Pronounced expression of extracellular matrix proteoglycans regulated by Wnt pathway underlies the parallel evolution of lip hypertrophy in East African cichlids

Nagatoshi Machii[1], Ryo Hatashima[1], Tatsuya Niwa[1,2], Hideki Taguchi[1,2], Ismael A. Kimirei[3], Hillary D. J. Mrosso[4], Mitsuto Aibara[1], Tatsuki Nagasawa[1], Masato Nikaido[1,*]

[1]School of Life Science and Technology, Tokyo Institute of Technology, Tokyo, Japan

[2]Cell Biology Center, Institute of Innovative Research, Tokyo Institute of Technology, Yokohama, Japan

[3]Tanzania Fisheries Research Institute (TAFIRI), Dar es Salaam, Tanzania

[4]Tanzania Fisheries Research Institute (TAFIRI), Mwanza Fisheries Research Center, Mwanza, Tanzania

*Corresponding author: Masato Nikaido

Email: mnikaido@bio.titech.ac.jp

Tel: +81-3-5734-2659, Fax: +81-3-5734-02946

Abstract

Cichlid fishes inhabiting the East African Great Lakes, Victoria, Malawi, and Tanganyika, are textbook examples of parallel evolution, as they have acquired similar traits independently in each of the three lakes during the process of adaptive radiation. In particular, “hypertrophied lip” has been highlighted as a prominent example of parallel evolution. However, the underlying molecular mechanisms remain poorly understood. In this study, we conducted an integrated comparative analysis between the hypertrophied and normal lips of cichlids across three lakes based on histology, proteomics, and transcriptomics. Histological and proteomic analyses revealed that the hypertrophied lips were characterized by enlargement of the proteoglycan-rich layer, in which versican and periostin proteins were abundant. Transcriptome analysis revealed that the expression of extracellular matrix-related genes, including collagens, glycoproteins and proteoglycans, was higher in hypertrophied lips, regardless of their phylogenetic relationships. In addition, the Wnt signaling pathway, which is involved in promoting proteoglycan expression, was highly expressed in both the juvenile and adult stages of hypertrophied lips. Our comprehensive analyses showed that hypertrophied lips of the three different phylogenetic origins can be explained by similar proteomic and transcriptomic profiles, which may provide important clues into the molecular mechanisms underlying phenotypic parallelisms in East African cichlids.

Introduction

Parallel evolution is evidence of natural selection that drives adaptive solutions to the environment (Losos, 2011). Acquisition of similar traits in distantly related lineages can be observed in multiple lineages, such as the loss of eyes in cave fish, patterns of wing colors in Heliconius butterflies, and body armor of sticklebacks (Krishnan and Rohner, 2017; Stern, 2013). It is considered that the same direction of natural selection drives phenotypic changes among species since it is unlikely that these complex phenotypes have been acquired repeatedly just by neutral evolution (Losos, 2011; Stern, 2013). Therefore, research on parallel evolution enables us to compare the commonalities and differences in the mechanisms behind the acquisition of novel traits and elucidate how adaptation occurs in similar environments.

East African cichlids offer one of the best opportunities to examine the molecular basis of parallel evolution (Kocher et al., 1993; Muschick et al., 2012). There are hundreds of endemic cichlid species in the East African Great Lakes, Victoria, Malawi, and Tanganyika, as a result of explosive adaptive radiation (Salzburger, 2018; Santos et al., 2023). As a result of independent adaptations in each lake, parallel evolution is often observed among lakes, such as that of the nose snout (Duenser et al., 2023), jaw morphology, and body shape (Kocher et al., 1993; Muschick et al., 2012). To reveal their evolutionary mechanisms, genetic bases of the phenotypes have been investigated. For example, the parallel evolution of body stripes in cichlids of Lake Victoria and Lake Malawi was explained by relatively simple changes in pigmental cell arrangement caused by mutations in the cis-regulatory element of agouti-related peptide 2 (agrp2) (Kratochwil et al., 2018). Although this parallel evolution of body stripe patterns can be explained by a single gene, most traits are polygenic, hindering elucidation of the parallel evolution mechanism.

Cichlid fishes have acquired a huge diversity in their oral morphologies, reflecting ecological and trophic specializations (Albertson and Kocher, 2006; Kocher, 2004; Lecaudey et al., 2021, 2019). Hypertrophied lips, one of the most prominent examples of oral diversity, have been independently acquired in cichlids inhabiting each of the East African Great Lakes, as well as in the Nicaraguan crater lakes and South American rivers (Henning et al., 2017; Kautt et al., 2020; Masonick et al., 2023, 2022). Several hypotheses have been proposed for the adaptive roles of hypertrophied lips, such as for effectively sucking prey out from narrow crevices (Lukas et al., 2015), reducing mechanical stress during foraging in rocky habitats (Fryer, 1972; Greenwood, 1974), and facilitating prey detection by providing an enlarged area for taste receptors (Arnegard and Snoeks, 2001; Oliver and Arnegard, 2010). Thus, hypertrophied lips are considered an adaptive trait for a specific trophic niche, making them an ideal model for understanding the mechanism of parallel evolution.

The molecular mechanisms underlying hypertrophied lips have been investigated in various lineages of cichlids. For example, transcriptome comparisons between hypertrophied and normal lips of cichlids in Lake Tanganyika (Colombo et al., 2013) and South American crater lakes (Manousaki et al., 2013) have identified some differentially expressed genes (DEGs) as candidates for lip hypertrophy; however, their primary regulators remain to be elucidated. In addition, quantitative trait locus (QTL) mapping (Henning et al., 2017) and genome-wide association study (GWAS) (Masonick et al., 2023) have been conducted to elucidate the genomic loci responsible for hypertrophied lips in the cichlids of Lake Victoria and Lake Malawi. These genome-wide studies concluded that lip hypertrophy was a polygenic trait involving several candidate genes. However, the molecular mechanism underlying lip hypertrophy has not yet been elucidated.

87 In this study, we comprehensively compared the hypertrophied lips of cichlids
88 across all East African Great Lakes using histology, proteomics, and transcriptomics.
89 Histological and proteomic analyses revealed a distinct microstructure of hypertrophied
90 lips compared to normal lips, and primary candidate proteins were identified.
91 Transcriptome analysis at different developmental stages showed that the Wnt signaling
92 pathway was highly expressed in cichlids with hypertrophied lips at both the juvenile and
93 adult stages. It is noteworthy that the distinct expression profiles observed in the
94 proteome and transcriptome analyses of hypertrophied lips were similar among cichlids
95 from each of the East African Great Lakes. The present study, which integrates
96 comprehensive analyses for cichlids from all East African Great Lakes, provides insight
97 for a better understanding of the molecular basis of a typical example of parallel
98 evolution.
99

Materials and Methods

Aminals

Three species of Victoria cichlids (*Haplochromis chilotes* with hypertrophied lips; *H. sauvagei,* and *H. pyrrhocephalus* with normal lips , four species of Malawi cichlids (*Placidochromis milomo* with hypertrophied lips; *Maylandia lombardoi, M. zebra,* and *Labidochromis caeruleus* with normal lips), and two species of Taganyika cichlids (*Lobochilotes labiatus* with hypertrophied lips; *Tropheus moorii* with normal lips) were used in the experiments. *H. chilotes, H. sauvagei*, and *H. pyrrhocephalus* were collected from the Mwanza gulf of Lake Victoria during an exploration that was led by the Nikaido Laboratory in 2018. The other fish were obtained from a commercial supplier. Fish were maintained and bred at 27°C on a 12/12 h light/dark cycle. One to 12 individuals were kept in a plastic tank (40 cm × 25 cm × 36 cm) and fed food pellets 2-3 times a day. The sample information is provided in Table S1. All experiments were conducted in accordance with the Institutional Animal Experiment Committee of the Tokyo Institute of Technology.

Histological observation

Cichlids were anesthetized with 0.02% ethyl p-aminobenzoate. The lower lips were then dissected, along with the lower jaw, and were fixed in 4% paraformaldehyde (PFA, Fujifilm-Wako, Chuo-ku, Osaka, Japan) in phosphate-buffered saline (PBS) at 4°C for 24 h. Tissues were demineralized in 0.34 M ethylenediaminetetraacetic acid (EDTA, Fujifilm-Wako, Chuo-ku, Osaka, Japan) in 100 mM, pH7.0 tris (hydroxymethyl) aminomethane (Tris, Fujifilm-Wako, Chuo-ku, Osaka, Japan) buffer at 4°C for 24 h. After demineralization, tissues were treated with 20% sucrose/PBS at 4°C for 24 h. Tissues were then embedded in Tissue-Tek O.C.T. compound (Sakura Finetek, Chuo-ku, Tokyo, Japan) and frozen with liquid nitrogen. Embedded blocks were sliced in the sagittal plane into 20 μm sections and placed on a coated glass slide (MAS-01, Matsunami, Bunkyo-

125 ku, Tokyo, Japan). Sections were stored at −80°C until they were used for experiments.
126 To observe the histological structure, hematoxylin–eosin (HE) staining, Alcian blue (AB)
127 staining, and Van Gieson (VG) staining were conducted. Mayer's hematoxylin solution
128 (Fujifilm-Wako, Chuo-ku, Osaka, Japan), Eosin Y solution (Fujifilm-Wako, Chuo-ku,
129 Osaka, Japan), ALCIAN BLUE 8GX (Sigma-Aldrich, St. Louis, MO, USA), Van Gieson
130 solution P (Fujifilm-Wako, Chuo-ku, Osaka, Japan), and Van Gieson solution F (Fujifilm-
131 Wako, Chuo-ku, Osaka, Japan) were used for each staining. Experiments were
132 performed according to the manufacturer's protocol and previous studies (Arnegard and
133 Snoeks, 2001; Oliver and Arnegard, 2010). For AB staining, the pH was adjusted to 2.5
134 with 3% acetic acid (Fujifilm-Wako, Chuo-ku, Osaka, Japan). For VG staining, Van
135 Gieson solution P and Van Gieson solution F were prepared at a ratio of 7.55:1. The
136 stained area was calculated using ImageJ version 1.5.1 (Schneider et al., 2012) after
137 trimming the image to include only the lips. The images were captured using a Zeiss
138 Axioplan fluorescence microscope and a Zeiss Axiocam 503 color CCD camera (Carl
139 Zeiss, Oberkochen, Germany). Images were level-corrected, contrast adjusted, and tiled
140 using Adobe Photoshop CC 2022.

