Agouti-related peptide 2 facilitates convergent evolution of stripe patterns across cichlid fish radiations

Claudius F. Kratochwil1,3,5, Yipeng Liang1, Jan Gerwin1,3, Joost M. Woltering1, Sabine Urban1,3, Frederico Henning1,4, Gonzalo Machado-Schaffino1,4, C. Darrin Hulsey1,3, Axel Meyer1,3,5

The color patterns of African cichlid fishes provide notable examples of phenotypic convergence. Across the more than 1200 East African rift lake species, melanic horizontal stripes have evolved numerous times. We discovered that regulatory changes of the gene agouti-related peptide 2 (agrp2) act as molecular switches controlling this evolutionarily labile phenotype. Reduced agrp2 expression is convergently associated with the presence of stripe patterns across species flocks. However, cis-regulatory mutations are not predictive of stripes across radiations, suggesting independent regulatory mechanisms. Genetic mapping confirms the link between the agrp2 locus and stripe patterns. The crucial role of agrp2 is further supported by a CRISPR-Cas9 knockout that reconstitutes stripes in a nonstriped cichlid. Thus, we unveil how a single gene affects the convergent evolution of a complex color pattern.

Stephen Jay Gould famously posited that if it were possible to rerun the “tape of life,” outcomes would be different (1). The relative importance of determinism and contingency during evolution is still far from settled (2, 3). But for particular groups of organisms, one can now test Gould’s hypothesis. For instance, in less than 8 million to 12 million years, more than 1200 species of cichlid fishes have evolved to form repeated adaptive radiations in the East African Rift Valley lakes, such as Lakes Victoria, Tanganyika, and Malawi (Fig. 1B) (4–8). These adaptive radiations have given rise to a large diversity of species displaying various color patterns (Fig. 1, C to N), including the repeated occurrence of melanic horizontal stripes (Fig. 1A and supplementary text). Convergent evolution is prevalent in the East African cichlid radiations (9–11), providing a replicated natural experiment whereby distantly related species from independent adaptive radiations can be used to determine what mechanisms have generated these recurrent phenotypes (12–16). More specifically, we address whether horizontal stripes, a convergent phenotype, have an identical, similar, or different molecular bases across the independent adaptive radiations of cichlid fishes.

Previously (17)—using a genetic mapping panel of two Lake Victoria species, Pandamillus nyererei (Pnye, nonstriped, Fig. 1E) and Haplochromis sawwagei (Hsaу, striped, Fig. 1C)—we found that horizontal stripes (Fig. 2C) are inherited as a recessive Mendelian trait mapping to chromosome 18 (Fig. 2A). This was confirmed by a second cross involving the same nonstriped species and another striped species, H. chilotes (Hehi, striped) (Fig. 2A and supplementary text).

To more precisely isolate the causal genetic interval for stripe presence, we fine-mapped the trait using recombinant F2 individuals of the Pnye × Hsaу cross and reduced the causal interval from 600 to 25 kb (fig. S1). This interval contained the genes agouti-related peptide 2 (agrp2), v-type proton ATPase subunit d 2 (atp6vd02), and an unknown gene (unk) (Fig. 2B). The resequencing of all coding regions revealed no fixed missense or nonsense mutations (fig. S2), suggesting that cis-regulatory variation determines stripe presence.

The teleost-specific agrp2 (Fig. 3) is a strong candidate gene for stripes because its paralogs have been previously associated with pigmentation phenotypes (16–20). To test for agrp2 expression differences between nonstriped (Pnye) and striped (Hsaу) Lake Victoria cichlids, we performed quantitative polymerase chain reaction (qPCR; Fig. 2D and fig. S5) on a number of adult tissues, including skin (supplementary text). Here, agrp2 showed a significantly higher expression in the skin of Pnye (Fig. 2D and fig. S4). The lack of consistent expression variation between melanic and nonmelanic regions and generally across dorsoventral and anterior-posterior positions suggests that agrp2 does not shape pigmentation patterns through local expression-level variation but rather acts as a general stripe pattern inhibitor (fig. S6). Whereas qPCR revealed no such expression differences for paralogs and neighboring genes (supplementary text), qPCRs on F2 Pnye/Hsaу hybrid individuals confirmed that expression differences are linked to the agrp2 locus and exhibit an allelic dosage effect as expected for cis-regulatory mutations (fig. S4).

