Molecular bases of morphological diversity in cichlid fishes

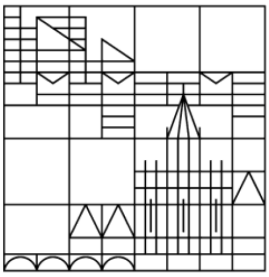
Dissertation submitted for the degree of Doctor of Natural Sciences

Presented by

Ralf Friedrich Schneider

at the

Universität Konstanz



Faculty of Sciences

Department of Biology

Date of the oral examination: 10.03.2017

First supervisor: Prof. Axel Meyer

Second supervisor: Prof. Manfred Schartl

The love for all living creatures is the most noble attribute of man.

- Charles Darwin

Table of Contents

List of Figures

Figure **page**

List of Tables

Summary

Key innovations are characters promoting species richness in the monophyletic group sharing it, relative to a sister taxon not exhibiting it. Key innovations can even induce adaptive radiations for example by increasing the available new niche space to the lineage, facilitating reproductive isolation or reducing extinction rates.

In **chapter I** the specialized morphology of the tiger tail seahorse, *Hippocampus comes*, exhibiting numerous characteristics and key innovations, is investigated using a full genome sequencing and *de novo* assembly approach. Derived traits in seahorses (and some allies) include a toothless tubular mouth, bony plates covering their whole body, male pregnancy within a brooding pouch and the loss of caudal and pelvic fins. Using comparative genomics, increased amino-acid and nucleotide evolution rates were identified in the seahorse genome compared with other teleost fish. Expansion of an astacin metalloprotease gene family was identified that is highly expressed at different stages of the seahorse male pregnancy in the brood pouch. Furthermore, seahorses lost enamel matrix protein-coding proline/glutamine-rich secretory calcium-binding phosphoprotein genes that might explain the lack of mineralized teeth. In addition, *tbx4*, a major regulatory of hind limb development in tetrapods, could not be found in the seahorse genome. Knock-out of *tbx4* in zebrafish led to a loss of pelvic fins – a phenotype resembling that of the seahorse.

Key innovations may lead to species proliferation and, in the most extreme cases, adaptive radiations, i.e. the extremely rapid emergence of new species corresponding to ecologically distinct niches from a single common ancestral population. Some of the largest and most rapid adaptive radiations are found in East African haplochromine cichlids, radiating in the three Great Rift lakes, Lake Victoria, Lake Malawi and Lake Tanganyika. Cichlid fishes possess a second set of functional jaws, pharyngeal jaws, that presumably facilitated their trophic diversification, which is why these jaws are considered a key innovation in cichlids. In some species, pharyngeal jaws were found to respond plastically to varying diets, raising the question about the evolutionary consequences of phenotypic plasticity in this key ecological trait.

Chapter II reviews the increasing evidence that phenotypic plasticity can facilitate population divergence by promoting phenotypic diversification and, eventually, genetic divergence. Phenotypic plasticity is the ability of organisms with a given genotype to develop different phenotypes according to environmental stimuli, resulting in individuals that are better adapted to local conditions. The chapter illustrates how a plastic 'ancestral' lineage, after colonizing a new habitat, phenotypically diversifies and how these diverse phenotypes can then be genetically fixed via a process called 'genetic assimilation' (a 'flexible-stem' scenario). Additionally, molecular mechanisms are reviewed that illustrate

how genetic fixation of a formerly plastic phenotype may work, solely by random mutations and without the need for a 'cost of plasticity'. It is described how genetic assimilation contributes to cryptic genetic variation, but also how it can lead to non-adaptive responses. Predictions about expected phenotypic, genetic and transcriptional patterns induced during a flexible-stem radiation are formulated and illustrated. Furthermore, it is noted that the degree of inducible adaptive and non-adaptive plasticity is expected to vary across lineages at different stages of genetic assimilation. Analyses of these patterns can inform on the state of genetic assimilation in candidate lineages. It is reasoned that, depending on the environment, phenotypic plasticity can promote lineage diversification and divergence, and increase the rate of evolution. The chapter also exemplifies proposed patterns and conclusions using the cichlids as a model system. It is concluded that available evidence supports a flexible stem scenario for at least some cichlid radiations.

In spite of their ecological importance, the developmental regulatory networks underlying plastic phenotypes often remain uncharacterized. In **chapter III** the regulatory basis of phenotypic plasticity in the lower pharyngeal jaw of the cichlid *Astatoreochromis alluaudi* is investigated, a model species in the study of adaptive plasticity. By raising juvenile *A. alluaudi* on either soft or hard diets for between one to eight months, the temporal regulation of previously identified candidate genes could be monitored during the plastic response. Morphological divergence of phenotypes of the two diet groups could be observed between three to five months of treatment, which are preceded by a consistent change in candidate gene expression patterns. It is concluded that investigated genes are likely contributing to the plastic response and pharyngeal jaw bone remodeling in this cichlid. Candidate genes were found to be strikingly co-expressed according to functional categories and transcription factor binding site analysis was performed to examine the prospective regulatory basis of this co-regulation. Based on these results a candidate gene regulatory network putatively underlying lower pharyngeal jaw plasticity is proposed, including evidence for a modular organization but also cross-talk among these modules, which presumably facilitates the plastic remodeling of this highly integrated morphological structure.

Chapter IV investigates whether adaptive diversity in pharyngeal jaw phenotypes found in one of the most extensive adaptive radiations of cichlids, the modern haplochromines, is likely to have originated in a flexible stem. Juveniles of five cichlid species from within the modern haplochromines, representing 'basal' non-radiating generalist and 'derived' radiating specialists, were fed on either a soft or a hard diet to induce a plastic response in the lower pharyngeal jaws. The measured morphological adaptive plastic response was determined to be most pronounced in the most basal generalist, while the more specialized species had considerably lower levels of plasticity.

This suggests that plasticity was reduced during trophic specialization in this radiation via genetic assimilation. In contrast, non-adaptive plastic responses were identified to be more pronounced in specialized species, coinciding with predictions made in chapter II of this thesis. Two candidate genes that potentially have undergone genetic assimilation are identified. It is concluded that in this cichlids' radiation the degree of adaptive phenotypic plasticity was reduced by genetic assimilation during trophic specialization to suit progressively the more narrow ecological niches of each species.

Besides the wide diversity of trophic characteristics, cichlids are famously known for their outstanding diversity in body colorations across species, but also across development within species and between sexes. The evolution of the latter was suggested to be driven by Fisher's run-away selection, which could be evident by sexual dimorphism not only in body coloration, but also the visual system. In **chapter V**, sexually monomorphic and dimorphic cichlid species, both from the Afrotropics as well as the Neotropics, are investigated for being sexually dimorphic in visual systems, however, no evidence was found supporting this. Nonetheless, rod opsin expression was highly variable across all species, while interspecies variations in cone opsin expression were limited to Afrotropic species. By predicting candidate cichlids' effective retina sensitivities and their body colorations in their corresponding habitats, evidence is found indicating that both abiotic factors (such as the available ambient light spectrum and its brightness) as well as biotic factors (here: conspecific body colorations) effecting opsin expression.

Zusammenfassung

Schlüsselinnovationen sind Merkmale, die die Artenvielfalt innerhalb einer monophyletischen Gruppe mit diesem gemeinsamen Merkmal fördern, relativ zu einem Schwester-Taxon, welches das spezifische Merkmal nicht aufweist. Schlüsselinnovationen können sogar adaptive Radiationen induzieren, zum Beispiel durch Erhöhung der verfügbaren neuen Nischenräume für diese Abstammungslinie, Förderung der reproduktiven Isolation oder Reduzierung der Extinktionsraten.

In **Kapitel I** wird mithilfe einer vollständigen Genomsequenzierung mittels „*de novo*“ Genomerstellung die spezialisierte Morphologie des Tigerschwanz-Seepferdchens, *Hippocampus comes*, untersucht, welches zahlreiche besondere Charakteristika und Schlüsselinnovationen aufweist. Abgeleitete Merkmale in Seepferdchen (und einigen verwandten Gruppen) umfassen einen zahnlosen röhrenförmigen Mund, knöcherne Platten, die ihren ganzen Körper bedecken, Schwangerschaft der Männchen mithilfe einer Bruttasche und den Verlust der Schwanz- und Bauchflossen. Mittels vergleichender Genomik wurden im Seepferdchen-Genom erhöhte Aminosäure- und Nukleotid-Evolutionsraten im Vergleich zu anderen Teleostei gemessen. Zudem wurde eine Ausweitung der Astacin-Metalloprotease-Gen-Familie identifiziert, die während verschiedener Stadien der Schwangerschaft in der Bruttasche stark exprimiert wird. Darüber hinaus weisen Seepferdchen keine Zahnschmelzmatrixprotein-codierenden prolin-/glutaminreichen sekretorischen Kalzium-bindenden Phosphoproteingene auf, was das Fehlen von mineralisierten Zähnen erklären könnte. Auch konnte *tbx4*, ein wichtiger Regulator der Entwicklung der hinteren Gliedmaßen in Tetrapoden, nicht im Seepferdchen-Genom gefunden werden. Der Knock-out von *tbx4* in Zebrafischen führte zu einem Verlust der Bauchflossen - einem Phänotyp ähnlich dem des Seepferdchens.

Schlüsselinnovationen können zur Ausbreitung von Arten und in den extremsten Fällen zu adaptiven Radiationen führen, d.h. dem extrem schnellen Entstehen neuer Arten, die ökologisch unterschiedlichen Nischen einer gemeinsamen Stammpopulation entsprechen. Einige der umfassendsten und schnellsten adaptiven Radiationen finden sich bei den Ostafrikanischen Haplochromini-Cichliden in den drei großen Seen, dem Victoria-See, Malawi-See und Tanganjika-See. Cichliden besitzen einen zweiten Satz von funktionalen Kiefern, die Schlundkiefer, die vermutlich ihre trophische Diversifizierung maßgeblich erleichtern, weshalb sie als Schlüsselinnovation in Cichliden angesehen werden. Bei manchen Arten wurde festgestellt, dass diese Kiefer plastisch auf unterschiedliche Ernährung reagieren, was die Frage nach den evolutionären Konsequenzen der phänotypischen Plastizität in diesem wichtigen ökologischen Merkmal aufwirft.

Kapitel II beleuchtet die immer zahlreicher werdenden Hinweise darauf, dass phänotypische Plastizität die Auftrennung von Populationen erleichtern kann, indem sie die phänotypische Diversifizierung und letztendlich die genetische Divergenz begünstigt. Phänotypische Plastizität ist die Fähigkeit von Organismen mit einem gegebenen Genotyp, je nach Umwelteinfluss verschiedene Phänotypen zu entwickeln und so Individuen hervorzubringen, die besser an die lokalen Bedingungen angepasst sind. Das Kapitel veranschaulicht, wie eine plastische Ahnenlinie nach der Besiedlung eines neuen Lebensraumes phänotypisch diversifiziert und wie diese verschiedenen Phänotypen über einen Prozess genetisch fixiert werden können, der als genetische Assimilation bekannt ist („flexible-stem"-Szenario). Zusätzlich werden molekulare Mechanismen untersucht, die veranschaulichen, wie die genetische Fixierung eines früheren plastischen Phänotyps allein durch zufällige Mutationen und ohne "inhärente Kosten von Plastizität" funktionieren kann. Es wird beschrieben, wie die genetische Assimilation zur kryptischen genetischen Variation beiträgt, aber auch, wie sie zu nicht-adaptiven Reaktionen führen kann. Es werden Vorhersagen über erwartete phänotypische, genetische und transkriptionale Muster formuliert und beleuchtet, die während einer „flexible-stem"-Radiation induziert werden. Darüber hinaus wird erklärt, warum der Grad der induzierbaren adaptiven und nicht-adaptiven Plastizität zwischen den Abstammungslinien verschiedener Stadien der genetischen Assimilation variiert. Analysen dieser Muster können über den genetischen Assimilationszustand ausgewählter Abstammungslinien Auskunft geben. Es wird argumentiert, dass die phänotypische Plastizität, abhängig von der Umwelt, die Diversifizierung und Divergenz von Abstammungslinien fördern und die Evolutionsrate erhöhen kann. Das Kapitel veranschaulicht zudem die vorgeschlagenen Muster und Schlussfolgerungen anhand der Cichliden als Modellsystem. Es wird gefolgert, dass die verfügbaren Hinweise zumindest für einige Cichliden-Radiationen ein „flexible-stem"-Szenario unterstützen.

Trotz ihrer ökologischen Bedeutung, bleiben die Entwicklungsregulationsnetze, die den plastischen Phänotypen zugrunde liegen, häufig ungeklärt. In **Kapitel III** wird die regulatorische Basis der phänotypischen Plastizität im unteren Schlundkiefer des Buntbarsches *Astatoreochromis alluaudi* untersucht, einer Modellart in der Erforschung der adaptiven Plastizität. Durch das Heranzüchten von juvenilen *A. alluaudi* mit weichem oder hartem Futter für ein bis acht Monate, konnte die zeitliche Regulation von zuvor identifizierten Kandidatengenen während der plastischen Reaktion beobachtet werden. Dabei konnte im Zeitraum von drei bis fünf Monaten nach Beginn der Behandlung eine morphologische Divergenz der Phänotypen der beiden Ernährungsgruppen beobachtet werden, denen eine gleichartige Veränderung der Kandidatengenexpressionsmuster vorausging. Daraus wird gefolgert, dass die untersuchten Gene wahrscheinlich zu der

plastischen Reaktion und den Schlundkieferveränderungen in diesen Buntbarschen beitragen. Die Kandidatengene wurden gemäß funktioneller Kategorien auffallend ko-exprimiert, weshalb Transkriptionsfaktorbindungsstellenanalysen durchgeführt wurden, um die regulatorische Basis dieser Ko-Regulierung zu untersuchen. Auf Grundlage dieser Ergebnisse wird ein Kandidaten-Genregulationsnetzwerk vorgeschlagen, das der Plastizität des unteren Schlundkiefers zugrunde liegen könnte. Dieses schließt Hinweise auf eine modulare Organisation, aber auch einer Koordination zwischen diesen Modulen ein, welche vermutlich die plastische Umgestaltung dieser hochintegrierten morphologischen Struktur erleichtern.

Kapitel IV untersucht, ob die adaptive Diversität bei Schlundkiefer-Phänotypen, die in einer der umfangreichsten adaptiven Radiationen der Buntbarsche - den modernen Haplochrominen - gefunden wurde, aus einem „flexible-stem“ entstanden sein könnte. Juvenile Buntbarsche von 5 verschiedenen modernen Haplochrominen-Arten, die "basale" nicht radiierende Generalisten und auch abgeleitete radiierende Spezialisten repräsentieren, wurden entweder mit weichem oder hartem Futter gefüttert, um eine plastische Reaktion in den unteren Schlundkiefern zu induzieren. Die gemessene morphologische adaptive plastische Reaktion war in den basalsten Generalisten am stärksten ausgeprägt, während die spezialisierteren Arten eine wesentlich geringere plastische Reaktion aufwiesen. Dies deutet darauf hin, dass die Plastizität während der trophischen Spezialisierung in dieser Radiation durch genetische Assimilation reduziert wurde. Im Gegensatz dazu waren nicht-adaptive plastische Reaktionen bei spezialisierten Arten stärker ausgeprägt, was mit den Voraussagen übereinstimmt, die in Kapitel II dieser Arbeit gemacht wurden. Es werden zwei Kandidatengene identifiziert, die sich möglicherweise in fortgeschrittenen Stadien der genetischen Assimilierung befinden. Vor diesem Hintergrund wird der Schluss gezogen, dass bei dieser Buntbarsch-Radiation der Grad der adaptiven phänotypischen Plastizität durch genetische Assimilation während der trophischen Spezialisierung reduziert wurde, um die Schlundkieferphänotypen den zunehmend enger werdenden ökologischen Nischen jeder Art anzupassen.

Neben der großen Vielfalt hinsichtlich trophischer Merkmale sind Buntbarsche besonders bekannt für ihre außergewöhnliche Vielfalt in der Körperfärbung – sowohl zwischen den Arten als auch innerhalb einer Art zwischen verschiedenen Entwicklungsstadien und Geschlechtern. Bei der Evolution der Geschlechtsunterschiede wurde postuliert, dass sie durch die “Fisher-Runaway“-Selektion angetrieben wurde, was nicht nur durch sexuellen Dimorphismus in der Körperfärbung, sondern auch im visuellen System ersichtlich werden könnte. In **Kapitel V** werden sexuell monomorphe und dimorphe Buntbarscharten aus der Afrotropis und der Neotropis hinsichtlich sexuellen Dimorphismus in ihrem visuellen System untersucht – jedoch wurden keine unterstützenden Beweise

dafür gefunden. Nichtsdestotrotz war die Expression von Stäbchen-Opsinen zwischen den Arten sehr variabel, während nennenswerte zwischenartliche Variationen in den Zapfen-Opsinen ausschließlich auf afrotropische Arten beschränkt waren. Durch die Untersuchung der wirksamen Retina-Empfindlichkeiten und Körperfärbungen der Kandidaten-Buntbarsche in ihren jeweiligen Lebensräumen werden Hinweise darauf gefunden, dass sowohl abiotische Faktoren (wie etwa das verfügbare Spektrum und die Helligkeit des Umgebungslichts) als auch biotische Faktoren (hier: Körperfärbungen der Artgenossen) die Opsin-Expression beeinflussen.

General Introduction

The extraordinary diversity of biological diversity on this planet amazed generations of evolutionary biologists and exploring the mechanisms through which such diversity did arise remains the designated challenge of biologists since Darwin's "The Origin of Species" was published in 1859. The external forces shaping the fate of organismal lineages and ultimately leading to their success or decline are selective forces, most prominently natural selection and sexual selection. These act on organisms' phenotypes that in turn are a product of their genomes and environments. It was realized that large parts of modern organisms' genomes originated in duplication events of different extents (cosanOhno 1970; Ohno 2013). The most remarkable of which are certainly whole genome duplications - events that have been related to strongly increased subsequent diversification rates (Taylor *et al.* 2003). All duplication events initially lead to an increase of genetic material that can act as a substrate for evolution, e.g. in that new gene copies take over new functions or pseudogenize and thus contribute to a lineage's cryptic genetic variation (Ohno 1970). In addition, duplicated regulatory elements may give rise to new expression domains of downstream genes and thus foster diversification. The relative importance of such regulatory evolution compared to coding evolution is a topic of ongoing debate (Davidson & Erwin 2006; Hoekstra & Coyne 2007). The first chapter of this thesis presents the findings of a genome analysis of the seahorse *Hippocampus comes*, one of the most bizarre fishes, exemplifying that indeed many proposed evolutionary mechanisms contributed to its morphological divergence compared to most other fishes (Lin *et al.* 2016). These include gene duplication and loss, coding evolution as well as alterations in conserved non-coding elements with putative regulatory functions. It illustrates excellently how previously evolved traits can get lost, such as pelvic fins or teeth, but also how new traits can evolve, such as the dermal body plates shared by all Syngnathids and some allies.

Most pronounced differences in body plans are typically found when phylogenetically higher taxa are compared while differences often become subtler when lower taxa are considered (Galis 2001). As evolution and diversification - but also extinction rates - are often different across lineages and evolutionary times, some taxa might comprise diverse assemblages of a multitude of species whilst their sister taxa may contain only very few species with limited phenotypic differences. In some cases certain characteristics facilitate a lineage's evolutionary success (in terms of species richness), e.g. by increasing its diversification potential, reducing its vulnerability to extinction or promote reproductive isolation within populations (Galis 2001).

Lineage evolution and diversification rates can be driven by external (i.e. environmental) or internal (e.g. mutational) change. While external changes may allow a

population to access new habitats (i.e. increase ecological opportunity), internal changes include behavioral, physiological or morphological changes (Kocher *et al.* 1993; Galis 2001). It was realized that some evolved (i.e. internal) changes affect diversification rates more positively than others, allowing lineages to proliferate beyond the range of a sister taxon. Such key changes with a disproportionally positive effect are often referred to as *key innovations* or *key ecological traits* (Meyer 1993a; Hunter 1998; Stiassny & Meyer 1999; Galis 2001; Schön & Martens 2004). Key innovations can facilitate the invasion of new areas of the adaptive landscape, decrease the probability of extinction or support ecological specialization and reproductive isolation, which had made them the stepping-stones for today's higher taxa to evolve. Therefore, key innovations are a major causative factor for their evolutionary success (in terms of species richness) (Hoogerhoud 1984; Hunter 1998; Galis 2001; Badyaev *et al.* 2005). It should also be noted that the evolution of a specific character may be a key innovation in one clade and environment, leading to a massive diversification, while the same character's affect on diversification rates may be very subtle in another clade. For example, the fish families *Cichlidae*, *Pomacentridae*, *Labridae* and *Embiotocidae* share a specifically modified pharyngeal jaw apparatus (Liem 1986). While this modification is thought to have critically driven the massive diversification of the first three families (with >1500, ~380 and ~600 species, respectively) the latter family comprises only about 25 species (Liem 1974, 1986; Chapman *et al.* 2001).

Pharyngeal jaws in cichlids are considered a key innovation, as they constitute a second set of functional jaws. They took over food portioning and processing – tasks that have to be taken care of by the oral jaws in most other fishes. In this way, pharyngeal jaws allowed the oral jaws to evolve less constraint into specialized and diverse tools for food extraction from the environment, which importantly drove cichlid trophic diversification (Liem 1974). Pharyngeal jaws phenotypes, similarly to oral jaws, typically reflect the trophic niche of the species having it: while molluscivores have massive, robust pharyngeal jaws made for crushing hard snail shells, zooplanktivores have much more slender and delicate pharyngeal jaws, with numerous fine teeth instead of fewer massive ones (Hoogerhoud 1986a). Further variants can be found in algae browsers, insectivores, piscivores and other food specialists (Hoogerhoud 1984). Surprisingly, this diversity in phenotypes can be found even in some of the very young adaptive radiations, such as in Lake Victoria where *Haplochromine* cichlids flourish and diversified in probably less than 100,000yrs and formed >500 species (Meyer 1990; Stiassny & Meyer 1999). This raises the question how this phenotypic diversity could originate within such very short evolutionary time-spans – much faster than plausibly could be explained by coding evolution (West-Eberhard 2003).

Work pioneered by Greenwood (1964) and followed by many others found that some cichlid lineages have remarkably plastic pharyngeal jaws, i.e. the pharyngeal jaws develop in response to diet robustness either molariform (robust pharyngeal jaws with molar-like teeth) or papilliform (more slender, with numerous fine teeth) phenotypes. Interestingly, pharyngeal jaw plasticity was found in riverine cichlids, i.e. those that approximate best the lineages that seeded East Africa's Great lakes and led to some of the most outstanding adaptive radiations known. This stands in harsh contrast to classical views that plasticity would reduce evolutionary pace and diversification rates, as phenotypes and genotypes are uncoupled and thus the effects of natural selection would be dampened. More recently it was realized that phenotypic plasticity might actually facilitate diversification under certain circumstances (e.g. West-Eberhard 2003; Pfennig & McGee 2010; Schneider & Meyer 2016). Chapter II of this thesis reviews recent advances in theoretical and empirical studies discussing how phenotypic plasticity can facilitate evolution, lineage divergence and even adaptive radiations. It also extents the theoretical framework previously published by incorporating non-adaptive plastic responses in it. Chapter II also illustrates the discussed concepts using the cichlids as a model system. Then, chapter III empirically investigates the morphological and transcriptional characteristics of pharyngeal jaw plasticity by measuring pharyngeal jaw morphology and candidate gene expression across the plastic response. Similarly to chapter III, chapter IV investigates the pharyngeal jaw plastic response in a comparative framework, utilizing 'basal' riverine species and more 'derived' specialists from within the Lake Victoria and the Tanganyika *Tropheini* radiation. These studies give first phylogenetically informed insights in the occurrence of pharyngeal jaw adaptive and non-adaptive plasticity and its underlying transcriptional patterns across Haplochromine cichlids, suggesting that ancestral plasticity has likely promoted pharyngeal jaw diversity in radiating Haplochromine lineages.

