine, Malaysian, and central African populations presenting the DSWS phenotype, it was reported that these groups showed an early onset reproduction in response to a hostile environment when compared to other hunter-gatherer groups (Migliano et al., 2007). Moreover, as described by Coall & Chisholm (2003), a hostile

environment during childhood generated environmental (mental) stress that was positively associated with early menarche.

Considering all the evidence described above, we can fairly hypothesize that the hostile environment in the rainforest acted as a selective pressure on DSWS populations, generating a positive selection on traits that induced the early onset of sexual maturation to increase reproductive success. In this light, the calcium-sensing receptor R990G substitution could have been adaptive by enhancing the adipose tissue accumulation, which specifically in the hips increased the leptin levels anticipating the menarche in women. Consequently, the early onset of sexual maturation caused the early closure of the growth plate resulting in the reduction of the body size.

As previously mentioned, the early onset of sexual maturation related to an enhanced fat accumulation (Lassek & Gaulin, 2007) and generated by a stressful environment (Coall & Chisholm, 2003) was described in populations from central Europe and the Americas suggesting that this is a plastic mechanism present in all humans that allows for a faster reproduction during critical times. However, our results suggest that this mechanism was further positively selected in nowadays populations from SEA carrying the DSWS phenotype as an adaptive response to their particularly stressful rainforest environment.

6.1. Carrying a specific adaptive variant in a western-centric world

This work shows that studying human diversity and adaptation allows us to have a complete understanding of how evolutionary forces can shape human biology and appreciate how natural selection can induce drastic phenotypic changes. The experimental validation of a putative adaptive variant into an animal model resulted crucial in understanding how a non-synonymous substitution behaves in a physiological environment, allowing us to go beyond a simple *in silico* prediction. Our results demonstrate that the R990G substitution in the *CASR* gene

produces an impaired lipidic accumulation that could have been positively selected in DSWS populations to increase their energetic storage in the form of fat depots. Furthermore, we hypothesize that the enhanced fat accumulation could be influencing the bases of an anticipated sexual maturation and diminished physical growth as well as the development of the steatopygic phenotype. It is worth noticing that in the last decades several studies have highlighted the increasing tendency for developing overweight, obesity, and diabetes in Indigenous peoples from South East Asia and the Andaman Islands living in urban areas and experiencing a dietary shift toward consuming higher caloric food. Specifically, the vulnerability of suffering from overweight was detected to be higher in the age group 20-39 in men, while women of every age group developed overweight. The authors suggested that the cultural shift toward a more sedentary life in governmental homes (as previously reported in chapter 2.1) together with the reduction of proteins and the increase in carbohydrates intake could be the causal factor for the high obesity risk (Chowdhury & Chakraborty, 2017; Sahani et al., 2010; Sharma, 2016; Vennu et al., 2019). In this light, it is possible that the higher 990G-CaSR activity could today result maladaptive within the new dietary environment being the Calcium-sensing receptor involved in lipid handling as well as playing a metabolically active role in the glucose/insulin balance (Babinsky et al., 2017; Oh et al., 2016; Rybczyńska et al., 2017; Squires et al., 2000).

7. How can we talk about genetic differences between populations in a time of major historical changes?

7.1. The never-ending story of the unknown's fear

This thesis has been written during the most inexplicable human invention: a new war. Once again, we found ourselves in a situation where our daily routine could change in a heartbeat. Even during the two years of the still ongoing COVID pandemic, we managed to keep our priorities in line, achieving whatever goals we previously planned and keeping our schedule full. But, on February 24th we woke up in a world where some young people who the day before had been planning to go to work, study, or travel, were embracing weapons to defend their homes in an anachronistic fashion. For decades we had lived through many wars, and we were more or less aware of what was happening on the other side of the sea. Despite that, we went on with our lives, with our plans and dreams. The threat only became tangible when the people in danger were the ones that we recognized as our equals[11], while on the other hand considering threatening what is different from us. The stranger is an enemy[12] that we cannot understand, we are different, we do not speak the same language, we do not have the same traditions, we do not have the same colors, we do not like the same things, we do not like the same food, we do not have the same blood, we just don't. It is fascinating our surgical ability to search for all the things that make us (*self*) different from others (*non-self*). As if the differences were the only things that legitimate our own group and that make us feel as if we belong.

---

[11] The press could not help but declare that "This [Ukraine] is not a place, with all due respect, like Iraq or Afghanistan who has seen conflict rage for decades. This is a relatively civilized, relatively European - I have to choose those words carefully - city where you would not expect that or hope that it was going to happen." Charlie D'Agata, CBS news; and also "what's compelling is just looking at them, the way they're dressed. These are prosperous middle-class people, these are not obviously refugees trying to get away from areas in the Middle East [or] North Africa, they look like any European family you would live next door to", Peter Dobbie, Al Jazeera.

[12] The Romans called their enemies with the word *hostis* which also indicated foreign people.

7.2. Searching for non-self to find the self: an archaic instinct

A great example of the antithesis' rhetoric between the *self* and the *non-self* can be found in the *De Bello Gallico* where Julius Caesar celebrated his victories and the Roman superiority by playing with the description of his enemies to his own advantage: sometimes describing them as fierce and heroic, while other times as ignorant and barbaric (Siedler, 1956). As previously mentioned (chapter 2.1), the same kind of dichotomic rhetoric was used during the Victorian age for describing the colonized people to legitimize the European campaigns of the slave trade and land exploitation (Opubor, 2020). However, this dualistic approach in defining the *self* in contraposition to the *non-self* appears to be something so ancestral that could be even identified in how populations call themselves in some hunter-gatherer groups. In fact, the self-assigned names *Onge* and *Khoekhoe* simply mean "human beings", with the only purpose to separate the humans from the non-humans (Pauls, 2021; Venkateswar, 2004). On the contrary, some populations having their name assigned from an external group often are defined by their belonging to the *non-self*. For instance, the word *Jarawa* in the Aka-bea-da language[13] means "stranger" (Venkateswar, 2004), as well as the words *Agta, Ayta/Aeta* in the Philippines languages, mean "outsiders" (R. Blust, 2013).

7.3. The unexpected and underestimated impact of science on society

Since the inclination in interpreting the *non-self* as a threat seems to be deeply rooted in humans, when talking about the differences among human populations from a genetic perspective it is particularly important to contextualize and put the genetic differences in perspective. It is necessary to acknowledge that the level of genetic diversity among human populations is estimated to be less than 10% and that genetic variation within continents is clinal (Lee et al., 2008; Mallick et al., 2016), therefore the strong phenotypic impact of

[13] Language spoken by one group of the Great Andamanese.

particular variants cannot justify any kind of marginalization, racial essentialism, or biological hierarchy (Panofsky et al., 2021). However, we must be aware that even though genetics has clearly demonstrated that human populations cannot be subdivided into biologically defined groups, racialized[14] people are still a reality, and the 18th century's legacy of stereotypes and discrimination is still present (Rutherford, 2021). Indeed, the misinterpretation and manipulation of the genetic ancestry tests' results have been used by different groups of White nationalists in the last years to undermine the scientific evidence and reinforce racial realism (Panofsky et al., 2021). Sadly, also in the academic world, a small group of scientists is still attached to the old-fashioned eugenics ideas that were so popular in the fascist propaganda. They are organized in small circles, publishing their biased ideas in small open-source journals spreading false racist, sexist and eugenics ideas such as the so-called "Woodley effect"[15] (Canlorbe, 2019b, 2019a; Meisenberg, 2009; Saini, 2020). Furthermore, scientific data are often quickly politicized, especially when talking about human populations (Lee et al., 2008). An example of what this can cause is the recent "Buffalo shooting" that took place in a supermarket in a predominantly Black community on May 14th, 2022 taking the life of 10 people. The killer described himself as a fascist, White supremacist, and anti-Semite supporting the "Great Replacement" conspiracy theory, and some days before the shooting published a "manifesto" where he cited numerous scientific publications to support his biased ideas on the one hand by cherry-picking and decontextualizing some scientific results, and on the other hand, relying on Mr. Woodley's independently founded and quite questionable research (Prokupecz et al., 2022; Pronczuk & Ryckewaert, 2022).

However, even if we are not biased by the unmissed resurrected nationalism, we still must take into serious consideration the 18th-century legacy of stereotypes that is widespread in our society. One example is the misrepresentation of the hunter-gatherer people as some magical, mysterious, and fascinating creatures that we inherited directly

[14] Here understood as historically excluded groups.
[15] Trend indicating a population-level decline in general cognitive ability.

from colonial propaganda. Sadly, this conception is still very much common in every society layer, unconscious or not. In fact, a simple search of Papua New Guinea in the Google Arts&Culture's section yields "a magical place that is incredibly rich in culture and far removed from modernity" as the first definition. Furthermore, even in the academic literature, when talking about the Indigenous DSWS peoples of a colonized area some authors still refer to them as people with "visually seductive features of phenotype" (scientific article from 2013). This logic is still dangerous because places the hunter-gatherer groups in the same fantasized and oneiric idealization of the wildlife detached from reality and fixated in time, once again reducing them to a "state of nature" directly counterposed to our idea of "modernity" that they cannot escape.

7.4. The responsibility of the correct scientific communication

If the latest discoveries in biology (including genetics, epigenetics, transcriptomic, proteomic, physiology, etc.) have taught us something is that the living organisms are the result of a delicate multifactorial equilibrium that is almost impossible to predict. For this reason, when researchers try to understand the patterns behind these complex dynamics we need to simplify and focus on a small part of the actual reality. On top of that, if the research is about human evolution and social traits, we need to consider that the social and economic environment has the strongest impact on human populations and, therefore, bear in mind that any kind of study trying to identify the genetics behind complex traits such as intelligence (IQ)[16], criminality or athleticism is just wrong (Blakey, 2020; Gómez-Valdés et al., 2013; Lee et al., 2008; Montana, 2013). As scientists, we must also take into consideration our own biases. We are people socialized in a segregated white environment bearing the cultural Enlightenment heritage where

[16] The so-called "Flynn effect" (the tendency of every human group to achieve increasingly higher IQ scores through the generations) is one of the most compelling evidence demonstrating that the cultural habits have the stronger impact on getting higher IQ scores rather than having a particular genetic asset (Saini, 2020).

science was elevated to universal truth and scientific discoveries defined as an unchangeable reality (Blakey, 2021; DiAngelo, 2011).

The consequences of a simplistic data interpretation

One clear example of the deterministic interpretation and naturalization of the scientific results is the patenting of the first drug approved by the US Food and Drug Administration for a specific racialized group: the BiDil. It was, in fact, shown that the drug administrated to 49 self-defined black patients together with the standard treatment reduced heart failure. Even if the group-specific association was controversial, the drug was approved resurrecting the claims that humans are subdivided into biologically defined groups also in academia (Barbujani et al., 2013; Montana, 2013). Heart disease is complex, and it depends on stress, income, diet, and access to health care, therefore, the fact that it affects a specific group tells more about the society' structure than the group's genetics. Furthermore, no specific genetic marker was considered when analyzing the BiDil efficacy in the group under study, but rather the criterium for the inclusion was the sole self-identification to a specific ethnic group that is variable over time according to the sociopolitical environment (Krimsky, 2012). On top of that, Hammermeister et al. (2009) analyzed more than 76,500 patients having different ancestries, revealing no significant association between BiDil and mortality within the ancestry groups. As previously mentioned, genetic variation among human populations is clinal, however, two populations placed at the two extremes of a cline tend to present different genetic assets, especially in complex traits. Several genetic studies have, indeed, reported different frequencies among the populations for alleles associated with complex disease and drug response, contributing to inter-population variability (Guha et al., 2015; Piel et al., 2010; Shoily et al., 2021). However, this cannot be interpreted as a biological confirmation of the race's existence, it is rather the result of a complex balance between adaptation and demography. For all the reasons listed here, it is important to increase the representation of non-European descendants in clinical research studies, to include a higher

genetic variability in pharmaceutical trials, while a transdisciplinary approach is necessary to take into consideration the social environment as a strong variable that can certainly influence the individual's health (Batai et al., 2021; Brody & Hunt, 2006; Montana, 2013).

7.5. Laying the foundation for a more conscious research

For all the reasons reported above, there is a growing number of scientists engaging in the application of specific good practice guidelines when talking about human populations. In 2008, Lee et al. proposed some principles to follow when characterizing group differences based on human genetic variation. Among them, while recognizing the inexistence of biological racial categories but also acknowledging the correlation between human genetic variation and geographic distribution, the authors focused on the importance of having clear and open communication between scientists and the general public. Both when analyzing the individual's genetic profile and when establishing human categories for a study, Lee et al. (2008) recommend carefully detailing why the categorization is important for the study and putting the results into a broader context. Furthermore, they recommend taking into consideration the sociopolitical consequence of particular genetic explanations for group differences in complex traits and avoiding overemphasizing the genetic contribution of the group. Furthermore, the importance of collaboration with the communities has been assuming a central role both in aDNA research and in large-scale genomic initiatives in developing countries (Patrinos et al., 2020; Wagner et al., 2020). While on one hand, the tight collaboration with the communities reinforces their trust in the research, on the other hand, it helps scientists to increase the quality of their work. However, community interests must be prioritized especially when the results could have an impact on their sociopolitical status. For this reason, scientists are starting to lean towards addressing questions that are important for the community rather than deciding them *a priori*. Moreover, community engagement in the research should continue throughout the process, from the discussion on how the community should benefit from the research when defining the

project's objectives, to the analysis of the final results. During this last step of the work, many researchers are emphasizing the importance of the community's feedback for the results' interpretation leaving to them the final decision on what should be published. Finally, an outreach program for the data returning to the population has been shown to enhance community engagement increasing both the people's curiosity about scientific-related questions and the probability of future collaborations (Ávila-arcos et al., 2022; Dávalos et al., 2022). Many scholars are also pointing out the need to discuss the problems related to the language used in science communication and the possible historical (and potentially colonialistic) background behind an evolutionary hypothesis, considering the dangerous possibility to fall into deterministic conclusions that could further stigmatize and marginalize vulnerable groups (M. Hernandez & Perry, 2021; Khan, 2021).

As reported in chapter 2.7, the populations carrying the DSWS phenotype have a particularly complicated historical background having suffered invasions, rape, kidnapping, and dehumanization "justified" by their different physical appearance. Although there is much room for improvement, understanding the history of the population under study here, the mechanisms used for their stigmatization as *alter*, and the social repercussions suffered by them, have been of critical importance in this work. For instance, to avoid the mythological representation of the hunter-gatherer groups and put the results interpretation in a wider context, it was important to specify that the plastic mechanism for the anticipated reproduction in stressful environments is present in all humans and that being under selection in populations carrying the DSWS phenotype is only the natural consequence of a specific environmental selective pressure and that therefore, it does not have any other type of implication. Furthermore, by renaming the phenotypic features with the DSWS acronym I tried to avoid any kind of deterministic naturalization and objectification, focusing only on the biological mechanisms and never identifying the people only with one of their many characteristics. As researchers, it is my belief that we should work towards making the process of questioning ourselves at every step of the way and recognizing our biases and limitations by being open to different perspectives, a core value.

CONCLUSIONS

Using genomic data we show that rainforest populations from Malaysia and the Philippines carrying the DSWS phenotype show signals of positive selection in the *CASR* gene as previously reported in the Andamanese genomes. Furthermore, the analysis of the rs1042636 worldwide frequency distribution revealed that populations from East Asia and particularly from South East Asia carry the highest frequency of the derived allele of this human polymorphism.

The CaSR protein is a G-coupled receptor that modulates intracellular cAMP levels. The rs1042636, causing the R990G substitution, is gain-of-function and generates a higher sensitive receptor to extracellular calcium. At a systemic level, CaSR's main function is to control calcium homeostasis by regulating the Parathyroid Hormone expression while balancing the Ca^{2+} excretion in kidneys and controlling its levels in serum through bone reabsorption and intestinal uptake. Besides, CaSR is also expressed in the bone growth plate where it regulates chondrocyte differentiation, and in osteoblasts and osteoclasts regulating bone synthesis and digestion. In the skin, the *CASR* gene is involved in keratinocytes differentiation, while in adipose tissue it regulates, on one hand, adipogenesis and on the other the whole adipose tissue homeostasis through the balancing of the lipogenic-lipolytic pathways.

The mouse model for the R990G substitution shows differential phenotypes between the homozygotes of the ancestral and derived alleles. At a systemic level, 990G homozygotes, in comparison to their littermates R990, show higher weight, higher body mass index, the tendency to have higher fat percentage and lower lean percentage, as well as a lower concentration of serum cholesterol and a higher tendency to develop liver steatosis in females. Furthermore, when performing morphometrical analysis on the adipose tissue histology, 990G homozygotes display smaller cell size both in subcutaneous and in visceral fat, while showing at the same time a downregulation of the lipolytic-related genes in visceral fat when performing quantitative PCR. Overall, our mouse model shows an enhanced fat accumulation probably due to the higher activity of CaSR during preadipocyte

maturation and adipocyte differentiation, generating a higher number of adipocyte cells in 990G animals as well as because of both enhanced adipogenesis and lipolysis downregulation.

A phenotype of greater fat accumulation and body weight in Indigenous hunter-gatherers from South East Asia could easily facilitate an earlier sexual maturation. These results agree with the hypothesis that some of the DSWS phenotypes (the higher fat accumulation resulting in steatopygia and the short stature) could result from selection for an early onset of reproduction in a particularly hostile environment such as the rainforest.

Since "nothing in Biology makes sense except in the light of evolution" (Dobzhansky, 1973), studying human adaptation and understanding the biological processes underneath the human differences has always fascinated evolutionary biologists. Furthermore, from a biological, medical, and ethical perspective, it is important to increase the representation in biomedical studies of understudied populations to achieve a more complete and accurate description of human variation for the development of new drugs. However, the ghosts of the nationalistic, colonial, and racist 18th-century legacy are very much alive in every societal layer including the scientific world. Using the words of Dr. Richard Cooper "Racialized thinking is such a deep part […] of our psychology that we can't just by conscious effort free ourselves from it completely. It keeps popping out in ways when we're unprepared or not vigilant" (Saini, 2020). For these reasons, it is our responsibility as scientists to be self-critical and conduct respectful, inclusive, and transdisciplinary research projects in collaboration with the socially excluded groups making sure to give the produced knowledge back to them. Besides, when communicating our research in scientific circles, it is our duty to be unequivocally clear about our results and conclusions putting them in a wider and multidisciplinary context to reject any possible misinterpretation.

Nevertheless, finding a solution to such complex issues is a work in progress. Like Michelangelo's slaves are constantly fighting the marble

that traps them, in the same way, we should constantly fight the prejudice that entangles us, questioning ourselves, challenging and deconstructing our most profound beliefs to overcome (conscious or unconscious) biases and build a fairer society through a more inclusive science.

REFERENCES

Aghakhanian, F., Yunus, Y., Naidu, R., Jinam, T., Manica, A., Hoh, B. P., & Phipps, M. E. (2015). Unravelling the genetic history of Negritos and Indigenous populations of Southeast Asia. *Genome Biology and Evolution*, *7*(5), 1206–1215. https://doi.org/10.1093/gbe/evv065

Akey, J. M. (2009). Constructing genomic maps of positive selection in humans: Where do we go from here? *Genome Research*, *19*(5), 711–722. https://doi.org/10.1101/gr.086652.108

Allen, J., & O'connell, J. F. (2020). A different paradigm for the initial colonisation of Sahul. *Archaeology in Oceania*, *55*(1), 1–14. https://doi.org/10.1002/arco.5207

Anderson, A. (2005). Crossing the Luzon Strait: archaeological chronology in the Batanes Islands, Philippines and the regional sequence of Neolithic dispersal. *Journal of Austronesian Studies*, *1*(2), 25–45.

