Journal of Dental Research
2018, Vol. 97(1) 41–48

DOI: 10.1177/0022034517729819
journals.sagepub.com/home/jdr

Loss-of-Function GRHL3 Variants Detected in African Patients with Isolated Cleft Palate

M.A. Eshete[1,2,3*], H. Liu[4,5*], M. Li[6], W.L. Adeyemo[7], L.J.J Gowans[8], P.A. Mossey[9], T. Busch[6], W. Deressa[1], P. Donkor[8], P.B. Olaitan[10], B.S. Aregbesola[11], R.O. Braimah[12], G.O. Oseni[10], F. Oginni[11], R. Audu[13], C. Onwuamah[13], O. James[7], E. Augustine-Akpan[6], L.A. Rahman[14], M.O. Ogunlewe[7], F.K.N. Arthur[8], S.A. Bello[15], P. Agbenorku[8], P. Twumasi[8], F. Abate[2], T. Hailu[2], Y. Demissie[2,3], A. Hailu[2,3], G. Plange-Rhule[8], S. Obiri-Yeboah[8], M.M. Dunnwald[4], P.E. Gravem[16], M.L. Marazita[17], A.A. Adeyemo[18], J.C. Murray[19], R.A. Cornell[4], and A. Butali[6]

Abstract

In contrast to the progress that has been made toward understanding the genetic etiology of cleft lip with or without cleft palate, relatively little is known about the genetic etiology for cleft palate only (CPO). A common coding variant of grainyhead like transcription factor 3 (*GRHL3*) was recently shown to be associated with risk for CPO in Europeans. Mutations in this gene were also reported in families with Van der Woude syndrome. To identify rare mutations in *GRHL3* that might explain the missing heritability for CPO, we sequenced *GRHL3* in cases of CPO from Africa. We recruited participants from Ghana, Ethiopia, and Nigeria. This cohort included case-parent trios, cases and other family members, as well as controls. We sequenced exons of this gene in DNA from a total of 134 nonsyndromic cases. When possible, we sequenced them in parents to identify de novo mutations. Five novel mutations were identified: 2 missense (c.497C>A; p.Pro166His and c.1229A>G; p.Asp410Gly), 1 splice site (c.1282A>C p.Ser428Arg), 1 frameshift (c.470delC; p.Gly158Alafster55), and 1 nonsense (c.1677C>A; p.Tyr559Ter). These mutations were absent from 270 sequenced controls and from all public exome and whole genome databases, including the 1000 Genomes database (which includes data from Africa). However, 4 of the 5 mutations were present in unaffected mothers, indicating that their penetrance is incomplete. Interestingly, 1 mutation damaged a predicted sumoylation site, and another disrupted a predicted CK1 phosphorylation site. Overexpression assays in zebrafish and reporter assays in vitro indicated that 4 variants were functionally null or hypomorphic, while 1 was dominant negative. This study provides evidence that, as in Caucasian populations, mutations in *GRHL3* contribute to the risk of nonsyndromic CPO in the African population.

Keywords: targeted sequencing, GWAS, zebrafish, missense mutation, orofacial clefts, Van der Woude

Introduction

Orofacial clefts are common in humans and are the most common craniofacial birth defects, with a worldwide prevalence of 1 in 700 live births (Mossey and Little 2002). They are usually divided into syndromic and nonsyndromic clefts. Nonsyndromic clefts, which account for about 70% of all clefts, occur without additional visible clinical features. Syndromic clefts are associated with additional structural abnormalities. The cleft spectrum that includes cleft palate only (CPO), cleft lip only, and cleft lip with or without palate can be found in either syndromic or nonsyndromic clefts. About 50% of CPO cases are classified as syndromic clefts (Marazita et al. 2002). Regardless of the category or types of clefts, the rehabilitation of affected individuals requires a multidisciplinary team approach. For instance, in the United States, the lifetime cost of managing a child with cleft may total as much as $100,000 (Waitzman et al. 1994). Therefore, orofacial clefts place a huge personal and financial burden on families, society, and health care systems, particularly in developing countries, where social and physical infrastructures are limited.

The etiology of nonsyndromic cleft palate is complex, and factors including genetics, environmental exposures, and gene-environment interaction are plausible causes (Beaty et al. 2016). Unlike nonsyndromic cleft lip with or without palate where >26 risk loci and several candidate genes have been implicated through some genome-wide association studies (Beaty et al. 2010; Ludwig et al. 2012; Sun et al. 2015; Leslie, Carlson, et al. 2016; Leslie et al. 2017; Yu et al. 2017), little is known about the genetic causes of CPO. In 2011, Ghassibe-Sabbagh et al. reported a role for FAS-associated factor 1 (FAF1) in CPO when they demonstrated that haploinsufficiency of FAF1

causes CPO in a family with Pierre Robin sequence. Pierre Robin sequence is characterized by CPO, micrognathia, and glossoptosis. In another study, FitzPatrick et al. (2003) reported 2 de novo CPO-associated translocations involving 2q32-q33 where 1 breakpoint obstructs the transcription unit of special AT-rich sequence-binding protein 2 (SATB2). Also, mutations in T-box 22 (TBX22) have been reported in X-linked CPO cases (Braybrook et al. 2001; Braybrook et al. 2002; Marçano et al. 2004).

Two recent independent studies identified a common missense mutation, p.Thr454Met, in *GRHL3* that is associated with CPO. The first was a case-control/case-triad genome-wide association study (Leslie, Liu, et al. 2016), and the second was a combination of targeted sequencing and an association study (Mangold et al. 2016). Mutations in *GRHL3* have been reported in Van der Woude syndrome cases with CPO (Peyrard-Janvid et al. 2014). *GRHL3* encodes the transcription factor grainyhead-like 3, which is necessary for formation of the epidermal permeability barrier in mice (Ting et al. 2005). Both it and the transcription factor interferon regulatory factor 6 (IRF6) are necessary for differentiation of oral periderm, the most superficial layer of oral epithelium, which covers palate shelves during morphogenesis of the face (Ingraham et al. 2006; Peyrard-Janvid et al. 2014). Loss of oral periderm results in adhesions between palate shelves in mice (Richardson et al. 2009), leading to the proposal that defective differentiation of oral periderm is the cellular event that ultimately results in cleft palate. In zebrafish, 2 homologues of *GRHL3* acting redundantly (*Grhl1* and *Grhl3*) and the ortholog of *IRF6* are required for differentiation of periderm with the Grhl factors acting downstream of Irf6 (Sabel et al. 2009; de la Garza et al. 2013). Injection of RNA encoding human GRHL3 into zebrafish eggs induced ectopic expression of keratin 4 (*krt4*), a marker of the periderm, in deep blastomeres at the gastrula stage (de la Garza et al. 2013); this assay was used to show that Van der Woude syndrome–associated and nonsyndromic CPO–associated variants of GRHL3 lack normal function (Peyard-Janvid et al. 2014; Leslie, Liu, et al. 2016). In both cases, disease-associated variants disrupted differentiation of periderm in wild-type embryos, indicating that they had dominant-negative activity.