Besides their extreme trophic diversity, cichlid fishes are particularly well-known for their outstanding diversity in body colorations, which may only be rivaled by the most colorful clades of coral reef fish. Whilst natural selection was suggested to have critically driven the evolution of this wide range of trophic diversity (Keenleyside 1991), sexual selection is thought to have facilitated the outstanding diversity of body coloration (Deutsch 1997; Allender *et al.* 2003), that itself was suggested to promote reproductive isolation and thus lineage divergence (Seehausen *et al.* 1997; Carleton *et al.* 2005).

Body coloration diversity in cichlids, which is particularly manifold in East African Rift Lake species, can be found across different species, but also within: across development or between sexes. Sexually dimorphism in body coloration was previously suggested to be driven by Fisherian run-away sexual selection (Fisher 1930; Kelber *et al.* 2003; Sabbah *et al.* 2010). In this scenario, female cichlids (the 'choosy' sex) would prefer a subtle

characteristic in male body coloration that may or may not initially be connected to an adaptive advantage. Sexual selection would lead to males with more pronounced versions of this characteristic and females would, in turn, evolve visual capacities that allow them to differentiate more effectively between males. This positive feedback loop may then lead to exaggerated male characters, way beyond adaptiveness. Interestingly, this scenario puts the female visual systems into a different selective regime compared to the one of the males, as females may tune theirs specifically for discriminating male character expression (Fisher 1930). Thus, sexual dimorphism in body coloration could be reflected in sexual dimorphism in visual capacities. Chapter V investigates whether sexual dimorphism in visual systems of sexually dimorphic (in body coloration) species can be found, which would represent strong support for a contribution of run-away sexual selection to cichlid body color dimorphism. As no evidence for visual dimorphism in any investigated species is found, other potential determinants of visual sensitivity are considered. By modeling fish body coloration in the corresponding habitat, as well as visual sensitivities, effects of abiotic and biotic factors on visual tuning can be identified. Therefore, it is concluded that both natural selection, as well as sexual selection are likely to affect cichlid visual sensitivity and contribute the body coloration diversity. Nonetheless, as no sexual dimorphism in visual systems in studied cichlids is identified, no evidence for a contribution of Fisher's run-away sexual selection to sexual body coloration dimorphism in cichlids is found.

Chapter I

The seahorse genome and the evolution of its specialized morphology

QIANG LIN*, SHAOHUA FAN*, YANHONG ZHANG1*, MENG XU*, HUIXIAN ZHANG*, YULAN YANG*, ALISON P. LEE, JOOST M. WOLTERING, VYDIANATHAN RAVI, HELEN M GUNTER, WEI LUO, ZEXIA GAO, ZHI WEI LIM, GENG QIN, RALF F SCHNEIDER, XIN WANG, PEIWEN XIONG, GANG LI, KAI WANG, JIUMENG MIN, CHI ZHANG, YING QIU, JIE BAI, WEIMING HE, CHAO BIAN, XINHUI ZHANG, DAI SHAN, HONGYUE QU, YING SUN, QIANG GAO, LIANGMIN HUANG, QIONG SHI, AXEL MEYER & BYRAPPA VENKATESH

NATURE, Vol. 540, pages 395–399, 2016

Abstract

Seahorses are a beautiful example of Darwin's 'endless forms most beautiful' that have fascinated biologists for centuries. Besides unique features such as male pregnancy, their specialized morphology includes a toothless tubular mouth, a body covered with bony plates, and absence of caudal and pelvic fins. To understand the genetic basis of their iconic morphology, we sequenced and analyzed the genome of a seahorse (*Hippocampus comes*). The seahorse genome is the fastest evolving fish genome sequenced so far and has lost a substantially higher number of potential *cis*-regulatory elements than other teleosts. An expanded astacin metalloprotease gene family is highly expressed in the male brood pouch. The seahorse has lost enamel matrix protein-coding P/Q-rich SCPP genes, which might have led to the loss of mineralized teeth. A master control limb development gene, *tbx4*, is lost in seahorse. Knockout of *tbx4* in zebrafish recapitulated the 'pelvic fin-loss' phenotype of seahorses, linking genotype to phenotype.


INSIGHT
Polymers
nature
THE INTERNATIONAL WEEK
OF SCIENCE
The rapidly evolving genome of the tiger tail seahorse PAGE 395
EVOLUTION AT A GALLOP
ECONOMICS
A CURE FOR POVERTY
Can Howard Rosling's statistics make the difference
PAGE 330
MEDICINE
XXXX XXXX BIOMATERIALS
Two-line explanation goes in here too
PAGE 335
BIOTECHNOLOGY
TWO LINE TEASER HEERE
Two-line explanation goes in here too
PAGES XXX & 414
NATURE.COM/NATURE
15 December 2016 £10
Vol. 540, No. 7633
9 770028 083095 28

Introduction

The natural world abounds with astonishing morphological innovations resulting in 'endless forms most beautiful' (Leysen *et al.* 2011). As such, these endless forms are excellent models for understanding the genetic basis of evolutionary novelties. Members of the teleost family *Syngnathidae* (seahorses, pipefishes and seadragons) (Fig. S.I.1), comprising approximately 300 species, are one such group exhibiting a complex array of morphological innovations and reproductive behaviors. They exhibit specialized phenotypes such as an elongated snout with a small terminal mouth, fused jaws, absent pelvic and caudal fins, and an extended body covered with an armor of bony plates instead of scales (Stölting & Wilson 2007) (Fig. I.1a). Syngnathids are also unique among vertebrates due to their 'male pregnancy' whereby males nourish developing embryos in a brood pouch until hatching and parturition occurs (Wilson *et al.* 2001). In addition, members of the subfamily *Hippocampinae* (seahorses) exhibit other derived features such as the lack of a caudal fin, a characteristic prehensile tail, and a vertical body axis (Near *et al.* 2013) (Fig. I.1a). To understand the genetic basis of seahorse's specialized morphology and reproductive system, we sequenced the genome of a seahorse (tiger tail seahorse, *Hippocampus comes*) and carried out comparative analyses with the genomes of other ray-finned fishes.

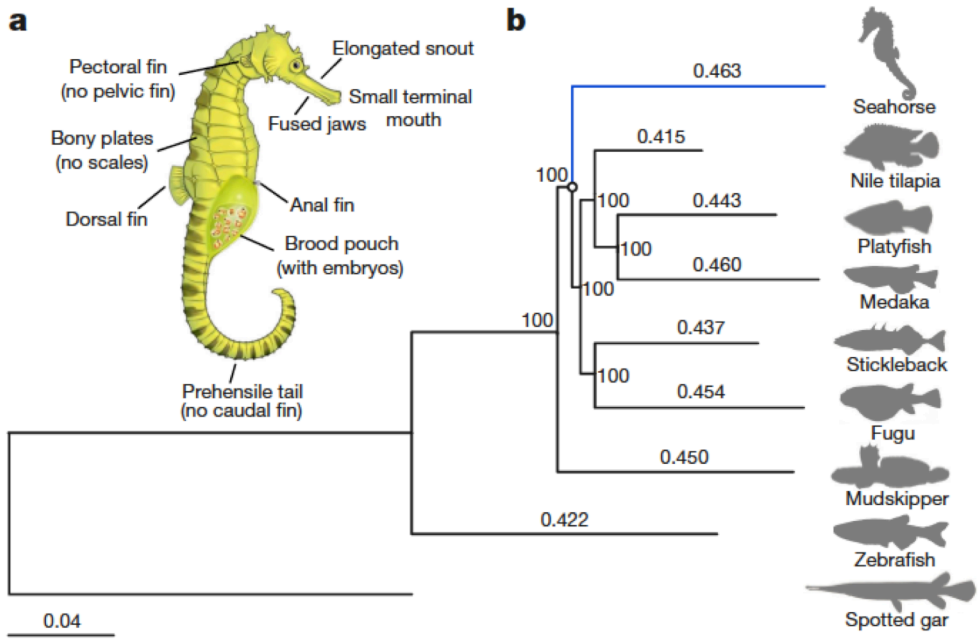


Fig. I.1 Adaptations and evolutionary rate of *H. comes*. a, Schematic diagram of a pregnant male seahorse. b, The phylogenetic tree generated using protein sequences. The values on the branches are the distances (number of substitutions per site) between each of the teleost fishes and the spotted gar (outgroup).

Genome assembly and annotation

The *H. comes* genome was sequenced from a single male individual using the Illumina HiSeq 2000 platform. After filtering low-quality and duplicate reads, 132.13 Gb (approximately 190-fold coverage of the estimated 695 Mb genome) of filtered reads from libraries with insert sizes ranging from 170 bp to 20 kb were retained for assembly. The filtered reads were assembled using SOAPdenovo (v2.04) to yield a 501.6 Mb assembly with an N50 contig size and N50 scaffold size of 34.7 kb and 1.8 Mb, respectively. Total RNA from combined soft tissues of *H. comes* was sequenced using RNA-seq and assembled de novo. The *H. comes* genome assembly is of high quality, as >99% of the *de novo* assembled transcripts (76,757 out of 77,040) could be mapped to the assembly; and 243 out of 248 CEGMA genes are complete in the assembly.

We predicted 23,458 genes in the genome of *H. comes* based on homology and by mapping the RNA-seq data of *H. comes* and a closely related species, the lined seahorse (*Hippocampus erectus*), to the genome assembly (Online Supplementary Information). More than 97% of the predicted genes (22,941 genes) either have homologs in public databases (Swissprot, Trembl and KEGG) or are supported by assembled RNA-seq transcripts. Analysis of gene family evolution using a maximum likelihood framework revealed an expansion of 25 gene families (261 genes; 1.11%) and contraction of 54 families (96 genes; 0.41%) in the *H. comes* lineage (Fig. S.I.2 and Online Supplementary Table 4.1 and 4.2).

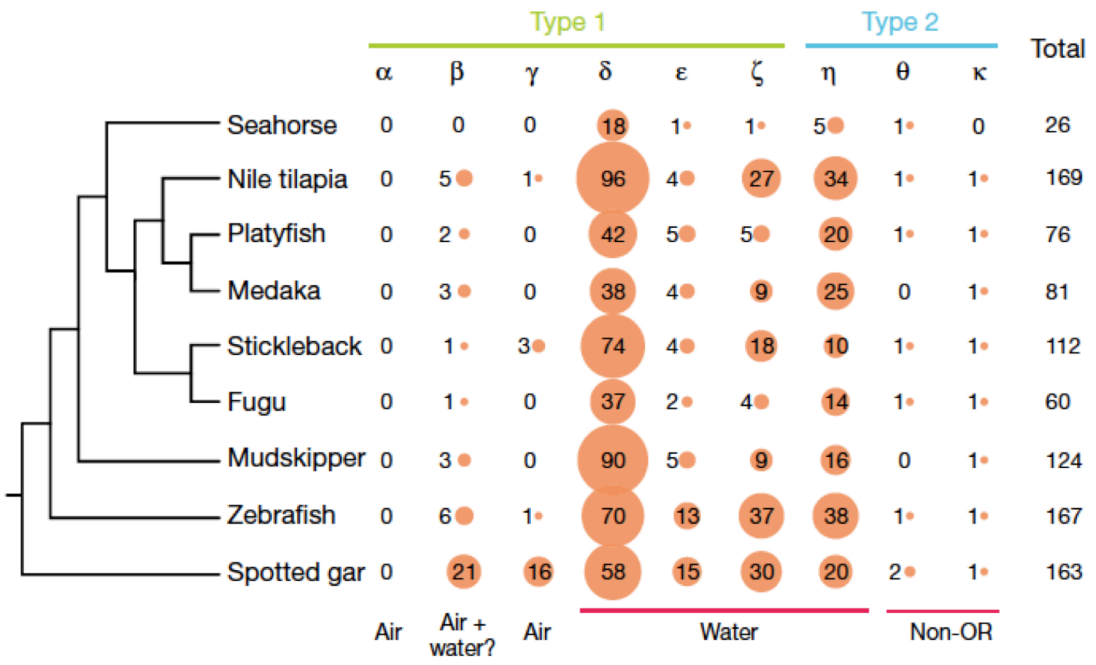


Fig. I.2 OR genes in *H. comes* and other ray-finned fishes. 'Air' and 'water' refer to the detection of airborne and water-soluble odorants, respectively. The sizes of the orange circles represent the number of OR genes of a particular category.

Transposable elements (TEs) comprise around 24.8% (124.5 Mb) of the *H. comes* genome, with class II DNA transposon being the most abundant class (9%; 45 Mb). Only one wave of TE expansion was identified with no evidence for a recent TE burst (Kimura divergence <= 5) (Online Supplementary Fig. 3).

Phylogenomics and evolutionary rate

The phylogenetic relationship between *H. comes* and other teleosts were determined using a genome-wide set of 4,122 one-to-one orthologous genes (Online Supplementary Note 4.2). The phylogenetic analysis (Fig. I.1b) showed that *H. comes* is a sister group to other percomorph fishes analyzed (stickleback, *Gasterosteus aculeatus*; medaka, *Oryzias latipes*; Nile tilapia, *Oreochromis niloticus*; fugu, *Takifugu rubripes*; and platyfish, *Xiphophorus maculatus*) with the exception of blue-spotted mudskipper (*Boleophthalmus pectinirostris*), a member of the family *Gobiidae*. Our inference, which placed the mudskipper as the outgroup, differs from that of a previous phylogenetic analysis based on fewer protein-coding genes that had placed syngnathids as an outgroup (Tajima 1989). Estimated divergence times of *H. comes* and other teleosts calculated using MCMCTree suggest that *H. comes* diverged from the other percomorphs approximately 103.8 million years ago, during the Cretaceous Period (Fig. S.I.2). Interestingly, the branch length of *H. comes* is longer than that of other teleosts indicating a faster evolutionary rate of its proteins compared to other teleosts analyzed in this study (Fig. I.1b). This result was found to be statistically significant by both relative rate test (Nei & Kumar 2000) and two cluster analysis (Bailly *et al.* 2003) (Online Supplementary Table 4.3 and Table 4.4). To determine whether the neutral nucleotide substitution rate of *H. comes* is also higher, we generated a neutral tree based on four-fold degenerate sites and calculated the pairwise distance of each teleost to spotted gar (outgroup) (Online Supplementary Fig. 4.4). The pairwise distance of *H. comes* was again greater than other teleosts, indicating that the neutral evolutionary rate of *H. comes* is also significantly higher than that of other teleosts. Thus, the overall nucleotide substitution rate of the sluggish seahorse is significantly higher than those of other teleosts. However, the reasons for such an accelerated evolutionary rate are unclear.

Loss of genes

Gene loss or loss of function can contribute to evolutionary novelties and can be positively selected for (MacArthur *et al.* 2007; Kawasaki 2011). We identified several genes that are lost in the seahorse compared to other sequenced teleost genomes. Secretory calcium-binding phosphoprotein (SCPP) genes encode extracellular matrix (ECM) proteins that are involved in the formation of mineralized tissues such as bone, dentin, enamel and enameloid. Bony vertebrate genomes encode multiple SCPP genes that can be divided into

two groups, the acidic and the proline/glutamine (P/Q)-rich SCPP genes. Acidic SCPPs regulate the mineralization of collagen scaffolds in bone and dentin whereas the P/Q-rich SCPPs are primarily involved in enamel or enameloid formation (Louchart & Viriot 2011). Analysis of the *H. comes* genome and the transcriptomes of *H. comes* and *Hippocampus erectus* revealed that both contain two acidic SCPP genes, *Scpp1* and *Spp1* (Fig. S.I.4). However, no intact P/Q-rich gene could be identified. The only P/Q-rich gene present in the *H. comes* genome assembly, *Scpp5*, is represented by only three out of ten exons indicating that it has become a pseudogene. Seahorses and pipefish (Family *Syngnathidae*) are toothless, a phenomenon known as edentulism. Besides syngnathids, edentulism has occurred convergently in several other vertebrate lineages (Meredith *et al.* 2014), the most notable ones being birds (Deméré *et al.* 2008), turtles, and some mammals such as baleen whales, pangolins and anteaters (Zhang *et al.* 2014). The loss of teeth in birds, turtles and mammals has been attributed to inactivating mutations in one or more P/Q-rich enamel-specific *SPCC* genes such as *ENAM*, *AMEL*, *AMBN* and *AMTN*, and the dentin-specific gene, *DSPP* (Deméré *et al.* 2008; Yamanoue *et al.* 2010). In the case of *H. comes*, the complete loss of functional P/Q-rich *SCPP* genes may explain the loss of mineralized teeth. Animals use their sense of smell or olfaction for finding food, mates and avoiding predators. Olfaction is mediated by olfactory receptors (ORs) that constitute the largest family of G-protein coupled receptors. The *H. comes* genome contains a significantly smaller repertoire of OR genes than other teleosts (P-value < 0.05, Wilcoxon rank sum test). Our sensitive search pipeline (based on TblastN and Genewise) and manual inspection were able to identify only 26 OR genes in the *H. comes* genome - the smallest OR repertoire identified in any ray finned fish genome analyzed so far (60 to 169 OR genes) (Fig. I.2 and Fig. S.I.5). This finding suggests that seahorses may not rely on their sense of smell as extensively as other fishes.

A derived phenotype of seahorse and other syngnathids is the complete lack of pelvic fins (Harris 1936; Kuiter 2000). Pelvic fins are homologous to tetrapod hind limbs and primarily serve a role in body trim and subtle swimming maneuvers during teleost locomotion (Gosline 1980; Tanaka *et al.* 2005; Standen 2008). In addition, pelvic spines play an important role in protection against predators (Kuiter 2000). Pelvic fin loss has occurred independently in several teleost lineages including *Tetraodontidae* (e.g. pufferfishes), *Anguillidae* (eels) and *Gasterosteidae* (some populations of sticklebacks), and is frequently associated with a reduced pressure from predators and/or the evolution of an elongated body plan (Kuiter 2000). In pufferfish (*Fugu*), pelvic fin loss is associated with a change in the expression pattern of *Hoxd9a* (Tanaka *et al.* 2005). In freshwater populations of stickleback the loss of pelvic fins has been demonstrated to be due to deletions in the pelvic fin-specific enhancer of *pitx1* (Chan *et al.* 2010).

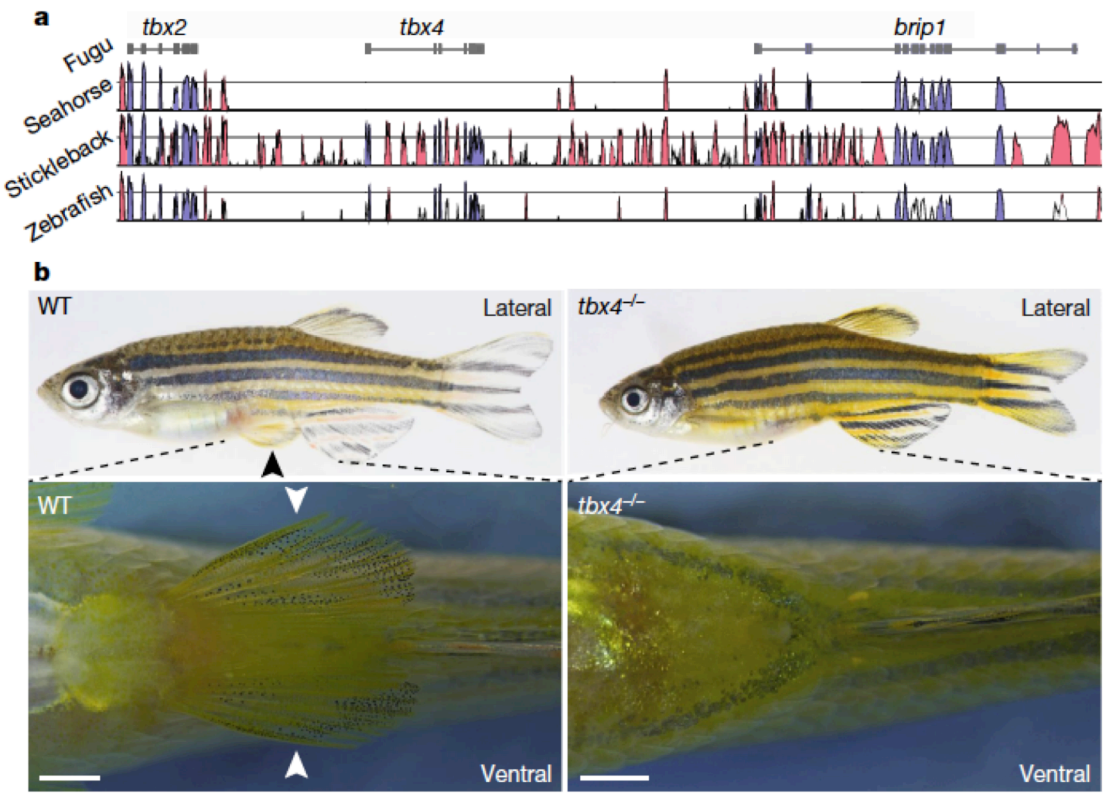


Fig. I.3 Pelvic fin loss in *H. comes* is associated with loss of *tbx4*. a, Vista plot of conserved elements in the *tbx2b-tbx4-brip1* syntenic region in fugu (reference genome), seahorse (*H. comes*), stickleback and zebrafish showing that *tbx4* is missing from this locus in seahorse. The blue and red peaks represent conserved exonic and non-coding sequences, respectively. b, Lateral (top) and ventral view (bottom) of wild-type (WT) and a representative (one out of five) F3 homozygous *tbx4*-null mutant (*tbx4*−/−) zebrafish. Bottom panel shows a close-up of the pelvic region (dashed lines indicate the approximate zoom region). Scale bar, 1 mm. Pelvic fins are indicated with black or white arrowheads in the wild-type fish. Homozygous *tbx4*-null mutants entirely lack pelvic fins without showing any other gross morphological defects.

Analysis of the *H. comes* genome and the transcriptomes of *H. comes* and *Hippocampus erectus* (see Online Supplementary Information, section 2), suggested that *tbx4*, a master control gene universally conserved in jawed vertebrates, is missing in the seahorse (Fig. I.3a) (Online Supplementary Information, section 9). To verify this, we carried out degenerate PCR using genomic DNA from *H. comes* and several other species of syngnathids and some nonsyngnathids. While the degenerate primers amplified a fragment of *tbx4* from nonsyngnathids, they failed to amplify a *tbx4* fragment from syngnathid fishes (see Online Supplementary Information, section 9). *Tbx4* is a T-box DNA-binding domain-containing transcription factor that acts as a master regulator for hind limb formation in mammals (Rodriguez-Esteban *et al.* 1999; Tamura *et al.* 1999; Arora *et al.* 2012). Loss of function of

this gene in mouse leads to a failure of hind limb formation (Rodriguez-Esteban *et al.* 1999; Tamura *et al.* 1999) as well as strong pleiotropic defects in lung (Don *et al.* 2016) and placental development (Rodriguez-Esteban *et al.* 1999). Expression of zebrafish *tbx4* specifically in pelvic fins suggests a similar role in appendage patterning in fishes (Arora *et al.* 2012). Given the major role of *tbx4* in hind limb formation in mammals, we hypothesized that its absence in *H. comes* might be associated with the loss of pelvic fins. To test this hypothesis, we generated a CRISPR/Cas9 tbx4-knock out mutant zebrafish line. The homozygous mutants are viable (unlike homozygous mouse *tbx4* mutants which fail to develop a functional allantois (Rodriguez-Esteban *et al.* 1999)) and completely lack pelvic fins without exhibiting any other gross morphological abnormalities in pectoral or median fins (Fig. I.3c and Fig. S.I.6; also see Online Supplementary Information Section 9.3, in particular Online Supplementary Fig 9.6 for additional phenotype analysis). This finding is consistent with the results of a recent study that showed that mutations in *tbx4* are associated with the loss of pelvic fins in a naturally occurring pelvic finless zebrafish strain (Kawaguchi *et al.* 2006) (also see Online Supplementary Information Section 9.3). These results show that *tbx4* plays a major role in pelvic fin formation in teleosts and indicate that the loss of pelvic fins in *H. comes* may be a direct result of the loss of *tbx4*. As syngnathids possess a body plan that is elongated as well as protected by bony plates, it is possible that relaxed selection on the underlying genetic pathway(s) allowed for the loss of this gene.

