

# Gizakiaren aurpegiaren ezaugarri fenotipikoei loturiko aldakortasun genetikoaren analisisa

(Analysis of the genetic variability linked to the phenotypic characteristics of the human face)

Amaia Cabrejas, Santos Alonso, Neskuts Izagirre\*

Genetika, Antropologia Fisikoa eta Animalien Fisiologia Saila.  
Zientzia eta Teknologia Fakultatea. UPV/EHU. Leioa, Bizkaia

**LABURPENA:** Aurpegiaren morfologia ezinbestekoa da banakoak euren artean ezberdindu ahal izateko. Lan honetan, aurpegiko ezaugarrietan eragina duten SNPak aztertu ditugu. Populazioen historia demografikoak aurpegiaren morfologiaren aldakortasunaren proportzio handia azaltzen duela aurkitzen dugu. Hala ere, aurpegiaren aldakortasunean hautespen sexuala bezalako beste faktore batzuek ere eragina izan dezakete, nahiz eta maila txikiagoan. Sudurraren morfologian arreta jarritz, Asia ekialdeko eta Afrikako populazioek sudur zabalagoa dute oro har, eta europarrek estuagoa. Bestalde, aurpegiaren ezaugarri loturiko SNPen azterketa auzitegi-zientzietan ere erabilgarria izan daiteke, nahiz eta gaur egun DNAtik abiatuta banako baten identifikazio fenotipikoa konplexua den. Lan honen helburua da giza populazioak aurpegiaren morfologiari loturiko datu genetikoetatik (SNPak) abiatuta desberdintzea posible den egiaztatzea.

**HITZ GAKOAK:** SNP, aurpegi morfologia, historia demografikoa.

**ABSTRACT:** Facial morphology is of great value for individual identification. In our project, we studied the geographical distribution of SNPs that influence facial morphology. Our PCA analysis shows that it is difficult to disentangle facial morphology from demographic history of mayor human populations. From PRS statistics, we show that regarding nasal morphology, populations of East Asia and Africa seems to have a wider nose, whereas populations of Europe show a narrower nose. These results can be explained as a climatic adaptation. The investigation of facial morphology related SNPs could also be useful in forensic sciences; however, the phenotypic identification of morphological traits of an individual is complex and would require population-specific SNPs associated to anthropometric traits.

**KEYWORDS:** SNP, facial morphology, demographic history.

\* **Harremanetan jartzeko / Corresponding author:** Neskuts Izagirre. Genetika, Antropologia Fisikoa, Animalien Fisiologia Saila. Zientzia eta Teknologia Fakultatea. – neskuts.izagirre@ehu.eus – <https://orcid.org/0000-0002-3345-9825>

**Nola aipatu / How to cite:** Cabrejas, Amaia; Alonso, Santos; Izagirre, Neskuts (2023). «Gizakiaren aurpegiaren ezaugarri fenotipikoei loturiko aldakortasun genetikoaren analisisa». *Ekaia*, 44, 2023, 291-317. (<https://doi.org/10.1387/ekaia.22949>).

Jasotze-data: 2021, ekainak 29; Onartze-data: 2022, azaroak 9.

ISSN 0214-9001 - eISSN 2444-3255 / © 2023 UPV/EHU

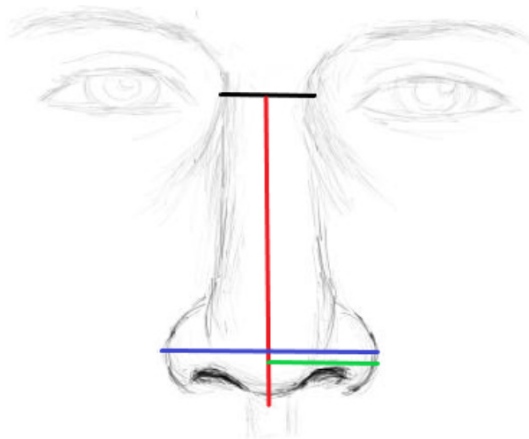


Lan hau Creative Commons Aitortu-EzKomertziala-LanEratorririkGabe 4.0 Nazioartekoa lizentzia baten mende dago

## 1. SARRERA

Gizakion aurpegiaren forma berariazkoa da pertsona bakoitzean, eta oso aldakorra populazioetan zehar. Hori dela eta, paper esanguratsua jotzen du banakoan identifikazioan, eta, horrenbestez, zientzia forentsean ezinbestekoa da [1]. Baina ageriko ezaugarria izateaz gain, aurpegiaren morfologiaren oinarri genetikoak ukaezina da; bai herentziazko zenbait ikerketek, bai familien arteko antzekotasun fenotipiko handiek hala frogatzen dute [2]. Beraz, posible litzateke DNA aztertuz hainbat ezaugarri morfologikoren fenotipoa berreraikitzea. Hori dela eta, genetika forentsean gero eta ikerketa gehiago egiten ari dira, arlo honek ekar ditzaken aurrerapenak ikaragarriak izan daitezkelako.

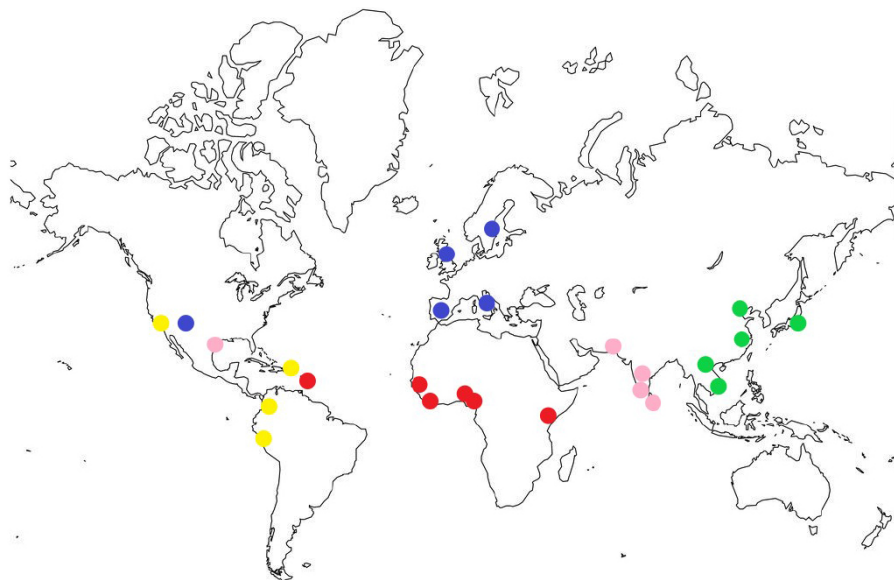
Hala ere, populazio baten barnean aurpegiaren morfologiari lotutako ezaugarri zehatz batzuk nabarmenagoak dira beste populazio batzuekin alderatuta, bakoitzak eboluzio ezberdin bat izan duelako. Horrela, nahiz eta denak duela  $\approx 300.000$  urte Afrikan sorturiko hominini talde baten ondorengoak izan [3], Afrikatik irteterakoan norabide ezberdinak hartu zituzten [4]. Ondorioz, *Homo sapiens* espezieko populazioak, neurri bateraino, euren artean desberdinduz joan dira hainbat ezaugarritan gaur egun arte. Desberdintze-prozesu horretan, aurpegiaren morfologia ez zen salbuespena izan, eta populazio ezberdinetan bestelako ezaugarriak nabarmendu ziren.



**1. irudia.** 1. taulako SNPek eragina duten ezaugarrien irudikapena. Beltzez, sudur zubiaren zabalera; gorritz, sudurraren luzera; urdinez, sudurraren zabalera; eta berdez, sudur-hegalen luzera.

Haatik, ez da ezagutzen aurpegiaren morfologiaren eragile-desberdintzaile zuzenik. Alde batetik, aurpegiaren morfologiaren zenbait ezaugarri

moldapenak izan daitezke. Horietariko moldapen batzuk hautespen naturalak faboratu ditu, gizakiak ingurune berrien kolonizazioan aurkitutako egoera berriak gailentzeko, klimarekiko egokitzapenak besteak beste [6]. Horretaz gain, hautespen sexualaren ondorio ere izan liteke, zeinak gainerrako primateen morfologian izugarritzko presioa eragin duen eboluzioan zehar [5]. Eta beste batzuk zoriz desberdinu zitezkeen. Ahal den heinean, gure lanean, 1. irudian azaltzen diren sudurarraren ezaugarri morfologikoe-kin erlazionatzen diren SNP eragileen banaketa 2. irudian agertzen diren populazioetako 1000 Genoma Proiektuan (1000G) analizaturiko gizabana-koetan aztertu da.