141 **Chondroitinase digestion**

142 Chondroitinase ABC from *Proteus vulgaris* (Sigma-Aldrich, St. Louis, MO, USA)
143 was used for chondroitinase digestion. 2 mUnits/0.1% bovine serum albumin
144 solutionBSA in 20 mmol/L Tris-HCl pH 7.7 were prepared as the enzyme solution. In the
145 experiment, the sections were first treated with 100 μL of 0.1% BSA solution for 5 min.
146 Next, 100 μL of enzyme solution, or BSA solution for the control, was added. During the
147 reaction, the sections were placed in a moist chamber. The enzymatic reaction was
148 performed at 37°C for 1 h. After the reaction, the samples were washed with water for
149 15 min, and then AB staining was examined.

150

151 **Identification of extracellular matrix (ECM)-related genes in cichlids**

152 To compare the transcriptome of cichlid lips, cichlid extracellular matrix (ECM)-
153 related genes were identified using the data from a previous study (Hynes and Naba,
154 2012). The amino acid sequences of human, mouse, and zebrafish ECM-related genes
155 registered in Human GENCODE 41, Mouse GENCODE M30, Ensembl, and UniProtKB
156 were used as queries, and those of M. zebra and O. niloticus were identified from their
157 protein sequences registered in Ensembl and Uniprot using OrthoFinder2 v.2.5.4. The
158 longest amino acid sequence of each isoform was selected to represent the gene. After
159 Orthfinder2 analysis, ortholog correspondences were corrected manually, if necessary.
160 A list of genes is provided in Table S2.

161

162 **Sample preparation for proteomic analysis**

163 Proteomic analysis was conducted on the lips of cichlids, *H. chilotes*
164 (hypertrophied) and *H. sauvagei* (normal), from Lake Victoria. The lips of *H. chilotes*
165 were analyzed separately for the base and tip. For proteomics analysis of pellet fractions,
166 dissected samples were homogenized with a pestle in 1% sodium dodecyl sulfate in
167 PBS supplemented with a protease inhibitor cocktail (cOmpleteTM mini, EDTA free,
168 Roche, Basel, Switzerland). The resultant lysate was centrifuged at 20,000× g for 15
169 min, and the supernatant was extracted. A fourfold volume of methanol, an equal volume
170 of chloroform, and a threefold volume of distilled water were added to the supernatants,
171 and the suspension was vortexed vigorously for 1 min after each addition. After vortexing,
172 the supernatants were centrifuged at 15,000× g for 2 min and the upper water layer was
173 removed. A threefold volume of methanol was added, and the samples were centrifuged
174 at 15,000× g for 2 min. The supernatant was removed by aspiration and the pellet
175 fraction was obtained.

176 Subsequently, sample preparation for LC−MS/MS was performed according to
177 a previous study (Masuda et al. 2009), with some modifications. Proteins in the pellet
178 fraction were solubilized in phase transfer surfactant (PTS) solution (12 mM sodium
179 deoxycholate, 12 mM sodium lauryl sulfate, 100 mM Tris-HCl, pH 9.0), and the protein
180 concentration was determined using a Pierce BCA Protein Assay Kit (Thermo Fisher
181 Scientific, Waltham, MA). Then, 50 μg of protein in 50 μL of PTS was prepared for the
182 following procedure. First, the proteins were reduced by treatment with 5 mM
183 dithiothreitol (DTT) at room temperature for 30 min, and then alkylated with 50 mM
184 iodoacetamide in the dark at room temperature for 30 min. The reaction was stopped by
185 a fivefold dilution with 50 mM ammonium bicarbonate. Denatured proteins in the solution
186 were digested into peptide fragments in two steps: first, incubation with 0.5 μg of
187 Trypsin/Lys-C Mix (Promega, Madison, WI) at room temperature for 3 h, followed by
188 incubation with 1.0 μg of Trypsin/Lys-C Mix at 37°C overnight. After digestion, an equal
189 volume of ethyl acetate and 0.5% trifluoroacetic acid (final concentration) were added
190 and the mixture was shaken vigorously for 2 min. The mixture was then centrifuged at
191 15,700 × g for 2 min and the upper ethyl acetate layer was removed. The solvent was
192 dried using a centrifugal evaporator. The residual pellet was redissolved in 600 μL of
193 0.1% TFA and 2% acetonitrile and then desalted using a handmade StageTip composed
194 of an SDB-XC Empore Disk (3 M, Maplewood, MN, USA). The solution was applied to
195 a StageTip equilibrated with 0.1% TFA and 2% acetonitrile, washed with 0.1% TFA and
196 2% acetonitrile, and eluted with 0.1% TFA and 80% acetonitrile. After the desalting, the
197 solvent was dried using a centrifugal evaporator, and the residual peptides were
198 redissolved in 120 μL of 0.1% TFA and 2% acetonitrile. The solution was centrifuged at
199 20,000 × g for 5 min, and the supernatant was collected and subjected to LC–MS/MS
200 measurement.
201

202 **LC–MS/MS measurement and data processing**

203 LC–MS/MS measurements and SWATH-MS acquisition (Gillet et al., 2012) were
204 performed using an Eksigent NanoLC Ultra and TripleTOF 4600 tandem-mass
205 spectrometer or an Eksigent nanoLC 415 and TripleTOF 6600 mass spectrometer (AB
206 Sciex, USA). The trap column used for nanoLC was a 5.0 mm × 0.3 mm ODS column
207 (L-column2, CERI, Bunkyo-ku, Tokyo, Japan), and the separation column was a 12.5
208 cm × 75 μm capillary column packed with 3 μm C18- silica particles (Nikkyo Technos,
209 Bunkyo-ku, Tokyo, Japan). The flow rate of nanoLC was 300 nL/min, and peptides were
210 separated by applying a 10-40% linear acetonitrile gradient over 70 min in the presence
211 of 0.1% formic acid. Data were acquired in DIA/SAWTH (Gillet et al., 2012) mode
212 controlled by Analyst software (version 1.7, AB Sciex, USA) with the following settings:
213 15 Da fixed width window in 350–1100 m/z (50 cycles), 45 msec acquisition time, and
214 100–1600 fragment ion m/z range. The measurements were performed three times for
215 each sample. Quantitative values were calculated by DIA-NN software (version 1.8.1)
216 with default settings (Demichev et al., 2020). The library for DIA/SWATH quantification
217 was generated using DIA-NN with the amino acid sequences of *M. zebra* total proteins
218 obtained from UniProtKB. The P-value and fold change were calculated by Python
219 libraries (scipy.stats and pandas) using the proteins detected in all three measurements
220 for three parts of the lips. Quantitative values were scaled by log10 before statistical
221 tests (Welch's t-test, two sided).

222

223 **RNA-seq analysis**

224 RNA was extracted from dissected lower lips using TRI Reagent (Molecular
225 Research Center, Inc.). Three species of Lake Victoria cichlids (*Haplochromis chilotes*
226 [n = 4], *H. sauvagei* [n = 4], *H. pyrrhocephalus* [n = 2]), three species of Lake Malawi

227 cichlids (*Placidochromis milomo* [n = 5], *Maylandia zebra* [n = 2], *Labidochromis*
228 *caeruleus* [n = 5]), and two species of Lake Taganyika cichlids (*Lobochilotes labiatus* [n
229 = 4], *Tropheus moorii* [n = 4]) were used for RNA-seq (Table S1). The extracted total
230 RNA was sequenced at 100 bp paired-end reads on a NovaSeq 6000 by Macrogen
231 Japan Corp., using a TruSeq stranded mRNA Library Prep Kit. The raw FASTQ data
232 were qualified using fastp version 0.20.1 (Chen et al., 2018) in default option. The high-
233 quality reads were mapped to *M. zebra* genome (Conte and Kocher, 2015) using STAR
234 version 2.7.10a (Dobin et al., 2013) and quantified using rsem-calculate-expression
235 version 1.3.1 (Li and Dewey, 2011). Principal component analysis (PCA) was conducted
236 using Python scikit-learn package. Low-expression genes (average transcripts per
237 kilobase million (TPM) below 5.0) were removed, and expression rates were normalized
238 by Z score. The resulting 12,231 genes were used for PCA analysis. For PCA of ECM-
239 related genes, we used 186 genes as inputs, which were extracted in the same manner.
240 DEGs were detected using the TCC package (Sun et al., 2013) with the TMM-edgeR-
241 edgeR combination. Enrichment analysis was conducted using DAVID version 2023q3
242 (Huang et al., 2009a, 2009b) after the gene IDs of *M. zebra* were converted to human
243 gene IDs using ensembl biomart. The annotation terms used for functional annotation
244 clustering were collected from biological process (Gene Ontology), KEGG, and
245 Reactome.
246

247 **Results**

248 **Comparative histology of hypertrophied and normal lips among cichlids from the**
249 **East African Great Lakes**

250 We compared the histological structures of hypertrophied and normal lips
251 among cichlids from Lake Victoria, Lake Malawi, and Lake Tanganyika. For an accurate
252 comparison, we first observed serial sections of the lower lips with the jaws to determine
253 the midline of the lips in each species, which was defined by the shape of the lower
254 dentary bone (Figure S1). Based on this definition, we compared the distribution of
255 collagens and proteoglycans, the main ECM components, at the midline of the lower lips
256 using VG stain (red) and AB stain (blue). Hypertrophied lips of cichlids from all lakes
257 generally consisted of two mutually exclusive structures in the hypodermis: a collagen-
258 rich layer at the base and a proteoglycan-rich layer at the tip. Moreover, comparison of
259 hypertrophied lips and normal lips showed that the area of the proteoglycan-rich (AB
260 positive) layer was notably enlarged in hypertrophied lips, which was statistically
261 significant in cichlids from Lake Victoria. (Figure 1; Figure S2).
262

