In the lakes of East Africa, cichlid fishes have formed adaptive radiations that are each composed of hundreds of endemic, morphologically stunningly diverse, but genetically extremely similar species. In the past 20 years, it became clear that their extreme phenotypic diversity arose within very short time spans, and that phenotypically radically different species are exceptionally similar genetically; hence, they could be considered to be ‘natural mutants’. Many species can be hybridized and, therefore, provide a unique opportunity to study the genetic underpinnings of phenotypic diversification. Comparative large-scale genomic analyses are beginning to unravel the patterns and processes that led to the formation of the cichlid species flocks. Cichlids are an emerging evolutionary genomic model system for fundamental questions on the origin of phenotypic diversity.

Introduction
Cichlid fishes are one of the most well-known models in evolutionary biology [1–4]. The adaptive radiations of cichlids in East African lakes (Figure 1) are composed of several hundred endemic species each that have diversified within extremely short time spans into phenotypically astonishingly diverse species flocks. In the case of Lake Victoria, more than 500 species arose within less than 100,000 years [5–7]. Since their discovery over 100 years ago, this exuberant diversity of cichlid fish species swarms has peaked the interest of evolutionary biologist. The diversity of this group of fishes is so much larger than that of the other groups of fish that also inhabit the large East African lakes that ecologists and evolutionary biologists alike soon began to ask by what evolutionary mechanisms their diversity arose and how it can be ecologically maintained. It was even suggested that these fishes might be defying biological theory [8] and, hence, their conspicuous diversity was dubbed the ‘cichlid problem’. Now, it would appear that the unusually diverse cichlids might end up providing an unusually informative system in which to study the genetic basis of adaptation and phenotypic diversification, as well as parallel evolution of phenotypes. The information obtained from the cichlid system is likely to be of relevance for many groups of organisms.

What are cichlids? Their phylogeny and timescale
Cichlids are teleost fishes that belong to the family Cichlidae. Recently, a molecular phylogenetic analysis using whole mitochondrial DNA sequences suggested close relationships of fishes of the families Pomacentridae (damselfishes) and Embiotocidae (surfperches) with the Cichlidae [9] (Figure 2). Among teleost species whose genome sequence is available, medaka is the phylogenetic closest one to Cichlidae.

In contrast to relatively young age of the oldest fossil cichlid (~45 Mya), recent molecular evidence suggests that cichlids are a rather ancient family that probably has a Gondwanan origin [10–15]. Since cichlids are likely to be well over 100 million years old (Figure 2), they probably diverged into several ancient lineages quite some time ago and their diversity might not be all that surprising. Furthermore, not all lineages of cichlids are especially species rich or have undergone explosive rates of speciation, even those that are part of the species flock differ in their speciation rates [16]. One lineage of cichlids clearly dominates in terms of diversity: the haplochromine cichlids. They are a rather young lineage of cichlids that is only ±4 million years of age [17]. It arose as part of the Lake Tanganyika cichlid species flock, was able to leave the confines of that lake and then gave rise to the adaptive radiations of both Lakes Malawi and Victoria [6,7,18,19] — they are entirely composed of the haplochromine lineage of cichlids (Figure 3). Haplochromines are the, by far, most species-rich lineage of cichlids with more than 1800 species that belong to this group of cichlids alone. This means that about 8% of all known species of fish belong to this one lineage of cichlids.

As the case of the haplochromines shows, surely several factors contribute to the species richness of cichlids. In the case of the haplochromines factors such as the habitat (cichlids thrive in lakes much more so than in rivers), the evolution of egg-spots on the anal fin in males in conjunction with the evolution of a maternal mouthbrooding mating system seems to have contributed to their speciation and diversification [18,20**]. Maternal
mouthbrooding limits the number of eggs a female can incubate which in turn might limit the effective population size of species and might thereby, in combination with their small size and short generation times of these cichlids, lead to an acceleration of speciation rates. Such differences in phenotypic traits even among cichlid lineages highlight which comparisons might be helpful in determining the relative effects of those traits that contributed to the diversification and speciation of cichlids. Then it becomes particularly interesting and potentially illuminating to investigate the genetic basis of those phenotypic traits.

