

FaceSNPs: Identifying Face-Related SNPs from the Human Genome - Supplementary Materials

Somadina Mbadiwe, Jeremy Dawson, Donald Adjeroh
Lane Department of Computer Science and Electrical Engineering
West Virginia University Morgantown WV, USA
 {somadina.mbadiwe, jeremy.dawson, donald.adjeroh}@mail.wvu.edu

SUPPL I. RESULTS FROM THE PRUNING STAGES

Table I shows the breakdown per chromosome from all the pruning stages.

TABLE I
 PRUNING SUMMARY USING DATA FROM THE 1000 GENOME PROJECT

Chr	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
1	6,196,151	6,189,941	1,636,512	583	283
2	6,786,300	6,779,891	1,490,802	567	262
3	5,584,397	5,579,072	1,276,695	165	80
4	5,480,936	5,475,541	1,039,230	490	210
5	5,037,955	5,034,476	1,126,224	512	268
6	4,800,101	4,796,486	1,141,240	919	436
7	4,517,734	4,514,054	1,207,843	201	99
8	4,417,368	4,413,083	944,929	252	118
9	3,414,848	3,411,936	909,246	871	419
10	3,823,786	3,819,395	1,015,610	1,812	731
11	3,877,543	3,872,914	1,049,921	121	69
12	3,698,098	3,694,785	928,208	924	463
13	2,727,881	2,724,583	547,922	121	69
14	2,539,149	2,535,345	590,453	375	160
15	2,320,474	2,318,207	760,828	323	164
16	2,596,072	2,593,765	703,361	302	156
17	2,227,080	2,174,863	761,605	1,469	685
18	2,171,378	2,169,126	457,391	1,249	514
19	1,751,878	1,750,709	532,527	240	113
20	1,739,315	1,737,721	411,308	1,239	513
21	1,054,447	1,051,751	303,088	379	178
22	1,055,454	1,054,440	410,326	1,066	558
X	3,223,927	3,222,780	271,338	406	212
Y	60,505	60,371	539	122	56
Total	81,102,777	80,975,235	19,517,146	14,708	6,816

Stage 1: data after removing all variants that are not SNPs, or with no REF, or with more than one REF, or where the REF allele is also in the ALT alleles. **Stage 2:** data after those where all individuals have the same allele were removed. **Stage 3:** data after those SNPs that are not part of the genes found from literature search were removed. **Stage 4:** data after pruning by iterative clustering. **Stage 5:** data after pruning by mean mean distance.

SUPPL II. DATASET BREAKDOWN

Table II shows the 26 ethnicities across 5 continents which form the basis of our classification, as well as the number of samples in each group.

TABLE II
 POPULATIONS AND CONTINENTS REPRESENTED IN SAMPLES FROM 1000 GENOME PROJECT

Code	Ethnicity	Continent	Ethnicity Sample Size	Continent Sample Size
PUR	Puerto Rican	America	104	
CLM	Colombian	America	94	
PEL	Peruvian	America	85	347
MXL	Mexican-American	America	64	
GBR	British	Europe	91	
FIN	Finnish	Europe	99	
IBS	Spanish	Europe	107	503
CEU	CEPH	Europe	99	
TSI	Tuscan	Europe	107	
CHS	Southern Han Chinese	E. Asia	105	
CDX	Dai Chinese	E. Asia	93	
KHV	Kinh Vietnamese	E. Asia	99	504
CHB	Han Chinese	E. Asia	103	
JPT	Japanese	E. Asia	104	
PJL	Punjabi	S. Asia	96	
BEB	Bengali	S. Asia	86	
STU	Sri Lankan	S. Asia	102	489
ITU	Indian	S. Asia	102	
GIH	Gujarati	S. Asia	103	
ACB	African-Caribbean	Africa	96	
GWD	Gambian	Africa	113	
ESN	Esan	Africa	99	
MSL	Mende	Africa	85	661
YRI	Yoruba	Africa	108	
LWK	Luhya	Africa	99	
ASW	African-American SW	Africa	61	

NB: Codes used for the continents in text: AMR for America; EUR for Europe; EAS for East Asia; SAS for South Asia; and AFR for Africa

SUPPL III. ITERATIVE CLUSTERING ALGORITHM

The iterative clustering algorithm is captured more succinctly in Algorithm 1.