**2. irudia.** Lan honetarako erabili ditugun populazio taldeen irudikapena munduko mapan, jatorri geografikoaren arabera (Material Osagarrien 2. taula): AFR gorritz, AMR, horiz, EAS berdez, EUR urdinez eta SAS arrosaz.

Arlo forentsearen harira, krimen-eszenetan lortutako lagin biologikoe-tatik erauzitako DNAtik abiatuz, material genetiko horren jabe den pertso-naren aurpegia nolakoa den jakiteko gai izango bagina, asko erraztuko lira-teke identifikazioak. Hau da, hain zuzen ere, genetika forentsearen arloan azken urteetan ikertzen ari dena. Ikerketa hauetan, SNPen (Single Nucleo-tide Polymorphism) garrantzia azpimarragarria da [7].

SNPak nukleotido bakarreko polimorfismoak dira, hots, genomaren base bakar batean eragina duen aldaketa. Zenbait ikertzailek SNP ezberdi-nek aurpegiko zer ezaugarritan duten eragina identifikatu dute [8-12]. Lan

honetan, SNP horien efektuzko aleloak munduko populazioetan nola banatzen diren jakinda, PRS (Poligenic Risk Score) [13-14] balioaren bidez, populazio ezberdinetan aurpegiaren eta, adibide zehatz moduan, sudurraren morfologiari loturiko ezaugarriak aztertuz, aztertu dugu ea populazioak bereizten diren.

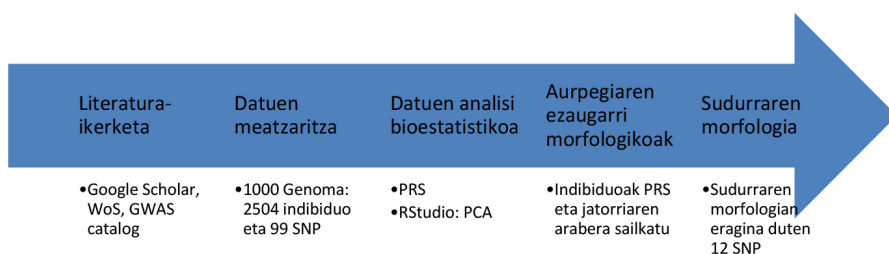
### 1.1. Helburua

Lan honen helburua da giza populazioak aurpegiaren morfologiari loturiko datu genetikoetatik (SNPak) abiatuta desberdintzea posible den egiaztatzea.

## 2. MATERIALAK ETA METODOAK

### 2.1. Literatura-ikerketa

3. irudian, lan honetan jarraitu den lan-fluxua adierazten da. Lehenik eta behin, bilaketa bibliografikoa egin dugu, Google Scholar (<https://scholar.google.es/>), NCBI (<https://www.ncbi.nlm.nih.gov/>) eta WoS (<https://www.recursoscientificos.fecyt.es/>) plataformak erabiliz. Hitz gakoak «SNP» eta «facial morphology» izan dira. Lortutako artikuluetatik SNPa, kokapen kromosomikoa, alelo eragilea («risk allele»), p-balioa,  $\beta$ -balioa eta asoziatutako aurpegiko ezaugarriarekin taula bat osatu dugu (Material Osagarrien 1. taula). Soilik  $p < 0,05$  duten SNPak hautatu dira. Horretaz gain, artikuluetan falta den informazioa betetzeko, gehienetan SNPen alelo eragilea eta  $\beta$ -balioa, GWAS catalog (<https://www.ebi.ac.uk/gwas/>) [15] datu-basea erabili dugu.



### 3. irudia. Analisisian jarraitutako lan-fluxuaren eskema.

### 2.2. Datuen meatzaritzza

1000-Genoma proiektuko 3. faseko datu-basea osatzen duten 2.504 lagunetan (<https://www.internationalgenome.org/>), hasieran aukeratutako SNPen genotipoak aztertu ditugu VCFtools paketearen bidez [16] eta posible izan

dugu 99 SNPetarako genotipoen informazioa aurkitzea. Horiek dira, hain zuzen ere, lan hau egiteko erabili izan direnak (Material Osagarrian 1. taula). Horiei erreferentzia egiteko «test-SNP» terminoa erabiliko dugu.

Aurpegiaren morfologiarekin erlazionaturiko SNPen genotipoak jasozten dituen VCF (Variant Call Format) fitxategi bat eraiki dugu 1000 Genomako indibiduen informazio genetiko laburbiltzen duena (Material Osagarrien 2. taula). Banako horiek euren jatorriaren arabera populazio eta azpipopulazioetan sailkatu dira (2. irudia).

### 2.3. Datuen analisi bioestatistikoa

Aurreko datuak abiapuntutzat hartuta, lehenik, indibiduen PRS balioa kalkulatu dugu. PRS balioak pertsona batek ezaugarri edo gaixotasun bat aurkezteko duen arrisku genetiko estimatzen du, indibiduoaren genotipoaren eta Genome Wide Association study-ko (GWAS) informazioaren arabera [17]. Horretarako, SNPen  $\beta$ -balioa eta indibiduo bakoitzak SNP horiekiko duen genotipoa (alelo eragilearekiko homozigotoa edo heterozigotoa den) kontuan hartu behar dira honako formula honetan [17] ( $p$ : SNP kopurua,  $\beta_j$ :  $j$  SNPren  $\beta$ -balioa eta  $SNP_{ij}$ :  $i$  indibiduoak  $j$  SNPrekiko aurkezten duen genotipoa):

$$PRS_i = \sum_{j=1}^p \beta_j x_{SNP_{ij}}$$

RStudio [17] plataformaren bidez 1000 Genoma proiektuko datu-baseko banakoetan guri interesatzen zaizkigun SNP bildumarekin aldakortasun genetikoaren analisi multibariantea egin dugu, Oinarritzko Osagaien Analisia hain zuzen ere (PCA, Principal Component Analysis), SNPReLATE [19] paketea erabilia. «ggplot2» [20] paketea erabili dugu banakoen bi dimentsioko ardatz edo Eigenvektorekiko (EV) banaketa irudikatzeke.

### 2.4. Aurpegiaren ezaugarri morfologikoak

Aurpegiaren ezaugarri morfologikoei loturiko SNPak erabilia, hots, test-SNPak, indibiduoak taldeetan nola sailkatzen ziren ikusteko, lehenbizi, Mclust [21], Stat-ellipse [22] eta ggplot2 [20] paketeak erabiliz, indibiduen taldekapenak eta talde hoiaren irudikapenak egin genituen. Taldekatze horiek banakoen taldekapenak dira, zeinak PCA argiago ikusten laguntzen diguten.

Talde bakoitzeko banakoen sailkapena egiteko, bi irizpide erabili ditugu: alde batetik PRS balioa eta bestetik indibiduen jatorri geografikoa. PRS balioaren araberrako sailkapena behatzeko, lau PRS klase bereiztu di-

tugu kuartilen arabera: lehengo klasea % 0-25 bitartean dauden PRS balioak, bigarren klasea % 25-50 tarteko balioak, hirugarren klasea % 50-75 bitartekoak eta, azkenik, laugarren klasea % 75-100 artekoak. Jatorri geografikoaren arabera sailkapena ikusteko, banakoen populazio eta azpi-populazioen jatorri geografikoa kontuan hartu ditugu (2. irudia eta Material Osagarrien 2. taula).

## 2.5. Sudurraren morfologia

Munduko populazioak euren artean bereizten dituzten aurpegiko ezaugarri morfologikoen adibide bat hautatu dugu, sudurraren morfologia, hain zuzen ere. Aztertutako SNPetik multzo handi batek eragiten zuelako ezaugarri horretan. Lortutako test-SNP guztietatik, hamabik sudurraren ezaugarri ezberdinetan dute eragina (1. taula eta 1. irudia).

**1. taula.** Sudurraren morfologian eragina duten SNPak, kokapen kromosomikoa, alelo eragilea eta zer ezaugarritan duten eragina. Ezaugarrien irudikapen grafikoa 2. irudian ikus daiteke.