Expansion of the patristacin gene family

Male pregnancy is an evolutionary innovation unique to syngnathids. In teleost fishes, the C6AST subfamily of astacin metalloproteases such as high choriolytic enzyme (HCE) and low choriolytic enzyme (LCE) are involved in lysing the chorion surrounding the egg leading to hatching of embryos (Harlin-Cognato *et al.* 2006). A member of this subfamily, named as patristacin (*pastn*), was found to be highly expressed in the brood pouch of pregnant males of the Gulf pipefish (*Syngnathus scovelli*), leading to the suggestion that this gene has been co-opted during the evolution of male pregnancy (Whittington *et al.* 2015). A *pastn* gene was also found to be highly expressed in the brood pouch of the male big belly seahorse (*H. abdominalis*) during mid- and late pregnancy (Kawaguchi *et al.* 2015), suggesting a shared role for this gene in male pregnancy in syngnathids.

The *H. comes* genome contains six pastn genes (*pastn1* to *pastn6*; Fig. I.4a) organized in a cluster. To determine whether these genes are expressed in the brood pouch, we carried out RNA-seq analysis at different stages of brood pouch development (see Online Supplementary Information, section 2) in a related species, *Hippocampus erectus*, which is easy to obtain and breed in the laboratory. *H. comes* and *Hippocampus erectus* exhibit very similar reproductive cycles and their coding sequences are highly similar (average identity

of 93.3%; determined by aligning *Hippocampus erectus* RNA-seq transcripts to the *H. comes* genome assembly). We could identify orthologues for five of the *H. comes pastn* genes (*pastn1*, *pastn2*, *pastn3*, *pastn5* and *pastn6*) in the RNA-seq transcripts of *Hippocampus erectus* (Online Supplementary Fig. 2). Quantitative real-time PCR (qRT-PCR) analysis of these genes showed that some of them are expressed at significantly higher levels in early- and late-pregnant stages (Fig. I.4c). For example, *pastn2* is expressed at significantly higher levels in early- and late-pregnant stages compared to the non-pregnant stage, whereas

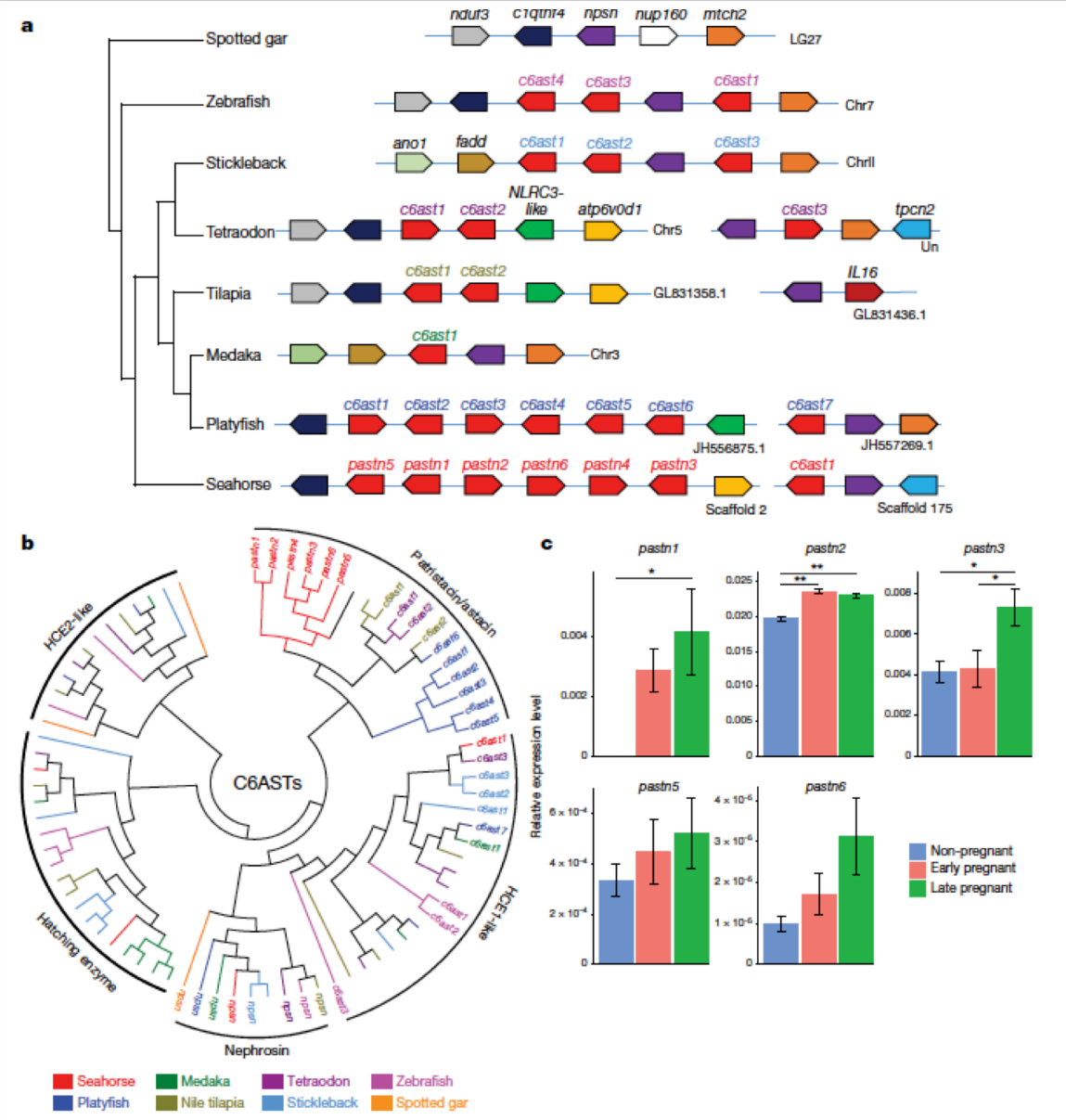


Fig. I.4 Astacin metalloproteinase gene family in ray-finned fishes. a, Astacin gene loci in various ray-finned fish genomes showing expansion of *pastn* genes in seahorse (*H. comes*) and *c6ast* genes in platyfish. Chr, chromosome. b, The phylogeny of the astacin gene family in ray-finned fishes. Only *pastn* or *c6ast* genes shown in a are labeled. Online Supplementary Fig. 10.1 shows an expanded version of the tree with all the genes

labeled. **c**, Expression patterns of *pastn* genes in relation to 18S ribosomal RNA genes in the brood pouch of male *H. erectus* determined by qRT–PCR. All data are expressed as mean Å} standard error of mean (n = 5) and evaluated by one-way analysis of variance (ANOVA) followed by Tukey's honestly significant difference test for adjusting P values from multiple comparisons (see Methods and Online Supplementary Information for details of methods). The average duration of pregnancy (from fertilization to parturition) is 17 days41. The y axis represents expression level in relation to 18S rRNA genes. *pastn1* is expressed at low levels at the non-pregnant stage, which is not clearly visible in the figure due to the large scale used. Non-pregnant: no embryos in the brood pouch; early pregnant: 2–4 days post-fertilization; late pregnant: 12–14 days postfertilization. * $P < 0.05$, * * $P < 0.01$. Note that *pastn4* is not expressed in these stages of brood pouch.

pastn1 and *pastn3* are expressed at significantly higher levels during the late-pregnant stage compared to non-pregnant stage (Fig. I.4c). This expression pattern suggests a role for these *pastn* genes in brood pouch development and/or hatching of embryos within the brood pouch prior to parturition.

Interestingly, the platyfish (*Xiphophorus maculatus*), in which fertilization and hatching of eggs occur within the maternal body (ovoviviparity), contains a cluster of six *c6ast* genes (Fig. I.4a), with potential hatching enzyme-like activity (Bejerano *et al.* 2004). Phylogenetic analysis of *c6ast* family genes in *H. comes*, platyfish and other fishes showed that *H. comes pastn* genes and platyfish *c6ast* genes form separate clades (Fig. I.4b) indicating that they have expanded independently in the two lineages. Thus, this is an interesting instance of a gene family (C6AST subfamily of astacin metalloproteases) that has undergone expansion independently in different teleost lineages and acquired novel expression patterns and functions associated with similar evolutionary innovations (i.e. ovoviviparity in platyfish and male pregnancy in seahorse).

Loss of conserved noncoding elements

Vertebrate genomes contain thousands of noncoding elements that are under purifying selection (Venkatesh *et al.* 2006; Navratilova *et al.* 2009; Lindblad-Toh *et al.* 2011). Many of these conserved noncoding elements (CNEs) function as *cis*-regulatory elements such as enhancers, repressors and insulators (Visel *et al.* 2008; Attanasio *et al.* 2013). Evolutionary loss of CNEs plays important roles in phenotypic differences and morphological innovations (Naiche & Papaioannou 2003; Sabherwal *et al.* 2007; McLean *et al.* 2011). To determine the extent of loss of CNEs in seahorse, we predicted genome-wide CNEs in *H. comes* and four other percomorph fishes (stickleback, fugu, medaka and Nile tilapia) using zebrafish as the reference genome (see Online Supplementary Information). We identified 239,976 CNEs

(average size of 168 bp) that are conserved in zebrafish and at least one of the five percomorph fishes (Online Supplementary Table 6.1). To determine the extent to which CNEs are lost in *H. comes*, we searched for CNEs that are uniquely lost in each of the percomorph fishes. We restricted our analyses to a high-confidence set of CNEs situated in gap-free syntenic intervals (Online Supplementary Table 6.5). Interestingly, *H. comes* was found to have lost a substantially higher number of CNEs (1,612 CNEs) compared to other percomorphs (fugu, 1,050 CNEs; stickleback, 843 CNEs; medaka, 335 CNEs; Nile tilapia, 281 CNEs) (Online Supplementary Table 6.6).

Analysis of zebrafish CNEs that are lost in *H. comes* indicated that they are present in the neighborhood of 728 genes enriched in functions such as regulation of transcription, regulation of the fibroblast growth factor receptor signaling pathway, embryonic pectoral fin morphogenesis, steroid hormone receptor activity and O-acetyltransferase activity (Online Supplementary Tables 6.8 and 6.9). The top 20 genes adjacent to regions with the highest number of CNEs lost in *H. comes* include *Sall1a*, *Shox* and *Irx5a* (Online Supplementary Tables 6.10 and 6.11) which are involved in the development of limbs, nervous system, kidney, heart and skeletal system. Altered expression patterns of these genes can potentially lead to altered morphological phenotypes. For example, loss of regulatory regions of the human *SHOX* gene is the cause of Leri-Weill dyschondrosteosis, a dominantly inherited skeletal dysplasia that is characterized by moderate short stature caused by short mesomelic limb segments (Superti—Furga *et al.* 1998; Indjeian *et al.* 2016).

To verify the potential cis-regulatory functions of CNEs that were lost in *H. comes*, we assayed the function of seven selected zebrafish CNEs that were uniquely lost in *H. comes*. Of the seven CNEs assayed in transgenic zebrafish, four CNEs drove reproducible patterns of reporter gene expression in F1 embryos (Fig. S.I.7; Online Supplementary Table 6.12). Thus, our transgenic assay suggests that many of the CNEs lost in *H. comes* are indeed *cis*-regulatory elements and their loss can potentially alter the expression pattern of their target genes. Thus, the loss of a large number of CNEs in the seahorse lineage might have played a significant role in its distinct phenotype.

Summary

Seahorses possess one of the most highly specialized morphologies and reproductive behaviours. We sequenced the genome of the tiger tail seahorse and performed comparative analysis with other teleost fishes. Our genome-wide analysis highlights several aspects that may have contributed to the highly specialized body plan and male pregnancy of seahorses. These include a higher protein and nucleotide evolutionary rate, loss of genes and expansion of gene families, with duplicated genes exhibiting new expression patterns, and loss of a selection of potential *cis*-regulatory elements. It is becoming recognized that

evolutionary changes in *cis*-regulatory elements, particularly the loss and gain of enhancers, might play a major part in the evolution of morphological innovations and phenotypic changes across species (Chan *et al.* 2010; McLean *et al.* 2011; Attanasio *et al.* 2013; Indjeian *et al.* 2016).

Male pregnancy is a unique developmental feature of seahorses and pipefishes (family Syngnathidae, comprising 57 genera and approximately 300 species). In the seahorse genome, the astacin subfamily of *c6ast* metalloprotease genes has undergone tandem duplications giving rise to six genes. This subfamily of metalloprotease includes the hatching enzyme (also known as choriolysin), HCE-like and HCE2-like enzymes that are responsible for hatching of embryos in fishes. Of the six duplicated genes in seahorse, five are highly expressed in the male brood pouch, suggesting that they may be involved in male pregnancy, possibly through rewiring of their regulatory network. The loss of pelvic fins in seahorse is associated with the evolution of an armour-like covering of its body and gain of an elongated, flexible, substrate-gripping tail. By combining comparative genomics and gene-knockout experiments in zebrafish, we suggest that loss of the *tbx4* gene may have a role in this phenotype in seahorse. The loss of mineralized teeth in seahorse is associated with the fusion of the jaws into a tube-like snout and a small mouth, which is extremely efficient in sucking small food items that are abundant in the benthic environment. In teleosts, P/Q-rich SCPP genes are involved in the mineralization of enameloid, which is the equivalent of enamel in tetrapods10. The seahorse genome does not contain any intact P/Q-rich SCPP genes that code for enamel matrix proteins, suggesting that the loss of these genes could have played a part in the loss of its mineralized teeth. Our analyses of the *H. comes* genome sequence and comparative genomics with other teleosts highlighted several genetic changes that may be involved in the evolution of the unique morphology of seahorses.

Methods

Genome sequencing and assembly

Genomic DNA of a single male *H. comes* was used to construct eleven libraries including short-insert (170 bp, 500 bp, 800 bp) and mate-paired (2 kb, 5 kb, 10 kb, 20 kb) libraries and sequenced on the Illumina HiSeq 2000 sequencing platform. In total, we obtained around 218 Gb of raw sequence data (Online Supplementary Table 1.1). The genome was assembled using SOAPdenovo2.04 (Luo *et al.* 2012) with default parameters. No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

RNA sequencing and analysis

In total, 19 RNA-seq libraries were constructed, including two libraries from combined soft tissues (brain, gills, intestine, liver and muscle) from a male and a female *H. comes* (Online Supplementary Table 2.1); and 17 libraries of five developmental stages of embryos and different stages of brood pouch development such as the juvenile stage, rudimentary stage, pre-pregnancy stage, pregnancy stage, and post pregnancy stage, using RNA from the lined seahorse (*Hippocampus erectus*) (Online Supplementary Information, section 2). All libraries were prepared using Illumina TruSeq RNA sample preparation kit according to the manufacturer's instructions (Illumina, San Diego, CA, USA) and sequenced using Illumina HiSeq 2000 platform. The RNA-seq reads were either *de novo* assembled using Trinity (Grabherr *et al.* 2011) or mapped to the *H. comes* genome using Tophat (Trapnell *et al.* 2009) with default parameters, and subsequently analysed using in-house Perl scripts. The differential expression of genes at different stages of brood pouch development was determined using the method developed previously (Yu *et al.* 2006). The RNA-seq results were validated using qRT–PCR, with five biological replicates for each stage. All data were expressed as mean +/- standard error of mean and were evaluated by one-way ANOVA followed by Tukey's honestly significant difference test for adjusting *P* values from multiple comparisons. Results were considered to be statistically significant for *P* values < 0.05.

Genome annotation

Annotation of the *H. comes* genome was carried out using the Ensembl gene annotation pipeline which integrated *ab initio* gene predictions and evidence-based gene models. Briefly, protein sequences of *D. rerio*, *G. aculeatus*, *O. latipes*, *T. rubripes* and *T. nigroviridis* were downloaded from Ensembl (release 75) and mapped to the genome using TblastN (Kent 2002) with the parameter "-evalue 1E-5". Second, high scoring segment pairs (HSPs) from blast were concatenated using Solar (in-house software, version 0.9.6). Third, the concatenated segments were aligned using GeneWise (Birney *et al.* 2004) to refine the gene models. Finally, we filtered the alignments that showed alignment rates less than 50% of the full-length copies and filtered redundant alignments based on the GeneWise score. In addition, *H. comes* transcripts (female_transcript and male_transcript) and *H. erectus* transcripts (Juv_brain, Juv_body, Rud_testis and PreP_pouch) were used to assist in the gene model prediction. We annotated the predicted gene models using Swiss-Prot, TrEMBL, NCBI NR database, and KEGG databases (Online Supplementary Table 3.4).

Expansion and contraction of gene families

We used CAFE (version 2.1), a program for analysing gene family expansion and contraction under maximum likelihood framework. The gene family results from TreeFam pipeline and

the estimated divergence time between species were used as inputs. We used the parameters "-p 0.01, -r 10000, -s" to search the birth and death parameter (λ) of genes, calculated the probability of each gene family with observed sizes using 10000 Monte Carlo random samplings, and reported birth and death parameters in gene families with probability less than 0.01. For the gene family expansion and contraction analysis in *H. comes*, we first filtered out gene families without homology in the SWISS-PROT database to reduce the potential false positive expansions or contractions caused by gene prediction. Besides, the families that contained sequences that have multiple functional annotations were also removed (Online Supplementary Tables 4.1 and 4.2).

Phylogenetic analysis

We obtained 4,122 one-to-one orthologous genes from the gene family analysis (Online Supplementary Information, section 4.1). The protein sequences of one-to-one orthologous genes were aligned using MUSCLE (Edgar 2004) with the default parameters. We then filtered the saturated sites and poorly aligned regions using trimAl (Capella-Gutiérrez *et al.* 2009) with the parameters "-gt 0.8 –st 0.001 –cons 60". After trimming the saturated sites and poorly aligned regions in the concatenated alignment, 2,128,000 amino acids were used for the phylogenomic analysis. The trimmed protein alignments were used as a guide to align corresponding coding sequences (CDSs). The aligned protein and the fourfold degenerate sites in the CDSs were each concatenated into a super gene using an in-house Perl script.

The phylogenomic tree was reconstructed using RAxML version 8.1.19 (Stamatakis 2006) based on concatenated protein sequences. Specifically, we used the PROTGAMMAAUTO parameter to select the optimal amino acid substitution model, specified spotted gar as the outgroup, and evaluated the robustness of the result using 100 bootstraps. To compare the neutral mutation rate of different species, we also generated a phylogeny based on fourfold degenerate sites. The phylogenomic topology was used as input and the "-f e" option in RAxML was used to optimize the branch lengths of the input tree using the alignment of fourfold degenerate sites under the general time reversible (GTR) model as suggested by ModelGenerator version 0.85 (Stamatakis *et al.* 2008). We calculated the pairwise distances to the outgroup (spotted gar) based on the optimized branch length of the neutral tree using the 'cophenetic.phylo' module in the R-package 'ape' (Paradis *et al.* 2004). The Bayesian relaxed-molecular clock (BRMC) method, implemented in the MCMCTree program (Yang 2007), was used to estimate the divergence time between different species. The concatenated CDS of one-to-one orthologous genes and the phylogenomics topology were used as inputs. Two calibration time points based on fossil records, *O. latipes–T. nigroviridis* (~ 96.9–150.9 million years ago (Mya)), and *D. rerio–G.*

aculeatus (~ 149.85– 165.2 Mya) (http://www.fossilrecord.net/dateaclade/index.html), were used as constrains in the MCMCTree estimation. Specifically, we used the 'correlated molecular clock' and 'REV' substitution model in our calculation. The MCMC process was run for 5,000,000 steps and sampled every 5,000 steps. MCMCTree suggested that *H. comes* diverged from the common ancestor of stickleback, Nile tilapia, platyfish, fugu, and medaka approximately 103.8 Mya, which corresponds to the Cretaceous period.

Analysis of OR genes

We downloaded protein sequences of 1,417 OR gene family members from NCBI and mapped them to *H. comes* genome using Tblastn with "E-value ≤ 1e-10" and "alignment rate ≥ 0.5". Solar (in-house software, version 0.9.6) was used to join high-scoring segment pairs (HSPs) between each pair of protein mapping result. We retained alignments with an alignment rate of more than 70% and a mapping identity of more than 40%. Subsequently, the protein sequences were mapped to the genome using GeneWise and extended 280 bp upstream and downstream to define integrated gene models. For phylogenetic analysis, protein sequences were aligned using MUSCLE and a JTT+ gamma model was used in a maximum-likelihood analysis using PhyML to construct a phylogenetic tree.

Evidence for loss of tbx4 in H. comes

The synteny analysis of *tbx2b-tbx4-brip1* region of *H. comes*, stickleback, fugu and zebrafish using Vista shows that *tbx4* was lost in *H. comes* (Fig. I.3). To exclude the scenario that the absence of *tbx4* in the *H. comes* genome sequence is due to an assembly error, we first validated the micro-synteny region of *tbx2b-tbx4-brip1* region in *H. comes* using a PCR-based genomic walk strategy. Briefly, 28 primer pairs (Online Supplementary Table 9.1) were designed for overlapping amplicons to 'walk' from the end of *tbx2b* to the start of *brip1*. Amplicon size and partial end sequencing of these products did not indicate any anomalies in the assembly of the *H. comes tbx4* 'ghost locus'. In addition, we carried out the following analyses: (1) searched the *H. comes* genome (TblastN) using Tbx4 protein from zebrafish and Nile tilapia and were unable to find a *tbx4* gene; (2) searched the *H. comes* genome using only the domain sequence of Tbx4 protein but were unable to find a *tbx4* gene; (3) searched *H. comes* and *H. erectus* transcriptome data for *tbx4* (TblastN) using Tbx4 protein from zebrafish and Nile tilapia but were unable to find any matching transcript; (4) searched *H. comes* and *H. erectus* transcriptome data with the domain sequence as well and did not find any remnant of a *tbx4* gene; and (5) predicted CNEs in the 'ghost' *tbx4* locus of *H. comes* using the fugu *tbx4* locus as the reference (base) (Online Supplementary Fig. 9.3). We used the CNEs present in the other fish genome loci (that were absent in *H. comes*) to search the *H. comes* genome to rule out the possibility that they may be present elsewhere in the

genome. We were unable to find any of these CNEs in the *H. comes* genome. Finally, we conducted degenerate PCR experiments to ascertain if the *tbx4* gene is missing in *H. comes*. Using a combination of four forward and two reverse primers (Online Supplementary Table 9.1), we checked for the presence of *tbx4* in seven species of *Hippocampus* (including *H. comes* and *H. erectus*), five species of pipefish (four from the genus *Syngnathus* and one species of *Corythoichthys*) (all from the family Syngnathidae that lack pelvic fins); ghost pipefish (*Solenostomus*) and the trumpetfish (Aulostomidae) which are closely related to the Syngnathidae but possess pelvic fins; and five other teleost species that possess pelvic fins (Online Supplementary Figs 9.1 and 9.2).