Anderson, C. (2011). Colonization, kidnap and confinement in the Andamans penal colony, 1771-1864. *Journal of Historical Geography*, *37*(1), 68–81. https://doi.org/10.1016/j.jhg.2010.07.001

Anderson, C. (2018). The Andaman Islands Penal Colony: Race, Class, Criminality, and the British Empire. *International Review of Social History*, *63*(S26), 25–43. https://doi.org/10.1017/S0020859018000202

Atshaves, B. P., Martin, G. G., Hostetler, H. A., McIntosh, A. L., Kier, A. B., & Schroeder, F. (2010). Liver fatty acid-binding protein and obesity. *Journal of Nutritional Biochemistry*, *21*(11), 1015–1032. https://doi.org/10.1016/j.jnutbio.2010.01.005

Auton, A., Abecasis, G. R., Altshuler, D. M., Durbin, R. M., Bentley, D. R., Chakravarti, A., Clark, A. G., Donnelly, P., Eichler, E. E., Flicek, P., Gabriel, S. B., Gibbs, R. A., Green, E. D., Hurles, M. E., Knoppers, B. M., Korbel, J. O., Lander, E. S., Lee, C., Lehrach, H., … Schloss, J. A. (2015). A global reference for human genetic variation. *Nature*, *526*(7571), 68–74. https://doi.org/10.1038/nature15393

Ávila-arcos, M. C., De, C., Castro, F., & Nieves-colón, M. A. (2022). *Recommendations for Sustainable Ancient DNA Research in the Global South : Voices From a New Generation of Paleogenomicists*. *13*(April), 1–8. https://doi.org/10.3389/fgene.2022.880170

Babinsky, V. N., Hannan, F. M., Ramracheya, R. D., Zhang, Q., Nesbit, M. A., Hugill, A., Bentley, L., Hough, T. A., Joynson, E., Stewart, M., Aggarwal, A., Prinz-Wohlgenannt, M., Gorvin, C. M., Källay, E., Wells, S., Cox, R. D., Richards, D., Rorsman, P., & Thakker, R. v. (2017). Mutant mice with calcium-sensing receptor activation have hyperglycemia that is rectified by calcilytic therapy. *Endocrinology*, *158*(8), 2486–2502. https://doi.org/10.1210/en.2017-00111

Bamshad, M., & Wooding, S. P. (2003). Signatures of natural selection in the human genome. *Nature Reviews Genetics*, *4*(2), 99–111. https://doi.org/10.1038/nrg999

Barbujani, G., Ghirotto, S., & Tassi, F. (2013). Nine things to remember about human genome diversity. *Tissue Antigens*, *82*(3), 155–164. https://doi.org/10.1111/tan.12165

Barrangou, R., & Marraffini, L. A. (2014). CRISPR-cas systems: Prokaryotes upgrade to adaptive immunity. *Molecular Cell*, *54*(2), 234–244. https://doi.org/10.1016/j.molcel.2014.03.011

Batai, K., Hooker, S., & Kittles, R. A. (2021). Leveraging genetic ancestry to study health disparities. *American Journal of Physical Anthropology*, *175*(2), 363–375. https://doi.org/10.1002/ajpa.24144

Bellwood, P., & Hiscock, P. (2005). Australia and the Austronesians. *The Human Past*, 264–305. http://archanth.anu.edu.au/sites/default/files/documents/bellwood_hiscock.pdf

Benjamin, G. (1976). Austroasiatic Subgroupings and Prehistory in the Malay Peninsula. *Oceanic Linguistics Special Publications*, *13*, 37–128.

Bergström, A., Oppenheimer, S. J., Mentzer, A. J., Auckland, K., Robson, K., Attenborough, R., Alpers, M. P., Koki, G., Pomat, W., Siba, P., Xue, Y., Sandhu, M. S., & Tyler-Smith, C. (2017). A Neolithic expansion, but strong genetic structure, in the independent history of New Guinea. *Science*, *357*(6356), 1160–1163. https://doi.org/10.1126/science.aan3842

Biswas, S. (2020). India coronavirus : Covid strikes remote Greater Andamanese tribe. *BBC News*. https://www.bbc.com/news/world-asia-india-53925071

Biswas, S., & Akey, J. M. (2006). Genomic insights into positive selection. *Trends in Genetics*, *22*(8), 437–446. https://doi.org/10.1016/j.tig.2006.06.005

Blakey, M. L. (2020). On the biodeterministic imagination. *Archaeological Dialogues*, *27*(1), 1–16. https://doi.org/10.1017/S1380203820000021

Blakey, M. L. (2021). Understanding racism in physical (biological) anthropology. *American Journal of Physical Anthropology*, *175*(2), 316–325. https://doi.org/10.1002/ajpa.24208

Blust, R. (2013). Terror from the sky: Unconventional linguistic clues to the negrito past. *Human Biology*, *85*(1–3), 401–415. https://doi.org/10.3378/027.085.0319

Blust, R. A. (2018). Austronesian languages. *Encyclopedia Britannica*. https://doi.org/10.4324/9781315644936

Borja, M. R. de. (2005). *Basques in the Philippines* (W. A. Douglass, Ed.; 1st ed.). University of Nevada Press.

Bravo-Sagua, R., Mattar, P., Díaz, X., Lavandero, S., & Cifuentes, M. (2016). Calcium sensing receptor as a novel mediator of adipose tissue dysfunction: Mechanisms and potential clinical implications. *Frontiers in Physiology*, *7*(SEP), 1–7. https://doi.org/10.3389/fphys.2016.00395

Brody, H., & Hunt, L. M. (2006). BiDil: Assessing a race-based pharmaceutical. *Annals of Family Medicine*, *4*(6), 556–560. https://doi.org/10.1370/afm.582

Brown, E. M. (2013). Role of the calcium-sensing receptor in extracellular calcium homeostasis. *Best Practice and Research: Clinical Endocrinology and Metabolism*, *27*(3), 333–343. https://doi.org/10.1016/j.beem.2013.02.006

Brown, E. M., & mac Leod, J. R. (2001). Extracellular Calcium Sensing and Extracellular Calcium Signaling. *Physiological Reviews*, *81*(1), 239–297. https://doi.org/10.1038/nrm1154

Brucato, N., André, M., Tsang, R., Saag, L., Kariwiga, J., Sesuki, K., Beni, T., Pomat, W., Muke, J., Meyer, V., Boland, A., Deleuze, J. F., Sudoyo, H., Mondal, M., Pagani, L., Gallego Romero, I., Metspalu, M., Cox, M. P., Leavesley, M., & Ricaut, F. X. (2021). Papua New Guinean Genomes Reveal the Complex Settlement

of North Sahul. *Molecular Biology and Evolution*, *38*(11), 5107–5121. https://doi.org/10.1093/molbev/msab238

Buniello, A., Macarthur, J. A. L., Cerezo, M., Harris, L. W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E., Suveges, D., Vrousgou, O., Whetzel, P. L., Amode, R., Guillen, J. A., Riat, H. S., Trevanion, S. J., Hall, P., Junkins, H., … Parkinson, H. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Research*, *47*(D1), D1005–D1012. https://doi.org/10.1093/nar/gky1120

Bxown, A. K. (1909). The Religion of the Andaman Islanders. *Folklore*, *20*(3), 257–371. https://doi.org/10.1080/0015587X.1909.9719883

Canlorbe, G. (2019a). A Conversation with Gerhard Meisenberg. *Psych*, *1*(1), 364–374. https://doi.org/10.3390/psych1010026

Canlorbe, G. (2019b). A Conversation with Michael A. Woodley of Menie, Yr. *Psych*, *1*(1), 207–219. https://doi.org/10.3390/psych1010015

Capecchi, M. R. (1989). Altering the Genome by Homologous Recombination. *Science*, *244*(4910), 1288–1292.

Carroll, D. (2017). Genome editing: Past, present, and future. *Yale Journal of Biology and Medicine*, *90*(4), 653–659.

Chakravarti, B., Chattopadhyay, N., & Brown, E. M. (2012). Signaling through the extracellular calcium-sensing receptor (CaSR). In *Advances in Experimental Medicine and Biology* (Vol. 740). https://doi.org/10.1007/978-94-007-2888-2_5

Chang, W., Dvorak, M., & Shoback, D. (2010). Assessing constitutive activity of extracellular calcium-sensing receptors in vitro and in bone. *Methods in Enzymology*, *484*(C), 253–266. https://doi.org/10.1016/B978-0-12-381298-8.00013-7

Chang, W., Tu, C., Chen, T.-H., Bikle, D., & Shoback, D. (2013). The CaSR Is a Critical Modulator of Skeletal Development. *Sci. Signal.*, *1*(35), 1–23. https://doi.org/10.1126/scisignal.1159945.The

Chatterjee, A. K. (2020). The science of the andamans and the sign of the four: The distorted racial hierarchy of british imperial

anthropology. *Shima*, *14*(2), 214–234. https://doi.org/10.21463/SHIMA.14.2.14

Chaubey, G., & Endicott, P. (2013). The Andaman islanders in a regional genetic context: Reexamining the evidence for an early peopling of the archipelago from south Asia. *Human Biology*, *85*(1–3), 153–172. https://doi.org/10.3378/027.085.0307

Chen, S., Lee, B., Lee, A. Y. F., Modzelewski, A. J., & He, L. (2016). Highly efficient mouse genome editing by CRISPR ribonucleoprotein electroporation of zygotes. *Journal of Biological Chemistry*, *291*(28), 14457–14467. https://doi.org/10.1074/jbc.M116.733154

Choin, J., Mendoza-Revilla, J., Arauna, L. R., Cuadros-Espinoza, S., Cassar, O., Larena, M., Ko, A. M. S., Harmant, C., Laurent, R., Verdu, P., Laval, G., Boland, A., Olaso, R., Deleuze, J. F., Valentin, F., Ko, Y. C., Jakobsson, M., Gessain, A., Excoffier, L., … Quintana-Murci, L. (2021). Genomic insights into population history and biological adaptation in Oceania. *Nature*, *592*(7855), 583–589. https://doi.org/10.1038/s41586-021-03236-5

Chowdhury, S., & Chakraborty, P. pratim. (2017). Universal health coverage - There is more to it than meets the eye. *Journal of Family Medicine and Primary Care*, *6*(2), 169–170. https://doi.org/10.4103/jfmpc.jfmpc

Cifuentes, M., Albala, C., & Rojas, C. (2005). Calcium-sensing receptor expression in human adipocytes. *Endocrinology*, *146*(5), 2176–2179. https://doi.org/10.1210/en.2004-1281

Coall, D. A., & Chisholm, J. S. (2003). Evolutionary perspectives on pregnancy: Maternal age at menarche and infant birth weight. *Social Science and Medicine*, *57*(10), 1771–1781. https://doi.org/10.1016/S0277-9536(03)00022-4

Conigrave, A. D., & Ward, D. T. (2013). Calcium-sensing receptor (CaSR): Pharmacological properties and signaling pathways. *Best Practice and Research: Clinical Endocrinology and Metabolism*, *27*(3), 315–331. https://doi.org/10.1016/j.beem.2013.05.010

Corbetta, S., Eller-Vainicher, C., Filopanti, M., Saeli, P., Vezzoli, G., Arcidiacono, T., Loli, P., Syren, M. L., Soldati, L., Beck-Peccoz, P., & Spada, A. (2006). R990G polymorphism of the calcium-sensing receptor and renal calcium excretion in patients with

primary hyperparathyroidism. *European Journal of Endocrinology*, *155*(5), 687–692. https://doi.org/10.1530/eje.1.02286

Crabtree, S. A., White, D. A., Bradshaw, C. J. A., Saltré, F., Williams, A. N., Beaman, R. J., Bird, M. I., & Ulm, S. (2021). Landscape rules predict optimal superhighways for the first peopling of Sahul. *Nature Human Behaviour*, *5*(10), 1303–1313. https://doi.org/10.1038/s41562-021-01106-8

Cribbs, A. P., & Perera, S. M. W. (2017). Science and bioethics of CRISPR-CAS9 gene editing: An analysis towards separating facts and fiction. *Yale Journal of Biology and Medicine*, *90*(4), 625–634.

Cuadros-Espinoza, S., Laval, G., Quintana-Murci, L., & Patin, E. (2022). The genomic signatures of natural selection in admixed human populations. *American Journal of Human Genetics*, *109*(4), 710–726. https://doi.org/10.1016/j.ajhg.2022.02.011

Dávalos, D. I. C., Cárdenas, Y. O. A., Bravo-Lopez, M. J., Neuenschwander, S., Reis, S., Bastos, M. Q. R., Stenderup, J., Yediay, F. E., Villa-Islas, V., Reyna-Blanco, C. S., Rodrigues-Carvalho, C., Hünemeier, T., Allentoft, M. E., Amorim, C. E. G., Moreno-Mayar, J. V., Ávila-Arcos, M. C., & Malaspinas, A.-S. (2022). Indigenous peoples in eastern Brazil: insights from 19th century genomes and metagenomes. *BioRxiv*, 2022.01.27.477466. https://www.biorxiv.org/content/10.1101/2022.01.27.477466v1%0Ahttps://www.biorxiv.org/content/10.1101/2022.01.27.477466v1.abstract

Deng, L., Pan, Y., Wang, Y., Chen, H., Yuan, K., Chen, S., Lu, D., Lu, Y., Mokhtar, S. S., Rahman, T. A., Hoh, B. P., & Xu, S. (2022). Genetic Connections and Convergent Evolution of Tropical Indigenous Peoples in Asia. *Molecular Biology and Evolution*, *39*(2), 1–19. https://doi.org/10.1093/molbev/msab361

Denham, T. (2004). The roots of agriculture and arboriculture in New Guinea: Looking beyond Austronesian expansion, Neolithic packages and indigenous origins. *World Archaeology*, *36*(4), 610–620. https://doi.org/10.1080/0043824042000303791

DiAngelo, R. (2011). White Fragility. *International Journal of Critical Pedagogy*, *3*(3), 54–70. https://doi.org/10.1051/tsm/201204014

Diez-Fraile, A., Lammens, T., Benoit, Y., & G.M.A. D'Herde, K. (2013). The Calcium-Sensing Receptor as a Regulator of Cellular Fate in Normal and Pathological Conditions. *Current Molecular*

Medicine, *13*(2), 282–295. https://doi.org/10.2174/1566524013804810763

Diffloth, G. (2018). Austroasiatic Languages. *Encyclopedia Britannica*. https://doi.org/10.1016/B0-08-044854-2/02110-6

Ding, Y., Wang, K. F., Wang, W. J., Ma, Y. R., Shi, T. Q., Huang, H., & Ji, X. J. (2019). Increasing the homologous recombination efficiency of eukaryotic microorganisms for enhanced genome engineering. *Applied Microbiology and Biotechnology*, *103*(11), 4313–4324. https://doi.org/10.1007/s00253-019-09802-2

Dobzhansky, T. (1973). Nothing in Biology Makes Sense except in the Light of Evolution. *The American Biology Teacher*, *35*(3), 125–129. https://doi.org/10.2307/4444260

Donzé-magnier, M. (2017). Edward Said : Orientalism. *Geonum Ed.*, *February*.

Dou, H. X., Wang, T., Su, H. X., Gao, D. D., Xu, Y. C., Li, Y. X., & Wang, H. Y. (2020). Exogenous FABP4 interferes with differentiation, promotes lipolysis and inflammation in adipocytes. *Endocrine*, *67*(3), 587–596. https://doi.org/10.1007/s12020-019-02157-8

Dvorak, M. M., Chen, T. H., Orwoll, B., Garvey, C., Chang, W., Bikle, D. D., & Shoback, D. M. (2007). Constitutive activity of the osteoblast Ca2+-sensing receptor promotes loss of cancellous bone. *Endocrinology*, *148*(7), 3156–3163. https://doi.org/10.1210/en.2007-0147

Ehrke-Schulz, E., Schiwon, M., Leitner, T., Dávid, S., Bergmann, T., Liu, J., & Ehrhardt, A. (2017). CRISPR/Cas9 delivery with one single adenoviral vector devoid of all viral genes. *Scientific Reports*, *7*(1), 1–11. https://doi.org/10.1038/s41598-017-17180-w

Endicott, P. (2013). Introduction: Revisiting the “Negrito” hypothesis: A transdisciplinary approach to human prehistory in southeast Asia. *Human Biology*, *85*(1–3), 7–20. https://doi.org/10.3378/027.085.0301

Endicott, P., Gilbert, M. T. P., Stringer, C., Lalueza-Fox, C., Willerslev, E., Hansen, A. J., & Cooper, A. (2003). The genetic origins of the Andaman Islanders. *American Journal of Human Genetics*, *72*(1), 178–184. https://doi.org/10.1086/345487

Fan, S., Hansen, M. E. B., Lo, Y., & Tishkoff, S. A. (2016). *Going global by adapting local: A review of recent human adaptation*. *354*(6308), 54–59. https://doi.org/10.1126/science.aaf5098.Going

Fay, J. C., & Wu, C. I. (2000). Hitchhiking under positive Darwinian selection. *Genetics*, *155*(3), 1405–1413. https://doi.org/10.1093/genetics/155.3.1405

Frank-Raue, K., Leidig-Bruckner, G., Haag, C., Schulze, E., Lorenz, A., Schmitz-Winnenthal, H., & Raue, F. (2011). Inactivating calcium-sensing receptor mutations in patients with primary hyperparathyroidism. *Clinical Endocrinology*, *75*(1), 50–55. https://doi.org/10.1111/j.1365-2265.2011.04059.x

Fukumoto, S., Chikatsu, N., Okazaki, R., Takeuchi, Y., Tamura, Y., Murakami, T., Obara, T., & Fujita, T. (2001). Inactivating mutations of calcium-sensing receptor results in parathyroid lipohyperplasia. *Diagnostic Molecular Pathology*, *10*(4), 242–247. https://doi.org/10.1097/00019606-200112000-00006

Geng, Y., Mosyak, L., Kurinov, I., Zuo, H., Sturchler, E., Cheng, T. C., Subramanyam, P., Brown, A. P., Brennan, S. C., Mun, H. C., Bush, M., Chen, Y., Nguyen, T. X., Cao, B., Chang, D. D., Quick, M., Conigrave, A. D., Colecraft, H. M., McDonald, P., & Fan, Q. R. (2016). Structural mechanism of ligand activation in human calcium-sensing receptor. *ELife*, *5*(JULY), 1–25. https://doi.org/10.7554/eLife.13662

Gómez-Valdés, J., Hünemeier, T., Quinto-Sánchez, M., Paschetta, C., de Azevedo, S., González, M. F., Martínez-Abadías, N., Esparza, M., Pucciarelli, H. M., Salzano, F. M., Bau, C. H. D., Bortolini, M. C., & González-José, R. (2013). Lack of Support for the Association between Facial Shape and Aggression: A Reappraisal Based on a Worldwide Population Genetics Perspective. *PLoS ONE*, *8*(1), 1–8. https://doi.org/10.1371/journal.pone.0052317

Gregory, T. R. (2008). Evolution as Fact, Theory, and Path. *Evolution: Education and Outreach*, *1*(1), 46–52. https://doi.org/10.1007/s12052-007-0001-z

Grossman, S. R., Shylakhter, I., Karlsson, E. K., Byrne, E. H., Morales, S., Frieden, G., Hostetter, E., Angelino, E., Garber, M., Zuk, O., Lander, E. S., Schaffner, S. F., & Sabeti, P. C. (2010). A composite of multiple signals distinguishes causal variants in

regions of positive selection. *Science*, *327*(5967), 883–886. https://doi.org/10.1126/science.1183863

Grzegorzewska, A. E., Frycz, B. A., Świderska, M., Niepolski, L., Mostowska, A., & Jagodziński, P. P. (2019). Calcium-sensing receptor gene (CASR) polymorphisms and CASR transcript level concerning dyslipidemia in hemodialysis patients: A cross-sectional study. *BMC Nephrology*, *20*(1), 1–14. https://doi.org/10.1186/s12882-019-1619-0

Guha, M., Bankura, B., Ghosh, S., Pattanayak, A. K., Ghosh, S., Pal, D. K., Puri, A., Kundu, A. K., & Das, M. (2015). Polymorphisms in CaSR and CLDN14 genes associated with increased risk of kidney stone disease in patients from the eastern part of India. *PLoS ONE*, *10*(6), 1–14. https://doi.org/10.1371/journal.pone.0130790

Gupta, A. (2017). Onge Tribe through the Ages : Continuity of Change Onge Tribe through the Ages : Continuity of Change. *International Journal of Social Science and Humanities Research*, *3*, 495–500.