To identify causal mutations affecting both agrp2 expression and stripe phenotype, we sequenced the agrp2 locus in individuals (n = 10 individuals per population) from natural populations of the three hybrid-cross species. We screened for alternatively fixed, fully associated variants with the stripe phenotype in pairwise comparisons of each striped species (Hsaу or Hehi) versus the nonstriped Pnye. Our analyses indicated that a 1.1-kb interval within the first agrp2 intron (Fig. 2B) that exhibited shared, alternatively fixed alleles is a strong candidate region for a regulatory element controlling agrp2 expression (fig. S7). To test whether this 1.1-kb interval [enhancer of agrp2 in Pnye (Pnye.enh.a)] contains cis-regulatory elements that could influence interspecific differences between striped and nonstriped species, we tested the elements of both species in a green fluorescent protein (GFP) reporter assay in vivo (supplementary text). It showed that Pnye.enh.a efficiently modulates GFP expression (Tukey’s honest significant difference (HSD), P < 0.001) and is significantly more potent than the homologous sequence of the striped species Hsaу (Tukey’s HSD, P < 0.001) (Fig. 2, E and F, and fig. S8). Together, these results indicate that higher expression of agrp2, and thereby the suppression of stripe patterns, is indeed enhanced by Pnye.enh.a (fig. S9).

Our results reveal agrp2 as a major determinant of stripe presence that might be sufficient to suppress stripe patterns in Pnye. To further test this finding, we used CRISPR-Cas9 genome engineering to manipulate agrp2 and to thereby potentially derepress stripe patterns. Pnye eggs were injected with Cas9 and agrp2 guide RNAs, and we obtained four mutants, all of which had nonsense and frameshift mutations within agrp2 (fig. S10 and table S1). These CRISPR-Cas9 mutants developed a continuous middorsal stripe (Fig. 2H and fig. S10) yet no dorsolateral stripe (supplementary text). Because horizontal stripes were never observed in noninjected Pnye individuals (>100 observations; Fig. 2G), this strongly suggests that although species such as Pnye have no stripes, the genomic and developmental machinery for stripe pattern formation is in place, and stripes can reappear in this nonstriped species by experimental manipulation of agrp2.

Next, we tested if the expression levels of agrp2 and stripe patterns are generally associated across other cichlid species from the repeated species flocks of Lakes Victoria, Malawi, and Tanganyika, suggesting a shared molecular basis for convergent stripe phenotypes. Using qPCR on adult skins of striped and nonstriped species of each of the three major East African cichlid radiations (in total, 24 species; fig. S11), we revealed that nonstriped species commonly had higher agrp2 expression levels than striped species (Fig. 3B). This association was confirmed...
**Fig. 2. Agrp2 controls stripe loss in Lake Victoria cichlids.** (A) In two mapping crosses between Pnye and Hsau and Pnye and Hchi, horizontal stripes map to the same region on chromosome 18. (B) F2 recombinant fine-mapping isolates a 25-kb interval containing agrp2. (C and D) Skin biopsies show differential agrp2 expressions between skins from Hsau (C) and Pnye (D). Error bars indicate means ± SD. (E and F) An intronic 1.1-kb element of Pnye (Pnye.enh.a) is regulatory active (E) and shows stronger activity than the homologous sequence of the striped species Hsau (F) in a zebrafish larval GFP reporter assay. (G and H) CRISPR-Cas9 mediated knockouts of agrp2 in the normally nonstriped cichlid Pnye (G) develop stripes (H); midlateral stripe indicated by line). LOD score, logarithm of the odds score; N/N, homozygous for the Pnye agrp2 allele; S/S, homozygous for the Hsau agrp2 allele; C/C, homozygous for the Hchi agrp2 allele; tg, transgenic construct; gRNA, guide RNA; mls, midlateral stripe; dls, dorsolateral stripe; d, dorsal; dm, dorsomedial; vm, ventromedial; v, ventral.

by comparative phylogenetic analyses (Fig. 3A), demonstrating a significant evolutionary association between low agrp2 expression and stripe presence [phylogenetic analysis of variance (ANOVA); mean $P < 0.001$; supplementary text].