263 **Proteoglycans and their related proteins were abundant in the hypertrophied lips**
264 To explore the proteins associated with lip hypertrophy, we used shotgun
265 proteomics to compare the protein profiles of hypertrophied lips and normal lips in Lake
266 Victoria cichlids. Using LC–MS/MS in DIA/SWATH acquisition mode, we detected 2,720
267 proteins in total from three types of lip samples: the tip of hypertrophied lips
268 (proteoglycan-rich loose connective tissue), the base of hypertrophied lips (collagen-rich
269 loose connective tissue), and normal lips. First, we observed few differences between
270 the tip of hypertrophied lips and the base of the hypertrophied lips (Figure S3). On the
271 other hand, we identified 138 differentially accumulated proteins between the tip of

272 hypertrophied and normal lips (133 proteins upregulated and 5 proteins downregulated)
273 under the criteria (P-value < 0.05; log2 fold change > 1) (Figure 2A). Compared to
274 normal lips, hypertrophied lips are abundant in proteoglycans and their related proteins
275 such as versican a/b (vcan), brevican (bcan), periostin (postn), hyaluronan and
276 proteoglycan link protein 1 (hapln1), and hyaluronan binding protein 2 (habp2), which
277 were highly accumulated in hypertrophied lips with statistical significance (Figures 2B-
278 D). The results of the above proteomic analyses were consistent with the histological
279 observation that proteoglycan-rich loose connective tissue was enlarged in
280 hypertrophied lips compared to that in normal lips.
281

282 **Chondroitin sulfate proteoglycans such as vcan are the main component of**
283 **hypertrophied lips**

284 Next, we performed glycosaminoglycan digestion experiments to examine
285 whether loose connective tissue was composed of the proteoglycans, which were
286 detected as highly accumulated proteins in hypertrophied lips. Among the proteoglycans,
287 we focused on vcan, which showed a clear difference in protein amount between
288 hypertrophied and normal lips. Vcan is a high-molecular-weight core protein
289 (approximately 370 kDa) that is glycosylated with chondroitin sulfate (Iozzo and
290 Schaefer, 2015; Wight et al., 2020). Therefore, we stained sections of hypertrophied lips
291 with AB stain after the treatment with chondroitinase, which specifically digests
292 glycosylated chondroitin sulfate. As a result, AB staining in loose connective tissue
293 became drastically weak in treated sections, which was common among lakes (Figures
294 2E-G). Our data suggested that the main components of enlarged loose connective
295 tissue in hypertrophied lips are chondroitin sulfate proteoglycans such as vcan.
296

297 **Convergent expression in ECM-related genes**

298 Given that proteomic analysis is not well suited for detecting high-turnover
299 proteins (Decaris et al., 2014), we additionally conducted transcriptomic comparisons
300 between hypertrophied and normal lips across Lake Victoria, Lake Malawi and Lake
301 Tanganyika. PCA with all expressed genes clustered according to each lake, reflecting
302 genetic closeness (Figure 3A). Next, we compared the gene expression of ECM-related
303 genes, for which the proteomics profiles were clearly distinct between hypertrophied and
304 normal lips. A PCA plot was drawn using only ECM-related genes in cichlids, which were
305 identified based on a previous study (Hynes and Naba, 2012) (Figure 3B). In contrast to
306 the PCA plot of all expressed genes, that of ECM-related genes reflected differences of
307 lip morphology (hypertrophied and normal), regardless of the genetic closeness. To
308 evaluate the results of these PCAs, we calculated the average silhouette (AS) score,
309 which assesses the cohesiveness of the cluster. In the PCA of all genes, the AS score
310 of the cluster by lake (0.1601) was higher than that of the cluster by lip morphology
311 (0.0849). In contrast, in the PCA of ECM-related genes, the AS score of the cluster by
312 lip morphology (0.1798) was higher than that of the cluster by lake (0.0774). These
313 results indicate that hypertrophied lips are characterized by the expression of ECM-
314 related genes.

315 Next, we compared the expression of each ECM-related gene between
316 hypertrophied and normal lips of cichlids from the three lakes (Figures 3C, D). As a result,
317 several ECM-related genes showed higher expression in hypertrophied lips than in
318 normal lips, such as type-I collagen fibrils (col1a1a, col1a1b, and col1a2), hyaluronan-
319 binding proteoglycans (vcanb, hapln1a/b, and hapln3), periostins (postna/b), small
320 leucine-rich proteoglycans (dcn, lum, and aspn), and microfibril-related genes (mfap2,
321 emilin1b, emilin2a, emilin2b, and emilin3b). Consistent with the proteomics comparison,
322 some proteoglycans and their related proteins (vcanb, hapln1a, and postna) showed

323 significantly higher expression in hypertrophied lips than in normal lips. In conclusion,
324 we observed convergent expression of ECM-related genes in hypertrophied lips of
325 cichlids among the East African Great Lakes.
326

327 **High expression of ECM-related genes in the juvenile and adult stages of**
328 **hypertrophied lips**

329 Next, we performed transcriptome analysis by comparing juvenile and adult
330 cichlids from Lake Victoria to examine the possible expression changes in ECM-related
331 genes during development (Figure 4A). In the adult stage, the expression profiles of
332 ECM-related genes between hypertrophied and normal lips showed obvious differences
333 (low correlation value 0.75; orange square in Figure 4B), which was consistent with
334 previous analyses (Figure 3B). In the juvenile stage, the difference was weaker than in
335 adult stage (0.85; blue square in Figure 4B). Interestingly, the difference in the
336 expression profiles between the juvenile and adult stages of hypertrophied lips (0.90;
337 Figure 4B black square) was lower than that of normal lips (0.83; Figure 4B black dotted
338 square). To examine the genes contributing to the differences in expression profiles, we
339 performed hierarchical clustering analysis for the relative expression levels of each
340 ECM-related gene (Figure 4C, D). The analyses consistently showed the higher
341 expression of proteoglycans and their related genes (e.g. postna, vcanb, and hapln1b)
342 in hypertrophied lips at both the juvenile and adult stages. These results indicated that
343 the accumulation of proteoglycans, which leads to the lip hypertrophy, begins at the
344 juvenile stage.
345

346 **The Wnt pathway-related genes are highly expressed in hypertrophied lips**

347 To explore the molecular pathways involved in lip hypertrophy, we performed
348 an enrichment analysis using DAVID (Huang et al., 2009a, 2009b) against DEGs
349 obtained from the transcriptome. The results of the enrichment analysis suggested that
350 the genes that were categorized into both canonical and non-canonical Wnt signaling
351 pathways, were highly expressed in the hypertrophied lips of juvenile and adult cichlids.
352 Comparison of the developmental stages revealed differences in the expression
353 patterns between the canonical and noncanonical Wnt pathways: the genes of the
354 canonical Wnt pathway were highly expressed in both stages (Figures 5C–I), whereas
355 those of the non-canonical pathway were highly expressed only in the adult stage
356 (Figures 5J–N). We further compared the expression patterns of these Wnt signaling
357 pathways in cichlids from Lake Malawi and Lake Tanganyika. Notably, the expression of
358 genes for the Wnt pathway tend to be higher in the hypertrophied-lip cichlids in Lake
359 Malawi and Lake Tanganyika, although the differences were not as distinct as those
360 found in cichlids from Lake Victoria. These results were consistent with those of previous
361 studies that showed that the expression of ECM-related genes, such as vcan, was
362 promoted by the Wnt signaling pathway (Rahmani et al., 2005). Continuous activation
363 of the Wnt signaling pathway throughout all developmental stages may lead to lip
364 hypertrophy via the deposition of ECM-related proteins.

365

366 **Discussion**

367 In this study, we investigated the differences between hypertrophied and normal
368 lips using histology, proteomics, and transcriptomics. Histological comparisons revealed
369 that cichlids with hypertrophied lips from all three East African Great Lakes exhibited an
370 enlargement of the proteoglycan-rich loose connective tissue. By integrating proteomics
371 and transcriptomics, we identified several proteins that were significantly accumulated
372 in the hypertrophied lips. Among them, proteoglycans and their related proteins such as
373 vcan and postn, could be a primary factor that explains lip hypertrophy. Interestingly,
374 these proteins are known to be involved in the formation of human keloids, a skin
375 disease characterized by an overabundance of ECM in the dermis (Deng et al., 2021;
376 Naitoh et al., 2005). It has been discussed that excessive deposition of vcan retains
377 large amounts of water in the sulfate or carbonyl moieties, contributing to the increase
378 in keloid volume (Yagi et al., 2013). In addition, postn is involved in fibrosis through its
379 ability to interact with multiple ECM proteins. For example, upregulation of postn leads
380 to strengthening of the ECM by crosslinking with collagens and excessive deposition of
381 ECM proteins by differentiating quiescent fibroblasts into activated myofibroblasts
382 (Ashley et al., 2017; Yamaguchi et al., 2013). In cichlid, vcan and postn may have a
383 similar function, leading to hypertrophic loose connective tissue.

