Parallel molecular mechanisms underlie convergent evolution of the exaggerated snout phenotype in East African cichlids

Anna Duenser[1], Pooja Singh[1,2,3], Laurène Alicia Lecaudey[4], Christian Sturmbauer[1]*, Craig Albertson[5], Wolfgang Gessl[1], Ehsan Pashay Ahi[1,6]*

1. Institute of Biology, University of Graz, Universitätsplatz 2, A-8010 Graz, Austria
2. Aquatic Ecology and Evolution, Institute of Ecology and Evolution, University of Bern, Bern, Switzerland
3. EAWAG, Swiss Federal Institute of Aquatic Science and Technology, Kastanienbaum, Switzerland
4. Department of Natural History, NTNU University Museum, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway
5. Department of Biology, University of Massachusetts, Amherst MA 01003, USA
6. Organismal and Evolutionary Biology Research Programme, University of Helsinki, Viikinkaari 9, 00014, Helsinki, Finland

*Authors of Correspondence:

Ehsan Pashay Ahi[1,6], Email: ehsanpashayahi@gmail.com

Christian Sturmbauer[1], Email: christian.sturmbauer@uni-graz.at

Abstract

Studying instances of convergent evolution of novel phenotypes can shed light on the evolutionary constraints that shape morphological diversity. Cichlid fishes from the East African Great Lakes are a prime model to investigate convergent adaptations. However, most studies on cichlid craniofacial morphologies have primarily considered bony structures, while soft tissue adaptations have been less intensely scrutinised. A rare example of an exaggerated soft tissue phenotype is the formation of a snout flap. This tissue flap develops from the upper lip and has evolved in only one cichlid genus from Lake Malawi and one genus from Lake Tanganyika. To investigate the molecular basis of snout flap convergence, we used mRNA sequencing to compare two species with snout flap (*Labeotropheus trewavasae* and *Ophthalmotilapia nasuta*) to their close relatives without snout flaps (*Tropheops tropheops* and *Ophthalmotilapia ventralis*) from Lake Tanganyika and Malawi. Our analysis revealed a greater complexity of differential gene expression patterns underlying the snout flap in the younger adaptive radiation of Lake Malawi than in the older Lake Tanganyika radiation. We identified 201 genes that were repeatedly differentially expressed between species with and without the snout flap in both lakes, suggesting that the pathway that gives rise to snout flaps is evolutionarily constrained, even though the flaps play very different functions in each species. The convergently expressed genes are involved in proline and hydroxyproline metabolism, which have been linked to human skin and facial deformities. Additionally, we also found enrichment for transcription factor binding sites upstream of differentially expressed genes such as members of the FOX transcription factor family, especially *foxf1* and *foxa2*, which also showed an increased expression in the flapped snout and are linked to nose morphogenesis in mammals, as well as *ap4* (tfap4), a transcription factor showing reduced expression in the flapped snout with an unknown role in the development of craniofacial soft tissues. As genes involved in cichlids snout flap development are associated with many

human mid-line facial dysmorphologies, our findings imply a conservation of genes involved in mid-line patterning across vastly distant evolutionary lineages of vertebrates.

Key words: RNA-seq, mid-line patterning, soft tissue adaptation, Lake Malawi, Lake Tanganyika, snout flap, cichlids

Significance statement

The study of the evolution of similar physical traits across taxa can give insight into the molecular architecture underlying shared phenotypes. This has mostly been studied in bony structures, while soft tissue traits have been less intensely covered. We investigated the exaggerated snout in cichlid species from Lake Malawi and Lake Tanganyika and found that many genes involved in the development of the snout flap are also associated with mid-line dysmorphologies in humans, implying a conservation across distant vertebrate lineages.

Introduction

Convergent evolution of phenotypes, reflecting particular ecological specializations, is a ubiquitous characteristic of adaptive radiations (Schluter & Nagel 1995; Losos et al. 1998; Rundle et al. 2000; Rüber et al. 1999). Cichlid adaptive radiations from the East African Great lakes display an impressive array of repeated morphological traits (Kocher et al. 1993), including a few dramatic examples such as nuchal humps and hypertrophied lips (Machado-Schiaffino et al. 2014; Manousaki et al. 2013; Colombo et al. 2013; Baumgarten et al. 2015; Lecaudey et al. 2019, 2021). However, the evolution of such phenotypic novelties is not well understood, but a comparative approach can shed light on the genetic and developmental mechanisms that reconfigure the body plan and give rise to complex traits. Cases of repeated

evolution of such phenotypic novelties can thus also help us to understand the constraints that shape morphological diversity.

The overgrowth of craniofacial soft tissues in cichlid fishes has presented some intriguing examples of exaggerated phenotypes in various anatomical regions such as lips, frontal head (nuchal hump) and nose snout (or nose flap) (Colombo et al. 2013; Manousaki et al. 2013; Henning et al. 2017; Lecaudey et al. 2019; Concannon & Albertson 2015; Conith et al. 2018). The exaggerated snout flap is a pronounced projection that emanates from a flap of fibrous tissue just above the upper lip. It is a rare morphological innovation that has only evolved in two tribes of cichlid fishes from East Africa, the modern Haplochromines in Lake Malawi and the Ectodini in Lake Tanganyika (Concannon & Albertson 2015). When this snout is sexually monomorphic, it is thought to be a trophic adaptation that improves feeding efficiency (Konings 2007). When the snout is sexually dimorphic, it is hypothesised to be involved in sexual selection (Konings 2007; Concannon & Albertson 2015). The cichlid snout flap has been studied at the molecular level only in the genus *Labeotropheus* from Lake Malawi where it is sexually monomorphic and functions as a trophic adaptation to efficiently leverage algae from rocks (Concannon & Albertson 2015; Conith et al. 2018). A similar snout structure has also been described in two species from the Ectodini tribe (*Ophthalmotilapia nasuta* and *Asprotilapia leptura*) from Lake Tanganyika. In *A. leptura* it is sexually monomorphic and likely involved in increased foraging efficiency (similar to *Labeotropheus*), whereas in *O. nasuta* it is only found in mature males and is likely a secondary sexual character (Hanssens et al. 1999; Conith et al. 2019). Thus, the exaggerated snout is a convergent phenotype that evolved independently in two cichlid lineages that diverged > 9 MYA (Irisarri et al. 2018; Conith et al. 2019).

In *Labeotropheus*, the snout is evident histologically by the time the yolk is absorbed and exogenous feeding occurs (~1 month post-fertilization) (Conith et al. 2018; Concannon &

Albertson 2015), and early formation and growth of the snout is linked to the transforming growth factor beta (TGFβ) signalling pathway (Conith et al. 2018). However, it remains unclear whether (1) similar molecular players are involved in maintaining this phenotype at later life-history stages and if (2) the molecular mechanism of snout formation is the same in other cichlid species that possess a snout. Furthermore, while previous research focused on the TGFβ signalling pathway, a more extensive molecular interaction map of the formation and maintenance of this exaggerated phenotype remains to be unravelled. A transcriptome-wide overview is particularly important since it is well-known that there is molecular cross-talk between the TGFβ signalling pathway and several other pathways which all play a pivotal role in craniofacial morphogenesis and adaptive evolutionary divergence in teleost fishes (Ahi 2016).

In this study, we set out to investigate the molecular mechanisms that underlie the convergent evolution of the exaggerated snout phenotype, in two non-sister cichlid lineages from lakes Tanganyika and Malawi (Figure 1A). We compared two species that develop the snout; (1) *Labeotropheus trewavasae* (tribe Haplochromini) from Lake Malawi and (2) *Ophthalmotilapia nasuta* (tribe Ectodini) from Lake Tanganyika (Figure 1). As controls, we used two closely related species within each tribe that do not develop such a structure; (1) the Lake Malawi mbuna species *Tropheops tropheops* (Haplochromini) and (2) the Lake Tanganyika featherfin cichlid *Ophthalmotilapia ventralis* (Ectodini) (Figure 1). We used mRNA-sequencing to quantify gene expression differences between the exaggerated snout and non-snout tissues for each lake. Altogether, we identified parallel and non-parallel molecular mechanisms that underlie the evolution of the snout flap in Lake Malawi and Lake Tanganyika cichlids. Our study design provides valuable information on convergent regulatory mechanisms underlying the morphogenesis of a unique hypertrophic facial soft

tissue in cichlids, which also exhibit striking similarity to those mechanisms driving craniofacial development and mid-line patterning in other vertebrates including humans.

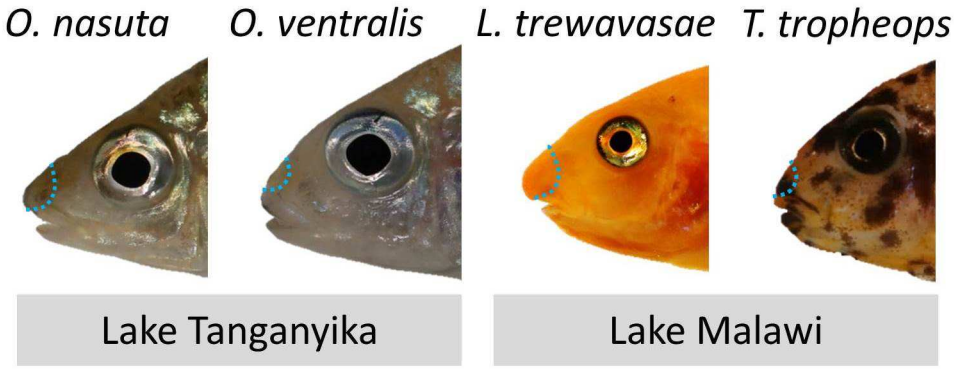

Fig. 1. Convergent cases of snout evolution. East African cichlid species used in this study. The area of the soft tissue that was dissected is depicted by blue dashed lines.

Results

RNA-sequencing, gene expression and downstream analyses

The RNA-sequencing resulted in between 6.7 and 15.8 million reads per sample and after filtering of low-quality reads, between 4.6 and 11.1 million reads were retained for each sample (Supplementary File S1). The raw data of sequence reads have been deposited in the Sequencing Read Archive (SRA) of NCBI (accession number: PRJNA770252). The pairwise comparisons between species of each lake radiation resulted in 832 differentially expressed genes for the comparison of *O. nasuta* versus *O. ventralis*, while the comparison between *L. trawavasae* versus *T. tropheops* yielded 4292 differentially expressed genes. Between these both results we identified an overlapping list of 201 differentially expressed (DE) genes which were distinct between the flapped snout versus the non-flapped snout regions in both lakes (Figure 2A) (Supplementary File S2). Among the DE genes, 74 genes showed

upregulation and 96 genes showed downregulation in snout tissues in both comparisons, whereas 31 genes showed expression differences in opposite directions across the comparisons for each Lake (Figure 2B-D). The heatmap clustering of the DE genes showed that there are at least two major branches in each group of up- or down-regulated gene sets, while the clustering of the DE gene with opposite expression pattern also revealed the presence of two major branches (Figure 2B-D). These clustering structures indicate distinct transcriptional regulations within each group which potentially originated from the effects of different upstream regulators.

