Am J Transl Res 2022;14(5):3464-3471
www.ajtr.org /ISSN:1943-8141/AJTR0141165

Original Article

Empirical assessment of allele frequencies of genome wide association study variants associated with obstructive sleep apnea

Mohit[1,2*], Indu Sharma[3*], Varun Sharma[3*], Sachin Kumar[3], Garima Rastogi[4], Pranjali Dutt[1], Ashutosh Shrivastava[2], Niraj Rai[3], Pooran Chand[1]

[1]Department of Prosthodontics, King George's Medical University, Lucknow 226003, Uttar Pradesh, India; [2]Center for Advance Research, Faculty of Medicine, King George's Medical University, Lucknow 226003, Uttar Pradesh, India; [3]Ancient DNA Laboratory, Birbal Sahni Institute of Palaeosciences, Lucknow, Uttar Pradesh, India; [4]NMC Genetics India Pvt Ltd., Gurugram 122002, Haryana, India. [*]Equal contributors.

Received December 16, 2021; Accepted April 26, 2022; Epub May 15, 2022; Published May 30, 2022

Abstract: Objective: Obstructive Sleep Apnea (OSA) is a heterogeneous disorder with a complex interplay of genetic and environmental factors. Over the years, with advancement in genotyping and sequencing techniques, various loci have shown an association with OSA. It is pertinent to understand the status of these associated variants in different ethnic groups. The aim of the study was to assess the genetic affinity among different population groups by evaluating the risk allele frequencies of variants associated with OSA. Method: The variants associated with OSA were obtained from the GWAS catalog with a significant *p* value of $<5 \times 10^{-7}$; 95 variants were obtained (www.ebi.ac.uk/gwas). Further, the variants were narrowed down on the basis of risk allele frequencies (>5%). The *fst* was calculated to assess the genetic affinity between super population groups and among the sub-population groups present in the 1000 genome project. Result: The *fst* values observed indicated all super populations were genetically related (SAS, AMR, EAS and EUR) except in the African (AFR) population group. Further, the closely related super population *i.e.,* SAS, AMR, EAS and EUR when bifurcated on the basis of sub-population groups shows population stratification and SAS population groups form separate clusters on the MDS plot. Conclusion: The study highlights genetic heterogeneity among different population groups that gets diluted and results are biased when the samples are pooled irrespective of their endogamous groups. Our results provide insight to researchers to target specific endogamous groups for future studies on OSA.

Keywords: Obstructive sleep apnea, genetic affinity, genome-wide association studies, genetic variants, ethnicities, geographic origin

Introduction

OSA is a serious respiratory sleep disorder that is associated with systemic and cardiovascular, respiratory, and neurological dysfunction [1]. The condition is diagnosed with repetition in full or partial episodes of collapsing of the upper airway while sleeping, resulting in reduced airflow. Untreated OSA leads to health consequences in the long term [2]. A recent epidemiological study predicts that the frequency of OSA in middle-aged men and women is higher [3]. In India, a new community-based investigation found the prevalence of OSA to be 9.3% [4]. This prevalence rate is parallel to the western population [5]. Moreover, the prevalence of OSA is higher in patients with cardiovascular and metabolic disorders [6]. OSA has several classic signs and symptoms; the most prevalent are repetitive episodes of complete or fractional upper airway obstruction during sleep with snoring and daytime sleepiness. Intermittent hypoxia and nocturnal hypertension have also been reported as self-determining risk factors in these patients [7].

OSA patients suffer from daytime fatigue, excessive drowsiness, insomnia, and increased risk of road mishaps [8]. Risk factors of OSA include narrowed airway, diabetes, arterial hy-

pertension, stroke, systemic inflammation, cognitive impairment, depression, and metabolic diseases [6]. The etiology is multifaceted and is not confined to airway obstruction, but is the result of several interconnected consequences and supportive risk factors of which the most common factors for developing OSA are excess weight, neck circumference, age, and gender [8, 9]. These factors raise the prevalence of OSA as we are observing the global obesity pandemic and also due to the ageing of societies. The growing prevalence of the disorder is a challenge for diagnosis especially in developing countries as the diagnostic criteria requires overnight monitoring of sleep breathing at high cost [10].

Presently, the most widely used diagnostic tool is polysomnography, which diagnoses sleep disorder by observing sleep quality and sleep patterns. These patterns are measured by apnea-hypopnea index (AHI), when the apnea-hypopnea index (AHI) goes beyond a certain threshold, these parameters are often considered "severity of disease" [11, 12]. There are a number of treatments available (reversible and irreversible) depending on the patient's symptoms/severity and specific pathologic character. The most common treatment comprises the use of continuous positive airway pressure (CPAP) or a mandibular advancement device (MAD); depending on the severity of the disease, medical devices are used accordingly [13]. Behavioral therapies include weight loss and several lifestyle changes. Irreversible treatments include a variety of surgeries [14]. However, today, surgical choices are less in demand. In their place, the CPAP and MAD are treatment choices provided by clinicians. The gold standard treatment option is CPAP, nevertheless, acceptance and compliance are low, and oral device therapy (OA) has been found to be more pertinent than CPAP [15]. Likewise, intervention-based studies depicted favorable effects of MAD in improving AHI scores by reducing the rate of obstruction [16-18].