Importantly, the common variant p.Thr454Met in GRHL3 explains only a fraction of the heritability for CPO. To identify rare variants that could explain in part the remaining heritability for CPO, we conducted targeted sequencing of the *GRHL3* gene in individuals with CPO from Ghana, Ethiopia, and Nigeria. We tested novel variants found in patients in functional assays in cell lines and in zebrafish embryos.

Subjects and Methods

This study is part of a collaborative study that investigates the genetic and environmental causes of orofacial clefts in sub-Saharan Africans from Ghana, Ethiopia, and Nigeria. Local Institutional Review Boards approved sample and data collection: College of Health Sciences, Addis Ababa University, Addis Ababa, Ethiopia (3.10/027/2015); Kwame Nkrumah University of Science and Technology, Kumasi, Ghana (CHRPE/AP/217/13); and College of Medicine, University of Lagos, Nigeria (ADM/DCST/HREC/APP/1374). Patients born with CPO, controls, and parents of cases and controls were recruited and assessed at treatment centers in each country by the investigators and their collaborators. Informed consent was obtained from all participating families, and saliva samples were collected with Oragene saliva kits and sponges (DNA Genotek). These samples were sent to the Butali Laboratory at

[1]School of Public Health, Addis Ababa University, Addis Ababa, Ethiopia
[2]Yekatit 12 Hospital Medical College, Addis Ababa, Ethiopia
[3]Department of Surgery, School of Medicine, Addis Ababa University, Addis Ababa, Ethiopia
[4]Department of Anatomy and Cell Biology, University of Iowa, Iowa City, IA, USA
[5]State Key Laboratory Breeding Base of Basic Science of Stomatology (Hubei-MOST) and Key Laboratory for Oral Biomedicine of Ministry of Education, School and Hospital of Stomatology, Wuhan University, Wuhan, China
[6]Department of Oral Pathology, Radiology and Medicine, University of Iowa, Iowa City, IA, USA
[7]Department of Oral and Maxillofacial Surgery, University of Lagos, Lagos, Nigeria
[8]Kwame Nkrumah University of Science and Technology, Kumasi, Ghana
[9]Department of Orthodontics, University of Dundee, Dundee, UK
[10]Department of Plastic Surgery, Ladoke Akintola University of Science and Technology, Osogbo, Nigeria
[11]Department of Oral and Maxillofacial Surgery, Obafemi Awolowo University, Ile Ife, Nigeria
[12]Department of Oral and Maxillofacial Surgery, Usmanu Danfodiyo University, Sokoto, Nigeria
[13]Department of Virology, Nigerian Institute of Medical Research, Lagos, Nigeria
[14]Division of Pediatric Surgery, Department of Surgery, University of Ilorin, Ilorin, Nigeria
[15]State House Clinic, Abuja, Nigeria
[16]Plastic and Reconstructive Surgery Department, Haukeland University Hospital, Bergen, Norway
[17]Center for Craniofacial and Dental Genetics, Department of Oral Biology, University of Pittsburgh, Pittsburgh, PA, USA
[18]National Human Genomic Research Institute, Bethesda, MD, USA
[19]Department of Pediatrics University of Iowa, Iowa City, IA, USA
*Authors contributing equally to this article.

A supplemental appendix to this article is available online.

Corresponding Authors:
A. Butali, Department of Oral Pathology, Radiology and Medicine, College of Dentistry, University of Iowa, Iowa City, IA 52241, USA.
Email: Azeez-butali@uiowa.edu
R.A. Cornell, Department of Anatomy and Cell Biology, University of Iowa, Iowa City, IA 52242-1101, USA.
Email: robert-cornell@uiowa.edu

Table 1. Sequence Variations for GRHL3 Observed in Individuals with CPO from Sub-Saharan Africa.

HGVS[c]	HGVS[p]	Type	Ghana	Nigeria	1KG	EVS	ExAC	p	S	PS
c.332delC	p.Gly158Alafster55	Frameshift	0	1	0	0	0			
c.497C>A	p.Pro166His	Missense	1	0	0	0	0	PD	T	−0.64
c.1229A>G	p.Asp410Gly	Missense/splice site	0	1	0	0	0	PD	D	−6.7
c.1282A>C	p.Ser428Arg	Missense	0	1	0	0	0	B	T	−1.46
c.1677C>A	p.Tyr559Ter	Stop-gain	1	0	0	0	0			

"c" refers to the coding sequence position within the GRHL3 transcript NM_198173. "p" refers to amino acid substitutions. Amino acid substitution is for the GRHL3 isoform 2 transcript, NP_937816.
1Kg, 1000 Genomes; B, benign; D, deleterious; EVS, Exome Variant Server; ExAC, Exome Aggregate Consortium; HGVS, Human Genome Variation Society; P, Polyphen; PD, probably damaging; PS, Provean score; S, Sorting Intolerant from Tolerant; T, tolerated.

the University of Iowa for processing and analysis. We extracted DNA from all samples, followed by XY genotyping analysis to ensure that the sex of the sample matched the sex of the donor and for quality control purposes. A combination of 134 case-parent samples (dyads and triads), as well as 270 unrelated controls (90 from each of the 3 countries), were included for Sanger sequencing.

Sanger Sequencing

For sequencing DNA from human subjects, we used methods that we reported previously (Gowans et al. 2016). We used primers that were optimized for the amplification of 17 exons in the GRHL3 gene (NM_198174). We used 4 ng of DNA in a 10-µL reaction for the polymerase chain reaction (PCR). Two Yoruba HapMap samples and 2 water samples were added to the 96-well plates as template and nontemplate controls, respectively. The primers used and annealing temperatures are available from the Butali Laboratory upon request. The amplified DNA products were sequenced at Functional Biosciences (http://order.functionalbio.com/seq/index).

