Association Studies and Direct DNA Sequencing Implicate Genetic Susceptibility Loci in the Etiology of Nonsyndromic Orofacial Clefts in Sub-Saharan African Populations

L.J.J. Gowans, W.L. Adeyemo, M. Eshete, P.A. Mossey, T. Busch, B. Aregbesola, P. Donkor, F.K.N. Arthur, S.A. Bello, A. Martinez, M. Li, E.A. Augustine-Akpan, W. Deressa, P. Twumasi, J. Olutayo, M. Deribew, P. Agbenorku, A.A. Oti, R. Braimah, G. Plange-Rhule, M. Gesses, S. Obiri-Yeboah, G.O. Oseni, P.B. Olaitan, L. Abdur-Rahman, F. Abate, T. Hailu, P. Gravem, M.O. Ogunlewe, C.J. Buxó, M.L. Marazita, A.A. Adeyemo, J.C. Murray, A. Butali

Appendix Methods

Eligible subjects or participants

Eligible subjects were individuals with NSOFCs and their families, born to indigenous Ghanaian, Ethiopian and Nigerian parents. These families were recruited at the cleft clinics and during surgical missions. Births from Caucasians and Asians were excluded. Controls were recruited in Ghana, Nigeria and Ethiopia at the immunization clinics and dental clinics to match cases recruited from each of these countries. Controls were Africans born alive without any congenital birth defects in Ghana, Ethiopia and Nigeria. In Nigeria, two different centers (Lagos and Ife) coordinated patient recruitment. Only one center each coordinated patient recruitment in Ghana and Ethiopia. We have previously described individuals that are involved in recruitments for our cleft studies in Africa (Butali et al. 2011; Butali et al. 2015). In summary, recruitment is done by surgeons (i.e. plastic surgery, ear nose and throat surgeons, pediatric surgeons, maxillofacial surgeons and dental surgeons).

DNA Collection and processing

We collected saliva and cheek swab samples from participants using Oragene DNA Collection Kits (http://www.dnagenotek.com). We extracted DNA from both saliva and cheek swab samples using the Oragene Saliva processing protocol

(http://genetics.uiowa.edu/protocols.php). We then determined the concentration of DNA using Quibit Assay that employed Quibit 2.0 Fluorometer (http://www.invitrogen.com/site/us/en/home/brands/Product-Brand/Qubit.html). We finally performed XY-Genotyping on all samples to validate the sexes and sanctity of the samples (http://genetics.uiowa.edu/protocols.php).

SNP Genotyping

The detailed protocol is available at Murray Laboratory (http://genetics.uiowa.edu/protocols.php) but a summary is presented here. We selected these SNPs based on GWAS and candidate gene studies. We randomly assigned each sample to a well in a labeled 96-well microplate to form a Plate Map, using sample concentration of 2ng/ul. Each of these microplates also contained two template controls, NA18856 (male) and NA18855 (female). These two template controls are Yoruba HapMap samples. They therefore served as a guide in calling the genotype of individuals genotyped in this study. Each microplate also had provision for at least two No Template Controls (NTCs), which was dH_2O; however, NTCs were not added unto Microplates until the running of the chips. SNPs were designed based on human genome assembly GRCh37/hg19, 2009 (http://genome.ucsc.edu) and were obtained from ABi/Life Technologies (www.lifetechnologies.com).

DNA sequencing and DNA sequence analyses

The protocols for primer design and optimization as well as DNA amplification by PCR and electrophoresis have been described earlier (Butali et al. 2014). We shipped PCR products to Functional Biosciences, Madison, Wisconsin (http://order.functionalbio.com/seq/index) where

they were sequenced using an ABI 3730XL (http://www.appliedbiosystems.com/absite/us/en/home.html). Chromatograms were then transferred to a Unix workstation, base-called with PHRED (http://www.phrap.org/phredphrapconsed.html, v.0.961028), assembled with PHRAP (http://www.phrap.org/, v.0.960731), scanned by POLYPHRED (http://droog.gs.washington.edu/polyphred/, v. 0.970312) and viewed with CONSED programme (http://www.phrap.org/consed/consed.html, v. 4.0).

We ascertained the genomic location of each variant revealed by CONSED by employing the “Blat” function of UCSC Genome Browser (https://genome.ucsc.edu/). We predicted the functional effect of a coding variant on protein using Polyphen-2 (http://genetics.bwh.harvard.edu/pph2/), SIFT (http://sift.jcvi.org/) and Ensemble (http://www.ensembl.org/Homo_sapiens/Tools/VEP). Effect of a variant on mRNA splicing was ascertained using Human Splicing Finder 3.0 (http://www.umd.be/HSF3/). Finally, we predicted the effect of a mutation on a regulatory region using RegulomeDB (http://regulomedb.org/).

We ascertained the Minor Allele Frequencies (MAF) or novelty of a mutation by comparing it to variants in 1000 Genomes (http://browser.1000genomes.org/index.html), Exome Variant Server (http://evs.gs.washington.edu/EVS/), dbSNP (www.ncbi.nlm.nih.gov/SNP/), ExAC Browser (http://exac.broadinstitute.org/) and other literature on OFCs. We classified mutations as “novel” if they have never been reported in any of these databases or literature.