Generation of mutant tbx4 zebrafish

We used a CRISPR–Cas9 strategy to generate a *tbx4* mutant zebrafish line. Two guide RNAs (gRNAs) were designed targeting zebrafish *tbx4* in the 5′ end of the sequence that is before or inside the DNA-binding TBOX domain (Online Supplementary Fig. 9.4). gRNAs were cloned using synthesized oligonucleotides into the pT7gRNA vector as described previously54 (oligonucleotide sequences given in Online Supplementary Table 9.2). gRNAs were synthesized from this vector after linearization with BamH1-HF (NEB R3136T), transcribed using the MEGAscript T7 Transcription Kit (Thermo Fischer Scientific AM1334) and purified using the mirVana miRNA isolation kit (Thermo Fischer Scientific AM1560). Cas9 mRNA was synthesized from the Cs2+ Cas9 vector using the mMessage mMachine Sp6 Transcription Kit (Thermo Fischer Scientific AM1340) and purified using the RNA cleanup protocol from the RNAeasy mini kit (Qiagen 74104).

Zebrafish from a wild caught strain were injected at the one-cell stage with ~ 50 ng gRNA and ~ 90 ng Cas9 RNA. These F0 fish were raised to maturity and genotyped using fin clipping, DNA isolation and PCR spanning the target site (genotyping primers given in Online Supplementary Table 9.2). PCR products were analysed for mutations as described previously54 using T7 endonuclease (NEB M0302L). Mosaic mutant F0 fish were outcrossed to AB wild-type fish and embryos were batch genotyped for transmission of the mutation using PCR and T7 endonuclease. Mutant PCR products were cloned into the pGEM-T vector (Promega, Madison, WI) and sequenced to identify carrier fish transmitting a frameshift mutation. These carrier fish were crossed again to AB wild type and the resulting F1 fish were raised to maturity. The F1 were genotyped using fin clipping, DNA isolation, PCR, T7 endonuclease to identify heterozygous mutant fish followed by cloning and sequencing of the mutant PCR products to validate presence of the frameshift allele. The CRISPR–Cas9 mutation strategy is schematically shown in Fig. S.I.5.

In the F0 mutant *tbx4* fish we observed pelvic fin loss at low frequency. gRNA#1 gave 3/42 fish with either double- or single-sided pelvic fin loss whereas 1/34 had single-

sided pelvic fin loss for gRNA#2 (Fig. S.I.5). We observed mutant allele transmission for both gRNA#1 and gRNA#2 but failed to identify a deletion leading to a frameshift mutation for gRNA#2 so no stable line was generated for this CRISPR. For gRNA#1 we identified several frameshift mutants, one of which was further analyzed. This mutant has a deletion/replacement mutation in which eight nucleotides are replaced by three nucleotides, leading to an effective 5 bp deletion and the introduction of a frameshift mutation (Fig. S.I.5). This mutation introduces a downstream STOP codon leading to a severely truncated protein lacking the DNA binding domain (Online Supplementary Figs 9.4 and 9.5). The mutant line is maintained on an AB wild-type background.

Loss of CNEs

Using zebrafish as the reference genome, whole-genome alignments of six teleost fishes were generated. The soft-masked genome sequence for zebrafish (Zv9, April 2010) was downloaded from the Ensembl release-75 FTP site. The following soft-masked genome sequences were downloaded from the UCSC Genome Browser: stickleback (gasAcu1, February 2006), fugu (fr3, October 2011), medaka (oryLat2, October 2005), Nile tilapia (oreNil2, February 2012). The *H. comes* genome sequence (hipCom0) was repeat-masked using WindowMasker (from NCBI BLAST+ package v.2.2.28) with additional parameter "-dust true". About 32% (158.1/501.6 Mb) of the *H. comes* genome was masked using this method.

Only chromosome sequences of zebrafish were aligned while unplaced scaffolds were excluded. The reference (zebrafish) genome was split into 21 Mb sequences with 10-kb overlap, while the percomorph fish genomes (*H. comes*, stickleback, fugu, medaka and Nile tilapia) were split into 10 Mb sequences with no overlap. Pairwise alignments were carried out using Lastz v.1.03.54 (Harris 2007) with the following parameters: –strand = both–seed = 12of19–notransition–chain–gapped– gap = 400,30–hspthresh = 3000–gappedthresh = 3000–inner = 2000–masking = 50– ydrop = 9400–scores = HoxD55.q–format = axt. Coordinates of split sequences were restored to genome coordinates using an in-house Perl script. The alignments were reduced to single coverage with respect to the reference genome using UCSC Genome Browser tools 'axtChain' and 'chainNet'. Multiple alignments were generated using Multiz.v11.2/roast.v3 (Blanchette *et al.* 2004) with the tree topology "(Zv9 (hipCom0 ((fr3 gasAcu1) (oryLat2 oreNil2))))".

Fourfold degenerate (4D) sites of zebrafish genes (Ensembl release-75) were extracted from the multiple alignments. These 4D sites were used to build a neutral model using PhyloFit in the rphast v.1.5 package (Hubisz *et al.* 2010) (general reversible "REV" substitution model). PhastCons was then run in rho-estimation mode on each of the zebrafish chromosomal alignments to obtain a conserved model for each chromosome.

These conserved models were averaged into one model using PhyloBoot. Subsequently, conserved elements were predicted in the multiple alignments using PhastCons with the following inputs and parameters: the neutral and conserved models, target coverage of input alignments = 0.3 and average length of conserved sequence = 45 bp. To assess the sensitivity of this approach in identifying functional elements, the PhastCons elements were compared against zebrafish protein-coding genes. Eighty per cent of protein-coding exons (197,508/245,556 exons) were overlapped by a conserved element (minimum coverage 10%), indicating that the identification method was fairly sensitive.

A CNE was considered present in a percomorph genome if it showed coverage of at least 30% with a zebrafish CNE in Multiz alignment. To identify CNEs that could have been missed in the Multiz alignments due to rearrangements in the genomes, or due to partitioning of the CNEs among teleost fish duplicate genes, we searched the zebrafish CNEs against the genome of the percomorph using BLASTN (E < 1 Å~ 10−10; ≥ 80% identity; ≥ 30% coverage). Those CNEs that had no significant match in a percomorph genome were considered as missing in that genome. To account for CNEs that might have been missed due to sequencing gaps, we identified gap-free syntenic intervals in zebrafish and the percomorph genomes, and generated a set of CNEs that were missing from these intervals. These CNEs represent a high-confidence set of CNEs missing in the percomorph fishes and thus were used for further analysis. Functional enrichment of genes associated with CNEs was carried out using the GREAT software (McLean *et al.* 2010) with each CNE assigned to the genes with the nearest transcription start site and within 1 Mb in the zebrafish genome, and significantly enriched functional categories identified based on a hypergeometric test of genomic regions (false discovery rate (FDR) q value < 0.05). We identified the statistically significant gene ontology biological process terms, molecular function terms and zebrafish phenotype descriptions of the genes that are associated with CNEs.

We also predicted CNEs in the Hox clusters of *H. comes* and other representative teleost fishes using the global alignment program MLAGAN. Orthologous Hox clusters were aligned using MLAGAN with zebrafish as the reference sequence and CNEs were predicted using VISTA.

Functional assay of CNEs

Seven representative zebrafish CNEs that are lost in *H. comes* (the largest among the lost CNEs) were assayed for enhancer activity in transgenic zebrafish using GFP as the reporter gene. The CNEs were amplified by PCR using zebrafish genomic DNA as template. The products were cloned into a miniTol2 transposon donor plasmid linked to the mouse *cFos* (McFos) basal promoter and the coding sequence of GFP. Transposase mRNA was generated by transcribing cDNA *in vitro* using the mMESSAGE mMACHINE T7 kit (Ambion; Life

Technologies). The CNE-containing McFos-miniTol2 construct and transposase mRNA were co-injected into the yolk of zebrafish embryos at the one to two-cell stage. Each CNE construct was injected into 250–350 embryos and the injections were repeated on two days. The embryos were reared at 28 °C, and GFP was observed at 24, 48 and 72 h post-fertilization (hpf). The survival rate of the embryos post-injection was 70–80%. Consistent GFP expression in at least 20% of F0 embryos was considered as specific expression driven by a CNE. Such embryos were reared to maturity and mated with wild type zebrafish to produce F1 lines. The expression of GFP in F1 embryos was observed under a compound microscope fitted for epifluorescence (Axio imager M2; Carl Zeiss, Germany) and photographed using an attached digital microscope camera (Axiocam; Carl Zeiss, Germany). Pigmentation was inhibited by maintaining zebrafish embryos in 0.003% N-phenylthiourea (Sigma-Aldrich, Sweden) from 8 hpf onwards. Consistent GFP expression observed in at least three lines of F1 fishes was considered as the specific expression driven by a CNE.

All animals were cared for in strict accordance with National Institutes of Health (USA) guidelines. The zebrafish gene knockout protocol was approved by the Institutional Animal Care and Use Committee of Sun Yat-Sen University. The zebrafish transgenic assay protocol was approved by the Institutional Animal Care and Use Committee of Biological Resource Centre, A* STAR, Singapore.

Data availability statement.

The tiger tail seahorse (*H. comes*) whole-genome sequence has been deposited in the DDBJ/EMBL/GenBank database under accession number LVHJ00000000. RNA-seq reads for *H. erectus* and *H. comes* have been deposited in the NCBI Sequence Read Archive under accession numbers SRA392578 and SRA392580, respectively.

Chapter II

How plasticity, genetic assimilation and cryptic genetic variation may contribute to adaptive radiations

RALF F. SCHNEIDER and AXEL MEYER

Accepted in MOLECULAR ECOLOGY, 2016

Abstract

There is increasing evidence that phenotypic plasticity can promote population divergence by facilitating phenotypic diversification and, eventually, genetic divergence. When a 'plastic' population colonizes a new habitat, it has the possibility to occupy multiple niches by expressing several distinct phenotypes. These initially reflect the population's plastic range but may later become genetically fixed by selection via the process of 'genetic assimilation' (GA). Through this process multiple specialized sister lineages can arise that share a common plastic ancestor – the 'flexible stem'. Here, we review possible molecular mechanisms through which natural selection could fix an initially plastic trait during GA. These mechanisms could also explain how GA may contribute to cryptic genetic variation that can subsequently be coopted into other phenotypes or traits, but also lead to nonadaptive responses. We outline the predicted patterns of genetic and transcriptional divergence accompanying flexible stem radiations. The analysis of such patterns of (retained) adaptive and nonadaptive plastic responses within and across radiating lineages can inform on the state of ongoing GA. We conclude that, depending on the stability of the environment, the molecular architecture underlying plastic traits can facilitate diversification, followed by fixation and consolidation of an adaptive phenotype and degeneration of nonadaptive ones. Additionally, the process of GA may increase the cryptic genetic variation of populations, which on one hand may serve as substrate for evolution, but on another may be responsible for nonadaptive responses that consolidate local allopatry and thus reproductive isolation.

Introduction

The ability of single genotypes to produce varying environmentally-tuned phenotypes during the individuals' development is called phenotypic plasticity (Box 1 – Glossary) (Bradshaw 1965; Colombo *et al.* 2013). Its influence on evolution and speciation has been debated controversially for decades (for reviews see: West-Eberhard 1989; Schlichting & Pigliucci 1998; Price *et al.* 2003; West-Eberhard 2003; DeWitt & Scheiner 2004; Pfennig & Pfennig 2012; Gilbert & Epel 2015; Oke *et al.* 2015; Sultan 2015). During the twentieth century, insights into how genes and allele frequencies underlie evolutionary change led to a strong focus of evolutionary research on the direct genetic effects on the fitness of individuals (Fisher 1930; Waddington 1942; Williams 1966; Smith 1989; Pigliucci & Murren 2003; Elmer *et al.* 2009; Sultan 2015). However, phenotypes are not only genetically determined (G) but can also be influenced by the environment (E) and their interaction (G x E). That is, organisms may respond phenotypically to their environment (Ghalambor *et al.* 2007; Colombo *et al.* 2013; Manousaki *et al.* 2013; Brawand *et al.* 2014; Gilbert & Epel 2015; Sultan 2015) and even shape their environments themselves (Odling-Smee *et al.* 1996; Sultan 2015). The classical view is that by 'uncoupling' the phenotype from the genotype to varying degrees, which are subject to selection themselves, plasticity shields the genotype from the direct effects of natural selection (Schlichting 2004). However, theoretical research and simulation models increasingly and convincingly predict that under certain conditions, phenotypic plasticity might nonetheless facilitate evolution, speciation and even adaptive radiations (Box 2) (e.g. West-Eberhard 2005; Ghalambor *et al.* 2007; Pfennig & McGee 2010; Ehrenreich & Pfennig 2015; Gilbert & Epel 2015). Gilbert *et al.* (2015) suggested that three main hurdles must be overcome to fully incorporate plasticity into evolutionary biology: understanding how commonly plasticity contributes to evolutionary transitions, learning about the molecular mechanisms underlying plasticity, and constructing a theoretical framework to predict population-specific evolutionary consequences of plasticity. Here, we aim to help overcome these hurdles by complementing and elaborating recently published work on genetic assimilation (e.g. Ehrenreich & Pfennig 2015; Levis & Pfennig 2016) by expanding upon important concepts about the nature of phenotypic plasticity and then discuss how plasticity may contribute to diversification, speciation and adaptive radiations. In addition, we focus on proposing how molecular mechanisms potentially mediate genetic assimilation and aim to predict patterns of genetic, transcriptional and phenotypic diversity expected in extant derived lineages that have a common phenotypically plastic ancestor. Evidence is summarized concluding that gene regulatory networks underlying plastic traits might provide the developmental variability that can serve as the substrate for the evolution of phenotypic diversity during lineage divergence and establish macro-evolutionary patterns.

Box 1. Glossary

Adaptive radiation

A process by which a multitude of new species evolves rapidly from a single common ancestor by phenotypic segregation into different ecological niches. Adaptive radiations can be induced when a founder population colonizes a new habitat with open niches.

Developmental robustness

A system is developmentally robust if environmental perturbations or DNA sequence mutations have little effect on developing phenotypes.

Developmental trajectory (in this manuscript)

Simplified concept of the specific gene regulatory pathway that leads to a trait's (usually adaptive) phenotype. Nonplastic traits only have one developmental trajectory while plastic traits have several in response to an environmental cue. As long as developmental trajectories are regularly expressed in a population, they are expected to remain under stabilizing selection.

Diversification potential

A property of a population that reflects its probability and degree of diversifying evolution when experiencing ecological opportunity. Simplified, it is the product of the ability of a population to utilize available niches, population attributes that increase its phenotypic evolvability and the probability of evolving reproductive barriers to other populations.

Ecological opportunity

A population-specific property of an environment composed of niche availability (allowing the population to persist) and niche discordance (causing diversification by differing selective regimes across niches).

Evolvability

The rate at which DNA sequence mutations can induce phenotypic alterations that can be targeted by selection.

Flexible stem hypothesis

A developmentally flexible lineage may diverge into derived linages under specific conditions. Variation in derived lineages reflects the developmental flexibility of the ancestor (i.e. its flexible stem).

Genetic accommodation

A selection-driven process during which a trait becomes more or less plastic and thus less or more heritable, respectively.

Genetic assimilation

A selection-driven process during which a trait becomes less plastic through its genetic fixation, thus removing its environmental responsiveness.

Phenotypic plasticity

The capacity of a single genotype to produce varying phenotypes according to external (e.g. environmental) cues (also 'developmental plasticity', when induced phenotypes are irreversible and set during development). Adaptive and nonadaptive plasticity produce phenotypes that are closer to or further away from the fitness optimum, respectively.

Reaction Norm

A reaction norm depicts a trait's phenotypic expressions of a single genotype across a range of environments. Plotting multiple reaction norms together allows the comparison of the level of plasticity among genotypes. For more details, see, for example, Woltereck (1913) and Schlichting and Pigliucci (1998).

The nature of phenotypic plasticity

Phenotypic plasticity is widespread in nature and has been demonstrated for most evolutionary lineages in ecologically relevant contexts. A multitude of possible environmental cues are known to induce plastic phenotypes (e.g. as summarized in West-Eberhard 2003; Schlichting 2004; Gilbert & Epel 2015). As plasticity allows organisms to deal with variable environmental conditions, it is assumed to be particularly important for sessile organisms, such as plants (Schlichting 1986; Nicotra *et al.* 2010; Turcotte & Levine 2016). However, also in the animal kingdom phenotypic plasticity has been demonstrated in many different contexts such as predator avoidance, seasonal change, social interactions or food availability. Classical examples include Daphnia species that show increased or decreased average body and helmet size when confronted with predator soluble cues or raised at different temperatures, respectively (Ostwald 1904; Dodson 1988). In Hymenopteran insects, phenotypic plasticity was shown to be responsible for the development of the caste phenotypes found in many derived lineages, such as ants or many bees and wasps, have (Hölldobler 1990). Other well-known examples include swarming behaviour in locusts (Simpson *et al.* 1999), the induction of cannibalism in spadefoot toads (Pfennig *et al.* 1993) and sex change in labrid fishes (Warner & Swearer 1991).

Generally, inducible phenotypes may either vary continuously and depend on the quantity of an environmental stimulus, or they may be discrete and follow an on/off switch mechanism ('polyphenism') (Schlichting & Pigliucci 1998; Nijhout 2003). Plasticity can thus provide a large pool of phenotypic diversity, even within a population. In spite of plasticity's

near ubiquity, the mechanisms by which it may contribute to specialization, lineage diversification and speciation remain debated (but see e.g. Foster *et al.* 2015; Ghalambor *et al.* 2015) and some even feel that a new evolutionary paradigm is necessary (e.g. Gilbert & Epel 2009; Pigliucci & Müller 2010). It is argued that the level of phenotypic plasticity, as typically depicted in reaction norms, can be the target of selection and thus evolve (Schlichting 1986; Scheiner 1993; Nussey *et al.* 2005). This implies the existence of mechanisms that increase or reduce a trait's environmental sensitivity and thus reduce or increase its heritability, respectively. West-Eberhard (2003) coined the term genetic accommodation to cover processes in which the level of phenotypic plasticity is either reduced or increased, making the trait's phenotype more or less heritable, respectively (Schlichting & Wund 2014).

Phenotypic plasticity is common in nature because it often offers advantages to organisms that live in heterogeneous environments and have to deal with varying environmental conditions in which the associated cost of phenotype (i.e. the cost of having a phenotype with suboptimal fitness) is particularly high (West-Eberhard 2003; Murren *et al.* 2015). Models and empirical studies confirmed that phenotypic plasticity increases geographic distribution ranges and enhances species persistence and fitness, particularly in temporarily heterogeneous environments, as long as the changing environmental cues predict the environmental conditions reliably (Price *et al.* 2003; Zhang 2005; Chevin *et al.* 2010; Scheiner & Holt 2012; Le Vinh *et al.* 2016; Orizaola & Laurila 2016). This kind of environmentally induced flexibility may occur in any kind of trait, such as morphology (Greenwood 1964; Pfennig *et al.* 2002; Binning *et al.* 2010), life history (Spight & Emlen 1976; Nylin & Gotthard 1998; Visser *et al.* 2009; Foster *et al.* 2015), and even behavior (Cotman & Berchtold 2002; Torres‐Dowdall *et al.* 2012), such as the avoidance of this new environmental cue (e.g. predator smell). Both theoretical as well as empirical studies suggested that phenotypically plastic species are also particularly successful in colonizing new environments, as plasticity can facilitate the immediate exploitation of many different niches (Yeh & Price 2004; Richards *et al.* 2006; Thibert-Plante & Hendry 2011).

In addition to increased species persistence and colonization success, phenotypic plasticity may also indirectly affect the standing genetic variation of populations (Nussey *et al.* 2005; Pfennig & McGee 2010; Renn & Schumer 2013). As selection acts on phenotypes, genotypes that plastically produce the same adaptive phenotype may shield their genetic differences from selection, which can preserve both genotypic variation as well as newly arising alleles in populations (which contributes to the populations cryptic genetic variability) (Pfennig *et al.* 2010; Draghi & Whitlock 2012; Schlichting & Wund 2014).

Plasticity's major advantages hold true as long as the future environmental conditions can be predicted sufficiently reliably from perceived cues. Limits and costs of plasticity may

thus arise due to unreliable environmental cues (e.g. Zimova *et al.* 2016) as well as other factors such as an inherent cost of maintaining the often complex regulatory network necessary for plastic responses, although studies increasingly suggest that the costs of plasticity are probably negligible in most cases (DeWitt *et al.* 1998; Ancel & Fontana 2000; Price *et al.* 2003; Binning *et al.* 2010; Snell‐Rood *et al.* 2010; Hendry 2015; Murren *et al.* 2015). Although a certain adaptive phenotype may be (plastically) produced by various genotypes, natural selection will still naturally favor the genotype with lowest associated costs, which in turn may reduce genetic diversity in a population (DeWitt *et al.* 1998; Pfennig *et al.* 2010).

As plastic responses can be induced by a variety of environmental stimuli, it is important to distinguish between 'adaptive' and 'nonadaptive' plasticity (Ghalambor *et al.* 2007). In contrast, when nonanticipated environmental stimuli induce a plastic response, the resulting phenotypes may be nonadaptive and therefore confer no increased fitness or even reduce fitness relative to the generic phenotype of the organism (Newman 1992; Ghalambor *et al.* 2007). As natural selection disfavours nonadaptive phenotypes, genotypes with nonadaptive plastic responses are expected to disappear from a population rapidly if the inducing environmental stimulus is common. Nonetheless, genotypes with nonadaptive phenotypic plasticity can presumably be retained in a population as long as their nonadaptive plastic response is not induced. It is thus important to distinguish between adaptive and nonadaptive plastic responses, as they can have different ecological and evolutionary consequences (Ghalambor *et al.* 2007; Ghalambor *et al.* 2015). Throughout this manuscript, we refer to the adaptive type of plasticity, if not mentioned otherwise.

Phenotypic plasticity's effect on evolution and speciation

As plasticity provides increased phenotypic variation, it is likely to also be a prime driver of evolutionary processes, such as phenotypic diversification. However, in the era of the 'modern synthesis', phenotypic plasticity was typically seen as being confounding for experiments, as the plastic response altered the phenotypes in a way that is not directly determined by the genotypes (West-Eberhard 1989; Nijhout 2015; Sultan 2015). In the light of this property, the notion arose that plasticity generally reduces the pace of evolutionary change as it partially uncouples the genotype from the phenotype and thus shields the genotype from the force of natural selection. Hence, it would move a trait effectively to an 'evolutionary slower lane' (Stebbins Jr 1950; Williams 1966; Meyer 1987b; Bell & Aubin-Horth 2010; Nijhout 2015).