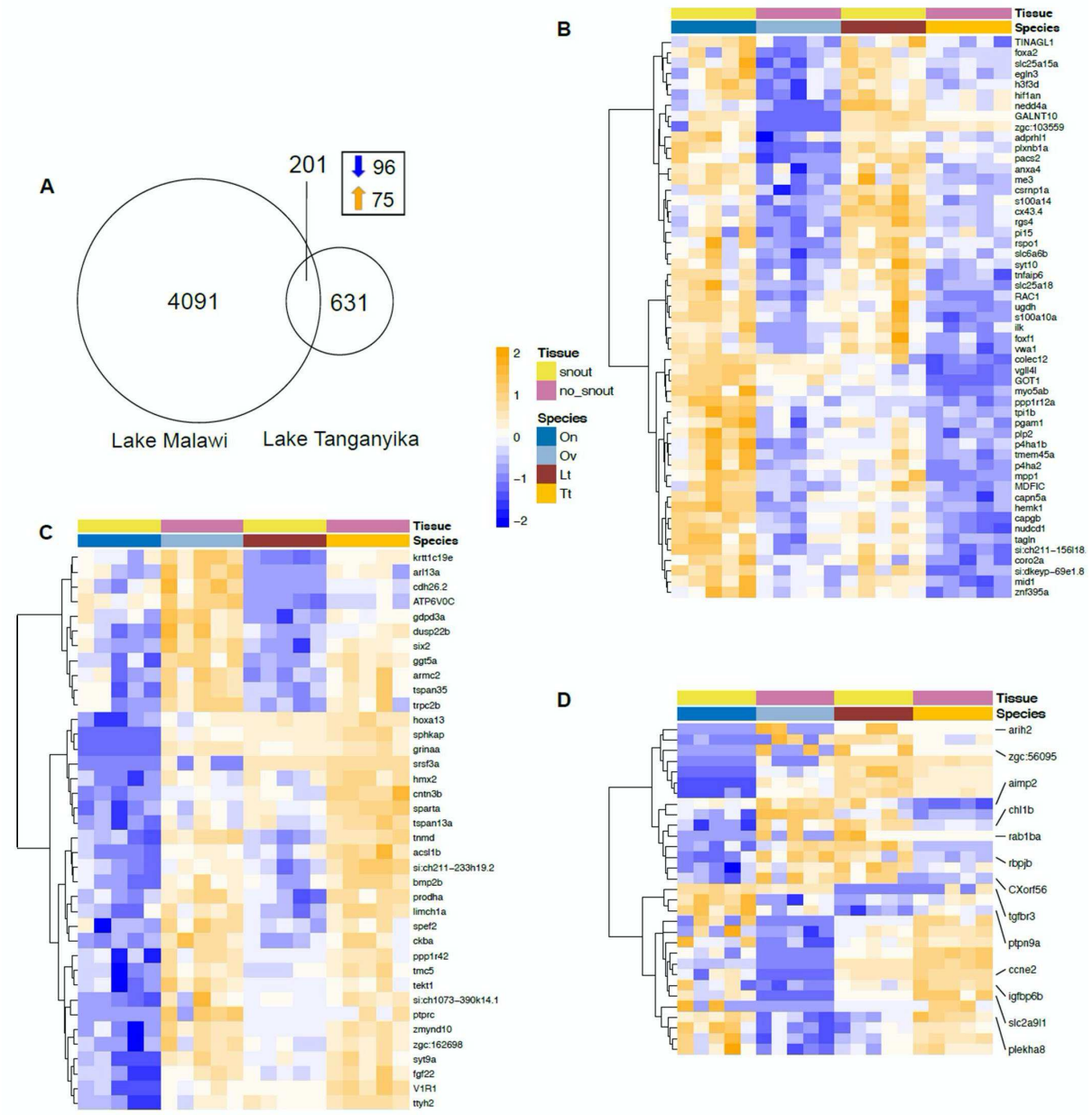

Fig. 2. Differentially expressed genes in the snout regions. (**A**) Venn diagram of 201 genes with differential expression between the snout regions ("snout" and "no snout") which overlap between the two comparisons. Dendrogram clusters of the overlapping annotated genes showing upregulation (**B**), and downregulation (**C**) in expression in the flapped snout tissue, as well as those showing differential expression in both comparisons but in opposite direction (**D**). Orange and blue shadings indicate higher and lower relative expression respectively. *Ophthalmotilapia ventralis* (Ov)*, Ophthalmotilapia nasuta* (On)*, Labeotropheus trewavasae* (Lt), *Tropheops tropheops* (Tt).

We next performed gene ontology enrichment analysis using the list of 201 DE genes as the input, and the result showed significant enrichment of GO terms for several biological processes such as amino acid metabolism (particularly proline related metabolic processes), Wnt and Notch signalling pathways, and regulation of cell adhesion (Figure 3A).

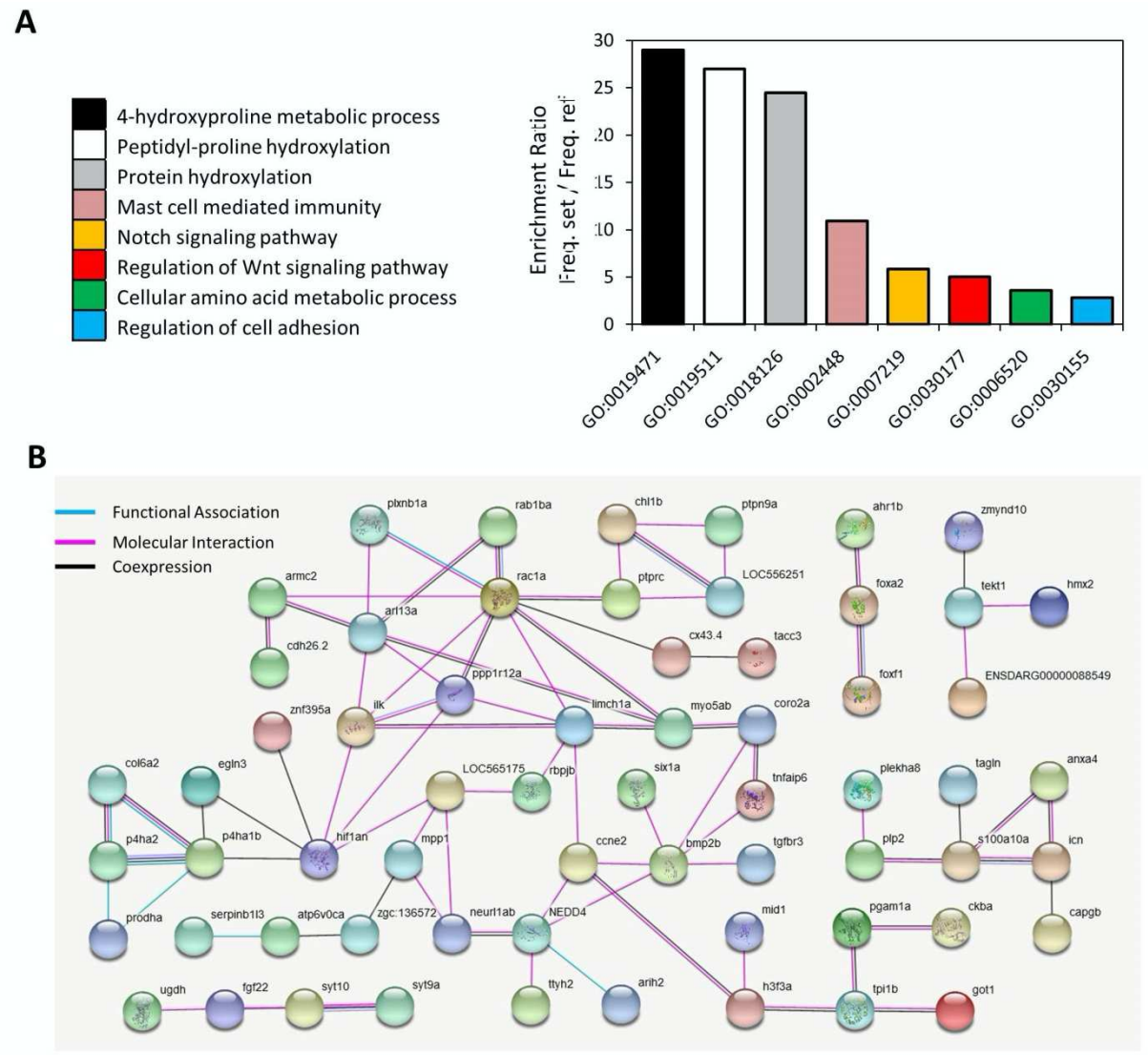

Fig. 3. Functional analyses of the overlapping differentially expressed genes in flapped snout. (**A**) Enrichment for gene ontologies of biological processes using the shared 201 differentially expressed genes and Manteia online tool. (**B**) Functional interactions between the differentially expressed genes predicted based on zebrafish databases in STRING v10 (http://string-db.org/).

We also applied the same list of DE genes for interactome analysis which demonstrated a large, interconnected network of genes with molecular and functional associations. Some of the genes in the network formed an interaction hub with the highest level of associations with other genes such as *bmp2b*, *hif1an* and *rac1a*, suggesting their more pivotal role in the formation of the flapped snout structure in cichlids (Figure 3B). Furthermore, we conducted TF binding motif overrepresentation analysis on the upstream regulatory sequences of the DE

175 genes through MEME tool (Bailey et al. 2009). In total, seven motifs were enriched on the
176 upstream regulatory sequences of at least 40 out of 201 DE genes (Table 1). Next, we
177 checked the similarities of the enriched motifs with known TF binding sites in vertebrates and
178 at least 11 TF candidates were identified to potentially bind to those motifs.

179

180 **Table 1. Predicted motifs and upstream regulators potentially binding to them.** Enriched motifs
181 on upstream regulatory sequences of the DE genes are presented in degenerated sequence format.
182 PWD IDs indicate positional weight matrix ID of predicted binding sites and E-values refer to
183 matching similarity between the predicted motif sequences and the PWD IDs. The count implies the
184 number of genes containing the predicted motif sequence on their regulatory region.