To recognize novel mutations, we compared mutations that we identified in patients with those in the 1000 Genomes database (http://www.1000genomes.org/), Exome Variant Server database (http://snp.gs.washington.edu/EVS/), and Exome Aggregation Consortium database (http://exac.broadinstitute.org/). These databases included exome data for >5,200 African and African American controls. We also sequenced 270 controls as an additional step to confirm novel mutations. To predict the functional effects of novel mutations on the protein, we used bioinformatics tools, such as Polymorphism Phenotyping (Polyphen; http://genetics.bwh.harvard.edu/pph2/; Adzhubei et al. 2010), Sorting Intolerant From Tolerant (SIFT; http://sift.jcvi.org/; Kumar et al. 2009), and Have Your Protein Explained (HOPE; http://www.cmbi.ru.nl/hope; Venselaar et al. 2010). To assess inheritance of novel mutations, we sequenced parent samples, when available.

Functional Test of Human GRHL3 Mutations in Zebrafish Embryos

Full-length, wild-type human *GRHL3* cDNA (GenBank: BC036890.2) was obtained from GE Healthcare. Novel *GRHL3* mutations were introduced into the wild type with PCR-mediated mutagenesis and the cDNAs shuttled into CS2+ vector (gift of David Turner, University of Michigan). The corresponding capped mRNAs were generated in vitro with the mMESSAGE mMACHINE SP6 kit (Ambion) and purified with the RNeasy mini kit (Qiagen). For the "rescue" experiment, 5 nL of *grhl1* AUG MO and/or *grhl3* E4I4 MO, each at 1 mg/mL, was injected into wild-type zebrafish embryos (NHGRI line) at the single-cell stage, after which approximately 1 ng of different human *GRHL3* mRNAs were injected into the same embryos. Embryos were fixed at 8 h postfertilization (hpf), and whole-mount in situ hybridization for *krt4* was performed as described previously (Sabel et al. 2009). For cryosection, after in situ hybridization, embryos were embedded in Tissue-Tek OCT Compound (Sakura Finetek) at −80 °C. Ten-micron sections were made with Microm Cryostat I HM505E (GMI). The Public Health Service Assurance approved the animal use protocols (protocol 6011616). Details of RNA extraction and luciferase assay are included in Appendix.

Results

Novel GRHL3 *Mutations Detected in Patients with CPO*

We sequenced *GRHL3* exons and splice sites in 134 cases with nonsyndromic CPO and in 270 controls. We identified 5 novel variants in patients that were absent from controls, the ExAc database of more than 100,000 whole exomes, and the 1000 Genomes database of whole genome sequences (accessed August 4, 2017). The mutations included 2 missense (p.Pro166His and p.Asp410Gly, referring to *GRHL3* isoform 2, NP_937816), 1 splice site (p.Ser428Arg), 1 deletion (p.Gly158Alafster55), and 1 nonsense (p.Tyr559Ter; Table 1, Fig. 1A). All of the identified mutations except for p.Ser428Arg were also detected in an unaffected parent. For the p.Ser428Arg mutation, we have samples only for the mother who did not have the mutation (Table 2). We also identified known missense mutations in cases (Appendix Table). The 5 novel variants were deposited into the Leiden Open Variation Database (submission 00100262).

CPO Patient–Derived Variants of GRHL3 Have Lower-Than-Normal Ability to Induce Ectopic Expression of a Periderm Marker in Zebrafish Embryos

We first tested the activity of CPO patient–derived *GRHL3* variants in an overexpression paradigm. We engineered each variant into full-length *GRHL3* cDNA, synthesized mRNAs in