**Why are there so many cichlids?**

Several kinds of key innovations that only cichlids possess have been suggested to provide at least part of the explanation for their extraordinary evolutionary success. One of these is their astonishingly precise adaptation to particular food items and ecological niches. This is achieved through a hugely flexible oral jaw and dentition and the evolution of novel arrangements of their pharyngeal jaws. Through the combination of the flexibility of the oral jaw morphology and dentition in combination with their second jaw, it is assumed that cichlid fishes have managed [21,22] to exploit many trophic resources that other fish could not. Therefore, they have managed to occupy many ecological niches that were not open to other fish lineages.

But there are several other explanations that have been offered to explain the evolutionary success of cichlids. The apt German word for the family Cichlidae is Buntbarsch, which translates to colorful perch. The conspicuous coloration of cichlids, in many species even females are almost as colorful as males, except in the haplochromine cichlid lineage, where a pronounced sexual color dimorphism exists in which females are drab and cryptically colored and only the males show their beautiful
colors [18,20**]. In several species it could be shown that females show preferences for particular color morphs of males. This has led to the suggestion that sexual selection in addition to more traditional types of mechanisms of speciation such as ecological speciation might be one of the major forces of diversification. Therefore, research has not only focused on studying the genes that underlie jaw and tooth shape (e.g. [23]), but recent research has also investigated genes responsible for color pigments and patterns [24,25], as well as genes involved in vision such as opsin genes [26**,27**], and gene possibly involved in fertilization [28]. But, other sensory modalities such as smell and sound and behavioral differences are very likely to contribute to mate choice and speciation as well, although, so far, they have not received as much attention as genes involved in colorational differences (sender) and visual pigment genes (receiver).

Other potential peculiarity of cichlid fishes that has been suggested to contribute their diversity is their purported propensity for hybridization [29]. Cichlids, possibly more often than other organisms, might also speciate through mechanisms other than allopatric speciation. In cichlids, sympatric speciation has been reported more than once and they are one of the few empirical examples where this mode of speciation has been widely acknowledged to occur, at least under certain environmental conditions [30]. However, it seems safe to suggest that allopatric speciation, because of the very patchy distribution of species that are closely associated with particular types of habitats even in the vast lakes of East Africa, in combination with limited gene flow — also owing to the aggression of males and their stable breeding territories — has made the main contribution to cichlid species ecomorphological diversity [1–4,19].

Cichlid resources for genomics and transcriptomics
The investigation of the genetics of phenotypic diversification and speciation in cichlids has included a number of methods. Because many of the species of the Lake Victoria and Lake Malawi cichlid species flocks are extremely closely related, it is often possible to produce fertile hybrids between them in laboratory settings. Hence, candidate gene approaches [31], microarray analyses [32], and quantitative trait loci (QTL) analyses [33] allowed the identification of genomic loci or even genes that appear to strongly contribute to differences in jaw and tooth shape, and those are species differences that contribute to ecological adaptation and possibly speciation. Other large-scale sequence resources that will contribute even more in the near future to an increasing understanding of the phenotype–genotype relationship are bacterial artificial chromosome (BAC) libraries [34–36], a number of genetic maps [37,38], and EST studies [39,40].

In the age of genomics, research on the diversification of cichlids has moved into large-scale molecular comparisons. Currently, a comparative genome project is underway at the BROAD Institute of MIT (URL: http://www.broad.mit.edu/models/tilapia/) that will sequence the genome of the tilapia at medium high coverage (7×) and will determine the genomes of three other
haplochromine cichlids at low coverage (2×) (Figure 3). Until those genomic efforts are completed other methodological approaches have been undertaken in an effort to learn about the genetics of phenotypic differentiation in cichlid fishes. Such genomic comparisons might also include more phylogenetically comprehensive comparisons among the major fish models such as medaka, stickleback, and fugu. Most recently a comparison of five cichlid genomes that were sequenced with a whole shotgun approach at low coverage (0.5×) yielded some interesting results nonetheless [41**]. These five species of Lake Malawi haplochromine cichlids were from as different lineages as can be found in this adaptive radiation and represented hugely different lifestyles and, yet, they were genetically more similar than humans of different ethnic groups or different laboratory strains of zebrafish. Because of the remarkable genetic homogeneity of cichlids, the large numbers of genetically extremely similar species of haplochromine cichlids have long been called natural experiments or ‘natural mutagenesis screens’. Of the large sets of single nucleotide polymorphisms (SNPs) that were collected, about 3–5% show statistical signs of possibly being associated with candidate genes that may have experienced positive Darwinian selection and may warrant further study. Such an approach will be useful for future genotype–phenotype association studies where representatives of an entire species flock are used as a mapping panel.