384 The Wnt signaling pathway was also elevated in both the adult and juvenile
385 stages of hypertrophied lips. The Wnt signaling pathway is known to be a key regulator
386 of the expression of ECM-related genes; for example, vcan is directly regulated by the
387 Wnt/b-catenin and tcf4 signaling pathways, leading to vcan upregulation (Rahmani et al.,
388 2005). Additionally, Wnt5a, a noncanonical pathway, plays a critical role in dermal
389 condensation during mouse and sugar glider development (Feigin et al., 2023). Cellular-
390 level localization of gene expression for the canonical and non-canonical Wnt pathway

391 is required for future examination of the involvement of the Wnt signaling pathway in lip
392 hypertrophy.
393 Although the parallel evolution of hypertrophied lips in cichlids has received
394 considerable attention in evolutionary biology, the commonality and diversity underlying
395 their molecular basis remain poorly understood. This might be primarily due to the fact
396 that many genetic factors are possibly involved in the dynamic organization of
397 hypertrophied lips. We uncovered the apparent relationships between hypertrophied lips
398 and the expression profiles of ECM proteins, in particularly proteoglycans. The trends
399 for the overall expression of ECM-related genes were similar across hypertrophied lip
400 species, but we rarely observed a specific gene that was commonly expressed at high
401 or low levels in all three examples of hypertrophied lips across all East African Great
402 Lakes. Furthermore, although we focused primarily on the relationship between the Wnt
403 signaling pathway and lip hypertrophy, several other pathways may be involved in the
404 development of hypertrophied lips. These findings imply that although enlargement of
405 proteoglycan-rich loose connective tissue is common in hypertrophied lips, the
406 developmental pathways to accomplish this are diverse in each lake. The pathways
407 responsible for lip hypertrophy therefore require further investigation. Our approach of
408 integrating histology, proteomics and transcriptomics revealed complex evolutionary
409 pathways for hypertrophied lips in cichlids of the East African Great Lakes, providing an
410 opportunity to further understand the molecular mechanism underlying the parallel
411 evolution of polygenic traits, such as jaw morphology or body shape.
412

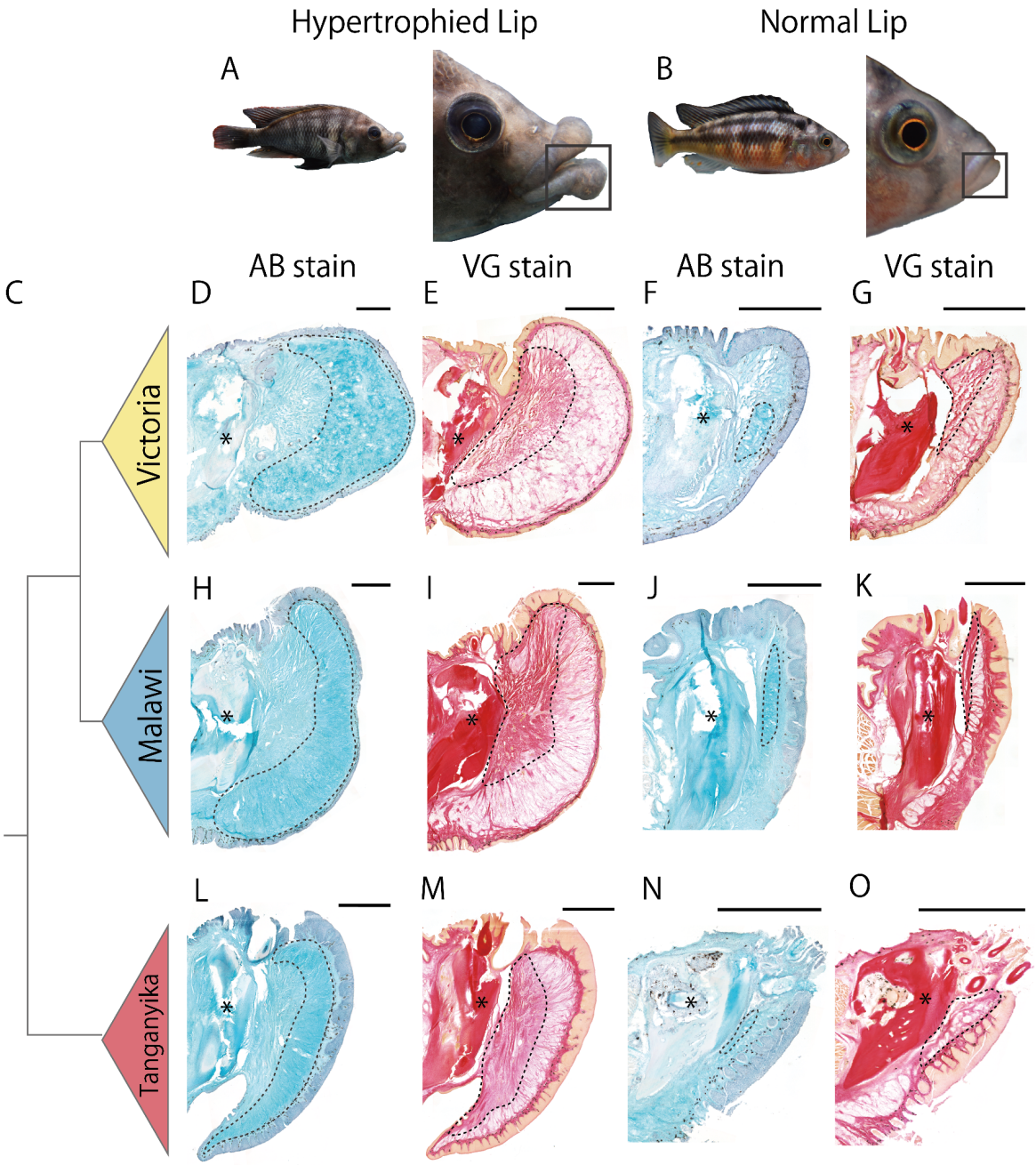


Fig.1 Histological comparisons of hypertrophied and normal lips of cichlids among lakes, focusing on proteoglycan-rich loose connective tissue. (A,B). Examples of typical cichlids with hypertrophied lips (A, *Haplochromis chilotes*) and normal lips (B, *H. sauvagei*). The lower jaw and lip are highlighted by black squares. (C) Phylogenetic relationships among cichlids from Lake Victoria, Lake Malawi, and Lake Tanganyika.

(D–O) The midline of the lips, based on the lower dentary bone of the lower jaw (marked by asterisks), in cichlids from Lake Victoria (D,E, *H. chilotes* with hypertrophied lips; F,G, *H. sauvagei* with normal lips), Lake Malawi (H,I, *Placidochromis milomo* with hypertrophied lips; J,K, *Maylandia lombardoi* with normal lips), and Lake Tanganyika (L,M, *Lobochilotes labiatus* with hypertrophied lips; N,O, *Tropheus moorii* with normal lips) were sectioned in sagittal, and stained with AB: Alcian blue (D, H, L, F, J, and N), or VG: Van Gieson (E, I, M, F, J, and N). Each section is shown with the dorsal side up and the anterior side to the right. The area surrounded by dotted lines shows proteoglycan-rich tissue (stained dense blue by AB) or collagen-rich tissue (stained dense red by VG). Each species showed mutually exclusive staining by AB and VG, and AB-positive (proteoglycan-rich) tissues were commonly enlarged in hypertrophied lips among the cichlids from all three lakes with statistical significance. Scale bars are 1mm.

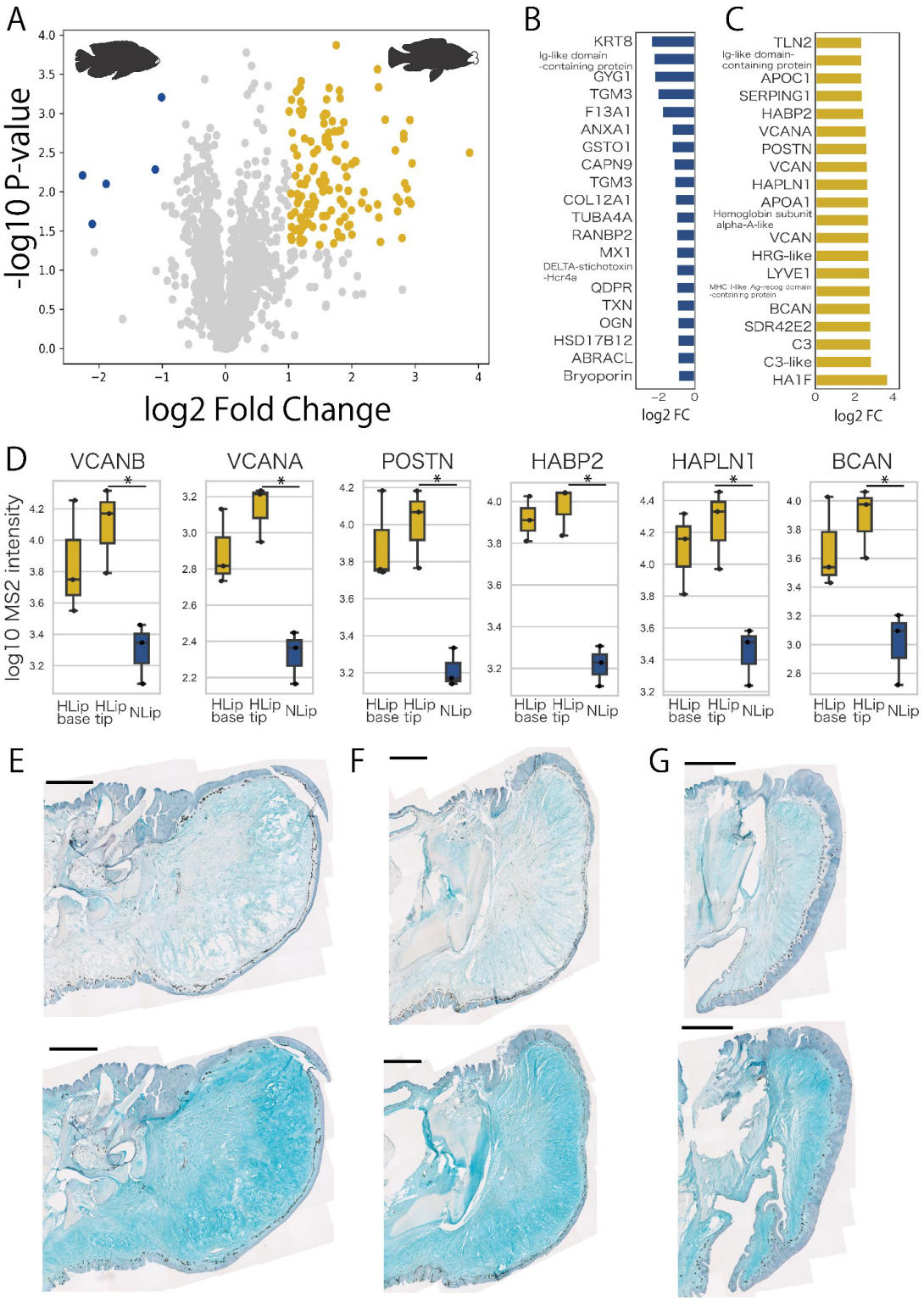


Fig.2 Chondroitin sulfate proteoglycans accumulate in hypertrophied lips. (A) Volcano plot of proteomics comparison between the tips of hypertrophied and normal lips.