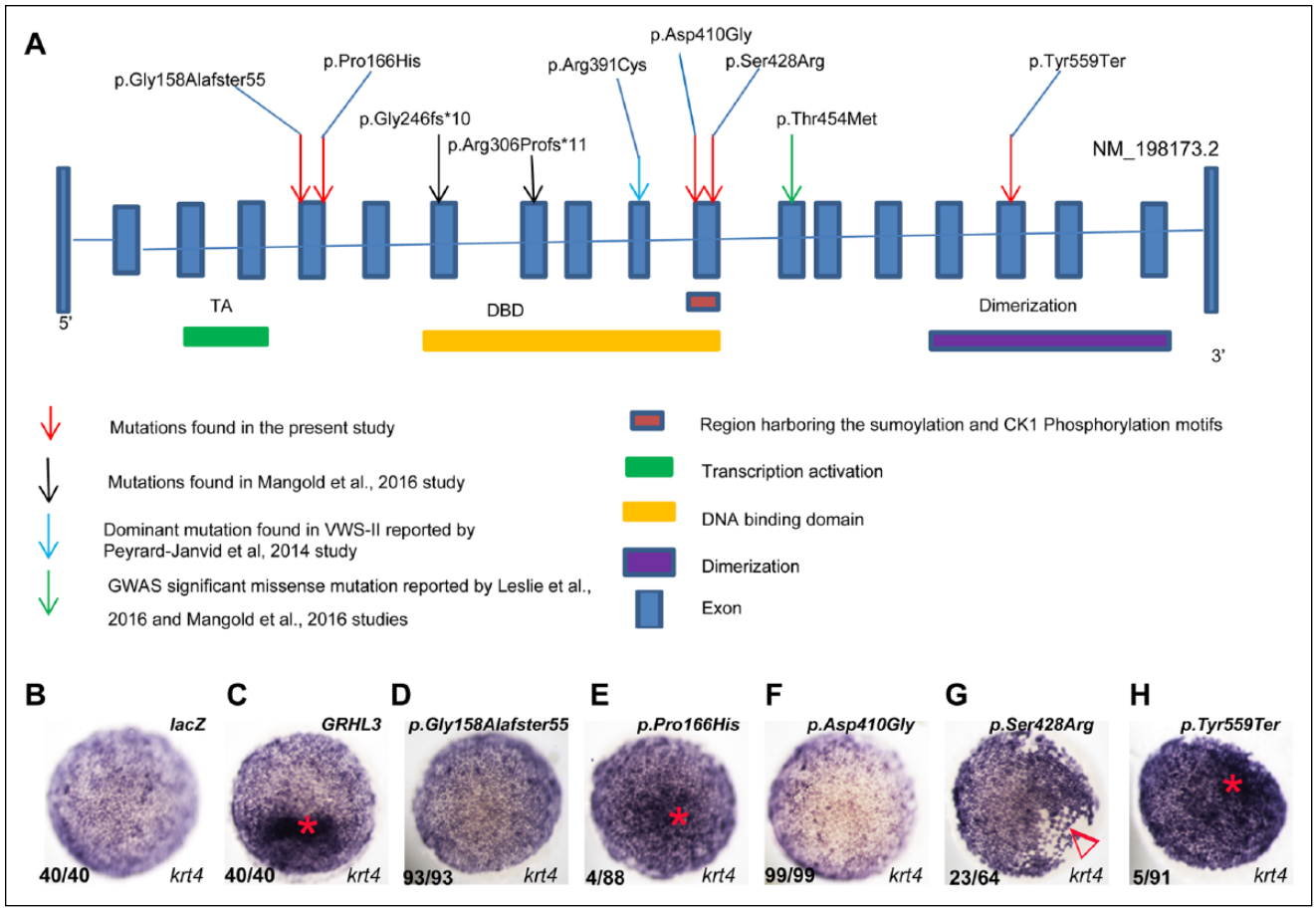

Figure 1. Rare variants in GRHL3 disrupt development of zebrafish enveloping layer upon overexpression in zebrafish embryos. (**A**) Schematic of the GRHL3 gene and the position of mutations identified in the present (red arrows) and previous (black, blue, and green arrows) studies (not to scale). All amino acid positions refer to GRHL3, variant 2 (NP_937816). The green arrow represents the missense mutation p.Thr454Met, which is a common variant associated with nonsyndromic CPO reported by Leslie, Liu, et al. (2016) and Mangold et al. 2016. The yellow bar represents the DNA binding domain where the sumoylation and CK1 phosphorylation motifs are harbored in exon 4. The purple bar represents the dimerization region, while the green bar represents the transcription activation domain. Blue boxes represent exons. (**B–H**) Animal pole views of embryos fixed at 7 h postfertilization and processed to reveal *krt4* expression. (B) Embryos injected with *lacZ* mRNA, a negative control, have relatively even expression of *krt4* in all enveloping layer cells. In 3 biological replicates of at least 40 injected embryos each, all had this appearance. None exhibited expression of *krt4* in deep blastomeres. (C) Embryos injected with mRNA encoding a reference variant of human GRHL3 (AAH36890.1) develop with intense foci of *krt4* expression in deep blastomeres, recognizable in whole-mount preparations as patches of dark staining (e.g., red asterisk). In 3 separate experiments of at least 40 injected embryos each, all exhibited patches of ectopic *krt4* expression. Embryos injected with mRNA encoding the patient variants (D) p.Gly158Alafster55 and (F) p.Asp410Gly lacked ectopic patches of *krt4* expression, consistent with them being strongly hypomorphic or null variants. Although *krt4* expression intensity appears lower than normal in this image, this was not consistently observed. In 3 biological replicates, with a total of 99 total embryos injected, none exhibited patches of ectopic *krt4* expression, and ≤10 noncontiguous deep cells expressing *krt4* in deep cells were observed in fewer than half of injected embryos. Rare embryos injected with (E) p.Pro166His and (H) p.Tyr559Ter variants exhibited patches of ectopic *krt4* expression (stars), indicating that these variants retained activity but at levels lower than those of wild type. In 3 experiments, 4 of 88 and 5 of 91 injected embryos, respectively, exhibited patches of ectopic *krt4* expression, as shown. (G) About one-third of embryos injected with p.Ser428Arg variant exhibited gaps in the expression of *krt4* (in 3 experiments, 23 of 64 injected embryos showed such gaps; red arrowhead). Fractions: numerator is the number of embryos resembling the one in the image; denominator is the number of embryos injected. The residue in each case resembles uninjected embryos. GWAS, genome-wide association study; VWS, Van der Woude syndrome.

vitro and injected them (or lacZ mRNA as a negative control) individually into single-cell zebrafish embryos, fixed embryos at 6 hpf (shield stage), and processed them to reveal *krt4* expression by in situ hybridization (Peyrard-Janvid et al. 2014; Leslie, Liu, et al. 2016). All embryos injected with *lacZ* mRNA exhibited *krt4* expression solely in the enveloping layer, which is the normal pattern of expression (40 of 40 embryos injected; Rauch et al. 2003; Fig. 1B; Appendix Fig.). In contrast, all embryos injected with wild-type human *GRHL3* mRNA exhibited large patches of ectopic expression of *krt4* in deep blastomeres readily visible in a whole-mount preparation (Fig. 1C) and confirmed by sectioned embryos (Appendix Fig.; 40 of 40 injected embryos). In embryos injected with the p.Gly158 Alafster55 (Fig. 1D), p.Asp410Gly (Fig. 1F), and p.Ser428Arg (Fig. 1G) variants, there were no ectopic patches of *krt4* expression (>60 embryos for each construct). In contrast, a small fraction (<10%) of embryos injected with the p.Pro166His (Fig. 1E) and p.Tyr559Ter (Fig. 1H) variants exhibited small

Table 2. Segregation Analyses of Novel Mutations in Case Families.

HGVS	HGVp	Individual	Genotypes
c.497C>A	p.Pro166His	Case child	CA
c.497C>A	p.Pro166His	Mother	CA
c.497C>A	p.Pro166His	Grandmother	CA
c.1677C>A	p.Tyr559Ter	Case child	CA
c.1677C>A	p.Tyr559Ter	Father	CA
c.1677C>A	p.Tyr559Ter	Mother	CC
c.1229A>G	p.Asp410Gly	Case child	AG
c.1229A>G	p.Asp410Gly	Mother	AG
c.1282A>C	p.Ser428Arg	Case child	AC
c.1282A>C	p.Ser428Arg	Mother	AA
c.470delC	p.Gly158Alafster55	Case child	C–
c.470delC	p.Gly158Alafster55	Mother	C–

patches of ectopic *krt4* expression (4 of 88 and 5 of 91 embryos injected, respectively), revealing that in this assay these variants are hypomorphic (i.e., less active than wild type). Moreover, in about half of embryos injected with the p.Ser428Arg variant, there were clear gaps in the expression of *krt4* (23 of 64 embryos injected and processed for *krt4* expression; Fig. 1G); 1 of the embryos injected with this construct ruptured just prior to fixation (i.e., at 8 hpf). Therefore, we repeated the experiment with this variant and permitted embryos to continue to develop. By 9 hpf, 36 of 83 embryos injected had ruptured, similar to embryos injected with a dominant-negative variant of *Grhl3* (de la Garza et al. 2013). In summary, the results indicate that in the overexpression assay, the p.Pro166His and p.Tyr559Ter variants are hypomorphic, the p.Gly158Alafster55 and p.Asp410Gly variants are strongly hypomorphic or null, and the p.Ser428Arg variant is dominant negative.