Ezaugarria	Kromosoma	SNP	Alelo eragilea
Sudurraren luzera (1. irudian gorritz)	1	rs647711	T
	8	rs113036800	T
	14	rs8007643	T
Sudurraren kartilago proportzioaren zabalera	16	rs62031988	C
Sudur zubiaren zabalera (1. irudian beltzez)	6	rs1852985	T
Sudurraren zabalera (1. irudian urdinez)	7	rs17640804	C
	9	rs10868138	C
	15	rs1258763	A
	17	rs227731	C
	20	rs2424399	C
	20	rs927833	T
	20	rs2206437	A

Taulan ageri diren SNPen alelo eragileek dagozkien ezaugarrien tamaina handiagoa eragiten dute eta alelo eragilea ez edukitzeak tamaina txikiagoa izatea esan nahi du. Hau da, esaterako, rs10868138 SNPean alelo eragilea denak, C kasu honetan, indibiduoak sudur zabalagoa izatea eragingo du. Alelo eragilea aurkezten ez dutenek, aldiz, sudur estuagoa izango dute. Gauza bera gertatuko da gainerako ezaugarrietan.

Sudurreko test-SNPak identifikatu ondoren, mundu osoko populazioetako banakoen PRS balioak (SNP horietarako) kalkulatu ditugu berriro ere.

Ondoren, populazioen batz besteko PRS balioak erabilita, sudurraren morfologia populazio batetik bestera (2. irudia) nola aldatzen den aztertu dugu.

## **2.6. SNP kontrolak**

Analisi hauen emaitzak balioztatzeko, prozedura bera errepikatu dugu zenbait SNP kontrolekin (hemendik aurrera «SNP-kontrolak»), hau da, ausazko SNPak. Bost ausazko SNP bilduma independente erabili ditugu. SNP-kontrol hauek SNPsnap online baliabidearen bidez lortu dira (<https://data.broadinstitute.org/mpg/snpSnap/>) [23]. Ondoren, 1000 Genoma proiektuko datu-baseko banakoek SNP hauetan dituzten genotipoen VCF fitxategia eraiki dugu. Modu honetan, SNP-kontrolen 5 bildumak eta gure lagina alderatuz (test-SNPak), lortutako test-SNPen genotipoak aztertutako ezaugarriekiko espezifikoak direla onar dezakegu; edo, aitzitik, bakarrik laginen historia demografikoaren ondorio soila direla.

## **3. EMAITZAK**

Bilaketa bibliografikoan aurkitutako test-SNP guztietatik honako baldintza hauek betetzen dituzten SNPak aukeratu ditugu: SNPen arteko independentzia, kromosoma autosomikoetan agertzea (ez kromosoma sexualetan edo genoma mitokondrialean) eta alelo eragilea aurpegiko zer ezaugarri morfologiko zehatzi lotzen zaion jakitea. Guztira, 99 SNP hautatzea lortu dugu (Material osagarrien 1. taula).

Test-SNP eta lehenengo SNP-kontrolen PCAk alderatuta (4. eta 5. irudiak, hurrenez urren, eta Material Osagarrien 2, 4, 6 eta 8. irudiak, gainontzeko SNP-kontrolerako), biek antzeko sakabanatze-ereduari jarraitzen diotela ikus dezakegu. Hau da, indibiduen taldekapena euren jatorri geografikoaren arabera da, eta populazio ezberdinak modu berdintsuan desberdinu daitezke, bai test-SNPetan, bai lehenengo SNP-kontrolerako ausaz aukeratu diren SNPetan ere. Horrela, lehendabiziko banapen garrantzitsua Afrika (4. eta 5. irudian puntu gorriak) eta gainontzeko populazioen artean gertatzen da. Afrikaz kanpoko populazioen artean, oso ondo bereizten dira ekialdeko asiarrak (4. eta 5. irudian puntu berdeak).



**4. irudia.** Test-SNPen EV1 eta EV2 ardatzekin irudikatutako Osagai Nagusien Analisisa (PCA). Banako bakoitza puntu batez adierazten da, eta euren jatorri geografikoa kolorearen arabera da (eskuineko leienduan adierazia). Taldekatzeak koloretzatutako biribilak dira, leiendako koloreen arabera bereizten dira.

Hala ere, banaketa hau ez da guztiz berdina test-SNP eta SNP-kontrolletan. Alde batetik, afrikar indibiduoek arreta jarritz, test-SNPen PCAn lehenengo ardatzak Asiaren ekialdeko populazioak gainontzeko populazioetatik bereizten ditu (EV2 % 8,24), eta SNP-kontrolletan, afrikar populazioa da gainontzeko populazioetatik bereizten dena (EV1 % 9,42). Gainontzeko SNP-kontrolak (Material osagarrien 2., 4., 6. eta 8. irudiak), lehenengo SNP-kontrolaren eredu berari jarraitzen diote, bariantzaren balio zehatzak aldatuz kasu bakoitzean.

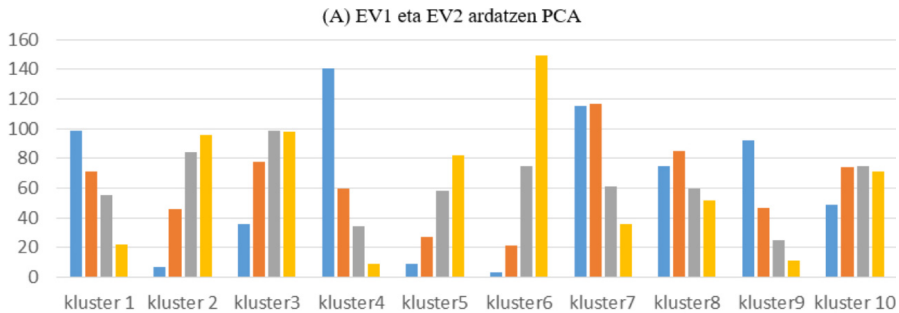




**5. irudia.** Lehenengo SNP-kontrolen EV1 eta EV2 ardatzekin irudikaturiko Osa-gai Nagusien Analisia (PCA). Banako bakoitza puntu batez adierazten da, eta euren jatorri geografikoa kolorearen arabera da (esku-biko leiendan adierazia). Taldekatzeak koloreztatutako biribilak dira, leiendako koloreen arabera bereizten dira.

Talde bakoitzaren PRS klase proportzioa (ikus material eta metodoak atala) ezberdina da barra-grafikoetan ikus daitekeen bezala (6. irudia). Taldekatze batetik bestera, PRS klaseen proportzioan aldaketa nabariak daudela ikus daiteke.

Desberdintasun horiek estatistikoki esanguratsuak direla frogatzeko, Fisherren Test Zehatza egin dugu. Hau da, 45 alderaketa (kluster guztiak binaka alderatu ziren euren artean) guztira. 2. taulan, egindako konparazio guztietatik lortutako baliorik esanguratsuenak azaltzen dira.



**6. irudia.** Test-SNPekin egindako PCA analisisen EV1 eta EV2 ardatzetan definitutako 10 taldekatzeak. Taldekatze bakoitzerako PRS balioen proportzioa kuartilen arabera: urdina, PRS balioa altuenak, i.e. goiko % 25 bitartean; laranja, PRS balioak % 25-% 50 bitartean; grisa, PRS balioak hirugarren kuartilean (% 50-% 75 bitartean); horia, PRS balio baxuenak, i.e. azken kuartilean.

Horrenbestez, p-balioak esangarriak izateko 0,0011 baino txikiagoak izan behar dira, 45 egiten ari baikara (0,05/45). 2. taulan ikus daitekeen modura, EV1/EV2 ardatzetan alderaketa guztiak esangarriak dira. Beraz, EV1/EV2 ardatzetan, taldekatze batetik bestera PRS balioaren proportzioetan desberdintasun estatistikoki esanguratsuak daudela onartzen dugu.

**2. taula.** Fisherren Test Zehatzean egindako konparaketak eta ardatz ezbedinen PCAn lortutako p-balioak. Beltzez esangarriak diren balioak (0.05/45 = 0.0011 ( $p < 0.0011$ )).