The genome sizes (haploid nuclear DNA content) of cichlids range from about 0.9–1.2 pg with some outliers, while chromosome numbers (2n) seem to vary only from 44 to 48 (Animal genome size database; URL: http://www.genomesize.com). These genome sizes and karyotypes resemble those of closely related families of fishes and do not suggest anything out of the ordinary for cichlids. On the basis of the still limited information on the genomes of cichlids it seems that there are no drastic change in their basic genomic organization (e.g. expansion/compaction of genome, whole genome duplications, number, and diversity of retrotransposons) compared to other lineage of fishes with many fewer species [42,43]. The evolution of regulatory elements is believed to be a particularly fast and effective means of very rapid phenotypic diversification [44**]. Larger, more representative data sets on regulatory elements and their evolution in cichlid genomes are still lacking, so it is not clear at this point as to whether there is anything special happening in the genomes of cichlids in regard to regulatory evolution. The limited information on this that has been collected so far would appear to suggest that the presence/absence of putative regulatory elements and even micro-RNA is variable and that those regulatory mechanisms are possibly rather quickly evolving, particularly in terms of neo-functionalization and the complementary fixation of regulatory elements in duplicated genes [45]. This is an avenue of research that will probably yield interesting insights as more comparative genomic sequences and functional genomic studies of cichlids will be conducted.

**Genetics of adaptive traits**

Many of the above-mentioned phenotypic features that are unique to cichlid fishes, namely, morphologies of craniofacial structures (e.g. lips, jaw-shapes, and tooth-shapes) and body color variation, can be attributed to the patterns of differentiation of neural crest cells. In vertebrate embryos, neural crest cells, that delaminate from dorsal neural fold, migrate to programmed sites, where they differentiate into cephalic skeletal element (e.g. jaws), color pigments such as melanocytes and so on. In general, neural crest cells strongly contribute to the
species-specific morphology of craniofacial regions of vertebrates (e.g. [46]). However, although the molecular regulatory factors for migration and differentiation of neural crest cells are relatively well studied, this aspect of cichlid biology has not been explored sufficiently. The first developmental studies about jaw and teeth development in cichlids through QTL analyses [33,47**] pointed toward a strong contribution of bone morphogenetic protein 4 (bmp4). These types of experimental approaches that use QTL or association analyses with genetic maps or entire genomic sequences promise in the near future to increase our understanding of molecular genetic basis of the rapid adaptive radiation of this fascinating group of organisms.

Empty morpho-space and massive parallel evolution through re-awakening of developmental programs?

Despite the impressive diversity of cichlids, not the entire theoretically available ‘morpho-space’ is taken up by them [48*]. For example, many forms (e.g. very large predators or eel-shaped ones) that are found in other families of fish were not invented by cichlids. Furthermore, only some, but not all lineages of cichlids diversified to a notable degree. Why that should be so is still a wide open question.

One of the most interesting features of cichlids is that the diversity of the independent radiations of cichlids is not
random, but rather that the same phenotypic solutions to similar ecological challenges have re-evolved repeatedly (Figure 4) [17,49,50]. Cichlids provide one of the hand textbook examples of exuberant and extremely fast speciation and phenotypic diversification and on the other the phylogenetic analyses discovered that a good portion of this diversity is accompanied by massive parallel evolution in and among these adaptive radiations. This raises very interesting questions the answer to which is likely to be of relevance to all organisms and not only cichlids. Did evolution reuse the same developmental pathways to come up independently with similar developmental outcomes or did it find alternative ways to respond to similar ecological challenges? Our bet would be that evolution reawakened [51] developmental pathways independently to come up with similar designs rather than evolved entirely new alternative genetic and phenotypic solutions in different lineages. But, at this point this is purely a guess, as the answers to these open questions are not in yet and are not easy to get as well.

