HHS Public Access

Novel GRHL3 variants in a South African cohort with cleft lip palate

Thirona Naicker[1,7], **Chinyere Adeleke**[2], **Azeez Alade**[2], **Peter A. Mossey**[3,8], **Waheed A. Awotoye**[2], **Tamara Busch**[2], **Mary Li**[2], **Joy Olotu**[4], **Lord J. J. Gowans**[5], **Colleen Aldous**[6], **Azeez Butali**[2,9]

[1]Genetics, Department of Paediatrics, University of KwaZulu-Natal, Durban, South Africa

[2]Department of Oral Pathology, Radiology and Medicine, College of Dentistry, University of Iowa, Iowa City, IA, USA

[3]Department of Orthodontics, University of Dundee, Dundee, UK

[4]Department of Anatomy, University of Port Harcourt, Nigeria

[5]Department of Biochemistry and Biotechnology, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

[6]School of Clinical Medicine, University of KwaZulu-Natal, Durban, South Africa

[7]Smile Train partner

[8]Member of Smile Train Global Medical Advisory Board, USA

[9]Member of Smile Train Research and Innovation Advisory Council, USA

Abstract

Objective: The etiology of cleft palate (CP) is poorly understood compared with that of cleft lip with or without palate (CL±P). Recently, variants in Grainyhead like transcription factor 3 (*GRHL3*) were reported to be associated with a risk for CP in European and some African populations including Nigeria, Ghana, and Ethiopia. In order to identify genetic variants that may further explain the etiology of CP, we sequenced *GRHL3* in a South African population to determine if rare variants in *GRHL3* are associated with the presence of syndromic or nonsyndromic CP.

Design: We sequenced the exons of *GRHL3* in 100 cases and where possible, we sequenced the parents of the individuals to determine the segregation pattern and presence of *de novo* variants.

Corresponding authors: 1. T. Naicker, Genetics, Department of Paediatrics, University of KwaZulu-Natal/Inkosi Albert Luthuli Central Hospital, 800 Vusi Mzimela Road, Umkumbaan, Durban, 4091, Kwa-Zulu Natal, South Africa. thironanaicker@gmail.com Tel: +27 31 240 2341. 2. A. Butali, Department of Oral Pathology, Radiology and Medicine/Dows Institute for Dental Research, College of Dentistry, University of Iowa, Butali Laboratory, ML2198, 500 Newton Road, Iowa City, IA 52242, USA. azeez-butali@uiowa.edu Fax: 319-384-1169. Tel: 319-335-8980.

Setting: The cleft clinics from two public, tertiary hospitals in Durban, South Africa (SA), namely Inkosi Albert Luthuli Central Hospital (IALCH) and KwaZulu-Natal Children's Hospital (KZNCH).

Patients, participants: 100 patients with CL±P and their parents.

Interventions: Saliva samples were collected.

Main outcome measures: To ascertain the genetic variants in the *GRHL3* gene in patients with CL±P in SA.

Results: Five variants in *GRHL3* were observed; three were novel and two were known variants. The novel variants were intronic variants (c.1062+77A>G and c.627+1G>A) and missense variant (p.Asp169Gly).

Conclusions: This study provides further evidence that variants in *GRHL3* contribute to the risk of nonsyndromic CP in African populations, specifically, in the South African population.

Introduction:

Orofacial clefts (OFC) are common craniofacial anomalies occurring with a worldwide prevalence of 1 in 700 live births (Mossey and Catilla, 2003). However, the prevalence of OFC differs according to ethnicity, race and geographical location. The highest reported prevalence of clefts occur in Asian populations, followed by Europeans, whereas African populations show the lowest prevalence (Mossey and Modell, 2012). The rates range from as high as 1 in 500 births in Asians to as low as 1 in 2000 births in Africans (Butali et al., 2014).

Although South Africa does not have a current national birth defect registry, studies conducted using a form called the Birth Defects Notification Tool shows that OFC represent one of the top five priority birth defects (Lebese, et al., 2016). The prevalence of OFCs in South Africa is estimated to be around 1 per 3333 live births (Hlongwa, et al., 2019).