Box 2. Cichlids as a model system for adaptive radiations & convergence

As one of the most species-rich families of vertebrates and a distribution across three continents, cichlid fishes represent a premier model to study adaptive radiation (Kocher 2004; Seehausen 2006; Muschick *et al.* 2012). In Lake Victoria alone, about 500 species have arisen within the last 200,000 years (Meyer *et al.* 1990; Wimberger 1994; Bell & Aubin-Horth 2010). The East African cichlids are particularly famous, as numerous ecological niches have been filled independently but in parallel across the lakes Victoria, Malawi and Tanganyika (Ohno 1970; Meyer *et al.* 1990; West-Eberhard 2003). Besides the adaptive radiations in these three lakes, there are numerous other examples of rapid evolution and adaptive radiations in cichlids in African cichlids, such as that of the Haplochromines in Lake Kyoga and Tilapiines in Lake Barombi Mbo (Trewavas *et al.* 1972; Muschick *et al.* 2012). Furthermore, it was demonstrated that the *Geophaginae* subfamily of the Neotropical cichlids exhibits all signs of an adaptive radiation, as many species emerged within a very short evolutionary time span and the radiating lineages are phenotypically specialists, occupying narrow ecological niches (Lopez-Fernández *et al.* 2005; Parsons *et al.* 2016). In particular, the genus *Crenicichla* was proposed to present an independent adaptive radiation within the *Geophaginae*, as they show astonishingly high levels of phenotypic diversification and niche specialization (Piálek *et al.* 2012). Another example is the Central American *Amphilophus spp.* complex. These fish inhabit the two geographically old great Nicaraguan lakes, Lake Nicaragua and Lake Managua, but also colonized an assemblage of very young crater lakes (Meyer 1989; Elmer *et al.* 2010). Within only a few thousand years, considerable phenotypic divergence has occurred between source populations (inhabiting the two great lakes) and the different satellite populations. Also, speciation has occurred within these satellite lineages, presumably in sympatry (Barluenga *et al.* 2006; Barluenga & Meyer 2010).

Convergence in cichlid fishes

By investigating morphological phenotypes of ecologically relevant traits across the phylogeny of cichlids, it has been repeatedly noted that adaptive phenotypes recur across cichlid radiations, most famously in the three East African Great Lakes (e.g. Kocher *et al.* 1993; Stiassny & Meyer 1999; Rüber & Adams 2001). Fish species from different lakes that occupy similar niches often have extremely similar morphological phenotypes. This is particularly noteworthy because these phenotypically similar species do not share a direct common ancestor in whom this morphology was already present, but originate from a single ancestral lineage per lake that also gave rise to all/many other species of the radiation and thus a plethora of phenotypes (Kocher *et al.* 1993; Salzburger *et al.*

2005). The coinciding assemblages of phenotypes in several radiations (and thus lakes) were thought to be the result of similar selection regimes in the respective similar ecological niches, leading to the independent evolution of very similar phenotypes, but with independent genetic bases – a textbook example for 'convergent evolution' (Stiassny & Meyer 1999; Wainwright *et al.* 2012). Cichlid convergent traits include thick lips, mouth brooding, body coloration and colour patterns, dentition and shape of the oral and pharyngeal jaws, gut lengths, body shape and many more (Muschick *et al.* 2012). The molecular underpinnings of the phenotypic convergence in these traits are mostly yet to be explored.

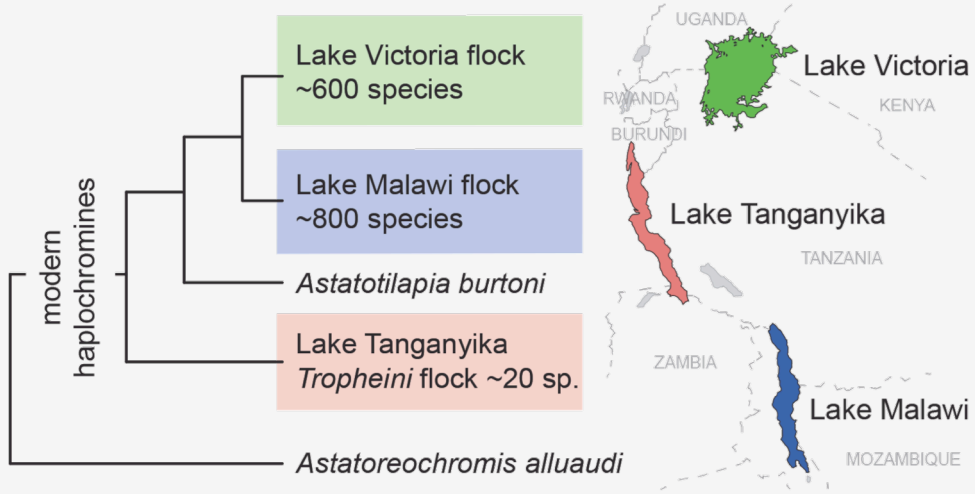


Fig. II.Box 2 Simplified cladogram of the 'modern Haplochromine' cichlid fishes. Species diversity mostly comprises the Lake Victoria (green), Lake Malawi (blue) and Lake Tanganyika Tropheini radiations (red) with age estimations approximately around ~0.1 Myrs, ~1.5 Myrs and ~2.5 Myrs, respectively (Salzburger *et al.* 2005). All three radiations experienced explosive phenotypic diversification after colonization. Additionally, the mostly riverine generalists *Astatotilapia burtoni* and *Astatoreochromis alluaudi* are included. Cladogram following Brawand *et al.* (2014).

Contrasting this view, plasticity in key ecological traits has often been found in taxa that are well known for their high rates of evolution and phenotypic diversification (Pfennig & McGee 2010). Phenotypic plasticity has been described in ecologically relevant traits such as leg length in Caribbean Anolis lizards (Losos *et al.* 2000), horn length in horned beetles (Emlen 1994; Moczek & Kijimoto 2014), castes in Hymenopterans (Cahan *et al.* 2004), body colour in spiny-legged spiders (Brewer *et al.* 2015), feeding ecology and life history decisions in nematodes (Gutteling *et al.* 2007; Serobyan *et al.* 2013), body shape and jaw size in Oceanic sticklebacks (Wund *et al.* 2008; Foster *et al.* 2015; Mazzarella *et al.* 2015) and pharyngeal jaw shape and dentition in cichlids (Box 3: Plasticity in cichlids) (e.g.

Greenwood 1964; Meyer 1993b; Muschick *et al.* 2012). These findings support the notion that phenotypic plasticity in ecologically relevant traits rather promotes evolution and phenotypic diversification, which eventually may even facilitate adaptive radiations (West-Eberhard 2003; Wund *et al.* 2008; Pfennig *et al.* 2010). Famously, early in the history of evolutionary biology, it was hypothesized that originally plastic traits can become genetically fixed in a homogeneous environment, where the advantage of being plastic is lost (Waddington 1942, 1953; West-Eberhard 2003). As whole sets of originally plastically induced phenotypes may be fixed, explosive phenotypic diversification may be explicable in this way (West-Eberhard 2003).

The first empirical support for a plastic trait becoming heritable was shown in the 1950s by Waddington (1953). He demonstrated that his results were most likely not due to selection on the genotype and proposed a new mechanism that he called 'genetic assimilation'. During the process of genetic assimilation, an originally environmentally induced phenotype becomes secondarily fixed and thus consistently expressed - that is, a trait loses its plasticity (Waddington 1942, 1953, 1961). It was proposed that due to the inherent cost of plasticity, variants with a reduced level of phenotypic plasticity that still includes the selected phenotype may have a fitness advantage and would therefore be selectively maintained in a homogeneous environment (DeWitt *et al.* 1998; Ancel & Fontana 2000; DeWitt & Scheiner 2004). Since then, however, only few studies have convincingly shown that genetic assimilation occurred (e.g. Sikkink *et al.* 2014) or that plasticity is generally connected to inherent maintenance costs. We are far from a complete understanding of the molecular mechanisms that might account for genetic assimilation (Ehrenreich & Pfennig 2015).

Ancestral phenotypic plasticity may promote explosive speciation events

It has been proposed for many adaptive radiations that the ancestral lineage might have been particularly phenotypically plastic in key evolutionary traits, that is adaptive traits that are especially diverse in derived lineages and probably contributed to diversification (Losos *et al.* 2000; Nylin & Wahlberg 2008; Wund *et al.* 2008; Tebbich *et al.* 2010; Muschick *et al.* 2012; Brewer *et al.* 2015; Susoy *et al.* 2015). This so-called flexible stem hypothesis is a key concept proposing that ancestral phenotypic plasticity can promote speciation and may even have contributed to adaptive radiations (West-Eberhard 2003). In the following paragraphs, we will describe why plastic ancestors are likely to facilitate the evolution of adaptive radiations and provide a model of how speciation patterns might be expected to proceed for adaptive radiations derived from an ancestral flexible stem lineage (Fig. II.1).

Ecological opportunity and ecotype formation

A novel environment colonized by a population offers ecological opportunity to this population, a driver of the adaptive diversification of lineages (Schluter 2000; Wellborn & Langerhans 2015). The magnitude of ecological opportunity is thought to be reflected by the number of available and utilizable niches (niche availability) and the degree to which they differ from each other in their selective regimes to facilitate diversifying selection (niche discordance) (Wellborn & Langerhans 2015). When comparing the ecological opportunity of a plastic and a nonplastic population that are about to colonize a new environment, the plastic population is predicted to be ready to utilize a broader range of niches if these overlap with their phenotypic plastic range (i.e. width of reaction norm). Both niche availability and niche diversity (and hence the chance for niche discordance) are thus expected to be higher for a plastic population compared to its nonplastic counterpart. Nonplastic lineages can only persist in the new habitat in niches similar or equal to its previous one (Pfennig *et al.* 2010), and the potential for adaptive diversification is therefore expected to be relatively low compared to a plastic population.

Lineage divergence and adaptive radiation – a flexible stem

For speciation and adaptive radiations to occur, the colonizer has to undergo (potentially multiple times) lineage divergence. After colonizing a new habitat, a plastic population may form subpopulations, each occupying one available niche and expressing the niche-specific most suitable (i.e. adaptive) plastic phenotype. However, whether a lineage is likely to undergo an adaptive radiation does not only depend on ecological opportunity, but also on the diversification potential of the founding population, which describes the ability of a population to respond to ecological opportunity (Wellborn & Langerhans 2015). According toWellborn and Langerhans (2015), the diversification potential of a population is the product of population properties that facilitate species persistence, properties that increase phenotypic variance when encountered ecological opportunity and properties facilitating reproductive isolation among subpopulations. As phenotypic plasticity promotes species persistence and leads to rapid phenotypic differentiation in different niches (if they correspond to the plastic range), two of three properties confer high diversification potential on a plastic population. The third property is the potential to establish reproductive isolation. Here, it is the plasticity itself that drives spatial segregation of subpopulations and therefore promotes reproductive isolation. Subsequently, further processes such as assortative mating (Doebeli & Dieckmann 2003), differences in reproductive timing or very high territoriality, may strengthen reproductive barriers (Dieckmann & Doebeli 1999). A plastic population is thus likely to have a high potential for

diversification and therefore is more likely to initiate an adaptive radiation in (macroscopic) sympatry.

Indeed, a considerable body of literature supports the notion that many adaptive radiations may have begun in sympatry (e.g. Kawecki 1997; Dieckmann & Doebeli 1999; Kondrashov & Kondrashov 1999; Doebeli & Dieckmann 2003; Levis & Pfennig 2016). Through the process of genetic assimilation, emerging lineages may then come to fix the most adaptive phenotype for their niche. Arising phenotypes of radiating lineages are thus expected to initially reflect different phenotypes of the ancestral range of plasticity (Levis & Pfennig 2016). Depending on the number of emerging lineages and the overall morphospace they occupy, a large number of new species can potentially arise within a short evolutionary time span. Indeed, the reduction of plasticity in a lineage has often been correlated with speciation events (Schwander & Leimar 2011), potentially reflecting plasticity-mediated diversification events and subsequent genetic assimilation of phenotypes. In contrast, many nonadaptive radiations arose presumably due to allopatric divergence of the ancestral lineage without a major initial niche shift in descending lineages (Nei *et al.* 1983; Gittenberger 1991; Rundell & Price 2009). Thus, the contribution of plasticity to diversification is expected to be reduced in nonadaptive compared to adaptive radiations, as a sympatric adaptive radiation is thought to depend on the availability of numerous exploitable niches in the same habitat. In addition, a plastic lineage may have increased standing genetic diversity (as discussed above), which can serve as the substrate for subsequent adaptations (Barrett & Schluter 2008).

Repeated colonization of similar habitats – convergence and parallel evolution

When a phenotypically plastic stem lineage (the 'source' population) persists over longer geographic time, it may encounter the opportunity to colonize further similar habitats. Such parallel colonizations may lead to repeated fixation of similar phenotypes if colonized habitats offer similar niches for which a specific, plastically induced adaptive phenotype exists. Such scenarios were proposed for oceanic stickleback populations that colonized freshwater lakes, whitefish in postglacial lakes and for Darwin's finches on the Galapagos Islands (Robinson & Parsons 2002; Wund *et al.* 2008; Tebbich *et al.* 2010; Lundsgaard‐Hansen *et al.* 2013; Oke *et al.* 2015). For example, marine sticklebacks are phenotypically plastic and can express phenotypes along a benthic (deeper body with larger eyes) to limnetic (elongated body with smaller eyes) phenotype gradient. While colonizing freshwater habitats, these sticklebacks were presumably confronted repeatedly with two different freshwater lake niches, a limnetic one in shallow lakes and a benthic one in deeper lakes, each with a specific selective regime (Schluter 1993; Wund *et al.* 2008).

Box 3. Plasticity in cichlids is widespread

Within the family of cichlid fishes, Greenwood was among the first to scientifically describe phenotypic plasticity in an adaptive trait (Greenwood 1964). He noticed that the morphology of the pharyngeal jaw bones in the generalist species Astatoreochromis alluaudi, which lives in Lake Victoria and the surrounding river system, varies in shape, robustness and dentition depending on population. The main difference between the populations turned out to be the stomach content composition of the fish, having populations with the highest proportion of food being insect larvae (soft food items) and some with a highest proportion being snails (a hard food item, as shells have to be crushed during ingestion). He conducted feeding experiments and found that the robustness of the food items individuals ingest determines the phenotypes of their pharyngeal jaws, which was confirmed in later studies (Greenwood 1964; Stauffer & van Snick Gray 2004).

Since then, phenotypic plasticity in cichlids has been reported in a number of traits. Importantly, plastic species were often reported to be riverine or being part of very recent intralacustrine adaptive radiations (Greenwood 1964; Peterson & Davidson 2000; Prud'homme *et al.* 2007; McLean *et al.* 2011; Peter & Davidson 2011).

Patterns of phenotypic plasticity in the pharyngeal jaws of modern Haplochromines

Astatoreochromis alluaudi belongs to a nonradiated lineage that is directly basal to the 'modern Haplochromines' (Salzburger *et al.* 2005) (Fig. II.Box 2). These comprise three adaptive radiations, two of them outstanding as they led to several hundred species within incredible short evolutionary time spans (Meyer 1993a; Salzburger *et al.* 2005). As mentioned before, the hyperdiverse phenotypes found across species in each of these radiations are famous and convergence was described in a suite of traits (O'Quin *et al.* 2010; Schneider *et al.* 2014). Previous studies showed that besides in *A. alluaudi*, pharyngeal jaw phenotypic plasticity can also be found in *Haplochromis iris* and *Haplochromis ishmaeli*, young, radiating species from the Lake Victoria flock that specialized on insects and snail cracking, respectively, and *Astatotilapia burtoni* (Smits *et al.* 1996). The latter is a generalist that inhabits Lake Tanganyika and associated river systems and is phylogenetically situated either between the *Tropheini* radiation and the Lake Malawi radiation (Brawand *et al.* 2014), or between the latter and the Lake Victoria radiation (Fig. II.Box 2) (Salzburger *et al.* 2005). Data recently collected in our laboratory suggest that the levels of inducible adaptive and nonadaptive plasticity vary considerably between representatives from the radiations and the rivers (H. M. Gunter, R. F. Schneider, I. Karner, C. Sturmbauer & A. Meyer, manuscript in prep.).

The morphologies of a limnetic and a benthic stickleback population that were collected in two Alaskan lakes were found to qualitatively reflect the plastically inducible limnetic and benthic phenotypes of a marine population that approximates the common marine ancestor of the freshwater populations (Wund *et al.* 2008). Repeated induction of ancestral plastic phenotypes and their subsequent (partial) fixation may thus provide an explanation for some instances of the phenomenon of recurring phenotypes among independent colonizations of similar habitats (Oke *et al.* 2015). In the following sections, we will explore possible molecular mechanisms, and their traces, that can potentially mediate the process of genetic assimilation and how it may overcome the uncoupling of genotype and phenotype that characterizes a plastic trait.

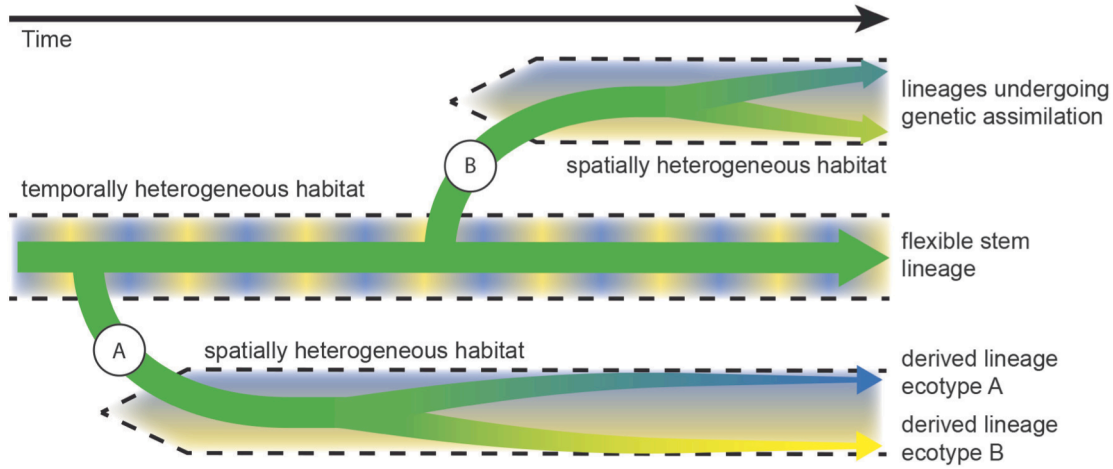


Fig. II.1 Illustration of the flexible stem model. A phenotypically plastic lineage (flexible stem) inhabiting a temporally heterogeneous habitat (e.g. heterogeneity in abundance of alternative food types in a river) colonizes a spatially heterogeneous habitat in which the two food sources are spatially separated (e.g. lake habitat with one food-type associated to the shore and the other to the open water) (A). Here, subpopulations displaying the alternative adaptive plastic phenotypes, and population divergence can occur accordingly. Gradually, selection fixes the expressed phenotype in each population via genetic assimilation, effectively reducing the level of plasticity in each emerging lineage. Later in time, the same flexible stem lineage might colonize another similar habitat, offering the same niches (e.g. another lake) (B). Again, lineage divergence can occur according to the two niches, leading to a phenotypically very similar species assemblage as in the other habitat (A) due to their common ancestor's phenotypic range. Convergent/parallel phenotypes between derived lineages of separate colonizations are thus initially expected to be very common.

Gene regulatory networks underlying plasticity may provide a prime substrate for phenotypic diversification

Environmentally induced plastic phenotypes are thought to be controlled by gene regulatory networks (GRNs) (Schlichting & Pigliucci 1993) that often remain poorly understood, particularly in non-model organisms (Pfennig & Ehrenreich 2014). GRNs appear to often have a modular structure, that is sets of genes that are more or less jointly up- or downregulated, for example during the plastic response (Shiga *et al.* 2007; Wagner *et al.* 2007; Aubin-Horth *et al.* 2009; Aubin-Horth & Renn 2009; Hinman *et al.* 2009; Schneider *et al.* 2014; Gilbert & Epel 2015). At some point during their evolution, GRNs underlying plastic traits have acquired responsiveness to external stimuli that trigger new regulatory cascades leading to new adaptive phenotypes (via alternative developmental trajectories). A particularly well-understood example for a gain of plasticity can be found in the nematode *Pristionchus pacificus*, in which a recent gene duplication of a sulfatase (*eud-1*) was identified to be the main switch between the developmental trajectories underlying two distinct plastic feeding morphologies (Ragsdale *et al.* 2013). Expression of the eud-1 gene is regulated via histone methylation and expression of an antisense RNA (Serobyan *et al.* 2016). These regulatory mechanisms are presumably connected to a hormone and its receptor, and hormone levels are directly influenced by pheromone concentrations in the environment (Bento *et al.* 2010).

Developmental switches that respond directly or indirectly to environmental stimuli allow a single genotype to evolve multiple developmental trajectories in parallel, leading to alternative phenotypes. GRNs of plastic traits are thus outstanding in that they not only allow different genotypes to produce the same phenotype (and thus contribute to developmental robustness), but also to comprise multiple developmental trajectories leading to multiple (plastic) phenotypes within a single genotype. Therefore, they provide a twofold opportunity for the evolution of phenotypic diversity during lineage divergence and specialization (e.g. during an adaptive radiation): (i) an ancestral GRN of a plastic trait can allow diverging lineages to immediately exhibit unique adaptive phenotypes fitting their respective niches by utilizing different developmental trajectories. Immediate phenotypic diversity among emerging lineages would thus reflect (partially) the plastic range of the ancestor (as predicted by the flexible stem hypothesis). (ii) Utilizing only one developmental trajectory (or a subset) in a specialized derived lineage renders all other trajectories of that genotype unused. However, with time, these 'dispensable' trajectories may provide an additional substrate for longer-term phenotypic change (as they are a source of cryptic genetic variation).

The following sections illustrate (i.) how selection can potentially genetically fix a plastic phenotype via genetic assimilation in a constant environment. (ii.) We then summarize predicted patterns on genetic, transcriptional and phenotypic variability that are expected to arise from diversification based on a plastic trait. (iii.) Finally, we speculate on how phenotypes can expand phenotypically from the ancestral phenotypic range in derived lineages, potentially more rapidly than classically assumed.

Step-wise degeneration of noninduced developmental trajectories – use it or lose it

In constant environments, plastic traits may undergo genetic assimilation (Fig. II.2): if an environmentally induced phenotype is constantly expressed (across generations), only the specific developmental trajectory underlying the specific phenotype is in usage and thus can be the target of stabilizing (or directional) selection (Price *et al.* 2003; Lahti *et al.* 2009). At this stage, mutations within the expressed developmental trajectory can only be selected for if these mutations lead to phenotypes beyond the ancestral plastic range and if these new phenotypes have higher fitness values. However, if the fitness optimum lies within the ancestral plastic range, plasticity is predicted to shield otherwise beneficial mutations from selection. In this simplified example, with no inherent cost of plasticity, natural selection can therefore not directly contribute to a trait's genetic fixation as long as the optimal phenotype is within the ancestral range. To explain how genetic assimilation can occur under such conditions, the remaining parts of the GRN have to be considered. The developmental trajectories underlying non-induced phenotypes are not expressed and thus released from stabilizing selection ('relaxed selection'), a prerequisite for cryptic genetic variation to be maintained (Rutherford & Lindquist 1998; Lahti *et al.* 2009). Thus, these unused developmental trajectories of the GRN (coding and regulatory sequence) are predicted to be prone to accumulate random mutations and thus degenerate (Fig. II.2D,F, H) (Masel *et al.* 2007). In addition, regulatory mutations consolidating the expression of the adaptive phenotype expressed in the given constant environment are not selected against, as would be the case in a heterogeneous environment, and can thus persist (Fig. II.2C,E,G). Therefore, the more extensive a GRN underlying a plastic trait is, the more cryptic genetic variation can accumulate when the trait undergoes genetic assimilation that is predicted to increase the population's evolvability (Iwasaki *et al.* 2013). This is in line with theoretical work generally indicating that developmental robustness increases the evolvability of populations (Wagner 2008; Hayden *et al.* 2011).