The recognition that cichlid species flocks also provide a textbook example of parallel evolution or convergence opens up very interesting future research directions that can be addressed only through comparative developmental and genomic approaches. These are not easy problems, but this line of research promises to yield insights into the genetics of phenotypic diversification that have obvious relevance beyond cichlids.

Conclusions

Clearly, more complex lake environments seem to contribute to or permit the diversification of cichlids since the species assemblages in lakes are always much more species rich than those of riverine communities. But, not all lineages of cichlids are equally prone to speciate, the champions being the haplochromine cichlids. This raises the question as to whether some genomic features of some or all cichlid lineages predispose them to radiate and diversify phenotypically. The investigation of the comparative developmental genetic basis of traits and genomic comparisons across different lineages and radiations will be necessary to get a handle on the long-standing ‘cichlid problem’. Comparative genomic information within cichlids and comparisons to other fish genomes are just beginning to be collected. Some efforts are underway to study changes in expression patterns of genes, investigations of micro-RNAs [45], retrotransposons [43], and other aspects of regulatory evolution. The question as to whether regulatory evolution in cichlids is, in some way, different, that is more effective, from other lineages of less species-rich organisms and particularly conducive to speciation remains open at this point. As recent work on Hox, ParaHox, and KCNA gene clusters suggests, the genomes of cichlids do not seem to differ all that much from those of other fishes [45,52,53]. SNP-based association studies, and whole genomic scans for conspicuous methylation patterns might provide some clues as to whether there is something peculiar in the genomes of these fishes that would suggest a genomic contribution to their particularly fast rates of speciation and phenotypic diversification. Finding a solution to the ‘cichlid problem’ has obvious implications for a deeper understanding of the genetic basis of phenotypic diversification that goes beyond a better grasp on cichlid fishes.

Acknowledgements

We thank the University of Konstanz and the Deutsche Forschungsgemeinschaft for financial support.

References and recommended reading

Papers of special interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

would suggest that this gene underwent positive selection and a change of the accelerated rate of evolution and the number and place of amino acid evolutionary analysis also shows a high Ka/Ks ratio for this gene and both developing and regenerating egg-spots on the anal fin of males. Molecular Using a candidate gene approach, csf1r is found to be expressed in 


Using a candidate gene approach, csf1r is found to be expressed in developing and regenerating egg-spots on the anal fin of males. Molecular evolutionary analysis also shows a high Ka/Ks ratio for this gene and both the accelerated rate of evolution and the number and place of amino acid substitutions occur at an evolutionary time and a position in the protein that would suggest that this gene underwent positive selection and a change of function when the egg-spots evolved in the haplochromine cichlids.


Using quantitative PCR and ontogenetic analyses of sensitivity spectra in opsin genes in several families of cichlids, the authors interpret differences between species in a phylogenetic context as heterochronic changes from an ancestral pattern seen in tilapia. The authors interpret differences between species in a phylogenetic context as heterochronic changes from an ancestral pattern seen in tilapia. The authors interpret differences between species in a phylogenetic context as heterochronic changes from an ancestral pattern seen in tilapia.


Albertson RC, Streelman JT, Kocher TD: Directional selection has shaped the oral jaws of Lake Malawi cichlid fishes. Proc Natl Acad Sci U S A 2003, 100:5252-5257.


Using partial (5× coverage) genomic sequences of five species of Lake Malawi cichlids the authors show that they differ by only about 0.25% and identify genomic regions that might contain differences that show signs of selection and contain genes that might have contributed to species differences and species.


The author develops verbal models of the genetic basis of phenotypic diversification.


Using crosses between two different Lake Malawi cichlid species, the authors use a QTL approach to identify the number and size of the effect of particular genomic regions and identify potential genes that might be responsible for particular differences in tooth shape.


Using a combination of morphometric and phylogenetic analysis, the authors analyze the patterns of diversification of the Lake Tanganyika cichlid adaptive radiation. They investigate similarities and differences within and among tribes of that species flock.