SUPPL IV. FACE PARTS AND ALTERNATIVE TERMS

In Table III, we show how each face region is represented in each chromosome. There we provide two numbers: the number of SNPs linked to only the given face region, and the number

Algorithm 1 Pruning by iterative clustering

Input: An $n \times m$ matrix, \mathbf{S} , of n SNPs with their population-wide distribution; chunk size, c ; minimum cluster size, d .

Output: \mathbf{S}_{rem} , a $p \times m$ matrix of p SNPs ($p \leq c$).

```
1: repeat
2:   chunks  $\leftarrow$  breakIntoChunks( $\mathbf{S}, c$ )
3:    $\mathbf{R} \leftarrow$  new collection() // list of remnants
4:   for each chunk in chunks do
5:     remnants  $\leftarrow$  pruneByClustering(chunk,  $d$ )
6:     append remnants to  $\mathbf{R}$ 
7:    $\mathbf{S} \leftarrow$  mergeRemnants( $\mathbf{R}$ )
8:    $n \leftarrow$  number of rows in  $\mathbf{S}$ 
9: until  $n \leq c$ 
10:  $\mathbf{S}_{rem} \leftarrow$  pruneByClustering( $\mathbf{S}, d$ )
    return  $\mathbf{S}_{rem}$ 
```

of SNPs linked to multiple face regions, including the given region.

SUPPL V. RESULTS FROM LITERATURE SEARCH

Table IV) shows the breakdown of the results from PubMed query search, grouped by chromosome.

SUPPL VI. FACE PARTS AND ALTERNATIVE TERMS

Table V shows the alternative terms we used for each face part when constructing queries for PubMed.

SUPPL VII. MEAN-MEAN-DISTANCE (MMD) ALGORITHM

Algorithm 2 captures the MMD algorithm succinctly.

Algorithm 2 Pruning by mean mean distance (MMD)

Input: An $n \times m$ matrix, \mathbf{S} , of n SNPs with their population-wide distribution.

Output: \mathbf{S}_{rem} , an $p \times m$ matrix of p SNPs ($p < n$), each of which has mean distance above the MMD threshold.

```
1:  $\mathbf{G} \leftarrow$  pairwiseDistance( $\mathbf{S}, \text{metric}=\text{JSDist}$ )
2: for  $i \leftarrow 1$  to  $n$  do
3:    $\mu_i \leftarrow$  computeMean( $\mathbf{G}_i$ )
4:  $\bar{\mu} \leftarrow$  computeMean( $\mu$ )
5:  $\mathbf{S}_{rem} \leftarrow$  new collection()
6: for  $i \leftarrow 1$  to  $n$  do
7:   if  $\mu_i > \bar{\mu}$  then
8:     append  $\mathbf{S}[i]$  to  $\mathbf{S}_{rem}$ 
9: sort  $\mathbf{S}_{rem}$  by mean,  $\mu$ , in descending order
    return  $\mathbf{S}_{rem}$ 
```

SUPPL VIII. MANUAL LITERATURE VALIDATIONS

Table VI shows a sample of the genes we found through our automated literature search strategy. For each gene, we show its function as reported by the relevant literature and provide references.