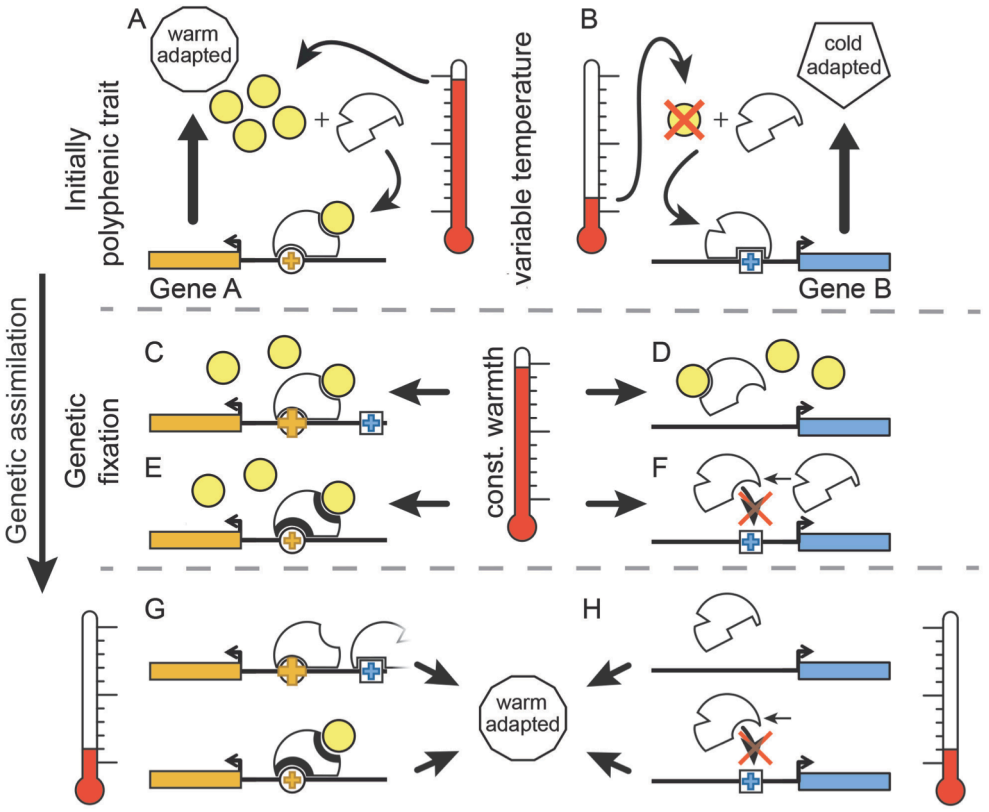


Fig. II.2 Simplified examples of how genetic assimilation may genetically fix a previously environmentally responsive trait. A plastic trait's expression is determined by gene products A and B, while A is inducing a phenotype adapted to warmth and B inducing the cold adapted one. A and B transcription is induced by the same transcription factor 'TF'. It binds to its binding site in the promoters of gene A, if a cofactor (yellow circle) is present 'TF-A' (A). If no cofactor is present, 'TF-B' will bind to another binding site in the promoter of gene B (B). Temperature (environmental stimulus) positively regulates the cofactor's abundance. Across time, constant warmth may lead to genetic assimilation via degeneration of the developmental trajectory of the cold adapted phenotype and consolidation of the warm adapted phenotype. Random mutations may, for example (C), increase binding affinity of TF-A to the gene A promoter via TF site modification or allow TF-B binding by the appearance of TF-B binding sites or (D) lead to loss of TF-B binding sites in the gene B promoter. Furthermore, random TF modifications may (E) increase binding affinity of the cofactor to the TF or the TF-A to its binding site or (F) lead to changes in the binding motives of TF-B, for example by allowing it to bind to the gene A promoter even without co-factor. After such events, even in low temperature environments, the warm adapted phenotype is produced due to consolidation of its developmental trajectory (G) and degeneration of the cold adapted phenotype's developmental trajectory. As the trait becomes genetically fixed, natural selection may increasingly act on the gene sequence of gene A.

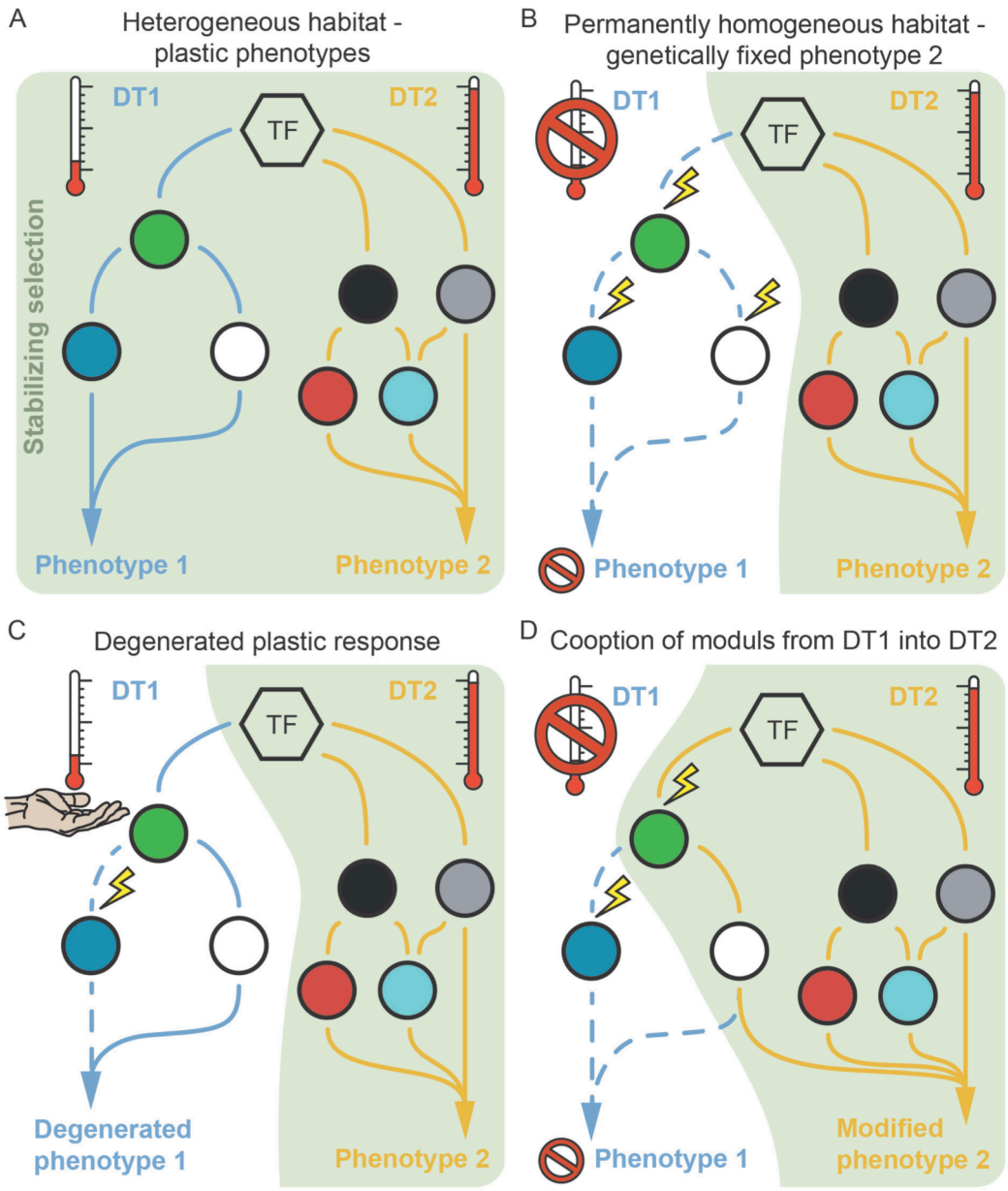


Fig. II.3 Illustration of how genetic assimilation may fix and alter a trait by the degeneration and reutilization of developmental trajectories. In a heterogeneous environment, the environmental cue varies (here: temperature) (A). In a phenotypically plastic population, the environmentally controlled transcription factor 'TF' controls a trait's phenotype: low temperatures induce expression of developmental trajectory 1 (DT1, blue lines, left side) and high temperatures induce developmental trajectory 2 (DT2, orange lines, right side) which lead to the adaptive phenotypes 1 and 2, respectively. The environmentally responsive portion of the trait's underlying regulatory network consists of different genes (or gene modules, illustrated as circles) in each DT. As both DTs are regularly expressed and adaptive, both are under stabilizing selection. After the environment changed to permanent warmth, only DT2 was induced and remained under stabilizing selection (B). In contrast, DT1 fell prone to random

mutations, which successively degenerated the circuitry of the involved gene regulatory network (GRN). When a population undergoing genetic assimilation by DT degeneration is confronted with the original plasticity cue (here: coldness), the induced phenotype 1 may be already degenerated and less adaptive, as parts of its underlying DT are not functional anymore (C). Parts of a degenerating DT may be reutilized by other DTs of the same or other traits, potentially facilitating the evolution of phenotypes beyond the ancestral range (D).

Generally, degeneration of regulatory circuitry may affect coding and noncoding sequences differently: as many genes are pleiotropic, coding regions presumably often remain under stabilizing selection due to their contribution to other traits (Jarvela & Hinman 2015). In contrast, regulatory regions are expected to be often less pleiotropic (Wray 2007), particularly if they contribute only to one trait's phenotype. Thus, if an inducible phenotype is not expressed, its specific regulatory sites are not under stabilizing selection and can more easily degenerate through the accumulation of neutral mutations (in the given homogeneous environment). Even single point mutations in regulatory sites of genes can have switch effects on the genes' sensitivity towards individual transcription factors, particularly in character loss (Van Laere *et al.* 2003; Mayo *et al.* 2006; Hoekstra & Coyne 2007). Alterations in cis- and trans-regulatory sites may have different effects, as cis-sites are typically thought to regulate the expression of one gene, while trans-sites can often regulate many genes (as discussed inEhrenreich and Pfennig (2015)). Considering the modular structure of many GRNs, a change in expression level of a single key transcription factor induced by mutations in regulatory sites may have tremendous effects, as whole gene modules can be affected (Fig. II.3). Across generations, even few of these mutations therefore have the potential to degenerate the noninduced phenotypes' trajectories or its 'inducibility' by environmental stimuli (Fig. II.3B,C). Populations undergoing genetic assimilation for a certain phenotype may therefore be particularly dependent on homogeneous environmental stimuli, as alternative stimuli may induce malfunctional developmental trajectories potentially leading to mal adaptive phenotypes. In this way, genetic assimilation may contribute to population reproductive isolation.

Although genetic assimilation may lead to rapid phenotypic divergence, genetic fixation by initially neutral mutations may not be a very fast process, as it relies on random mutations. However, faster genetic assimilation may occur if there are maintenance costs of plasticity: phenotypically plastic traits may require costly and elaborate regulatory networks that respond to variable environmental stimuli by inducing a fitter phenotype (DeWitt *et al.* 1998). If environmental conditions are rather homogeneous, maintaining such a regulatory network offers no fitness advantage and natural selection is thus expected to

select for genotypes with reduced 'maintenance costs', that is reduced plasticity (Pigliucci *et al.* 2006). Costs can thus promote genetic fixation in homogeneous environments. In addition, trans-generational epigenetic mechanisms may be another mechanism facilitating genetic assimilation: environmentally induced phenotypes in a parental generation may be epigenetically inherited by offspring generations, even in the absence of the inducing environmental cue (Whitelaw & Whitelaw 2006; Aubin-Horth & Renn 2009; Mirouze & Paszkowski 2011; Schlichting & Wund 2014; Beaty *et al.* 2016). Thus, it can be speculated that epigenetic inheritance of plastically acquired phenotypes may buffer against short-term environmental heterogeneity that would elsewise interfere with ongoing genetic assimilation (Jablonka & Lamb 2007).

Ancestral plasticity's traces in genetic, transcriptional and phenotypic patterns during lineage diversification

A promoting effect of ancestral phenotypic plasticity to lineage diversification seems plausible, for example in the context of colonization events, and it may indeed facilitate adaptive radiations (Levis & Pfennig 2016). Present adaptive radiations (and particularly repeated adaptive radiations) may offer the opportunity to test for the effects of ancestral plasticity, as the traces of (ongoing) genetic assimilation are predictable and should be observable in today's radiations.

The process of genetic assimilation and the accompanying degeneration of alternative phenotypes is a form of regulatory evolution. Generally, regulatory evolution is thought to be a main driver of phenotypic evolution, essentially by the rewiring of existing GRNs ('developmental recombination' sensu West-Eberhard 2005) (e.g. Britten & Davidson 1971; King & Wilson 1975; West-Eberhard 2005; Davidson & Erwin 2006; Prud'homme *et al.* 2007; Wray 2007; Carroll 2008). Regulatory evolution may include sequence evolution of cis-regulatory elements in the genome (such as enhancers, promoters, splice-site modules and microRNAs) but also coding sequence evolution that can affect the binding properties of transcription factors as well as changes of epigenetic patterning and processing (e.g. alterations in DNA or histone methylation or acetylation). These mechanisms can work synergistically to shape new developmental trajectories, as exemplified in the feeding polyphenism of *Pristionchus pacificus* mentioned before (Ragsdale *et al.* 2013; Serobyan *et al.* 2016). DNA sequence mutations can alter interactions between regulatory pathways and thus lead to phenotypic change (Alonso & Wilkins 2005). Thus, rewiring of existing developmental pathways, such as those underlying plastic responses, can potentially contribute to fast lineage divergence in phenotypes induced by only minor underlying genetic divergence (West-Eberhard 2005). Based on gene reaction norms, Renn and Schumer (2013) provided a theoretical framework that assists in classifying evolved gene

expression patterns found after genetic accommodation in plastic (behavioral) traits. Complementing this, we focus in the following sections on patterns of phenotypical and transcriptional variation predicted to arise between and within populations undergoing genetic assimilation.

Investigating the contribution of regulatory evolution to the diversification of target lineages can be challenging, as one has to rely on indirect evidence (Levis & Pfennig 2016). West-Eberhard (2005) described peculiar patterns in genes and their expression as well as morphology and behaviour that are expected to emerge across diverging lineages if regulatory evolution is involved in driving lineage divergence. Summarized, these are that (i) across diverging lineages, homologous genes are expected to be conserved over relatively distantly related taxa (although they might be used in different contexts), (ii) in emerging lineages, phenotypic divergence may be relatively large while genetic divergence remains subtle, (iii) among closely related species, homoplasy and parallelism are common, and (iv) phenotypic differences in key traits across emerging species can arise before reproductive isolation.

In addition to these four predictions that presumably apply to all cases in which regulatory evolution drives diversification, further predictions can be made if the underlying GRN is that of an initially plastic diversifying trait and if diversification occurs recurrently. Then, observable convergent/parallel phenotypes are expected to reflect the plastic range of the ancestor, as predicted by the flexible stem hypothesis (West-Eberhard 2003). Diverging lineages that emerge from a plastic ancestral population and occupy different niches thus may not only show marked phenotypic divergence before genetic divergence and reproductive isolation, but also pronounced transcriptional divergence as a result of alternative developmental trajectory usage (Fig. II.4A–C and II.1). Under certain conditions, a flexible stem lineage may have the opportunity to independently give rise to multiple lineages that undergo genetic accommodation for essentially the same ecological niche. Although such events may be rare in sympatry, it may be a common occurrence if the lineages live in allopatry. This may happen if a flexible stem lineage colonizes multiple new habitats with similar ecological niches, as was suggested for marine sticklebacks colonizing freshwater lakes and riverine cichlids colonizing lakes (Fig. II.1) (West-Eberhard 2003; Wund *et al.* 2008).

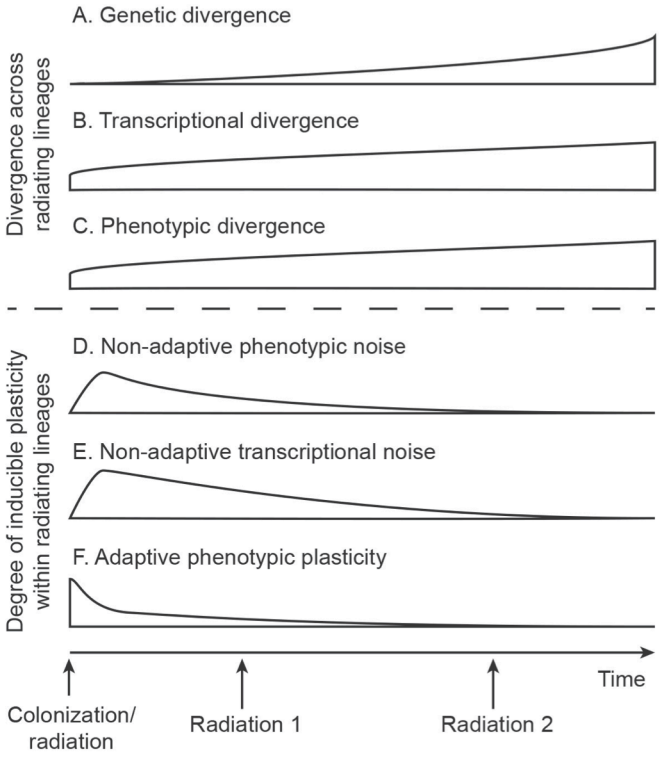


Fig. II.4 Patterns of phenotypic and transcriptional variability across diverging lineages expected due to genetic assimilation after a colonization event and during an adaptive radiation with a plastic ancestor. Genetic divergence arises gradually and relatively slowly as it relies on random mutations (A). Different environmental stimuli immediately lead to pronounced differences in transcription across lineages that underlie the respective distinct phenotypes (B,C). Genetic assimilation degenerates unused developmental trajectories gradually (D,E). Early in this process, some unused trajectories are still responding to the alternative stimuli, however, already low levels of degeneration may lead to nonadaptive transcriptional noise. Accordingly, the resulting phenotypes might only be partially functional and thus result in nonadaptive phenotype characteristics (phenotypic noise). The functionality of the regulatory network underlying the adaptive plastic response is strongly decreasing with the increasing degeneration of alternative developmental trajectories by random mutations in regulatory regions (F). Comparisons of such patterns in repeated radiations of different ages can allow conclusions about the existence of a common flexible stem. Thus, by investigating patterns of inducible plasticity in a trait, it is possible to estimate at which stage in genetic assimilation a radiation is, as indicated in the figure (Radiation 1 vs. Radiation 2).

As these lineages undergo genetic assimilation for the essentially the same phenotype, they are likely to utilize the same ancestral developmental trajectory as well. Thus, we expect that the variation in the transcriptional profiles of genes specific (or with specific expression patterns) to the focal trait is low among the lineages with parallel phenotypes compared to variation found across genes not specific to the focal trait's phenotype. Accordingly, expression profiles of genes specific to a focal trait's parallel phenotypes

originating in a common flexible stem are predicted to cluster together, when analysed comparatively with genes that do not contribute to that trait. Or, from another perspective: if transcriptional profiles of genes contributing to such a trait's phenotypes are analysed across radiating lineages (with differing phenotypes), a clustering according to phenotypes and not phylogeny is the expected outcome. Thus, we conclude that comparative analyses of gene expression profiles underlying – or not underlying – a focal trait's phenotypes can be a powerful approach to unveil regulatory relationships due to a common flexible stem.

One way to characterize the different developmental trajectories underlying the inducible phenotypes of a plastic trait are their gene expression profiles and their differences (as reviewed e.g. in Gibson 2008). During genetic assimilation, selection successively fixes a specific developmental trajectory by reinforcing its expression or degenerating the developmental trajectories underlying alternative non-induced phenotypes. However, as this process is gradual, we assume that the GRN retains some inducibility by alternative environmental stimuli until the phenotype is entirely fixed. If an individual belonging to a population undergoing genetic assimilation for a certain adaptive phenotype 'A' is confronted with an environmental stimulus (e.g. due to environmental perturbations) that was only experienced by their plastic ancestor and led here to phenotype 'B', the derived individual's GRN may still partially respond to that stimulus. We would then expect that the resulting phenotype may be considerably less adaptive than both phenotype A and B, as the plastic response may either be a degenerated version of phenotype B or a mixture of both phenotypes. Such non-adaptive phenotypic plasticity may be especially striking on the transcriptional level, when ancestral (adaptive responsive) and descendent lineages (less/ non-adaptive responsive) are compared under the influence of variable environmental stimuli (we summarized our predictions in Fig. II.4D–F). We thus encourage studies using comparative transcriptomics in candidate radiations that determine the level of adaptive and non-adaptive plasticity from which the stage of ongoing genetic assimilation could be inferred.

Patterns of non-adaptive plasticity around colonization or invasion events have been reported frequently in previous studies, for example in a study recently published on guppies (*Poecilia reticulata*) (Ghalambor *et al.* 2015). The authors found that a pronounced non-adaptive plastic response can be induced in a putative ancestral guppy lineage (living in a high-predation environment) by an atypical stimulus (low-predation). Furthermore, the authors found that artificially formed subpopulation from this ancestral population could rapidly genetically adapt to a low-predation environment within few generations. The authors conclude that non-adaptive plasticity (rather than adaptive plasticity) may facilitate evolution. Considering that the ancestral source population lives in a high-predation environment, it may also be plausible that its high-predation phenotype is undergoing

genetic assimilation. Thus, by confronting it with the low-predation stimulus, a degenerated low-predation phenotype that comes with the observed pronounced transcriptional noise may have been induced (compare to Fig. II.4E 'Radiation 10). Therefore, the possibility exists that the ancestral guppy population represents an already quite derived population (i.e. at an advanced stage of becoming genetically assimilated for a high-predation phenotype). Non-adaptive plasticity might thus reflect an unveiling of cryptic genetic variation by the atypical low-predation environment. Furthermore, the cryptic genetic variation has potentially originated in alternative developmental trajectories of a formerly plastic trait. Patterns of non-adaptive plastic responses in phenotypes and gene transcription can hence inform about the state of ongoing genetic assimilation. Recent investigations in cichlid fishes found evidence that predicted patterns (Fig. II.4) may also be observable across African cichlid radiations (Box 4). Further examples for non-adaptive plastic responses can be found in plants: in a meta-analysis, Davidson *et al.* (2011) found that invasive plant species show increased levels of plasticity when compared to noninvasive species, supporting the notion that plastic species are good colonizers. However, the plastic responses of these invaders were not necessarily adaptive. The non-adaptive responses may therefore simply reflect a degenerated plastic response when an ancestrally plastic trait undergoes genetic assimilation in a new environment, as predicted by our model.

Going beyond the ancestral phenotypic range

Thus far, this manuscript has presented and discussed studies suggesting that phenotypic plasticity can be a driver of phenotypic differentiation and lineage divergence. However, we have not yet discussed how phenotypes extend beyond ancestral plastic range (i.e. reaction norm). The modular nature of developmental trajectories (Schlosser & Wagner 2004) is likely to facilitate diversification. It was suggested that genes contributing to plastic responses are regulated in gene regulatory modules during the plastic response, maybe by relatively few master regulators (Aubin-Horth & Renn 2009; Schneider *et al.* 2014). Permanently un-induced developmental trajectories of a plastic trait that are undergoing successive degeneration can potentially be (partially) integrated into the expressed trajectory (and thus phenotype) of the trait undergoing genetic assimilation. Such cooption of regulatory circuitry may lead to phenotypes outside the range of the ancestor (Fig. II.3D). Thus, characteristics specific for originally separate phenotypes may be integrated into a single derived phenotype with altered fitness. The phenotype of such a trait would no longer reflect the phenotypic range of the plastic ancestral population.