The connection between physiologic sleep and its genetic association was first conveyed in the 1930s which concluded that genetic factors under certain situations can develop OSA [19]. Additionally, a family-based twin study put forward that there is a noteworthy association between OSA and inherited craniofacial disorder [20].

Certain gene and molecular biomarkers affect the sleep cycle and other crucial body functions [21]. The divergent etiology of OSA is associated with various factors such as age, sex, neck circumference, fat accumulation, and genes [22, 23]. Detecting the involved genetic factors contemplate as a superlative approach to conduct the study to acquire more about the pathogenesis of the OSA. However, very limited genomic studies have been accomplished on OSA to date [24]. Therefore, the Genome Wide Association (GWA) and candidate gene association study approach are useful for OSA to explore the association of genetic traits and related complications. Still, most of the genetic studies were conducted in Europeans, East Asians and Hispanics whereas such studies are lacking in other ethnicities. Therefore, considering ethnicity as one of the prominent aspects associated with the genetic predisposition of OSA, we aim to evaluate the genetic affinity of effective alleles associated with OSA among global populations.

Methodology

Data collection and statistical analysis

The variants specific to OSA were obtained from the GWAS catalog with a significant *p*-value of $P<5 \times 10^{-7}$; 95 variants were obtained (www.ebi.ac.uk/gwas) [25]. The variants were further narrowed down for the study with criteria to include the variants with frequency less than 5% in South Asian population groups. The reason to select variants with more than 5% minor allele frequency was to determine the common variants in South Asians. Additionally, significant expressions in the tissues associated with OSA were also considered and the expression of the genes was obtained by the Genotype-Tissue Expression (GTEx) (https://gtexportal.org/) [26]. The variants with NES (positive as well as negative) and significant *p*-value determined the functional relevance of the variant with respect to tissues that were associated with OSA and traits associated with OSA. Following these criteria, 21 variants were narrowed down (**Table 1**) and 74 variants were excluded from the study. The allele frequencies of the risk allele (GWAS; **Table 1**) were docu-

Table 1. List of selected GWAS variants with risk allele and their putative expression in human tissues

SNP ID	Risk Allele	Chromo-Location	GWAS *p* value	Region	Gene/Locus	Tissue expression in GtexPoratal (*p* value; NES)
rs1156250	G	Chr 2:204453088	8×10^{-6}	regulatory region variant	MPPED2	Tissue-Muscle Skeletal (4.2e-21; 0.294)
rs394452	A	Chr 2:219055315	4×10^{-6}	synonymous variant	IHH	Tissue-Adipose Subcutaneous (4.2e-4; 0.237)
rs158258	G	Chr 6:156880797	6×10^{-6}	intron variant	ARID1B	Tissue-Heart Appendage (0.5; 0.0476)
rs322818	A	Chr 7:128096993	9×10^{-6}	intergenic variant	SND1	Tissue-Esophagus Mucosa (8.8e-4; 0.0696)
rs10767859	T	Chr 11:30549800	6×10^{-6}	intron variant	MPPED2	Tissue-Muscle Skeletal (4.2e-21; 0.294)
rs1793596	A	Chr 11:131143730	6×10^{-6}	intergenic variant	RN7SL167P	----
rs1793821	T	Chr 11:131149976	5×10^{-6}	intergenic variant	RN7SL167P	----
rs2573647	A	Chr 15:99933493	5×10^{-6}	intergenic variant	AC084855.1, AC090825.1	-----
rs2581360	A	Chr 15:99951677	9×10^{-6}	regulatory region variant	AC084855.1, AC090825.1	----
rs6117669	A	Chr 20:7189144	2×10^{-7}	intron variant	LINC01428	Tissue-Testis (0.000030; -1.0)
rs5755038	T	Chr 22:34415848	7×10^{-6}	intergenic variant	LINC01643	Tissue-Testis (2.2e-8; 0.21)
rs11588454	T	Chr 1:191832003	3×10^{-6}	intron variant	LINC02770	----
rs999944	A	Chr 2:64822719	2×10^{-7}	intergenic variant	RN7SL341P	----
rs4849682	T	Chr 2:118185307	5×10^{-7}	intron variant	THORLNC	----
rs116791765	T	Chr 11:94347083	2×10^{-8}	intergenic variant	GPR83	Tissue-Heart left ventricle (0.6; 0.0344)
rs11897825	A	Chr 2:21471579	5×10^{-7}	intron variant	AC011752.1, AC067959.1	Tissue-Adipose Subcutaneous (1.4e-32; 0.470)
rs111942351	A	Chr 5:143693985	4×10^{-7}	intron variant	CTB-57H20.1	No significant tissue type found
rs11074782	T	Chr 16:26493690	2×10^{-6}	intergenic variant	AC130464.1	No significant tissue type found
rs2743173	T	Chr 20:8264646	2×10^{-7}	intron variant	PLCB1	Tissue-Adipose Subcutaneous (0.2; 0.0417)
rs72699765	A	Chr 15:25584050	5×10^{-7}	intergenic variant	ATP10A	Tissue-Esophagus Mucosa (9.3e-5; 0.0208)
rs2033354	T	Chr 2:15076492	4×10^{-7}	intergenic variant	NBAS	Tissue-Heart Appendage (1.8e-40; 0.457)

mented for further analysis from the 1000 genome database (http://www.internationalgenome.org) [27].