Konparaketa	EV1/EV2
Kluster1-Kluster2	<b>9,206E-25</b>
Kluster1-Kluster6	<b>2,2988E-31</b>
Kluster4-Kluster6	<b>3,227662E-51</b>
Kluster6-Kluster9	<b>1,76298E-39</b>

## Sudurreko morfologiaren analisisia

Sudurreko morfologiaren azterketari dagokionez, mundu osoko populazioetan sudurreko morfologiarekin erlazionatuta dauden hogeit hamar SNP (1. taulan izartxoaren dutenak) aztertu ditugu. Sudurreko morfologiaren PRS balioei dagokienez, populazioetan ezberdintasun nabariak ikusi ditugu (3. taula). Alde batetik, baliorik txikiak europarretan behatu dira. Baliorik handienak, berriz, Afrika eta Ekialdeko Asiako populazioetan.

**3. taula.** Sudur-morfologiaren araberako PRS balioak munduko populazio ezberdinetan.

Populazioa	PRS (Sudurreko morfologia)
EUR	-1,860645328
SAS	-1,071955256
AMR	-1,060773487
AFR	-0,259062390
EAS	0,251773254

Material eta metodoetan adierazi den bezala, sudurreko SNP morfologikoetako alelo eragileak ezaugarri horietan tamaina handiagoaren arduradun direnez, PRS handieneko populazioak sudur zabalena izango ditu. PRS balio txikiak dituztenek, berriz, sudur estuena. Sudurraren luzerarekin (1. irudian, gorri) loturiko SNP bakarra lortu dugunez (1. taula), ezin dugu erabateko ziurtasunez adierazi populazioen sudur luzera handiagoa edo txikiagoa denik. Beraz, zabalerari dagokionez, emaitzek iradokitzen dute europarrek sudur estuagoa dutela, eta, bai Afrika eta bai Asia Ekialdekoek, berriz, sudur zabalagoa.

#### 4. EZTABAIDA

Lan honetan, informazio genetiko erabiliz munduko populazioak morfologikoki desberdintzeko posibilitatea aztertu dugu. Horretarako, bibliografian aurpegiaren morfologiarekin erlazionatu diren 99 SNP autosomiko identifikatu ditugu, eta 1000 Genoma proiektuko 2.505 indibiduoek SNP hauentzako azaltzen dituzten genotipoak bildu ditugu. Horretaz gain, ausaz aukeraturiko SNP kopuru bereko (99 SNP) 5 bilduma independente eraiki ditugu, SNP-kontrolak deritzenak, alderaketak egiteko. Datu multzo horietan Osagai Nagusien Analisia burutu ondoren, aurpegiko morfologiarekin erlazionaturiko SNP eta kontrolako bildumen artean zera ikusi dugu (4. eta 5. irudiak eta Material Osagarrien 2, 4, 6 eta 8 irudiak): SNP-Kontrolatan EV1 ardatzak Afrikako populazioa gainontzeko populazioetatik bereizten du % 9,42 bariantzarekin, eta EV2 ardatzak % 4,94 bariantzarekin Asia ekialdeko populazioa bereizten du; SNP morfologikoen (test-SNP) egingandako analisisetan, aldiz, bereizketa bera ikusten da, baina EV1 ardatzean bariantza handiagoarekin (% 8,24) Asia ekialdeko populazioa banatzen da, eta EV2an, bariantza txikiagoarekin (% 6,51) afrikarrak. Beraz, Asia ekialdeko populazioetan SNP morfologikoen eragin handiagoa dute.

Aurpegiaren morfologiarekin erlazionaturiko SNPak erabilia, Asia ekialdeko populazioaren aurpegiaren morfologia gainontzeko populazioek

erakusten dutenaren ezberdina dela dirudi. Izan ere, Asia ekialdeko populazioaren bereizketak garrantzi handiagoa du test-SNPetan SNP-kontrolatan baino (SNP-kontrolatan EV2 ardatzak % 4,94 eta test-SNPetan EV1 ardatzak % 8,24). Hori dela eta, hautatutako SNPek benetan aurpegi-morfologia ezberdina eragiten dutela ondoriozta dezakegu. Dena den, ezin dugu ahaztu ikerketa honetan erabilitako test-SNPak (hots, morfologikoak) Asiako populazioetan aurkituak eta aztertuak izan direla batez ere, eta, beraz, baliteke PCA analisisietan ikusten ditugun desberdintasunak alborapen edo sesgo hori adieraztea.

Horrenbestez, historia demografikoak aurpegiaren morfologiarekin erlazioaturiko aldakortasun genetikoan eragin handia izan duen arren (4. irudiko EV1 eta EV2 ardatzek azaltzen dituzten bariantzak % 8,24 eta % 6,51 dira, hurrenez hurren), gaur egungo populazioetan agertzen den aurpegiko morfologiaren aldakortasuna beste faktore batzuen menpe ere egon daiteke. Gure datuekin ezin dugu jakin aldakortasun horren beste eragileak zehatz-mehatz zein diren. Hala ere, hautespen sexuala faktore bat izan daiteke, gainerako ezaugarrietan izugarritzko eragina izan baitu [25]. Baina hau espekulazio hutsa da.

Test-SNPen taldekapenari dagokionez 4. irudian argi ikusten da taldekatze bakoitzean gizabanakoen sailkapena jatorri geografikoaren arabera oinarrituta dagoela. EV1/EV2 ardatzen PCAn (4. irudian) 5. eta 6. taldekatzeetan, Asia Ekialdeko indibiduoak daude nagusiki, 3. eta 10. taldekatzeetan, aldiz, afrikarrak. Zenbait taldekatzek Amerika, Asia Hegoaldea eta Europako indibiduoak barneratzen dituzte. Asia ekialdekoak, Eurasiako populazioen artean historia demografiko desberdinena azaltzen dute, gaintzekoengandik isolatuta egon baitira, eta horren ondoriozko jito genetikoaren eragina dela eta, morfologikoki (eta beraz genetikoki) desberdin omen zitezkeen [25]. Azkenik, Europa, Amerika eta Asia hegoaldeko populazioek historia demografiko antzekoagoa dute, eta ondorioz, antzekotasun morfologiko handiago erakusten dute euren artean.

Bestalde, EV1/EV2 ardatzetako taldekatzeetako PRS klase proportzioa (6. irudia) eta taldekatze bakoitzean gailentzen den jatorri geografikoa kontuan hartuz (4. irudia), jatorri geografikoaren arabera PRS balio ezberdinak gailentzen direla esan daiteke. Beraz, PRS balio horri lotuta doazen aurpegiaren morfologiarekin erlazioatutako ezaugarriak izango dituzte populazio horretako indibiduoek. Hau da, lehengo taldekatzea adibidetzat hartuz, hango gehienak Asia Ekialdekoak dira (3. irudia) eta gailentzen den PRS klasea, lehengoa (6. irudia). Hortaz, PRSko lehengo klasea Asia Ekialdeko populazioarekin estuki erlazioatuta dago eta hango indibiduen aurpegi-morfologia bereizgarria da, gainerako populazioetatik desberdintzen dituzten ezaugarriak baitituzte.

Azkenik, sudurraren morfologiari dagokionez (4. taula), zehazki lortu ditugun emaitzei erreparatuta, hein handi batean behintzat, populazioetan

ikusten diren sudur morfologia ezberdinak moldapen morfologiko moduan azaldu daitezke. Izan ere, beste ezaugarri morfologiko batzuk bezala [27], sudurraren morfologia klimarekin estuki lotuta dagoela argitaratu da [28]. Izan ere, sudurreko muki mintzak arnasten den airea berotzen eta hezetzen du, arnasbideetara pasatzeko prestatuz. Horrela, sudur luze eta estuek ingurunekeo airea hezetzeko eta berotzeko azalera handiagoa izango dute. Klima hotz eta lehorretan oso garrantzitsua da prozesu hori, eta, beraz, klima konkretu horietan moldapen paregabea izango da.

Hala ere, bioklima berean habitat oso ezberdinak aurkitzen dira, eta zenbait kasutan faktore bakoitzak kontrako sudur motak hobesten omen baititu. Beraz, nahiz eta munduan zehar populazioak sudurraren morfologiaren arabera bereizteko eredu orokor bat planteatu, salbuespen asko daude. Bestalde, sudur-morfologian populazioen arteko desberdintasunak beste faktore baten edo faktore anitzen arteko elkarrekintzen ondoriozkoak ere izan litezke, hautespen sexuala barne, zenbait autorek proposatu duten bezala [28].