Next, in a rescue paradigm, we asked whether *krt4* expression could be restored by injection of mRNA encoding GRHL3 variants. We previously showed that simultaneous inhibition of *grhl1* and *grhl3* with morpholinos disrupts epiboly and *krt4* expression but that knockdown of either gene singly does not grossly affect periderm development (de la Garza et al. 2013; Fig. 2A–D). For rescue assays, we injected human *GRHL3* variants into embryos depleted of both *grhl1* and *grhl3*. Injection of wild-type *GRHL3* into embryos injected with *grhl1* MO and *grhl3* MO partially restored superficial *krt4* expression and induced ectopic expression of *krt4* in deep blastomeres (Fig. 2E). For quantitative comparisons of rescue efficiency, we evaluated *krt4* expression in whole embryo lysates by reverse transcription quantitative PCR (Fig. 2F). Embryos injected with *grhl1* MO and *grhl3* MO (i.e., double morphants) and with *LACZ* mRNA had significantly lower *krt4* expression than those injected control morpholino, whereas double morphants injected with RNA coding the reference (wild type) variant, the p.Pro166His variant, or the p.Tyr559Ter variant of GRHL3 had significantly higher levels of *krt4* expression than those injected with *LACZ* mRNA. *krt4* expression levels in double morphants injected with the other variants (p.Gly158Alafster55, p.Asp410Gly, and p.Ser428Arg) were not significantly higher than those in embryos injected with LACZ and were significantly lower than in those injected with wild-type GRHL3 ($P < 0.01$ by 1-way analysis of variance; Fig. 2F).

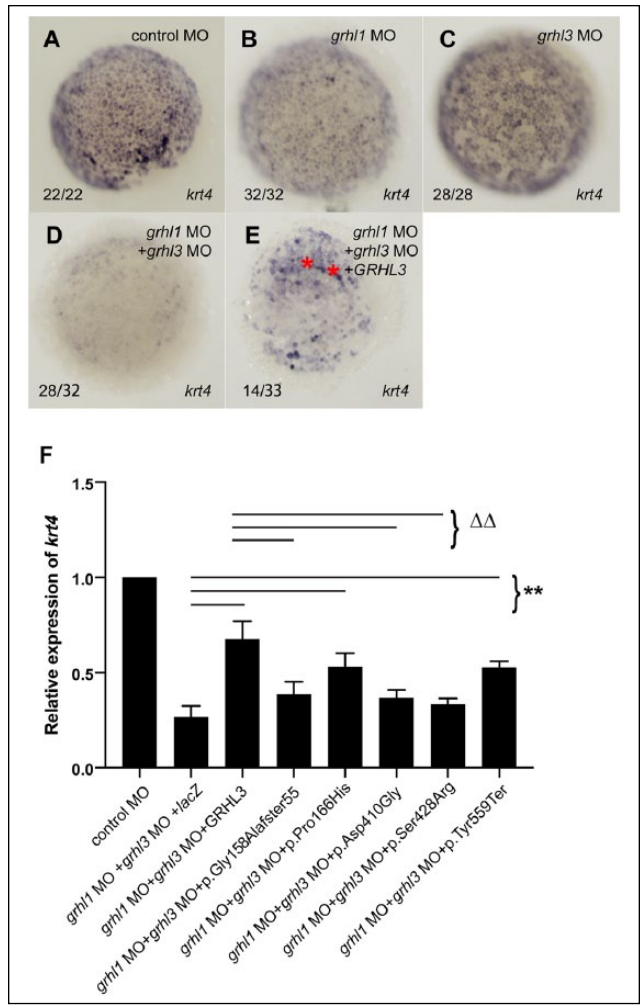

Figure 2. Injection of human *GRHL3* RNA partially restored *krt4* in the embryos injected with *grhl1* and *grhl3* MOs. (**A–E**) Animal pole views of embryos fixed with the injected reagents at the 1-cell stage, fixed at 6 h postfertilization, and processed to reveal *krt4* expression by in situ hybridization. In embryos injected with (A) control MO, (B) *grhl1* MO, or (C) *grhl3* MO, *krt4* expression is contiguous, whereas in embryos injected with (D) *grhl1* MO and *grhl3* MO, *krt4* expression is highly reduced. (E) In embryos injected with *grhl1* MO and *grhl3* MO and also with human *GRHL3* mRNA, there was mosaic rescue of *krt4* expression in the enveloping layer, in addition to ectopic *krt4* expression in the deep layer cells (red asterisks). Fractions: numerator is the number of embryos resembling the injected one (in panel E, includes ectopic *krt4*-expressing cells); denominator is the total number of embryos injected. In panel D, the residual embryos resembled control MO-injected ones. In panel E, the residual embryos resembled those injected with *grhl1* MO and *grhl3* MO pictured in panel D. (**F**) Reverse transcription quantitative real-time polymerase chain reaction analysis of *krt4* mRNA levels in embryos injected with the indicated MO and mRNA. $^{**}P < 0.01$ vs. "*grhl1* MO + *grhl3* MO + lacZ" group (1-way analysis of variance). $^{\Delta\Delta}P < 0.01$, vs. "*grhl1 MO+ grhl3 MO+GRHL3*" group. Error bars represent standard error.