TF binding site	PWM ID	Count	Predicted motif sequence	E-value
FOXP1	M00987	71 / 201	AMAMACAMAMAMAMAMACACACAMAMACA	3.85E-12
FOXJ1	M00742			3.52E-08
RREB1	MA0073.1			1.87E-07
FOXJ1	M00742	47 / 201	AAAAASAAAMAAAMWMWCWKT	8.69E-10
FOX	M00809			9.15E-07
FOXD3	MA0041.1			3.95E-07
SP1	MA0079.2	41 / 201	CHCCYCCYCCYCCSCYCTCCY	1.12E-08
IRF9	M00258	61 / 201	KTTTTTYTTTYYCWK	2.90E-09
MEF2	M00405	72 / 201	RTTAAAAAAAA	4.28E-08
AP4	M00927	93 / 201	CWGCTGCWGCTKSTS	7.38E-08
HEB/tcf12	M00698	66 / 201	NYYCTGCTGD	1.03E-06

185

186 **Expression analysis by qPCR**

187 Validation of DE genes from RNA-seq was accomplished via quantitative RT-PCR (qPCR),
188 normalised to stably expressed reference genes (Kubista et al. 2006). In our previous studies

of East African cichlids, we found that validation of reference gene(s) is an essential step for every species and tissues (Ehsan P. Ahi et al. 2020; Ehsan Pashay Ahi et al. 2020; Pashay Ahi & Sefc 2018; Ahi et al. 2019). We chose six candidate reference genes with a small log2 fold change and the lowest coefficient of variation (CV) throughout all the samples (Supplementary File S2). Based on the rankings by the three software tools, BestKeeper, geNorm and NormFinder, only one of the candidate reference genes, *pak2b*, showed consistent stability, i.e. always ranked among top two most stable reference genes (Table 2). Thus, we used the Cq value of *pak2b* in each sample to normalize the relative gene expression levels of our target genes.

Table 2. Ranking of reference genes in the nose tissue samples using three different algorithms.

BestKeeper				geNorm		NormFinder	
Ranking	**SD**	**Ranking**	**r**	**Ranking**	**M**	**Ranking**	**SV**
sp3	0.461	*pak2b*	0.94	*pak2b*	0.369	*nup58*	0.310
pak2b	0.471	*pphln1*	0.931	*flot2a*	0.386	*pak2b*	0.408
flot2a	0.491	*flot2a*	0.926	*pphln1*	0.397	*pphln1*	0.443
nup58	0.509	*vps26a*	0.916	*sp3*	0.418	*flot2a*	0.498
vps26a	0.551	*nup58*	0.889	*nup58*	0.427	*sp3*	0.518
pphln1	0.587	*sp3*	0.887	*vps26a*	0.428	*vps26a*	0.646

Abbreviations: SD = Standard deviation, r = Pearson product-moment correlation coefficient, SV = stability value, M = M value of stability.

Among the DE genes identified by RNA-seq, we chose 12 genes with a known role in nose morphogenesis and/or other related functions in craniofacial development mainly based on genetic studies in humans (Table 3), together with eight predicted upstream TFs (including *ap4*, *foxd3*, *foxj1*, *foxp1*, *irf9*, *mef2a*, *rreb1a* and *sp1*) for qPCR analysis (Figure 4).

Table 3. A selected set of differentially expressed genes in the flapped snout regions of studied cichlids with known related functions in nose morphogenesis in mammalian models.

Gene	Related function	Organism	References
adprhl1	Duplication of this gene is associated with prominent forehead, short and bulbous nose, and broad philtrum	Human	(De Pater et al. 2005)
bmp2	A ligand of the TGFβ signaling and its monoallelic deletion is associated with short upturned nose and long philtrum	Human	(Tan et al. 2017)
dusp22	Deletion at terminal end of this gene is associated with saddle shape nose morphogenesis	Human	(Hosono et al. 2020)
fgf22	Genomic rearrangement encompassing this gene is associated with elongation of nose with prominent nasal bridge	Human	(Quigley et al. 2004)
foxa2	Both deletion and missense variation in this gene causes hallow nasal bridge, short upturned nose and downturned nasolabial folds	Human	(Dines et al. 2019)
foxf1	Duplication and triplication causes bulbous nose and wide nasal bridge	Human	(Kucharczyk et al. 2014)
hmx2	Hemizygous deletion in this gene causes broad nasal bridge and prominent nose	Human	(Miller et al. 2009)
lyve1	Dysregulation of this gene is associated with cutaneous angiosarcoma on the nose	Human	(Mitteldorf et al. 2018)
prodh	Deletion and/or missense mutations in this gene causes frontal bossing, thin upper lip and short nose	Human	(Guilmatre et al. 2010)
rac1	Loss of function mutation in this gene causes failure in fusion of medial nasal processes and prominent nasal bridge	Human Mouse	(Thomas et al. 2010)(Reijnders et al. 2017)
six2	Deletion in this gene causes frontonasal dysplasia syndrome in human with nasal clefting and broad nasal tip, and developmental deformities in nasal bridge in mouse	Human Mouse	(Hufnagel et al. 2016)(Okello et al. 2017)
ugdh	Missense mutation in this gene causes bulbous nose and smooth philtrum	Human	(Alhamoudi et al. 2020)

Based on the RNA-seq results, six of these candidate genes displayed up-regulation in expression in the flapped snout (*adprhl1*, *foxa2*, *foxf1*, *lyve1a*, *rac1* and *ugdh*), while the six other candidate genes (*bmp2b*, *dusp22b*, *fgf22*, *hmx2*, *prodha* and *six2a*) showed a down-regulation in expression in the flapped snout. The results of qPCR analysis confirmed that almost all of the genes showed expression patterns similar to RNA-seq results, except for *bmp2b* and *hmx2* which showed no significant difference between the snout regions of L.t and T.t. Among the predicted TFs only *ap4* showed consistent differences across both comparisons displaying a slightly reduced expression in both species with protruded snouts (O.n and L.t). This indicates potential transcriptional repressor effects of *ap4* on the downstream genes in the hypertrophic snout region. Two members of FOX transcription factors, *foxj1* and *foxp1*, also showed expression differences but only in one of the comparisons (O.n vs O.v). Altogether, the qPCR results demonstrate consistency between RNA-seq and qPCR results confirming the validity of our transcriptome data analysis.

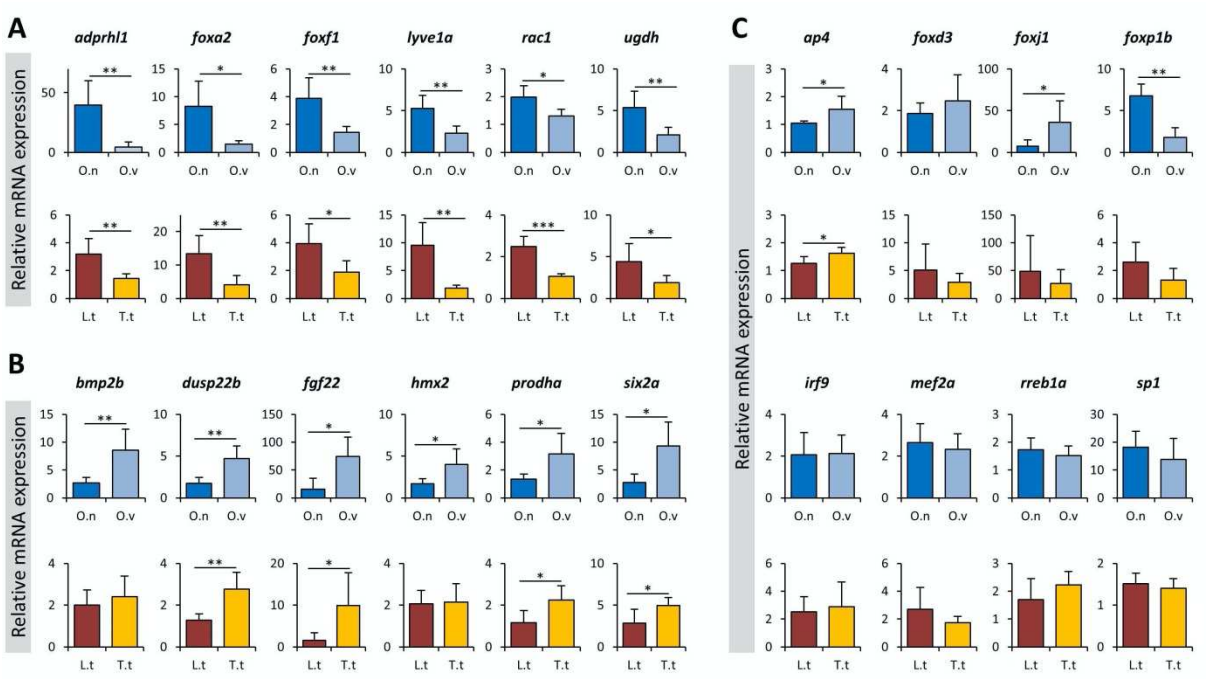

Fig. 4. qPCR expression analysis of a selected set of candidate genes. The bars indicate mean and standard deviation of RQ expression values for five biological replicates per species. The asterisks above the bar represent significant expression differences (*P < 0.05; **P < 0.01; ***P < 0.001). *Ophthalmotilapia ventralis* (O.v), *Ophthalmotilapia nasuta* (O.n), *Labeotropheus trewavasae* (L.t), *Tropheops tropheops* (T.t)

Discussion

Biologists have been fascianted with divergent as well as convergent morphological evolution as it contributes to our understanding of the interplay between ecological opportunity and genetic constraints. The snout flap of *L. trevawasae* is thought to have evolved under natural selection (Concannon & Albertson 2015), as it plays a distinct role in the foraging efficiency for algal scraping (Konings 2007; Conith et al. 2019). Additionally, no difference in snout flap size has been detected between male and female of *Labeotropheus* and its formation has been shown to coincide with the time point of independent foraging, further supporting its function (Concannon & Albertson 2015). In contrast, only *O. nasuta* males show the distinct snout flap, implicating its role in mate choice (Concannon &

Albertson 2015). Furthermore, both sexes of *O. nasuta* are planktivorous suction feeders, a feeding adaptation that is presumably not enhanced by a snout flap, although the snout of males continues to grow with increasing age (Hanssens et al. 1999). In a comparison of tissue types of the snout flap it has been found that the snout of *Labeotropheus fuelleborni* contains a lot more of intermaxillary ligament and loose connective tissue (80%) than the snout of *O. nasuta* (50%) (Conith et al. 2019). The morphological convergence of the snout flap across two cichlid radiations allows us to investigate if convergence can also be found at the transcriptional level, even if the morphologies probably possess different functions and differ in tissue composition and life-history.