REFERENCES

- [1] S. Shen *et al.*, “Seven-cpg-based prognostic signature coupled with gene expression predicts survival of oral squamous cell carcinoma,” *Clinical Epigenetics*, vol. 9, no. 1, p. 88, 2017.
- [2] L. Tosca *et al.*, “Genotype–phenotype correlation in 13q13.3–q21.3 deletion,” *European Journal of Medical Genetics*, vol. 54, no. 5, e489–e494, 2011.
- [3] N. F. Pedro *et al.*, “Candidate biomarkers for oral squamous cell carcinoma: Differential expression of oxidative stress-related genes,” *Asian Pacific journal of cancer prevention : APJCP*, vol. 19, no. 5, pp. 1343–1349, 2018, PMC6031819[pmcid].
- [4] N. Han, Z. Chen, and Q. Zhang, “Expression of klf5 in odontoblastic differentiation of dental pulp cells during in vitro odontoblastic induction and in vivo dental repair,” *International Endodontic Journal*, vol. 50, no. 7, pp. 676–684, 2017.
- [5] C. Chen *et al.*, “Genetic variants of mgmt, rhpn2, and fam49a contributed to susceptibility of nonsyndromic orofacial clefts in a chinese population,” *Journal of Oral Pathology & Medicine*, vol. 47, no. 8, pp. 796–801, 2018.
- [6] R. W. J. Collin *et al.*, “Mutations of esrrb encoding estrogen-related receptor beta cause autosomal-recessive nonsyndromic hearing impairment dfnb35,” *American journal of human genetics*, vol. 82, no. 1, pp. 125–138, 2008, S0002-9297(07)00014-6[PII].
- [7] A. L. Fletcher *et al.*, “Observations regarding retinopathy in mitochondrial trifunctional protein deficiencies,” *Molecular Genetics and Metabolism*, vol. 106, no. 1, pp. 18–24, 2012.
- [8] W. D’Souza, S. Pradhan, and D. Saranath, “Multiple single nucleotide polymorphism analysis and association of specific genotypes in fhit, samd4a, and ankrd17 in indian patients with oral cancer,” *Head & Neck*, vol. 39, no. 8, pp. 1586–1595, 2017.
- [9] M. Crompton *et al.*, “A mutation in nischarin causes otitis media via limk1 and nf- κ b pathways,” *PLOS Genetics*, vol. 13, no. 8, pp. 1–28, 2017.
- [10] R. L. Shah *et al.*, “A genome-wide association study of corneal astigmatism: The cream consortium,” *Molecular vision*, vol. 24, pp. 127–142, 2018, PMC5800430[pmcid].
- [11] L. Huang *et al.*, “Association of coding and utr variants in the known regions with wet age-related macular degeneration in han chinese population,” *Journal of Human Genetics*, vol. 63, no. 10, pp. 1055–1070, 2018.
- [12] N. Urraca *et al.*, “Significant transcriptional changes in 15q duplication but not angelman syndrome deletion stem cell-derived neurons,” *Molecular Autism*, vol. 9, no. 1, p. 6, 2018.
- [13] X. Li *et al.*, “Genome-wide linkage study suggests a susceptibility locus for isolated bilateral microtia on 4p15.32–4p16.2,” *PLOS ONE*, vol. 9, no. 7, pp. 1–8, 2014.
- [14] J. S. Bae *et al.*, “Genetic association analysis of ciita variations with nasal polyp pathogenesis in asthmatic patients,” *Molecular Medicine Reports*, vol. 7, no. 3, pp. 927–934, 2013.
- [15] J. L. Wiggs and L. R. Pasquale, “Genetics of glaucoma,” *Human Molecular Genetics*, vol. 26, no. R1, R21–R27, 2017.
- [16] M.-J. An *et al.*, “Transcriptome analysis for UVB-induced phototoxicity in mouse retina,” *Environmental Toxicology*, vol. 33, no. 1, pp. 52–62, 2018.
- [17] M. Hayashi and T. Suzuki, “Oculocutaneous albinism type 4,” in *GeneReviews® [Internet]*, M. Adam *et al.*, Eds., Updated 2017 Sep 7, Seattle (WA): University of Washington, Seattle, pp. 1993–2019.
- [18] Y. Peng *et al.*, “Downregulation and aberrant localization of forkhead box j1 in allergic nasal mucosa,” *International Archives of Allergy and Immunology*, vol. 176, no. 2, pp. 115–123, 2018.
- [19] Deciphering Developmental Disorders Study, “Large-scale discovery of novel genetic causes of developmental disorders,” *Nature*, vol. 519, no. 7542, pp. 223–228, 2015, nature14135[PII].
- [20] H. Kasama *et al.*, “Adenosine a2b receptor promotes progression of human oral cancer,” *BMC Cancer*, vol. 15, no. 1, p. 563, 2015.
- [21] C. Halgren *et al.*, “Haploinsufficiency of celf4 at 18q12.2 is associated with developmental and behavioral disorders, seizures, eye manifestations, and obesity,” *European Journal Of Human Genetics*, vol. 20, 1315 EP -, 2012, Short Report.
- [22] K. Ma *et al.*, “Clinical features and linkage analysis for a chinese family with autosomal dominant central areolar choroidal dystrophy,” *Chinese Medical Journal*, vol. 122, no. 22, 2009.