Gurumurthy, C. B., Sato, M., Nakamura, A., Inui, M., Kawano, N., Islam, M. A., Ogiwara, S., Takabayashi, S., Matsuyama, M., Nakagawa, S., Miura, H., & Ohtsuka, M. (2019). Creation of CRISPR-based germline-genome-engineered mice without ex vivo handling of zygotes by i-GONAD. In *Nature Protocols* (Vol. 14, Issue 8). Springer US. https://doi.org/10.1038/s41596-019-0187-x

Hammarström, H., Forkel, R., Haspelmath, M., & Bank, S. (2022). Glottolog 4.6. *Leipzig: Max Planck Institute for Evolutionary Anthropology*. http://glottolog.org

Hammermeister, K. E., Fairclough, D., Emsermann, C. B., Hamman, R., Ho, M., Phibbs, S., Plomondon, M., Valuck, R., West, D., & Steiner, J. F. (2009). Effectiveness of hydralazine/isosorbide dinitrate in racial/ethnic subgroups with heart failure. *Clinical Therapeutics*, *31*(3), 632–643. https://doi.org/10.1016/j.clinthera.2009.03.019

Hansen, M., Flatt, T., & Aguilaniu, H. (2013). Reproduction, fat metabolism, and life span: What is the connection? *Cell Metabolism*, *17*(1), 10–19. https://doi.org/10.1016/j.cmet.2012.12.003

Hashimoto, M., Yamashita, Y., & Takemoto, T. (2016). Electroporation of Cas9 protein/sgRNA into early pronuclear zygotes generates non-mosaic mutants in the mouse. *Developmental Biology*, *418*(1), 1–9. https://doi.org/10.1016/j.ydbio.2016.07.017

Hay, E. A., Khalaf, A. R., Marini, P., Brown, A., Heath, K., Sheppard, D., & MacKenzie, A. (2017). An analysis of possible off target effects following CAS9/CRISPR targeted deletions of neuropeptide gene enhancers from the mouse genome. *Neuropeptides*, *64*, 101–107. https://doi.org/10.1016/j.npep.2016.11.003

He, Y. H., He, Y., Liao, X. L., Niu, Y. C., Wang, G., Zhao, C., Wang, L., Tian, M. J., Li, Y., & Sun, C. H. (2012). The calcium-sensing receptor promotes adipocyte differentiation and adipogenesis through PPARγ pathway. *Molecular and Cellular Biochemistry*, *361*(1–2), 321–328. https://doi.org/10.1007/s11010-011-1118-5

He, Y. H., Song, Y., Liao, X. L., Wang, L., Li, G., Alima, Li, Y., & Sun, C. H. (2011). The calcium-sensing receptor affects fat accumulation via effects on antilipolytic pathways in adipose tissue of rats fed low-calcium diets. *Journal of Nutrition*, *141*(11), 1938–1946. https://doi.org/10.3945/jn.111.141762

He, Y. han, Kong, W. lan, Wang, G., Zhao, Y., Bi, M. xin, Na, L. xin, Wang, M. qing, Perry, B., & Li, Y. (2014). The calcium-sensing receptor R990G polymorphism is associated with increased risk of hypertriglyceridemia in obese Chinese. *Gene*, *533*(1), 67–71. https://doi.org/10.1016/j.gene.2013.09.122

Headland, T. N. (1991). How Negrito Foragers Live in a Philippine Rainforest: What They Eat and What They Don't Eat. *Food and Nutrition in the Tropical Forest: Biocultural Interactions and Applications to Development*, *November*, 1–14. https://doi.org/10.13140/RG.2.1.2579.6567

Hendy, G. N., D'Souza-Li, L., Yang, B., Canaff, L., & Cole, D. E. C. (2000). Mutations of the calcium-sensing receptor (CASR) in familial hypocalciuric hypercalcemia, neonatal severe hyperparathyroidism, and autosomal dominant hypocalcemia. *Human Mutation*, *16*(4), 281–296. https://doi.org/10.1002/1098-1004(200010)16:4<281::AID-HUMU1>3.0.CO;2-A

Hermisson, J., & Pennings, P. S. (2017). Soft sweeps and beyond: understanding the patterns and probabilities of selection footprints under rapid adaptation. *Methods in Ecology and Evolution*, *8*(6), 700–716. https://doi.org/10.1111/2041-210X.12808

Hernandez, M., & Perry, G. H. (2021). Scanning the human genome for "signatures" of positive selection: Transformative opportunities and ethical obligations. *Evolutionary Anthropology*, *30*(2), 113–121. https://doi.org/10.1002/evan.21893

Hernandez, R. D., Kelley, J. L., Elyashiv, E., Melton, S. C., Auton, A., McVean, G., Sella, G., & Przeworski, M. (2011). Classic selective sweeps were rare. *Science*, *331*(6019), 920–924. https://doi.org/10.1126/science.1198878.Classic

Hertzel, A. v., Smith, L. A., Berg, A. H., Cline, G. W., Shulman, G. I., Scherer, P. E., & Bernlohr, D. A. (2006). Lipid metabolism and adipokine levels in fatty acid-binding protein null and transgenic mice. *American Journal of Physiology - Endocrinology and Metabolism*, *290*(5), 814–823. https://doi.org/10.1152/ajpendo.00465.2005

Hough, T. A., Bogani, D., Cheeseman, M. T., Favor, J., Nesbit, M. A., Thakker, R. v., & Lyon, M. F. (2004). Activating calcium-sensing receptor mutation in the mouse is associated with cataracts and ectopic calcification. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(37), 13566–13571. https://doi.org/10.1073/pnas.0405516101

Hu, J., & Spiegel, A. M. (2007). Structure and function of the human calcium-sensing receptor: Insights from natural and engineered mutations and allosteric modulators: Ca2+-Sensing Receptor Review Series. *Journal of Cellular and Molecular Medicine*, *11*(5), 908–922. https://doi.org/10.1111/j.1582-4934.2007.00096.x

Indonesia attack: Gunmen kill 24 construction workers in Papua. (2018). *BBC News*. https://www.bbc.com/news/world-asia-46446719

Jeong, S., & Yoon, M. (2011). 17β-Estradiol inhibition of PPARγ-induced adipogenesis and adipocyte-specific gene expression. *Acta Pharmacologica Sinica*, *32*(2), 230–238. https://doi.org/10.1038/aps.2010.198

Jinam, T. A., Phipps, M. E., Aghakhanian, F., Majumder, P. P., Datar, F., Stoneking, M., Sawai, H., Nishida, N., Tokunaga, K., Kawamura, S., Omoto, K., & Saitou, N. (2017). Discerning the

origins of the negritos, first sundal and people: Deep divergence and archaic admixture. *Genome Biology and Evolution*, *9*(8), 2013–2022. https://doi.org/10.1093/gbe/evx118

Joosen, A. M. C. P., Bakker, A. H. F., Zorenc, A. H. G., Kersten, S., Schrauwen, P., & Westerterp, K. R. (2006). PPARγ activity in subcutaneous abdominal fat tissue and fat mass gain during short-term overfeeding. *International Journal of Obesity*, *30*(2), 302–307. https://doi.org/10.1038/sj.ijo.0803146

Journé, F., Dumon, J. C., Kheddoumi, N., Fox, J., Laïos, I., Leclercq, G., & Body, J. J. (2004). Extracellular calcium downregulates estrogen receptor alpha and increases its transcriptional activity through calcium-sensing receptor in breast cancer cells. *Bone*, *35*(2), 479–488. https://doi.org/10.1016/j.bone.2004.03.021

Junker, L. L. (2010). Hunter-gatherer prehispanic Philippines. *World*, *27*(3), 389–410.

Kapur, K., Johnson, T., Beckmann, N. D., Sehmi, J., Tanaka, T., Kutalik, Z., Styrkarsdottir, U., Zhang, W., Marek, D., Gudbjartsson, D. F., Milaneschi, Y., Holm, H., DiIorio, A., Waterworth, D., Li, Y., Singleton, A. B., Bjornsdottir, U. S., Sigurdsson, G., Hernandez, D. G., … Bergmann, S. (2010). Genome-wide meta-analysis for serum calcium identifies significantly associated SNPs near the calcium-sensing receptor (CASR) gene. *PLoS Genetics*, *6*(7), 1–12. https://doi.org/10.1371/journal.pgen.1001035

Keen, I. (2003). Aboriginal economy and society at the threshold of colonisation. *Before Farming*, *2003*(3), 1–24. https://doi.org/10.3828/bfarm.2003.3.2

Kersten, S. (2001). Mechanisms of nutritional and hormonal regulation of lipogenesis. *EMBO Reports*, *2*(4), 282–286. https://doi.org/10.1093/embo-reports/kve071

Khan, A. (2021). Equity, diversity and inclusion a call to eradicate non-inclusive terms from the life sciences. *ELife*, *10*, 1–5. https://doi.org/10.7554/eLife.65604

Krimsky, S. (2012). The short life of a race drug. *The Lancet*, *379*(9811), 114–115. https://doi.org/10.1016/S0140-6736(12)60052-X

Krušinová, E., & Pelikánová, T. (2008). Fatty acid binding proteins in adipose tissue: A promising link between metabolic syndrome and atherosclerosis? *Diabetes Research and Clinical Practice*, *82*(SUPPL. 2), 127–134. https://doi.org/10.1016/j.diabres.2008.09.023

Kubota, N., Terauchi, Y., Miki, H., Tamemoto, H., Yamauchi, T., Komeda, K., Satoh, S., Nakano, R., Ishii, C., Sugiyama, T., Kazuhiro, E., Tsubamoto, Y., Okuno, A., Murakami, K., Sekihara, H., Hasegawa, G., Naito, M., Toyoshima, Y., Tanaka, S., … Kadowaki, T. (1999). PPARγ mediates high-fat diet-induced adipocyte hypertrophy and insulin resistance. *Molecular Cell*, *4*(4), 597–609. https://doi.org/10.1016/S1097-2765(00)80210-5

Lassek, W. D., & Gaulin, S. J. C. (2007). Brief Communication: Menarche is Related to Fat Distribution. *American Journal of Physical Anthropology*, *133*, 1147–1151. https://doi.org/10.1002/ajpa

Lee, S. S. J., Mountain, J., Koenig, B., Altman, R., Brown, M., Camarillo, A., Cavalli-Sforza, L., Cho, M., Eberhardt, J., Feldman, M., Ford, R., Greely, H., King, R., Markus, H., Satz, D., Snipp, M., Steele, C., & Underhill, P. (2008). The ethics of characterizing difference: Guiding principles on using racial categories in human genetics. *Genome Biology*, *9*(7), 404.1-404.4. https://doi.org/10.1186/gb-2008-9-7-404

Li, H. L., Wu, X., Xu, A., & Hoo, R. L. C. (2021). A-fabp in metabolic diseases and the therapeutic implications: An update. *International Journal of Molecular Sciences*, *22*(17), 1–22. https://doi.org/10.3390/ijms22179386

Li, T., Guo, W., & Zhou, Z. (2022). Adipose triglyceride lipase in hepatic physiology and pathophysiology. *Biomolecules*, *12*(1), 1–17. https://doi.org/10.3390/biom12010057

Liljeblad, J. (2014). Human Safaris: A Foucauldian Alternative to the Law's Treatment of the Indigenous Andaman Jawara. *University of Southern California*, *18*(2), 6–18. http://heinonline.org/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/austindlr18§ion=17

Lipson, M., Loh, P. R., Patterson, N., Moorjani, P., Ko, Y. C., Stoneking, M., Berger, B., & Reich, D. (2014). Reconstructing

Austronesian population history in Island Southeast Asia. *Nature Communications*, *5*, 1–7. https://doi.org/10.1038/ncomms5689

Lowell, B. B. (1999). PPARγ: An essential regulator of adipogenesis and modulator of fat cell function. *Cell*, *99*(3), 239–242. https://doi.org/10.1016/S0092-8674(00)81654-2

Lundholm, L., Zang, H., Hirschberg, A. L., Gustafsson, J. Å., Arner, P., & Dahlman-Wright, K. (2008). Key lipogenic gene expression can be decreased by estrogen in human adipose tissue. *Fertility and Sterility*, *90*(1), 44–48. https://doi.org/10.1016/j.fertnstert.2007.06.011

Magno, A. L., Ward, B. K., & Ratajczak, T. (2011). The calcium-sensing receptor: A molecular perspective. *Endocrine Reviews*, *32*(1), 3–30. https://doi.org/10.1210/er.2009-0043

Malaspinas, A. S., Westaway, M. C., Muller, C., Sousa, V. C., Lao, O., Alves, I., Bergström, A., Athanasiadis, G., Cheng, J. Y., Crawford, J. E., Heupink, T. H., MacHoldt, E., Peischl, S., Rasmussen, S., Schiffels, S., Subramanian, S., Wright, J. L., Albrechtsen, A., Barbieri, C., … Willerslev, E. (2016). A genomic history of Aboriginal Australia. *Nature*, *538*(7624), 207–214. https://doi.org/10.1038/nature18299

Mallick, S., Li, H., Lipson, M., Mathieson, I., Gymrek, M., Racimo, F., Zhao, M., Chennagiri, N., Nordenfelt, S., Tandon, A., Skoglund, P., Lazaridis, I., Sankararaman, S., Fu, Q., Rohland, N., Renaud, G., Erlich, Y., Willems, T., Gallo, C., … Reich, D. (2016). The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature*, *538*(7624), 201–206. https://doi.org/10.1038/nature18964

Manickham, S. K. (2009). Africans in Asia: The Discourse of "Negritos" in Early Nineteenth -century Southeast Asia. *Amsterdam University Press.*, 69–79.

Matta, C., & Mobasheri, A. (2014). Regulation of chondrogenesis by protein kinase C: Emerging new roles in calcium signalling. *Cellular Signalling*, *26*(5), 979–1000. https://doi.org/10.1016/j.cellsig.2014.01.011

McAllister, P., Nagle, N., & Mitchell, R. J. (2013). The Australian barrineans and their relationship to southeast Asian negritos: An investigation using mitochondrial genomics. *Human Biology*, *85*(1–3), 485–494. https://doi.org/10.3378/027.085.0322

McColl, H., Racimo, F., Vinner, L., Demeter, F., Gakuhari, T., Moreno-mayar, J. V., Driem, G. van, Wilken, U. G., Seguin-orlando, A., De, C., & Castro, F. (2018). The prehistoric peopling of Southeast Asia. *Science*, *92*(July), 88–92.

Meisenberg, G. (2009). The belief that genes cannot be changed is now outdated. *Nature*, *458*(7235), 145. https://doi.org/10.1038/458145a

Meyer, M., Kircher, M., Gansauge, M. T., Li, H., Racimo, F., Mallick, S., Schraiber, J. G., Jay, F., Prüfer, K., de Filippo, C., Sudmant, P. H., Alkan, C., Fu, Q., Do, R., Rohland, N., Tandon, A., Siebauer, M., Green, R. E., Bryc, K., … Pääbo, S. (2012). A high-coverage genome sequence from an archaic Denisovan individual. *Science*, *338*(6104), 222–226. https://doi.org/10.1126/science.1224344

Migliano, A. B., Romero, I. G., Metspalu, M., Leavesley, M., Pagani, L., Antao, T., Huang, D. W., Sherman, B. T., Siddle, K., Scholes, C., Hudjashov, G., Kaitokai, E., Babalu, A., Belatti, M., Cagan, A., Hopkinshaw, B., Shaw, C., Nelis, M., Metspalu, E., … Kivisild, T. (2013). Evolution of the pygmy phenotype: Evidence of positive selection from genome-wide scans in African, Asian, and Melanesian pygmies. *Human Biology*, *85*(1–3), 251–284. https://doi.org/10.3378/027.085.0313

Migliano, A. B., Vinicius, L., & Lahr, M. M. (2007). Life history trade-offs explain the evolution of human pygmies. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(51), 20216–20219. https://doi.org/10.1073/pnas.0708024105

Minter, T. (2010). *The Agta of the Northern Sierra Madre. Livelihood strategies and resilience among Philippine hunter-gatherers*. Universiteit Leiden the Netherlands. https://hdl.handle.net/1887/15549

Miranda, V. P. N., Faria, F. R. de, Faria, E. R. de, & Priore, S. E. (2014). Somatic maturation and body composition in female healthy adolescents with or without adjustment for body fat. *Revista Paulista de Pediatria*, *32*(1), 78–84. https://doi.org/10.1590/s0103-05822014000100013

Modzelewski, A. J., Chen, S., Willis, B. J., Lloyd, K. C. K., Wood, J. A., & He, L. (2018). Efficient mouse genome engineering by CRISPR-EZ technology. *Nature Protocols*, *13*(6), 1253–1274. https://doi.org/10.1038/nprot.2018.012

Mondal, M., Casals, F., Xu, T., Dall'Olio, G. M., Pybus, M., Netea, M. G., Comas, D., Laayouni, H., Li, Q., Majumder, P. P., & Bertranpetit, J. (2016). Genomic analysis of Andamanese provides insights into ancient human migration into Asia and adaptation. *Nature Genetics*, *48*(9), 1066–1070. https://doi.org/10.1038/ng.3621

Montana, A. (2013). Color blind. *Elements*, *9*(3), 165. https://doi.org/10.1177/0739456x08317171

Nakagawa, Y., Sakuma, T., Sakamoto, T., Ohmuraya, M., Nakagata, N., & Yamamoto, T. (2015). Production of knockout mice by DNA microinjection of various CRISPR/Cas9 vectors into freeze-thawed fertilized oocytes. *BMC Biotechnology*, *15*(1), 1–10. https://doi.org/10.1186/s12896-015-0144-x

News, U. (2022). Indonesia: Shocking abuses against indigenous Papuans, rights experts report. *United Nation News*, *16*, 3–5. https://news.un.org/en/story/2022/03/1113062

Nielsen, R., Hellmann, I., Hubisz, M., Bustamante, C., & Clark, A. G. (2007). Recent and ongoing selection in the human genome. *Nature Reviews Genetics*, *8*(11), 857–868. https://doi.org/10.1038/nrg2187

O'Connell, J. F., & Allen, J. (2010). Pre-LGM Sahul (Pleistocene Australia-New Guinea) and the Archaeology of Early Modern Humans. *Islands of Inquiry: Colonization, Seafaring and the Archaeology of Maritime Landscapes*, *March*, 395–410. https://www.researchgate.net/publication/239569942_Pre-LGM_Sahul_Pleistocene_Australia-New_Guinea_and_the_Archaeology_of_Early_Modern_Humans

Oh, Y. S., Seo, E. H., Lee, Y. S., Cho, S. C., Jung, H. S., Park, S. C., & Jun, H. S. (2016). Increase of calcium sensing receptor expression is related to compensatory insulin secretion during aging in mice. *PLoS ONE*, *11*(7), 1–15. https://doi.org/10.1371/journal.pone.0159689

Okuno, A., Tamemoto, H., Tobe, K., Ueki, K., Mori, Y., Iwamoto, K., Umesono, K., Akanuma, Y., Fujiwara, T., Horikoshi, H., Yazaki, Y., & Kadowaki, T. (1998). Troglitazone increases the number of small adipocytes without the change of white adipose

tissue mass in obese Zucker rats. *Journal of Clinical Investigation*, *101*(6), 1354–1361. https://doi.org/10.1172/JCI1235

Olivero, J., Fa, J. E., Farfán, M. A., Lewis, J., Hewlett, B., Breuer, T., Carpaneto, G. M., Fernández, M., Germi, F., Hattori, S., Head, J., Ichikawa, M., Kitanaishi, K., Knights, J., Matsuura, N., Migliano, A., Nese, B., Noss, A., Ekoumou, D. O., … Nasi, R. (2016). Distribution and numbers of Pygmies in central African forests. *PLoS ONE*, *11*(1), 1–16. https://doi.org/10.1371/journal.pone.0144499

Ong, K. T., Mashek, M. T., Bu, S. Y., Greenberg, A. S., & Mashek, D. G. (2011). Adipose triglyceride lipase is a major hepatic lipase that regulates triacylglycerol turnover and fatty acid signaling and partitioning. *Hepatology*, *53*(1), 116–126. https://doi.org/10.1002/hep.24006

Opubor, M. Z. (2020). Victorian punitive expeditions : Orientalist rationalisations behind the looting of colonial artefacts. *VIDES*, *8*(March), 192–204.