We found many of DE genes, both upregulated and downregulated, that are associated with craniofacial development and involved in human dysmorphologies, many with mid-line facial defects including those that effect the nose in literature. Among the upregulated genes with related functions were *adprhl1* (De Pater et al. 2005), *angptl2* (Ehret et al. 2015), *colec12* (Zlotina et al. 2016), *cx43* (McLachlan et al. 2005), *foxa2* (Dines et al. 2019), *foxf1* (Kucharczyk et al. 2014), *galnt10* (Starkovich et al. 2016), *got1* (Tomkins et al. 1983), *lyve1* (Mitteldorf et al. 2018), *mdfic* (Kosho et al. 2008), *mid1* (Preiksaitiene et al. 2015)(Hüning et al. 2013), *nudcd1* (Selenti et al. 2015), *pacs2* (Holder et al. 2012), *plxnb1* (Haldeman-Englert et al. 2009), *rac1* (Thomas et al. 2010)(Reijnders et al. 2017), *rspo1* (Wieacker & Volleth 2007), *s100a10* (Sawyer et al. 2007), *slc25a18* (Chen et al. 2013), *slc6a6* (Kariminejad et al. 2015), *ugdh* (Alhamoudi et al. 2020), *vgll4* (Czeschik et al. 2014)(Barrionuevo et al. 2014), and *vwa1* (Giannikou et al. 2012). Among the downregulated genes we also found the following candidates to have such roles; *acsl1* (Yakut et al. 2015), *adgb* (Alazami et al. 2016), *arl13* (Brugmann et al. 2010), *ATP6v0c* (Mucha et al. 2019; Tinker et al. 2021), *bmp2* (Tan et al. 2017), *cntn3* (Ţuţulan-Cuniţă et al. 2012), *dusp22* (Hosono et al. 2020)(Martinez-Glez et al. 2007), *fgf22* (Quigley et al. 2004), *gdpd3* (Dell'Edera et al. 2018), *grina* (Bonaglia

et al. 2005), *hmx2* (Miller et al. 2009), *hoxa13* (Fryssira et al. 2011), *il23r* (Rivera-Pedroza et al. 2017), *ppp1r42* (Mordaunt et al. 2015), *prodh* (Guilmatre et al. 2010), *six2* (Hufnagel et al. 2016)(Okello et al. 2017), *srsf3* (Pillai et al. 2019), *syt9* (Sofos et al. 2012), and *trpc2* (Sansone et al. 2014)(Zhang et al. 2010). Interestingly, one of the downregulated genes, *pi15*, is known as an important molecular player in beak formation in birds (Nimmagadda et al. 2015). Even among the overlapping DE genes which showed opposing expression patterns between the two comparisons, we still found at least four genes to have been associated with craniofacial mid-line defects in other vertebrates, including *ccne2* (Jain et al. 2010), *plekha8* (Schulz et al. 2008), *rab1b* (Alwadei et al. 2016), *rbpj* (Nakayama et al. 2014) and *tgfbr3* (Lopes et al. 2019). These findings demonstrate that similar sets of genes are involved in mid-line patterning and growth across evoulutionary distant vertebrates. Thus, functional studies investigating their specific role in divergent morphogenesis of snout structures in fish can provide valuable information about the conserved molecular mechanisms underlying the formation of facial soft tissues (Powder & Albertson 2016).

Conducting gene ontology enrichment analysis on the list of DE genes also revealed the involvement of several biological processes such as proline and hydroxyproline metabolisms, regulation of cell adhesion, as well as Notch and Wnt signalling pathways in the formation of the flapped snout in these cichlid species. Interestingly, a defective proline and hydroxyproline metabolisms has been already associated with a range of skin and facial deformities including abnormal nose morphogenesis in humans (Kiratli & Satilmiş 1998; Kretz et al. 2011; Zaki et al. 2016; Baumgartner et al. 2016). A defective proline metabolism is known to severely affect collagen formation and extracellular matrix integrity, and subsequently cell adhesion (Velez et al. 2019; Karna et al. 2020; Xinjie et al. 2001; Javitt et al. 2019; Noguchi et al. 2020). The lack of proline hydroxylation in collagen by reduced prolyl 4-hydroxylase activity can directly affect integrin binding and cell adhesion

295 mechanisms (Sipila et al. 2018). Interestingly, we found genes involved in 4-hydroxyproline
296 metabolic process as the most enriched biological process, which suggests changes in proline
297 hydroxylation as top candidate of metabolic changes during the formation of exaggerated
298 snout in cichlids. In addition, it has been recently shown that the biosynthesis of proline is
299 tightly regulated by transforming growth factor-beta (Tgfβ) (Schwörer et al. 2020), a
300 transcription factor that is also playing an important role in the early development of the
301 flapped snout structure in cichlids (Conith et al. 2018). Although, we did not detect
302 differential expression of *Tgfβ1* itself, components of this pathway were identified (e.g.,
303 *tgfbr3*), and both of the enriched signalling pathways, Wnt and Notch, have evolutionary
304 conserved crosstalk with Tgfβ mediated signal in regulation of various molecular, cellular
305 and developmental events (Attisano & Labbé 2004; Chesnutt et al. 2004; Arnold et al. 2019;
306 Klüppel & Wrana 2005; Ahi 2016). In addition, both Wnt and Notch signalling pathways are
307 known to play a pivotal role in craniofacial development and morphogenesis including the
308 formation of middle structures including nasal structures (Penton et al. 2012; Pakvasa et al.
309 2020; Brugmann et al. 2007; Wang et al. 2011).
310 A proposed model for Tgfβ and Notch crosstalk suggests that induction of Tgfβ signalling is
311 required for the early establishment of cell-cell contacts in different tissues, whereas later
312 induction of Notch signal stabilizes the Tgfβ mediated effects (Klüppel & Wrana 2005). This
313 allows the cells to react to a new environment through induction of alternative genetic
314 programs controlling differentiation or migration (Klüppel & Wrana 2005). In the context of
315 the snout, it is possible that activation of Tgfβ is required for early snout induction (Conith et
316 al. 2018) and that continued snout growth is maintained via Notch signalling. This potential
317 time dependent crosstalk may be mediated through the downstream targets of Notch and Tgfβ
318 signals, since it is shown that both signals can regulate similar target genes (Klüppel &
319 Wrana 2005; De Jong et al. 2004), including *foxa2*, a member of the FOX family of

transcription factors (both signals suppress *foxa2* expression) (Kondratyeva et al. 2016; Liu et al. 2012). In our study, we found upregulation of *foxa2* in the flapped snout region, and interestingly, a recent study in human shows that a deletion in *Foxa2* can cause specific facial deformities including a shallow nasal bridge, a short upturned nose, and a downturned nasolabial fold (Dines et al. 2019). In addition, we found *rbpjb*, a major transcription factor mediating canonical Notch signal in various cell types (Tanigaki et al. 2002) to be differentially expressed in the flapped snout of both species. *rbpj* is shown to directly regulate a receptor of Tgfβ signal (*Tgfbr1*) in mice, thus making a reciprocal positive regulatory loop between the two pathways (Valdez et al. 2012). We also found another receptor of Tgfβ signal (*tgfbr3*) to show a similar expression pattern as *rbpjb* raising the possibility of the existence of such a reciprocal regulatory loop in flapped snout cichlids. In human, a deletion in *Rbpj* gene has been linked to developmental defects in brain and abnormal thickening of the nose and lip (Nakayama et al. 2014). Another genome-wide study revealed that *Rbpj* acts as a direct upstream regulator of Wnt signalling in mammalian stem cells (Li et al. 2012). On the other hand, Bmp2 signal which is regarded as a molecular cross point between Smad/Tgfβ and Notch pathways (De Jong et al. 2004), mediates its signal through *Tgfbr3* (Hill et al. 2012). It is already known that Bmp2 can regulate Notch signal and its downstream target genes (De Jong et al. 2004). Previous studies in cichlids had proposed variations in Bmp expression as a molecular player in adaptive morphological divergence in different skeletal structures (Gunter et al. 2013; Albertson et al. 2005; Ahi et al. 2017; Hulsey et al. 2016). We found down-regulation of *bmp2b* expression suggesting that a key regulator linking both pathways is affected in the flapped snout region of the cichlid species in this study. Furthermore, deletion of *Bmp2* in human has been reported to cause a range of facial deformities including shortened nose, anteverted nares, elongation of the philtrum and changes in the thickness of lips (Tan et al. 2017). Taken together, these findings suggest

345 potentially complex interactions between Notch and Tgfβ signals in the formation and
346 possibly the maintenance of the flapped snout structure in cichlids.
347 Finally, we also conducted enrichment for TF binding sites on regulatory sequences of DEGs
348 and found several potential binding sites for TFs that may play a role in the formation of a
349 flapped snout. The most represented TF binding sites belonged to members of FOX
350 transcription factor family, e.g. *foxd3*, *foxj1* and *foxp1*. In previous studies of East African
351 cichlids, both *foxd3* and *foxp1* were found to act upstream of a gene network involved in
352 exaggerated fin elongation (Pashay Ahi & Sefc 2018; Ahi et al. 2019). Additionally, *foxp1*
353 was recently identified as a key upstream regulator of genes involved in the formation of the
354 hypertrophic lip in another East African cichlid species from Lake Tanganyika (Lecaudey et
355 al. 2021). None of the predicted FOX members (*foxd3*, *foxj1* and *foxp1*) displayed consistent
356 differential expression across both comparisons, but interestingly, the consensus binding site
357 for FOX transcription factor family was also among the predicted binding sites. It is,
358 therefore, possible that the two other FOX members identified by RNA-seq and qPCR, *foxf1*
359 and *foxa2*, are the key regulators of the entire list of DEGs, since they might bind to the
360 consensus FOX binding site. In addition, both *foxf1* and *foxa2* displayed consistently
361 increased expression in the flapped snout in both comparisons, and are also implicated in the
362 nose morphogenesis in mammals (Dines et al. 2019; Kucharczyk et al. 2014). We have
363 recently found *foxf1* among the regulators of lip hypertrophy in an East African cichlid
364 species as well (Lecaudey et al. 2021), suggesting a potentially important and general role of
365 *foxf1* in soft tissues exaggeration in cichlids, and given the potential role of *foxd3* and *foxp1*
366 in cichlid fin exaggeration, FOX genes may play a more general role in tissue elaboration in
367 vertebrates. Among the other predicted TF binding site we found overrepresentation of
368 binding motif for *tcf12*, a transcription factor with known roles in cranio-skeletal
369 development; particularly in the morphogenesis of the frontal bone and cranial vault

370 thickening in mammals (Piard et al. 2015; Sharma et al. 2013). Moreover, we have previously
371 identified *tcf12* as a potential key player in the formation of a nuchal hump in an East African
372 cichlid (Lecaudey et al. 2019). In this study, however, we did not detect its expression in the
373 snout region by differential gene expression or qPCR analysis. The only predicted TF with
374 consistent expression difference in both comparisons was *ap4* (or *tfap4*), i.e. showing slight
375 but significant reduced expression in the flapped snout. *ap4* encodes a member of the basic
376 helix-loop-helix-zipper (bHLH-ZIP) family, and can act as a transcriptional activator or
377 repressor on a variety of downstream target genes mediating cell fate decisions (Wong et al.
378 2021). The exact role of *ap4* in craniofacial morphogenesis of soft tissues is unclear,
379 however, deletions in a genomic region containing this gene appeared to cause facial
380 dysmorphisms in human such as prominent beaked nose and micrognathia (Gervasini et al.
381 2007). Future functional studies are required to verify the role of *ap4* in formation and
382 morphogenesis of craniofacial soft tissues in fish.
383