To determine if this convergence at the phenotypic and agrp2 gene expression level is also paralleled at the sequence level (16), we comparatively analyzed homologous enh.a sequences across cichlids from Lakes Victoria, Malawi, and Tanganyika. A tree of enh.a revealed substantial sequence variation and resolved striped Lake Victoria species as monophyletic, suggesting a single origin of the striped alleles, whereas striped species of other lakes were not monophyletic (Fig. 3C). None of the nine mutations within enh.a that showed complete association with stripes in Lake Victoria cichlids showed similar stripe association in cichlids of Lakes Malawi or Tanganyika (figs. S12 and S13). Consequently, independent mutations must be affecting agrp2 expression and thereby stripe patterns across the three major cichlid radiations (Fig. 4D).

Lastly, we tested whether the same locus is responsible for stripe-pattern variation outside of Lake Victoria cichlids using a hybrid cross between the nonstriped Lake Malawi species Pseudotropheus demasoni (Pdem; Fig. 1M) and striped species *Pcya* (striped) and *Pcyx* (nonstriped) individuals that were genotyped at the agrp2 locus and phenotyped regarding their stripe patterns (Fig. 4, B and C; fig. S15; and supplementary text). The results revealed significant linkage between the agrp2 allele and stripe presence (Fisher’s exact test, $P = 7.6 \times 10^{-8}$; table S2). The allelic variation at the agrp2 locus explains more than 50% of the phenotypic variance in stripe patterns [Cox-Snell or Nagelkerke pseudo-$R^2$ from ordered logistic regression; table S3]. Nevertheless, the phenotypic distribution of F2 individuals (49 nonstriped and 221 striped individuals) differed from the Mendelian 3:1 ratio observed in the Lake Victoria crosses (chi-square test, $P < 0.001$), providing evidence for additional minor modifier loci. These results strongly suggest that agrp2 acts as a major determinant of stripe pattern absence or presence in Lake Malawi (as in the younger Lake Victoria radiation), but additional minor stripe modifiers have evolved or were recruited in the older Lake Malawi radiation (Fig. 4D).

The repeated evolution of horizontal stripes in East African cichlid radiations is facilitated by cis-regulatory evolution of agrp2. Despite its described role and function in the brain (18), we have discovered a hitherto unknown function for this gene in the skin that highlights notable functional similarities between Agrp2 and the mammalian Agouti (Asip) as well as telost Asipl (19, 20). From what is known about proteins of the Agouti family, Agrp2 likely acts as an antagonist for the melanocortin receptors Mc1r and/or Mc5r (21). Low Agrp2 levels would trigger stripe melanophore proliferation, pigment
loss as well as reevolution of stripes within East African cichlids. Although the presence of stripes appears to be controlled by differential expression of the same gene (agrp2), causal genetic variants must differ among the independent radiations of Lakes Victoria, Malawi, and Tanganyika (Fig. 4D and figs. S12 and S13). The intermediate phenotypes obtained from the Malawi cross (Fig. 4A), together with the lack of the dorso-lateral stripe in the CRISPR-Cas9 mutants (Fig. 2H), provide evidence for additional modifier loci determining stripe presence (Fig. 4D). However, those seem generally less prominent in the young (<15,000 years old) Lake Malawi radiation compared to the older (2 million to 4 million years old) Lake Malawi radiation (supplementary text).

Regulatory variation of agrp2 provides a molecular basis for the repeated evolution and loss of stripe patterns across cichlid species flocks. Recent regulatory evolution at the agrp2 locus constitutes an example of regulatory tinkering (22, 23) that might have facilitated the ease and speed of the evolution of both converged and diverged phenotypes that characterize the East African cichlid radiations. The simplicity of such a threshold mechanism might have permitted the phylogenetically observed rapid losses and reevolutions of stripe patterns. Therefore, Stephen Jay Gould’s predictions (I) appear questionable at this evolutionary scale, and if one were to replay the evolution of cichlid adaptive radiations, the results might be surprisingly similar: striped and nonstriped cichlids evolving again and again through regulatory evolution at the agrp2 locus.

REFERENCES AND NOTES

Fig. 3. Regulatory changes of agrp2 are predictive of convergent stripe evolution. (A) Densitree representation of phylogeny for the 24 examined species, including divergence time estimates (B). Dendrogram inferred with maximum likelihood (C). Error bars indicate means ± SD (D). A gene (locus) tree of the cis-regulatory element enh.a supports a single origin of striped alleles in Lake Victoria. Numbers present posterior probabilities >0.9. Ajac, Aulonocara jacobfreibergi; Ceeu, Cheilochromis echiulus; Dicom, Dimidiochromis compressiceps; Hmel, Haplochromis melanopterus; Hser, Haplochromis serranus; Hthe, Haplochromis thereuterion; Jmar, Julidochromis marlieri; Jreg, Julidochromis regani; Lmul, Lamprologinus multifasciatus; Mzeb, Maylandia zebra; Ncau, Neolamprologus caudopunctatus; Onl, Oreochromis niloticus; Pjoh, Placidochromis johnstoni; Ppun, Puntamilia pundamilia.