An exciting candidate for an originally plastic network undergoing genetic assimilation and subsequent cooption of degenerating developmental trajectories in a

derived species can be found among the horned beetles of the *Onthophagus* genus. Horn size in horned beetles can be very variable, not only among species, but also between sexes and even within a sex. *O. taurus* exemplifies the general patterns found in the genus: males have considerably larger horns than females. However, the relative size of male horns critically depends on the food supply during an individual's larval stage: limited food supply leads to small males with disproportionally smaller horns, while food in excess induces larger males with relatively larger horns (Moczek & Emlen 1999). The horn phenotypes are controlled in these beetles by the highly conserved *doublesex* (*dsx*) gene pathway: alternative splice variants of *dsx* determine both the sex of an individual and the expression of secondary sexual traits, such as the horn phenotypes (Kijimoto *et al.* 2009; Kijimoto *et al.* 2012). Interestingly, a particular *Onthophagus* species, *O. sagittarius*, shows entirely altered horn development patterns compared to other species of the genus. Most notably in the context of this manuscript is that *O. sagittarius* males probably lost their ancestral plastic responsiveness towards food supply during larval stages, leaving all males with the relatively same horn size independent of body size (Kijimoto *et al.* 2012). Instead, they evolved two further, evolutionarily novel sets of horns: one anterior to the ancestral one that is only present in males, and one posterior to it on the thorax that is only present in females. Kijimoto *et al.* (2012) also found that the novel horn sets and phenotypes are due to novel functions of *dsx* splice variants and an alteration of *dsx* expression patterns. Horned beetles thus constitute a prime candidate to investigate 'whether' and 'how' genetic assimilation and subsequent cooption of degenerating developmental trajectories may lead to evolutionary novelty.

Diversification in plastic traits is in line with studies finding that modeled GRNs with a plastic history are more capable of adapting to new environments than those with a static history (Fierst 2011). In addition, fragments of the regulatory circuitry of degenerating developmental trajectories may start to contribute to other traits and, if a link to environmental sensitivity is maintained, facilitate the evolution of environmental sensitivity in that trait. Early in the process of such gene regulatory rewiring, involved genes and regulatory sites might experience strong selection that shapes the traits' modified phenotypes. It was previously described that relaxed selection on genes, followed by rapid evolution after being coopted in a new trait may be a mechanism of evolving phenotypic plasticity (Hunt *et al.* 2011; Leichty *et al.* 2012). This scenario thus constitutes another possible explanation for the frequent cooption of regulatory circuitry in traits that originated in completely different ones (e.g., Lee *et al.* 2003; Shubin *et al.* 2009; Peter & Davidson 2011). In spite of the potentially positive effects of cryptic genetic variation due to non-induced degenerating developmental trajectories on the rate of evolution, such cryptic genetic variation can also have maladaptive consequences if expression is induced, e.g. by

becoming a source of genetically linked diseases, as generally suggested for cryptic genetic variation (Paaby & Rockman 2014).

In conclusion, phenotypic plasticity offers considerable advantages in species persistence in many heterogeneous environments and plastic traits can evolve, particularly via regulatory evolution. However, in unpredictable or very stable environments, genetically fixed phenotypes were shown to outperform plasticity as natural selection can act directly on the genes underlying a phenotype (Le Vinh *et al.* 2016). In the natural world, however, environments are rarely heterogeneous but perfectly predictable, nor perfectly stable or completely unpredictable. Therefore, selection holds most traits at a trait- and environment-specific 'equilibrium state' of genetic accommodation, which is why phenotypes of most traits can only be predicted when both genetic and environmental effects (and their interactions) are considered: P = G + E + GxE. Some adaptive radiations may have been the result of a flexible stem diversifying after experiencing high ecological opportunity, while others may have been driven by strong diversifying selection on a genetically fixed phenotype and diversity arose, for example, from standing genetic variation. By investigating the genetic, transcriptional and phenotypic patterns observable in the lineages that emerged from explosive speciation events, comparative studies can investigate the relative importance of plasticity vs. genetic determination in driving this extraordinary diversification.

Box 4. Patterns of morphological and transcriptional plasticity in cichlids may coincide with predictions of a flexible stem in cichlids

A direct or indirect contribution of phenotypic plasticity to the adaptive radiations across cichlids has been suggested (e.g. Meyer 1987b) but few have provided specific hypotheses of the underlying processes and mechanisms (but see Wimberger 1994). Recently, it was hypothesized that similar phenotypes across radiations of the East African cichlids are not independently evolved, but rather share a common developmental trajectory that was presumably already present in their plastic common ancestor – an example of 'parallel evolution' (O'Quin *et al.* 2010). Insights into the molecular underpinnings of the plastic response in cichlid pharyngeal jaws as well as an increase in the knowledge of its phenotypic distributions across cichlids allow us to more specifically investigate how plasticity in key evolutionary traits has affected cichlid evolution (Muschick *et al.* 2012). The hypothesis that a phenotypically plastic ancestral trait might have provided the baseline variability reflected today by hundreds of diverse cichlid species across several radiations (West-Eberhard 2003) challenges our understanding of how trait diversity generally evolves.

Do predicted morphological and transcription patterns of plasticity fit observations in cichlids?

An adaptive radiation initiated by a plastic ancestor would lead to a sudden increase in phenotypic diversity, followed by a gradual increase in genetic diversity, as described before. This pattern coincides with observations in modern Haplochromines: even the youngest radiation of a Great Lake (Lake Victoria) displays a plethora of phenotypes, but genetically, the species are almost indistinguishable, with fragile reproductive barriers that can easily be circumvented when conditions are manipulated (e.g. in an aquarium) (Stelkens *et al.* 2009). Verheyen *et al.* (2003) noted that the phenotypic diversity found in the Lake Victoria species flock is likely to have emerged in the very early stages of the radiation – a phenomenon that was also suggested for the Neotropical *Geophagini* cichlids (Lopez-Fernández *et al.* 2005; López-Fernández *et al.* 2012). Such rapid phenotypic divergence is hardly explainable assuming classical gene-mutation-based gradual phenotypic change that is assumed to require much longer time spans (Miller 1949; Gavrilets 2003). Furthermore, our preliminary data on the level of inducible adaptive and non-adaptive phenotypic plasticity in the pharyngeal jaws of radiating East African cichlids suggest that, while two riverine species show the highest levels of adaptive plasticity, a radiating lineage from Lake Victoria and one from Lake Tanganyika show reduced and no adaptive plasticity anymore, respectively. Total morphological variation, however, is comparable between the species, suggesting that the latter shows higher degrees of non-adaptive plasticity, which is in line with the flexible stem hypothesis. Within Lake Tanganyika, the *Tropheini* provide further evidence for a possible flexible stem in cichlids: Kerschbaumer *et al.* (2011) studied four morphologically distinct populations of Tropheus moori, which have been split from each other for about 100,000 years (Sturmbauer *et al.* 2005). By raising broods of all four populations in a common environment, they found that plastically induced morphological variation exceeded naturally occurring population differences by a factor of 2.4. In addition, a small proportion of the variation could be linked to the genetic background. The authors also noted that plastically induced morphological patterns might have been at least partially adaptive, as experimental fish morphologies roughly coincided with the morphology of the population that lives in environments similar to the common garden. These lineages of *Tropheus moorii* may therefore be at an early stage of divergence where morphological niche segregation is still mostly due to adaptive plasticity and genetic assimilation has just begun to fix particular developmental trajectories.

In spite available data in line with a flexible stem in modern Haplochromines, the inferences from only morphological and genetic data are limited as long as the GRN

underlying the plastic response in the cichlid pharyngeal jaw remains elusive. First efforts have been undertaken to describe this GRN in *A. alluaudi* (Schneider *et al.* 2014). Studies investigating gene expression within pharyngeal jaw bone tissue during the plastic response proposed that, throughout development, involved genes are orchestrated in gene modules in all investigated 'plastic' species, as was found for other GRNs (Schneider *et al.* 2014). These gene modules are sets of functionally related genes that were observed to be jointly up- and downregulated by, potentially, only few master regulators, such as the transcription factor AP1 or the second messenger cAMP. Although data are very limited to date, the available evidence suggests that at least parts of the GRN putatively underlying ancestral pharyngeal jaw plastic development in the modern Haplochromines are preserved among plastic species. This is in line with the notion that the regulatory modules may be an effective way to manage the complexity of GRNs underlying complex traits, as only relatively few master regulators are potentially necessary to control a large number of directly downstream genes (Teske *et al.* 2004; Schneider *et al.* 2014). Thus, during genetic assimilation, relatively few regulatory changes could have dramatic phenotypic effects, and single developmental trajectories may retain functionality while other trajectories are corrupted by mutational alterations in regulatory sites. Another very recent paper found further support for a flexible stem in Lake Malawi cichlid oral jaw morphology (West-Eberhard 2003; Parsons *et al.* 2016; Schneider & Meyer 2016). The authors also found that plasticity had a specific genetic signature using a QTL approach. Combined with their previously published work, the authors identified a specific gene, ptch1, to be a likely regulator of mouth plasticity.

The general importance of regulatory evolution during cichlid diversification is also reflected in their genomes, which show increased rates of regulatory evolution particularly associated with pleiotropic genes and such underlying morphological traits (Brawand *et al.* 2014; Schneider & Meyer 2016). Our current work is focusing on understanding 'whether' and 'how' GRN degeneration may have occurred in derived cichlid lineages.

Summary & concluding remarks

In recent years, our awareness of how multifaceted and consequential the biotic and abiotic environment influences the phenotypes of all living organisms has greatly increased (Gilbert *et al.* 2015; Gilbert & Epel 2015; Sultan 2015). Phenotypic plasticity is a key response mechanism to organisms' environments. Plasticity's effects on evolutionary rates have been intensely discussed for more than a century and received renewed attention in the last decade as the molecular mechanisms become better understood and new

evolutionary models have been developed (West-Eberhard 2005; Pfennig *et al.* 2010; Ghalambor *et al.* 2015; Hendry 2015; Oke *et al.* 2015; Levis & Pfennig 2016).

We reviewed recent studies that suggest that phenotypic plasticity itself might be the substrate for rapid phenotypic diversification after colonization of a new environment, as (i.) it facilitates a population's persistence in a broad range of environments, (ii.) developmental trajectories can potentially be fixed, even by random mutations and (iii.) non-induced developmental trajectories that undergo degeneration contribute to population's cryptic genetic variation that can act as substrate for the evolution of further phenotypic diversity. Specifically, we propose that at least some non-adaptive plastic responses are a side product from a trait undergoing genetic assimilation. While only one developmental trajectory is fixed during genetic assimilation, alternative trajectories, which have been part of an adaptive response in an ancestor, are increasingly degenerating. Degenerating developmental trajectories might retain some of their inducibility and, if induced by atypical environmental stimulus, their plastic response becomes increasingly non-adaptive until environmental sensitivity is lost entirely. Adaptive radiations are illustrated here as emerging from a flexible stem using the model system of cichlid fishes in the boxes. Due to their diversity, common parallelism and convergence and well-described examples of phenotypic plasticity in key ecological traits, cichlids are likely to have a flexible stem in some traits that promoted their phenotypic diversity (Box 2–4).

We encourage future studies examining the occurrence and role of adaptive and non-adaptive phenotypic plasticity in traits that contributed to lineage diversification at different stages of lineage divergence and genetic assimilation. Comparative studies examining patterns of phenotypic, genetic and transcriptional variation associated with a diversifying trait across radiating lineages will further our understanding of the multifaceted contributions of phenotypic plasticity to the evolution of biological diversity.

Chapter III

Regulatory gene networks that shape the development of adaptive phenotypic plasticity in a cichlid fish

RALF F. SCHNEIDER, YUANHAO LI, AXEL MEYER and HELEN M. GUNTER

MOLECULAR ECOLOGY, Vol. 23, pages 4511-4526, 2014

Abstract

Phenotypic plasticity is the ability of organisms with a given genotype to develop different phenotypes according to environmental stimuli, resulting in individuals that are better adapted to local conditions. In spite of their ecological importance, the developmental regulatory networks underlying plastic phenotypes often remain uncharacterized. We examined the regulatory basis of diet-induced plasticity in the lower pharyngeal jaw (LPJ) of the cichlid fish *Astatoreochromis alluaudi*, a model species in the study of adaptive plasticity. Through raising juvenile *A. alluaudi* on either a hard or soft diet (hard-shelled or pulverized snails) for between 1 and 8 months, we gained insight into the temporal regulation of 19 previously identified candidate genes during the early stages of plasticity development. Plasticity in LPJ morphology was first detected between 3 and 5 months of diet treatment. The candidate genes, belonging to various functional categories, displayed dynamic expression patterns that consistently preceded the onset of morphological divergence and putatively contribute to the initiation of the plastic phenotypes. Within functional categories, we observed striking co-expression, and transcription factor binding site analysis was used to examine the prospective basis of their coregulation. We propose a regulatory network of LPJ plasticity in cichlids, presenting evidence for regulatory crosstalk between bone and muscle tissues, which putatively facilitates the development of this highly integrated trait. Through incorporating a developmental time-course into a phenotypic plasticity study, we have identified an interconnected, environmentally responsive regulatory network that shapes the development of plasticity in a key innovation of East African cichlids.

Introduction

Traditionally, research on adaptive morphological variation has focused predominantly on genetic variability (e.g. Clausen *et al.* 1941; Morgan 1947; Lewontin 1974). However, throughout recent decades, investigations have demonstrated that a broad range of environmental variables, both internal and external to the organism, play an instructive role in the development of adaptive phenotypes (Adler & Harvell 1990; Janzen & Paukstis 1991; Gilbert 2005; West-Eberhard 2005; Beldade *et al.* 2011). Phenotypic variability can arise when environmental cues modulate the developmental trajectories encoded by a single genotype through a process termed phenotypic plasticity (Schlichting & Pigliucci 1998; Pfennig *et al.* 2010). Integrative approaches such as 'ecological developmental biology', which incorporate evolutionary biology, developmental biology and ecology, have the power to shed new light on the role of phenotypic plasticity in evolution (Gilbert 2001; West-Eberhard 2005; Gilbert & Epel 2009). While a historical lack of strong molecular evidence has hampered the acceptance that phenotypic plasticity would promote (rather than impede) speciation, logically it stands to reason that it should promote speciation under certain ecological circumstances (Price *et al.* 2003; Nicotra *et al.* 2010). Indeed, one of the basic concepts of evolutionary biology is that selection acts on an organism's phenotype (Mayr 1963; Lande & Arnold 1983; Schwander & Leimar 2011; Scheiner & Holt 2012) regardless of whether this phenotype is genetically encoded, opening the possibility that plastic phenotypes might represent an important intermediate step in speciation. Moreover, it has been recognized that phenotypic plasticity itself can be the target of selection (Pigliucci 2007; Wund *et al.* 2008).

In spite of its potential importance in generating adaptive phenotypes, understanding the molecular mechanisms underlying phenotypic plasticity has only become technically possible in recent years. To date, several ground-breaking studies have shown that alternative phenotypes may be induced through the alteration of gene expression by epigenetic means, as well as alterations of hormonal and enzymatic activity (Denver 1997; Gilbert 2005; Aubin-Horth & Renn 2009; Jablonka & Raz 2009; Snell‐Rood *et al.* 2010; Beldade *et al.* 2011; Sommer & Ogawa 2011). Such studies typically focus on the molecular bases of polyphenisms, discrete alternative phenotypes that arise through environmentally mediated switches in developmental pathways (Abouheif & Wray 2002). Nonetheless, trait values for environmentally induced phenotypes often form a continuous distribution, the reaction norm (Woltereck 1913).

To determine the molecular bases of phenotypic plasticity, as a crucial step to evaluating its effects on species evolution, it is essential to investigate the temporal expression dynamics of environmentally responsive genes that mediate the plastic response,

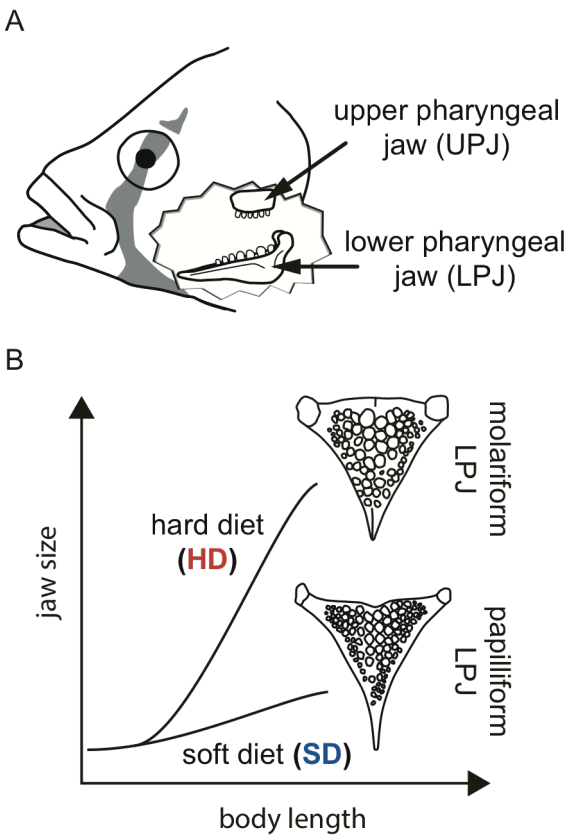


Fig. III.1 Cichlid fishes possess a mechanically robust pharyngeal jaw apparatus. (a) The PJA is comprised of a pair of upper pharyngeal jaws (UPJ) that articulate directly with the neurocranium and the lower pharyngeal jaw (LPJ), which is formed by the suturing of the 5th ceratobranchial arches. (b) For the cichlid *Astatoreochromis alluaudi*, the mechanical properties of the diet influence LPJ development, whereby individuals fed a soft diet develop a smaller, more slender papilliform LPJ, the baseline condition, while individuals fed a hard diet develop a larger, more robust molariform LPJ that withstands increased biting forces. Image modified from Hoogerhoud (1984).

for which developmental time-course experiments are a powerful tool (Aubin-Horth & Renn 2009). For organisms that respond plastically to an environmental cue, genes that are upstream in a regulatory cascade underlying a plastic trait are expected to alter their expression earlier in the presence of this inductive cue compared to genes that are more downstream (Aubin-Horth & Renn 2009). Thus, it is necessary to interpret changes in gene expression in a plastic trait throughout development, to be able to delineate the roles of specific genes throughout the plastic response.

In a previously published analysis, we characterized the molecular basis of diet-induced phenotypic plasticity in the lower pharyngeal jaw (LPJ) in an East African fish, *Astatoreochromis alluaudi* (Gunter *et al.* 2013). The LPJ forms part of the pharyngeal jaw apparatus (PJA), a key innovation of cichlid fishes that contributed significantly to their explosive adaptive radiation within <2 Myr (Liem 1974; Meyer *et al.* 1990; Elmer *et al.* 2009). Through evolving highly specialized feeding morphologies and behaviours, many East African cichlids have been able to exploit narrow and creative trophic niches, such as algae scraping, Aufwuchs plucking and insect picking, often in parallel across the three main African lakes (Meyer *et al.* 1990; Kocher *et al.* 1993; Meyer 1993a; Stiassny & Meyer 1999; Schön & Martens 2004; Young *et al.* 2009). In contrast, *A. alluaudi* is a generalist species, basal to the modern haplochromines, which inhabits Lake Victoria, its satellite lakes and

associated river systems (Greenwood 1964; Hoogerhoud 1984; Salzburger *et al.* 2005). Its' plastic LPJ allows this cichlid to exploit the available food resources efficiently across varying habitats: if their preferred diet of soft food (such as insects) is sufficiently available, they develop a slender, 'papilliform' LPJ, bearing numerous fine teeth (Slootweg *et al.* 1994). However, if soft food is scarce, individuals feed on hard-shelled mollusks that induce the formation of a robust molariform LPJ, bearing fewer, more molar-like teeth (Fig. III.1) (Greenwood 1964; Huysseune *et al.* 1994; Slootweg *et al.* 1994; Huysseune 1995). Further cases of trophic plasticity have been investigated amongst Neotropical cichlids, leading to the hypothesis that these might be more phenotypically plastic than most African cichlids (Meyer 1987b, 1989; Meyer *et al.* 1990; Meyer 1993b; Muschick *et al.* 2011). However, the general relationship between phenotypic plasticity and speciation rates is still unclear and remains hotly debated (Pfennig *et al.* 2010; Thibert-Plante & Hendry 2011; Landry & Aubin-Horth 2014).

In this study, we investigate the molecular basis of environmentally induced morphological divergence that results in the papilliform and molariform LPJ phenotypes in *A. alluaudi*. A time-course experiment was employed, exposing naive juvenile fish to experimental diets for variable time periods, allowing us to analyze gene expression across 8 months of plastic development in *A. alluaudi*. We investigated the expression of 19 candidate genes that were previously found to be associated with induced plasticity in the LPJs of *A. alluaudi* (Gunter *et al.* 2013). Of these 19 candidates, 16 belong to six main functional categories, including 'immediate early genes', 'haem pathway genes', 'matrix-related genes', 'bone-related genes', 'muscle-related genes' and 'calcium pathway genes '. We present evidence that most of these genes are not only associated with the plastic phenotypes, but that they also contributed to the early and on-going development of LPJ plasticity in *A. alluaudi*. Finally, we shed light on the putative regulatory network underlying LPJ plasticity by utilizing an analysis of transcription factor binding sites on the genome of a closely related cichlid.

Materials and methods

Fish husbandry and experimental set-up

A developmental time-course experiment was conducted on an inbred strain of *Astatoreochromis alluaudi* using modifications to the methods described inGunter *et al.* (2013), assuring minimal genetic variation between individuals. Two broods (which share the same grandparents) comprising a total of ~40 individuals were raised on Artemia salina nauplii and Tetramin flake food until they reached a standard length (SL) of ~30 mm. These individuals were split randomly in two equally sized groups, and each group was raised in a substrate-free 100-l tank for up to 8 months. Tanks were set up in a temperature-controlled

room set to 25 °C with artificial illumination, regular water changes and weekly water quality assessments ensuring minimal heterogeneity between the tanks. Comparisons to an independent study with a similar experimental design (H. M. Gunter unpublished results) demonstrated that there were no systemic tank effects (Fig. S.III.1). Each of the two groups was fed Melanoides spp. snails, with one group receiving whole snails, which the fish had to crack with their pharyngeal jaw apparatus (PJA) (the hard diet (HD) group), and the second group receiving an equivalent amount of snails that had been finely pulverized (the soft diet (SD) group). After 1, 3, 5 and 8 months of treatment, fish from each of the two groups were randomly selected and sacrificed within 30 min of feeding using a lethal dose of Tricaine (MS222). These time points were chosen to evenly cover the treatment period, which was predicted to include size of plastic divergence, previously determined to be 55 mm SL (Huysseune *et al.* 1994). Mean SL did not differ between the two diet groups for any of the four developmental time points (Fig. S.III.2). A total of 33 individuals were sampled across all developmental stages. Sex was not taken into account as previous investigations indicated that diet and SL are far greater determinants of LPJ plasticity in *A. alluaudi* (Gunter *et al.* 2013). The fish's SLs were measured, LPJs were dissected and stored in RNAlater (Qiagen) at -20 °C.