384 ## Conclusions

385 The snout flap in *Labeotropheus trewavasae* and *Ophthalmotilapia nasuta* is a striking and
386 rare example of an exaggerated soft tissue trait that has evolved repeatedly in the cichlid
387 radiations of Lake Malawi and Lake Tanganyika, albeit for different functions. Comparing
388 the transcriptional landscape of the snout flap tissue of these two species with the snout of
389 close relatives that do not develop such a structure, we identified 201 genes that were
390 repeatedly recruited to give rise to the snout flap phenotype even after > 9 MYA of
391 divergence. Our study provides support for a change in proline hydroxylation, a mechanism
392 also linked to human facial deformations, to be a mechanism for metabolic changes involved
393 in the formation of the snout flap in fish. Additionally, we found indications of complex
394 interactions between the transforming growth factor-beta (Tgfβ), regulating the biosynthesis

of proline, and Notch signalling, associated with morphogenesis and craniofacial development, in the formation and maintenance of the snout flap. Upstream of the differentially expressed genes we identified transcription factors belonging to the FOX family (especially *foxf1* and *foxa2*) which are both linked to the morphogenesis of the nose in mammals and *ap4* a transcription factor that showed reduced expression in the species with snout flap, but with an unknown role in craniofacial soft tissue formation. We want to emphasise that the identification of genes involved in snout morphogenesis in fish can shed light on the conserved molecular mechanisms crucial for the development and shaping of facial soft tissue. In the future it would be important to build on these findings and confirm the reuse of these genes and pathways across more distant teleost groups.

Methods

Fish rearing and tissue sampling

Five captive bred males of each *O. nasuta*, *O. ventralis*, and five captive bred females of *L. trewavasae* and *T. tropheops* were raised and kept in a large tank (approximately 450 litres) containing multiple stony shelters to decrease competition stress. All specimens were at the young adult stage and have been fed with the same diet, Tropical multi-ingredient flakes suitable for omnivorous cichlids. The two species in each comparison were sampled at the same time when the protrusion of the flapped snout had already appeared (Figure 1). To perform the dissections, we used a solution with 0.3 g MS222 per 1L water to euthanize the fish, and similar snout regions, an area above the upper lip encompassing the nostrils which includes epidermis, dermis and the underlying soft connective tissues, were sampled for each fish (Figure 1). The sampled snout tissues for each individual were placed into separate tubes containing RNAlater (Qiagen) and stored at -20 C°. The sacrificing of fish followed the

guidelines of the Federal Ministry of Science, Research and Economy of Austria according to the regulations of the BMWFW.

RNA extraction and cDNA synthesis

Total RNA was extracted from 20 dissected snout tissue samples (5 biological replicates per species) following the TRIzol method (Thermo Fischer Scientific). Each dissected sample included epidermis, dermis, and the underlying fibrous/connective tissues of the specified nose regions (Figure 1). Tissue samples were placed into tubes containing 1 ml of TRIzol with a ceramic bead (1.4mm) and homogenized using a FastPrep-24 Instrument (MP Biomedicals, CA, USA). RNA extraction followed the protocol of TRIzol RNA extraction from Thermo Fischer Scientific. A DNA removal step with DNase followed the extraction (invitrogen). The total RNAs were dissolved in 50 µl nuclease-free water and their concentrations were quantified through a Nanophotometer (IMPLEN GmbH, Munich, Germany). We measured the quality of RNAs with the R6K ScreenTape System using an Agilent 2200 TapeStation (Agilent Technologies) and RNA integrity numbers (RIN) above 7 were aimed at for all samples. To synthesize cDNA for qPCR analysis, we used 500 ng of the total RNA per sample and followed the manufacturer's protocol of the High Capacity cDNA Reverse Transcription kit (Applied Biosystems), and the resulted cDNAs were diluted 1:4 to be used for the qPCR reaction.

RNA-seq library preparation and gene expression quantification

To attain transcriptome data of the snout tissues, we conducted RNA-seq library preparation with 1000 ng of total RNA per tissue sample as input and following the protocol of the Standard TruSeq Stranded mRNA Sample Prep Kit (Illumina) with indexing adapters. The library qualities were assessed using D1000 ScreenTape and reagents (Agilent) on a

TapeStation 2200 machine (Agilent). In order to reach an optimal quantity recommended for sequencing, we diluted the libraries and pooled them with equal molar concentration for each library. The RNA-sequencing was conducted in the NGS Facility at Vienna Biocenter Core Facilities (VBCF, Austria) on an Illumina HiSeq2500 and generated between 6.7 and 15.8 million paired-end reads with 125bp length per sample (Supplementary File S1). Raw reads were de-multiplexed based on unique barcodes by the same facility. The quality of the reads was assessed with FastQC (v0.11.8) (Andrews 2012), and reads were filtered for a quality > 28 and a minimum length of 70 bp with Trimmomatic (v0.3.9) (Bolger et al. 2014). Reads were aligned to the *O. niloticus* reference genome (Conte et al. 2017) of the University of Maryland using RNAstar (v2.7.3.a) (Dobin et al. 2013). To check the mapping statistics, we used samtools idxstats (v1.9) (Danecek et al. 2021) and further merged the single files for species and Lake with picard (v2.21.7) (Picard Toolkit. 2019. Broad Institute, GitHub Repository. https://broadinstitute.github.io/picard/). We used StringTie (v.2.0.6) (Pertea et al. 2015) to assemble the alignments into potential transcripts without a reference. This step was conducted separately for single files (per biological replicate) and the merged files (per species and per Lake). The single files per biological replicate were further merged into species. This process of repeated merging steps was implemented to reduce the probability of false positives. To assess the accuracy of the mapping we used gffcompare (v0.11.2) (Pertea & Pertea 2020) to compare our annotations to the reference annotation. Subsequently we filtered for monoexonic transcripts that were not contained in our reference and the transcripts assigned the class code 'possible polymerase run-on' by gffcompare. As the maximum intron length of the *O. niloticus* reference is 200000 bp, we also filtered for that in the produced annotation. The expression estimates for each transcript were based on these annotations and generated with StringTie (v.2.0.6) with no multimapping allowed. The final raw count matrices were produced from the expression estimates with a Perl script from the

griffith lab (https://github.com/griffithlab/rnaseq_tutorial/blob/master/scripts/stringtie_expression_matrix.pl) and the code used in this analysis is available at this github repository (https://github.com/annaduenser/snout_flap_RNAseq).

Differential expression analysis was conducted using DESeq2 (Love et al. 2014) in R (R Core Team 2017) running comparisons for each Lake separately. DESeq2 estimates variance-mean dependence based on a model using negative binomial distribution using the raw counts (Love et al. 2014). A false discovery rate of $p < 0.05$ was chosen as the cutoff. For the downstream analysis, an enrichment step for gene ontology (GO) terms of biological processes was conducted using the intersection list of DE genes between the Lakes and Manteia, a free online tool for data mining of vertebrate genes (Tassy & Pourquié 2014). In addition, we investigated the functional interactions between the products of DE genes through STRING v10 (http://string-db.org/), a knowledge based interaction prediction tool, and zebrafish databases for protein interactomes (Szklarczyk et al. 2017).

Primer design and qPCR

We designed the qPCR primers on conserved regions of the selected candidate genes by aligning their assembled sequences to their already available homologous mRNA sequences from *Ophthalmotilapia ventralis* (Böhne et al. 2014), *Metriaclima zebra*, *Pundamilia nyererei*, *Neolamprologus brichardi* and *Astatotilapia burtoni* (Brawand et al. 2014), as well as *Oreochromis niloticus*. After aligning the conserved sequence regions across the abovementioned East African cichlids, we identified the exon junctions (using CLC Genomic Workbench, CLC Bio, Denmark, and annotated genome of *Astatotilapia burtoni* in the Ensembl database, http://www.ensembl.org). The primer designing steps were conducted as

494 described previously (Pashay Ahi & Sefc 2018; Ahi et al. 2019) using Primer Express 3.0
495 (Applied Biosystems, CA, USA) (Supplementary file S3). The qPCR was performed based
496 on the protocol provided by Maxima SYBR Green/ROX qPCR Master Mix (2X) (Thermo
497 Fisher Scientific, Germany) following the guidelines for optimal experimental set-up for each
498 qPCR run (Hellemans et al. 2007). The qPCR program was set for 2 min at 50°C, 10 min at
499 95°C, 40 cycles of 15 sec at 95°C and 1 min at 60°C, followed by an additional step of
500 dissociation at 60°C – 95°C. The primer efficiency (E values) for each gene was calculated
501 through standard curves generated by serial dilutions of pooled cDNA samples. The standard
502 curves were run in triplicates and calculated using the following formula: $E = 10[-1/slope]$
503 (Supplementary file S3).
504 In order to select stably expressed candidate reference genes, we filtered for genes with a low
505 log2 fold change and subsequently ranked the remaining genes according to low coefficient
506 of variation. The top six most stable genes shared across the transcriptome comparisons were
507 selected as candidate reference genes. After qPCR expression analysis of the six genes across
508 all samples, we ranked them based on their expression stability by three different algorithms:
509 BestKeeper (Pfaffl et al. 2004), NormFinder (Andersen et al. 2004) and geNorm
510 (Vandesompele et al. 2002). We used the Cq values of the top most stable reference genes to
511 normalize Cq values of target genes in each sample (ΔCq target = Cq target – Cq reference).
512 The relative expression levels (RQ) were calculated by $2^{-\Delta\Delta Cq}$ method (Pfaffl 2001) and the
513 log-transformed RQ values were used for t-tests to calculate the statistical differences.
514

515 ## Acknowledgements

516 The authors thank Holger Zimmermann and Stephan Koblmüller for sharing their precious
517 knowledge on cichlid fishes of Lake Tanganyika, Sylvia Schäffer for sharing her experience
518 on RNA-seq library preparation, and Martin Grube and his lab for technical assistance and

519 access to their real-time PCR System. The authors acknowledge the financial support by the
520 University of Graz.

521

522 Author Contributions

523 EPA, CA and CS conceived the project. WG contributed to fish husbandry and photography,
524 and EPA and AD conducted the sampling and tissue dissection. AD, EPA and LL conducted
525 the RNA lab work. AD, PS, LL, EPA contributed to the analyses and all authors to
526 manuscript writing. CS and EPA contributed to funding. This work was supported by the
527 Austrian Science Fund [project number P29838] awarded to CS. All authors approved the
528 final version of the manuscript.

529

530 Competing financial interests

531 The authors declare no competing interests.

532

533 Ethical approval

534 No experiments were conducted on the fish prior to sampling, so an ethics approval is not
535 required according to the Austrian animal welfare law. Fish keeping and sacrifice was carried
536 out in our certified aquarium facility in accordance with the Austrian animal welfare law.