TABLE III
IDENTIFIED FACE SNPs COUNTED PER FACE REGION AND PER CHROMOSOME

Chromosome	eye	ear	nose	mouth	cheek	chin	forehead	face shape	Count*
1	49 (239)	4 (112)	2 (99)	2 (149)	0 (37)	20 (151)	0 (65)	1 (31)	283
2	60 (211)	5 (113)	2 (74)	12 (129)	0 (90)	6 (102)	3 (57)	0 (37)	262
3	11 (50)	0 (28)	0 (22)	16 (45)	2 (3)	7 (28)	1 (14)	2 (3)	80
4	17 (119)	2 (96)	5 (104)	13 (159)	0 (46)	6 (101)	4 (48)	0 (13)	210
5	32 (187)	4 (127)	3 (92)	20 (150)	2 (54)	9 (141)	0 (94)	0 (31)	268
6	51 (354)	2 (217)	5 (167)	24 (243)	5 (164)	32 (229)	5 (147)	0 (32)	436
7	14 (76)	2 (31)	3 (35)	8 (61)	0 (16)	0 (46)	0 (26)	0 (6)	99
8	8 (92)	0 (39)	0 (47)	9 (95)	0 (6)	9 (44)	1 (35)	0 (14)	118
9	45 (342)	9 (224)	0 (158)	12 (252)	0 (81)	21 (271)	0 (83)	0 (17)	419
10	115 (558)	3 (283)	0 (196)	14 (370)	0 (139)	17 (354)	0 (132)	0 (46)	731
11	3 (55)	0 (32)	0 (33)	1 (47)	0 (14)	2 (36)	0 (25)	0 (2)	69
12	42 (277)	2 (159)	9 (198)	39 (290)	0 (72)	42 (206)	8 (134)	0 (20)	463
13	6 (54)	4 (33)	0 (24)	4 (41)	0 (6)	5 (37)	1 (13)	0 (0)	69
14	38 (123)	7 (50)	10 (50)	8 (56)	0 (4)	2 (49)	1 (17)	0 (1)	160
15	32 (147)	3 (55)	1 (49)	3 (86)	0 (26)	5 (91)	1 (53)	0 (1)	164
16	31 (97)	9 (44)	2 (44)	13 (50)	0 (15)	18 (80)	1 (33)	0 (8)	156
17	70 (421)	7 (220)	34 (185)	54 (377)	0 (129)	67 (326)	26 (165)	0 (37)	685
18	96 (371)	7 (152)	3 (156)	72 (257)	2 (74)	13 (211)	0 (101)	0 (9)	514
19	14 (67)	3 (39)	1 (27)	20 (59)	0 (16)	20 (57)	0 (20)	0 (2)	113
20	115 (424)	3 (187)	18 (202)	10 (263)	0 (57)	26 (173)	16 (136)	3 (41)	513
21	10 (118)	0 (89)	0 (94)	10 (120)	0 (38)	32 (130)	0 (48)	0 (12)	178
22	95 (374)	1 (166)	20 (217)	9 (262)	0 (25)	64 (276)	7 (125)	0 (19)	558
X	29 (155)	0 (69)	5 (73)	1 (112)	0 (37)	17 (144)	1 (61)	5 (14)	212
Y	15 (45)	0 (8)	0 (15)	0 (8)	0 (8)	3 (26)	0 (0)	0 (0)	56
Total	998 (4956)	77 (2573)	123 (2361)	374 (3681)	11 (1157)	443 (3309)	76 (1632)	11 (396)	6816

Table shows how the identified face SNPs (see *Stage 5* column in Table I) are distributed across face regions. Some SNPs feature in more than one face parts. For each face part, SNPs are counted when they match only the face part (shown by the number **not** in parenthesis) and when they match the face part and at least one more (shown by the number in parenthesis). The last column is reproduced from *Stage 5* column in Table I and shows the total number of face-related SNPs identified for the given chromosome.

