

Understanding explosive diversification through cichlid fish genomics

Walter Salzburger 

Abstract | Owing to their taxonomic, phenotypic, ecological and behavioural diversity and propensity for explosive diversification, the assemblages of cichlid fish in the East African Great Lakes Victoria, Malawi and Tanganyika are important role models in evolutionary biology. With the release of five reference genomes and many additional genomic resources, as well as the establishment of functional genomic tools, the cichlid system has fully entered the genomic era. The in-depth genomic exploration of the East African cichlid fauna — in combination with the examination of their ecology, morphology and behaviour — permits novel insights into the way organisms diversify.

Model organisms

Non-human species studied in detail in the context of a particular research question with the motivation to be able to make more general statements about the functioning of organisms.

Clades

Branches on an evolutionary tree, consisting of a common ancestor and all its descendants (a clade is, hence, equivalent to a monophyletic group).

Why is species richness so unequally distributed across the tree of life? Why did some organismal lineages diversify into new forms in a seemingly explosive manner, whereas others have lingered phenotypically unvaried over millions of years? These questions have puzzled generations of biologists ever since Darwin and Wallace jointly introduced their theory of evolution by natural selection¹. 160 years of scholarly study later, there is a reasonable understanding of how and under which circumstances new species can originate^{2–7}. However, the causal factors that determine species richness and the rate at which new species form remain largely elusive^{8,9}. Particularly in light of the global biodiversity crisis that our planet is currently facing¹⁰, it is no longer of purely academic interest to know how novel species form and, consequently, how biodiversity arises.

Unravelling how variation at the genomic level is interlinked with phenotypic evolution is key to understanding organismal diversification^{2,11–13}. To this end, we must understand how organisms evolve, how they function and how they interact with other organisms and the environment. The problem is that many widely used model organisms provide limited insights into the underpinnings of rapid — by way of comparison — organismal diversification; many traditional laboratory-based model organisms tell us little about how organisms adapt, behave and diversify in the wild, while model species in ecology and evolution often lack tractability in the laboratory and fundamental data on genomics and development. Importantly, most established model organisms do not belong to extensively diversifying clades.

Instances of adaptive radiation — that is, the rapid (sometimes ‘explosive’) origin of taxonomic, ecological

and morphological diversity as a consequence of adaptation to novel or hitherto underutilized ecological niches^{14,15} — combine the advantages of laboratory and natural model species in the context of the genesis of biodiversity. Therefore, iconic examples of adaptive radiation, such as Darwin’s finches on the Galapagos archipelago, anole lizards on the islands of the Caribbean, threespine stickleback fish in post-glacial rivers and lakes, and cichlid fish in East Africa (BOX 1), have long been recognized as essential models to study organismal diversification^{12,16–18}. Scientific interest in many of these radiations can be traced back to the 19th century, such that, for a long time, these groups have been intensely investigated with respect to their evolution, ecology, ontogenetic development and behaviour. The close relatedness of the species emerging from adaptive radiations facilitates genetic and genomic investigations¹², for example, on the basis of hybrid crosses or divergence mapping. Moreover, representatives of these adaptive radiations were among the first vertebrates to have their genomes sequenced^{19–23}.

The species flocks of cichlid fish in the East African Great Lakes Victoria, Malawi and Tanganyika represent the most species-rich and phenotypically diverse adaptive radiations in vertebrates and are characterized by exceptionally fast diversification rates^{18,24,25} (BOX 1; FIG. 1). To put cichlid radiations into a temporal context, during the evolutionary time span of our own species, starting with the split between chimpanzees and humans some 5–7 million years ago, approximately 2,000 species of cichlid fish evolved in East Africa, the geographic region where the chimpanzee–human split initially occurred. Within the time span that it took for 14 species

Zoological Institute,
Department of Environmental
Sciences, University of Basel,
Basel, Switzerland.

e-mail: walter.salzburger@unibas.ch

<https://doi.org/10.1038/s41576-018-0043-9>

Box 1 | What is a cichlid?

The family Cichlidae (Bonaparte, 1835) is one of 536 families in fish¹¹⁷. Cichlidae, together with only two species of convict blennies in the family Pholidichthyidae, make up the percomorph order Cichliformes in the series Ovalentaria^{118,119}. Etymologically, the family name descends from the Greek word κίχλη ('kiche') for thrushes and some marine wrasses^{120,121}. The oldest known written record of a cichlid is a hieroglyph in ancient Egypt specifically assigned to the Nile tilapia (*Oreochromis niloticus*)¹²². The first cichlid artefacts also appeared in ancient Egypt, displaying the mouthbrooding tilapia as a symbol of rebirth¹²³. Another tilapia, *Sarotherodon galilaeus*, embodies the biblical St. Peter's fish in the Sea of Galilee.

Cichlids inhabit fresh and brackish waters in — from East to West — southern India and Sri Lanka, Madagascar, the Middle East, Africa, the West Indies and the Americas (South and Central America, Mexico and Texas). Wherever they occur, cichlids show a strong tendency to form species flocks with high levels of endemism (the exclusive occurrence of a species or higher-level taxon in a confined geographic area)^{31,124–126}. Species flocks are conceptually related to the phenomenon of adaptive radiation because most, if not all, species flocks are the outcome of such radiations^{35,85,127}. Cichlids show substantial diversity across different levels of biological organization:

Taxonomic diversity. Approximately 1,700 cichlid species in more than 200 genera are recognized to date^{117,121}, but many more species have yet to be described. The total number of cichlid species has been estimated in the range of 3,000–4,000, suggesting that roughly every tenth teleost fish species is a cichlid. More than half of this estimated taxonomic diversity is allocated to the cichlid faunas of the African Great Lakes Tanganyika (250 species), Malawi (800–1,000 species) and Victoria (700 species, including its surrounding water bodies)^{25,39,85,128}.

Morphological diversity. Morphologically, cichlids are characterized by a single nostril on each side of their snout rather than two as in most other fish; their lateral line is interrupted; their dorsal and anal fins consist of spines (7–25 and 3–15, respectively) and soft rays (3–15 and 4–30, respectively); and they possess a characteristic pharyngeal jaw apparatus, a second set of jaws in the pharynx that is functionally decoupled from the oral jaw apparatus, for processing food^{117,121}. Although none of these characteristics is unique to cichlids, their combination is. Cichlids vary widely in body shape — from rounded, as exemplified by the freshwater angelfish (*Pterophyllum scalare*) or the discus fish (*Symphysodon discus*), to elongated, as in *Crenicichla percna* — as well as in body size — from less than 3 cm and 0.1 g in, for example, *Telmatochromis bifrenatus* to >80 cm and >3 kg in *Boulengerochromis microlepis*, both from Lake Tanganyika. Above all, cichlids differ in ecologically relevant traits, such as the mouth apparatus and head shape^{17,61}.

Ecological diversity. Cichlids occupy a wide range of feeding niches, ranging from eating algae, sponges, invertebrates or other fish to mollusc crushing or scale eating, and habitat types, ranging from small streams or rivers to lakes and, within lakes, from the shallow to the deep benthic and limnetic zones. Cichlids occur on rocky, sandy, weedy and muddy substrates, yet others are found in caves or use empty snail shells as shelter^{17,18}. As in other cases of adaptive radiation, there is a strong correlation between phenotype and environment in cichlids^{17,61}.

Colour diversity. Cichlids are highly diverse with respect to body coloration and pigmentation, hence their German name 'Buntbarsche' (colourful perches). Coloration in cichlids has been implicated in sexual selection via female choice with respect to male nuptial coloration¹²⁹ but also with mimicry^{17,130} and camouflage¹⁷. Not least because of their conspicuous coloration, cichlids are popular with aquarists.

Behavioural diversity. Cichlids show a very broad behavioural repertoire. Many cichlid species are territorial, others form large schools. All cichlids show some kind of brood-care behaviour, ranging from substrate spawning to maternal, paternal or biparental mouthbrooding^{17,18,120}.

of Darwin's finches to evolve on the Galapagos archipelago²², about 1,000 cichlid species evolved in Lake Malawi alone^{26–28}. In addition, since the last ice age, which is when sticklebacks began to diverge into replicate species pairs in the Northern hemisphere²⁰, hundreds of cichlid species evolved in Lake Victoria^{29,30}.

In this Review, I discuss how the examination of recently available genome-wide sequence data of East African cichlids has deepened our understanding of the phenomena of adaptive radiation and explosive diversification in general and in cichlids in particular. I start with a discussion of the explosive nature of species formation in East African cichlids and the resultant difficulties in delineating species. Then, I focus on the challenges that emerge in the reconstruction of the evolutionary history of rapidly diversifying clades at the interface between population genetics and phylogenetics. Finally, I summarize what we have learned about the genomes of East African cichlids thus far and discuss which features in their genomes are potentially linked to their propensity to diversify explosively.

Explosive diversification in cichlids

The adaptive radiations of cichlid fish in Lakes Victoria, Malawi and Tanganyika (FIG. 1) differ from all other cases of adaptive radiation in vertebrates — including those of cichlids elsewhere — by their unparalleled degree of phenotypic and taxonomic diversity in sympatry^{24,31}. That cichlids are unusual had been realized already over the course of the earliest biological explorations of the African Great Lakes in the late 19th century^{32,33}. Ever since, researchers repeatedly came to the conclusion that the current understanding of speciation was insufficient to explain the plethora of cichlid species in East Africa^{17,34,35}. With specific reference to the cichlids, Woltereck in 1931 (REF.³⁴) introduced the term 'Artexplosion' for outbursts of endemic diversity on islands and in some ancient lakes. More widely known as explosive speciation, this term refers to a substantial increase in speciation rate in a clade relative to a comparable group, irrespective of ecological and morphological differentiation, and is thus distinct from the phenomenon of adaptive radiation³⁶.

Divergence mapping

Genetic-marker-based search for genomic regions exhibiting exceptionally strong differentiation between different biological entities, such as populations or species.

Species flocks

Unusually taxon-rich assemblages of closely related species that coexist in the same area, such as an island, a lake or a section of a river.

Lateral line

A sensory system along the body in aquatic vertebrates that consists of sensory cells.

Sympatry

The existence of two or more species in the same geographic area such that they encounter each other frequently.

Reproductive isolation

Any property that prevents (or reduces the probability of) members of one species breeding successfully with members of another species.

Gene flow

The movement or exchange of genes into or through a population by interbreeding or by migration and subsequent interbreeding.

Speciation continuum

The bandwidth of variation between diversifying populations ranging from virtually no variation (panmixia) through partially discontinuous variation (incipient barriers to gene flow) to strongly discontinuous variation (complete reproductive isolation).

Tribes

The taxonomic rank between the genus and the family level.

Sister taxa

Reciprocally closest relatives of one another.

Ring species

Two reproductively isolated populations are connected through a geographic ring of populations that interbreed; no morphological character can be used, except arbitrarily, to divide the ring into discrete taxonomic units.