537

538 Data availability

539 The data underlying this article are available in the Sequencing Read Archive (SRA) of NCBI
540 at https://www.ncbi.nlm.nih.gov/ and can be accessed with PRJNA770252.

541

542

543

544

References

Ahi E. 2016. Signalling pathways in trophic skeletal development and morphogenesis: Insights from studies on teleost fish. Dev. Biol. 420:11–31. doi: http://dx.doi.org/10.1016/j.ydbio.2016.10.003.

Ahi Ehsan Pashay et al. 2020. Expression levels of the tetratricopeptide repeat protein gene ttc39b covary with carotenoid-based skin colour in cichlid fish . Biol. Lett. doi: 10.1098/rsbl.2020.0629.

Ahi Ehsan P., Duenser A, Singh P, Gessl W, Sturmbauer C. 2020. Appetite regulating genes may contribute to herbivory versus carnivory trophic divergence in haplochromine cichlids. PeerJ. 8:e8375. doi: 10.7717/peerj.8375.

Ahi EP, Richter F, Lecaudey LA, Sefc KM. 2019. Gene expression profiling suggests differences in molecular mechanisms of fin elongation between cichlid species. Sci. Rep. 9. doi: 10.1038/s41598-019-45599-w.

Ahi EP, Richter F, Sefc KM. 2017. A gene expression study of ornamental fin shape in Neolamprologus brichardi, an African cichlid species. Sci. Rep. doi: 10.1038/s41598-017-17778-0.

Alazami AM et al. 2016. Expanding the clinical and genetic heterogeneity of hereditary disorders of connective tissue. Hum. Genet. 135:525–540. doi: 10.1007/s00439-016-1660-z.

Albertson RC, Streelman JT, Kocher TD, Yelick PC. 2005. Integration and evolution of the cichlid mandible: the molecular basis of alternate feeding strategies. Proc. Natl. Acad. Sci. U. S. A. 102:16287–92. doi: 10.1073/pnas.0506649102.

Alhamoudi KM et al. 2020. A Missense Mutation in the UGDH Gene Is Associated With Developmental Delay and Axial Hypotonia. Front. Pediatr. 8. doi: 10.3389/fped.2020.00071.

Alwadei AH et al. 2016. Loss-of-function mutation in *RUSC2* causes intellectual disability and secondary microcephaly. Dev. Med. Child Neurol. 58:1317–1322. doi: 10.1111/dmcn.13250.

Andersen CL, Jensen JL, Ørntoft TF. 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res. 64:5245–50. doi: 10.1158/0008-5472.CAN-04-0496.

Andrews S. 2012. FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Arnold CP, Benham-Pyle BW, Lange JJ, Wood CJ, Sánchez Alvarado A. 2019. Wnt and TGFβ coordinate growth and patterning to regulate size-dependent behaviour. Nature. 572:655–659. doi: 10.1038/s41586-019-1478-7.

Attisano L, Labbé E. 2004. TGFβ and Wnt pathway cross-talk. Cancer Metastasis Rev. 23:53–61. doi: 10.1023/A:1025811012690.

Bailey TL et al. 2009. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 37:W202-8. doi: 10.1093/nar/gkp335.

Barrionuevo MG, Aybar MJ, Tríbulo C. 2014. Two different vestigial like 4 genes are differentially expressed during Xenopus laevis development. Int. J. Dev. Biol. 58:369–377. doi: 10.1387/ijdb.130353ct.

587 Baumgarten L, Machado-Schiaffino G, Henning F, Meyer A. 2015. What big lips are good
588 for: On the adaptive function of repeatedly evolved hypertrophied lips of cichlid fishes. Biol.
589 J. Linn. Soc. 115:448–455. doi: 10.1111/bij.12502.

590 Baumgartner MR, Valle D, Dionisi-Vici C. 2016. Disorders of Ornithine and Proline
591 Metabolism. In: Inborn Metabolic Diseases. Springer Berlin Heidelberg pp. 321–331. doi:
592 10.1007/978-3-662-49771-5_21.

593 Böhne A, Sengstag T, Salzburger W. 2014. Comparative transcriptomics in East African
594 cichlids reveals sex- and species-specific expression and new candidates for sex
595 differentiation in fishes. Genome Biol. Evol. 6:2567–2585. doi: 10.1093/gbe/evu200.

596 Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina
597 sequence data. Bioinformatics. 30:2114–2120. doi: 10.1093/bioinformatics/btu170.

598 Bonaglia MC et al. 2005. A 2.3 Mb duplication of chromosome 8q24.3 associated with
599 severe mental retardation and epilepsy detected by standard karyotype. Eur. J. Hum. Genet.
600 13:586–591. doi: 10.1038/sj.ejhg.5201369.

601 Brawand D et al. 2014. The genomic substrate for adaptive radiation in African cichlid fish.
602 Nature. 513:375–381. doi: 10.1038/nature13726.

603 Brugmann SA et al. 2007. Wnt signaling mediates regional specification in the vertebrate
604 face. Development. 134:3283–95. doi: 10.1242/dev.005132.

605 Brugmann SA, Cordero DR, Helms JA. 2010. Craniofacial ciliopathies: A new classification
606 for craniofacial disorders. Am. J. Med. Genet. Part A. 152A:2995–3006. doi:
607 10.1002/ajmg.a.33727.

608 Chen CP, Ko TM, Chen YY, Su JW, Wang W. 2013. Prenatal diagnosis and molecular
609 cytogenetic characterization of mosaicism for a small supernumerary marker chromosome
610 derived from chromosome 22 associated with cat eye syndrome. Gene. 527:384–388. doi:
611 10.1016/j.gene.2013.05.061.

612 Chesnutt C, Burrus LW, Brown AMC, Niswander L. 2004. Coordinate regulation of neural
613 tube patterning and proliferation by TGFβ and WNT activity. Dev. Biol. 274:334–347. doi:
614 10.1016/j.ydbio.2004.07.019.

615 Colombo M et al. 2013. The ecological and genetic basis of convergent thick-lipped
616 phenotypes in cichlid fishes. Mol. Ecol. 22:670–684. doi: 10.1111/mec.12029.

617 Concannon MR, Albertson RC. 2015. The genetic and developmental basis of an exaggerated
618 craniofacial trait in East African cichlids. J. Exp. Zool. B. Mol. Dev. Evol. 324:662–70. doi:
619 10.1002/jez.b.22641.

620 Conith MR et al. 2018. Genetic and developmental origins of a unique foraging adaptation in
621 a Lake Malawi cichlid genus. Proc. Natl. Acad. Sci. 115:7063–7068. doi:
622 10.1073/PNAS.1719798115.

623 Conith MR, Conith AJ, Albertson RC. 2019. Evolution of a soft-tissue foraging adaptation in
624 African cichlids: Roles for novelty, convergence, and constraint. Evolution (N. Y). 73:2072–
625 2084. doi: 10.1111/evo.13824.

626 Conte MA, Gammerdinger WJ, Bartie KL, Penman DJ, Kocher TD. 2017. A high quality
627 assembly of the Nile Tilapia (Oreochromis niloticus) genome reveals the structure of two sex
628 determination regions. BMC Genomics. 18:341. doi: 10.1186/s12864-017-3723-5.

629 Czeschik JC et al. 2014. A patient with a de-novo deletion 3p25.3 and features overlapping
630 with Rubinstein–Taybi syndrome. Clin. Dysmorphol. 23:67–70. doi:
631 10.1097/MCD.0000000000000035.

632 Danecek P et al. 2021. Twelve years of SAMtools and BCFtools. Gigascience. 10. doi:
633 10.1093/gigascience/giab008.

634 Dell'Edera D et al. 2018. 16p11.2 microdeletion syndrome: A case report. J. Med. Case Rep.
635 12:1–6. doi: 10.1186/s13256-018-1587-1.

636 Dines JN et al. 2019. Expanding phenotype with severe midline brain anomalies and
637 missense variant supports a causal role for *FOXA2* in 20p11.2 deletion syndrome. Am. J.
638 Med. Genet. Part A. 179:ajmg.a.61281. doi: 10.1002/ajmg.a.61281.

639 Dobin A et al. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 29:15–21.
640 doi: 10.1093/bioinformatics/bts635.

641 Ehret JK et al. 2015. Microdeletions in 9q33.3-q34.11 in five patients with intellectual
642 disability, microcephaly, and seizures of incomplete penetrance: Is STXBP1 not the only
643 causative gene? Mol. Cytogenet. 8:1–14. doi: 10.1186/s13039-015-0178-8.

644 Fryssira H et al. 2011. Severe Developmental Delay in a Patient with 7p21.1–p14.3
645 Microdeletion Spanning the <i>TWIST</i> Gene and the
646 <i>HOXA</i> Gene Cluster. Mol. Syndromol. 2:45–49. doi: 10.1159/000334313.

647 Gervasini C et al. 2007. High frequency of mosaic CREBBP deletions in Rubinstein-Taybi
648 syndrome patients and mapping of somatic and germ-line breakpoints. Genomics. 90:567–
649 573. doi: 10.1016/j.ygeno.2007.07.012.

650 Giannikou K et al. 2012. Further delineation of novel 1p36 rearrangements by array-CGH
651 analysis: Narrowing the breakpoints and clarifying the 'extended' phenotype. Gene. 506:360–
652 368. doi: 10.1016/j.gene.2012.06.060.

653 Guilmatre A et al. 2010. Type I hyperprolinemia: genotype/phenotype correlations. Hum.
654 Mutat. 31:961–965. doi: 10.1002/humu.21296.

655 Gunter HM et al. 2013. Shaping development through mechanical strain: the transcriptional
656 basis of diet-induced phenotypic plasticity in a cichlid fish. Mol. Ecol. 22:4516–4531. doi:
657 10.1111/mec.12417.

658 Haldeman-Englert CR et al. 2009. A 3.1-Mb microdeletion of 3p21.31 associated with
659 cortical blindness, cleft lip, CNS abnormalities, and developmental delay. Eur. J. Med. Genet.
660 52:265–268. doi: 10.1016/j.ejmg.2008.11.005.

661 Hanssens M, Snoeks J, Verheyen E. 1999. A morphometric revision of the genus
662 Ophthalmotilapia (Teleostei, Cichlidae) from Lake Tanganyika (East Africa). Zool. J. Linn.
663 Soc. 125:487–512. doi: 10.1111/j.1096-3642.1999.tb00602.x.

664 Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J. 2007. qBase relative
665 quantification framework and software for management and automated analysis of real-time
666 quantitative PCR data. Genome Biol. 8:R19. doi: 10.1186/gb-2007-8-2-r19.

667 Henning F, Machado-Schiaffino G, Baumgarten L, Meyer A. 2017. Genetic dissection of
668 adaptive form and function in rapidly speciating cichlid fishes. Evolution (N. Y). 71:1297–
669 1312. doi: 10.1111/evo.13206.