## 5. ONDORIOAK

Lan honetan giza populazioak aurpegiaren morfologiari loturiko datu genetikoetatik abiatuz euren artean desberdintzen saiatu gara. Hala ere, ez dugu behar adina SNP kopuru giza aurpegiaren morfologia populazioen banapen geografikoaren arabera aurreratzeko. Izan ere, gure lanean gehien jorratu dugun gaia aurpegiaren morfologiaren eboluzioa izan da. Hala eta guztiz ere, SNP multzo bat aztertuta populazioen sudur-morfologia nolakoa den aurreratzeko gai izan garen moduan, antzekoa egin daiteke gainerako ezaugarri morfologikoekin, baldin eta sakontasun eta baliabide gehiagorekin aztertzen badira. Kontuan hartu behar da, ordea, ezaugarri konplexuak ikertzen ari garela, eta, beraz, horien efektua neurtzea zailagoa da. Gainera, aurpegiko ezaugarri guztietan eragina duten SNP guztien identifikazioa oso zaila izateaz gain, ikerketa asko soilik populazio konkretu batzuetan egin dira, besterik ez; horrenbestez, informazio urria dugu. Bestalde, populazio bakoitzak SNP eragile ezberdinak izan ditzake eta hori kontuan ez edukitzeak berrespen-alborapena izateko arriskua sortzen du. Hala ere, nahiz eta gaur egun banako baten DNAn oinarrituz bere aurpegiko morfologia aurreratea oso zaila izan, geroz eta ugariagoak diren datu genomikoek eta GWASek arlo honen etorkizuna argitzen dute.

## BIBLIOGRAFIA

- [1] TOME, P., FIERREZ, J., VERA-RODRIGUEZ, R. eta RAMOS, D. 2013. «Identification using face regions: Applications and assesment in forensic scenarios». *Forensic Science International*, **233** (1-3), 75-53.

- [2] COLE, J. B., MANYAMA, M., LARSON, J. R., LIBERTON, D. K., FERRARA, T. M., RICCARDI, S. L., ... HALLGRÍMSSON, B. 2017. «Human Facial Shape and Size Heritability and Genetic Correlations». *Genetics*, **205**(2), 967-978. Doi: 10.1534/genetics.116.193185.
- [3] GALWAY-WITHAM, J., & STRINGER, C. 2018. «How did Homo sapiens evolve?». *Science*, **360**(6395), 1296-1298.
- [4] O'SHEA, N. eta DELSON, E. 2018. «The Origin of Us: Archaeological and fossils finds shed light on the emergence and spread of modern humans». *Natural History*, **126**(8), 19-22.
- [5] LINDENFORS, P. eta TULLBERG, B. S. 1998. «Phylogenetic analyses of primate size evolution: the consequences of sexual selection». *Biological Journal of the Linnean Society*, **64** (4), 413-447. Doi: 10.1111/j.1095-8312.1998.tb00342.x.
- [6] EVTSEV, A. A. eta MOVSESIAN, A. A. 2016. «Testing the Association between Human Mid-Facial Morphology and Climate Using Autosomal, Mitochondrial, Y Chromosomal Polymorphisms and Cranial Non-Metrics». *American Journal of Physical Anthropology*, **159**(3), 517-522.
- [7] SOBRINO, B., BRIÓN, M. eta CARRACEDO, A. 2005. «SNPs in forensic genetics: a review on SNP typing methodologies». *Forensic Science International*, **154** (2-3), 181-194. Doi: 10.1016/j.forsciint.2004.10.020.
- [8] ADHIKARI, K., FUENTES-GUAJARDO, M., QUINTO-SÁNCHEZ, M., MENDOZA-REVILLA, J., CHACÓN-DUQUE, J. C., ACUÑA-ALONZO, V., ... GÓMEZ-VALDÉS, J. 2016. «A genome-wide association scan implicates *DCHS2*, *RUNX2*, *GLI3*, *PAX1* and *EDAR* in human facial variation». *Nature communications*, **7**(1), 1-11. Doi: 10.1038/ncomms11616
- [9] CHA, S., LIM, J. E., PARK., A. Y., DO, J. H., LEE, S. W., SHIN, C., ... WOO, K. M. 2018. «Identification of five novel loci related to facial morphology by genome-wide associations studies». *BCM Genomics*, **19** (1), 481.
- [10] CROUCH, D. J., WINNEY, B., KOPPEN, W. P., CHRISTMAS, W. J., HUTNIK, K., DAY, T., ... SPECTOR, T. D. 2018. «Genetics of the human face: Identification of large-effect single gene variants». *Proceedings of the National Academy of Sciences*, **115** (4), E676-E685. Doi: 10.1073/pnas.1708207114.
- [11] PICKRELL, J. K., BERISA, T., LIU, J. Z., SÉGUREL, L., TUNG, J. Y. eta HINDS, D. A. 2016. «Detection and interpretation of shared genetic influences on 42 human traits». *Nature Genetics*, **48**(7), 709-717.
- [12] INDENCLEEF, K., ROOSENBOOM, J., HOSKENS, H., WHITE, J. D., SHRIVER, M. D., RICHMOND, S., ... CLAES, P. 2018. «Six NSCL/P loci show associations with normal-range craniofacial variation». *Frontiers in Genetics*, **9**, 502.
- [13] ANDRADE, C. 2015. «Understanding Relative Risk, Odds Ratio, and Related Terms: As simple as It Can Get». *The Journal of clinical Psychiatry*, **76**(7), 857-861.
- [14] DUNCAN, L., SHEN, H., GELAYE, B., MEIJSEN, J., RESSLER, K., FELDMAN, M. eta DOMINGUE, B. 2019. «Analysis of polygenic risk score usage and performance in diverse human populations». *Natural communications*, **10**(1), 1-9. Doi: 10.1038/s41467-019-11112-0.

- [15] WELTER, D., MACARTHUR, J., MORALES, J., BURDETT, T., HALL, P., JUNKINS, H., ... PARKINSON, H. 2014. «The NHGRI GWAS Catalog, a curated resource of SNP-trait associations». *Nucleic acids research*, **42**(D1), D1001-D1006. Doi: 10.1093/nar/gkt1229.
- [16] DANECEK, P., AUTON, A., ABECASIS, G., ALBERS, C. A., BANKS, E., DEPRISTO, M. A., ... & 1000 Genomes Project Analysis Group. 2011. «The variant call format and VCFtools». *Bioinformatics*, **27**(15), 2156-2158.
- [17] RAI BHATNAGAR, S. McGill University. Polygenic Risks Scores with data.table in R, <https://sahirbhatnagar.com/blog/2017/08/11/polygenic-risks-scores-with-data.table-in-r/>, 2020/06/17.
- [18] R CORE TEAM. 2019. «R: A language and environment for statistical computing». R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- [19] ZHENG, X., LEVINE, D., SHEN, J., GOGARTEN S. M., LAURIE, C. eta WEIR, B. S. 2012. «A High-performance Computing Toolset for Relatedness and Principal Component Analysis of SNP Data». *Bioinformatics*, **28**(24), 3326-3328. Doi: 10.1093/bioinformatics/bts606
- [20] WICKHAM, H. 2016. «ggplot2: Elegant Graphics for Data Analysis». Springer-Verlag New York.
- [21] SCRUCICA, L., FOP, M., MURPHY, T. B. eta RAFTERY A. E. 2016. «mclust 5: clustering, classification and density estimation using Gaussian finite mixture models». *The R Journal*, **8**(1), 205-233.
- [22] FOX, J. eta WEISBERG, S. 2011. «An R Companion to Applied Regression», Second Edition. Thousand Oaks CA: Sage. URL: <http://socserv.socsi.nmaster.ca/jfox/Books/Companion>.
- [23] PERS, T. H., TIMSHEL, P. eta HIRSCHHORN, J. N. 2014. «SNPsnap: a Web-based tool for identification and annotation of matched SNPs». *Bioinformatics*. Doi: 10.1093/bioinformatics/btu6555.
- [24] YU, Y., ZUO, X., HE, M., GAO, J., FU, Y., QIN, C., ... BIAN, Z. 2017. «Genome-wide analyses of non-syndromic cleft lip with palate identify 14 novel loci and genetic heterogeneity». *Nature. Communications*, **8**(1), 1-11. Doi: 10.1038/ncomms14364.
- [25] PUTS, D. A. 2010. «Beauty and the beast: mechanisms of sexual selection in humans». *Evolution and Human Behaviour*, **31**(3), 175-175.
- [26] KEINAN, A., MULLIKIN, J. C., PATTERSON, N. eta REICH, D. 2007. «Measurement of the human allele frequency spectrum demonstrates greater genetic drift in East Asians than in Europeans». *Nature Genetics*, **39**(10), 1251-1255. Doi: 10.1038/ng2116.
- [27] KATZAMARZYK, P. T. eta LEONARD, W. R. 1998. «Climatic Influences on Human Body Size and Proportions: Ecological Adaptations and Secular Trends». *American Journal of Physical Anthropology*, **106**(4), 483-503.
- [28] ZAIDI, A. A., MATTERN, B. C., CLAES, P., MCECOY, B., HUGHES, C. eta SHRIVER, M. D. 2017. «Investigating the case of human nose shape and climate adaptation». *PLOS Genetics*, **13**(3): e1006616. Doi: 10.1371/journal.pgen.1006616.