Biological species concept

A concept positing that species are groups of actually or potentially interbreeding natural populations that are reproductively isolated from other such groups.

Haplotype

A stretch of DNA on a single molecule (chromosome, plastid or mitochondrion) that is inherited as a single unit.

Phylogenetic species concept

A concept positing that species are 'tips' on a phylogeny, that is, the smallest set of organisms that share an ancestor and can be distinguished from other such sets.

The process of species formation

Generally speaking, speciation is the formation of a new species that is distinct from all other species; it is commonly defined as the build-up of reproductive isolation between an ancestral species and the newly formed species². Most of the earlier work on this topic has focused on the role of geographical separation in the origin of species (reviewed in REF.²). More recently, the field has shifted towards a more process-oriented approach, emphasizing the importance of ecology in speciation^{5,6} — through divergent natural selection in distinct environments — and establishing that speciation is possible, perhaps even common, in light of some levels of gene flow between the diversifying lineages^{37,38}. Both of these features, ecological speciation and gene flow, are particularly common in adaptive radiations, in the course of which new species typically form in the absence of geographical barriers^{14,15}.

Speciation is often (but not always) a gradual process that has a clearly defined starting point — a single species — and is completed when at least one new species (if the ancestral species continues to exist) or a minimum of two species (if the ancestral species becomes extinct) has emerged. In between, along the so-called speciation continuum, it is impossible to conclusively determine whether there is one or more than one species. Any consideration of the process of speciation is therefore inextricably interwoven with the questions of what a species is and how it can be distinguished from other, closely related ones.

Species delineation in cichlids

The delineation of species is not an easy undertaking in cichlids. Although species belonging to different lineages (that is, tribes or genera) in the longer persisting cichlid radiations of Lakes Tanganyika and Malawi are phenotypically and ecologically clearly distinct from one another (FIG. 1), the status of species within such lineages as well as within younger species flocks is often unclear. This is, in part, due to the sheer number of species, which makes it difficult for taxonomists to keep track of taxonomic units and to sort species according to diagnostic characteristics^{39,40}. However, the main issues in deciding what a species is in cichlids emerge from their close relatedness and the fact that radiations are still ongoing: morphologically distinct sister taxa are typically of very recent origin^{4,30,41,42} (in the range of a few hundred to a few thousand years), and, in many cases, it is not clear whether they have reached the end of the speciation continuum. In other cases, geographically isolated sister taxa are connected through intermediate forms^{43–45}, much as in the case of ring species, rendering it impossible to define clear boundaries between them. By contrast, closely related cichlids of uncertain taxonomic status often mate assortatively with respect to their own source population^{46–48}, that is, individuals mate more frequently with members of their own population than expected under a random pattern, or even occur in sympatry in parts of their distribution ranges^{43,44} (FIG. 2), suggesting that these sister taxa are valid species.

The classic species concepts provide little guidance for species delineation in cichlids. The most widely used definition in biology for the category species, the biological species concept^{49,50}, is not very practical when

applied to cichlids and often fails on the basis that reproductive isolation is usually incomplete between sister taxa. In fact, many East African cichlid species are intercrossable^{51,52}, even when belonging to distinct phylogenetic lineages and being derived from different adaptive radiations^{53,54}, and cichlids do interbreed in the wild as evidenced by occasionally observed hybrid specimens⁵⁵ as well as molecular analyses demonstrating substantial levels of gene flow between species^{41,56,57}. Delineating cichlid species by means of genetic markers is problematic too. DNA barcoding, a widely used method for identifying species on the basis of the mitochondrial *COX1* gene⁵⁸, performs poorly when applied to cichlids⁵⁹. This is not surprising, given the high levels of DNA sequence similarity in cichlids (for example, the average genome-wide sequence divergence between Lake Malawi cichlids is only 0.1–0.25%)²⁷ as well as mitochondrial haplotype sharing between species^{42,43}. Defining species according to the phylogenetic species concept is equally problematic, as there is no a priori level of genetic distinctiveness above which two sister taxa should be considered different species and because reciprocal monophyly of sister taxa does not help in deciding whether these are populations of the same species or different species. Grouping individuals (or lineages thereof) into species according to their shared ecology — as suggested in the ecological species concept — is difficult in cichlids, as there is substantial niche and resource overlap and, hence, little competitive exclusion (whereby two species cannot stably coexist in the same ecological niche) between some species^{17,60,61}. In practice, to facilitate the expedient naming of distinct taxonomic units in cichlids, species are seen as clusters of individuals that are morphologically and ecologically similar and distinct from other such clusters (that is, the vernacular species concept)⁶².

Taken together, there is no straightforward way of species delineation in cichlids. Phenotypically distinct yet closely related cichlids within species flocks may perhaps best be characterized as multispecies, that is, sets of closely related species that co-occur and that occasionally exchange genes⁶³. This is exemplified by two closely related *Pundamilia* species from Lake Victoria, *Pundamilia nyererei* and *Pundamilia pundamilia*, which, when co-occurring, are genetically more similar to one another than two geographically separated (allopatric) populations of the same species^{4,41}. Importantly, what makes species delineation so difficult in cichlid species flocks — namely, that there are many species, that species are very young and that speciation is often 'caught in the act' — is exactly what makes cichlids such a useful model for the study of the diversification process. Note that to account for the situation in which, in many cases, speciation is not yet complete, one should use the more general term 'diversification' instead of 'speciation' in the context of the adaptive radiations of East African cichlids, as the latter designation requires completion of the process.

Reconstructing cichlid evolution

The availability of accurate, comprehensively sampled and time-calibrated phylogenetic hypotheses is crucial to understanding the progression of adaptive radiations and explosive diversification⁶⁴, as well as to interpret



Nile tilapia
(*Oreochromis niloticus*)

Haplochromis sp. *Haplochromis* sp. *Pundamilia* sp. *Paralabidochromis sauvagei* *Haplochromis aeneocolor*

Haplochromis paropiis *Pundamilia nyererei* *Haplochromis thereuterion* *Yssichromis piceatus* *Haplochromis* sp. 'Victoria Nile'

Lake Victoria region | ~700 species

Lake Tanganyika | 250 species

Altolamprologus sp. 'Sumbu shell' *Neolamprologus brichardi* *Ctenochromis horei*

Lamprologus signatus *Neolamprologus longior* *Petrochromis* sp. 'Texas'

Lepidolamprologus elongatus *Astatotilapia burtoni* *Julidochromis transcriptus*

Asprotilapia leptura *Greenwoodochromis christyi* *Tropheus moorii*

Eretmodus cyanostictus *Boulengerochromis microlepis*



Lake Malawi | 800-1,000 species

Cheilochromis euchilus *Dimidiochromis compressiceps* *Placidochromis* sp. *Exochromis anagenys* *Maylandia zebra* *Nimbochromis livingstonii*

Astatotilapia calliptera *Rhamphochromis esox* *Labeotropheus fuelleborni* *Melanochromis auratus* *Pseudotropheus johannii* *Lichnochromis acuticeps*

◀ Fig. 1 | **The species flocks of cichlid fish in East Africa.** The lakes and rivers in the area harbour diverse assemblages of endemic cichlids, of which the species flocks in Lakes Tanganyika and Malawi and in the Lake Victoria region are, by far, the most species-rich. Dotted boxes highlight the five species that were selected for the initial round of genome sequencing¹⁹. Images courtesy of W. Gessl, University of Graz, Austria.

general patterns of these phenomena (BOX 2). At the same time, the close relatedness of species within adaptive radiations, occasional gene flow between species and the pace at which new species form pose considerable challenges to the phylogenetic reconstruction of rapidly diversifying clades^{22,27,64–66} (FIG. 3), even if genome-wide data are available.

The first cichlid genomes

The initial round of genome sequencing in cichlids involved a set of five phylogenetically representative African species¹⁹: the Nile tilapia (*Oreochromis niloticus*), as a member of a less species-rich yet geographically widespread sister lineage to the cichlid radiations in the East African Great Lakes; the Princess of Burundi (*Neolamprologus brichardi*), from the most species-rich tribe within Lake Tanganyika, the Lamprologini; and three members of the tribe Haplochromini, namely, Burton's cichlid (*Astatotilapia burtoni*) from Lake Tanganyika and nearby rivers, the Zebra mbuna (*Metriaclima zebra*) from the mbuna clade of Lake Malawi and *P. nyererei* from Lake Victoria (FIGS. 1, 4). To facilitate genome annotations, the study of Brawand et al.¹⁹ also established reference transcriptomes for these five species. In the meantime, many more genomes of East African cichlids have been sequenced at low coverage with short-read (Illumina) sequencing approaches^{27,41,57,67,68}. The assemblies of two of the reference genomes, the Nile tilapia and the Zebra mbuna, have been much improved with long-read (PacBio) sequencing^{69,70}, and additional transcriptomes from more species and more tissues have been generated^{71–75}. These advances have confirmed that the genomes of rapidly diversifying cichlids are frequently subjected to incomplete lineage sorting and introgression (FIG. 3c,d), resulting in mosaic genomes that consist of small segments with different evolutionary histories^{12,19,27,57} (FIG. 3g).

The phylogeny of East African cichlids

Incongruence between gene trees and between gene trees and the species tree is particularly pronounced in East African cichlids. For example, analysis of the initially sequenced cichlid reference genomes revealed that more than 40% of all single nucleotide polymorphisms (SNPs) support topologies among the three representatives of the tribe Haplochromini that are in conflict with the species tree¹⁹. Phylogenetic analyses of thousands of segments in the genomes of 5 closely related *Neolamprologus* species from Lake Tanganyika showed that all of the 15 possible topologies connecting the 5 species received support from at least a few dozen segments in the genome, whereby one topology stood out as being supported by about half of all segments⁵⁷. In addition, the analysis of 2,543 non-overlapping windows, each containing 8,000 SNPs, across the genomes of 73 Lake Malawi cichlid species produced 2,542 different topologies²⁷. Obviously, in such situations, a single tree-like phylogeny can no longer capture the

entire evolutionary history of a group, making the concept of a clade and the quest for bifurcating branching diagrams questionable.

The analysis of genome-wide markers nevertheless provides novel insight regarding evolutionary relationships in cichlids, not least because incomplete lineage sorting and (introgressive) hybridization can now be looked at on a genome-wide scale and across many taxa^{26,27,76}. By and large, the phylogeny of the East African cichlids reflects the dynamic geological history of the area (FIG. 4). Lake Tanganyika, the oldest and geologically most stable of the African Great Lakes³¹, is home to the phenotypically most diverse cichlid species flock¹⁷. Taxonomically, the Tanganyikan cichlids have been grouped into 14 tribes, which differ substantially in species number (ranging from 1 species in, for example, Boulengerochromini to about 100 species in Lamprologini)^{61,77}. Some tribes — even if endemic to Lake Tanganyika today — must have evolved elsewhere, that is, before the formation of the present lake some 9–12 million years ago (notably, the Bathybatini, Boulengerochromini and Trematocarini), whereas the origin of other tribes is compatible with a scenario of in situ evolution^{26,78,79}. Although the respective monophyly of these tribes is usually well supported in phylogenetic analyses using genome-wide data^{26,78–80}, there is strong evidence for past gene flow between some of them^{26,78,79} (FIG. 4).