Statistical significance was examined using Welch's t-test. Yellow plots highlight highly accumulated proteins in hypertrophied lips, and blue plots highlight highly accumulated proteins in normal lips with the criteria of P-value < 0.05 & log2 fold change > 1. (B,C) The top 20 proteins highly accumulated in normal (B) and hypertrophied lips (C) are shown along with log2 fold change relative to normal lips. (D) The relative amount of each proteoglycan-related protein. MS2 intensity indicates the intensity of peptides obtained from LC–MS/MS and DIA/SWATH analysis using DIA-NN software by comparing three parts of the lip: the base of hypertrophied lips (HLip), the tip of HLip, and normal lips (NLip). Welch's t-test was used to examine statistical significance. (E–G) Digestion of chondroitin sulfate by treatment with (upper panels) and without (lower panels) chondroitinase in hypertrophied lips of cichlids from Lake Victoria (E), Lake Malawi (F), and Lake Tanganyika (G). Each section is shown with the dorsal side up and the anterior side to the right. Scale bars are 1mm.

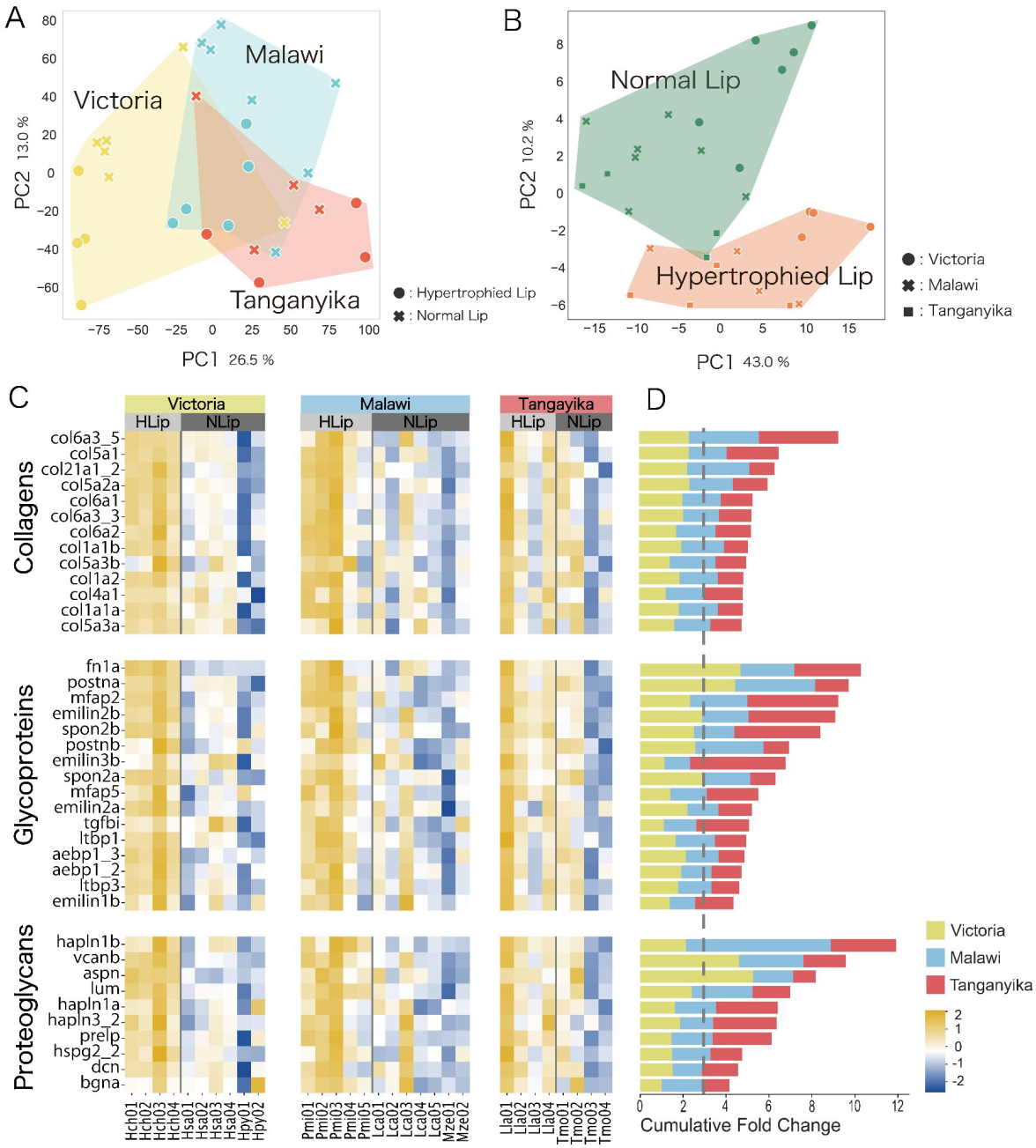
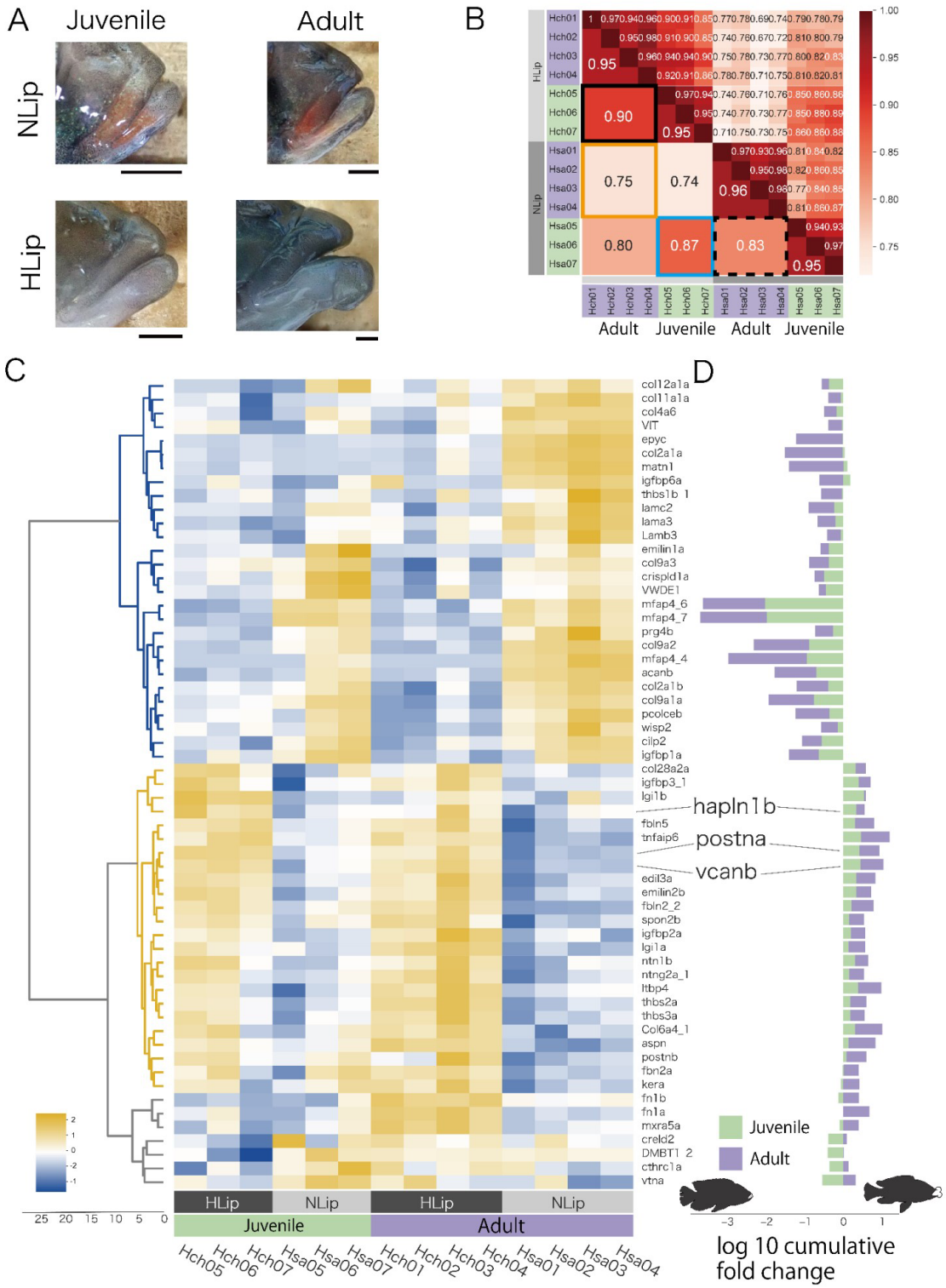


Fig.3 Hypertrophied lips of cichlids are characterized by higher expression of ECM-related genes. Transcriptomic analysis using three species of Lake Victoria cichlids (*H. chilotes* (Hch) with hypertrophied lips; *H. sauvagei* (Hsa) and *H. pyrrhocephalus* (Hpy) with normal lips), three species of Lake Malawi cichlids (*Placidochromis milomo* (Pmi) with hypertrophied lips; *Maylandia zebra* (Mze) and *Labidochromis caeruleus* (Lca) with

454 normal lips), and two species of Lake Tanganyika cichlids (*Lobochilotes labiatus* (Lla)
455 with hypertrophied lips; *Tropheus moorii* (Tmo) with normal lips). (A, B) PCA plots of all
456 expressed genes showed differences depending on the lakes (A), while those of ECM-
457 related genes showed differences depending on lip morphology: hypertrophied or
458 normal (B). (C) Heatmap of ECM-related gene expression between hypertrophied and
459 normal lips. The expression rate (TPM) normalized by Z score of each gene is compared
460 for Lake Victoria (yellow; Hch, Has and Hpy), Lake Malawi (blue; Pmi, Lca and Mze),
461 and Lake Tanganyika (red; Lla and Tmo). The intensity of the relative expression rate is
462 shown in the bottom right of the figure. (D) Cumulative fold change of the ECM-related
463 genes comparing hypertrophied and normal lips. The gene symbol corresponds to
464 Figure 3C. The color represents the lake information. The gray dotted line represents
465 the value at which higher expression was observed in the hypertrophied lips.
466



467

468 Fig.4 Comparison of juvenile and adult cichlids to examine gene expression changes

469 during the development of hypertrophied lips. (A) The lips of the cichlids used in this