The populations selected for analyzing pooled and stratified analyses frequencies were: Africa (AFR), South Asia (SAS), Admixed American (AMR), East Asia (EAS) and Europe (EUR), SAS {(Sri Lankan Tamil from UK (STU), Bengali from Bangladesh (BEB), Gujarati Indian from Houston (GIH), Indian Telugu from UK (ITU), Punjabi from Lahore Pakistan (PJL))}, EAS {(Southern Han Chinese (CHS), Chinese Dai in Xishuangbanna, China (CDX))}, and EUR {(Finnish (FIN), Japanese in Tokyo (JPT), British in England and Scotland (GBR))}. Pooled allele frequencies in five super populations and stratified frequencies in sub-populations were obtained from a 1000 Genome database. Further statistical analysis was performed on the basis of the allele frequencies. *FST* was calculated using Arlequin version.3.5.2.2 using 10,000 permutations to avoid the false discovery rate [28]. Slatkin's linearized *FST*'s model was used to calculate the average number of pairwise differences within and between populations. The *fst* values determine the genetic distance among the different population groups and are a key component in studying population structure. The heatmap and frequencies were plotted using R packages [29]. Multi Dimensional Scaling was performed to visualize the similarities among the selected population groups using the *fst* matrix. The detailed strategy adopted for the study is presented in **Figure 1**.

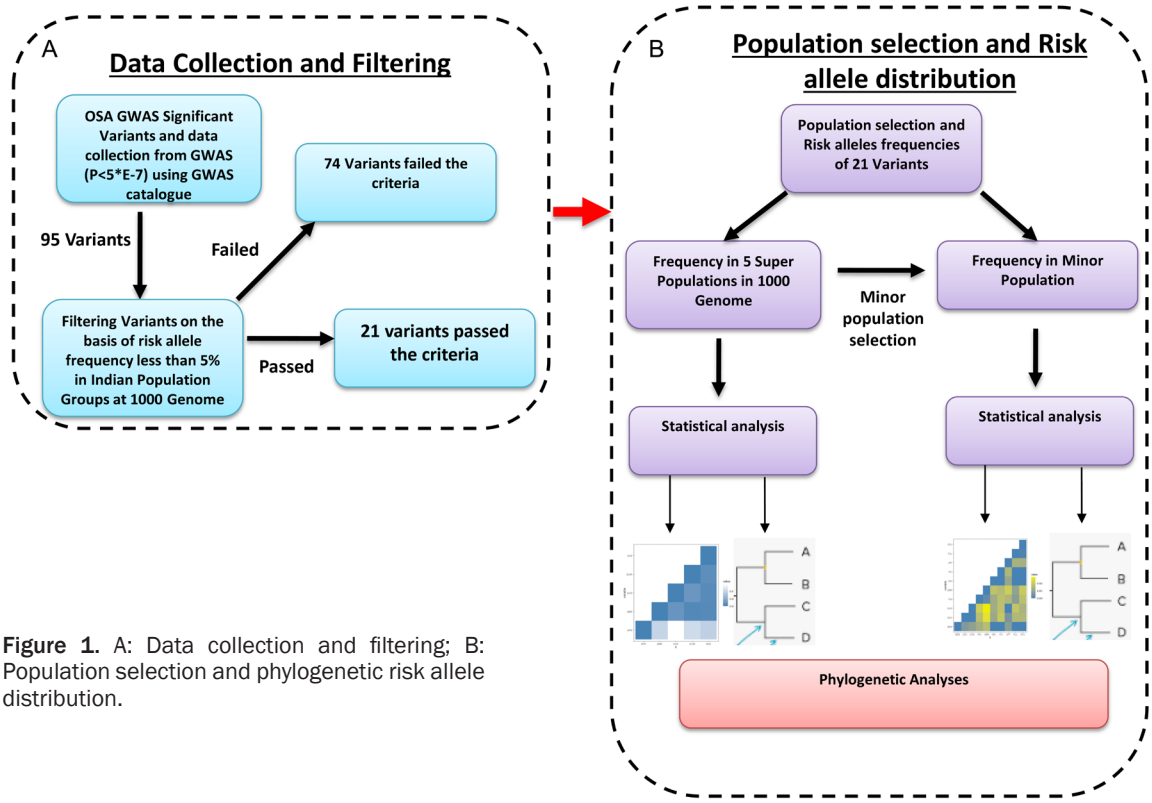


Figure 1. A: Data collection and filtering; B: Population selection and phylogenetic risk allele distribution.

Results

With the increasing incidence and sighting role of genetics in OSA [30], it is pertinent to evaluate the established candidate genes in the study population group. In order to estimate the risk, population based case-control association studies play a vital role [30]. Further, replication of variants identified by GWAS in different ethnic groups is also an essential step in establishing that the associations revealed by initial GWAS are genuine. Failure to replicate is usually interpreted as false-positive interactions and warrants further validation by independent replication studies [31]. It has been observed that many studies of different complex disorders showed redundant results when replicated in other ethnicities, especially in Indian population groups [32-34].