Panofsky, A., Dasgupta, K., & Iturriaga, N. (2021). How White nationalists mobilize genetics: From genetic ancestry and human biodiversity to counterscience and metapolitics. *American Journal of Physical Anthropology*, *175*(2), 387–398. https://doi.org/10.1002/ajpa.24150

Patrinos, G. P., Pasparakis, E., Koiliari, E., Pereira, A. C., Hünemeier, T., Pereira, L. v., & Mitropoulou, C. (2020). Roadmap for Establishing Large-Scale Genomic Medicine Initiatives in Low- and Middle-Income Countries. *American Journal of Human Genetics*, *107*(4), 589–595. https://doi.org/10.1016/j.ajhg.2020.08.005

Pauls, E. P. (2021). Khoekhoe. *Britannica Online Encyclopedia*, 1–2. https://www.britannica.com/topic/Khoekhoe

Pawley, A., Attenborough, R., Golson, J., & Hide, R. (2005). *Papuan Pasts: Cultural, linguistic and biological histories of Papuan-speaking peoples* (A. Pawley, R. Attenborough, J. Golson, & R. Hide, Eds.). Pacific Linguistics.

Pedro, N., Brucato, N., Fernandes, V., André, M., Saag, L., Pomat, W., Besse, C., Boland, A., Deleuze, J. F., Clarkson, C., Sudoyo, H., Metspalu, M., Stoneking, M., Cox, M. P., Leavesley, M., Pereira, L., & Ricaut, F. X. (2020). Papuan mitochondrial genomes and

the settlement of Sahul. *Journal of Human Genetics*, *65*(10), 875–887. https://doi.org/10.1038/s10038-020-0781-3

Perry, G. H., Foll, M., Grenier, J. C., Patin, E., Nédélec, Y., Pacis, A., Barakatt, M., Gravel, S., Zhou, X., Nsobya, S. L., Excoffier, L., Quintana-Murci, L., Dominy, N. J., & Barreiro, L. B. (2014). Adaptive, convergent origins of the pygmy phenotype in African rainforest hunter-gatherers. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(35), E3596–E3603. https://doi.org/10.1073/pnas.1402875111

Philip, P., Matsumura, H., & Bulbeck, D. (2017). *New Perspectives in Southeast Asian and Pacific Prehistory.* ANU Press. http://doi.org/10.22459/TA45.03.2017

Piel, F. B., Patil, A. P., Howes, R. E., Nyangiri, O. A., Gething, P. W., Williams, T. N., Weatherall, D. J., & Hay, S. I. (2010). Global distribution of the sickle cell gene and geographical confirmation of the malaria hypothesis. *Nature Communications*, *1*(8), 1–7. https://doi.org/10.1038/ncomms1104

Prokupecz, B. S., Maxouris, C., Andone, D., Beech, S., & Vera, A. (2022). What we know about Buffalo supermarket shooting suspect Payton Gendron. *CNN.* https://edition.cnn.com/2022/05/15/us/payton-gendron-buffalo-shooting-suspect-what-we-know/index.html

Pronczuk, M., & Ryckewaert, K. (2022). A Racist Researcher, Exposed by a Mass Shooting. *The New York Times.* https://www.nytimes.com/2022/06/09/world/europe/michael-woodley-buffalo-shooting.html

Pybus, M., Dall'Olio, G. M., Luisi, P., Uzkudun, M., Carreño-Torres, A., Pavlidis, P., Laayouni, H., Bertranpetit, J., & Engelken, J. (2014). 1000 Genomes Selection Browser 1.0: A genome browser dedicated to signatures of natural selection in modern humans. *Nucleic Acids Research*, *42*(D1), 903–909. https://doi.org/10.1093/nar/gkt1188

Randhawa, S., & Sengar, S. (2021). The evolution and history of gene editing technologies. In *Progress in Molecular Biology and Translational Science* (1st ed., Vol. 178). Elsevier Inc. https://doi.org/10.1016/bs.pmbts.2021.01.002

Ranieri, M., Tamma, G., di Mise, A., Vezzoli, G., Soldati, L., Svelto, M., & Valenti, G. (2013). Excessive signal transduction of gain-

of-function variants of the calcium-sensing receptor (CaSR) are associated with increased ER to cytosol calcium gradient. *PLoS ONE*, *8*(11), 1–10. https://doi.org/10.1371/journal.pone.0079113

Rees, J. S., Castellano, S., & Andrés, A. M. (2020). The Genomics of Human Local Adaptation. *Trends in Genetics*, *36*(6), 415–428. https://doi.org/10.1016/j.tig.2020.03.006

Reich, D., Patterson, N., Kircher, M., Delfin, F., Nandineni, M. R., Pugach, I., Ko, A. M. S., Ko, Y. C., Jinam, T. A., Phipps, M. E., Saitou, N., Wollstein, A., Kayser, M., Pääbo, S., & Stoneking, M. (2011). Denisova admixture and the first modern human dispersals into Southeast Asia and Oceania. *American Journal of Human Genetics*, *89*(4), 516–528. https://doi.org/10.1016/j.ajhg.2011.09.005

Reid, B. N., Ables, G. P., Otlivanchik, O. A., Schoiswohl, G., Zechner, R., Blaner, W. S., Goldberg, I. J., Schwabe, R. F., Chua, S. C., & Huang, L. S. (2008). Hepatic overexpression of hormone-sensitive lipase and adipose triglyceride lipase promotes fatty acid oxidation, stimulates direct release of free fatty acids, and ameliorates steatosis. *Journal of Biological Chemistry*, *283*(19), 13087–13099. https://doi.org/10.1074/jbc.M800533200

Relethford, J. H., & Harding, R. M. (2001). Population Genetics of Modern Human Evolution. *ELS*, 1–6. https://doi.org/10.1038/npg.els.0001470

Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J., & Kircher, M. (2018). CADD: Predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Research*, *47*(D1), D886–D894. https://doi.org/10.1093/nar/gky1016

Reyes, M., Rothe, H. M., Mattar, P., Shapiro, W. B., & Cifuentes, M. (2012). Antilipolytic effect of calcimimetics depends on the allelic variant of calcium-sensing receptor gene polymorphism rs1042636 (Arg990Gly). *European Journal of Human Genetics*, *20*(4), 480–482. https://doi.org/10.1038/ejhg.2011.221

Richard, C., Huo, R., Samadfam, R., Bolivar, I., Miao, D., Brown, E. M., Hendy, G. N., & Goltzman, D. (2010). The calcium-sensing receptor and 25-hydroxyvitamin D-1α-hydroxylase interact to modulate skeletal growth and bone turnover. *Journal of Bone and*

Mineral Research, *25*(7), 1627–1636. https://doi.org/10.1002/jbmr.58

Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L., & Corn, J. E. (2016). Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nature Biotechnology*, *34*(3), 339–344. https://doi.org/10.1038/nbt.3481

Riley, I. D. (1983). Population change and distribution in Papua New Guinea: an epidemiological approach. *Journal of Human Evolution*, *12*(1), 125–132. https://doi.org/10.1016/S0047-2484(83)80017-7

Roberts, P., & Petraglia, M. (2015). Pleistocene rainforests: barriers or attractive environments for early human foragers? *World Archaeology*, *47*(5), 718–739. https://doi.org/10.1080/00438243.2015.1073119

Roca-Umbert, A., Vogel-Gonzales, M., Fierro-Villegas, A., Garcia-Calleja, J., Bosnjak, A., Ill-raga, G., Herrera-fernández, V., Muntané, G., Campelo, F., Vicente, R., & Bosch, E. (2022). Functional characterization of the Met50Val substitution in SLC30A9 as a novel case of adaptive introgression in humans. *BioRxiv*. https://doi.org/https://doi.org/10.1101/2022.06.29.498106

Rocha, G., Villalobos, E., Fuentes, C., Villarroel, P., Reyes, M., Díaz, X., Mattar, P., & Cifuentes, M. (2015). Preadipocyte proliferation is elevated by calcium sensing receptor activation. *Molecular and Cellular Endocrinology*, *412*, 251–256. https://doi.org/10.1016/j.mce.2015.05.011

Ronen, R., Udpa, N., Halperin, E., & Bafna, V. (2013a). Learning natural selection from the site frequency spectrum. *Genetics*, *195*(1), 181–193. https://doi.org/10.1534/genetics.113.152587

Ronen, R., Udpa, N., Halperin, E., & Bafna, V. (2013b). Learning natural selection from the site frequency spectrum. *Genetics*, *195*(1), 181–193. https://doi.org/10.1534/genetics.113.152587

Rutherford, A. (2021). Race, eugenics, and the canceling of great scientists. *American Journal of Physical Anthropology*, *175*(2), 448–452. https://doi.org/10.1002/ajpa.24192

Rybczyńska, A., Marchwińska, A., Dyś, A., Boblewski, K., Lehmann, A., & Lewko, B. (2017). Activity of the calcium-sensing receptor influences blood glucose and insulin levels in rats. *Pharmacological Reports*, *69*(4), 709–713. https://doi.org/10.1016/j.pharep.2017.01.034

Sabeti, P. C., Schaffner, S. F., Fry, B., Lohmueller, J., Varilly, P., Shamovsky, O., Palma, A., Mikkelsen, T. S., Altshuler, D., & Lander, E. S. (2006a). Positive natural selection in the human lineage. *Science*, *312*(5780), 1614–1620. https://doi.org/10.1126/science.1124309

Sabeti, P. C., Schaffner, S. F., Fry, B., Lohmueller, J., Varilly, P., Shamovsky, O., Palma, A., Mikkelsen, T. S., Altshuler, D., & Lander, E. S. (2006b). Positive natural selection in the human lineage. *Science*, *312*(5780), 1614–1620. https://doi.org/10.1126/science.1124309

Sahani, R., Chakrabarty, S., & Bharati, P. (2010). Temporal trends in overweight and obesity among Nicobarese adults in Nicobar Islands, India, 1960s-1999. *Obesity Research and Clinical Practice*, *4*(2), e119–e125. https://doi.org/10.1016/j.orcp.2009.11.001

Saini, A. (2020). *Superior*. Fourth Estate.

Sandip, R. (2014). 10 years since Tsunami : Tribes survived disaster , but their languages are doomed. *FirstPost*. https://www.firstpost.com/world/10-years-since-tsunami-tribes-survived-disaster-but-their-languages-are-doomed-2016429.html

Santa Maria, C., Cheng, Z., Li, A., Wang, J., Shoback, D., Tu, C. L., & Chang, W. (2016). Interplay between CaSR and PTH1R signaling in skeletal development and osteoanabolism. *Seminars in Cell and Developmental Biology*, *49*, 11–23. https://doi.org/10.1016/j.semcdb.2015.12.004

Scully, P., & Crais, C. (2008). Race and Erasure: Sara Baartman and Hendrik Cesars in Cape Town and London. *Journal of British Studies*, *47*(2), 301–323. https://doi.org/10.1086/526552

Sen, S. (2009a). Savage bodies, civilized pleasures: M. V. Portman and the andamanese. *American Ethnologist*, *36*(2), 364–379. https://doi.org/10.1111/j.1548-1425.2009.01140.x

Sen, S. (2009b). Savage bodies, civilized pleasures: M. V. Portman and the andamanese. *American Ethnologist*, *36*(2), 364–379. https://doi.org/10.1111/j.1548-1425.2009.01140.x

Sharma, N. (2016). Lack of physical activity increases obesity among Onges in Andaman. *The Economic Times*. https://economictimes.indiatimes.com/news/politics-and-nation/-lack-of-physical-activity-increases-obesity-among-onges-in-andaman/articleshow/54840131.cms

Shaw, B., Field, J. H., Summerhayes, G. R., Coxe, S., Coster, A. C. F., Ford, A., Haro, J., Arifeae, H., Hull, E., Jacobsen, G., Fullagar, R., Hayes, E., & Kealhofer, L. (2020). Emergence of a Neolithic in highland New Guinea by 5000 to 4000 years ago. *Science Advances*, *6*(13), 1–10. https://doi.org/10.1126/sciadv.aay4573

Shoily, S. S., Ahsan, T., Fatema, K., & Sajib, A. A. (2021). Common genetic variants and pathways in diabetes and associated complications and vulnerability of populations with different ethnic origins. *Scientific Reports*, *11*(1), 1–14. https://doi.org/10.1038/s41598-021-86801-2

Siedler, C. W. (1956). Rhetorical Devices in Caesar's Commentaries. *The Classical Weekly*, *50*(2), 28–31.

Siervogel, R. M., Demerath, E. W., Schubert, C., Remsberg, K. E., Chumlea, W. C., Sun, S., Czerwinski, S. A., & Towne, B. (2003). Puberty and body composition. *Hormone Research*, *60*(SUPPL. 1), 36–45. https://doi.org/10.1159/000071224

Singh, A., Chakraborty, D., & Maiti, S. (2016). CRISPR/Cas9: A historical and chemical biology perspective of targeted genome engineering. *Chemical Society Reviews*, *45*(24), 6666–6684. https://doi.org/10.1039/c6cs00197a

Sorgoni, B. (2003). "Defending the race": The Italian reinvention of the Hottentot Venus during Fascism. *Journal of Modern Italian Studies*, *8*(3), 411–424. https://doi.org/10.1080/09585170320000113761

Spiegelman, B. M., & Flier, J. S. (1996). Adipogenesis and obesity: Rounding out the big picture. *Cell*, *87*(3), 377–389. https://doi.org/10.1016/S0092-8674(00)81359-8

Squires, P. E., Harris, T. E., Persaud, S. J., Curtis, S. B., Buchan, A. M. J., & Jones, P. M. (2000). The extracellular calcium-sensing

receptor on human β-cells negatively modulates insulin secretion. *Diabetes*, *49*(3), 409–417. https://doi.org/10.2337/diabetes.49.3.409

Stock, J. T., & Migliano, A. B. (2009). Stature, mortality, and life history among indigenous populations of the Andaman Islands, 1871-1986. *Current Anthropology*, *50*(5), 713–725. https://doi.org/10.1086/605429

Szpak, M., Xue, Y., Ayub, Q., & Tyler-Smith, C. (2019). How well do we understand the basis of classic selective sweeps in humans? *FEBS Letters*, *593*(13), 1431–1448. https://doi.org/10.1002/1873-3468.13447

Tajima, F. (1989). Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics Society of America*, *3*(9), 607–612.

Talukdar, A. (2017). a Study on Tribal Development of the Jarawa in the Andaman and Nicobar Islands. *International Journal of Advanced Research*, *5*(12), 1998–2010. https://doi.org/10.21474/ijar01/6173

Tang, Q. Q., & Lane, M. D. (2012). Adipogenesis: From Stem Cell to Adipocyte. *Annual Review of Biochemistry*, *81*(1), 715–736. https://doi.org/10.1146/annurev-biochem-052110-115718

Toshihiro, N. (2009). Living on the Periphery: Development and Islamization among the Orang Asli in Malaysia. *Kyoto University Press*, *14*, 1–402. https://www.coac.org.my/dashboard/modules/cms/cms~file/93c38c2f6837049ec87607013c0c5404.pdf

Tu, C. L., & Bikle, D. D. (2013). Role of the calcium-sensing receptor in calcium regulation of epidermal differentiation and function. *Best Practice and Research: Clinical Endocrinology and Metabolism*, *27*(3), 415–427. https://doi.org/10.1016/j.beem.2013.03.002

Tu, C. L., Chang, W., & Bikle, D. D. (2001). The Extracellular Calcium-sensing Receptor Is Required for Calcium-induced Differentiation in Human Keratinocytes. *Journal of Biological Chemistry*, *276*(44), 41079–41085. https://doi.org/10.1074/jbc.M107122200

Tu, C. L., Crumrine, D. A., Man, M. Q., Chang, W., Elalieh, H., You, M., Elias, P. M., & Bikle, D. D. (2012). Ablation of the calcium-

sensing receptor in keratinocytes impairs epidermal differentiation and barrier function. *Journal of Investigative Dermatology*, *132*(10), 2350–2359. https://doi.org/10.1038/jid.2012.159

Turksen, K., & Troy, T. C. (2003). Overexpression of the calcium sensing receptor accelerates epidermal differentiation and permeability barrier formation in vivo. *Mechanisms of Development*, *120*(6), 733–744. https://doi.org/10.1016/S0925-4773(03)00045-5

Venkateswar, S. (2004). *DEVELOPMENT AND ETHNOCIDE: COLONIAL PRACTICES IN THE ANDAMAN ISLANDS* (C. Erni & S. Stidsen, Eds.). IWGIA.