There are biological and social effects with OFC and treatment aims should address the appearance, speech and hearing of an affected individual. This condition also places a significant psychological and economic burden on the individual, family and the healthcare service. However, good treatment outcomes are possible, especially in centres where a multidisciplinary OFC team approach is utilised.

The clinical classification of OFC includes cleft lip with or without cleft palate (CL±P) and cleft palate (CP). The etiology of OFC is complex and multifactorial with influences from genetics, genomics, the environment and possible gene-environment interactions (Dixon, et al., 2011).

The genetic etiology of CP alone is least understood and approximately 50% of all CP patients are nonsyndromic (Dixon et al., 2011). To date, very little research has been

done on the genetic etiology of OFCs in South Africa (SA). The current study is part of a comprehensive project to identify the genetic and genomic causes of OFC in the South African population through the African Craniofacial Anomalies Network (AfriCRAN) (Butali et al., 2015). In this study, we screened the Grainyhead like transcription factor 3 (*GRHL3*) gene for variants in a cohort from South Africa.

GRHL3 is a transcription factor that plays an important role in primary neurulation and in the differentiation of stratified epithelia of ectodermal or endodermal origin (Boglev et al., 2011). *GRHL3* has been previously associated with CP. Mutations were found in a Van Der Woude Syndrome (VWS) family with CP in a linkage and targeted sequencing study (Peyrard-Janvid et al., 2014). Leslie et al. (2016) performed genome wide association studies (GWAS) of nonsyndromic CP and found an association with a missense variant in *GRHL3*. A parallel study, (Mangold et al., 2016) found a significant association between the p.Thr454Met variant in *GRHL3* and nonsyndromic CP. In both studies, they concluded that the p.Thr454Met variant is an etiologic variant for nonsyndromic CP (Leslie et al., 2016)·(Mangold et al., 2016).

Recently, variants of *GRHL3* were found to be associated with risk for CP in African populations from Nigeria, Ghana and Ethiopia (Eshete et al., 2018). The current study aims to ascertain the genetic variants in the *GRHL3* gene in patients with CL±P in SA.

Participants & Methods

Study population

Patients born with various OFC types and the parents of these cases were recruited and assessed at the cleft clinics from two Durban hospitals in South Africa, namely Inkosi Albert Luthuli Central Hospital (IALCH) and KwaZulu-Natal Children's Hospital (KZNCH). The assessment was conducted by the principal investigator, a paediatrician (TN); and all findings in the 100 participants were verified by plastic surgeons and an orthodontist at the multi-disciplinary cleft clinic. This was an observational clinical study to ascertain the genetic causes of cleft lip and palate in the African population in KwaZulu-Natal. Patients were consecutively selected from the cleft clinics. To be included in the study, eligible children were of African ancestry and had a syndromic or non-syndromic CL±P. Patients belonging to other South African population groups were excluded to avoid confounding due to population structure.

Informed consent was obtained from all participating families. A case report form was completed, and saliva samples were collected with Oragene saliva kits and sponges (DNA Genotek). These samples were sent to the Butali Laboratory at the University of Iowa for processing and analysis. DNA from all samples was extracted, followed by XY genotyping analysis for quality control purposes to ensure that the sex of the sample matched the sex of the actual donor. A combination of 100 case-parent samples (dyads and triads) were collected and included for Sanger sequencing. Biomedical Research Ethics Committee clearance (BE309/18) was obtained from the University of KwaZulu-Natal.

Sanger Sequencing

Methods previously reported by our team were used to sequence DNA from humans (Gowans et al., 2016). We optimized 13 primers for the amplification of 16 exons in the *GRHL3* gene (NM_198174). We used 1ul of DNA in a 10-μL reaction for the polymerase chain reaction (PCR). Two Nigerian CEPH samples (NA18855 and NA18856) were added as template controls. The primers used and annealing temperatures are available from the Butali Laboratory upon request. The amplified DNA products were sequenced at Functional Biosciences (http://functionalbio.com).