The radiations in Lakes Malawi and Victoria involve only one of the African cichlid tribes, the Haplochromini (note that in all three African Great Lakes, a few members of the Oreochromini are found, which did not radiate). In the early stages of the cichlid adaptive radiation in Lake Malawi, three major clades emerged in closely timed lineage-splitting events: a pelagic clade formed by *Diplotoxodon* and *Rhamphochromis*; a clade including shallow and deep benthic species as well as the utaka lineage; and the mbuna clade. The presumed ancestor of the radiation, the generalist species *Astatotilapia calliptera*, continued to exist in rivers and lakes in the area (including Lake Malawi)²⁷ (FIG. 4); this species is, hence, phylogenetically nested within the Lake Malawi cichlid radiation^{27,81}. The inspection of genome-wide data further revealed multiple events of gene flow within and between the main cichlid clades in Lake Malawi²⁷. The current age estimates for the onset of the Lake Malawi cichlid radiation^{26,27} are compatible with a time-calibrated paleoecological record⁸², which revealed that the lake transitioned into its current state as a more or less closed system with deepwater habitats approximately 800,000 years ago, making cichlid diversification into deepwater habitats possible only afterwards²⁸.

The situation in Lake Victoria differs somewhat given that the cichlid fauna of the lake is part of a geographically more extended species assemblage, the so-called Lake Victoria region super-flock, which includes the adaptive radiations of cichlids in Lakes Victoria, Edward, Albert and Kivu, among others^{42,76} (FIG. 1). The onset of the diversification of the super-flock has been estimated at 100,000–200,000 years ago^{26,42,76,83}, and most, if not all, species within Lake Victoria must have evolved within the past 15,000 years following its refill after complete

Monophyly

The condition of forming a monophyletic group (that is, a clade).

Ecological species concept

A concept positing that species are sets of organisms that are adapted to a particular set of resources, that is, to the same ecological niche.

Incomplete lineage sorting

The imperfect segregation of a gene into all evolutionary lineages, that is, a gene fails to coalesce within the duration of a species.

Introgression

Also referred to as introgressive hybridization. The transfer of genetic material from one species into another via repeated and asymmetrical backcrossing with one of the parental lineages after a hybridization event.

Mosaic genomes

Genomes heterogeneous in ancestry, emerging from introgression, incomplete lineage sorting or lateral gene transfer.

Gene trees

Actual evolutionary relationships between the versions of a gene as present in different taxa.

Species tree

Actual evolutionary relationships between a set of species. The species tree reflects the 'true' evolutionary history of a clade.

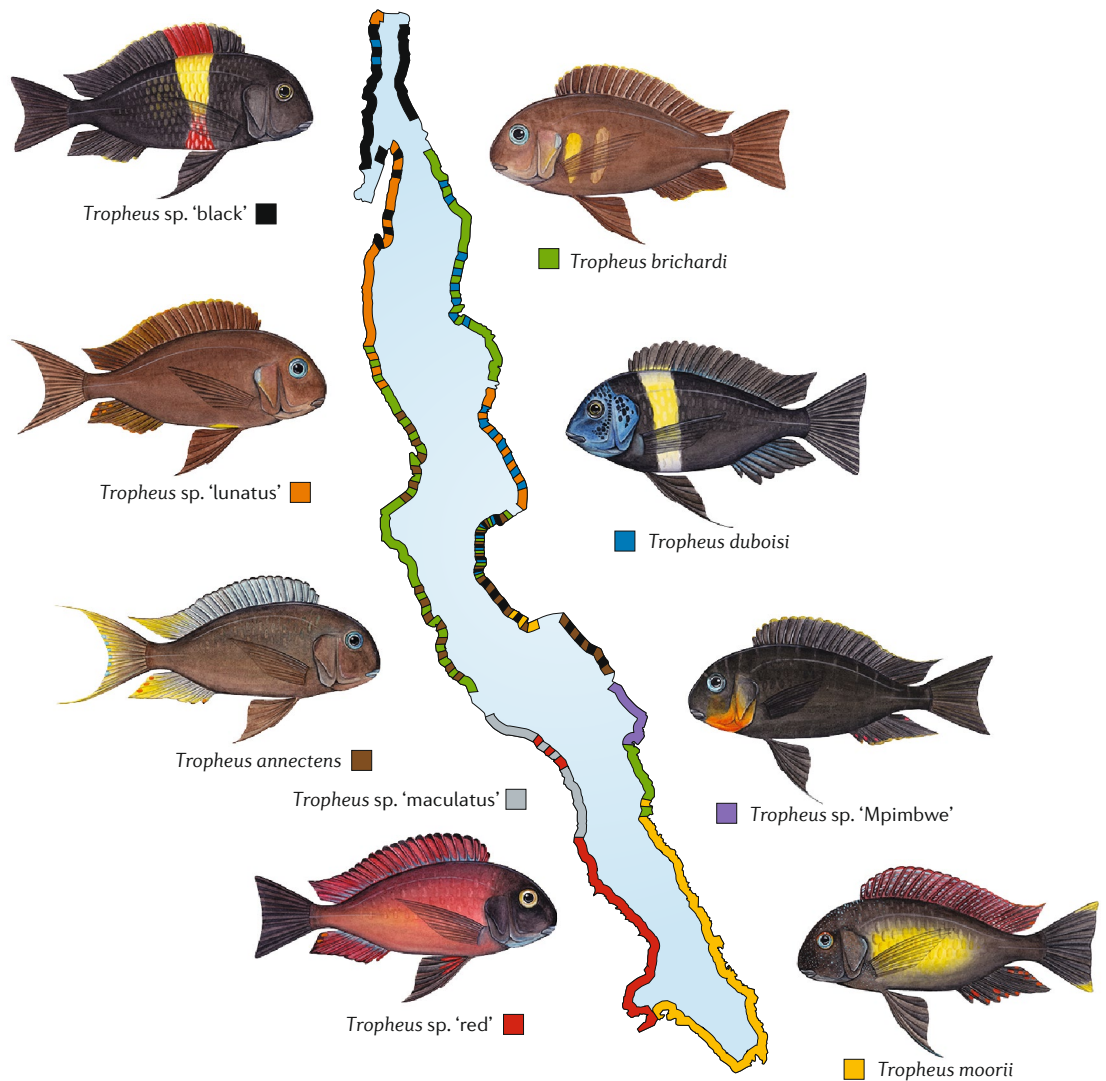


Fig. 2 | The *Tropheus* species complex from Lake Tanganyika. At Lake Tanganyika, basically every stretch of rocky shoreline is inhabited by its own colour morph of *Tropheus*, with more or less sharp boundaries of gene flow between adjacent populations^{44,135}. Six nominal species are currently described in this genus, but a few more await formal description¹³⁶. Up to four different taxonomic units (formally described species, undescribed species and local colour morphs) occur in sympatry at certain locations, suggesting that these are valid species. Distribution ranges are according to REFS^{45,136}. Drawings courtesy of J. Himes, University of Basel, Switzerland.

desiccation in the late Pleistocene^{29,30}. The extremely young age of the roughly 700 species in this super-flock thus makes classic phylogenetic analyses difficult. The most thorough analysis so far on the basis of restriction-site associated DNA (RAD) sequencing suggests that not all radiations in the lakes of the Lake Victoria region are reciprocally monophyletic and that hybridization played a key part in triggering these radiations⁷⁶.

Genomic basis of cichlid diversification

The phenomenon of explosive diversification in cichlids has long been implicated with the particular environment in which cichlid evolution has taken place, as well as with intrinsic features of the cichlids themselves^{17,35,84–86}. Without a doubt, the ample ecological opportunity provided by the African Great Lakes is conducive for diversification through adaptive radiation, not only in cichlids but also across different

groups of animals³¹, which is corroborated by the scaling of the number of endemic species in a lake with such a size, relative stability and depth^{31,84,87}. Still, the number of endemic cichlid species in Lakes Victoria, Malawi and Tanganyika is at least an order of magnitude higher than the number of endemic species in any other family resident in these lakes (except, perhaps, ostracods in Lake Tanganyika)³¹, and none of approximately 20 non-cichlid fish families that also occur in each of these lakes has brought forth more than a handful of endemic species^{17,88}. These findings suggest that there is something special about cichlids — some kind of ‘cichlidness’ — that permits these fish to diversify explosively. That cichlids in general feature an intrinsic propensity for diversification is substantiated by numerous examples of smaller-scale (in comparison to the ones in the African Great Lakes) adaptive radiations in rivers and lakes throughout their distribution range^{84,89–93}, of which the

Ecological opportunity
The availability of ecologically accessible resources that may be exploited by an evolutionary lineage.

Box 2 | General patterns in (cichlid) adaptive radiations

Adaptive radiations in different groups of organisms are thought to follow the same general patterns, which include an early burst of phenotypic divergence, a taxonomic overshooting and a progression in stages (reviewed in REF.¹⁵). In the adaptive radiations of cichlids in East Africa, early bursts of divergence are a likely scenario (given the generally young age of the species flocks) but are difficult to prove in the absence of independent sources of information, such as fossils in chronological resolution. The phenomenon of overshooting due to decreasing speciation and/or increasing extinction rates along the course of an adaptive radiation might explain why there are more cichlid species in the lake of intermediate age, Lake Malawi, than in the younger (Lake Victoria) and the older (Lake Tanganyika) lakes³⁰ (FIG. 1). However, differences in (estimated) species numbers between the African Great Lakes also reflect the different taxonomic strategies that have been deployed when assigning taxa into species, and overall morphological disparity is greatest in Lake Tanganyika^{17,31}. Similar to the early-burst scenario, it is difficult to test the overshooting effect on the basis of phylogenomic reconstructions alone. The so-called stage model of adaptive radiation posits that diversification initially occurs with respect to macrohabitat adaptations and then with respect to microhabitat specializations (trophic adaptations), after which disruptive sexual selection based upon mating traits becomes the dominant mode of speciation^{15,131,132}. More recent work has cast doubts on the validity of the stage model in this chronological order, suggesting instead that trophic traits diversified earlier than traits implicated in macrohabitat adaptation^{97,133}. One of the most fascinating aspects of adaptive radiation is the frequent occurrence of convergent forms¹². This phenomenon is particularly pronounced within cichlids, where convergent evolution is prevalent and occurs not only between^{75,134} but also within⁶¹ species flocks.

ones in small volcanic crater lakes in Africa and Central America are the most widely perceived^{3,67,94–96}. Early on, the extraordinary diversity of cichlids has spurred speculations that particular genomic features might underlie their propensity to diversify³⁵. Examining this ‘genomic substrate’ for diversification has become one of the main research targets within the cichlid genome project¹⁹.