- [23] B. Coste *et al.*, "Gain-of-function mutations in the mechanically activated ion channel *piezo2* cause a subtype of distal arthrogyrosis," *Proceedings of the National Academy of Sciences*, vol. 110, no. 12, pp. 4667–4672, 2013.
- [24] M. Rodríguez Pulido *et al.*, "Foot-and-mouth disease virus infection induces proteolytic cleavage of *ptb*, *eif3a,b*, and *pabp* rna-binding proteins," *Virology*, vol. 364, no. 2, pp. 466–474, 2007.
- [25] A. Sultana *et al.*, "Olfactomedin 2: Expression in the Eye and Interaction with Other Olfactomedin Domain-Containing Proteins," *Investigative Ophthalmology & Visual Science*, vol. 52, no. 5, pp. 2584–2592, 2011.
- [26] K. Szczaluba *et al.*, "High-resolution array comparative genomic hybridization utility in polish newborns with isolated cleft lip and palate," *Neonatology*, vol. 107, no. 3, pp. 173–178, 2015.
- [27] Y. Chen, P. Sternberg, and J. Cai, "Characterization of a Bcl-XL-Interacting Protein FKBP8 and Its Splice Variant in Human RPE Cells," *Investigative Ophthalmology & Visual Science*, vol. 49, no. 4, pp. 1721–1727, 2008.
- [28] S. G. Younkin *et al.*, "A genome-wide study of inherited deletions identified two regions associated with nonsyndromic isolated oral clefts," *Birth Defects Research Part A: Clinical and Molecular Teratology*, vol. 103, no. 4, pp. 276–283, 2015.
- [29] G. Jedraszak *et al.*, "Clinical and molecular characterization of the 20q11.2 microdeletion syndrome: Six new patients," *American Journal of Medical Genetics Part A*, vol. 167, no. 3, pp. 504–511, 2015.
- [30] Z. W. E. Yong *et al.*, "Genetic alterations of chromosome 8 genes in oral cancer," *Scientific Reports*, vol. 4, 6073 EP -, 2014, Article.
- [31] G. A. Stamatiou and K. M. Stankovic, "A comprehensive network and pathway analysis of human deafness genes," *Otology & Neurotology*, vol. 34, no. 5, 2013.
- [32] C. Luna *et al.*, "Regulation of trabecular meshwork cell contraction and intraocular pressure by mir-200c," *PLOS ONE*, vol. 7, no. 12, pp. 1–9, 2012.
- [33] H. Huang *et al.*, "Expression of *microrna-10a*, *microrna-342-3p* and their predicted target gene *tiam1* in extranodal nk/t-cell lymphoma, nasal type," *Oncology letters*, vol. 11, no. 1, pp. 345–351, 2016, OL-0-0-3831[PII].
- [34] G. Giroto *et al.*, "Hearing function and thresholds: A genome-wide association study in european isolated populations identifies new loci and pathways," *Journal of Medical Genetics*, vol. 48, no. 6, pp. 369–374, 2011.
- [35] C.-M. Chung *et al.*, "Combined genetic biomarkers and betel quid chewing for identifying high-risk group for oral cancer occurrence," *Cancer Prevention Research*, vol. 10, no. 6, pp. 355–362, 2017.
- [36] D. Fine *et al.*, "A syndrome of congenital microcephaly, intellectual disability and dysmorphism with a homozygous mutation in *frmd4a*," *European Journal Of Human Genetics*, vol. 23, 1729 EP -, 2014, Short Report.
- [37] Y.-Z. Kuo *et al.*, "Mir-99a exerts anti-metastasis through inhibiting myotubularin-related protein 3 expression in oral cancer," *Oral Diseases*, vol. 20, no. 3, e65–e75, 2014.
- [38] N. Okamoto *et al.*, "Deletion at chromosome 10p11.23-p12.1 defines characteristic phenotypes with marked midface retrusion," *Journal Of Human Genetics*, vol. 57, 191 EP -, 2012, Original Article.
- [39] V. R. M. Chavali *et al.*, "Association of oct derived drusen measurements with amd associated-genotypic snps in amish population," *Journal of clinical medicine*, vol. 4, no. 2, pp. 304–317, 2015, PMC4398021[pmcid].
- [40] A. T. Fung *et al.*, "New *best1* mutations in autosomal recessive bestrophinopathy," *Retina (Philadelphia, Pa.)*, vol. 35, no. 4, pp. 773–782, 2015, PMC4425987[pmcid].