In order to identify novel and rare variants, every variant found in our isolated population were compared to those in the 1000 Genomes database, Exome Variant Server database (http://snp.gs.washington.edu/EVS/), and exomes in the Genome Aggregation Database (gnomAD; http://gnomad.broadinstitute.org). To predict the functional effects of these variants on the protein, we used bioinformatics tools, such as Combined Annotation Dependent Depletion (CADD; https://cadd.gs.washington.edu/; (Rentzsch, Witten, Cooper, Shendure, & Kircher, 2019), Polymorphism Phenotyping (Polyphen; http://genetics.bwh.harvard.edu/pph2/; (Adzhubei et al., 2010), Sorting Intolerant from Tolerant (SIFT; http://sift.jcvi.org/; (Kumar, Henikoff, & Ng, 2009), Have Your Protein Explained (HOPE; http://www.cmbi.ru.nl/hope; (Venselaar, Te Beek, Kuipers, Hekkelman, & Vriend, 2010), and Human Splicing Finder (HSF; http://www.umd.be/HSF/) (Desmet et al., 2009). To assess the inheritance patterns of these variants, we sequenced parent samples, when available. Novel variants were deposited into the ClinVar public database (https://www.ncbi.nlm.nih.gov/clinvar/).

Results:

We observed five variants in *GRHL3*, three were novel (two intronic and one missense variant) and two were known variants. The novel variants were missense variant, p.Asp169Gly and intronic variants, c.1062+77A>G and c.627+1G>A (Table 1). These novel variants were absent from all public exome and genome databases including over 8000 individuals from Africa.

On analysis of the multiplex family pedigree for the novel missense variant, p.Asp169Gly, we note that 3 family members had cleft palates, namely our index patient (who had a soft palate cleft), her half sister and her paternal cousin (Figure 1). Segregation analyses and the chromatograph for this variant showed that the index patient and her father both carried the single nucleotide change, however the father did not have a CP, implying that the phenotype was not completely penetrant in him (Figure 2).

Phenotypic data for remaining 2 novel variants were as follows: the female patient with the intronic variant c.1062+77A>G had a bilateral cleft lip and palate (BCLP). There was an incomplete alveolar cleft on the right whereas the left side was complete and extended into the hard and soft palates. The last novel variant c.627+1G>T was an intronic, splice donor variant and was detected in a little girl with a cleft of the hard and soft palate. Both these patients had no family history of orofacial clefts. All patients with reported variants were non-syndromic.

The missense mutation (p.Arg298Cys) was found in the 1000 Genomes database and was previously reported twice; in a Non-Finnish European population and a Southeast Asian population. However, this is the first report of the same mutation in an African population with phenotypic data. Our patient had a cleft of the hard and soft palate.

The known missense variant (p.Arg16Gln) was identified in an individual with a cleft lip and alveolus. This variant was previously reported in East and West African populations. However, this is the first report in a South African population.

According to HOPE (http://www.cmbi.ru.nl/hope), the novel missense mutation (p.Asp169Gly) affected the structure and function of the protein for a number of biochemical reasons. There was a difference in charge, size and hydrophobicity between the wild-type and variant amino acid. The charge of the wild-type residue was lost. This could cause loss of interactions with other molecules or residues. The mutant residue was smaller (Figure 3) and this may lead to loss of interactions. Lastly, the mutation introduced a more hydrophobic residue in this position which could result in the loss of hydrogen bonds thereby disturbing the correct folding of the *GRHL3* protein.

Discussion:

In our variant screen study of *GRHL3* families with OFC from South Africa, we identified a novel missense variant (p.Asp169Gly) in a multiplex family with CP. The variant segregates in the family with the father carrying the variant. However, the father does not have a CP suggesting that there is incomplete penetrance. We strongly suspect that this family may have Van Der Woude Syndrome, where some cases are accounted for by the *GRHL3* gene (Peyrard-Janvid et al., 2014).