470 analysis are shown with the dorsal side up and the anterior side to the right: the juvenile
471 and adult stages of *Haplochromis chilotes* (Hch, representative of hypertrophied lips:
472 HLip) and *H. sauvagei* (Hsa, representative of normal lips: NLip) from Lake Victoria.
473 Hypertrophied lips at juvenile stage are not notably enlarged. Scale bars are 0.5 mm.
474 (B) Spearman correlation matrix of the ECM-related gene expression rate (TPM)
475 normalized by Z score. The raw correlation values are indicated on the right side and
476 the average correlation values within each stage and species are indicated on the left
477 side. The orange frame highlights the correlations between the hypertrophied and
478 normal lips at the adult stage, and the blue frame highlights those at the juvenile stage.
479 The black frame highlights the correlations between juvenile and adult samples of each
480 species. (C) Cluster heatmap of the ECM-related gene expression rate (TPM)
481 normalized by Z score in each gene. Each gene is clustered by similarity of expression
482 rates. The intensity of the relative expression rate is shown on the bottom left of the
483 figure. (D) The log10 cumulative fold change of the ECM-related genes comparing
484 hypertrophied and normal lips. The gene symbol corresponds to Figure 4C. The positive
485 values indicate high expression in hypertrophied lips, whereas the negative values
486 indicate low expression in hypertrophied lips. The colors of the bars indicate the stage
487 of the samples.
488

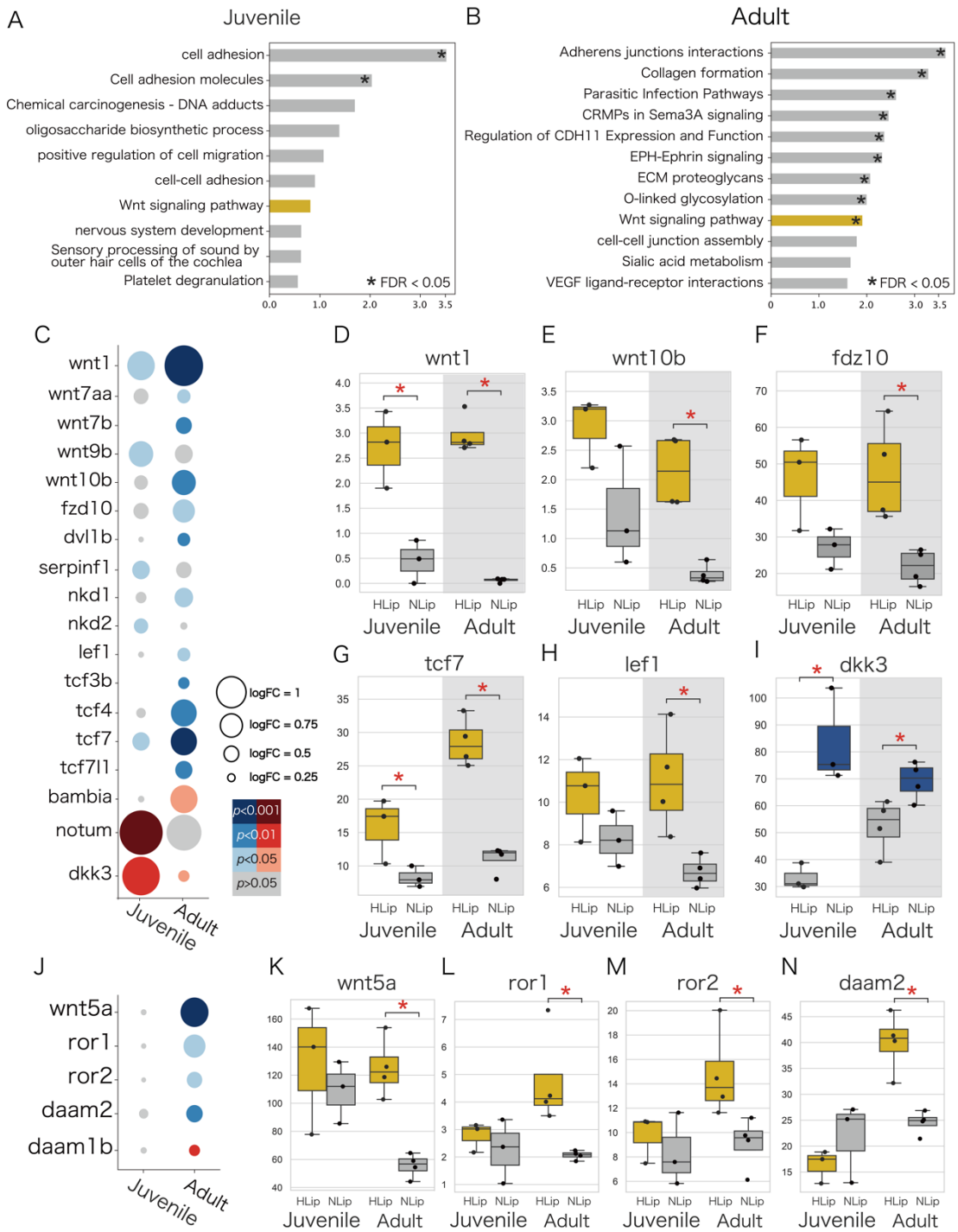
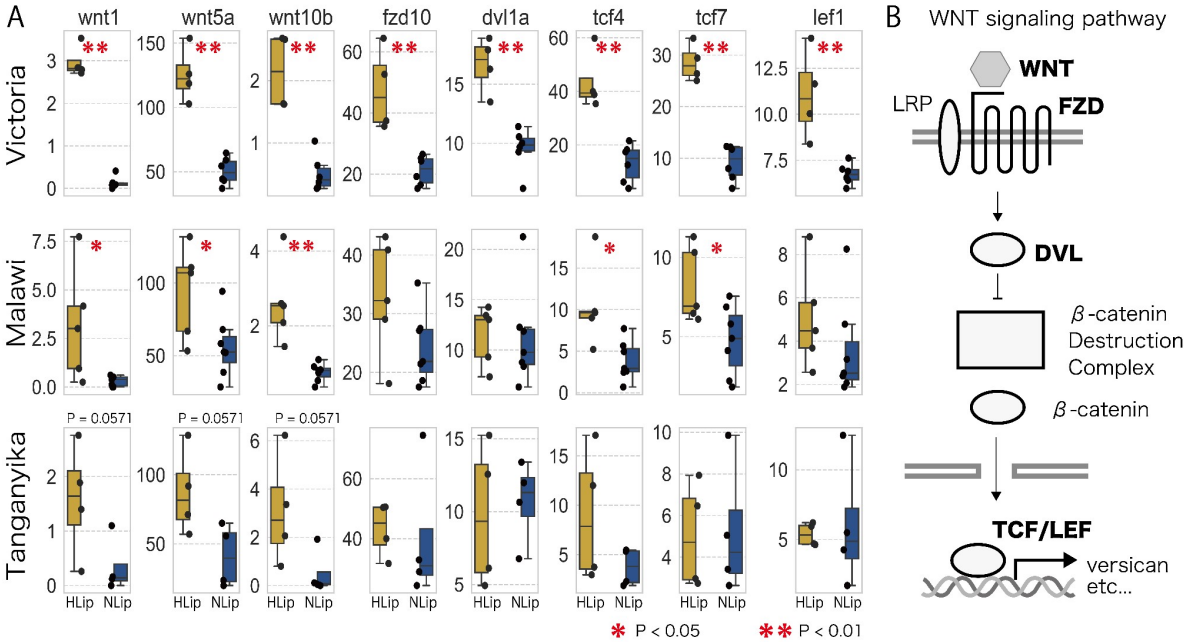


Fig.5 Wnt pathway-related genes are highly expressed in hypertrophied lips of juvenile and adult cichlids. (A, B) Enrichment analysis of highly expressed genes in the hypertrophied lips. Labels are according to the top annotation in DAVID annotation

clusters. Asterisks indicate statistically significant annotation clusters (FDR < 0.05). The Wnt-related annotations are highlighted in yellow. (C–N) Expression of the specific Wnt pathway-related genes. (C, J) The size of the dots represents the log fold change, and the color of the dots represents the P-value. Blue dots represent genes that are highly expressed in hypertrophied lips (HLip), and red dots represent genes that are highly expressed in normal lips (NLip). (D–I, K–N) Box plots show the expression rate (TPM) of Wnt pathway-related genes. Yellow boxes represent genes that are highly expressed in hypertrophied lips and blue boxes represent genes that are highly expressed in normal lips, respectively. Mann–Whitney U-test was used to examine the statistical significance (*: P-value < 0.05). (C–I) The canonical Wnt pathway-related genes, and (J–N) show the non-canonical Wnt pathway-related genes.



505

506 Fig.6 The trend of higher expression of Wnt-related genes in hypertrophied lips. (A) Box

507 plots of expression rate (TPM) of Wnt related genes of hypertrophied lips (HLip: yellow)

508 and normal lips (NLip: blue). Mann–Whitney U-test was used to examine the statistical

509 significance (*: P-value < 0.05, **: P-value < 0.01). (B) Illustration of the canonical Wnt

510 pathway. DEGs are marked in the bold.

511

Albertson RC, Kocher TD. 2006. Genetic and developmental basis of cichlid trophic diversity. *Heredity* **97**:211–221.

Arnegard ME, Snoeks J. 2001. New Three-Spotted Cichlid Species with Hypertrophied Lips (Teleostei: Cichlidae) from the Deep Waters of Lake Malaŵi/Nyasa, Africa. *The American Society of Ichthyologists and Herpetologists*. doi:10.1643/0045-8511(2001)001[0705:NTSCSW]2.0.CO;2

Ashley SL, Wilke CA, Kim KK, Moore BB. 2017. Periostin regulates fibrocyte function to promote myofibroblast differentiation and lung fibrosis. *Mucosal Immunol* **10**:341–351.

Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**:i884–i890.

Colombo M, Diepeveen ET, Muschick M, Santos ME, Indermaur A, Boileau N, Barluenga M, Salzburger W. 2013. The ecological and genetic basis of convergent thick-lipped phenotypes in cichlid fishes. *Mol Ecol* **22**:670–684.

Conte MA, Kocher TD. 2015. An improved genome reference for the African cichlid, Metriaclima zebra. *BMC Genomics* **16**:724.