The genetic heterogeneity of candidate gene variants of complex disorders always raises an alarm as the variants observed to be associated with the trait provide redundant results when replicated in other ethnic groups. For instance, variant rs1137101 of LEPR is observed to be significantly associated with OSA and its associated traits like increased sleeping efficiency, less wakefulness after sleep, *etc.* in European Americans. In spite of the evidence, the same variant was not associated with OSA in Asians [35]. Additionally, when we observe the variant frequency in the database for the single nucleotide polymorphisms, there is heterogeneity among the allele frequency of alternate allele and reference allele in different population groups (https://www.ncbi.nlm.nih.gov/snp/rs1137101?horizontal_tab=true).

Further, another variant rs1409986 of *PTGR* was significantly associated with OSA in Europeans [36], whereas it provides contrary results in a study conducted on Mycode biobank participants [37]. To validate this varying frequency we plotted the selected variants and observed that there is shifting of allele frequencies within and among global super populations and in the South Asian population.

The most plausible reason for the genetic heterogeneity and its implication for the genetic studies is a difference in allele frequency that could be an outcome of ancient admixture, or founder effect due to strict endogamous prac-

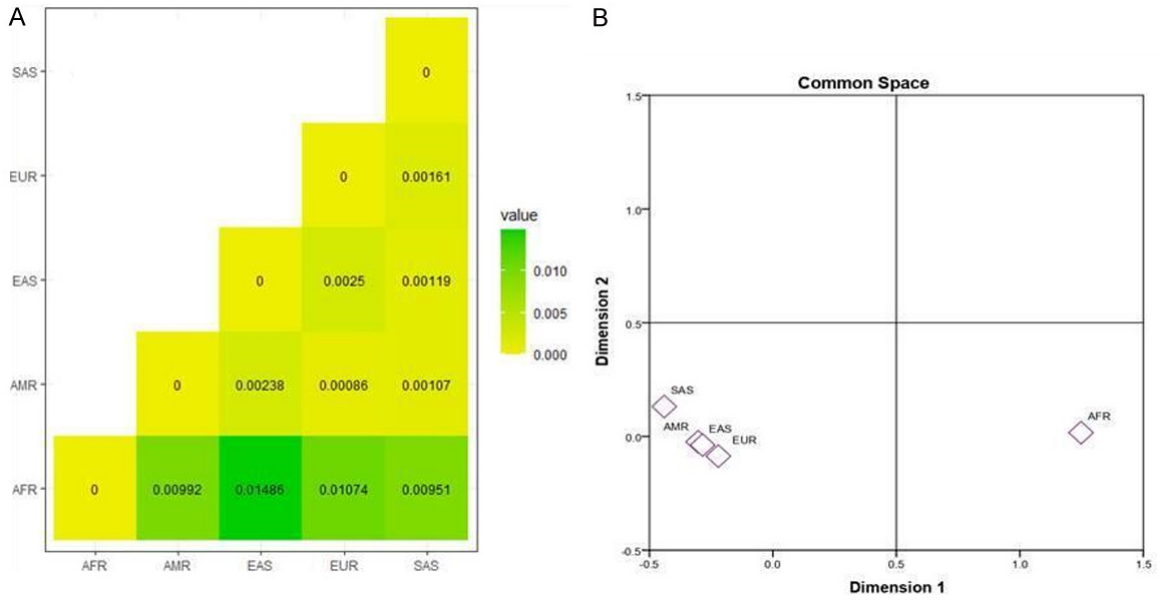


Figure 2. (A) Heat map of pairwise *Fst* and (B) Multi-Dimensional Scaling calculated from the risk variants of GWAS of OSA indicates that the genetic affinity of super populations except Africa is the same when analyzed together.

tices or due to modern human migrations that took place in the past [38, 39].

The frequency distribution of effective variants in super population and other population groups obtained from 1000 genomes is given in Supplementary Figures 1 and 2, respectively. The present study's strategy may aid in the evaluation of genetic differences in population groups with respect to risk alleles in OSA.

In order to visualize the genetic affinity of the variants among the selected population groups, genetic distance matrices were generated on the basis of allele frequencies of risk variants (**Figures 1** and **2**). Using these matrices Multi Dimensional Scaling (MDS) was plotted using *Fst* values [40]. The pairwise *Fst* values were generated for five super populations were Africa (AFR), South Asia (SAS), Admixed from the Americas (AMR), East Asia (EAS) and Europe (EUR), where the four population groups have closer affinity SAS, AMR, EAS and EUR indicating that the variants have an alike effect on the population sets, whereas AFR population set diverged out in the plot indicating genetic differentiation from the other population sets (**Figure 2**).

To have a clearer idea about the genetic affinity in sub-population groups like Sri Lankan Tamil from UK (STU), Bengali from Bangladesh (BEB), Gujarati Indian from Houston (GIH), Indian Telugu from UK (ITU), Punjabi from Lahore Pakistan (PJL), Southern Han Chinese (CHS), Chinese Dai in Xishuangbanna, China (CDX), Japanese in Tokyo (JPT), Finnish (FIN), British in England and Scotland (GBR), when plotted with genetic distances clearly showed differentiation of GIH, ITU, BEB and PJB sub-population groups having different genetic affinity from the rest of the population subsets (**Figure 3**).