This novel missense variant (p.Asp169Gly) was not found in control databases that cumulatively has over 16,000 African alleles. PolyPhen-2 predicts the (p.Asp169Gly) variant is probably damaging while SIFT predicts that it is tolerated. The difference in effect prediction between the two in silico bioinformatics tools may be due to the design of the algorithms used for prediction of variant pathogenicity utilized by each tool. SIFT predicts whether an amino acid substitution is likely to affect protein function based on sequence homology and physico-chemical similarity between the alternate amino acids.

The CADD score for this novel missense variant (p.Asp169Gly) is 26 suggesting that this variant is in the top 1% of deleterious variants in the genome. The PROVEAN score of −0.36 predicts that the variant has a neutral effect. The variant was absent from the 1000 Genomes, Exome Variant Server and gnomAD databases.

The novel splice donor variant (c.627+1G>A) was found in intron 4 of a patient with CP. The variant has a high CADD score of 34, placing it in the top 0.1% of deleterious mutations in the genome. It caused alteration to the wild type donor site, which most probably affected splicing. The novel intronic variant (c.1062+77A>G) has a low CADD score of 0.104 and was found in a patient with BCLP with the right side having an incomplete cleft lip. This variant may not be associated with phenotype since variants in *GRHL3* have only been previously associated with CP.

The rare known missense variant (p.Arg298Cys) has been reported in the gnomAD database, however it is uncertain if these individuals had isolated CP. It has a CADD score of 26.7 and was predicted to be probably damaging and deleterious. The absence of phenotype in the control database makes it difficult to interpret this finding. Lastly, the known splicing variant (p.Arg16Gln) with a CADD score of 22.7 may contribute to CP as it will potentially affect splicing.

This study provides further evidence that mutations in *GRHL3* contribute to the risk of nonsyndromic CP in African populations, specifically, in the South African population. This is consistent with findings in a previous study from Africa (Eshete et al., 2018). More genetics and genomics studies need to be conducted in South Africa to determine additional genetic etiology of orofacial clefts.

Limitations:

Usually only 1 parent accompanies a child to the cleft clinic, and this influences the collection of triad samples, making it difficult to analyse segregation in most families. The current COVID-19 pandemic hampered saliva sample collection and hence study numbers were reduced. Nonetheless, our study and findings represent the need for genetic studies of these complex traits in South Africa.

Acknowledgements:

Funding:

Abbreviations:

BCLP	Bilateral cleft lip and palate
CADD	Combined Annotation Dependent Depletion score
CEPH	Human Polymorphism Study Center
CL±P	Cleft lip with or without cleft palate
CP	Cleft palate
DNA	Deoxyribonucleic Acid
EVS	Exome Server Variant
gnomAD	Genome Aggregation Database
GRHL3	Grainyhead-like transcription factor 3
GWAS	Genome wide association studies

HGVS	Human Genome Variation Sequence
HOPE	Have Your Protein Explained
HSF	Human Splicing Finder
IALCH	Inkosi Albert Luthuli Central Hospital
KZNCH	KwaZulu-Natal Children's Hospital
OFC	Orofacial clefts
PCR	Polymerase chain reaction
Polyphen	Polymorphism Phenotyping
SA	South Africa
SIFT	Sorting Intolerant from Tolerant
TN	Thirona Naicker, principal investigator
VWS	Van Der Woude Syndrome

References:

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR (2010). A method and server for predicting damaging missense mutations. Nat Methods, 7(4), 248–249. doi:10.1038/nmeth0410-248 [PubMed: 20354512]

Boglev Y, Wilanowski T, Caddy J, Parekh V, Auden A, Darido C, Hislop NR, Cangkrama M, Ting SB, Jane SM (2011). The unique and cooperative roles of the Grainy head-like transcription factors in epidermal development reflect unexpected target gene specificity. Dev Biol, 349(2), 512–522. doi:10.1016/j.ydbio.2010.11.011 [PubMed: 21081122]