Comparative cichlid genomics

The in-depth comparative analysis of the initial cichlid genome sequence data¹⁹ identified several distinctive features in the genomes of the radiating East African cichlids that could potentially — individually or jointly — be responsible for explosive diversification in this group, thereby confirming earlier assumptions. First, the cichlid genomes turned out to be genetically more diverse than expected in light of the very recent origin of the species flocks, which was attributed mainly to the accumulation of standing genetic variation before the radiations (as opposed to new mutations). Second, the four genomes of the explosively diversifying cichlid lineages from Lakes Victoria, Malawi and Tanganyika showed an increased rate of gene duplications compared with the Nile tilapia and other teleosts, and about one-fifth of the duplicated genes showed evidence of neo-functionalization, the gain of a new function in one of the gene copies. Third, the same four cichlid genomes are characterized by a greater dynamic in gene regulatory processes than observed in other fish species, as evidenced by increased rates of regulatory element evolution and novel and functionally diverse microRNAs. Fourth, these genomes exhibit accelerated coding sequence evolution, as evidenced by elevated rates of non-synonymous to synonymous (dN/dS) substitutions compared with the rates in the Nile tilapia. Finally, three waves of transposable

element expansions were detected in cichlids. Many more cichlid genomes have been inspected since, calling for a revisit to the question of what the genomic underpinnings of adaptive radiation and explosive diversification in cichlids are.

Sources of genetic variation

There is a general consensus that divergent natural selection, partly in combination with sexual selection, has played a key part in the diversification of cichlids^{18,19,31,84,97}. Heritable phenotypic variation in fitness-related traits is a prerequisite for selection to operate. This variation can arise from mutations at the level of single nucleotides, of genes or of chromosomes and via the reshuffling of existing genetic material during meiosis (recombination), making these sources of variation at the molecular level prime targets in the quest for the genomic basis of explosive diversification in cichlids.

Mutation rate. Interestingly, the nucleotide mutation rate is not very high in East African cichlids. Quite to the contrary, the nucleotide mutation rate has been estimated — using trio sequencing — at 3.5×10^{-9} per bp per generation (TABLE 1), which is 3–4-fold lower than the rate in humans²⁷. Therefore, the exceptionally high speciation rates in East African cichlids cannot be linked to an elevated mutation rate.

Gene duplications. The rate of lineage-specific gene duplications, by contrast, has been found to be 4.5–6-fold higher in the common ancestor of the explosively diversifying East African cichlid lineages than in other fish; it is highest in the common ancestor of the most species-rich lineage in cichlids, the Haplochromini¹⁹, suggesting a link between gene duplication and diversification.

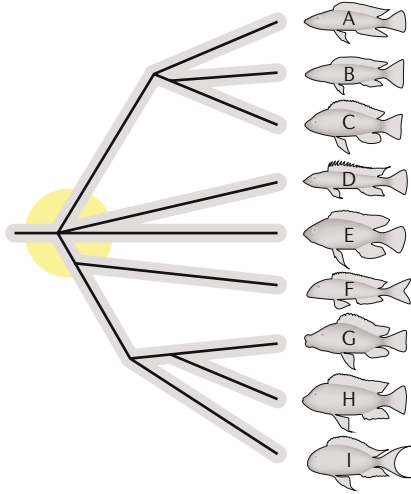
Transposable elements. Transposable element insertions provide another source of genetic variation, and there is evidence from the cichlid genomes that the insertion of transposable elements near genes has altered their expression¹⁹. This is exemplified by a case study showing that the insertion of a short interspersed nuclear element (SINE) upstream of a previously unknown colour gene, *fhl2b*, in the ancestor of the Haplochromini lineage led to a gain of expression in iridophores (a specific type of pigment cells), which in turn has been linked to the origin of a new pigmentation trait in this group⁷¹.

Chromosome numbers. At the level of entire chromosomes, there is little variation among the East African cichlids, with chromosome numbers ranging from $2n = 40$ to $2n = 46$ according to karyotyping; most species have $2n = 44$ (REFS^{98,99}) (TABLE 1). Differences in chromosome number, thus, do not seem to have an important role in the origin or maintenance of cichlid species. The contribution of smaller-scale chromosomal rearrangements (such as inversions) to cichlid diversification is less clear to date, as its examination is currently hampered by the insufficient quality of genome assemblies based on short-read sequence data, and too few genome assemblies based on long-read data in combination with genetic mapping are currently available.

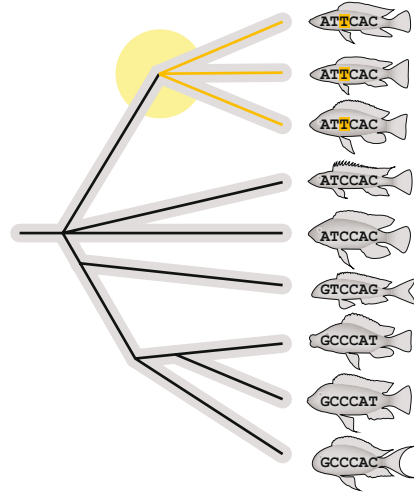
Trio sequencing

Whole-genome sequencing of two parents and one of their offspring, allowing accurate phasing (and the determination of the nucleotide mutation rate if sequence coverage is high enough).