Vennu, V., Abdulrahman, T. A., & Bindawas, S. M. (2019). The prevalence of overweight, obesity, hypertension, and diabetes in India: Analysis of the 2015–2016 national family health survey. *International Journal of Environmental Research and Public Health*, *16*(20), 1–22. https://doi.org/10.3390/ijerph16203987

Vezzoli, G., Soldati, L., & Mora, S. (2017). Calcium-Sensing Receptor Polymorphisms and Human Disease. In *Molecular, Genetic, and Nutritional Aspects of Major and Trace Minerals*. Elsevier Inc. https://doi.org/10.1016/B978-0-12-802168-2.00001-4

Vezzoli, G., Terranegra, A., Arcidiacono, T., Biasion, R., Coviello, D., Syren, M. L., Paloschi, V., Giannini, S., Mignogna, G., Rubinacci, A., Ferraretto, A., Cusi, D., Bianchi, G., & Soldati, L. (2007). R990G polymorphism of calcium-sensing receptor does produce a gain-of-function and predispose to primary hypercalciuria. *Kidney International*, *71*(11), 1155–1162. https://doi.org/10.1038/sj.ki.5002156

Vilar, M. G., Kaneko, A., Hombhanje, F. W., Tsukahara, T., Hwaihwanje, I., & Lum, J. K. (2008). Reconstructing the origin of the Lapita Cultural Complex: mtDNA analyses of East Sepik Province, PNG. *Journal of Human Genetics*, *53*(8), 698–708. https://doi.org/10.1007/s10038-008-0301-3

Villarroel, P., Reyes, M., Fuentes, C., Segovia, M. P., Tobar, N., Villalobos, E., Martínez, J., Hugo, E., Ben-Jonathan, N., & Cifuentes, M. (2013). Adipogenic effect of calcium sensing receptor activation. *Molecular and Cellular Biochemistry*, *384*(1–2), 139–145. https://doi.org/10.1007/s11010-013-1791-7

Villarroel, P., Villalobos, E., Reyes, M., & Cifuentes, M. (2014). Calcium, obesity, and the role of the calcium-sensing receptor. *Nutrition Reviews*, *72*(10), 627–637. https://doi.org/10.1111/nure.12135

Vitti, J. J., Grossman, S. R., & Sabeti, P. C. (2013). Detecting natural selection in genomic data. *Annual Review of Genetics*, *47*, 97–120. https://doi.org/10.1146/annurev-genet-111212-133526

Wagner, J. K., Colwell, C., Claw, K. G., Stone, A. C., Bolnick, D. A., Hawks, J., Brothers, K. B., & Garrison, N. A. (2020). Fostering Responsible Research on Ancient DNA. *American Journal of Human Genetics*, *107*(2), 183–195. https://doi.org/10.1016/j.ajhg.2020.06.017

Walsh, S., Izquierdo-Serra, M., Acosta, S., Edo, A., Lloret, M., Moret, R., Bosch, E., Oliva, B., Bertranpetit, J., & Fernández-Fernández, J. M. (2020). Adaptive selection drives TRPP3 loss-of-function in an Ethiopian population. *Scientific Reports*, *10*(1), 1–16. https://doi.org/10.1038/s41598-020-78081-z

Wang, W., Kutny, M. P., Byers, L. S., Longstaff, J. C., DaCosta, J. M., Pang, C., Zhang, Y., Taft, R. A., Buaas, F. W., & Wang, H. (2016). Delivery of Cas9 protein into mouse zygotes through a series of electroporation dramatically increased the efficiency of model creation. *Journal of Genetics and Genomics*, 319–327. https://doi.org/doi:10.1016/j.jgg.2016.02.004

Weeber, C. (2020). Why Capitalize "Indigenous"? *Sapiens*, *1*, 3–5. https://www.sapiens.org/language/capitalize-indigenous/

Weigand, H., & Leese, F. (2018). Detecting signatures of positive selection in non-model species using genomic data. *Zoological Journal of the Linnean Society*, *184*(2), 528–583. https://doi.org/10.1093/zoolinnean/zly007

White, P. (2011). Revisiting the 'Neolithic Problem' in Australia. *Records of the Western Australian Museum, Supplement*, *79*(1), 86. https://doi.org/10.18195/issn.0313-122x.79.2011.086-092

Wohlers, L. M., & Spangenburg, E. E. (2010). 17β-estradiol supplementation attenuates ovariectomy-induced increases in ATGL signaling and reduced perilipin expression in visceral adipose tissue. *Journal of Cellular Biochemistry*, *110*(2), 420–427. https://doi.org/10.1002/jcb.22553

Zhang, X. H., Tee, L. Y., Wang, X. G., Huang, Q. S., & Yang, S. H. (2015). Off-target effects in CRISPR/Cas9-mediated genome engineering. *Molecular Therapy - Nucleic Acids*, *4*(11), 1–8. https://doi.org/10.1038/mtna.2015.37

ANNEX

ANNEX: supplementary material

Figure S1....166
Figure S2....168
Figure S3....170
Figure S4....171
Figure S5....172
Figure S6....173
Figure S8....176
Figure S9....180
Figure S10....184
Figure S11....185
Figure S12....186
Figure S13....187
Supplementary File S8....188

Supplementary Figure, Tables, and Files

All the supplementary tables and files can be accessed through the online resources [https://bit.ly/3RpkYHh]

Figure S7. Wound healing after a 5 mm punch skin biopsy**.** Wound healing images after a skin biopsy at 10 (A) or 20 weeks (B) of age (animals from line 7).

Table S1. Functional annotation of SNPs in the *CASR* gene region.

Table S2. Mosaic mice generated by CRISPR Cas9.

Table S3. Genes with potential off-target sites for each corresponding gRNAs used, as predicted by the Benchling designing tool, and primers used for Sanger sequencing.

Table S4. PAM, off-target site sequences, primers IDs, conditions, and nomenclature summary for all corresponding amplicon sequencing analysis.

Table S5. Primers for qPCR expression analysis and tissues analysed.

File S1. Selection statistics computed around the *CASR* gene region for Andamanese, Malaysian and Philippine rainforest hunter-gatherers, and Papua New Guinea populations.

File S2. Data and statistical analysis for weekly body weight and body mass index (BMI) analysis. The 990G genotype is indicated as HO, the R990 is indicated as WT.

File S3. Blood and urine biochemical data. The 990G genotype is indicated as HO, the R990 is indicated as WT.

File S4. Data and statistical analysis for body composition at week 14 (weight, fat and lean) and micro-architecture parameters of the tibial cortical bone and the trabecular bone at the proximal tibial metaphysis obtained from micro-computed tomography performed ex-vivo at endpoint (week 25). The 990G genotype is indicated as HO, the R990 is indicated as WT.

File S5. Wound Healing and Skin Biopsies data and analysis. The 990G genotype is indicated as HO, the R990 is indicated as WT.

File S6. Liver steatosis data and analysis. The 990G genotype is indicated as HO, the R990 is indicated as WT.

File S7. Visceral and subcutaneous adipocyte size data and statistical analysis. The 990G genotype is indicated as HO, the R990 is indicated as WT.

File S9. Gene expression results. The 990G genotype is indicated as HO, the R990 is indicated as WT.

File S10. DNA extraction protocol.

File S11. Alignment of sequences in off-target gene predicted sites as obtained from all line 7 and line 15 founders.

The QR code can be used for quick access with mobile devices to the online resources files and tables.

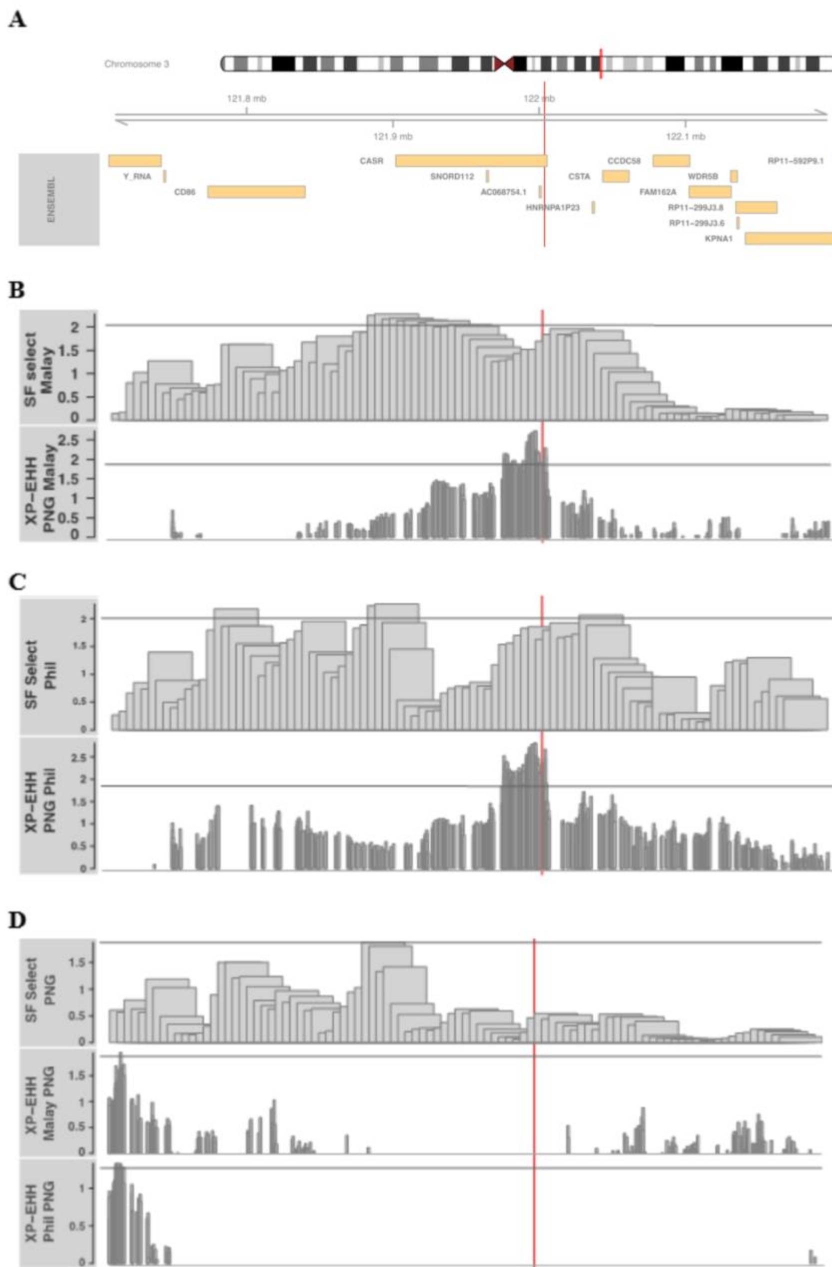


Figure S1.

Figure S1. Signals of positive selection in the *CASR* gene region. (A) On the top, the location of the *CASR* gene region on chromosome 3. Below, zoom in with Ensembl gene tracks within 200 Kb upstream and downstream of the *CASR* gene. (B) Tests of positive selection in the Malaysian DSWS group (SF select, and XP-EHH comparing Malaysian rainforest hunter-gathers versus the Papua New Guinea group). (C) Tests of positive selection in the Philippine DSWS group (SF select, and XP-EHH comparing Philippine rainforest hunter-gathers versus the Papua New Guinea group). (D) Tests of positive selection in the Papua New Guinea group (SF select, XP-EHH comparing the Papua New Guinea group versus Malaysian rainforest hunter-gathers, and XP-EHH comparing the Papua New Guinea group versus Philippine rainforest hunter-gathers). For each selection statistic, the horizontal bar indicates the top or bottom 5th percentile of their respective genome-wide distribution (z-scores are available in File S1). The position of the CaSR R990G polymorphism is represented by the red vertical line across the panels.

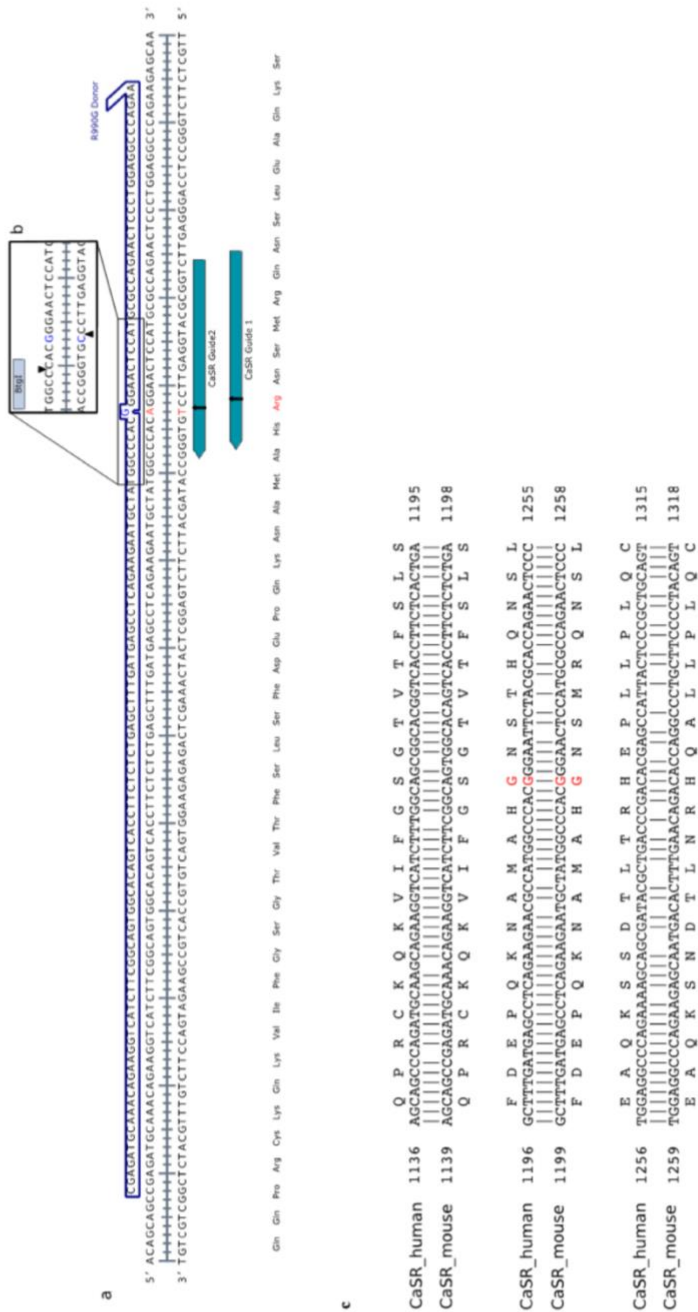


Figure S2.

Figure S2. Experimental design. (A) R990G donor carrying the G allele to generate the R990G substitution by homology-directed repair. The arrows within the CaSR guides indicate the Cas9 cleavage site. (B) Creation of a consensus *BtgI* recognition site for RFLP analysis. (C) Mouse-human nucleotide/aminoacid alignment for part of *CASR* exon 7 displaying in red the corresponding derived G/Gly allele of the CaSR R990G polymorphism.

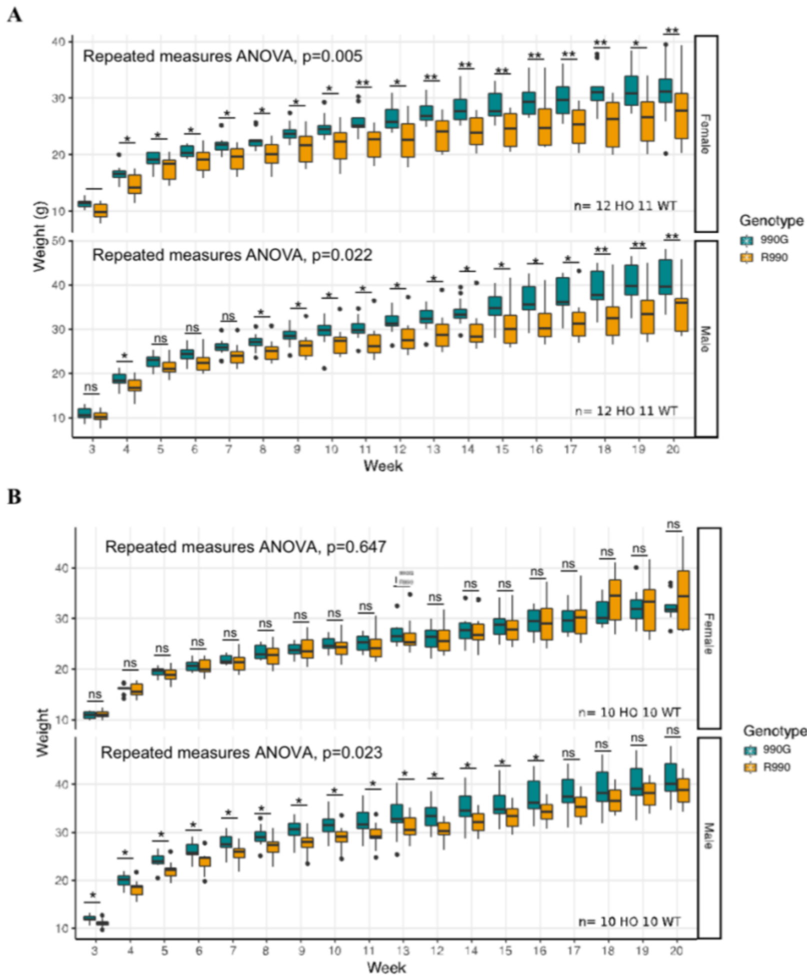


Figure S3. **Body weight growth curves of CaSR R990G knock-in mice.** (A) Boxplots showing the differences in body weight from 3 to 20 weeks of age between the homozygote carriers of the ancestral (abbreviated as HO and indicated in green) and derived (WT, in orange) alleles of the CaSR R990G polymorphism in line 7. * $p < 0.05$; ** $p < 0.005$; ns, not significant (one-way ANOVA). (B) Boxplots showing the differences in mice body weight from 3 to 20 weeks of age between the homozygote carriers of the derived (abbreviated as HO and indicated in green) and ancestral (WT, in orange) alleles of the CaSR R990G polymorphism in line 15. *$p < 0.05$; **$p < 0.005$; ns, not significant (one-way ANOVA). For all boxplots, the line in the middle of each box represents the median for each group examined, and the edges of the box represent the first and third quartiles. Filled circles represent outliers (values greater than 1.5 and less than 3 interquartile ranges from the edge of the box). Detailed significance levels for all tests are available in File S2.

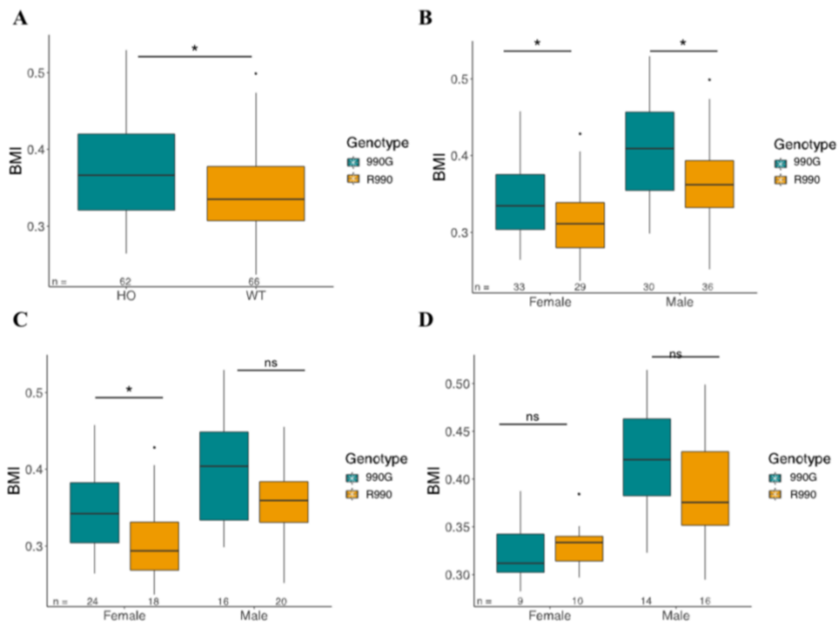


Figure S4. **Body Mass Index (BMI) differences between genotypes at the final point (weeks 20 - 26).** (A) BMI differences considering all animals together. (B) Body Mass Index (BMI) differences per sex. (C) Body Mass Index (BMI) differences per sex in line 7. (D) Body Mass Index (BMI) differences per sex in line 15. *p < 0.05 (Kruskal-Wallis test when analysing all animals together; otherwise Mann-Whitney U test). For all boxplots, the line in the middle of each box represents the median for each group examined, and the edges of the box represent the first and third quartiles. Filled circles represent outliers (values greater than 1.5 and less than 3 interquartile ranges from the edge of the box). Detailed significance levels for all tests are available in File S2.

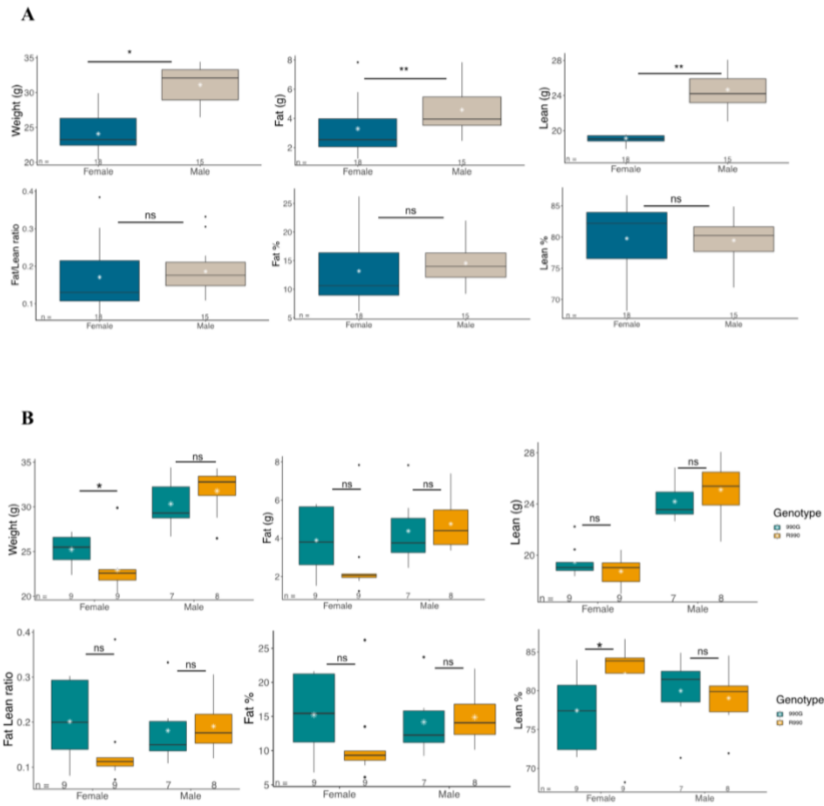


Figure S5. Body composition differences (weight, fat, and lean) between sexes and genotypes as determined after EchoMRI analysis at week 14. For all boxplots, the line in the middle of each box represents the median for each group examined, and the edges of the box represent the first and third quartiles. Filled circles represent outliers (values greater than 1.5 and less than 3 interquartile ranges from the edge of the box). Detailed significance levels are available in File S4.