Butali A, Adeyemo WL, Mossey PA, Olasoji HO, Onah II, Adebola A, Awoyale TA, et al. (2014). Prevalence of orofacial clefts in Nigeria. Cleft Palate Craniofac J, 51(3), 320–325. doi:10.1597/12-135 [PubMed: 23557093]

Butali A, Mossey P, Tiffin N, Adeyemo W, Eshete M, Mumena C, Ogunlewe M, Adeyamo A, et al. (2015). Multidisciplinary approach to genomics research in Africa: the AfriCRAN model. Pan African Medical Journal, 21(1). doi: 10.11604/pamj.2015.21.229.7380

Desmet FO, Hamroun D, Lalande M, Collod-Béroud G, Claustres M, & Béroud C (2009). Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res, 37(9), e67. doi:10.1093/nar/gkp215 [PubMed: 19339519]

Dixon MJ, Marazita ML, Beaty TH, & Murray JC (2011). Cleft lip and palate: understanding genetic and environmental influences. Nat Rev Genet, 12(3), 167–178. doi:10.1038/nrg2933 [PubMed: 21331089]

Eshete MA, Liu H, Li M, Adeyemo WL, Gowans LJJ, Mossey PA, Butali A, et al. (2018). Loss-of-Function GRHL3 Variants Detected in African Patients with Isolated Cleft Palate. J Dent Res, 97(1), 41–48. doi:10.1177/0022034517729819 [PubMed: 28886269]

Gowans LJJ, Adeyemo WL, Eshete M, Mossey PA, Busch T, Aregbesola B, Butali A, et al. (2016). Association Studies and Direct DNA Sequencing Implicate Genetic Susceptibility Loci in the Etiology of Nonsyndromic Orofacial Clefts in Sub-Saharan African Populations. Journal of Dental Research, 95(11), 1245–1256. doi:10.1177/0022034516657003 [PubMed: 27369588]

Hlongwa P, Levin J, & Rispel LC (2019). Epidemiology and clinical profile of individuals with cleft lip and palate utilising specialised academic treatment centres in South Africa. PLoS One, 14(5), e0215931. doi:10.1371/journal.pone.0215931 [PubMed: 31071123]

Kumar P, Henikoff S, & Ng PC (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc, 4(7), 1073–1081. doi:10.1038/nprot.2009.86 [PubMed: 19561590]

Lebese V, Aldous C, & Malherbe HL (2016). South African congenital disorders data, 2006 – 2014. S Afr Med J, 106(10), 992–995. doi:10.7196/SAMJ.2016.v106i10.11314 [PubMed: 27725019]

Leslie EJ, Liu H, Carlson JC, Shaffer JR, Feingold E, Wehby G, Marazita ML, et al. (2016). A Genome-wide Association Study of Nonsyndromic Cleft Palate Identifies an Etiologic Missense Variant in GRHL3. Am J Hum Genet, 98(4), 744–754. doi:10.1016/j.ajhg.2016.02.014 [PubMed: 27018472]

Mangold E, Böhmer AC, Ishorst N, Hoebel AK, Gültepe P, Schuenke H, Ludwig KU, et al. (2016). Sequencing the GRHL3 Coding Region Reveals Rare Truncating Mutations and a Common Susceptibility Variant for Nonsyndromic Cleft Palate. Am J Hum Genet, 98(4), 755–762. doi:10.1016/j.ajhg.2016.02.013 [PubMed: 27018475]

Mossey PA, & Catilla EE (2003). Global registry and database on craniofacial anomalies: Report of a WHO Registry Meeting on Craniofacial Anomalies. Available at: https://apps.who.int/iris/handle/10665/42840

Mossey PA, & Modell B (2012). Epidemiology of oral clefts 2012: an international perspective. Front Oral Biol, 16, 1–18. doi:10.1159/000337464 [PubMed: 22759666]