Fig. 4. A shared genomic basis of stripes across cichlid radiations. (A) Stripes are associated with agrp2 alleles in Lake Malawi hybrid F2 individuals (Pdem × Pcy). Inheritance is not Mendelian, suggesting that additional genetic modifiers exist in Lake Malawi cichlids. (B and C) F2 hybrids homozygous for the Pcy agrp2 allele (C/C, showing stripes, black lines) and Pdem allele (D/D, no stripes) exhibit clear stripe pattern differences. Parental species are shown in Fig. 1K (striped) and Fig. 1M (nonstriped). (D) Summary of the known (black) and unknown (gray) aspects of the genetic control of cichlid horizontal stripes.

dispersion, and/or pigment production (stripe patterns present), whereas high levels would block these processes (no stripe patterns) (22). Expression levels of agrp2 thereby act as a switch controlling stripe presence and absence. In Lake Victoria, expression-level differences seem to be caused by several mutations in a 1.1-kb intronic regulatory region (enh.a; Fig. 2B) that push the expression of agrp2 levels above or below a threshold that determines the stripe phenotype. Such a threshold-based molecular on-off switch may have permitted the frequent

ACKNOWLEDGMENTS
The authors thank S. Finkernagel for her exceptional technical assistance; A. Kautt, A. Nater, J. Torres-Dowdall, and the anonymous referees for insightful comments that improved the manuscript; and the University of Konstanz bioimaging center and animal facility. Funding: This work was supported by the Swiss National Science Foundation (P3BSP3_154629), the Marie Curie Zukunftskolleg Incoming Fellowship Program (291784), the Baden-Württemberg Foundation, the Deutsche Forschungsgemeinschaft (DFG, KR 4670/2-1), and the Young Scholar Fund of the University of Konstanz (all to C.F.K.). Y.L. is funded by the China Scholarship Council (CSC). A.M. and C.D.H. are funded by DFG and University of Konstanz. Partial funding came from an ERC Advanced Grant GenAdap (no. 293700) to A.M.

Author contributions: C.F.K. conceptualized the study; provided funding; planned and supervised all experiments; collected, analyzed, and visualized all data, if not otherwise indicated; and wrote the manuscript. Y.L. performed qPCRs and in situ hybridization. J.G. and S.U. conducted the Lake Malawi cross and assisted with the sequencing and enhancer assay. J.M.W. designed and generated the CRISPR-Cas9 knockouts and helped with the enhancer assay. F.H. and G.M.-S. provided data for Lake Victoria crosses and contributed to their analysis. C.D.H. conducted the phylogenetic analysis and edited the manuscript. A.M. conceptualized the study, provided funding, and revised the manuscript. All authors gave comments on the manuscript and approved the final version.

Competing interests: The authors declare no competing interests. Data and materials availability: Additional data have been deposited on Dryad (24) and are available in the supplementary materials.

SUPPLEMENTARY MATERIALS
www.sciencemag.org/content/362/6413/457/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S15
Tables S1 to S5
References (25–63)
16 October 2017; resubmitted 19 April 2018
Accepted 12 September 2018
Agouti-related peptide 2 facilitates convergent evolution of stripe patterns across cichlid fish radiations
Claudius F. Kratochwil, Yipeng Liang, Jan Gerwin, Joost M. Woltering, Sabine Urban, Frederico Henning, Gonzalo Machado-Schiaffino, C. Darrin Hulsey and Axel Meyer

Science 362 (6413), 457-460.
DOI: 10.1126/science.aao6809

Showing their stripes
The adaptive radiation of East African cichlids has led to more than 1200 species across a number of lakes. Across these species, many convergent traits have emerged, including the presence or absence of horizontal stripes. Kratochwil et al. show that the appearance or loss of stripes is related to changes in the agouti-related peptide 2 gene, which acts as a kind of on-off switch for stripe generation (see the Perspective by Gante). This action has enabled rapid and repeated evolution of stripes across this speciose radiation.
Science, this issue p. 457; see also p. 396