Decaris ML, Gatmaitan M, FlorCruz S, Luo F, Li K, Holmes WE, Hellerstein MK, Turner SM, Emson CL. 2014. Proteomic analysis of altered extracellular matrix turnover in bleomycin-induced pulmonary fibrosis. *Mol Cell Proteomics* **13**:1741–1752.

Demichev V, Messner CB, Vernardis SI, Lilley KS, Ralser M. 2020. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods* **17**:41–44.

Deng C-C, Hu Y-F, Zhu D-H, Cheng Q, Gu J-J, Feng Q-L, Zhang L-X, Xu Y-P, Wang D, Rong Z, Yang B. 2021. Single-cell RNA-seq reveals fibroblast heterogeneity and increased mesenchymal fibroblasts in human fibrotic skin diseases. *Nat Commun* **12**:3709.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**:15–21.

Duenser A, Singh P, Lecaudey LA, Sturmbauer C, Albertson RC, Gessl W, Ahi EP. 2023. Conserved Molecular Players Involved in Human Nose Morphogenesis Underlie Evolution of the Exaggerated Snout Phenotype in Cichlids. *Genome Biol Evol* **15**. doi:10.1093/gbe/evad045

Feigin CY, Moreno JA, Ramos R, Mereby SA, Alivisatos A, Wang W, van Amerongen R, Camacho J, Rasweiler JJ 4th, Behringer RR, Ostrow B, Plikus MV, Mallarino R. 2023. Convergent deployment of ancestral functions during the evolution of mammalian flight membranes. *Sci Adv* **9**:eade7511.

Fryer G. 1972. The cichlid fishes of the Great Lakes of Africa. *Their Biology and Evolution* **641**.

Gillet LC, Navarro P, Tate S, Röst H, Selevsek N, Reiter L, Bonner R, Aebersold R. 2012. Targeted Data Extraction of the MS/MS Spectra Generated by Data-independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis*. *Mol Cell Proteomics* **11**:O111.016717.

Greenwood PH. 1974. Cichlid fishes of Lake Victoria, East Africa: the biology and evolution of a species flock. *Bull Brit Mus (Nat Hist) Zool Suppl* **6**:1–134.

Henning F, Machado-Schiaffino G, Baumgarten L, Meyer A. 2017. Genetic dissection of adaptive form and function in rapidly speciating cichlid fishes. *Evolution* **71**:1297–1312.

Huang DW, Sherman BT, Lempicki RA. 2009a. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37**:1–13.

563 Huang DW, Sherman BT, Lempicki RA. 2009b. Systematic and integrative analysis of
564 large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**:44–57.
565 Hynes RO, Naba A. 2012. Overview of the matrisome--an inventory of extracellular
566 matrix constituents and functions. *Cold Spring Harb Perspect Biol* **4**:a004903.
567 Iozzo RV, Schaefer L. 2015. Proteoglycan form and function: A comprehensive
568 nomenclature of proteoglycans. *Matrix Biol* **42**:11–55.
569 Kautt AF, Kratochwil CF, Nater A, Machado-Schiaffino G, Olave M, Henning F, Torres-
570 Dowdall J, Härer A, Hulsey CD, Franchini P, Pippel M, Myers EW, Meyer A. 2020.
571 Contrasting signatures of genomic divergence during sympatric speciation. *Nature*
572 **588**:106–111.
573 Kocher T, Conroy JA, McKaye KR, Richard SJ. 1993. Similar Morphologies of Cichlid
574 Fish in Lakes Tanganyika and Malawi Are Due to Convergence.
575 Kocher TD. 2004. Adaptive evolution and explosive speciation: the cichlid fish model.
576 *Nat Rev Genet* **5**:288–298.
577 Kratochwil CF, Liang Y, Gerwin J, Woltering JM, Urban S, Henning F, Machado-
578 Schiaffino G, Hulsey CD, Meyer A. 2018. Agouti-related peptide 2 facilitates
579 convergent evolution of stripe patterns across cichlid fish radiations. *Science*
580 **362**:457–460.
581 Krishnan J, Rohner N. 2017. Cavefish and the basis for eye loss. *Philos Trans R Soc*
582 *Lond B Biol Sci* **372**. doi:10.1098/rstb.2015.0487
583 Lecaudey LA, Singh P, Sturmbauer C, Duenser A, Gessl W, Ahi EP. 2021.
584 Transcriptomics unravels molecular players shaping dorsal lip hypertrophy in the
585 vacuum cleaner cichlid, Gnathochromis permaxillaris. *BMC Genomics* **22**:506.
586 Lecaudey LA, Sturmbauer C, Singh P, Ahi EP. 2019. Molecular mechanisms underlying
587 nuchal hump formation in dolphin cichlid, Cyrtocara moorii. *Sci Rep* **9**:20296.
588 Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data
589 with or without a reference genome. *BMC Bioinformatics* **12**:323.
590 Losos JB. 2011. Convergence, adaptation, and constraint. *Evolution* **65**:1827–1840.
591 Lukas B, Gonzalo M-S, Frederico H, Axel M. 2015. What big lips are good for: on the
592 adaptive function of repeatedly evolved hypertrophied lips of cichlid fishes. *Biol J*
593 *Linn Soc Lond* **115**:448–455.
594 Manousaki T, Hull PM, Kusche H, Machado-Schiaffino G, Franchini P, Harrod C, Elmer
595 KR, Meyer A. 2013. Parsing parallel evolution: ecological divergence and differential
596 gene expression in the adaptive radiations of thick-lipped Midas cichlid fishes from
597 Nicaragua. *Mol Ecol* **22**:650–669.
598 Masonick P, Meyer A, Darrin Hulsey C. 2023. A Kiss of Deep Homology: Partial
599 Convergence in the Genomic Basis of Hypertrophied Lips in Cichlid Fish and
600 Human Cleft Lip. *Genome Biol Evol* **15**. doi:10.1093/gbe/evad072
601 Masonick P, Meyer A, Hulsey CD. 2022. Phylogenomic Analyses Show Repeated
602 Evolution of Hypertrophied Lips Among Lake Malawi Cichlid Fishes. *Genome Biol*
603 *Evol* **14**. doi:10.1093/gbe/evac051
604 Muschick M, Indermaur A, Salzburger W. 2012. Convergent evolution within an adaptive
605 radiation of cichlid fishes. *Curr Biol* **22**:2362–2368.
606 Naitoh M, Kubota H, Ikeda M, Tanaka T, Shirane H, Suzuki S, Nagata K. 2005. Gene
607 expression in human keloids is altered from dermal to chondrocytic and osteogenic
608 lineage. *Genes Cells* **10**:1081–1091.
609 Okuda S, Watanabe Y, Moriya Y, Kawano S, Yamamoto T, Matsumoto M, Takami T,
610 Kobayashi D, Araki N, Yoshizawa AC, Tabata T, Sugiyama N, Goto S, Ishihama Y.
611 2017. jPOSTrepo: an international standard data repository for proteomes. *Nucleic*
612 *Acids Res* **45**:D1107–D1111.

613 Oliver MK, Arnegard ME. 2010. A new genus for Melanochromis labrosus, a problematic
614 Lake Malawi cichlid with hypertrophied lips (Teleostei: Cichlidae). *Ichthyol Explor*
615 *Freshwaters* **21**:209–232.
616 Rahmani M, Read JT, Carthy JM, McDonald PC, Wong BW, Esfandiarei M, Si X, Luo Z,
617 Luo H, Rennie PS, McManus BM. 2005. Regulation of the Versican Promoter by the
618 β-Catenin-T-cell Factor Complex in Vascular Smooth Muscle Cells*. *J Biol Chem*
619 **280**:13019–13028.
620 Salzburger W. 2018. Understanding explosive diversification through cichlid fish
621 genomics. *Nat Rev Genet* **19**:705–717.
622 Santos ME, Lopes JF, Kratochwil CF. 2023. East African cichlid fishes. *Evodevo* **14**:1.
623 Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image
624 analysis. *Nat Methods* **9**:671–675.
625 Stern DL. 2013. The genetic causes of convergent evolution. *Nat Rev Genet* **14**:751–
626 764.
627 Sun J, Nishiyama T, Shimizu K, Kadota K. 2013. TCC: an R package for comparing tag
628 count data with robust normalization strategies. *BMC Bioinformatics* **14**:219.
629 Wight TN, Kang I, Evanko SP, Harten IA, Chang MY, Pearce OMT, Allen CE, Frevert CW.
630 2020. Versican-A Critical Extracellular Matrix Regulator of Immunity and
631 Inflammation. *Front Immunol* **11**:512.
632 Yagi Y, Muroga E, Naitoh M, Isogai Z, Matsui S, Ikehara S, Suzuki S, Miyachi Y, Utani A.
633 2013. An ex vivo model employing keloid-derived cell-seeded collagen sponges for
634 therapy development. *J Invest Dermatol* **133**:386–393.
635 Yamaguchi Y, Ono J, Masuoka M, Ohta S, Izuhara K, Ikezawa Z, Aihara M, Takahashi
636 K. 2013. Serum periostin levels are correlated with progressive skin sclerosis in
637 patients with systemic sclerosis. *Br J Dermatol* **168**:717–725.
638