**MATERIAL OSAGARRIAK**

**1. taula osagarria.** Adierazten dira aurpegiaren morfologiaren ezaugarriei loturiko 99 SNP autosomikoen zerrenda, kokapen kromosomikoa, euren base-pare kokapena, kromosoma eta bakoitzaren alelo eragilea. Izartxoan duten 37 SNPak Asia Ekialdeko populazioetan soilik aztertu direnak dira.

Kromosoma	Base-pare posizioa	SNP	Alelo eragilea
1	3261516	rs4648379	T
	18972776	rs9439713	A
	18979874	rs742071	T
	94553438	rs560426	G
	94558110	rs66515264	T
	94570016	rs481931*	C
	94582293	rs4147803	G
	110218761	rs619686	G
	117122288	rs647711	T
	209959614	rs1044516*	G
	209960925	rs742214*	C
	209964080	rs2235371*	C
	209977111	rs861020	A
	209989270	rs642961*	A
	209993801	rs596731*	G
	210048819	rs2064163*	C
	210050794	rs9430019*	A
2	9971366	rs287980*	G
	16729357	rs7566780	G
	16733054	rs10172734*	G
	16733928	rs7552	G
	19732291	rs7567283	G
	42181679	rs6740960	T
	43540125	rs7590268	G
	177111819	rs970797	A
	223005314	rs974448	G
	223068286	rs7559271	G
	234394487	rs1048603	A
234432017	rs838543	A	
3	89534377	rs7632427	T
	189549423	rs17447439	G
	189553372	rs76479869	T



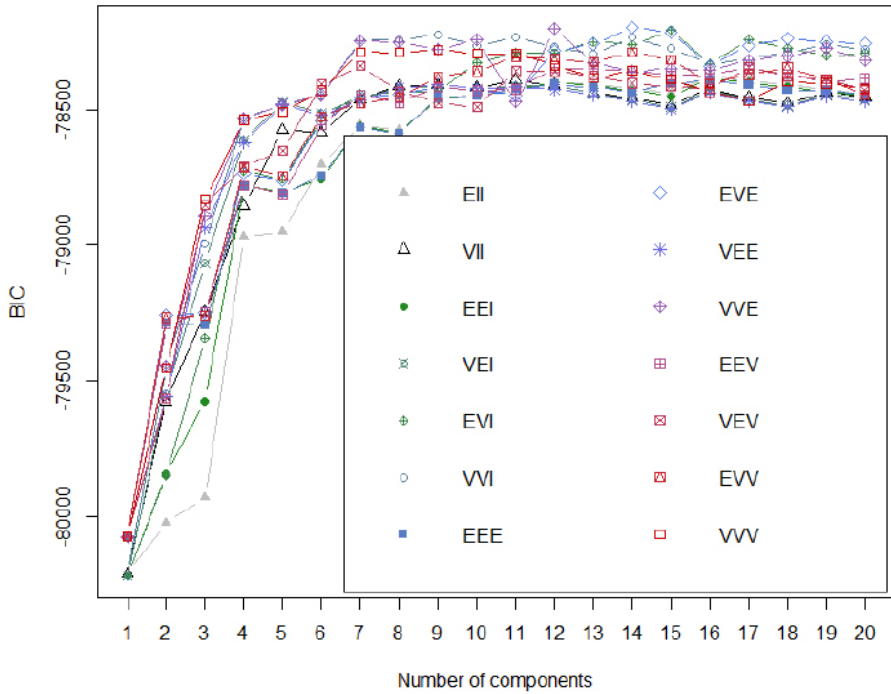
*Gizakiaren aurpegiaren ezaugarri fenotipikoei loturiko aldakortasun  
genetikoaren analisia*

Kromosoma	Base-pare posizioa	SNP	Alelo eragilea
4	4794195	rs34246903*	C
	4818925	rs1907989*	G
	124906257	rs908822*	A
	154831899	rs2045323	A
5	44068846	rs10462065*	A
	171128464	rs6555969	T
6	9469238	rs9381107*	G
	45329489	rs1852986	G
	45329656	rs1852985	T
	162590018	rs9456748	G
	169957964	rs3736712	C
7	42131390	rs17640804	C
8	38269514	rs13317*	A
	58449027	rs113036800	T
	88868340	rs12543318	C
	88918331	rs1034832*	T
	95401265	rs12681366*	A
	95541302	rs957448*	A
	129645234	rs16903544	C
	129888794	rs7845615*	A
	129946154	rs987525	A
	129950844	rs7017252*	A
129976136	rs55658222	A	
9	14883254	rs72713618	A
	86917301	rs10868138	C
	92209587	rs7871395	A
	98259703	rs10512248	A
10	105810879	rs805723	T
	118827560	rs7078160	A
	118893231	rs6585429*	A
11	101394765	rs12786942	T
12	56435412	rs705704*	A
	72080272	rs2304269*	A
	72089040	rs7967428*	G
	119414564	rs117438382	A
13	80639405	rs9545308*	A
	80668874	rs9574565	C
	80679302	rs11841646	A

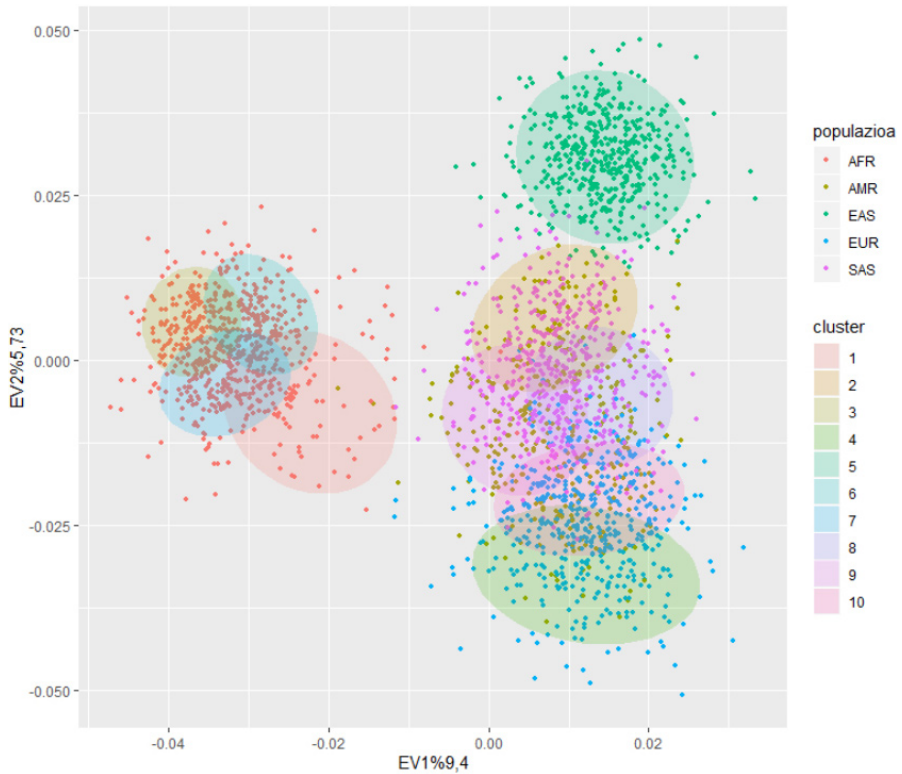
Kromosoma	Base-pare posizioa	SNP	Alelo eragilea
14	21365801	rs8007643	T
	38038468	rs17106852	G
	51839645	rs7148069*	A
	95379583	rs1243573*	G
15	33050423	rs1258763	A
	63312632	rs1873147	C
	74889163	rs11072494	T
	75005575	rs28689146	A
16	3969886	rs2283487*	A
	3980445	rs8049367	C
	3996282	rs17136624*	A
	27017495	rs62031988	C
17	7521915	rs12150660	T
	7533423	rs1799941	A
	8914693	rs2872615*	A
	8929845	rs1880646*	A
	8932119	rs4791774	G
	44988703	rs4968247*	A
	45008935	rs1838105*	A
	54615617	rs17760296	G
	54773238	rs227731	C
70025809	rs2193054	C	
19	2050823	rs3746101	T
20	21632545	rs2424399	C
	22041577	rs927833	T
	37992741	rs2206437	A
	38904203	rs6129564	A
	39269074	rs13041247	T
	39275603	rs6129653*	A

**2. taula Osagarria.** 1.000 Genoma datu-basean aztertutako populazio ezberdinetan dauden azpipopulazioak eta siglen esanahiak.