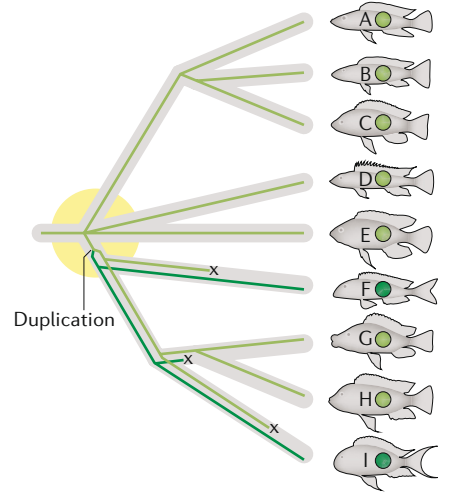
a Hard polytomy



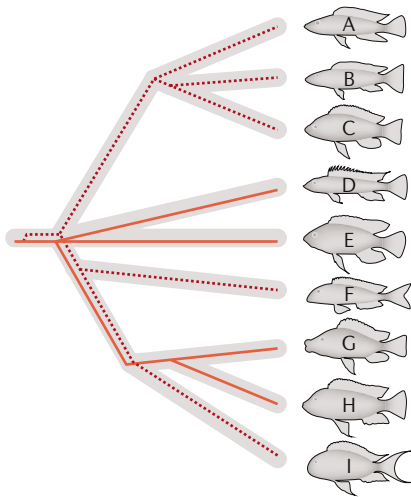
b Soft polytomy



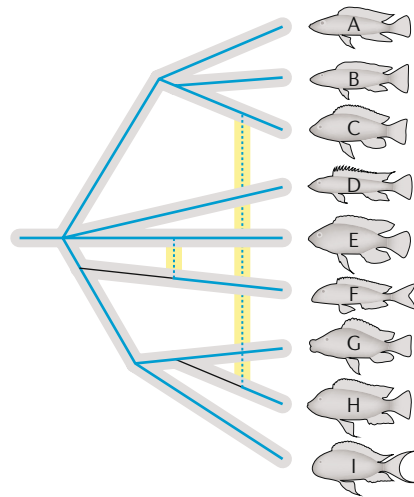
c Gene duplication or loss



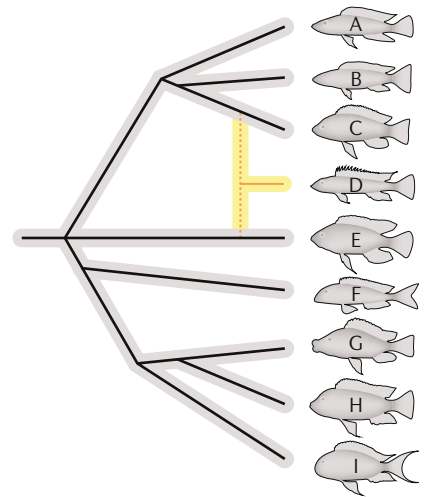
d Incomplete lineage sorting



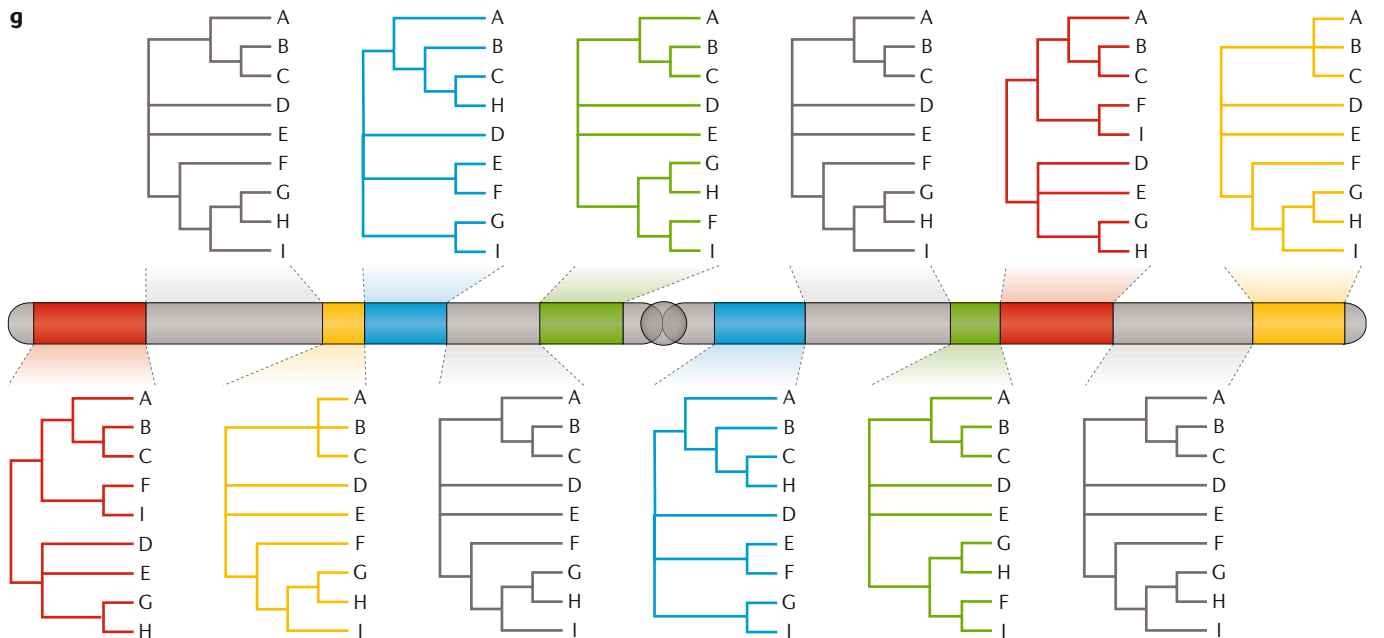
e Introgression



f Hybridization



g



◀ Fig. 3 | **Phylogenetic inference in rapidly diversifying clades.** A number of factors contribute to discordance among individual gene trees and between gene trees and the species tree, resulting in mosaic genomes. **a** | In the event of an adaptive radiation, particularly if novel species emerge through explosive diversification, the species tree is not a priori expected to be strictly bifurcating (neither are the underlying gene trees). This is because an ancestor may give rise to more than two daughter species at a given time and continue to persist as its own species. In this case, polytomies in the species tree and in gene trees are 'hard' and reflect true multifurcations. **b** | In the course of an adaptive radiation, new species form rapidly (in some cases explosively). As a consequence, there is often insufficient time for the accumulation of branch-specific diagnostic nucleotide substitutions between consecutive speciation events, leading to non-bifurcating gene trees ('soft' polytomies) or gene trees with very short internal branches^{64,137}; in many such cases, more mutations accumulate after rather than during speciation events, resulting in bottom-heavy phylogenies¹³⁸. **c** | The duplication of genes and the subsequent (stochastic) loss (marked with x) of one of the gene copies — a phenomenon that is known to occur in adaptive radiations¹³⁹ — can also lead to phylogenetic incongruence. **d** | When speciation events are closely timed, which is prototypical of both adaptive radiation and explosive diversification, ancestral genetic variation can stochastically persist as polymorphisms between speciation events. This phenomenon — called incomplete lineage sorting — causes discordance between individual gene trees as a function of when in the past the respective genes coalesce; incomplete lineage sorting also causes discordance between the gene trees affected by incomplete lineage sorting and the species tree. **e** | Closely related species that coexist in the same area — a defining feature of species flocks, adaptive radiations and instances of explosive diversification — are prone to the exchange of genetic material via introgression^{12,22,140}. Introgression (or introgressive hybridization) leads to incongruence between gene trees in such a way that a larger proportion of gene trees is equivalent to the species genealogy, while a smaller fraction of gene trees reveals the donor lineage of the genetic material that introgressed^{57,141}. **f** | In some instances, hybridization may lead to the formation of a new species⁷, in which case about half of the gene trees place the hybrid species together with one of the parental lineages, whereas the other half of the gene trees reveals the second parental lineage. For this reason, the species tree becomes reticulate. **g** | As a result of rapid divergence (parts **a** and **b**), duplication or deletion of genes (part **c**), incomplete lineage sorting (part **d**) and introgressive hybridization (part **e**), chromosomes become a mosaic of segments with different evolutionary histories (as shown by local phylogenies), whereby the size of each segment is primarily determined by the number of recombination events that have occurred. Depending on the number of generations since divergence, population sizes and the recombination rate, segments might become excessively small and plentiful (for example, for the cichlids of the Lake Victoria region, it has been estimated that such ancestry blocks are in the size of 500–1,000 bp only).⁷⁶

Recombination. Besides mutational change, recombination is the other major factor that can lead to an increase in genetic variation. The effect of recombination on genetic variation is expected to be especially strong when chromosomes from genetically more distinct parents are involved, as is the case when two species hybridize. In this way, hybridization can instantaneously boost genetic variation¹⁰⁰. In East African cichlids, hybridization has been identified as an important factor for establishing and maintaining genetic variation in species flocks^{27,57,76}. For example, it has been shown that explosive diversification in cichlids of the Lake Victoria region was predated by a hybridization event involving two distantly related riverine lineages, creating a hybrid swarm, whereby fixed differences in the parental lineages recombined to form many new combinations of alleles in the emerging species⁷⁶.

Obvious signatures of introgression were found in the cichlid assemblages of all three large lakes^{26,27,41,57,79} (FIG. 4). Based on the analysis of genome-wide data, it has further been suggested that there is occasional gene exchange between the cichlid faunas in rivers and the African Great Lakes^{26,101}. Recurrent introgression events, potentially triggered by large-scale environmental

changes in the form of lowstands of the lake level that lead to increased water turbidity and/or bring together populations in dense contact^{27,28,44,82}, incomplete lineage sorting and, at least in some cases, hybridization at the onset of adaptive radiation can explain why cichlid genomes are genetically diverse and why much of the variation is shared between cichlid species (FIG. 3). For example, within Lake Malawi, more than 80% of the heterozygous sites are shared between species²⁷, and it has been estimated that more than half of the SNPs that are polymorphic in Lake Malawi are also polymorphic in cichlids outside this lake¹⁰¹. Taken together, it seems that, in East African cichlids, the reshuffling of existing genetic variation has played a major part in creating the raw genetic material for selection to act upon.

Multifarious selection

Comparative genomic analysis in East African cichlids suggests that selection is multifarious during explosive diversification and targets many loci across the genome, as evident from the multiple genomic regions with elevated levels of divergence — so-called outlier regions — between sister taxa^{19,67,102,103}, which is in turn a probable consequence of syndrome selection¹². For example, in a comparison between two *A. calliptera* ecomorphs that are at the verge of speciation in a small crater lake near Lake Malawi, Malinsky et al.⁶⁷ identified 55 genomic regions that were characterized by high divergence (in both F_{ST} and D_{XY}) relative to the genome-wide background. Interestingly, these 'islands of speciation' are not randomly distributed across the genome; approximately half of these islands cluster on only five chromosomes⁶⁷, that is, they form archipelagos¹⁰⁴. Note that the fairly large number of outlier regions is somewhat contradictory to theoretical models, which predict that rapid diversification is most probable if the number of underlying loci is small¹⁵ and that genomic regions of high divergence are not necessarily indicative of divergent selection¹⁰⁵.

Gene candidates emerging from divergence mapping — that is, genes showing elevated levels of divergence between diversifying populations or more rapid coding sequence evolution — include genes with functions in morphogenesis^{19,67}, cytoskeleton development⁶⁷, the sensory system (including the visual opsin genes)^{4,19,27,67,76,102,103}, oxygen transport^{27,103}, hormone signalling^{27,67}, protein translation^{27,67}, pigmentation¹⁹ and the immune system^{27,102}. To date, very few of these gene candidates have been studied in sufficient detail to allow their assignment to discrete functions and to varying phenotypes in cichlids, and a direct link between any of these genes and explosive diversification is currently lacking. Future work should thus focus on the identification and in-depth functional characterization of the genes involved in diversification in cichlids, making use of divergence mapping and laboratory crosses.

Conclusions and outlook

The first wave of genomic exploration of the exceptionally diverse cichlid species flocks of the East African Great Lakes Victoria, Malawi and Tanganyika has identified gene duplication, accelerated coding sequence evolution, transposable element insertion and regulatory

Syndrome selection

Selection on a combination of traits (for example, body and mouth morphology and coloration in cichlids) in a given environment.

F_{ST}

Known as the fixation index, F_{ST} is a measure of differentiation between two populations due to genetic structure.

D_{XY}

Average number of differences between two individuals randomly sampled from two populations.

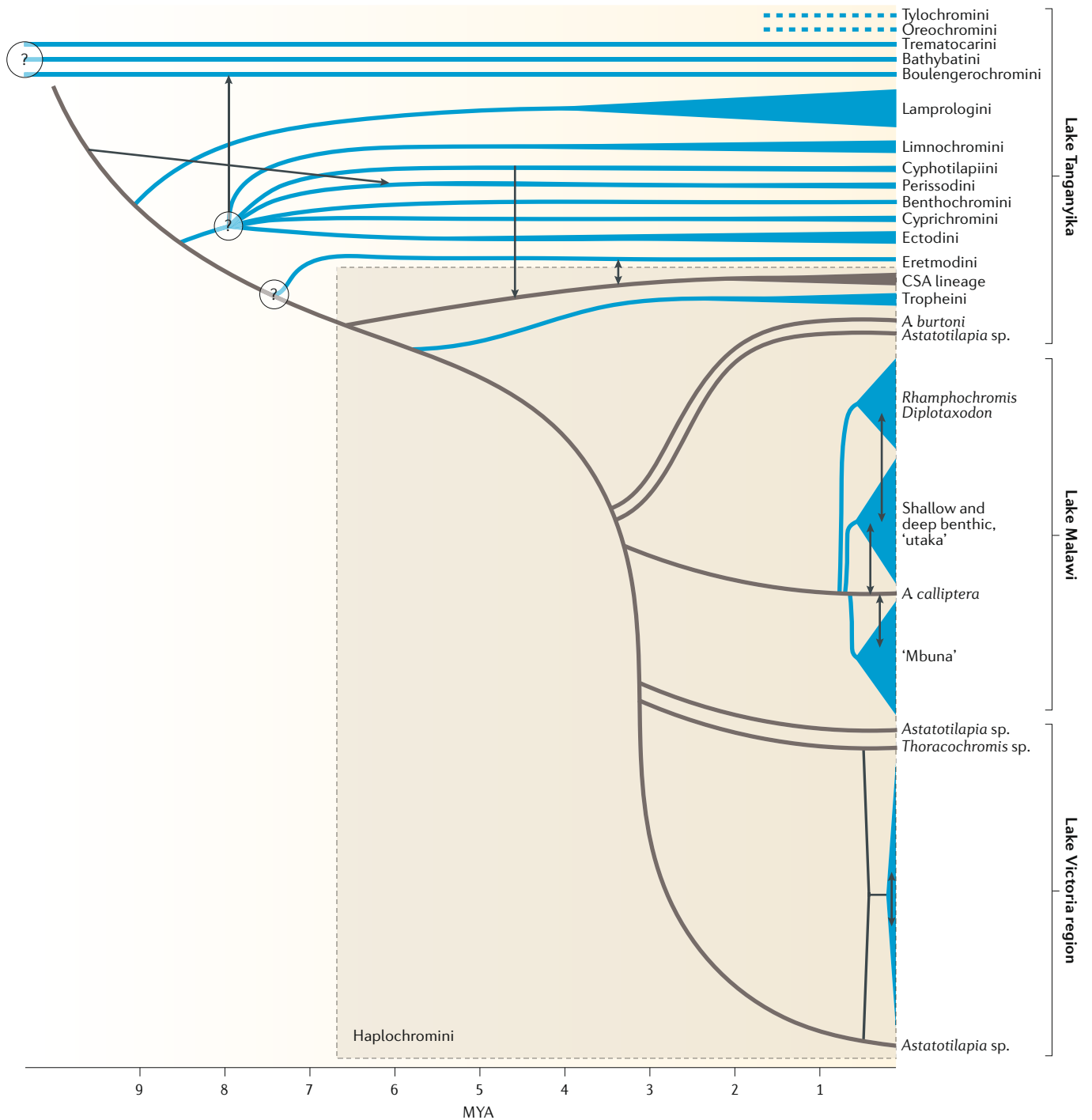


Fig. 4 | **Phylogenetic relationships among the cichlid faunas in the East Africa Great Lakes Victoria, Malawi and Tanganyika.** Phylogenetic hypotheses are based on data in REFS^{26,27,76,78,80}. Lacustrine clades are indicated in blue and scaled according to species numbers. The arrows highlight gene exchange via (introgressive) hybridization (according to REFS^{26,27,41,55–58,76,79}). Dotted lines indicate more recent intruders from more ancestral lineages (Oreochromini and Tylochromini). The grey box indicates the members of the most species-rich cichlid tribe, Haplochromini. *A. burtoni*, *Astatotilapia burtoni*; *A. calliptera*, *Astatotilapia calliptera*; CSA, Congolese/South African; MYA, million years ago

evolution, but not increased nucleotide mutation rate, as candidate genomic features underlying explosive diversification^{19,27}. The reshuffling of existing allelic variation via hybridization at the onset or in the form of introgression in the course of adaptive radiation has further been identified as an important factor boosting genetic

variation^{27,57,76}, which can be tied directly to novel phenotypes via transgressive segregation^{106,107}. Occasional gene exchange between evolutionary lineages in cichlids is possibly facilitated by the relative overall stability of their genomes in terms of genome size and the number of chromosomes as well as comparably low nucleotide