This highlights genetic heterogeneity among different population groups which gets diluted, and results are biased when the samples are pooled irrespective of their endogamous groups.

Discussion

To have a better understanding of the susceptibility of a particular variant in the population group it is necessary to carry out association studies in individual endogamous groups irrespective of the linguistic group and geographic origin. Since our population is heterogeneous and is stratified, the results will vary when combined population groups are targeted [32].

The results signify that in order to perform a genetic study on OSA, it is required to evaluate a particular population group independently rather than pooling samples on the basis of

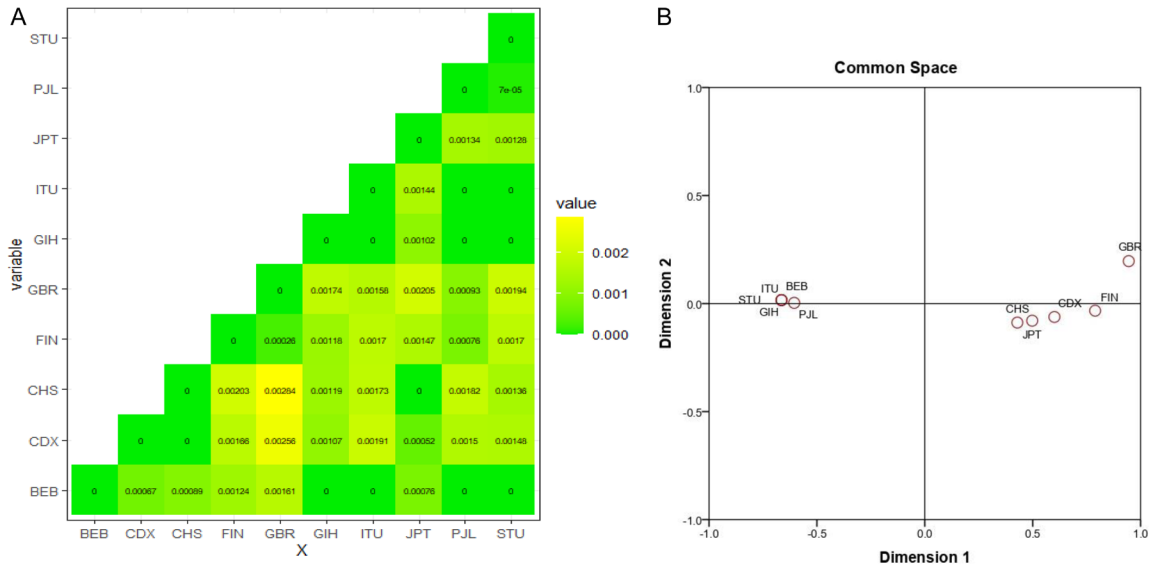


Figure 3. (A) Heat map of pairwise *Fst* and (B) MultiDimensional Scaling of ten sub-population groups confirm the population stratification when analyzed independently.

geography or linguistic groups as each population possesses its own selection and adaptation towards any trait (OSA). Keeping in view the heterogeneity of different ethnic population groups and results observed, it is pertinent to evaluate the risk-causing variants in specific endogamous groups without grouping the samples, and this is a model for future studies on OSA.

We are trying to highlight how the future studies or the studies being conducted on different population groups need to be carried out regarding OSA to have a list of concrete markers affecting certain population groups. This will help in the long run to have a better understanding of the disease and treatment in the future. Further, it is anticipated that there is a dire need for a genetic biobank of South Asian population groups to show the missing heritability of complex disorders including OSA.

Acknowledgements

Mohit thanks to Indian Council of Medical Research, Government of India, for Senior Research Fellowship (ID: 2020-3690).

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Pooran Chand, Department of Prosthodontics, Faculty of Dental Sciences, King George's Medical University, Lucknow 226003, Uttar Pradesh, India. E-mail: dr.pooran_chand@yahoo.com; Dr. Ashutosh Shrivastava, Center for Advance Research, Faculty of Medicine, King George's Medical University, Lucknow 226003, Uttar Pradesh, India. E-mail: shrivastava.as@gmail.com; ashutoshshrivastava@kgmcindia.edu; Niraj Rai, Ancient DNA Laboratory, Birbal Sahni Institute of Palaeosciences, Lucknow 226020, Uttar Pradesh, India. E-mail: nirajrai1@bsip.res.in

References

[1] Young T, Palta M, Dempsey J, Skatrud J, Weber S and Badr S. The occurrence of sleep-disordered breathing among middle-aged adults. N Engl J Med 1993; 328: 1230-1235.

[2] Guilleminault C and Quo SD. Sleep-disordered breathing. A view at the beginning of the new Millennium. Dent Clin North Am 2001; 45: 643-656.