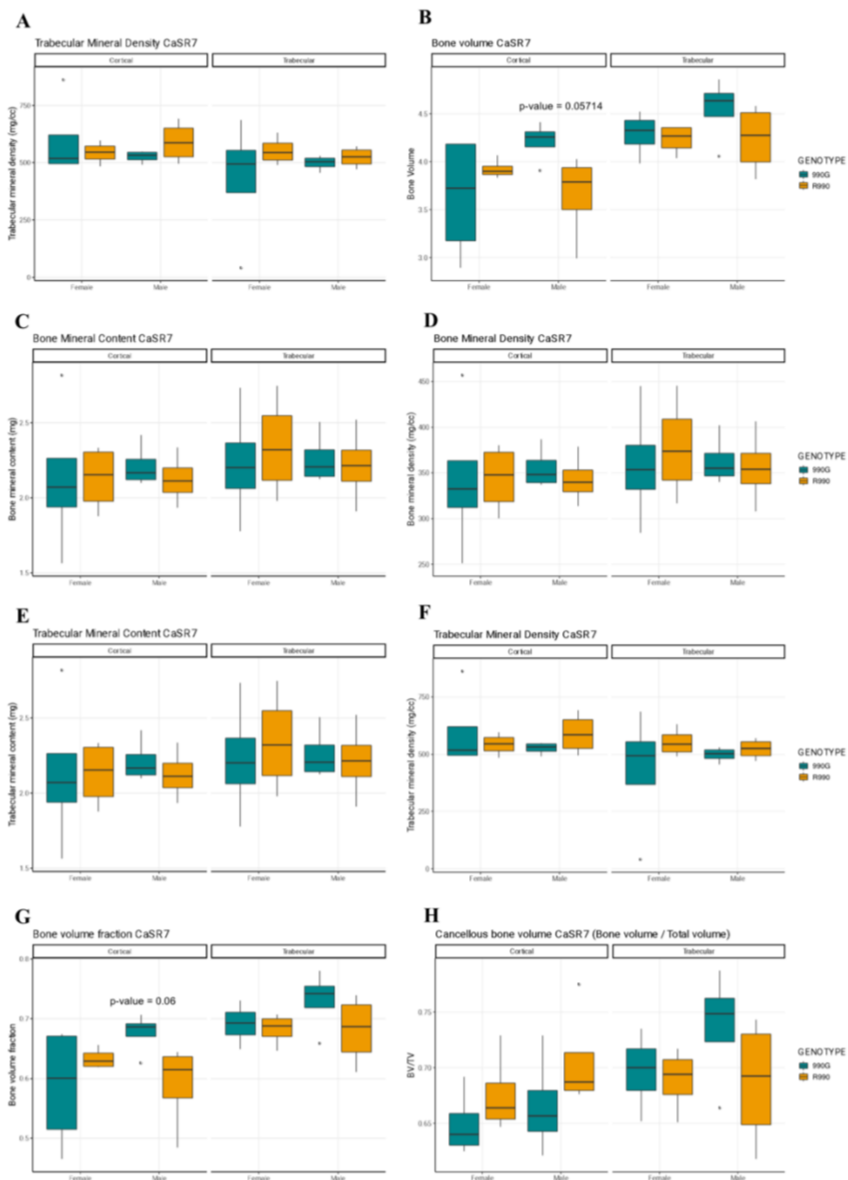
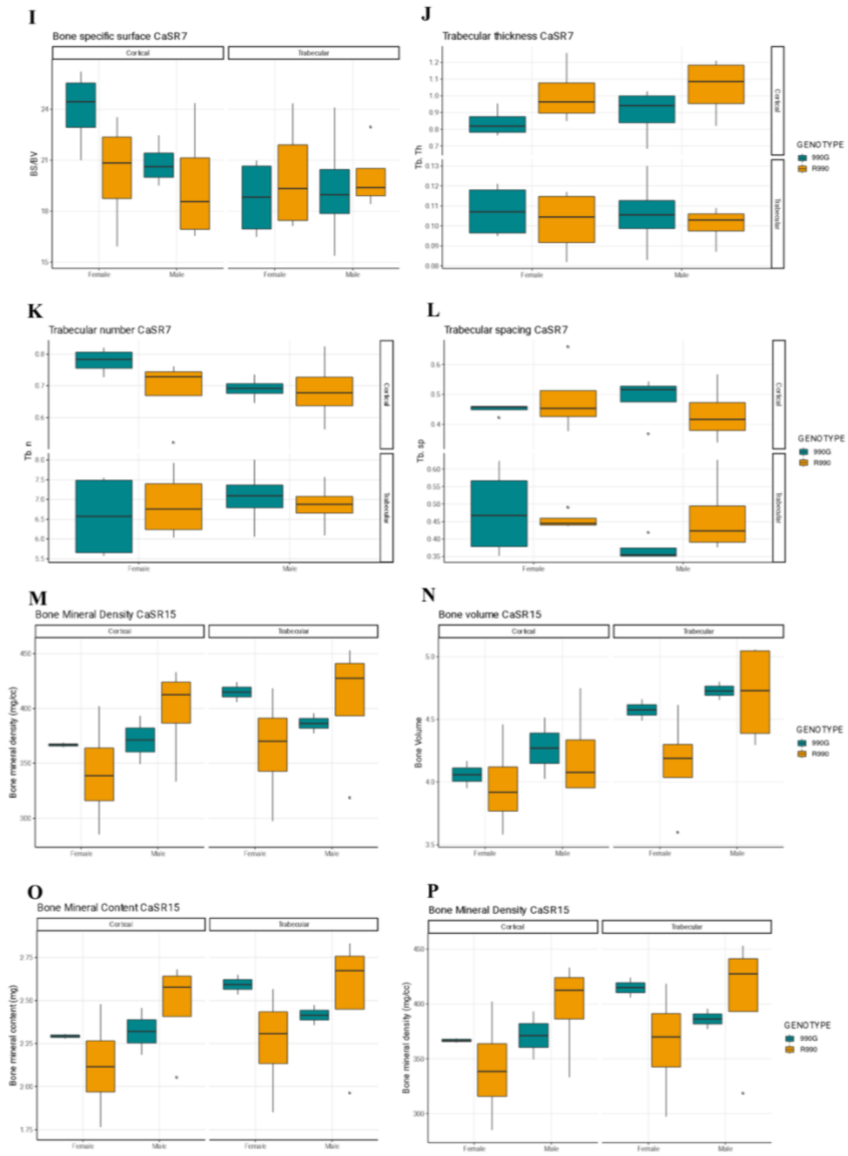


Figure S6.


I
Bone specific surface CaSR7
Cortical
Trabecular
BS/BV
Female
Male
J
Trabecular thickness CaSR7
Tb. Th
GENOTYPE
990G
R990
K
Trabecular number CaSR7
Tb. n
L
Trabecular spacing CaSR7
Tb. sp
M
Bone Mineral Density CaSR15
Bone mineral density (mg/cc)
N
Bone volume CaSR15
Bone Volume
O
Bone Mineral Content CaSR15
Bone mineral content (mg)
P
Bone Mineral Density CaSR15

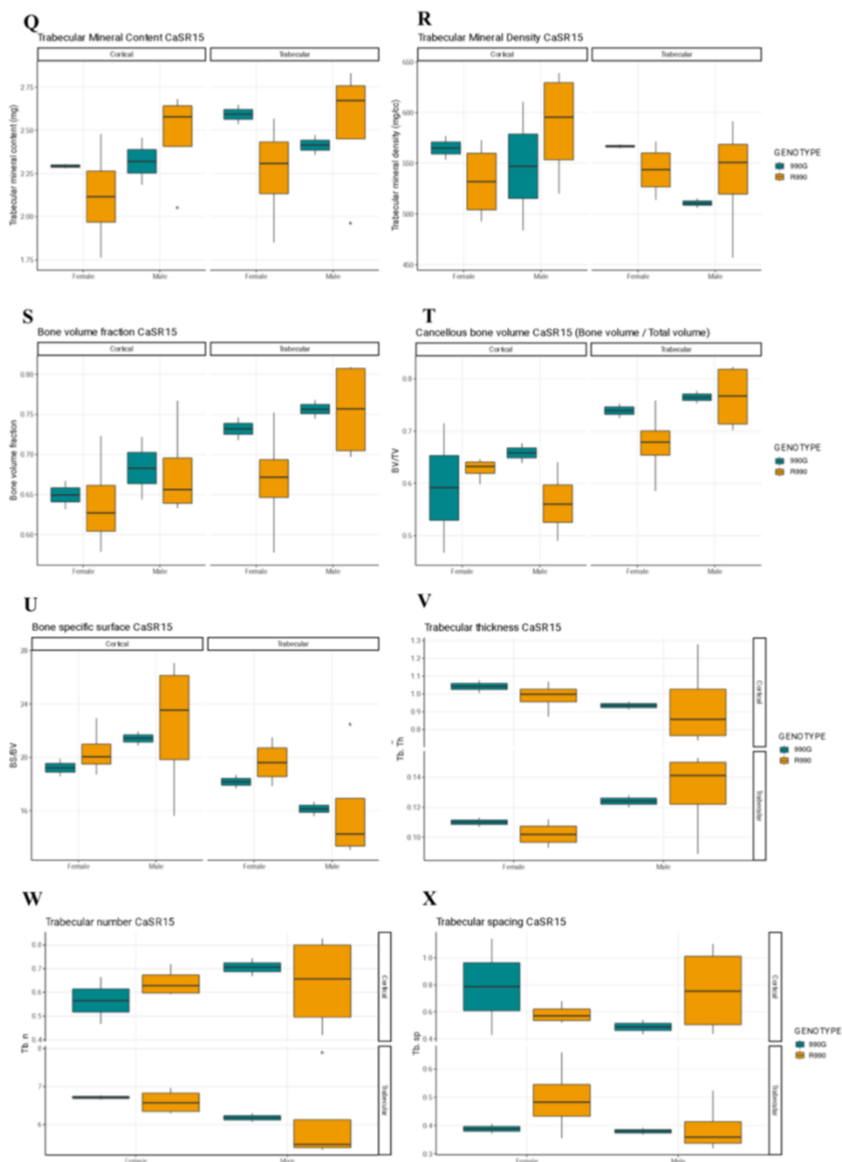
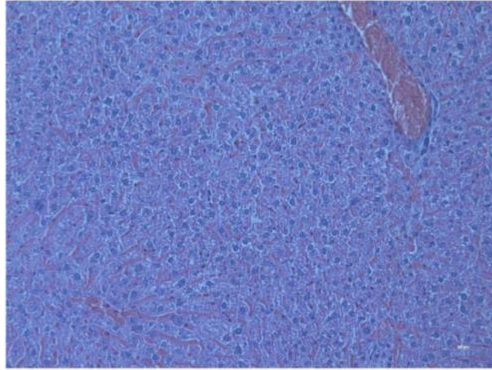
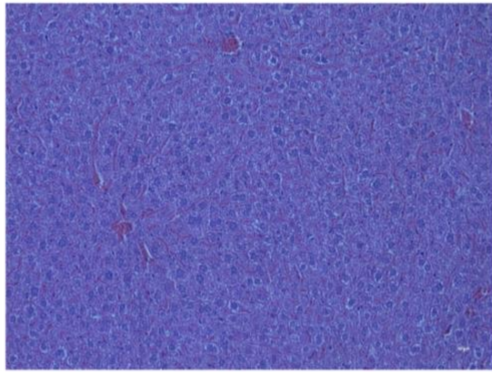


Figure S6. Micro-architecture parameters of the tibial cortical bone and the trabecular bone at the proximal tibial metaphysis obtained from micro-computed tomography performed *ex-vivo* at endpoint (week 25). For all boxplots, the line in the middle of each box represents the median for each group examined, and the edges of the box represent the first and third quartiles. Filled circles represent outliers (values greater than 1.5 and less than 3 interquartile ranges from the edge of the box). No comparison between genotypes was significant (Wilcoxon Rank sum test).

A





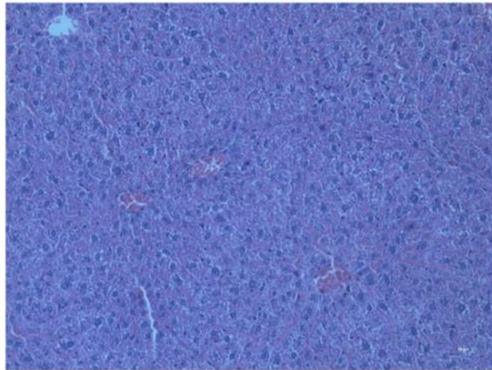
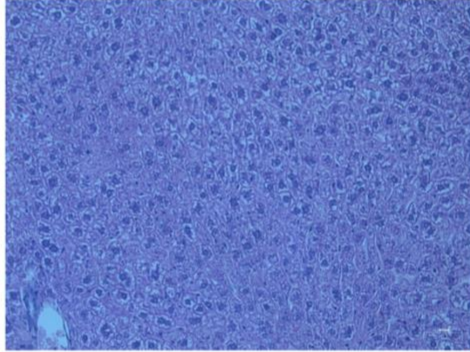
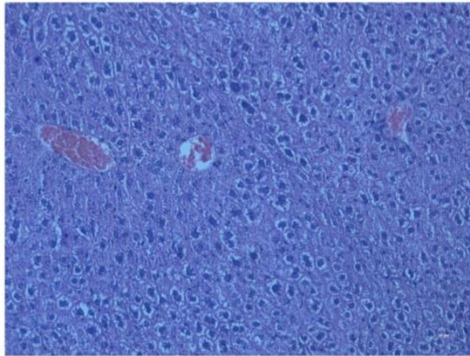
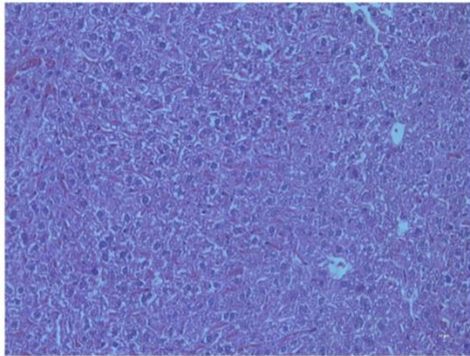


Figure S8.

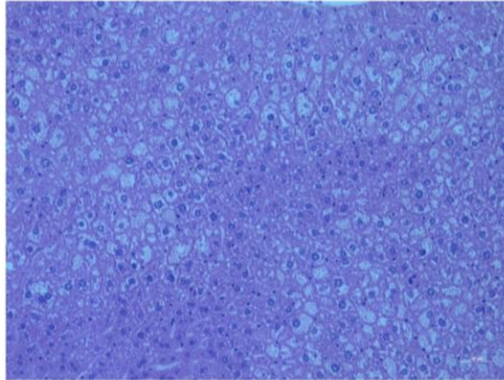
B

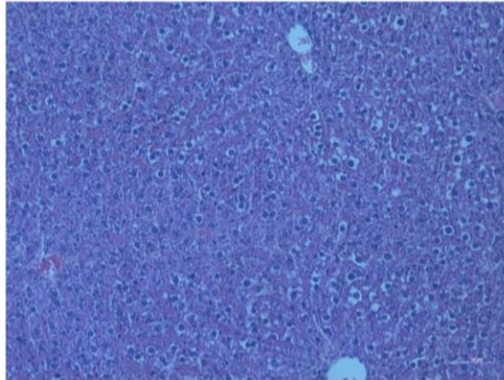


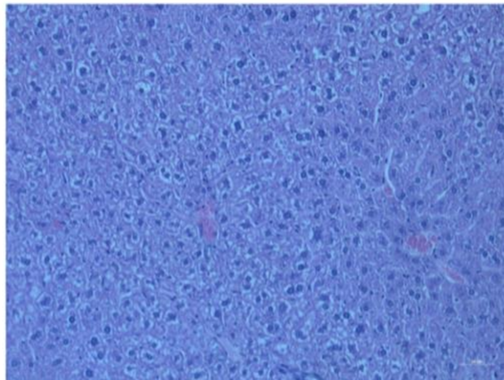




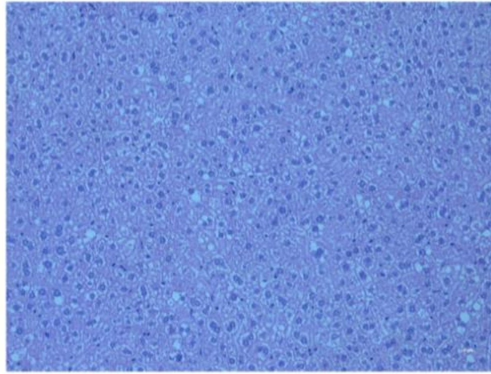
C

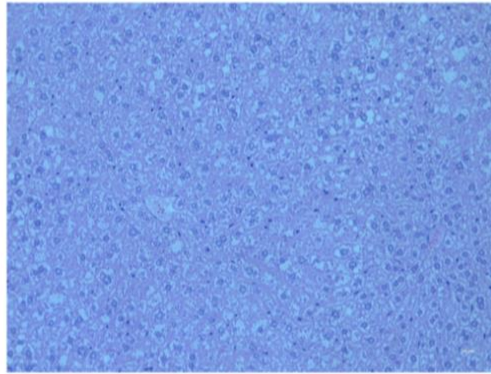






D





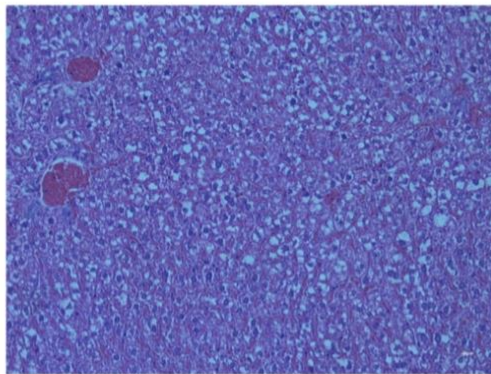


Figure S8. Liver steatosis categorization. Liver biopsies were categorized into a normal or a steatosis phenotype, depending on whether they presented normal (A) to mildly phenotype with lipid droplets (B) or moderate (C) to severe fatty liver (D), respectively.