Table 1 | The genomic features of East African cichlids

Characteristic	Value or range	Refs
Genome size (in Gb)	0.92 _{Astbur} –1.01 _{Orenil}	19
Anchored sequence information (in bp)	868,591,263 _{Orenil}	69
Number of chromosomes	2n = 40 _{Astbur} –46 _{Erecya} (44 _{Orenil} ; 44 _{Mayzeb})	98,99
Number of genes and pseudogenes	38,412 _{Orenil}	69
Number of protein-coding genes	29,249 _{Orenil}	69
Percentage of repeat elements	37.6% _{Orenil} <ul style="list-style-type: none"> • DNA transposons: 13% • LINE: 8% • SINE: 0.7% • LTR: 3.2% • Low complexity (satellite and simple repeats): 2.7% • Rolling-circle (Helitrons): 0.3% • Unknown: 9.7% 	69
Nucleotide mutation rate	3.5 × 10 ⁻⁹ per bp per generation _{Astcal,Alnstu,Letlet} (95% CI: 1.6 × 10 ⁻⁹ to 4.6 × 10 ⁻⁹)	27
Mean recombination rate	2.3 cM/Mb _{Punpun, Pspread}	113
Heterozygosity ^a	<ul style="list-style-type: none"> • 1/365_{Neobri}–1/1,800_{Astcal} (natural populations) • 1/976_{Astbur}–1/4,365_{Orenil} (inbred strains) 	19,27
Effective population size (N _e)	50,000–130,000 breeding individuals	27

Alnstu, *Aulonocara stuartgranti*; Astbur, *Astatotilapia burtoni*; Astcal, *Astatotilapia calliptera*; Erecya, *Eretmodus cyanostictus*; Letlet, *Lethrinops lethrinus*; LINE, long interspersed nuclear elements; LTR, long terminal repeats; Mayzeb, *Maylandia zebra*; Neobri, *Neolamprologus brichardi*; Orenil, *Oreochromis niloticus*; Punpun, *Pundamilia pundamilia*; Pspread, *Pundamilia* sp. 'red head'; SINE, short interspersed nuclear elements. ^aNote that some of the reference genomes were generated from inbred individuals, which necessarily results in lower levels of heterozygosity.

mutation rates (TABLE 1), allowing species to hybridize across fairly large phylogenetic distances (see REFS^{53,54}). It is less clear whether hybridization has a causal role in the generation of taxonomic and phenotypic diversity throughout the course of cichlid adaptive radiations or whether hybridization is rather a by-product of the proliferation and subsequent coexistence of numerous recently diverged taxa that are reproductively isolated primarily based on pre-mating isolation mechanisms.

An important question is then what maintains the taxonomic and phenotypic diversity in the cichlid species flocks of the African Great Lakes in light of (occasional) gene flow between lineages and whether assortative mating alone is sufficient to keep species

apart stably or whether other mechanisms, be they genetic or ecological ones, are involved. Clearly, more experimental work — ideally under (semi-)natural conditions — is needed to disentangle the relative roles of signal and mate choice evolution, genetic incompatibilities, habitat preferences and/or spatiotemporal isolation in cichlid diversification, which in turn would expand our knowledge of the nature of cichlid species. Similarly, more experimental work is needed to clarify the relative contribution of phenotypic plasticity to cichlid evolution^{108,109} and the potential role that the high turnover rates in sex determination systems might have^{110–113}.

Ultimately, it will of course be important to know more about the genes underlying explosive diversification in cichlids, as well as their exact functions. The possibility of generating and rearing artificial hybrids for genetic mapping of particular traits^{51,52,113} in combination with divergence mapping in natural populations^{67,102,114} and the prospect of CRISPR–Cas9 gene editing in cichlids¹¹⁵ promise exciting new insights into this topic, especially when these strategies are combined. Overall, only a small fraction of the East African cichlid fauna has been subjected to genome (re-)sequencing to date. Many more genomes from many more species and in much better quality would be needed to fully understand the genomic underpinnings of adaptive radiation and explosive diversification in cichlids. The cichlid system offers an ideal comparative framework in this context, as excessively proliferating lineages can be compared with their non-radiating sister lineages living outside the lakes as well as with comparatively species-poor lineages that diversified side by side to species-rich ones in the same lake and that descend from the same common ancestor (FIG. 4). Importantly, such considerations make sense only in an integrative framework, that is, when similar efforts, such as with genome sequencing, are undertaken to scrutinize the morphology, ecology, physiology and behaviour of the East African cichlids, rather than with a sole focus on ecomorphological traits in the study of cichlid divergence. These empirical efforts should be accompanied by the development of new theory tailored to the phenomenon of explosive diversification and the establishment of novel analytical pipelines to handle these data. Finally, more emphasis should be devoted to better incorporate information provided by the fossil and paleoecological record buried in the lakes' sediments¹¹⁶, including the examination of ancient DNA.

Published online 15 August 2018

- Darwin, C. & Wallace, A. On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection. *J. Proc. Linnean Soc. Zool.* **3**, 45–62 (1858).
- Coyne, J. A. & Orr, H. A. *Speciation*. (Sinauer Associates, 2004).
- Barluenga, M., Stolting, K. N., Salzburger, W., Muschick, M. & Meyer, A. Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature* **439**, 719–723 (2006).
- Seehausen, O. et al. Speciation through sensory drive in cichlid fish. *Nature* **455**, 620–626 (2008).
- Schluter, D. Evidence for ecological speciation and its alternative. *Science* **323**, 737–741 (2009).
- Nosil, P. *Ecological Speciation* (Oxford Univ. Press, 2012).
- Lamichhane, S. et al. Rapid hybrid speciation in Darwin's finches. *Science* **359**, 224–227 (2018).
- Pennisi, E. What determines species diversity? *Science* **309**, 90 (2005).
- Schluter, D. & Pennell, M. W. Speciation gradients and the distribution of biodiversity. *Nature* **546**, 48–55 (2017).
- Barnosky, A. D. et al. Has the Earth's sixth mass extinction already arrived? *Nature* **471**, 51–57 (2011).
- Nosil, P. & Schluter, D. The genes underlying the process of speciation. *Ecol. Evol.* **26**, 160–167 (2011).
- Berner, D. & Salzburger, W. The genomics of organismal diversification illuminated by adaptive radiations. *Trends Genet.* **31**, 491–499 (2015).
- Wolf, J. B. W. & Ellegren, H. Making sense of genomic islands of differentiation in light of speciation. *Nat. Rev. Genet.* **18**, 87–100 (2017).
- Schluter, D. *The Ecology of Adaptive Radiation* (Oxford Univ. Press, 2000).
- Gavrilets, S. & Losos, J. B. Adaptive radiation: contrasting theory with data. *Science* **323**, 732–737 (2009).
- Brakefield, P. M. Evo-devo and constraints on selection. *Trends Ecol. Evol.* **21**, 362–368 (2006).
- Fryer, G. & Iles, T. D. *The cichlid fishes of the Great Lakes of Africa: Their biology and Evolution* (Oliver & Boyd, 1972).

This book is by far the most extensive treatise of the biology of East African cichlid fish.

- Salzburger, W. The interaction of sexually and naturally selected traits in the adaptive