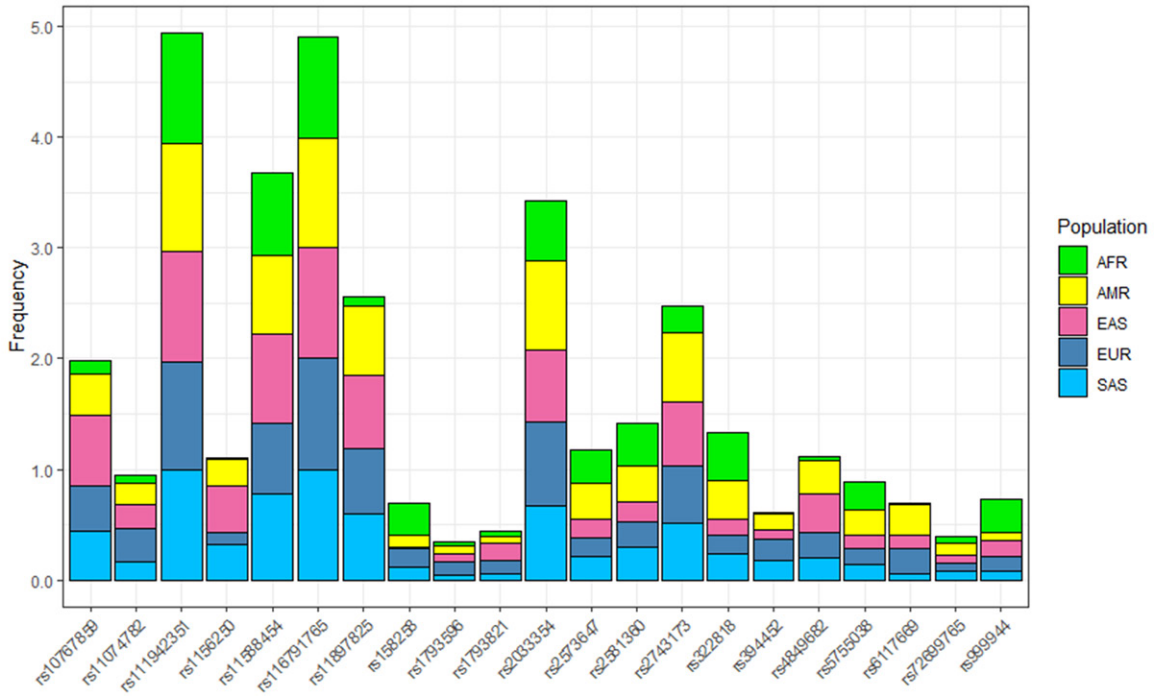
[3] Peppard PE, Young T, Barnet JH, Palta M, Hagen EW and Hla KM. Increased prevalence of sleep-disordered breathing in adults. Am J Epidemiol 2013; 177: 1006-1014.

[4] Reddy EV, Kadhiravan T, Mishra HK, Sreenivas V, Handa KK, Sinha S and Sharma SK. Prevalence and risk factors of obstructive sleep apnea among middle-aged urban Indians: a community-based study. Sleep Med 2009; 10: 913-918.

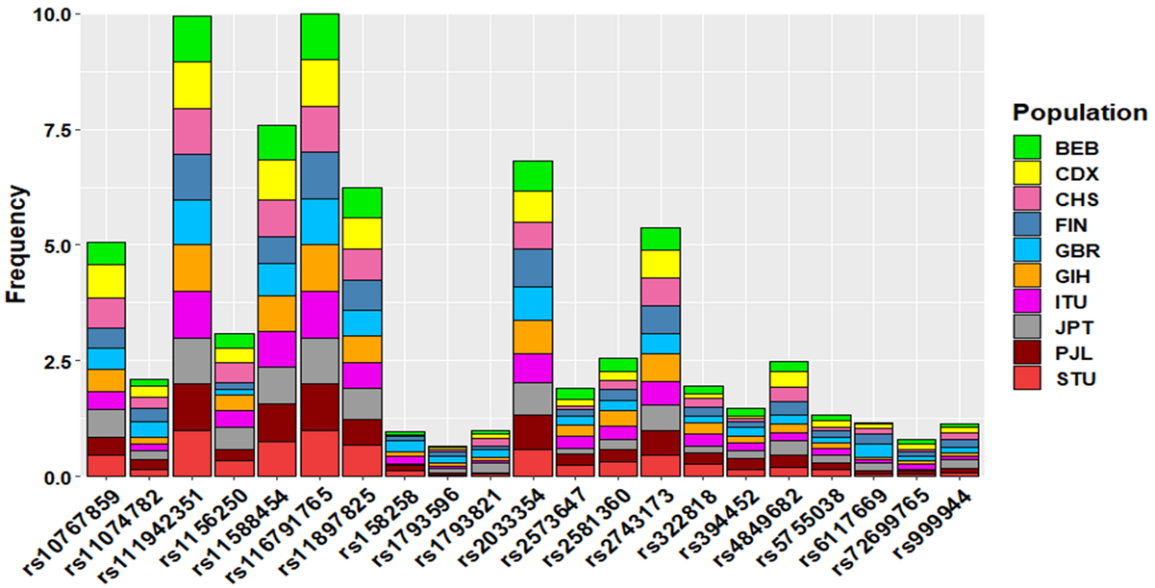
[5] Lam JC, Sharma SK and Lam B. Obstructive sleep apnoea: definitions, epidemiology & natural history. Indian J Med Res 2010; 131: 165-170.