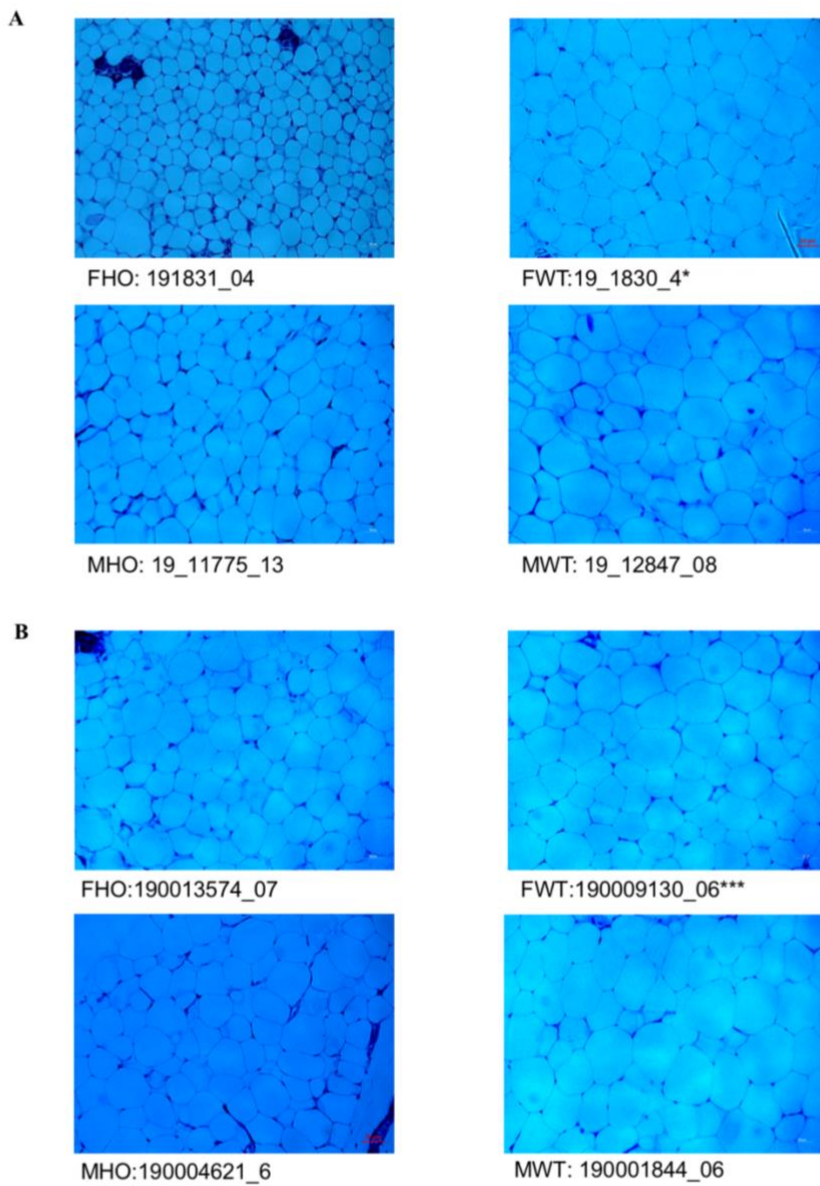
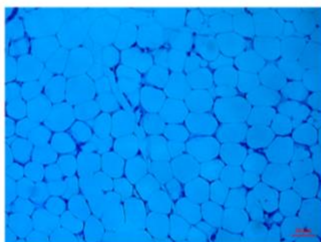
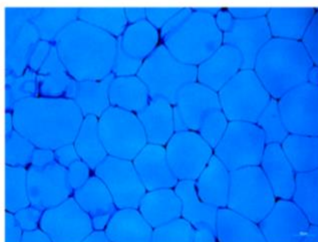


Figure S9.

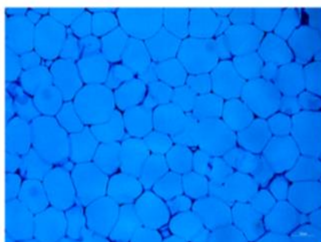
C



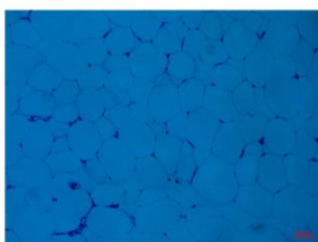
FHO: 19_978_2



FWT: 19_7217_07

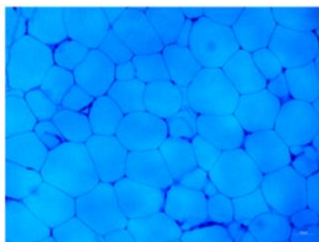


MHO: 19_11774_02

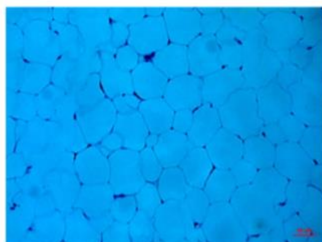


MWT: 19_970_5

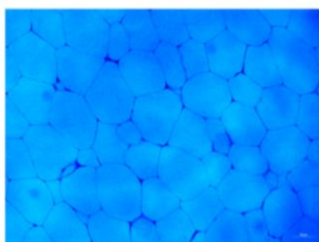
D



FWT: 190009130_08

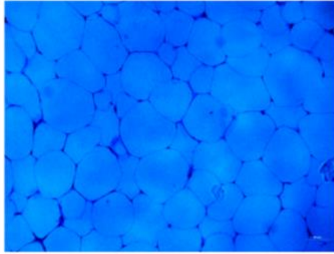


MHO: 190004621_12

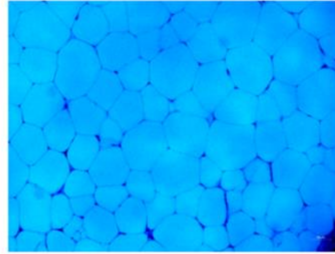


MWT: 190001844_09

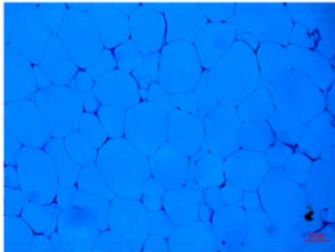
E



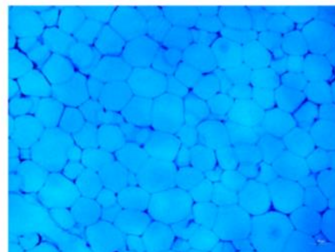
FHO: 190014859_12



FWT: 190009129_04

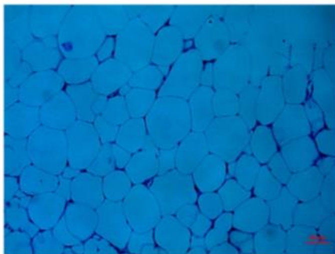


MHO: 190008028_13

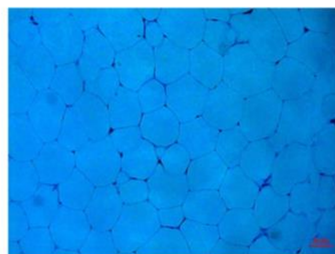


MWT: 190004622_06

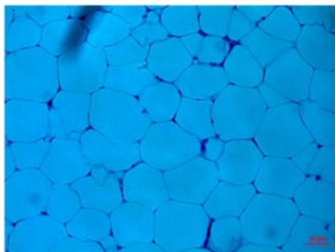
F



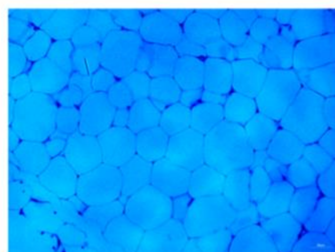
FHO: 190008031_01



FWT: 190004624_3**



MHO: 190008028_06



MWT: 190004620_05

Figure S9. Histology of subcutaneous and visceral fat. (A) Subcutaneous fat in line 7 individuals without liver steatosis. (B) Subcutaneous fat in line 15 individuals without liver steatosis. (C) Visceral fat in line 7 individuals without liver steatosis. (D) Visceral fat in line 15 individuals without liver steatosis. (E) Subcutaneous fat in line 15 individuals with liver steatosis. (F) Visceral fat in line 15 individuals with liver steatosis. FHO, female homozygote for the derived allele of the R990G substitution; FWT, female homozygote for the ancestral allele of the R990G substitution; MHO, male homozygote for the derived allele of the R990G substitution; MWT, male homozygote for the ancestral allele of the R990G substitution

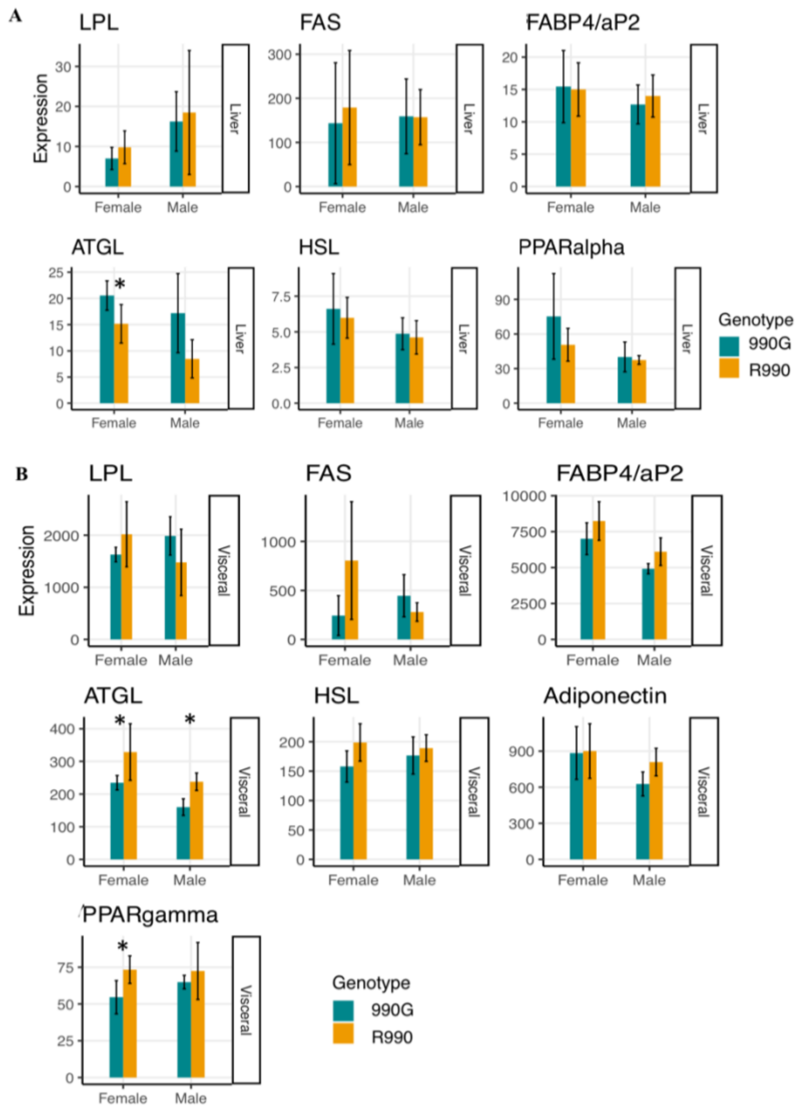


Figure S10. Graphical representation of gene expression differences between the genotypes of the R990G substitution in (A) Liver and (B) Visceral adipose tissue (line 7, liver without steatosis).

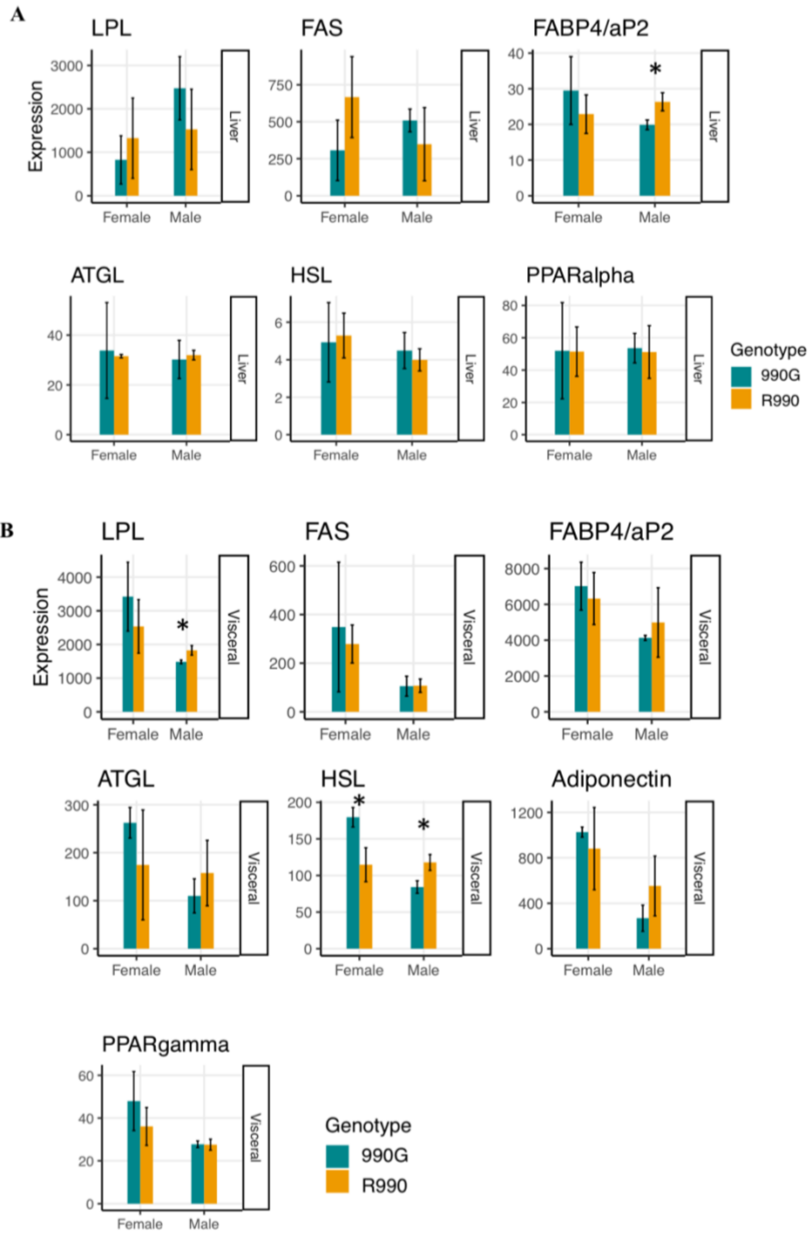


Figure S11. Graphical representation of gene expression differences between the genotypes of the R990G substitution in (A) Liver and (B) Visceral adipose tissue (line 15, liver with steatosis).

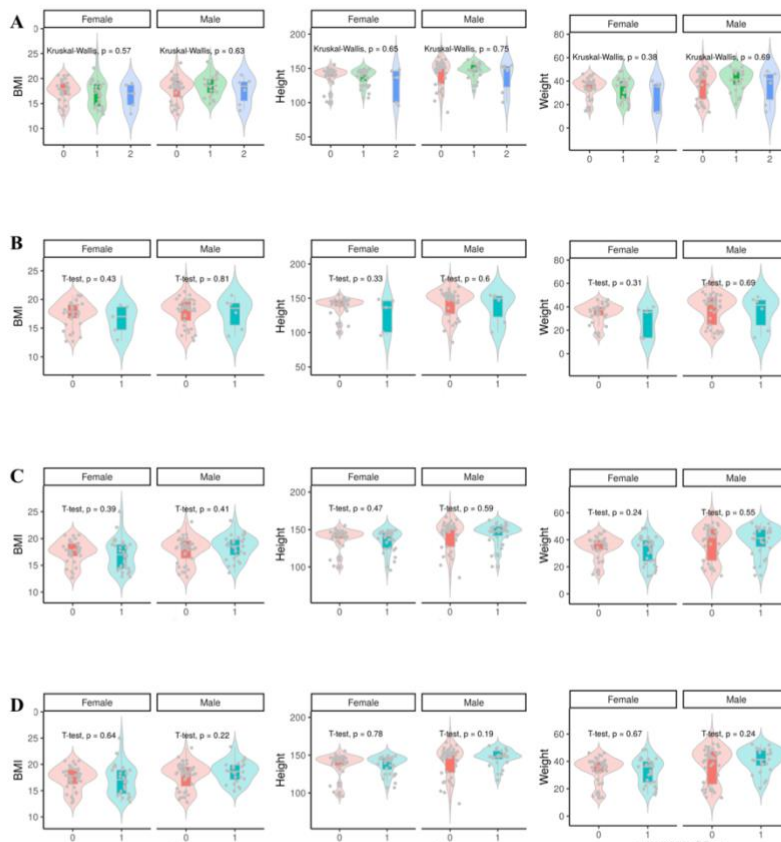


Figure S12. BMI, height and weight phenotype distribution by sexes in the Agta population. (A) Phenotypical differences between genotypes tested according to an additive model (0 = GG, 1 = GA, 2 = AA), (B) an homozygote only model (0 = GG, 1 = AA), (C) a recessive model (0 = GG, 1 = GA/ AA) or (D) an over-dominant model (0 = GG/AA, 1 = GA).

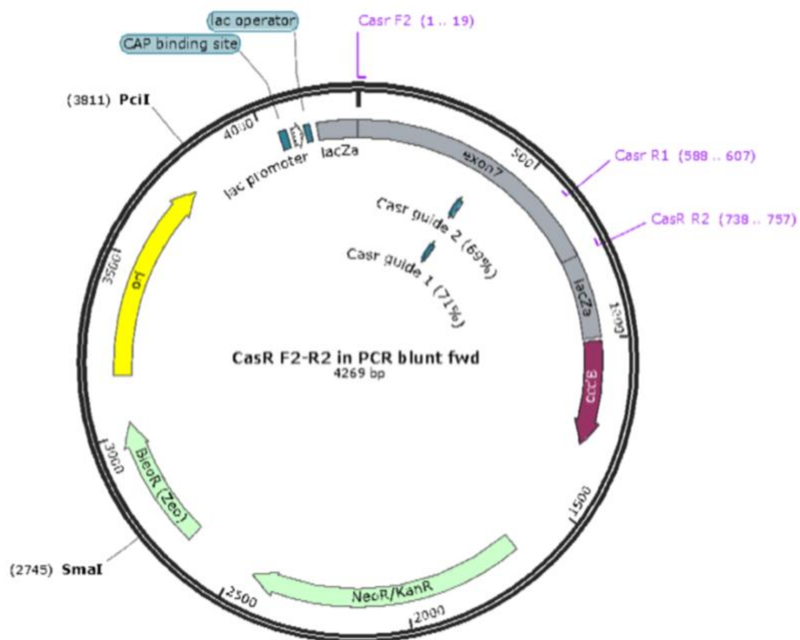
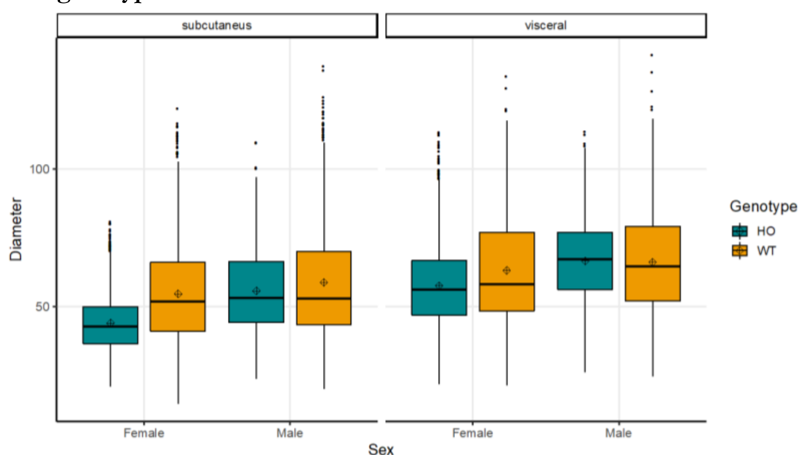


Figure S13. PCR blunt vector used for cloning with neomycin and kanamycin resistance.

Supplementary File S8: Visceral and subcutaneous adipocyte size results. The 990G genotype is indicated as HO, the R990 is indicated as WT

File S8 Figure 1. Analysis of adipocyte diameter and adipocyte size distribution in animals without hepatic steatosis (line 7)

S8Fig1A. Boxplots of subcutaneous and visceral adipocyte diameters by sex and genotype.



S8Fig1B. Two samples T-tests comparing adipose tissue cell's diameters between R990G genotypes per tissue and sex.