TABLE IV
LITERATURE SEARCH FOR GENES AND FACE PARTS

Chr	# variants (Ensembl)	# variants (1KGP)	# genes queried	# genes with hits
1	13,660,159	6,468,094	5,317	1,139
2	12,573,660	7,081,600	4,006	724
3	10,526,448	5,832,276	3,050	621
4	8,476,803	5,732,585	2,510	445
5	9,325,049	5,265,763	2,878	494
6	9,259,954	5,024,119	2,907	577
7	9,774,376	4,716,715	2,917	497
8	7,679,737	4,597,105	2,386	415
9	7,377,552	3,560,687	2,275	407
10	8,013,735	3,992,219	2,240	411
11	8,646,594	4,045,628	3,285	669
12	7,824,609	3,868,428	2,969	608
13	4,488,878	2,857,916	1,335	188
14	4,880,784	2,655,067	2,230	336
15	6,239,380	2,424,689	2,179	325
16	5,611,467	2,697,949	2,513	440
17	6,574,456	2,329,288	3,024	648
18	3,635,781	2,267,185	1,183	173
19	4,481,178	1,832,506	2,959	661
20	3,320,546	1,812,841	1,397	300
21	2,400,725	1,105,538	833	118
22	3,317,760	1,103,547	1,353	262
X	5,076,300	3,468,093	2,375	456
Y	21,277	62,042	518	15
Total	163,187,208	84,801,880	58,639	10,929

TABLE V
FACE REGIONS AND ALTERNATIVE TERMS

Face region	Alternative terms
eye	iris, pupil, sclera, eyebrow, eyelid, orbit, cornea, lens, epicanthic, eyelash, orbicularis oculi
nose	nostril, ala nasi, nasalis, nasal, glabella
mouth	rima oris, lip, philtrum, orbicularis oris
ear	earlobe, pinna, auricle, auricula, tragus, antitragus
cheek	nasolabial, buccal, cheekbone
chin	mentolabial sulcus, mandible, jaw, mental region, mandible
forehead	temple, frontalis, brow ridge, supraorbital ridge, superciliary arch
face shape	facial shape

mous cell carcinoma," *Head & Neck*, vol. 37, no. 12, pp. 1816–1822, 2015.

- [45] A. G. Fincham *et al.*, "Isolation and partial characterization of a human amelogenin from a single fetal dentition using HPLC techniques," *Calcified tissue international*, vol. 47, no. 2, pp. 105–111, 1990.
- [46] A. M. Heikkinen *et al.*, "Pilot study on the genetic background of an active matrix metalloproteinase-8 test in finnish adolescents," *Journal of Periodontology*, vol. 88, no. 5, pp. 464–472, 2017.
- [47] J. L. Llorente *et al.*, "Nasosinusal adenocarcinoma: Molecular and genetic analysis by MLPA," *Acta Otorrinolaringologica (English Edition)*, vol. 59, no. 4, pp. 151–158, 2008.
- [48] National Center for Biotechnology Information, *Gene*, <https://www.ncbi.nlm.nih.gov/gene>.

- [41] J.-D. Cha, H. J. Kim, and I.-H. Cha, "Genetic alterations in oral squamous cell carcinoma progression detected by combining array-based comparative genomic hybridization and multiplex ligation-dependent probe amplification," *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontics*, vol. 111, no. 5, pp. 594–607, 2011.
- [42] M. L. Kwee *et al.*, "An autosomal dominant high bone mass phenotype in association with craniosynostosis in an extended family is caused by an *lrp5* missense mutation," *Journal of Bone and Mineral Research*, vol. 20, no. 7, pp. 1254–1260, 2005.
- [43] S. David *et al.*, "Hypotonia and delayed motor development as an early presentation of lowe syndrome: Case report and literature review," *Acta Clinica Belgica*, vol. 0, no. 0, pp. 1–5, 2018, PMID: 30501482.
- [44] K. Duś-Szachniewicz *et al.*, "Protein tyrosine phosphatase receptor r and z1 expression as independent prognostic indicators in oral squa-