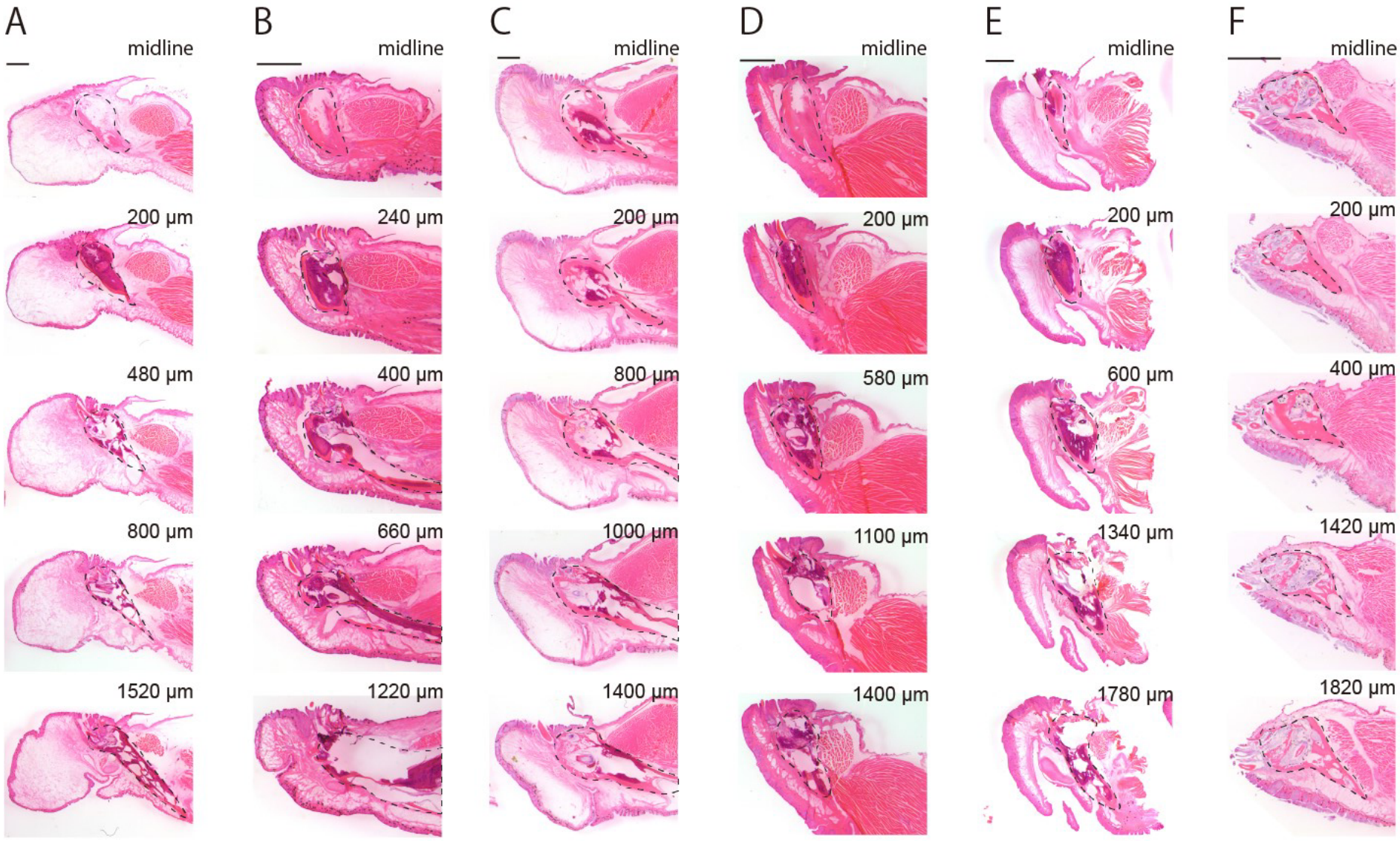

A
midline
200 µm
480 µm
800 µm
1520 µm
B
midline
240 µm
400 µm
660 µm
1220 µm
C
midline
200 µm
800 µm
1000 µm
1400 µm
D
midline
200 µm
580 µm
1100 µm
1400 µm
E
midline
200 µm
600 µm
1340 µm
1780 µm
F
midline
200 µm
400 µm
1420 µm
1820 µm

Fig S1 The midlines of the lips were defined by the shape of the lower dentary bone. Hematoxylin–eosin stain (HE stain) was used to examine serial sections of the lower lips and lower jaws of cichlids from Lake Victoria (A, *H. chilotes* with hypertrophied lips; B, *H. sauvagei* with normal lips), Lake Malawi (C, *Placidochromis milomo* with hypertrophied lips; D, *Maylandia lombardoi* with normal lips), and Lake Tanganyika (E, *Lobochilotes labiatus* with hypertrophied lips; F, *Tropheus moorii* with normal lips). Each section is shown with the dorsal side up and the anterior side to the right. The lower dentary bone became smaller as it approached the midline, and just at the midline, the bone was composed of cartilage. Scale bars are 1mm.

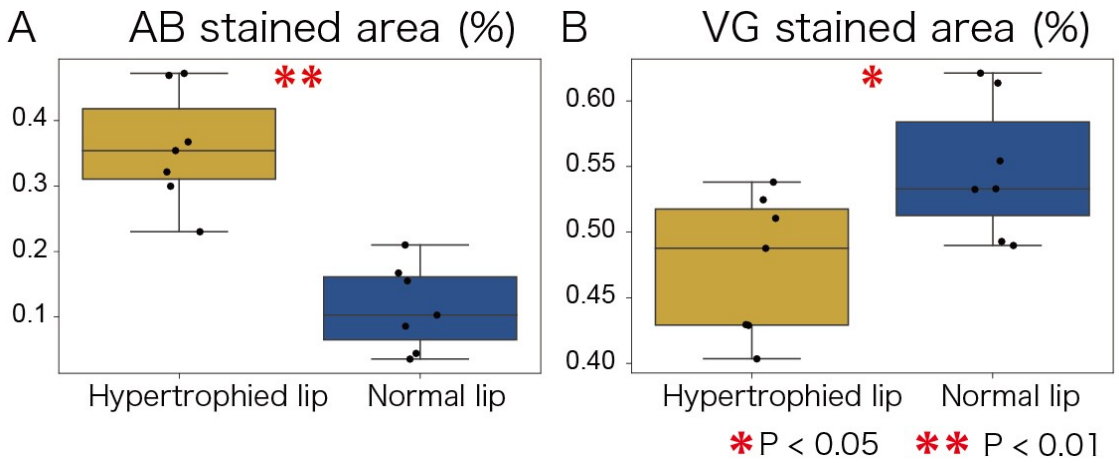


Fig. S2 The relative area of the proteoglycan-rich layer was significantly enlarged in the hypertrophied lips of Victoria cichlids. The stained area percentage of alcian blue stain (AB stain; A) and Van Gieson stain (VG stain; B) relative to the whole-lip area was calculated using ImageJ software. *H. chilotes* with hypertrophied lips (yellow box) and *H. sauvagei* with normal lips (blue box) were used in the experiment. Each plot is an average value calculated by three sections near the midline of the lips. Mann–Whitney U-test was used to examine the statistical significance (*: P-value < 0.05, **: P-value < 0.01).

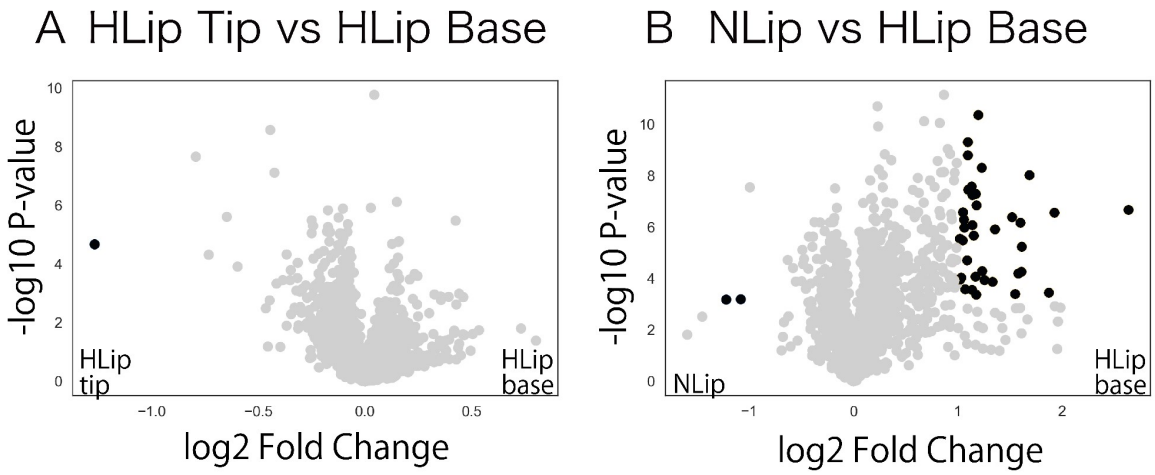


Fig. S3 Volcano plot of proteomics comparison between the tip of hypertrophied lips and the base of hypertrophied lips (A) and the base of hypertrophied lips and normal lips (B). Statistical significance was examined by Welch's t-test. Differentially accumulated proteins in each comparison are highlighted in the black plot with the following criteria: P-value < 0.05, log2 fold change > 1.

666 Table. S1 Basic information of the cichlids used in this study.

667

668 Table. S2 List of ECM-related genes of cichlids identified in this study. The classification
669 of ECM-related genes (the column of matrisome_annotation) follows a previous study
670 (Hynes and Naba, 2012).

671

672 Table. S3 All differentially accumulated proteins detected by LC–MS/MS under the
673 criteria of P-value < 0.05, log2 fold change > 1).

674

675 Table. S4 Results of enrichment analysis of highly expressed genes in hypertrophied
676 lips of juvenile and adult cichlids. The annotation terms were clustered by DAVID
677 annotation clustering.

678

679 Table. S5 All differentially expressed genes between hypertrophied lips and normal lips.
680 We compared adult Victoria cichlids (*H. chilotes* with hypertrophied lips and *H. sauvagei*
681 with normal lips), juvenile Victoria cichlids (*H. chilotes* and *H. sauvagei*), Malawi cichlids
682 (*Placidochromis milomo* with hypertrophied lips and *Labidochromis caeruleus* with
683 normal lips), and Taganyika cichlids (*Lobochilotes labiatus* with hypertrophied lips and
684 *Tropheus moorii* with normal lips).

685

Acknowledgement

This study was supported by JSPS KAKENHI (23KJ0951 to N. M., 17H04606, 20KK0167 to M.N.) and JST SPRING (JPMJSP2106). Computations were partially performed on the NIG supercomputer at ROIS National Institute of Genetics. The authors thank the Open Research Facilities for Life Science and Technology, Tokyo Institute of Technology for LC–MS/MS analysis (for technical assistance). The authors thank the Hongoh laboratory members. Nagatoshi Machii thank my family, Nagashi Machii, Mistue M. and Sawa M.

Data availability

RNA-seq short reads and raw proteome data are stored at NCBI and jPOST (Okuda et al., 2017) (Accession: PRJDB3401, PRJDB17757, and JPST003042/PXD051571).