Peyrard-Janvid M, Leslie EJ, Kousa YA, Smith TL, Dunnwald M, Magnusson M, Schutte BC, et al. (2014). Dominant mutations in GRHL3 cause Van der Woude Syndrome and disrupt oral periderm development. Am J Hum Genet, 94(1), 23–32. doi:10.1016/j.ajhg.2013.11.009 [PubMed: 24360809]

Rentzsch P, Witten D, Cooper GM, Shendure J, & Kircher M (2019). CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res, 47(D1), D886–d894. doi:10.1093/nar/gky1016 [PubMed: 30371827]

Venselaar H, Te Beek TA, Kuipers RK, Hekkelman ML, & Vriend G (2010). Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC Bioinformatics, 11, 548. doi:10.1186/1471-2105-11-548 [PubMed: 21059217]
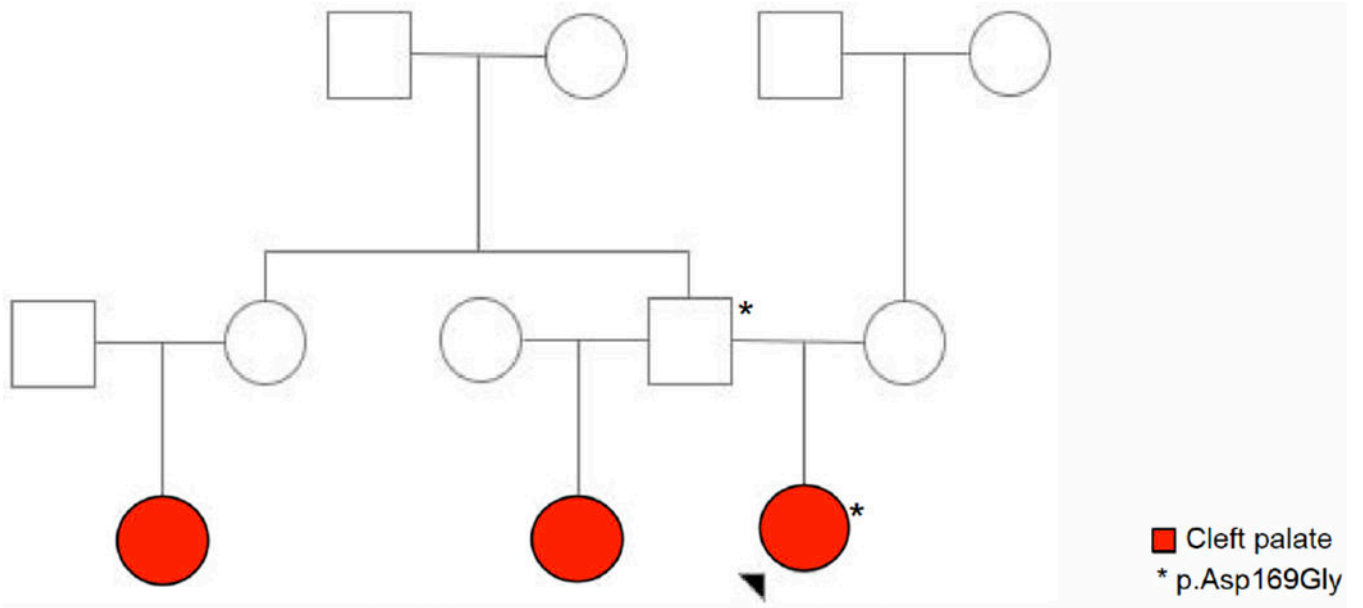
Figure 1.
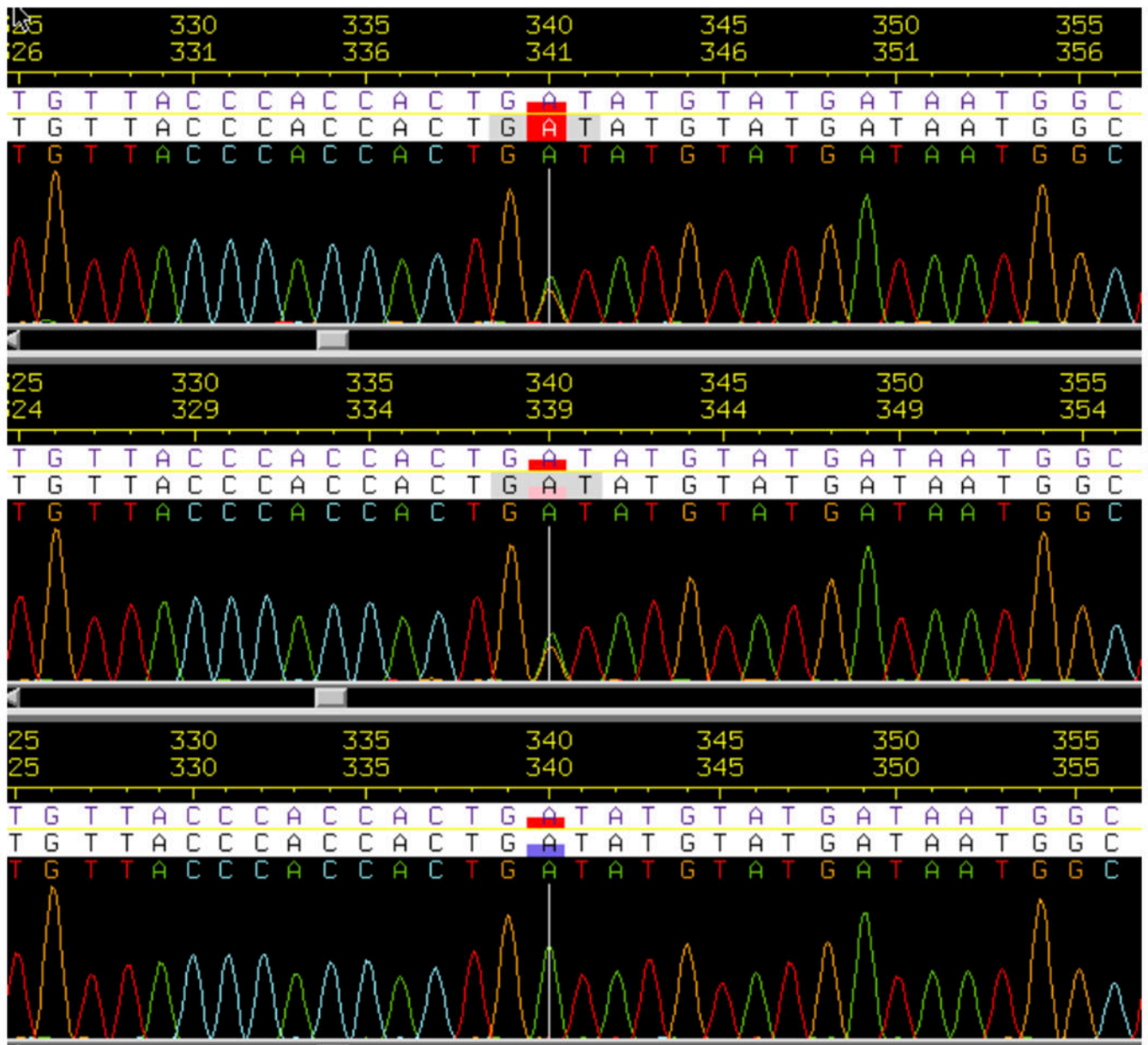Pedigree of family with novel p.Asp169Gly variant