TABLE VI
MANUAL LITERATURE VALIDATION OF SELECTED GENES

Chr	Genes	Function (from literature)	Region	Ref	Chr	Genes	Function (from literature)	Region	Ref
1	AJAP1	Associated with Oral Squamous Cell Carcinoma (OSCC)	mouth	[1]	13	DCLK1	Associated with macrocephaly, high forehead, hypertelism, large nose, large and malformed ears and retrognathia	all	[2]
	DHCR24	Down-regulated in Oral Squamous Cell Carcinoma cases	mouth	[3]		KLF5	Associated with odontoblastic differentiation of dental pulp cells	mouth	[4]
2	FAM49A	Associated with nonsyndromic orofacial clefts	mouth	[5]	14	ESRRB	Essential for inner-ear development and function	ear	[6]
	HADHB	Associated with retinopathy	eye	[7]		SAMD4A	Associated with oral cancer	mouth	[8]
3	NISCH	Mutation associated with chronic otitis media	ear	[9]	15	TNFAIP8L3	Candidate gene for corneal astigmatism	eye	[10]
	BBX	Associated with age-related macular degeneration	eye	[11]		HERC2	Associated with Angelman Syndrome	face shape, forehead	[12]
4	EVC2	Association with microtia, Ellis-van Creveld syndrome or Weyers acrocardial dysostosis	ear, mouth, face shape	[13]	16	CIITA	Associated with nasal polyposis	nose	[14]
	AFAP1	Associated with primary-open-angle glaucoma	eye	[15]		RNF40	May affect how the retina absorbs UV radiations	eye	[16]
5	SLC45A2	Associated with Oculocutaneous Albinism Type 4	eye	[17]	17	DNAH9	Associated with allergic rhinitis	nose	[18]
	COL4A3BP	Associated with several developmental disabilities	face shape	[19]		ADORA2B	May be a key regulator of tumoral progression in OSCCs	mouth	[20]
6	GMDS	Associated with primary-open-angle glaucoma	eye	[15]	18	CELF4	An important role in eye development	eye	[21]
	RIMS1	Associated with Central Areolar Choroidal Dystrophy	eye	[22]		PIEZO2	Mutation can cause a subtype of Distal Arthrogyposis Type 5	eye	[23]
7	EIF3B	Plays a role in Foot-and-mouth disease virus infection	mouth	[24]	19	OLFM2	Associated with glaucoma	eye	[25]
	CHN2	Candidate gene for cleft lip and palate	mouth	[26]		FKBP8	Associated with loss of retinal pigment epithelial cells	eye	[27]
8	ADAM3A	May be involved in the etiology of oral clefts.	mouth	[28]	20	GDF5	Associated with craniofacial dysmorphism	forehead	[29]
	CSGALNACT1	Associated with Oral Squamous Cell Carcinoma	mouth	[30]		HNF4A	Associated with genetic hearing loss	ear	[31]
9	LPAR1	Affects regulation of trabecular meshwork (TM) cell networks	eye	[32]	21	TIAM1	May contribute to the pathogenesis of extranodal natural killer (NK)/T-cell lymphoma, nasal type (ENKTCL)	nose	[33]
	PTRPD	Affects hearing functions	ear	[34]		HSPA13	Associated with OSCC occurrence	mouth	[35]
10	FRMD4A	Associated with congenital microcephaly	face shape	[36]	22	MTMR3	Possible therapeutic target for oral cancer treatment	mouth	[37]
	MPP7	Potentially associated with maxillofacial abnormalities	mouth	[38]		SYN3	Associated with larger area perifoveal ring and larger drusen volume	eye	[39]
11	BEST1	Mutations cause autosomal recessive bestrophinopathy	eye	[40]	X	CHRD1	Associated with Oral Squamous Cell Carcinoma	mouth	[41]
	LRP5	Missense mutation cause macrocephaly, craniosynostosis	all	[42]		OCRL	Associated with Lowe syndrome	eye	[43]
12	PTPRR	Potentially associated with oral cancer	mouth	[44]	Y	AMELY	Associated with tooth enamel development	mouth	[45]
	VDR	Polymorphisms linked to initial periodontitis	mouth	[46]		UTY	Associated with nasosinus adenocarcinoma	nose	[47]

NB: The functions are summaries from *NCBI Gene* [48]. The references our system relied on to associate the genes with the face regions are listed alongside.