- radiations of cichlid fishes. *Mol. Ecol.* **18**, 169–185 (2009).
19. Brawand, D. et al. The genomic substrate for adaptive radiation in African cichlid fish. *Nature* **513**, 375–381 (2014).
This paper reports the initial sequencing of the five reference genomes of African cichlid fish.
 20. Jones, F. C. et al. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **484**, 55–61 (2012).
 21. Alföldi, J. et al. The genome of the green anole lizard and a comparative analysis with birds and mammals. *Nature* **477**, 587–591 (2011).
 22. Lamichhaney, S. et al. Evolution of Darwin’s finches and their beaks revealed by genome sequencing. *Nature* **518**, 371–375 (2015).
 23. Zhang, G. et al. Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* **346**, 1311–1320 (2014).
 24. Seehausen, O. Process and pattern in cichlid radiations — inferences for understanding unusually high rates of evolutionary diversification. *New Phytol.* **207**, 304–312 (2015).
 25. Kocher, T. D. Adaptive evolution and explosive speciation: the cichlid fish model. *Nat. Rev. Genet.* **5**, 288–298 (2004).
 26. Meyer, B. S., Matschiner, M. & Salzburger, W. Disentangling incomplete lineage sorting and introgression to refine species-tree estimates for Lake Tanganyika cichlid fishes. *Syst. Biol.* **66**, 531–550 (2017).
 27. Malinsky, M. et al. Whole genome sequencing of Malawi cichlids reveal multiple radiations interconnected by gene flow. *bioRxiv* <https://doi.org/10.1101/143859> (2017).
This paper presents the largest analysis based on whole-genome data of any cichlid adaptive radiation thus far.
 28. Ivory, S. J. et al. Environmental change explains cichlid adaptive radiation at Lake Malawi over the past 1.2 million years. *Proc. Natl Acad. Sci. USA* **113**, 11895–11900 (2016).
This paper reports a time-calibrated paleoecological record for Lake Malawi covering the past 1.2 million years.
 29. Johnson, T. C. et al. Late Pleistocene desiccation of Lake Victoria and rapid evolution of cichlid fishes. *Science* **273**, 1091–1093 (1996).
This study shows that Lake Victoria dried out completely during the last glaciation cycle.
 30. Seehausen, O. African cichlid fish: a model system in adaptive radiation research. *Proc. Biol. Sci.* **273**, 1987–1998 (2006).
 31. Salzburger, W., Van Bocxlaer, B. & Cohen, A. S. Ecology and evolution of the African great lakes and their faunas. *Annu. Rev. Ecol. Syst.* **45**, 519–545 (2014).
 32. Moore, J. E. S. The fresh-water fauna of Lake Tanganyika. *Nature* **56**, 198–200 (1897).
 33. Boulenger, G. Catalogue of the fresh-water fishes of Africa in the British Museum (Natural History). Vol. 3 (British Museum for Natural History, 1898).
 34. Woltreck, R. Wie entsteht eine endemische Rasse oder Art? *Biol. Zent. Bl.* **51**, 231–253 (1931).
 35. Mayr, E. in *Evolution of Fish Species Flocks*. (eds Echelle, A. A. & Kornfield, I.) 3–12 (Univ. of Maine at Orono Press, 1984).
 36. Givnish, T. J. Adaptive radiation versus “radiation” and “explosive diversification”: why conceptual distinctions are fundamental to understanding evolution. *New Phytol.* **207**, 297–303 (2015).
 37. Wu, C.-I. The genic view of the process of speciation. *J. Evol. Biol.* **14**, 851–865 (2001).
 38. Feder, J. L., Egan, S. P. & Nosil, P. The genomics of speciation-with-gene-flow. *Trends Genet.* **28**, 342–350 (2012).
 39. Snoeks, J. How well known is the ichthyodiversity of the large East African lakes? *Adv. Ecol. Res.* **31**, 17–38 (2000).
 40. Kornfield, I. & Smith, P. F. African cichlid fishes: model systems for evolutionary biology. *Annu. Rev. Ecol. Syst.* **31**, 163–196 (2000).
 41. Meier, J. I. et al. Demographic modelling with whole-genome data reveals parallel origin of similar *Pundamilia* cichlid species after hybridization. *Mol. Ecol.* **26**, 123–141 (2017).
 42. Verheyen, E., Salzburger, W., Snoeks, J. & Meyer, A. Origin of the super-flock of cichlid fishes from Lake Victoria, East Africa. *Science* **300**, 325–329 (2003).
 43. Nevado, B., Fazalova, V., Bacheljau, T., Hanssens, M. & Verheyen, E. Repeated unidirectional introgression of nuclear and mitochondrial DNA between four congeneric Tanganyikan cichlids. *Mol. Biol. Evol.* **28**, 2253–2267 (2011).
 44. Egger, B., Koblmüller, S., Sturmbauer, C. & Sefc, K. M. Nuclear and mitochondrial data reveal different evolutionary processes in the Lake Tanganyika cichlid genus *Tropheus*. *BMC Evol. Biol.* **7**, 137 (2007).
 45. Konings, A. Tanganyika cichlids in their natural habitat. 3 edn. (CichlidPress, 2015).
 46. Blais, J. et al. Assortative mating among Lake Malawi cichlid fish populations is not simply predictable from male nuptial colour. *BMC Evol. Biol.* **9**, 53 (2009).
 47. Egger, B., Obermüller, B., Eigner, E., Sturmbauer, C. & Sefc, K. M. Assortative mating preferences between colour morphs of the endemic Lake Tanganyika cichlid genus *Tropheus*. *Hydrobiologia* **615**, 37–48 (2008).
 48. Svensson, O. et al. The genetics of mate preferences in hybrids between two young and sympatric Lake Victoria cichlid species. *Proc. Biol. Sci.* **248**, 20162332 (2017).
 49. Mayr, E. *Animal Species and Evolution* (Harvard Univ. Press, 1963).
 50. Dobzhansky, T. *Genetics and the origin of species* (Columbia Univ. Press, 1937).
 51. Henning, F., Machado-Schiaffino, G., Baumgarten, L. & Meyer, A. Genetic dissection of adaptive form and function in rapidly speciating cichlid fishes. *Evolution* **71**, 1297–1312 (2017).
 52. Albertson, R. C., Streebman, J. T. & Kocher, T. D. Directional selection has shaped the oral jaws of Lake Malawi cichlid fishes. *Proc. Natl Acad. Sci. USA* **100**, 5252–5257 (2003).
 53. Stelkens, R. B., Young, K. A. & Seehausen, O. The accumulation of reproductive incompatibilities in African cichlid fish. *Evolution* **64**, 617–632 (2010).
 54. Stelkens, R. B., Schmid, C. & Seehausen, O. Hybrid breakdown in cichlid fish. *PLoS ONE* **10**, e0127207 (2015).
 55. Koblmüller, S. et al. Reticulate phylogeny of gastropod-shell-breeding cichlids from Lake Tanganyika — the result of repeated introgressive hybridization. *BMC Evol. Biol.* **7**, 1 (2007).
 56. Smith, P. F., Konings, A. & Kornfield, I. Hybrid origin of a cichlid population in Lake Malawi: implications for genetic variation and species diversity. *Mol. Ecol.* **12**, 2497–2504 (2003).
 57. Gante, H. F. et al. Genomics of speciation and introgression in Princess cichlid fishes from Lake Tanganyika. *Mol. Ecol.* **25**, 6143–6161 (2016).
 58. Hebert, P. D., Cywinska, A., Ball, S. L. & deWaard, J. R. Biological identifications through DNA barcodes. *Proc. Biol. Sci.* **270**, 313–321 (2003).
 59. Breman, F. C., Loix, S., Jordaens, K., Snoeks, J. & Van Steenberge, M. Testing the potential of DNA barcoding in vertebrate radiations: the case of the littoral cichlids (Pisces, Perciformes, Cichlidae) from Lake Tanganyika. *Mol. Ecol. Res.* **16**, 1455–1464 (2016).
 60. Witte, F. in *Evolution of Fish Species Flocks* (eds Echelle, A. A. & Kornfield, I.) 155–168 (Univ. of Maine at Orono Press, 1984).
 61. Muschick, M., Indermaur, A. & Salzburger, W. Convergent evolution within an adaptive radiation of cichlid fishes. *Curr. Biol.* **22**, 2362–2368 (2012).
This study shows that convergent evolution is common not only between but also within a single cichlid adaptive radiation.
 62. Shapiro, B. J., Leducq, J. B. & Mallet, J. What is speciation? *PLoS Genet.* **12**, e1005860 (2016).
 63. Van Valen, L. Ecological species, multispecies, and oaks. *Taxon* **25**, 233–239 (1976).
 64. Glor, R. E. Phylogenetic insights on adaptive radiation. *Annu. Rev. Ecol. Syst.* **41**, 251–270 (2010).
 65. Mallet, J., Besansky, N. & Hahn, M. W. How reticulated are species? *Bioessays* **38**, 140–149 (2016).
 66. Edwards, S. V., Potter, S., Schmitt, C. J., Bragg, J. G. & Moritz, C. Reticulation, divergence, and the phylogeography-phylogenetics continuum. *Proc. Natl Acad. Sci. USA* **113**, 8025–8032 (2016).
 67. Malinsky, M. et al. Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. *Science* **350**, 1493–1498 (2015).
This study is the most thorough genomic examination of two cichlid ecomorphs at the verge of speciation in a small crater lake near Lake Malawi.
 68. McGee, M. D., Neches, R. Y. & Seehausen, O. Evaluating genomic divergence and parallelism in replicate ecomorphs from young and old cichlid adaptive radiations. *Mol. Ecol.* **25**, 260–268 (2016).
 69. Conte, M. A., Gammerding, W. J., Bartie, K. L., Penman, D. J. & Kocher, T. D. A high quality assembly of the Nile tilapia (*Oreochromis niloticus*) genome reveals the structure of two sex determination regions. *BMC Genom.* **18**, 341 (2017).
This paper reports a high-quality genome assembly based on PacBio sequencing of the Nile tilapia.
 70. Conte, M. A. & Kocher, T. D. An improved genome reference for the African cichlid. *Metriaclima zebra*. *BMC Genom.* **16**, 724 (2015).
 71. Santos, M. E. et al. The evolution of cichlid fish egg-spots is linked with a cis-regulatory change. *Nature Commun.* **5**, 5149 (2014).
 72. Singh, P., Boerger, C., More, H. & Sturmbauer, C. The role of alternative splicing and differential gene expression in cichlid adaptive radiation. *Genome Biol. Evol.* **9**, 2764–2781 (2017).
 73. Bohne, A., Sensttag, T. & Salzburger, W. Comparative transcriptomics in East African cichlids reveals sex- and species-specific expression and new candidates for sex differentiation in fishes. *Genome Biol. Evol.* **6**, 2567–2585 (2014).
 74. Lee, H. J. et al. Lateralized feeding behavior is associated with asymmetrical neuroanatomy and lateralized gene expressions in the brain in scale-eating cichlid fish. *Genome Biol. Evol.* **9**, 3122–3136 (2017).
 75. Colombo, M. et al. The ecological and genetic basis of convergent thick-lipped phenotypes in cichlid fishes. *Mol. Ecol.* **22**, 670–684 (2013).
 76. Meier, J. I. et al. Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nat. Commun.* **8**, 14363 (2017).
This is the most thorough genomic analysis of the Lake Victoria region super-flock of cichlids, uncovering a hybridization event at the onset of the adaptive radiation.
 77. Takahashi, T. & Koblmüller, S. The adaptive radiation of cichlid fish in Lake Tanganyika: a morphological perspective. *Int. J. Evol. Biol.* **2011**, 620754 (2011).
 78. Meyer, B. S., Matschiner, M. & Salzburger, W. A tribal level phylogeny of Lake Tanganyika cichlid fishes based on a genomic multi-marker approach. *Mol. Phylogenet. Evol.* **83**, 56–71 (2015).
 79. Weiss, J. D., Cotterill, F. P. D. & Schlieven, U. K. Lake Tanganyika — a ‘melting pot’ of ancient and young cichlid lineages (Teleostei: Cichlidae)? *PLoS ONE* **10**, e0125043 (2015).
 80. McGee, M. D. et al. Replicated divergence in cichlid radiations mirrors a major vertebrate innovation. *Proc. Biol. Sci.* **283**, 20151413 (2016).
 81. Joyce, D. A. et al. Repeated colonization and hybridization in Lake Malawi cichlids. *Curr. Biol.* **21**, R108–109 (2011).
 82. Malinsky, M. & Salzburger, W. Environmental context for understanding the iconic adaptive radiation of cichlid fishes in Lake Malawi. *Proc. Natl Acad. Sci. USA* **113**, 11654–11656 (2016).
 83. Bezault, E., Mwaiko, S. & Seehausen, O. Population genomic tests of models of adaptive radiation in Lake Victoria region cichlid fish. *Evolution* **65**, 3381–3397 (2011).
 84. Wagner, C. E., Harmon, L. J. & Seehausen, O. Ecological opportunity and sexual selection together predict adaptive radiation. *Nature* **487**, 366–370 (2012).
This study reports a meta-analysis across many different cichlid species flocks, showing that ecology and sexual selection jointly predict the extent of adaptive radiation in cichlids.
 85. Salzburger, W. & Meyer, A. The species flocks of East African cichlid fishes: recent advances in molecular phylogenetics and population genetics. *Naturwissenschaften* **91**, 277–290 (2004).
 86. Sturmbauer, C. Explosive speciation in cichlid fishes of the African Great Lakes: a dynamic model of adaptive radiation. *J. Fish. Biol.* **53** (Suppl. A), 18–36 (1998).
 87. Wagner, C. E., Harmon, L. J. & Seehausen, O. Cichlid species-area relationships are shaped by adaptive radiations that scale with area. *Ecol. Lett.* **5**, 583–592 (2014).
 88. Coulter, G. W. in *Lake Tanganyika and its life* (ed. Coulter, G. W.) 200–274 (Oxford Univ. Press, 1991).
 89. Burress, E. D. et al. Island- and lake-like parallel adaptive radiations replicated in rivers. *Proc. Biol. Sci.* **285**, 20171762 (2018).
 90. Alter, S. E., Munshi-South, J. & Stiassny, M. L. J. Genomewide SNP data reveal cryptic phylogeographic structure and microallopatric divergence in a rapidly-adapted clade of cichlids from the Congo River. *Mol. Ecol.* **26**, 1401–1419 (2017).
 91. Lopez-Fernandez, H., Winemiller, K. O. & Honeycutt, R. L. Multitool phylogeny and rapid radiations in Neotropical cichlid fishes (Perciformes: Cichlidae):