670 Hill CR et al. 2012. BMP2 signals loss of epithelial character in epicardial cells but requires

671 the Type III TGFβ receptor to promote invasion. Cell. Signal. 24:1012–1022. doi:
672 10.1016/j.cellsig.2011.12.022.

673 Holder JL, Lotze TE, Bacino C, Cheung S-W. 2012. A child with an inherited 0.31 Mb
674 microdeletion of chromosome 14q32.33: Further delineation of a critical region for the 14q32
675 deletion syndrome. Am. J. Med. Genet. Part A. 158A:1962–1966. doi:
676 10.1002/ajmg.a.35289.

677 Hosono K et al. 2020. A case of childhood glaucoma with a combined partial monosomy
678 6p25 and partial trisomy 18p11 due to an unbalanced translocation. Ophthalmic Genet.
679 41:175–182. doi: 10.1080/13816810.2020.1744019.

680 Hufnagel RB et al. 2016. A new frontonasal dysplasia syndrome associated with deletion of
681 the *SIX2* gene. Am. J. Med. Genet. Part A. 170:487–491. doi: 10.1002/ajmg.a.37441.

682 Hulsey CD, Fraser GJ, Meyer A. 2016. Biting into the genome to phenome map:
683 Developmental genetic modularity of cichlid fish dentitions. In: Integrative and Comparative
684 Biology.Vol. 56 Oxford University Press pp. 373–388. doi: 10.1093/icb/icw059.

685 Hüning I et al. 2013. Exon 2 duplication of the MID1 gene in a patient with a mild phenotype
686 of Opitz G/BBB syndrome. Eur. J. Med. Genet. 56:188–191. doi:
687 10.1016/j.ejmg.2013.01.004.

688 Irisarri I et al. 2018. Phylogenomics uncovers early hybridization and adaptive loci shaping
689 the radiation of Lake Tanganyika cichlid fishes. Nat. Commun. 9:1–12. doi: 10.1038/s41467-
690 018-05479-9.

691 Jain S, Yang P, Farrell SA. 2010. A case of 8q22.1 microdeletion without the Nablus mask-
692 like facial syndrome phenotype. Eur. J. Med. Genet. 53:108–110. doi:
693 10.1016/j.ejmg.2009.12.006.

694 Javitt G et al. 2019. *cis* -Proline mutants of quiescin sulfhydryl oxidase 1 with altered redox
695 properties undermine extracellular matrix integrity and cell adhesion in fibroblast cultures.
696 Protein Sci. 28:228–238. doi: 10.1002/pro.3537.

697 De Jong DS et al. 2004. Regulation of Notch signaling genes during BMP2-induced
698 differentiation of osteoblast precursor cells. Biochem. Biophys. Res. Commun. 320:100–107.
699 doi: 10.1016/j.bbrc.2004.05.150.

700 Kariminejad A et al. 2015. Intellectual disability, muscle weakness and characteristic face in
701 three siblings: A newly described recessive syndrome mapping to 3p24.3-p25.3. Am. J. Med.
702 Genet. Part A. 167:2508–2515. doi: 10.1002/ajmg.a.37248.

703 Karna E, Szoka L, Huynh TYL, Palka JA. 2020. Proline-dependent regulation of collagen
704 metabolism. Cell. Mol. Life Sci. 77:1911–1918. doi: 10.1007/s00018-019-03363-3.

705 Kiratli H, Satilmiş M. 1998. Prolidase deficiency associated with pathologic myopia.
706 Ophthalmic Genet. 19:49–53. doi: 10.1076/opge.19.1.49.2180.

707 Klüppel M, Wrana JL. 2005. Turning it up a Notch: cross-talk between TGFβ and Notch
708 signaling. BioEssays. 27:115–118. doi: 10.1002/bies.20187.

709 Kocher TD, Conroy JA, McKaye KR, Stauffer JR. 1993. Similar morphologies of cichlid fish
710 in lakes tanganyika and malawi are due to convergence. Mol. Phylogenet. Evol. 2:158–165.
711 doi: 10.1006/mpev.1993.1016.

712 Kondratyeva LG et al. 2016. Downregulation of expression of mater genes SOX9, FOXA2,

713 and GATA4 in pancreatic cancer cells stimulated with TGFβ1 epithelial–mesenchymal
714 transition. Dokl. Biochem. Biophys. 469:257–259. doi: 10.1134/S1607672916040062.

715 Konings A. 2007. *Malawi cichlids in their natural habitat 4th Edition*. 4th ed. Cichlid Press.
716 El Paso, Texas, USA.: El Paso, Texas, USA.

717 Kosho T et al. 2008. De-novo balanced translocation between 7q31 and 10p14 in a girl with
718 central precocious puberty, moderate mental retardation, and severe speech impairment. Clin.
719 Dysmorphol. 17:31–34. doi: 10.1097/MCD.0b013e3282f17688.

720 Kretz R et al. 2011. Defect in proline synthesis: Pyrroline-5-carboxylate reductase 1
721 deficiency leads to a complex clinical phenotype with collagen and elastin abnormalities. J.
722 Inherit. Metab. Dis. 34:731–739. doi: 10.1007/s10545-011-9319-3.

723 Kubista M et al. 2006. The real-time polymerase chain reaction. Mol. Aspects Med. 27:95–
724 125. doi: 10.1016/j.mam.2005.12.007.

725 Kucharczyk M et al. 2014. The first case of a patient with de novo partial distal 16q
726 tetrasomy and a data's review. Am. J. Med. Genet. Part A. 164:2541–2550. doi:
727 10.1002/ajmg.a.36686.

728 Lecaudey LA et al. 2021. Transcriptomics unravels molecular players shaping dorsal lip
729 hypertrophy in the vacuum cleaner cichlid, Gnathochromis permaxillaris. BMC Genomics.
730 22:506. doi: 10.1186/s12864-021-07775-z.

731 Lecaudey LA, Sturmbauer C, Singh P, Ahi EP. 2019. Molecular mechanisms underlying
732 nuchal hump formation in dolphin cichlid, Cyrtocara moorii. Sci. Rep. doi: 10.1038/s41598-
733 019-56771-7.

734 Li Y, Hibbs MA, Gard AL, Shylo NA, Yun K. 2012. Genome-Wide Analysis of
735 N1ICD/RBPJ Targets In Vivo Reveals Direct Transcriptional Regulation of Wnt, SHH, and
736 Hippo Pathway Effectors by Notch1. Stem Cells. 30:741–752. doi: 10.1002/stem.1030.

737 Liu M et al. 2012. IKKα Activation of NOTCH Links Tumorigenesis via FOXA2
738 Suppression. Mol. Cell. 45:171–184. doi: 10.1016/j.molcel.2011.11.018.

739 Lopes F et al. 2019. Genomic imbalances defining novel intellectual disability associated
740 loci. Orphanet J. Rare Dis. 14:1–13. doi: 10.1186/s13023-019-1135-0.

741 Losos JB, Jackman TR, Larson A, De Queiroz K, Rodríguez-Schettino L. 1998. Contingency
742 and determinism in replicated adaptive radiations of island lizards. Science (80-.). 279:2115–
743 2118. doi: 10.1126/science.279.5359.2115.

744 Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for
745 RNA-seq data with DESeq2. Genome Biol. 15:550. doi: 10.1186/s13059-014-0550-8.

746 Machado-Schiaffino G, Henning F, Meyer A. 2014. Species-specific differences in adaptive
747 phenotypic plasticity in an ecologically relevant trophic trait: hypertrophic lips in Midas
748 cichlid fishes. Evolution (N. Y). 68:2086–2091. doi: 10.1111/evo.12367.

749 Manousaki T et al. 2013. Parsing parallel evolution: ecological divergence and differential
750 gene expression in the adaptive radiations of thick-lipped Midas cichlid fishes from
751 Nicaragua. Mol. Ecol. 22:650–69. doi: 10.1111/mec.12034.

752 Martinez-Glez V et al. 2007. Clinical presentation of a variant of Axenfeld-Rieger syndrome
753 associated with subtelomeric 6p deletion. Eur. J. Med. Genet. 50:120–127. doi:
754 10.1016/j.ejmg.2006.10.005.

McLachlan E et al. 2005. Functional Characterization of Oculodentodigital Dysplasia-Associated Cx43 Mutants. Cell Commun. Adhes. 12:279–292. doi: 10.1080/15419060500514143.

Miller ND et al. 2009. Molecular (SNP) analyses of overlapping hemizygous deletions of 10q25.3 to 10qter in four patients: Evidence for HMX2 and HMX3 as candidate genes in hearing and vestibular function. Am. J. Med. Genet. Part A. 149A:669–680. doi: 10.1002/ajmg.a.32705.

Mitteldorf C et al. 2018. Deceptively bland cutaneous angiosarcoma on the nose mimicking hemangioma-A clinicopathologic and immunohistochemical analysis. J. Cutan. Pathol. 45:652–658. doi: 10.1111/cup.13275.

Mordaunt D et al. 2015. 8q13.1-q13.2 deletion associated with inferior cerebellar vermian hypoplasia and digital anomalies: A new syndrome? Pediatr. Neurol. 52:230-234.e1. doi: 10.1016/j.pediatrneurol.2014.09.002.

Mucha BE et al. 2019. A new microdeletion syndrome involving TBC1D24, ATP6V0C, and PDPK1 causes epilepsy, microcephaly, and developmental delay. Genet. Med. 21:1058–1064. doi: 10.1038/s41436-018-0290-3.

Nakayama T et al. 2014. RBPJ is disrupted in a case of proximal 4p deletion syndrome with epilepsy. Brain Dev. 36:532–536. doi: 10.1016/j.braindev.2013.07.009.

Nimmagadda S et al. 2015. Identification and functional analysis of novel facial patterning genes in the duplicated beak chicken embryo. Dev. Biol. 407:275–288. doi: 10.1016/j.ydbio.2015.09.007.

Noguchi Y, Iwasaki Y, Ueda M, Kakinoki S. 2020. Surfaces immobilized with oligo-prolines prevent protein adsorption and cell adhesion. J. Mater. Chem. B. 8:2233–2237. doi: 10.1039/d0tb00051e.

Okello DO et al. 2017. Six2 Plays an Intrinsic Role in Regulating Proliferation of Mesenchymal Cells in the Developing Palate. Front. Physiol. 8:955. doi: 10.3389/fphys.2017.00955.

Pakvasa M et al. 2020. Notch signaling: Its essential roles in bone and craniofacial development. Genes Dis. doi: 10.1016/j.gendis.2020.04.006.

Pashay Ahi E, Sefc KM. 2018. Towards a gene regulatory network shaping the fins of the Princess cichlid. Sci. Rep. 8:9602. doi: 10.1038/s41598-018-27977-y.