CPO Patient–Derived Variants Have Lower-Than-Normal Ability to Activate a GRHL3-Sensitive Reporter

We next assessed the ability of the GRHL3 variants to activate a synthetic GRHL3-sensitive reporter, which was composed of

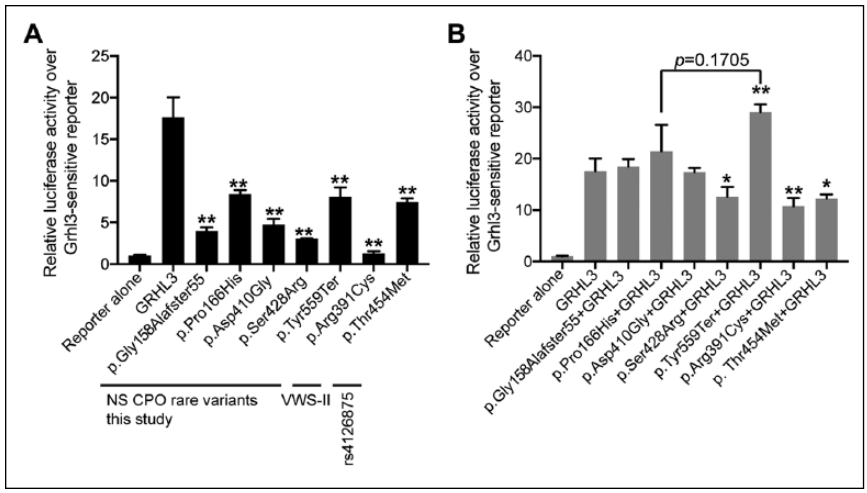

Figure 3. In vitro reporter analysis of transactivation activity of rare variants in GRHL3. (**A**) Bar chart representing firefly luciferase activity, normalized to Renilla luciferase activity, in 293FT cells cotransfected with the indicated (0.5 µg) GRHL3 expression constructs, or pCS2+ plasmid, with 0.5 µg of a GRHL3-sensitive firefly reporter plasmid and 0.05 µg of a constitutive Renilla reporter plasmid as a transfection control. (**B**) Bar chart representing firefly luciferase activity, normalized to Renilla luciferase activity, in 293FT cells cotransfected with the indicated (0.5 µg) GRHL3 expression constructs, or CS2+, with 0.5 µg of wild-type GRHL3 expression constructs, the GRHL3-sensitive firefly reporter plasmid, and the Renilla reporter plasmids. p.Arg345Cys is a rare GRHL3 variant found in Van der Woude syndrome (VWS; referred to as p.Arg391Cys in Peyrard-Janvid et al. 2014), and p.Thr408Met is the CPO-associated *GRHL3* variants reported previously (referred to as pThr454Met in Leslie, Liu, et al. 2016; Mangold et al. 2016). These 2 variants serve as positive controls for dominant negative activity for the current study. The data represent results from 3 separate experiments, and error bars represent standard error. *$P < 0.05$. **$P < 0.01$. CPO, cleft palate only; NS, nonsyndromic.

6 GRHL3 binding sites upstream of a minimal promoter and the firefly luciferase gene (Leslie, Liu, et al. 2016). We cotransfected 293FT cells with the GRHL3-sensitive reporter, expression vectors harboring different GRHL3 variants (separately), and a plasmid-driving constitutive Renilla luciferase expression as a transfection control (internal control). For controls, we additionally tested 2 other patient-derived GRHL3 variants, both previously shown to have dominant-negative activity in zebrafish: p.Arg391Cys, from an individual with Van der Woude syndrome (Peyrard-Janvid et al. 2014), and p.Thr454Met, a common variant associated with risk for nonsyndromic CPO (Leslie, Liu, et al. 2016; Mangold et al. 2016). Compared with cells transfected with the GRHL3-sensitive reporter alone, those additionally transfected with the expression vector containing wild-type GRHL3 had >6-fold-greater luciferase levels (Fig. 3). By contrast, expression vectors containing the p.Ser428Arg variant did not activate the reporter above background, and those containing p.Pro166His, p.Tyr559Ter, and p.Thr454Met variants did so but to a significantly lesser level than that of wild type (Fig. 3A). Of note, p.Pro166His and p.Tyr559Ter were the variants that also retained some ability to induce ectopic *krt4* expression in zebrafish embryos (Fig. 1E, H).

p.Ser428Arg GRHL3 Variant Shows Dominant Negative Effects in 293FT Cells

Finally, to test the patient-derived GRHL3 variants for dominant-negative activity, we cotransfected equal amounts of expression vectors containing wild-type and a patient-derived variant. p.Ser428Arg, p.Arg345Cys, and p.Thr408Met, which all had dominant-negative activity in the zebrafish assay, all decreased the reporter levels as compared with cells transfected with wild-type GRHL3 alone (Fig. 3B). The remaining variants all had additive effects when combined with wild-type GRHL3, indicating that they are loss-of-function but not dominant-negative variants. In summary, variants of GRHL3 detected in patients with CPO all had reduced function in comparison with the reference variant, and a subset had dominant-negative activity, consistent with each having pathogenic effects.

Discussion

Given evidence on the association of a variant of *GRHL3* with CPO in Europeans, we tested the hypothesis that rare-function blocking variants of *GRHL3* contribute to pathogenesis of nonsyndromic CPO in sub-Saharan Africa by sequencing the *GRHL3* exons in such patients. We detected 5 variants that were of interest because of their absence from the public domain catalogs of coding variants and because 1 or more in silico algorithms predicted that they were "possibly damaging" or worse. We subjected all of them to functional assays in zebrafish, including an overexpression and a rescue of GRHL3-loss-of-function paradigm, and to in vitro tests of their transactivation activity, all in comparison with the reference variant of GRHL3. These assays consistently indicated that 2 variants, p.Pro166H and p.Tyr559Ter, were modestly hypomorphic; 2 others, p.Gly158Alafster55 and p.Asp410Gly, were strongly hypomorphic or null; and 1, p.Ser428Arg, was dominant negative. These findings raise several issues.

First, the presence of some of these mutations in unaffected family members suggests that their effects are incompletely penetrant, as reported in previous genetic studies of clefting (Mangold et al. 2016). Modifier alleles in the parents may be protective for clefting but simultaneously elevate their risk for other chronic diseases, such as cancer, which has elevated prevalence in families with clefts (Christensen et al. 2004; Bille et al. 2005). Alternatively, affected children and unaffected parents may share the same genetic risk factors but be differentially exposed to environmental factors, such as maternal smoking during the periconceptional period, leading to clefts in the presence of risk alleles (Little et al. 2004).