- Cichlinae). *Mol. Phylogenet. Evol.* **55**, 1070–1086 (2010).
92. Arbour, J. H. & Lopez-Fernandez, H. Continental cichlid radiations: functional diversity reveals the role of changing ecological opportunity in the Neotropics. *Proc. Bio. Sci.* **283**, 20160556 (2016).
 93. Ford, A. G. P. et al. Niche divergence facilitated by fine-scale ecological partitioning in a recent cichlid fish adaptive radiation. *Evolution* **70**, 2718–2735 (2016).
 94. Schliwien, U. K., Tautz, D. & Paabo, S. Sympatric speciation suggested by monophyly of crater lake cichlids. *Nature* **368**, 629–632 (1994).
 95. Martin, C. H. et al. Complex histories of repeated gene flow in Cameroon crater lake cichlids cast doubt on one of the clearest examples of sympatric speciation. *Evolution* **69**, 1406–1422 (2015).
 96. Elmer, K. R. et al. Parallel evolution of Nicaraguan crater lake cichlid fishes via non-parallel routes. *Nature Commun.* **5**, 5168 (2014).
 97. Muschick, M. et al. Testing the stages model in the adaptive radiation of cichlid fishes in East African Lake Tanganyika. *Proc. Bio. Sci.* **281**, 20140605 (2014).
 98. Ozouf-Costaz, C. et al. New insights into the chromosomal differentiation patterns among cichlids from Africa and Madagascar. *Cybius* **41**, 35–43 (2017).
 99. Poletto, A. B. et al. Chromosome differentiation patterns during cichlid fish evolution. *BMC Genet.* **11**, 50 (2010).
 100. Mallet, J. Hybrid speciation. *Nature* **446**, 279–283 (2007).
 101. Loh, Y. H. et al. Origins of shared genetic variation in African cichlids. *Mol. Biol. Evol.* **30**, 906–917 (2013).
 102. Egger, B., Roesti, M., Bohne, A., Roth, O. & Salzburger, W. Demography and genome divergence of lake and stream populations of an East African cichlid fish. *Mol. Ecol.* **26**, 5016–5030 (2017).
 103. Hahn, C., Genner, M. J., Turner, G. F. & Joyce, D. A. The genomic basis of cichlid fish adaptation within the deepwater “twilight zone” of Lake Malawi. *Evol. Lett.* **1**, 184–198 (2017).
 104. Ronco, F. & Salzburger, W. Speciation: genomic archipelagos in a crater lake. *Curr. Biol.* **26**, R197–R199 (2016).
 105. Guerrero, R. F. & Hahn, M. W. Speciation as a sieve for ancestral polymorphism. *Mol. Ecol.* **26**, 5362–5368 (2017).
 106. Albertson, R. C. & Kocher, T. D. Genetic architecture sets limits on transgressive segregation in hybrid cichlid fishes. *Evolution* **59**, 686–690 (2005).
 107. Selz, O. M., Lucek, K., Young, K. A. & Seehausen, O. Relaxed trait covariance in interspecific cichlid hybrids predicts morphological diversity in adaptive radiations. *J. Evol. Biol.* **27**, 11–24 (2014).
 108. Schneider, R. F. & Meyer, A. How plasticity, genetic assimilation and cryptic genetic variation may contribute to adaptive radiations. *Mol. Ecol.* **26**, 330–350 (2017).
 109. Rajkov, J., Weber, A. A., Salzburger, W. & Egger, B. Adaptive phenotypic plasticity contributes to divergence between lake and river populations of an East African cichlid fish. *Ecol. Evol.* <https://doi.org/10.1002/ece3.4241> (2018).
 110. Roberts, R. B., Ser, J. R. & Köcher, T. D. Sexual conflict resolved by invasion of a novel sex determiner in Lake Malawi cichlid fishes. *Science* **326**, 998–1001 (2009).
 111. Roberts, N. B. et al. Polygenic sex determination in the cichlid fish *Astatotilapia burtoni*. *BMC Genomics* **17**, 835 (2016).
 112. Bohne, A., Wilson, C. A., Postlethwait, J. H. & Salzburger, W. Variations on a theme: genomics of sex determination in the cichlid fish *Astatotilapia burtoni*. *BMC Genom.* **17**, 883 (2016).
 113. Feulner, P. G. D., Schwarzer, J., Haesler, M. P., Meier, J. I. & Seehausen, O. A dense linkage map of Lake Victoria cichlids improved the *Pundamilia* genome assembly and revealed a major QTL for sex-determination. *G3* <https://doi.org/10.1534/g3.118.200207> (2018).
 114. Meier, J. I., Marques, D. A., Wagner, C. E., Excoffier, L. & Seehausen, O. Genomics of parallel ecological speciation in Lake Victoria cichlids. *Mol. Biol. Evol.* **35**, 1489–1506 (2018).
 115. Juntti, S. A. et al. A neural basis for control of cichlid female reproductive behavior by prostaglandin $F_{2\alpha}$. *Curr. Biol.* **26**, 943–949 (2016).
 116. Muschick, M. et al. Arrival order and release from competition does not explain why haplochromine cichlids radiated in Lake Victoria. *Proc. Biol. Sci.* **285**, 20180462 (2018).
 117. Nelson, J. S., Grande, T. C. & Wilson, M. V. H. *Fishes of the World* (John Wiley & Sons, 2016).
 118. Betancur-R., R. et al. Phylogenetic classification of bony fishes. *BMC Evol. Biol.* **17**, 162 (2017).
 119. Wainwright, P. C. et al. The evolution of pharyngognathy: a phylogenetic and functional appraisal of the pharyngeal jaw key innovation in labroid fishes and beyond. *Syst. Biol.* **61**, 1001–1027 (2012).
 120. Barlow, G. W. *The Cichlid Fishes. Nature's Grand Experiment in Evolution* (Perseus Publishing, 2000).
 121. Froese, R. & Pauly, D. Vol. 10/2017. *FishBase* <http://www.fishbase.org> (2017).
 122. Gardiner, A. *Egyptian Grammar*. 3rd edn (Griffith Institute, Ashmolean Museum, 1957).
 123. Robins, G. *Women in ancient Egypt* (British Museum Press, 1993).
 124. Greenwood, P. H. in *Evolution of Fish Species Flocks* (eds Echelle, A. A. & Kornfield, I.) 13–20 (Univ. of Maine at Orono Press, 1984).
 125. Chenuil, A. et al. Understanding processes at the origin of species flocks with a focus on the marine Antarctic fauna. *Biol. Rev.* **93**, 481–504 (2018).
 126. Schwarzer, J., Misof, B., Ifuta, S. N. & Schliwien, U. K. Time and origin of cichlid colonization of the Lower Congo rapids. *PLoS ONE* **6**, e22380 (2011).
 127. Herder, F., Pfaender, J. & Schliwien, U. K. Adaptive sympatric speciation of polychromatic “roundfin” sailfin silverside fish in Lake Matano (Sulawesi). *Evolution* **62**, 2178–2195 (2008).
 128. Turner, G. F., Seehausen, O., Knight, M. E., Allender, C. J. & Robinson, R. L. How many species of cichlid fishes are there in African lakes? *Mol. Ecol.* **10**, 793–806 (2001).
 129. Seehausen, O. & van Alphen, J. J. M. Can sympatric speciation by disruptive sexual selection explain rapid evolution of cichlids diversity in Lake Victoria? *Ecol. Lett.* **2**, 262–271 (1999).
 130. Boileau, N. et al. A complex mode of aggressive mimicry in a scale-eating cichlid fish. *Biol. Lett.* **11**, 20150521 (2015).
 131. Streebman, J. T. & Danley, P. D. The stages of vertebrate evolutionary radiation. *Trends Ecol. Evol.* **18**, 126–131 (2003).
 132. Danley, P. D. & Kocher, T. D. Speciation in rapidly diverging systems: lessons from Lake Malawi. *Mol. Ecol.* **10**, 1075–1086 (2001).
 133. Hulsey, C. D., Zheng, J., Faircloth, B. C., Meyer, A. & Alfaro, M. E. Phylogenomic analysis of Lake Malawi cichlid fishes: further evidence that the three-stage model of diversification does not fit. *Mol. Phylogenet. Evol.* **114**, 40–48 (2017).
 134. Kocher, T. D., Conroy, J. A., McKaye, K. R. & Stauffer, J. R. Similar morphologies of cichlid fish in lakes Tanganyika and Malawi are due to convergence. *Mol. Phylogenet. Evol.* **2**, 158–165 (1993).
 135. Koblmüller, S. et al. Separated by sand, fused by dropping water: habitat barriers and fluctuating water levels steer the evolution of rock-dwelling cichlid populations in Lake Tanganyika. *Mol. Ecol.* **20**, 2272–2290 (2011).
 136. Van Steenberge, M. *Species and speciation in Tropheus, Simochromis and Pseudosimochromis: a multidisciplinary approach to a cichlid radiation from Lake Tanganyika*. Thesis, KU Leuven (2014).
 137. Rokas, A. & Carroll, S. B. Bushes in the tree of life. *PLoS Biol.* **4**, 1899–1904 (2006).
 138. Gavrilets, S. & Vose, A. Dynamic patterns of adaptive radiation. *Proc. Natl Acad. Sci. USA* **102**, 18040–18045 (2005).
 139. Cortesi, F. et al. Ancestral duplications and highly dynamic opsin gene evolution in percomorph fishes. *Proc. Natl Acad. Sci. USA* **112**, 1493–1498 (2015).
 140. Consortium, T. H. G. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* **487**, 94–98 (2012).
 141. Martin, S. H. et al. Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Res.* **23**, 1817–1828 (2013).

Acknowledgements

The author is grateful for generous support from the European Research Council (ERC) and the Swiss National Science Foundation (SNF) and to three referees for valuable comments.

Competing interests

The author declares no competing interests.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Reviewer information

Nature Reviews Genetics thanks G. F. Turner and the other anonymous reviewer(s) for their contribution to the peer review of this work.