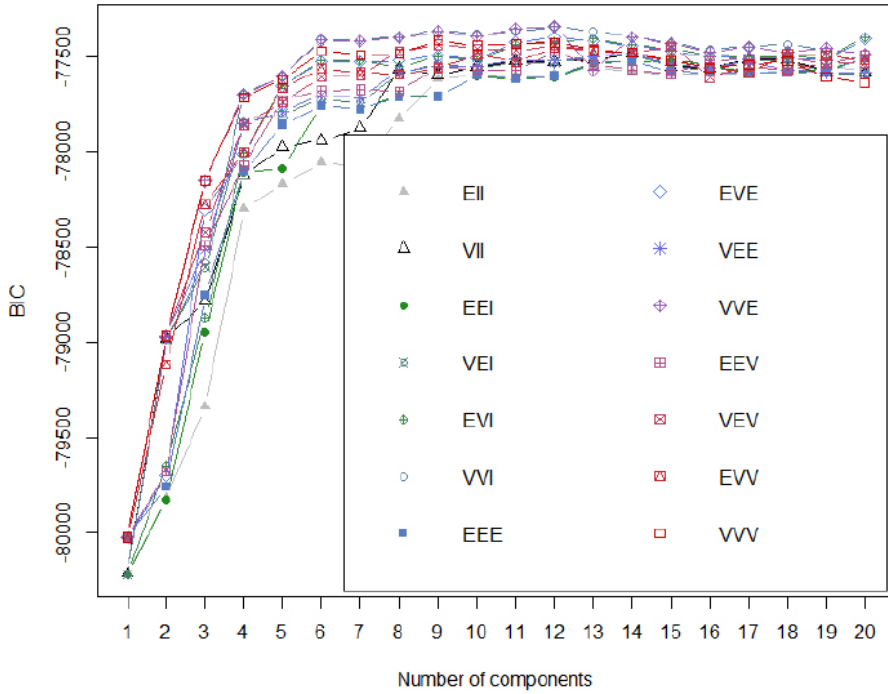
Populazioa	Azpipopulazioa	Laburpenaren esanahia
EAS (Asia ekialdea)	CHB	Han txinatarrak (Pekin, Txina)
	JPT	Japoniarrak (Tokio, Japonia)
	CHS	Hegoaldeko Han txinatarrak
	CDX	Dai txinatarrak (Xishuangbanna, Txina)
	KHV	Kinh Ho Chi Minh hirian, Vietnam
EUR (Europa)	CEU	Iparraldeko eta mendebaldeko europar jatorriko Utah hiritarrak
	TSI	Toscanakoak (Italia)
	FIN	Finlandiarrak (Finlandia)
	GBR	Britaniarrak (Ingalaterra eta Eskozia)
	IBS	Espainiako Iberiar populazioa
AFR (Afrika)	YRI	Yorubak (Ibadan, Nigeria)
	LWK	Luhyak (Webuye, Kenya)
	GWD	Gambiarrak mendebaldeko dibisioetan, Gambia
	MSL	Mendetarrak, Sierra Leonan
	ESN	Esantarrak Nigerian
	ASW	Afrikar jatorriko amerikarrak AEBtako hego-mendebaldean
	ACB	Afrikar karibetarrak Barbadosen
AMR (Amerika)	MXL	Jatorri mexikarra dutenak Los Angelesen, AEB
	PUR	Puertoricoarrak Puerto Ricon
	CLM	Kolonbiarrak (Medellin, Colombia)
	PEL	Perutarrak (Lima, Peru)
SAS (Asia hegoaldea)	GIH	Houstoneko Gujarati Indioak, Texas
	PJL	Lahoreko Punjabiak, Pakistan
	BEB	Bangladeshko bengaliak (India)
	STU	Erresuma Batuko Sri Lankako tamilak
	ITU	Erresuma Batuko Telugu indiarrak



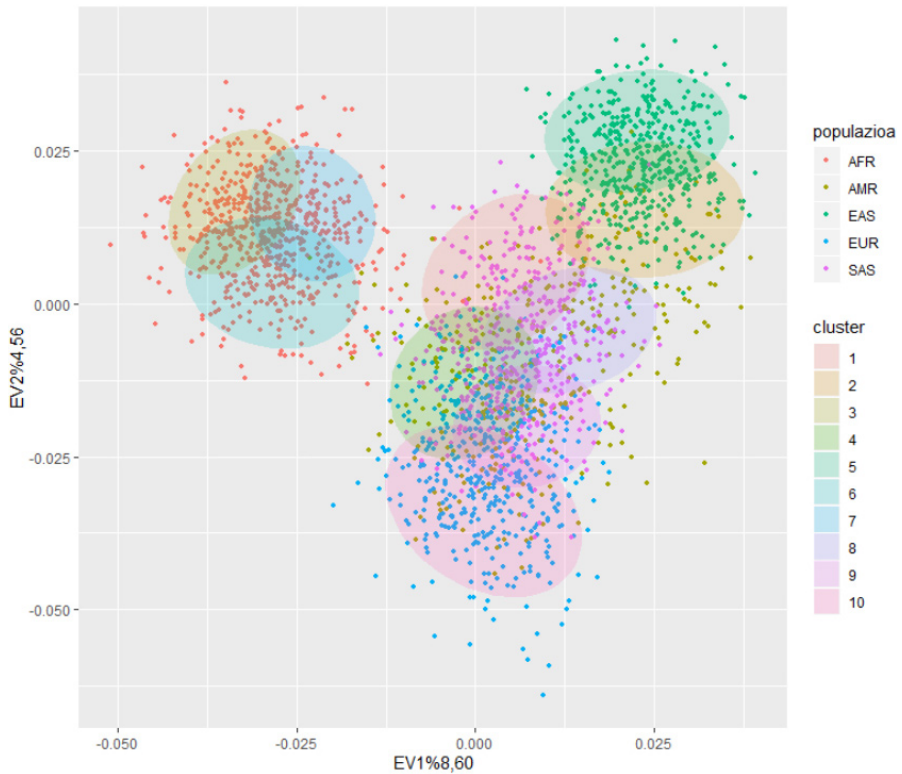
**1. irudi Osagarria.** Mclust paketea erabilia talde-kopuru hobeezina zenbatetsi da, 2. SNP-kontrola erabilia. X ardatzak talde-kopurua adierazten du, eta Y ardatzak taldekatzeak egiteko orduan metodo ezberdinek adierazten dituzten BIC irizpidea (Bayesian Information Criterion). Talde-kopuru hobeezina 12 da.