Figure 2:
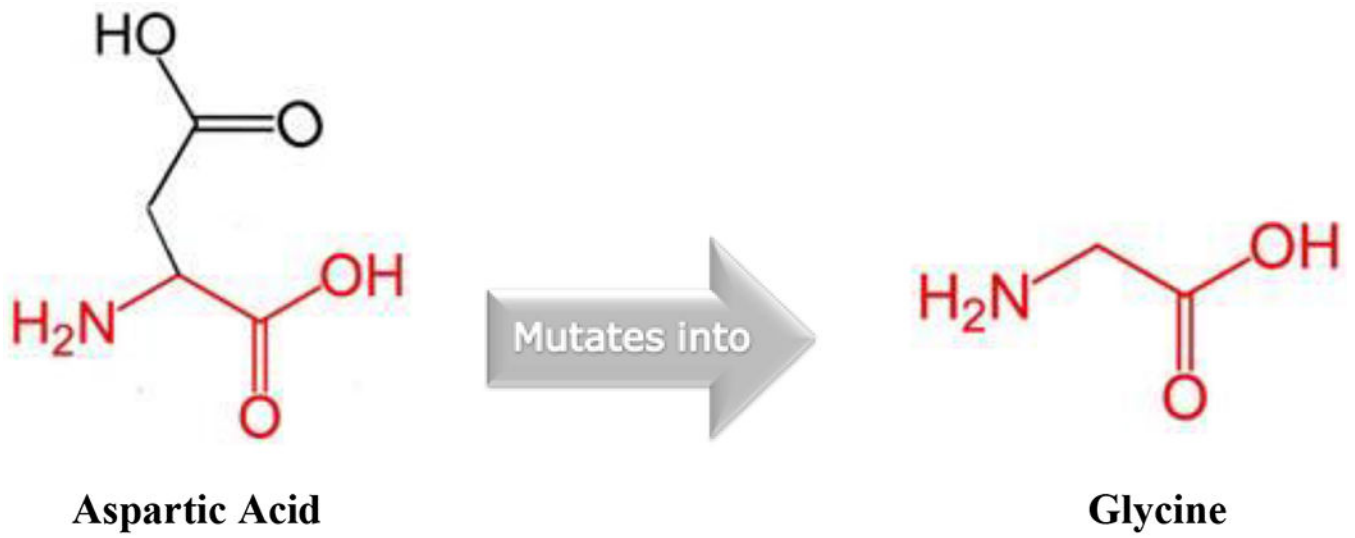Chromatograph of p.Asp169Gly variant in index patient, father and control showing A>G changes