De Pater JM et al. 2005. Striking Facial Dysmorphisms and Restricted Thymic Development in a Fetus with a 6-Megabase Deletion of Chromosome 14q. Pediatr. Dev. Pathol. 8:497–503. doi: 10.1007/s10024-005-0041-8.

Penton AL, Leonard LD, Spinner NB. 2012. Notch signaling in human development and disease. Semin. Cell Dev. Biol. 23:450–457. doi: 10.1016/j.semcdb.2012.01.010.

Pertea G, Pertea M. 2020. GFF Utilities: GffRead and GffCompare. F1000Research. 9. doi: 10.12688/f1000research.23297.2.

Pertea M et al. 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat. Biotechnol. 33:290–295. doi: 10.1038/nbt.3122.

Pfaffl MW. 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 29:e45.

797 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=55695&tool=pmcentrez&rendert
798 ype=abstract (Accessed July 31, 2014).

799 Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP. 2004. Determination of stable
800 housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--
801 Excel-based tool using pair-wise correlations. Biotechnol. Lett. 26:509–15.
802 http://www.ncbi.nlm.nih.gov/pubmed/15127793 (Accessed June 12, 2015).

803 Piard J et al. 2015. *TCF12* microdeletion in a 72-year-old woman with intellectual disability.
804 Am. J. Med. Genet. Part A. 167:1897–1901. doi: 10.1002/ajmg.a.37083.

805 Pillai NR et al. 2019. Novel deletion of 6p21.31p21.1 associated with laryngeal cleft,
806 developmental delay, dysmorphic features and vascular anomaly. Eur. J. Med. Genet.
807 62:103531. doi: 10.1016/j.ejmg.2018.08.012.

808 Powder KE, Albertson RC. 2016. Cichlid fishes as a model to understand normal and clinical
809 craniofacial variation. Dev. Biol. 415:338–346. doi: 10.1016/J.YDBIO.2015.12.018.

810 Preiksaitiene E et al. 2015. R368X mutation in MID1 among recurrent mutations in patients
811 with X-linked Opitz G/BBB syndrome. Clin. Dysmorphol. 24:7–12. doi:
812 10.1097/MCD.0000000000000059.

813 Quigley DI, Kaiser-Rogers K, Aylsworth AS, Rao KW. 2004. Submicroscopic deletion
814 9(q34.4) and duplication 19(p13.3): Identified by subtelomere specific FISH probes. Am. J.
815 Med. Genet. 125A:67–72. doi: 10.1002/ajmg.a.20457.

816 R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. doi:
817 10.1007/978-3-540-74686-7.

818 Reijnders MRF et al. 2017. RAC1 Missense Mutations in Developmental Disorders with
819 Diverse Phenotypes. Am. J. Hum. Genet. 101:466–477. doi: 10.1016/j.ajhg.2017.08.007.

820 Rivera-Pedroza CI et al. 2017. Chromosome 1p31.1p31.3 Deletion in a Patient with
821 Craniosynostosis, Central Nervous System and Renal Malformation: Case Report and
822 Review of the Literature. Mol. Syndromol. 8:30–35. doi: 10.1159/000452609.

823 Rüber L, Verheyen E, Meyer A. 1999. Replicated evolution of trophic specializations in an
824 endemic cichlid fish lineage from Lake Tanganyika. Proc. Natl. Acad. Sci. U. S. A.
825 96:10230–10235. doi: 10.1073/pnas.96.18.10230.

826 Rundle HD, Nagel L, Boughman JW, Schluter D. 2000. Natural selection and parallel
827 speciation in sympatric sticklebacks. Science (80-.). 287:306–308. doi:
828 10.1126/science.287.5451.306.

829 Sansone A, Syed AS, Tantalaki E, Korsching SI, Manzini I. 2014. Trpc2 is expressed in two
830 olfactory subsystems, the main and the vomeronasal system of larval Xenopus laevis. J. Exp.
831 Biol. 217:2235–2238. doi: 10.1242/jeb.103465.

832 Sawyer JR, Binz RL, Swanson CM, Lim C. 2007. De novo proximal duplication of
833 1(q12q22) in a female infant with multiple congenital anomalies. Am. J. Med. Genet. Part A.
834 143A:338–342. doi: 10.1002/ajmg.a.31604.

835 Schluter D, Nagel LM. 1995. Parallel speciation by natural selection. Am. Nat. 146:292–301.
836 doi: 10.1086/285799.

837 Schulz S, Volleth M, Muschke P, Wieland I, Wieacker P. 2008. Greig cephalopolysyndactyly
838 (GCPS) contiguous gene syndrome in a boy with a 14 Mb deletion in region 7p13-14 caused

by a paternal balanced insertion (5; 7). Appl. Clin. Genet. 1:19–22. doi: 10.2147/tacg.s4401.

Schwörer S et al. 2020. Proline biosynthesis is a vent for TGFβ-induced mitochondrial redox stress. EMBO J. 39. doi: 10.15252/embj.2019103334.

Selenti N et al. 2015. An interstitial deletion at 8q23.1-q24.12 associated with Langer-Giedion syndrome/Trichorhinophalangeal syndrome (TRPS) type II and Cornelia de Lange syndrome 4. Mol. Cytogenet. 8:64. doi: 10.1186/s13039-015-0169-9.

Sharma VP et al. 2013. Mutations in TCF12, encoding a basic helix-loop-helix partner of TWIST1, are a frequent cause of coronal craniosynostosis. Nat. Genet. 45:304–307. doi: 10.1038/ng.2531.

Sipila KH et al. 2018. Proline hydroxylation in collagen supports integrin binding by two distinct mechanisms. J. Biol. Chem. 293:7645–7658. doi: 10.1074/jbc.RA118.002200.

Sofos E et al. 2012. A novel familial 11p15.4 microduplication associated with intellectual disability, dysmorphic features, and obesity with involvement of the ZNF214 gene. Am. J. Med. Genet. Part A. 158A:50–58. doi: 10.1002/ajmg.a.34290.

Starkovich M, Lalani SR, Mercer CL, Scott DA. 2016. Chromosome 5q33 deletions associated with congenital heart defects. Am. J. Med. Genet. Part A. 170:3338–3342. doi: 10.1002/ajmg.a.37957.

Szklarczyk D et al. 2017. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. Nucleic Acids Res. 45:D362–D368. doi: 10.1093/nar/gkw937.

Tan TY et al. 2017. Monoallelic BMP2 Variants Predicted to Result in Haploinsufficiency Cause Craniofacial, Skeletal, and Cardiac Features Overlapping Those of 20p12 Deletions. Am. J. Hum. Genet. 101:985–994. doi: 10.1016/j.ajhg.2017.10.006.

Tanigaki K et al. 2002. Notch-RBP-J signaling is involved in cell fate determination of marginal zone B cells. Nat. Immunol. 3:443–450. doi: 10.1038/ni793.

Tassy O, Pourquié O. 2014. Manteia, a predictive data mining system for vertebrate genes and its applications to human genetic diseases. Nucleic Acids Res. 42:D882-91. doi: 10.1093/nar/gkt807.

Thomas PS, Kim J, Nunez S, Glogauer M, Kaartinen V. 2010. Neural crest cell-specific deletion of Rac1 results in defective cell-matrix interactions and severe craniofacial and cardiovascular malformations. Dev. Biol. 340:613–625. doi: 10.1016/j.ydbio.2010.02.021.

Tinker RJ et al. 2021. Haploinsufficiency of <scp> *ATP6V0C* </scp> possibly underlies 16p13.3 deletions that cause microcephaly, seizures, and neurodevelopmental disorder. Am. J. Med. Genet. Part A. 185:196–202. doi: 10.1002/ajmg.a.61905.

Tomkins DJ, Gitelman BJ, Roberts MH. 1983. Confirmation of a de novo duplication, dup(10) (q24→q26), by GOT1 gene dosage studies. Hum. Genet. 63:369–373. doi: 10.1007/BF00274764.

Ţuţulan-Cuniţă AC et al. 2012. 3p interstitial deletion: Novel case report and review. J. Child Neurol. 27:1062–1066. doi: 10.1177/0883073811431016.

Valdez JM et al. 2012. Notch and TGFβ form a reciprocal positive regulatory loop that suppresses murine prostate basal stem/progenitor cell activity. Cell Stem Cell. 11:676–688. doi: 10.1016/j.stem.2012.07.003.

881 Vandesompele J et al. 2002. Accurate normalization of real-time quantitative RT-PCR data
882 by geometric averaging of multiple internal control genes. Genome Biol.
883 3:RESEARCH0034.
884 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=126239&tool=pmcentrez&render
885 type=abstract (Accessed July 27, 2014).

886 Velez DO et al. 2019. 3D collagen architecture regulates cell adhesion through degradability,
887 thereby controlling metabolic and oxidative stress. Integr. Biol. 11:221–234. doi:
888 10.1093/intbio/zyz019.

889 Wang Y, Song L, Zhou CJ. 2011. The canonical Wnt/ß-catenin signaling pathway regulates
890 Fgf signaling for early facial development. Dev. Biol. 349:250–260. doi:
891 10.1016/j.ydbio.2010.11.004.

892 Wieacker P, Volleth M. 2007. <i>WNT4</i> and <i>RSPO1</i> Are
893 Not Involved in a Case of Male-to-Female Sex Reversal with Partial Duplication of 1p. Sex.
894 Dev. 1:111–113. doi: 10.1159/000100032.

895 Wong MM-K, Joyson SM, Hermeking H, Chiu SK. 2021. Transcription Factor AP4 Mediates
896 Cell Fate Decisions: To Divide, Age, or Die. Cancers (Basel). 13:676. doi:
897 10.3390/cancers13040676.

898 Xinjie L et al. 2001. Evaluation of the role of proline residues flanking the RGD motif of
899 dendroaspin, an inhibitior of platelet aggregation and cell adhesion. Biochem. J. 355:633–
900 638. doi: 10.1042/bj3550633.

901 Yakut S et al. 2015. A familial interstitial 4q35 deletion with no discernible clinical effects.
902 Am. J. Med. Genet. Part A. 167:1836–1841. doi: 10.1002/ajmg.a.37097.

903 Zaki MS et al. 2016. *PYCR2* Mutations cause a lethal syndrome of microcephaly and failure
904 to thrive. Ann. Neurol. 80:59–70. doi: 10.1002/ana.24678.

905 Zhang P, Yang C, Delay RJ. 2010. Odors activate dual pathways, a TRPC2 and a AA-
906 dependent pathway, in mouse vomeronasal neurons. Am. J. Physiol. Physiol. 298:C1253–
907 C1264. doi: 10.1152/ajpcell.00271.2009.

908 Zlotina A et al. 2016. Ring chromosome 18 in combination with 18q12.1 (DTNA) interstitial
909 microdeletion in a patient with multiple congenital defects. Mol. Cytogenet. 9:1–7. doi:
910 10.1186/s13039-016-0229-9.

911