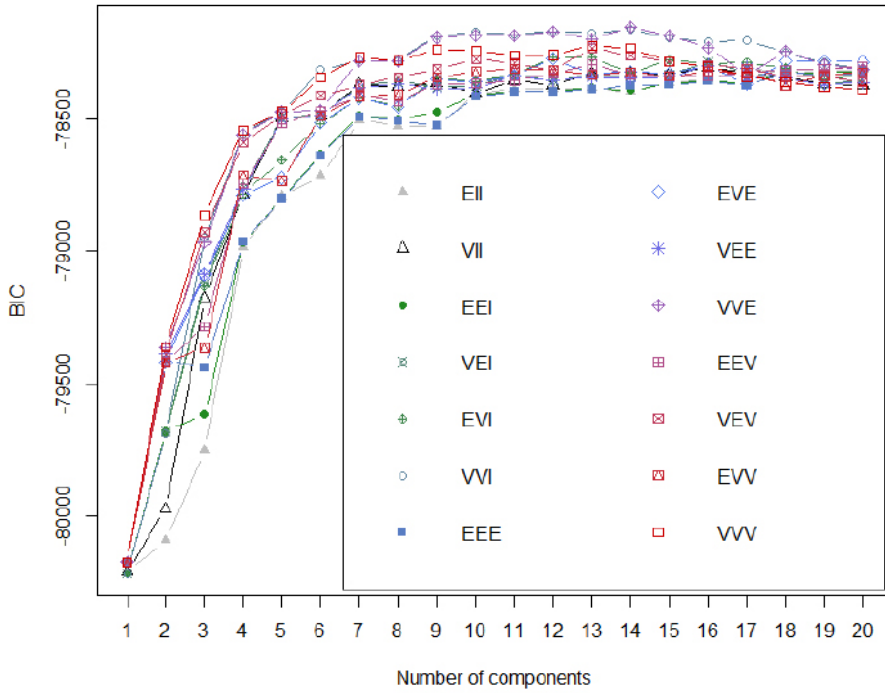
**2. irudi Osagarria.** Bigarren SNP-kontrolaren EV1 eta EV2 ardatzekin irudikatutako Osagai Nagusien Analisia (PCA). Indibiduo bakoitza puntu batez adierazten da eta euren jatorri geografikoa kolorearen arabera da (eskuineko leiendako adieraziaz). Taldekatzeak koloreztatutako biribilak dira, leiendako koloreen arabera bereizten dira.



**3. irudi Osagarria.** Mclust paketea erabilia talde-kopuru hobeezina zenbatetsi da, 3. SNP-kontrola erabilia. X ardatzak talde-kopurua adierazten du, eta Y ardatzak taldekatzeak egiteko orduan metodo ezberdinek adierazten duten BIC irizpi-dea (Bayesian Information Criterion). Talde-kopuru hobeezina 12 da.



**4. irudi Osagarria.** Hirugarren SNP-kontrolaren EV1 eta EV2 ardatzekin irudikatutako Osagai Nagusien Analisia (PCA). Indibiduo bakoitza puntu batez adierazten da, eta euren jatorri geografikoa kolorearen arabera da (eskuineko leiendako adierazia). Taldekatzeak koloretzatutako biribilak dira, leiendako koloreen arabera bereizten dira.

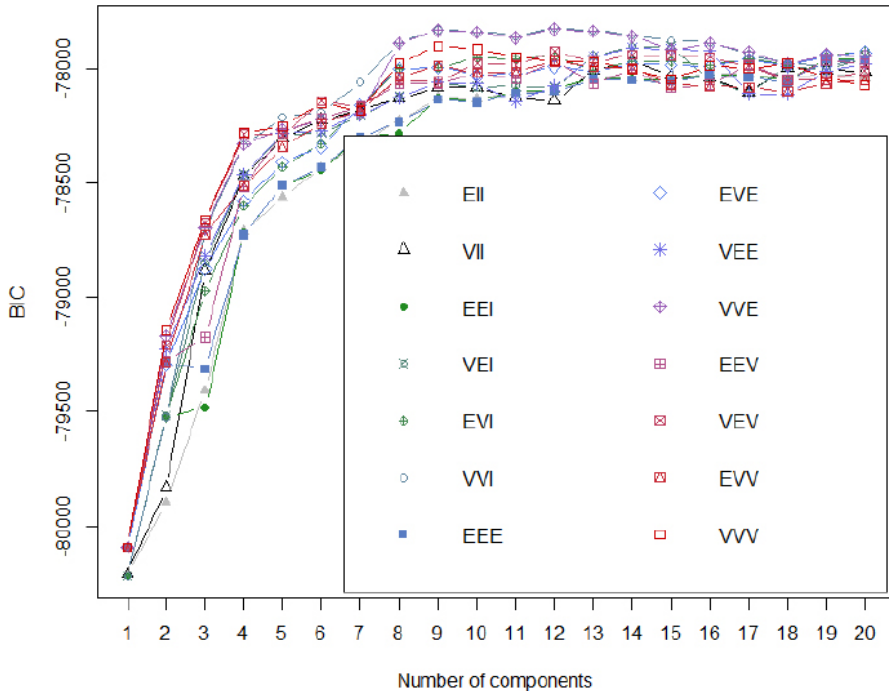


**5. irudi Osagarria.** Mclust paketea erabilia talde-kopuru hobeezina zenbatetsi da, 4. SNP-kontrola erabilia. X ardatzak talde-kopurua adierazten du, eta Y ardatzak taldekatzeak egiteko orduan metodo ezberdinek adierazten duten BIC (Bayesian Information Criterion). Talde-kopuru hobeezina 14 da.

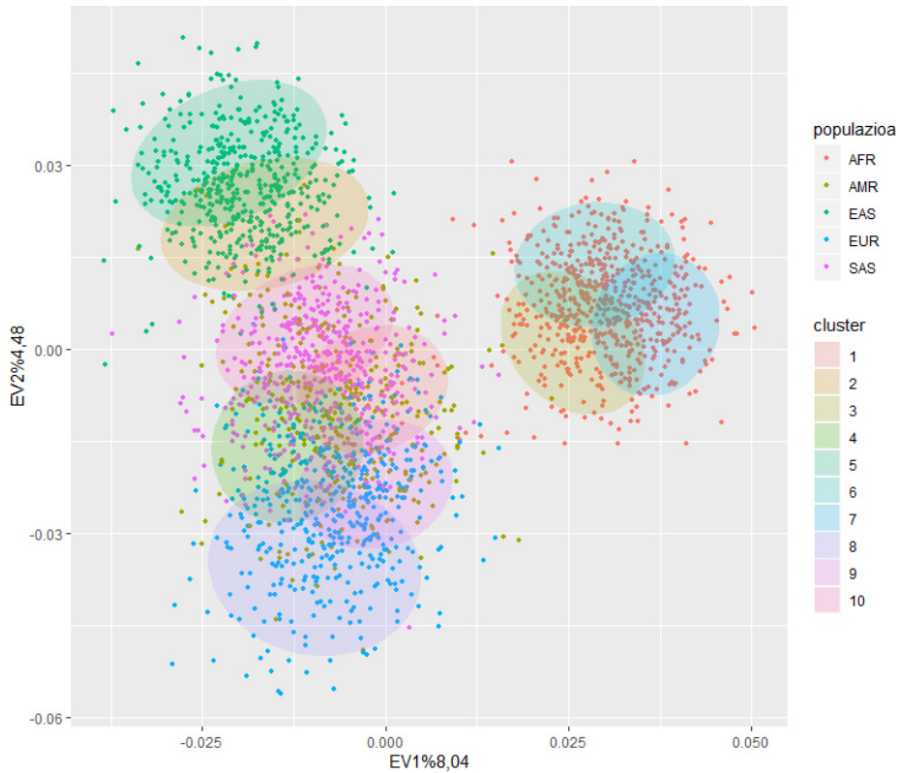




**6. irudi Osagarria.** Laugarren SNP-kontrolaren EV1 eta EV2 ardatzekin irudikatutako Osagai Nagusien Analisia (PCA). Indibiduo bakoitza puntu batez adierazten da, eta euren jatorri geografikoa kolorearen arabera da (eskuineko leiendako adierazia). Taldekatzeak koloretzatutako biribilak dira, leiendako koloreen arabera bereizten dira.



**7. irudi Osagarria.** Mclust paketea erabilia talde-kopuru hobeezina balioetsi da, 5. SNP-kontrola erabilia. X ardatzak talde-kopurua adierazten du, eta Y ardatzak taldekatzeak egiteko orduan metodo ezberdinek adierazten duten BIC (Bayesian Information Criterion). Talde-kopuru hobeezina 12 da.



**8. irudi Osagarria.** Bosgarren SNP-kontrolaren EV1 eta EV2 ardatzekin irudikatutako Osagai Nagusien Analisia (PCA). Indibiduo bakoitza puntu batez adierazten da, eta euren jatorri geografikoa kolorearen arabera da (eskuineko leiendak adierazia). Taldekatzeak koloreztatutako borobilak dira, leiendako koloreen arabera bereizten dira.