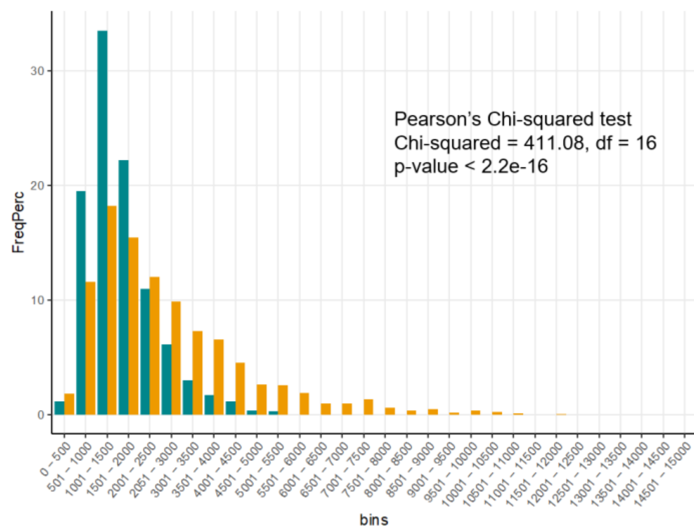
Sex	Genotype	Tissue	n[1]	n[2]	mean	sd	t-test	df	p-value
Female	HO	Subcutaneous fat	3	1465	44,071	± 10,524	-20,722	2974,1	**< 2.2e-16**
Female	WT	Subcutaneous fat	5	1808	54,643	± 18,273			
Male	HO	Subcutaneous fat	4	1231	55,753	± 15,147	-3,94	1994,9	**8,44E-05**
Male	WT	Subcutaneous fat	4	1113	58,774	± 21,145			
Female	HO	Visceral fat	5	1676	57,616	± 15,409	-7,972	2177,1	**2,50E-15**
Female	WT	Visceral fat	5	1227	63,163	± 20,502			
Male	HO	Visceral fat	4	1032	66,641	± 15,128	0,652	2362,1	0,514
Male	WT	Visceral fat	5	1335	66,185	± 18,899			

n[1]: number of animals

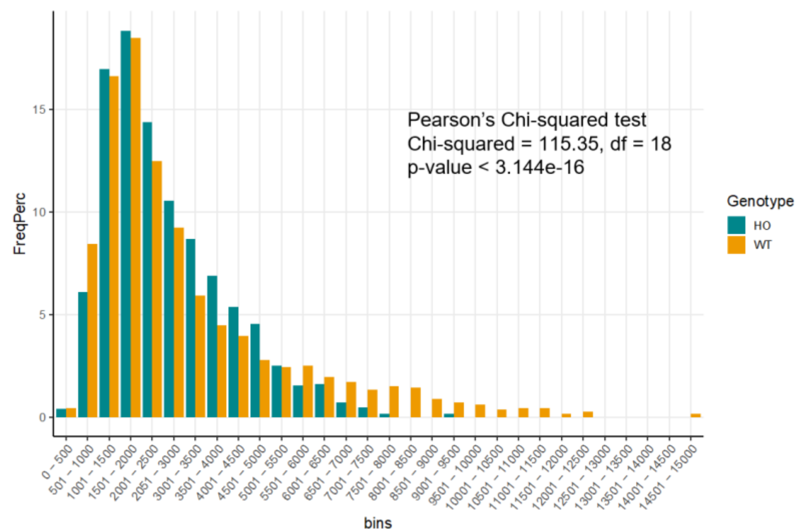
n[2]: number of cells

S8Fig1C. Adipocyte size distribution analysis.

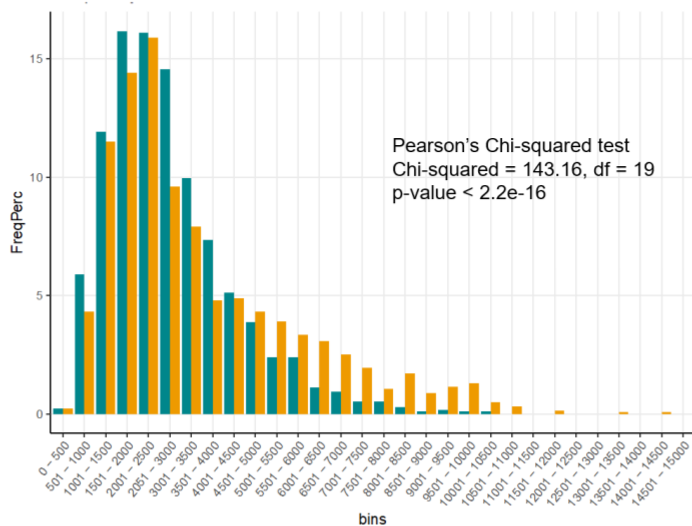
S8Fig1C1. Subcutaneous, females



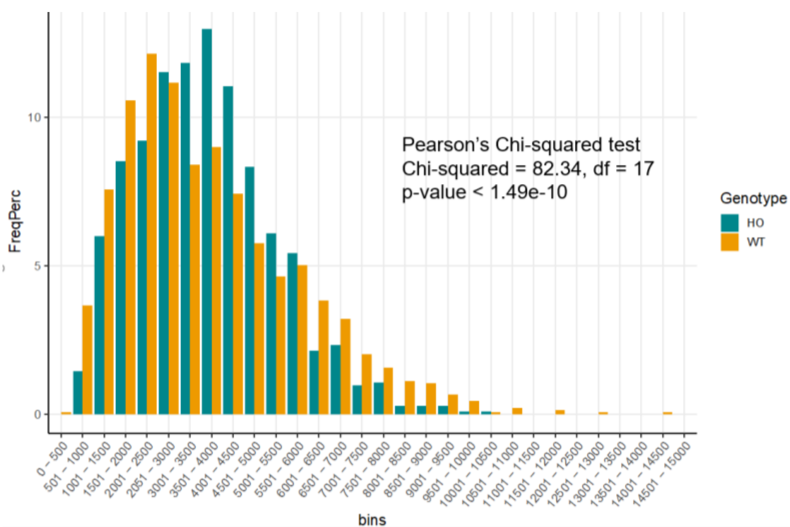
S8Fig1C2. Subcutaneous, males



S8Fig1C3. Visceral, females

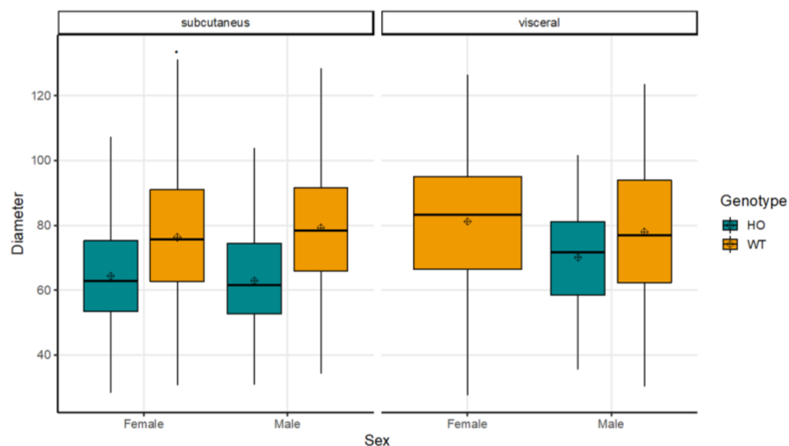


S8Fig1C4. Visceral, males



File S8 Figure 2. Analysis of adipocyte diameter and adipocyte size distribution in animals without hepatic steatosis (line 15)

S8Fig2A. Boxplots of subcutaneous and visceral adipocyte diameters by sex and genotype.



S8Fig2B. Two samples T-tests comparing adipose tissue cell's diameters between R990G genotypes per tissue and sex.

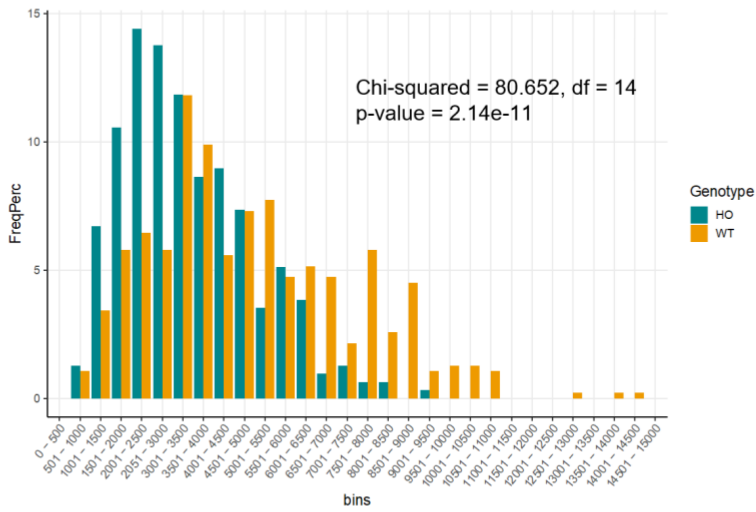
Sex	Genotype	Tissue	n[1]	n[2]	mean	sd	t-test	df	p-value
Female	HO	Subcutaneous fat	1	312	64,397	15,1159	-9,455	763,36	**< 2.2e-16**
Female	WT	Subcutaneous fat	2	465	76,303	19,9215			
Male	HO	Subcutaneous fat	1	287	62,93	15,5278	-11,945	626,92	**< 2.2e-16**
Male	WT	Subcutaneous fat	2	342	79,245	18,7259			
Female	HO	Visceral fat	0	0	NA	NA			
Female	WT	Visceral fat	2	428	81,223	18,7859			
Male	HO	Visceral fat	1	175	70,068	14,3142	-4,6726	411,61	**4,04E-06**
Male	WT	Visceral fat	2	239	77,992	20,19			

n[1]: number of animals

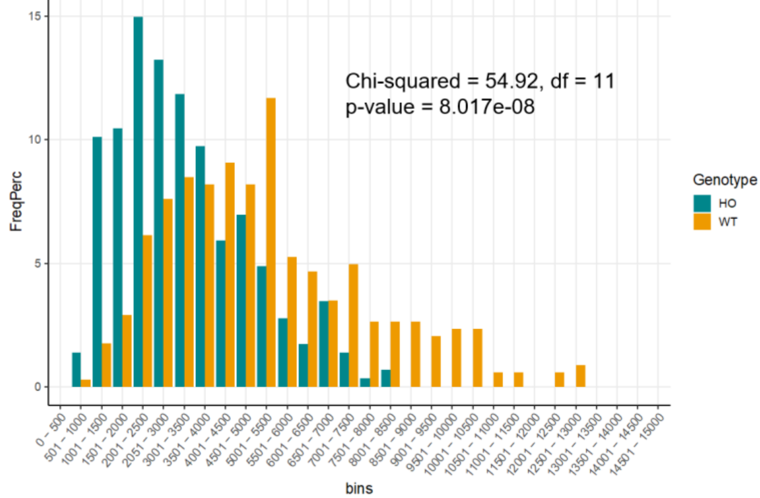
n[2]: number of cells

S8Fig2C. Adipocyte size distribution analysis.

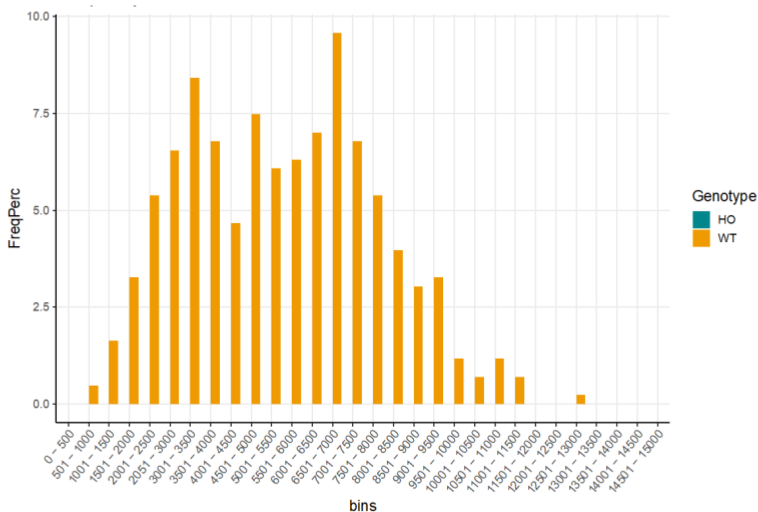
S8Fig2C1. Subcutaneous, females



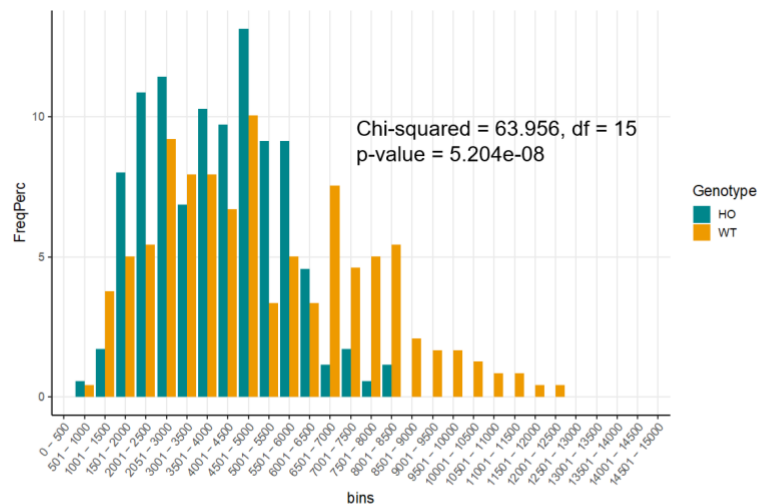
S8Fig2C2. Subcutaneous, males



S8Fig2C3. Visceral, females

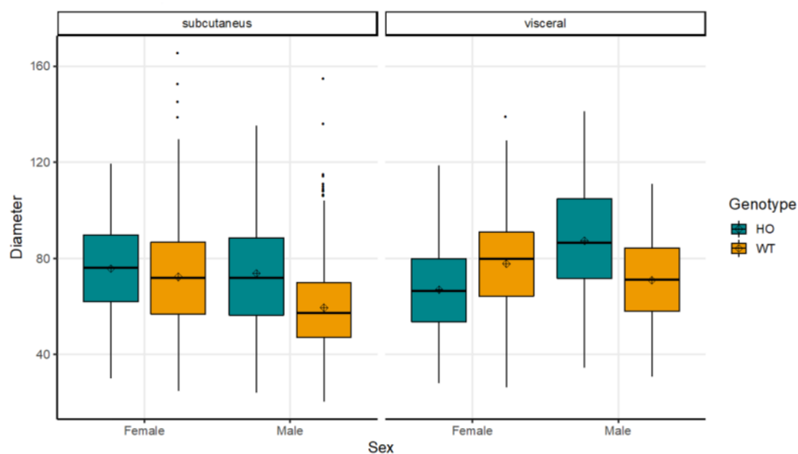


S8Fig2C4. Visceral, males



File S8 Figure 3. Analysis of adipocyte diameter and adipocyte size distribution in animals with hepatic steatosis (line 15)

S8Fig3A. Boxplots of subcutaneous and visceral adipocyte diameters by sex and genotype.



S8Fig3B. Two samples T-tests comparing adipose tissue cell's diameters between R990G genotypes per tissue and sex.

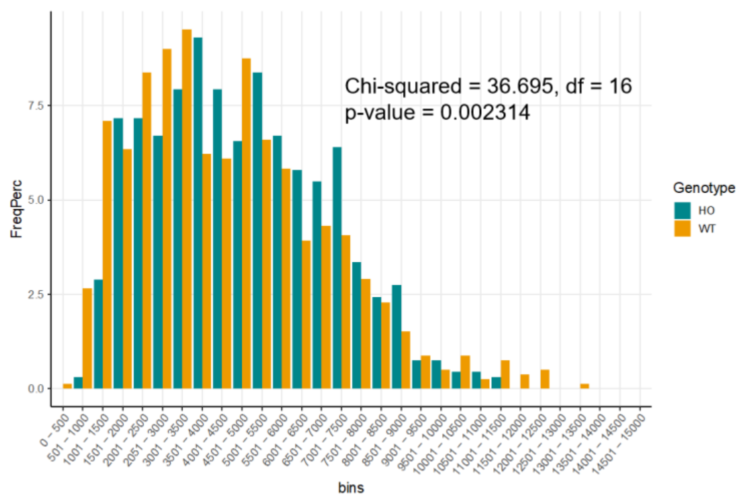
Sex	Genotype	Tissue	n[1]	n[2]	mean	sd	t-test	df	p-value
Female	HO	Subcutaneous fat	3	656	75,706	18,1448	3,2706	1445	**0,001098**
Female	WT	Subcutaneous fat	3	792	72,312	21,3481			
Male	HO	Subcutaneous fat	2	529	73,747	21,9999	13,01	870,95	**< 2.2e-16**
Male	WT	Subcutaneous fat	3	1105	59,48	17,8266			
Female	HO	Visceral fat	3	769	66,977	17,9761	-10,716	1351,4	**< 2.2e-16**
Female	WT	Visceral fat	3	667	77,796	19,9898			
Male	HO	Visceral fat	2	362	87,37	22,8112	12,128	567,78	**< 2.2e-16**
Male	WT	Visceral fat	3	770	70,942	17,4907			

n[1]: number of animals

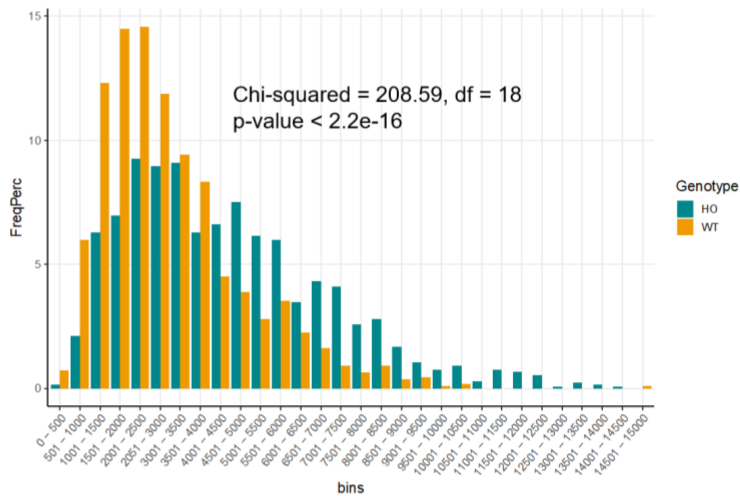
n[2]: number of cells

S8Fig3C. Adipocyte size distribution analysis.

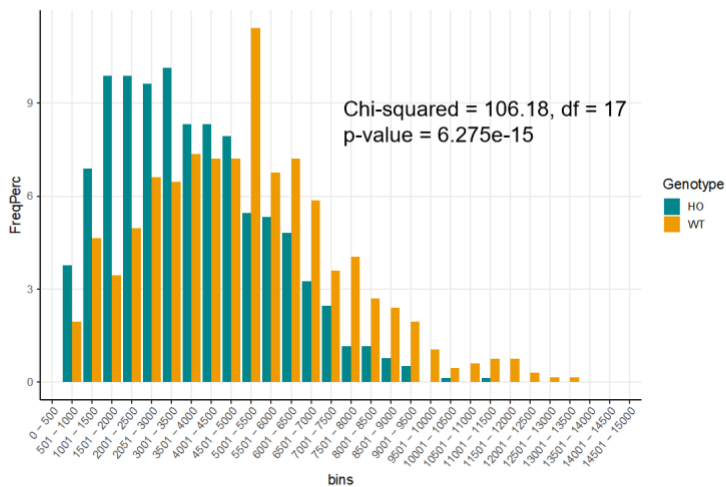
S8Fig3C1. Subcutaneous, females



S8Fig3C2. Subcutaneous, males



S8Fig3C3. Visceral, females



S8Fig3C4. Visceral, males