Figure 3.
Mutation of Aspartic Acid into Glycine in p.Asp169Gly

Table 1:

Sequence Variations for GRHL3 Observed in Individuals from an Isolated South African Population

Nucleotide position	Exon/Intron	HGVS^c	HGVS^p	Type of Variant	Novel/ Known	1KG	EVS	gAD	PolyPhen	SIFT	Provean Score	CADD
chr1:24336721	Exon 4/16	c.506A>G	p.Asp169Gly	Missense	Novel	0	0	0	PD	T	−0.36	26.0
chr1:24339839	Intron 8/15	c.1062+77A>G		Intronic	Novel	0	0	0	n/a	n/a	n/a	0.104
chr1:24336828	Intron 4/15	c.627+1G>A		Splice_Donor_Variant	Novel	0	0	0	n/a	n/a	n/a	34
chr1:24338043	Exon 7/16	c.892C>T	p.Arg298Cys	Missense	Known	0	0	7.99 E-06	PD	D	−5.50	26.7
chr1:24331440	Exon 2/16	c.47G>A	p.Arg16Gln	Missense	Known	4.50 E-03	2.72 E-03	3.26 E-03	B	T	n/a	22.7

"c" refers to the coding sequence position within the GRHL3 transcript, NM_198173.3 "p" refers to amino acid substitutions. Amino acid substitution is for the GRHL3 isoform 2 transcript, NP_937816. Human assembly build h38.

1KG, 1000 Genomes; B, benign; CADD, Combined Annotation Dependent Depletion score; D, deleterious; EVS, Exome Variant Server; gAD, Genome Aggregation Database; HGVS, Human Genome Variation Sequence; HSF, Human Splicing Finder; PAS, potential alteration of splicing; PD, probably damaging; SIFT, Sorting Intolerant from Tolerant; T, tolerated.

BCLP, bilateral cleft lip palate; CP, cleft palate; CL±A, cleft lip alveolus.