[6] Bradley TD and Floras JS. Obstructive sleep apnoea and its cardiovascular consequences. Lancet 2009; 373: 82-93.
[7] Guilleminault C and Bassiri A. Clinical features and evaluation of obstructive sleep apnea-hypopnea syndrome and upper airway resistance syndrome. 2005. pp. 1043-1052.
[8] Slater G and Steier J. Excessive daytime sleepiness in sleep disorders. J Thorac Dis 2012; 4: 608-616.
[9] Davies RJ and Stradling JR. The relationship between neck circumference, radiographic pharyngeal anatomy, and the obstructive sleep apnoea syndrome. Eur Respir J 1990; 3: 509-514.
[10] Randerath W, Bassetti CL, Bonsignore MR, Farre R, Ferini-Strambi L, Grote L, Hedner J, Kohler M, Martinez-Garcia MA, Mihaicuta S, Montserrat J, Pepin JL, Pevernagie D, Pizza F, Polo O, Riha R, Ryan S, Verbraecken J and McNicholas WT. Challenges and perspectives in obstructive sleep apnoea: report by an ad hoc working group of the Sleep Disordered Breathing Group of the European Respiratory Society and the European Sleep Research Society. Eur Respir J 2018; 52: 1702616.
[11] Punjabi NM, Sorkin JD, Katzel LI, Goldberg AP, Schwartz AR and Smith PL. Sleep-disordered breathing and insulin resistance in middle-aged and overweight men. Am J Respir Crit Care Med 2002; 165: 677-682.
[12] Taheri S and Mignot E. The genetics of sleep disorders. Lancet Neurol 2002; 1: 242-250.
[13] White DP and Shafazand S. Mandibular advancement device vs. CPAP in the treatment of obstructive sleep apnea: are they equally effective in short term health outcomes? J Clin Sleep Med 2013; 9: 971-972.
[14] Lee JS, Choi HI, Lee H, Ahn SJ and Noh G. Biomechanical effect of mandibular advancement device with different protrusion positions for treatment of obstructive sleep apnoea on tooth and facial bone: a finite element study. J Oral Rehabil 2018; 45: 948-958.
[15] Ferguson KA, Ono T, Lowe AA, al-Majed S, Love LL and Fleetham JA. A short-term controlled trial of an adjustable oral appliance for the treatment of mild to moderate obstructive sleep apnoea. Thorax 1997; 52: 362-368.
[16] Epstein LJ, Kristo D, Strollo PJ Jr, Friedman N, Malhotra A, Patil SP, Ramar K, Rogers R, Schwab RJ, Weaver EM and Weinstein MD; Adult Obstructive Sleep Apnea Task Force of the American Academy of Sleep Medcine. Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. J Clin Sleep Med 2009; 5: 263-276.
[17] Hoekema A, Stegenga B and De Bont LG. Efficacy and co-morbidity of oral appliances in the treatment of obstructive sleep apnea-hypopnea: a systematic review. Crit Rev Oral Biol Med 2004; 15: 137-155.
[18] Lim J, Lasserson TJ, Fleetham J and Wright J. Oral appliances for obstructive sleep apnoea. Cochrane Database Syst Rev 2006; 2006: CD004435.
[19] Szily M, Tarnoki AD, Tarnoki DL, Kovacs DT, Forgo B, Lee J, Kim E, Sung J, Kunos L, Meszaros M, Muller V and Bikov A. Genetic influences on the onset of obstructive sleep apnoea and daytime sleepiness: a twin study. Respir Res 2019; 20: 125.
[20] Varvarigou V, Dahabreh IJ, Malhotra A and Kales SN. A review of genetic association studies of obstructive sleep apnea: field synopsis and meta-analysis. Sleep 2011; 34: 1461-1468.
[21] Mullington JM, Abbott SM, Carroll JE, Davis CJ, Dijk DJ, Dinges DF, Gehrman PR, Ginsburg GS, Gozal D, Haack M, Lim DC, Macrea M, Pack AI, Plante DT, Teske JA and Zee PC. Developing biomarker arrays predicting sleep and circadian-coupled risks to health. Sleep 2016; 39: 727-736.
[22] Franklin KA and Lindberg E. Obstructive sleep apnea is a common disorder in the population-a review on the epidemiology of sleep apnea. J Thorac Dis 2015; 7: 1311-1322.
[23] Peppard PE, Young T, Palta M, Dempsey J and Skatrud J. Longitudinal study of moderate weight change and sleep-disordered breathing. JAMA 2000; 284: 3015-3021.
[24] Cade BE, Chen H, Stilp AM, Gleason KJ, Sofer T, Ancoli-Israel S, Arens R, Bell GI, Below JE, Bjonnes AC, Chun S, Conomos MP, Evans DS, Johnson WC, Frazier-Wood AC, Lane JM, Larkin EK, Loredo JS, Post WS, Ramos AR, Rice K, Rotter JI, Shah NA, Stone KL, Taylor KD, Thornton TA, Tranah GJ, Wang C, Zee PC, Hanis CL, Sunyaev SR, Patel SR, Laurie CC, Zhu X, Saxena R, Lin X and Redline S. Genetic associations with obstructive sleep apnea traits in Hispanic/Latino Americans. Am J Respir Crit Care Med 2016; 194: 886-897.
[25] MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E, Junkins H, McMahon A, Milano A, Morales J, Pendlington ZM, Welter D, Burdett T, Hindorff L, Flicek P, Cunningham F and Parkinson H. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res 2017; 45: D896-D901.
[26] Consortium GT. The Genotype-Tissue Expression (GTEx) project. Nat Genet 2013; 45: 580-585.
[27] Clarke L, Fairley S, Zheng-Bradley X, Streeter I, Perry E, Lowy E, Tasse AM and Flicek P. The International Genome Sample Resource (IG-