Second, it is noteworthy that our functional assays had only modest concordance with in silico algorithms that predict the functional consequence of missense variants (e.g., Polyphen, SIFT, and HOPE). The most extreme disagreement was for the p.Ser428Arg variant, which had dominant-negative activity in

our assays but was predicted to be benign by Polyphen and SIFT. The p.Pro166His mutation, which had hypomorphic activity in our in vivo and in vitro assays, was predicted to be "probably damaging" by Polyphen and "tolerated" by SIFT with a Provean score of −0.64. The p.Tyr559Ter variant, which had reduced transactivation activity in our in vitro assay, was predicted to be a gain-of-function mutation by HOPE, leading to increased activity of GRHL3. The p.Asp410Gly variant, which in our assays was loss of function, Polyphen predicted to be "probably damaging", and SIFT predicted to be deleterious. The strength of the zebrafish-based assays is that they tested the variant's function in the context of a gene regulatory network governing differentiation of a relevant tissue. The in vitro assays have the advantage of being more quantitative than the in vivo ones. The results from the in vivo assay is consistent with the 2 in vitro assays, which agreed with each other, all supporting their validity. Functional assays remain an important complement to in silico algorithms for discerning pathologic variants from innocuous ones.

Third, the specific biochemical deficits in each functional variant are unknown and are an interesting subject for future study. One or more of the variants may diminish the protein's stability, which is testable with appropriate antibodies or epitope-tagged variants. HOPE predicted that p.Ser428Arg disrupts a potential phosphorylation site of the kinase CK1, although it is unknown if GRHL3 is phosphorylated at this residue. The same algorithm predicted that the substitution of proline by histidine in the p.Pro166His variant will lead to a loss of hydrophobic interaction in the nucleus or surface of the protein, and prolines provide rigidity to local protein conformation. Finally, it predicted that p.Asp410Gly disrupts a potential sumoylation site; sumoylation of lysine residues generally alters protein stability, function, or subcellular localization (Hendriks and Vertegaal 2015). It remains to be determined if this residue of GRHL3 is subjected to sumoylation.

Conclusion

We identified novel GRHL3 variants in CPO patients and tested them in functional assays that had the ability to recognize wild-type, reduction-in-function, and dominant-negative activities. The results of the functional tests suggest that lower-than-normal levels of GRHL3 and not necessarily dominant-negative activity can confer increased susceptibility to CPO. The p.Thr454Met variant, first identified in a European genome-wide association study of common variants, was not found in our samples, but other rare variants of GRHL3 influence heritability for CPO in Africans. These findings expand understanding of the genetics and biology underpinning CPO.

Author Contributions

M.A. Eshete, contributed to conception, design, data acquisition, and analysis, drafted and critically revised the manuscript; H. Liu, contributed to conception, design, and data analysis, drafted and critically revised the manuscript; M. Li, contributed to data analysis, critically revised the manuscript; W.L. Adeyemo, L.J.J. Gowans, P.A. Mossey, M.L. Marazita, A.A. Adeyemo, J.C. Murray, contributed to conception, design, data acquisition, and analysis, critically revised the manuscript; T. Busch, contributed to conception, design, and data analysis, critically revised the manuscript; W. Deressa, P. Donkor, P.B. Olaitan, B.S. Aregbesola, R.O. Braimah, G.O. Oseni, F. Oginni, R. Audu, C. Onwuamah, O. James, E. Augustine-Akpan, L.A. Rahman, M.O. Ogunlewe, F.K.N. Arthur, S.A. Bello, P. Agbenorku, P. Twumasi, F. Abate, T. Hailu, Y. Demissie, A. Hailu, G. Plange-Rhule, S. Obiri-Yeboah, M.M. Dunnwald, P.E. Gravem, contributed to data acquisition and analysis, critically revised the manuscript; R.A. Cornell, contributed to conception, design, data analysis, and interpretation, drafted and critically revised the manuscript; A. Butali, contributed to conception, design, data acquisition, analysis, and interpretation, drafted and critically revised the manuscript. All authors gave final approval and agree to be accountable for all aspects of the work.

Acknowledgments

We appreciate the support of all members of the families who participated in this research. This work was supported by the National Institute of Dental and Craniofacial Research K99/R00 (grant DE022378), the Robert Wood Johnson Foundation (grant 72429; A.B.), the National Institutes of Health (grant R01 DE023575, R.A.C.; R37 grants DE-08559 and DE-016148, J.C.M.; AR067739 R01, M.M.D.), and the Ghana Cleft Foundation (G001, L.J.J.G.). The authors declare no potential conflicts of interest with respect to the authorship and/or publication of this article.

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 variants. Nat Methods. 7(4):248–249.

Beaty TH, Marazita ML, Leslie EJ. 2016. Genetic factors influencing risk to orofacial clefts: today's challenges and tomorrow's opportunities. F1000Res. 5:2800.

Beaty TH, Murray JC, Marazita ML, Munger RG, Ruczinski I, Hetmanski JB, Liang KY, Wu T, Murray T, Fallin MD, et al. 2010. A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. Nat Genet. 42(6):525–529.

Bille C, Winther JF, Bautz A, Murray JC, Olsen J, Christensen K. 2005. Cancer risk in persons with oral cleft—a population-based study of 8,093 cases. Am J Epidemiol. 161(11):1047–1055.

Braybrook C, Doudney K, Marcano ACB, Arnason A, Bjornsson A, Patton MA, Goodfellow PJ, Moore GE, Stanier P. 2001. The T-box transcription factor gene TBX22 is mutated in X-linked cleft palate and ankyloglossia. Nat Genet. 29(2):179–183.

Braybrook C, Lisgo S, Doudney K, Henderson D, Marçano AC, Strachan T, Patton MA, Villard L, Moore GE, Stanier P, et al. 2002. Craniofacial expression of human and murine TBX22 correlates with the cleft palate and ankyloglossia phenotype observed in CPX patients. Hum Mol Genet. 11(22):2793–2804.

Butali A, Mossey PA, Adeyemo WL, Eshete MA, Gaines LA, Even D, Braimah RO, Aregbesola SB, Rigdon JV, Emeka CI, et al. 2014. Novel IRF6 mutations in families with Van der Woude syndrome and popliteal pterygium syndrome from sub-Saharan Africa. Mol Genet Genomic Med. 2(3):254–260.

Christensen K, Juel K, Herskind AM, Murray JC. 2004. Long term follow up study of survival associated with cleft lip and palate at birth. BMJ. 328(7453):1405.

de la Garza G, Schleiffarth JR, Dunnwald M, Mankad A, Weirather JL, Bonde G, Butcher S, Mansour TA, Kousa YA, Fukazawa CF, et al. 2013. Interferon regulatory factor 6 promotes differentiation of the periderm by activating expression of Grainyhead-like 3. J Invest Dermatol. 133(1):68–77.

FitzPatrick DR, Carr IM, McLaren L, Leek JP, Wightman P, Williamson K, Gautier P, McGill N, Hayward C, Firth H, et al. 2003. Identification of SATB2 as the cleft palate gene on 2q32-q33. Hum Mol Genet. 12(19):2491–2501.

Ghassibe-Sabbagh M, Desmyter L, Langenberg T, Claes F, Boute O, Bayet B, Pellerin P, Hermans K, Backx L, Mansilla MA, et al. 2011. FAF1, a gene that is disrupted in cleft palate and has conserved function in zebrafish. Am J Hum Genet. 88(2):150–161.

Gowans LJJ, Adeyemo WL, Eshete M, Mossey PA, Busch T, Aregbesola B, Donkor P, Arthur FK, Bello SA, Martinez 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. J Dent Res. 95(11):1245–1256.

Hendriks IA, Vertegaal AC. 2015. SUMO in the DNA damage response. Oncotarget. 6(18):15734–15735.

Ingraham CR, Kinoshita A, Kondo S, Yang B, Sajan S, Trout KJ, Malik MI, Dunnwald M, Goudy SL, Lovett M, et al. 2006. Abnormal skin, limb and craniofacial morphogenesis in mice deficient for interferon regulatory factor 6 (Irf6). Nat Genet. 38(11):1335–1340.

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.

Leslie EJ, Carlson JC, Shaffer JR, Butali A, Buxo CJ, Castilla EE, Christensen K, Deleyiannis FW, Leigh Field L, Hecht JT, et al. 2017. Genome-wide meta-analyses of nonsyndromic orofacial clefts identify novel associations between FOXE1 and all orofacial clefts, and TP63 and cleft lip with or without cleft palate. Hum Genet. 136(3):275–286.

Leslie EJ, Carlson JC, Shaffer JR, Feingold E, Wehby G, Laurie CA, Jain D, Laurie CC, Doheny KF, McHenry T, et al. 2016. A multi-ethnic genome-wide association study identifies novel loci for non-syndromic cleft lip with or without cleft palate on 2p24.2, 17q23 and 19q13. Hum Mol Genet. 25(13):2862–2872.

Leslie EJ, Koboldt DC, Kang CJ, Ma L, Hecht JT, Wehby GL, Christensen K, Czeizel AE, Deleyiannis FW, Fulton RS, et al. 2016. IRF6 mutation screening in non-syndromic orofacial clefting: analysis of 1521 families. Clin Genet. 90(1):28–34.

Leslie EJ, Liu H, Carlson JC, Shaffer JR, Feingold E, Wehby G, Laurie CA, Jain D, Laurie CC, Doheny KF, et al. 2016. Genome-wide association study of nonsyndromic cleft palate identifies an etiologic missense variant in GRHL3. Am J Hum Genet. 98(4):744–754.

Little J, Cardy A, Arslan MT, Gilmour M, Mossey PA. 2004. Smoking and orofacial clefts: a United Kingdom-based case-control study. Cleft Palate Craniofac J. 41(4):381–386.

Ludwig KU, Mangold E, Herms S, Nowak S, Reutter H, Paul A, Becker J, Herberz R, Alchawa T, Nasser E, et al. 2012. Genome-wide meta-analyses of nonsyndromic cleft lip with or without cleft palate identify six new risk loci. Nat Genet. 44(9):968–971.

Mangold E, Böhmer AC, Ishorst N, Hoebel AK, Gültepe P, Schuenke H, Klamt J, Hofmann A, Gölz L, Raff R, 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.

Marazita ML, Field LL, Cooper ME, Tobias R, Maher BS, Peanchitlertkajorn S, Liu YE. 2002. Nonsyndromic cleft lip with or without cleft palate in China: assessment of candidate regions. Cleft Palate Craniofac J. 39(2):149–156.

Marçano AC, Doudney K, Braybrook C, Squires R, Patton MA, Lees MM, Richieri-Costa A, Lidral AC, Murray JC, Moore GE, et al. 2004. TBX22 mutations are a frequent cause of cleft palate. J Med Genet. 41(1):68–74.

Mossey PA, Little J. 2002. Epidemiology of oral clefts: an international perspective. In: Wyszynski DF, editor. Cleft lip and palate: from origin to treatment. Oxford (UK): Oxford University Press. p. 127–144.

Peyrard-Janvid M, Leslie EJ, Kousa YA, Smith TL, Dunnwald M, Magnusson M, Lentz BA, Unneberg P, Fransson I, Koillinen HK, 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.

Rauch GJ, Lyons DA, Middendorf I, Friedlander B, Arana N, Reyes T, Talbot WS. 2003. Submission and curation of gene expression data [ZFIN direct data submission]. http://zfin.org.

Richardson RJ, Dixon J, Jiang R, Dixon MJ. 2009. Integration of IRF6 and Jagged2 signalling is essential for controlling palatal adhesion and fusion competence. Hum Mol Genet. 18(14):2632–2642.

Sabel JL, d'Alencon C, O'Brien EK, Van Otterloo E, Lutz K, Cuykendall TN, Schutte BC, Houston DW, Cornell RA. 2009. Maternal interferon regulatory factor 6 is required for the differentiation of primary superficial epithelia in Danio and Xenopus embryos. Dev Biol. 325(1):249–262.

Sun Y, Huang Y, Yin A, Pan Y, Wang Y, Wang C, Du Y, Wang M, Lan F, Hu Z, et al. 2015. Genome-wide association study identifies a new susceptibility locus for cleft lip with or without a cleft palate. Nat Commun. 6:6414.

Ting SB, Caddy J, Wilanowski T, Auden A, Cunningham JM, Elias PM, Holleran WM, Jane SM. 2005. The epidermis of grhl3-null mice displays altered lipid processing and cellular hyperproliferation. Organogenesis. 2(2):33–35.

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.

Waitzman NJ, Romano PS, Scheffler RM. 1994. Estimates of the economic costs of birth defects. Inquiry. 31(2):188–205.

Yu Y, Zuo X, He M, Gao J, Fu Y, Qin C, Meng L, Wang W, Song Y, Cheng Y, et al. 2017. Genome-wide analyses of non-syndromic cleft lip with palate identify 14 novel loci and genetic heterogeneity. Nat Commun. 8:14364.