SR): a worldwide collection of genome variation incorporating the 1000 genomes project data. Nucleic Acids Res 2017; 45: D854-D859.
[28] Excoffier L, Laval G and Schneider S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform Online 2007; 1: 47-50.
[29] Wickham H. Reshaping data with the reshape package. J Stat Softw 2007; 21: 20.
[30] Benjafield AV, Ayas NT, Eastwood PR, Heinzer R, Ip MSM, Morrell MJ, Nunez CM, Patel SR, Penzel T, Pepin JL, Peppard PE, Sinha S, Tufik S, Valentine K and Malhotra A. Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis. Lancet Respir Med 2019; 7: 687-698.
[31] Kraft P, Zeggini E and Ioannidis JP. Replication in genome-wide association studies. Stat Sci 2009; 24: 561-573.
[32] Sharma V, Sharma I, Sethi I, Mahajan A, Singh G, Angural A, Bhanwer AJS, Dhar MK, Singh V, Rai E and Sharma S. Replication of newly identified type 2 diabetes susceptible loci in Northwest Indian population. Diabetes Res Clin Pract 2017; 126: 160-163.
[33] Sharma S, Lalrohlui F, Sharma V, Sharma I, Sharma S, Parihar TJ, Zohmingthanga J, Singh V, Sharma S, Senthil Kumar N and Rai E. Candidate gene association study of UCP3 variant rs1800849 with T2D in Mizo population of Northeast India. Int J Diabetes Dev Ctries 2020; 40: 513-517.
[34] Tabassum R, Chauhan G, Dwivedi OP, Mahajan A, Jaiswal A, Kaur I, Bandesh K, Singh T, Mathai BJ, Pandey Y, Chidambaram M, Sharma A, Chavali S, Sengupta S, Ramakrishnan L, Venkatesh P, Aggarwal SK, Ghosh S, Prabhakaran D, Srinath RK, Saxena M, Banerjee M, Mathur S, Bhansali A, Shah VN, Madhu SV, Marwaha RK, Basu A, Scaria V, McCarthy MI; DIAGRAM; INDICO, Venkatesan R, Mohan V, Tandon N and Bharadwaj D. Genome-wide association study for type 2 diabetes in Indians identifies a new susceptibility locus at 2q21. Diabetes 2013; 62: 977-986.
[35] Lv D, Tan L, Wu Y, Cao C and Deng Z. Leptin and leptin receptor gene polymorphisms in obstructive sleep apnea: a HuGE review and meta-analysis. Sleep Breath 2015; 19: 1073-1078.
[36] Patel SR, Goodloe R, De G, Kowgier M, Weng J, Buxbaum SG, Cade B, Fulop T, Gharib SA, Gottlieb DJ, Hillman D, Larkin EK, Lauderdale DS, Li L, Mukherjee S, Palmer L, Zee P, Zhu X and Redline S. Association of genetic loci with sleep apnea in European Americans and African-Americans: the Candidate Gene Association Resource (CARe). PLoS One 2012; 7: e48836.
[37] Veatch OJ, Bauer CR, Keenan BT, Josyula NS, Mazzotti DR, Bagai K, Malow BA, Robishaw JD, Pack AI and Pendergrass SA. Characterization of genetic and phenotypic heterogeneity of obstructive sleep apnea using electronic health records. BMC Med Genomics 2020; 13: 105.
[38] Lalrohlui F, Sharma V, Sharma I, Singh H, Kour G, Sharma S, Yuman, Zohmingthanga J, Vanlalhruaii, Rai E, Singh V, Kumar NS and Sharma S. MACF1 gene variant rs2296172 is associated with T2D susceptibility in Mizo population from Northeast India. Int J Diabetes Dev Ctries 2020; 40: 223-226.
[39] McClellan J and King MC. Genetic heterogeneity in human disease. Cell 2010; 141: 210-217.
[40] Willing EM, Dreyer C and van Oosterhout C. Estimates of genetic differentiation measured by F(ST) do not necessarily require large sample sizes when using many SNP markers. PLoS One 2012; 7: e42649.



Supplementary Figure 1. Allele frequency for effective variants in super populations obtained from 1000 genome database. AFR: African, AMR: Admixed American, EAS: East Asians, EUR: European, SAS: South Asians.



Supplementary Figure 2. Allele frequency for effective variants obtained for 1000 genome database for different population groups. BEB: Bengali from Bangladesh, CDX: Chinese Dai in Xishuangbanna, CHS: Southern Han Chinese, FIN: Finnish, GBR: British in England and Scotland, GIH: Gujarati Indian from Houston, ITU: Indian Telugu from UK, JPT: Japanese in Tokyo, PJL: Punjabi from Lahore Pakistan, STU: Sri Lankan Tamil from UK.