Morphometric measurements

To determine the stage at which LPJ plasticity could first be detected, the size of each LPJ was analyzed according to Gunter *et al.* (2013). These analyses were made from digital images, which were captured after the LPJs had been cleaned of surrounding connective tissue. Ten linear morphometric measurements of all LPJs were obtained following Gunter *et al.* (2013) (Fig. S.III.3). ANCOVAs were performed using R (RDC-Team 2005), utilizing each linear morphometric measure as a dependent variable, diet as a factor and SL as a covariate (Fig. S.III.4). SL was plotted against each linear measurement to evaluate the relationship between size and morphological divergence. As there was a significant linear relationship between SL and all other linear morphometric measurements (for linear regressions all $P < 0.05$), only the residuals of the respective variables were used for further analyses. The morphological differences amongst the diet groups were evaluated for each time point by performing Wilcoxon signed-rank tests in R. Multiple testing correction was not performed, but we chose rather to interpret our results with caution. All P-values are listed in the Supporting Information (Tables S.III.2,S.III.3).

RNA extraction and cDNA synthesis

RNA was extracted from the LPJ samples using the RNA Mini kit (Qiagen). In addition to the steps recommended by the manufacturer, LPJs were ground to a fine powder while

submerged in liquid nitrogen and homogenized in a FastPrep (MP Biomedicals). Polyacryl carrier (MRC) was added to increase RNA yield, and samples were treated with an optional on-column DNase treatment (Qiagen). Furthermore, additional wash steps with 80% ethanol were included to remove all traces of salt. RNA quantity was assessed using a fluorometer (Qbit 2.0), and its integrity was confirmed using a Bioanalyzer 2100 (Agilent). RNA was determined to be free of gDNA contamination according to noRT controls, and cDNA was synthesized using Invitrogen SuperscriptIII and oligo dT primer, according to Gunter et al. (2013).

qRT–PCR

We examined the expression of 19 candidate genes that were previously shown to be either up- or downregulated in the LPJs of *A. alluaudi* individuals that received a hard diet (HD), vs. the control, which received a soft diet (SD), for a period of 18 months (Gunter *et al.* 2013). The upregulated genes include: *cfos*, *ier2* and *rgs2*, termed immediate early genes (Nose & Shibanuma 1994; Versele *et al.* 1999; Ott *et al.* 2009); *klf4*, a pleiotropic transcription factor involved in differentiation of stem cells including osteoblasts and osteoclasts (Nose & Shibanuma 1994; Michikami *et al.* 2012); *ryr1*, *anxa6* and *srl*, termed calcium pathway genes (Leberer *et al.* 1989; Meissner 1994; Song *et al.* 2002; Treves *et al.* 2005); *tnnt*, *tpm4* and des, termed muscle-related genes (Perry 1998; Mantila Roosa *et al.* 2011); *col6*, *col12* and *thbs3*, termed matrix-related genes (Tucker *et al.* 1997; Gelse *et al.* 2003) and *runx2b* and *osx*, which orchestrate osteoblast proliferation and differentiation, termed bone-related genes (Nakashima *et al.* 2002; Li *et al.* 2009b). Downregulated genes include *alas1* and *c1ql*, which are putatively associated with haem biosynthesis and haematopoiesis, termed haem pathway genes (Sadlon *et al.* 1999; Mei *et al.* 2008); *abcb3* and *gif*, which are associated with MHC function (Edidin 1983; Karttunen *et al.* 2001; Nonaka & Nonaka 2010) and Vitamin B12 binding and transport, respectively (Greibe *et al.* 2012) (Table S.III.1).

Primers used in this study were either described previously in Gunter *et al.* 2013 or were derived from RNA-seq contigs generated by the same study, and their efficiency was tested using standard curves using a Bio-Rad CFX96 cycler (efficiencies ranged between 92% and 108%; Table S.III.1). As in Gunter *et al.* (2013), actinr and twinfilin were used as housekeeping genes, which were selected from an RNA-seq data set, on the basis of showing the lowest variability between diet treatments and individuals. Primer concentrations and annealing temperatures were optimized to ensure negligible dimer amplification, based on assessment of the melt curve and comparison to no template controls. All primers were used at a concentration of 0.3 pmol/lL, and annealing temperatures were 60 °C for *ier2*, *gif* and *klf4* and 55 °C for all other genes. For each gene, all samples were run in duplicate on a single 96-well plate, alongside a no template control. For each run, we ensured that

negligible/no dimer was produced based on assessment of the melt curve and comparison to the no template controls. Gene expression values were calculated for each sample and gene using the respective primer efficiencies. Relative quantification values (RQ-values) were calculated by scaling to the maximum expression value observed for each gene. Relative quantification values were further normalized to a normalization factor, calculated from the expression of the aforementioned housekeeping genes (Vandesompele *et al.* 2002).

Statistical analysis of qRT–PCR

Wilcoxon signed-rank tests were performed to analyze differences in candidate gene expression between diet groups across the four time points as data distributions were nonparametric. We did not perform correction for multiple testing (see above) and report all P-values in the Supporting Information (Table S.III.4). A PCA was performed on a co-variation matrix to explore patterns of co-expression within our gene expression data set. A hierarchical cluster analysis was then conducted using a Spearman rank correlation based matrix for specimen clustering and Euclidean distances for gene expression clustering in R, to independently examine the robustness of any identified clusters and to explore the clustering of specimens across treatment and diet groups (following Haas *et al.* (2013)). Throughout the study, we applied a significance level of $\alpha = 0.05$.

Analysis of transcription factor binding sites

To gain insight into the regulatory network of LPJ plasticity, we investigated the distribution of transcription factor binding sites (TFBSs) in the promoter regions of the candidate genes. As the genome of *A. alluaudi* has not been sequenced, the genome of the closely related cichlid *Astatotilapia burtoni* was chosen for this purpose. Both species are generalist feeders, inhabiting comparable trophic niches, and phenotypic plasticity can be induced by diet in the LPJs of each, although to a lesser extent in *A. burtoni* (H.M. Gunter, unpublished). Moreover, they display a high degree of genetic similarity, computed to be 99.13% based on exon similarity calculations performed using samtools. To calculate this similarity, we divided the total number of SNPs between the two species by the total number of aligned nucleotides (indels were masked – samtools-0.1.18; Li *et al.* 2009a).

We identified all 19 candidate genes in the genome of *A. burtoni* and used Jaspar to investigate the distribution of TFBSs in their promoters (version 5.0; http:// 129.177.120.189/cgi-bin/jaspar2010/jaspar_db.pl). The investigated interval spanned from 1 kb upstream from each gene's translation start site to the end of the first exon, encompassing both the putative promoter region and the 50-UTR (similar to Lerch *et al.* (2012)). In some cases (*cfos*, *rgs2*, *runx2b*, *alas1* and *anxa6*), more than one translation start site was identified, and we analyzed the promoters of each separately, accounting for

potential overlaps. TFBSs were selected above a conservative threshold of 0.85 (Kwon *et al.* 2012) of the relative matrix score, and results for threshold levels of 0.9 and 0.95 are also reported as an indicator of the distribution of TFBS qualities. Moreover, we focused on the distribution of binding sites for two transcription factors that have been shown to be mechanically responsive: AP1 (a heterodimer comprising a JUN and a FOS protein unit, referred to as 'JUN::FOS' in Jaspar) and CREB1 (cAMP response element-binding protein 1) (Davidovitch *et al.* 1984; Nomura & Takano-Yamamoto 2000). Finally, we examined TFBSs for *RUNX2B* ('RUNX2' in Jaspar) and *KLF4* to identify further potential regulatory pathways amongst our candidate genes.

Results

Diet influences morphological divergence in Astatoreochromis alluaudi LPJ development

Using a split brood experiment, we investigated the regulatory networks underlying the development of diet-induced adaptive plasticity in the LPJ of *Astatoreochromis alluaudi*. Our analyses indicated the on-going divergence of LPJ size between diet groups across the developmental time-course, with the majority of significant differences being detected after 5 and 8 months of treatment (Fig. III.2, Table S.III.2, Fig. S.III.3). At these time points, HD fish already displayed the typical molariform LPJ phenotype, with relative increases in jaw weight, average tooth size, centroid size and size of muscle attachment horns, in comparison with SD fish. The ANCOVA analyses on the linear morphometric data indicated a significant factor effect and a significant interaction for the log of jaw weight and suture width, suggesting morphological divergence between the two diet groups. In addition, average tooth size and centroid size also showed increased values in the HD treatment, and although they did not attain statistical significance, they displayed a trend towards morphological divergence ($0.05 < P < 0.1$) between diet treatments (Fig. III.2). Indeed, this trend was observed for most linear measurements and was supported by non-overlapping confidence intervals (Table S.III.3, Fig. S.III.4).

Our linear morphometric data show that morphological divergence in the LPJ of *A. alluaudi* was first detectable at 5 months of treatment. The confidence intervals of our scatterplots suggest that significant morphological divergence is likely to have occurred at a SL of 55–60 mm (which coincides with a treatment duration of between 3 and 5 months, Fig. S.III.2) for variables such as LPJ weight, suture width and average horn width (Fig. III.2, Fig. S.III.4). Such variables are indicative of a robust, functionally integrated molariform LPJ phenotype, where the LPJ is denser, the two 5th ceratobranchials are united by a stronger suture, and movement of the LPJ is controlled by larger, stronger muscles (e.g. Huysseune *et al.* 1994; Smits *et al.* 1996; Hulsey 2006). This divergence lies in the middle of the

developmental time-course employed by our study, enabling gene expression to be compared for stages prior to and after the onset of observable morphological divergence.

Diet influences gene expression during Astatoreochromis alluaudi LPJ development

Quantitative reverse transcription–PCR (qRT–PCR) was used to analyse the expression of previously identified candidate plasticity genes in a developmental time-course of diet-manipulated *A. alluaudi*. As determined by our morphological analyses, this time-course captured the period prior to and shortly following the development of observable

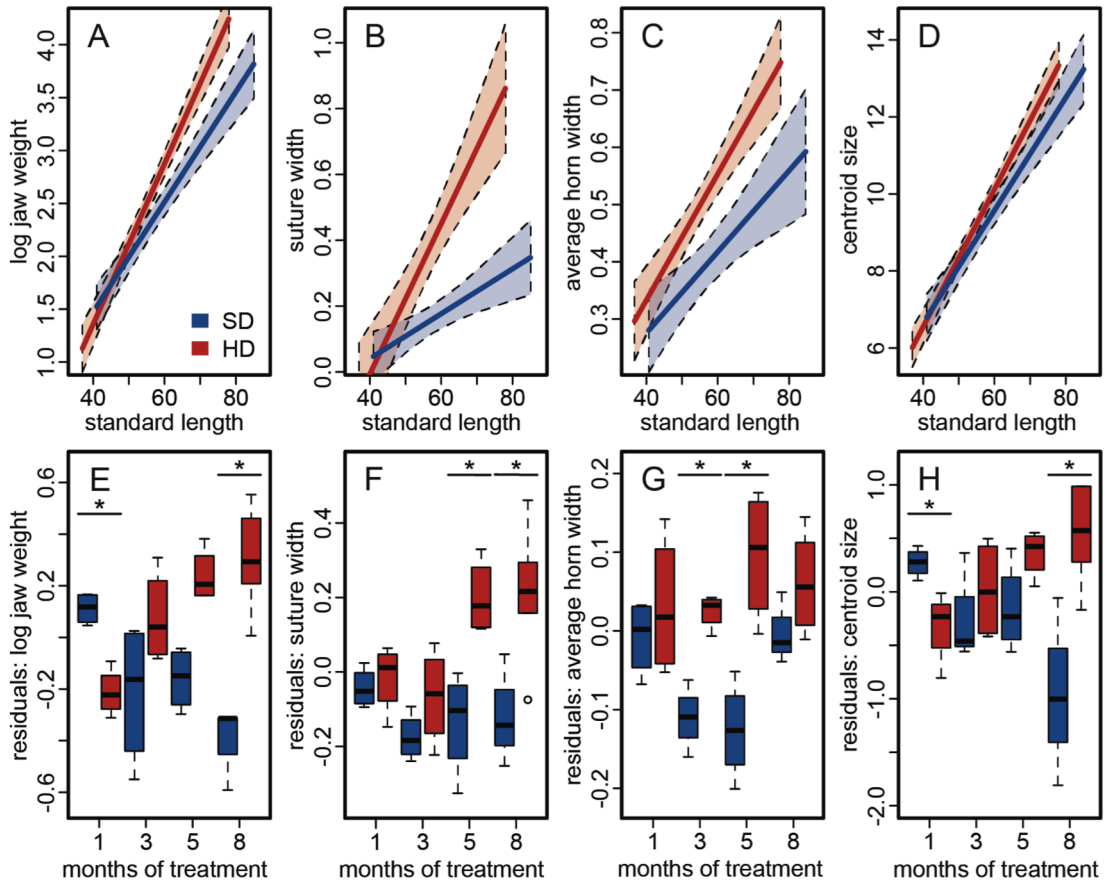


Fig. III.2 Developmental plasticity in LPJ size for *Astatoreochromis alluaudi* fed either a hard diet (HD) or soft diet (SD). Four linear morphometric measurements were plotted against (a-d) standard length and (e-h) months of treatment. Morphometric measures include (a,e) LPJ weight, (b,f) suture width, (c,g) horn width and (d,h) centroid size. (a-d) Different slopes were detected for each morphological measurement in each diet group, indicating that morphological divergence arose at ~55–60 mm. (e-h) Residuals of morphometric measurements tended to increase in HD compared with SD individuals across development (* = $P < 0.05$). Marked areas reflect 95% CIs (a-d) and boxplots show the median, the 1st and 3rd quartiles as hinges and upper and lower whiskers (e-h).

morphological differences that arose due to the plastic response. We observed significant gene expression differences between treatments after 3, 5 or 8 months of treatment (*rgs2*, *ryr1*, *anxa6* and *tnnt*; *osx* and *alas1*; *col12*, respectively) (Fig. III.3, Table S.III.4). Further trends of differential expression ($0.1 > P > 0.05$) were observed after 3, 5 or 8 months of treatment (*srl*, *tpm4* and *des*; *tpm4* and *gif*; *srl* and *osx*, respectively). After 1 month of treatment, mean expression levels were higher in SD than HD individuals for 17 of the 19 candidate genes (albeit not significantly); however, this pattern was inverted for most genes after 3 months of treatment (Fig. III.3, Table S.III.4). After treatment periods of 5 and 8 months, consistent trends were not observed amongst all genes, but we observed divergent gene expression patterns between the functional categories. Within functional categories, we observed a high degree of co-expression across most genes when diet-specific expression patterns were examined across all developmental time points. Additional analyses were thus employed to formally explore the degree of co-expression within the functional categories, including a principal component analysis (PCA) and hierarchical clustering.

Genes within functional categories display marked co-expression

PCA was used to analyze co-expression between the candidate genes, comparing gene expression values from HD and SD treatments across all developmental time points. Groups of co-expressed genes were identified through visual examination of the PCA plots (Fig. III.4, Table S.III.6). Three groups were clearly identifiable in each of the first three principle components (PCs) of the analysis, namely (i) the muscle-related genes and *srl* clustered on PC1; (ii) immediate early genes and *klf4* aligned on PC2; and (iii) matrix-related genes aligned on PC3. Additional genes grouped more loosely on the first three PCs. Specifically, *anxa6* and *ryr1* were associated with both the muscle-related genes on PC1, and the matrix-related genes on PC3. Also, the bone-related genes *runx2b* and *osx* aligned with the matrix-related genes on PC3, albeit less of their total variance was explained by this PC than for the matrix-related genes. Moreover, a few genes did not display detectable co-expression, namely the haem pathway genes, in addition to *abcb3* and *gif*. No further meaningful co-expression groups could be identified (Table S.III.6). A hierarchical cluster analysis was then conducted, using only the genes that displayed clustered co-expression according to the PCA (Fig. III.5, Fig. S.III.1, Fig. S.III.6). Similar to the PCA, three distinct clusters were detected, with gene compositions that concurred with results of the PCA. Immediate early genes and *klf4* formed one cluster, calcium pathway genes clustered with the muscle-related genes and the bone-related genes clustered with the matrix-related genes, suggesting that these clusters form three putative regulatory modules. When all expression patterns were considered, specimens clustered neither according to their treatment group nor the

treatment duration, suggesting that the treatment effects were rather subtle in our samples, consistent with our observation of subtle morphological differences.

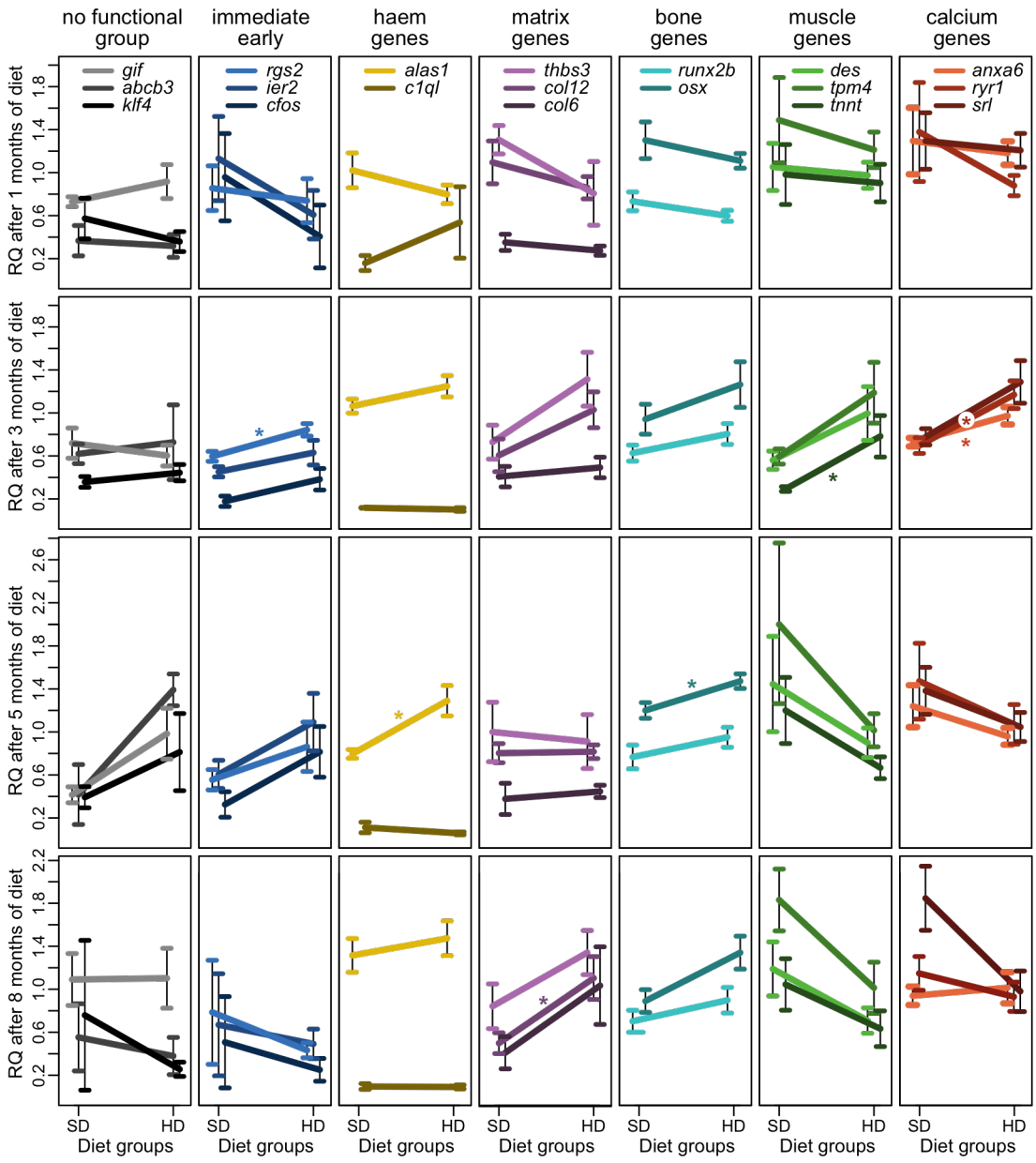


Fig. III.3 Expression of candidate genes displays a dynamic pattern during the development of diet-induced plasticity in the LPJ. Reaction norms of gene expression for SD vs. HD treatments, with expression plotted separately for treatment periods of 1, 3, 5 and 8 months, and for genes of each functional category. Genes within functional categories tend to display similar patterns of expression within each time point, while expression is highly differential between developmental stages, and between functional categories (* = $P < 0.05$). Shared line coloration hues reflect the respective functional categories.

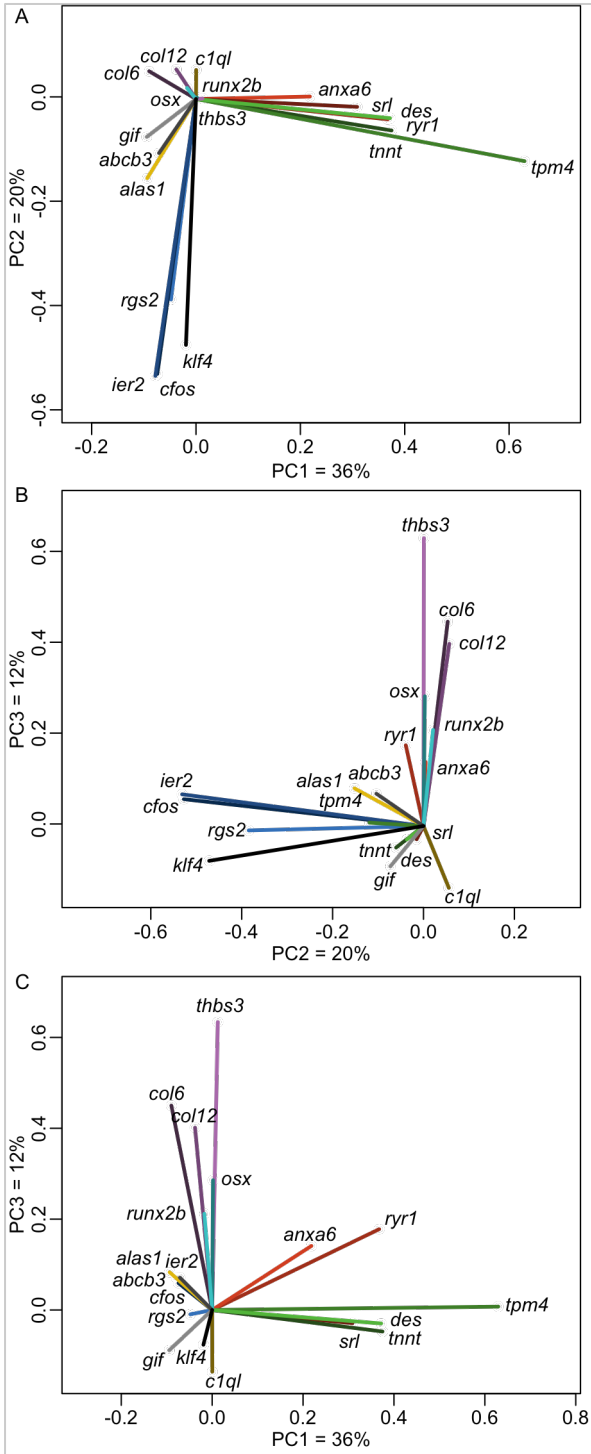


Fig. III.4 PCA on candidate gene expression patterns indicates coregulation amongst functionally related genes. (a,c) Muscle-related and calcium pathway genes cluster on PC1, which explains 36% of the total variation. (a,b) Immediate early genes and klf4 cluster on PC2, which explains 20% of the total variation. (b,c) Matrix- and bone-related genes cluster on PC3, which explains 12% of the total variation. Plot lines of the PCA are colour-coded according to the functional class of each gene: immediate early genes are blue, calcium pathway genes are red, muscle-related genes are green, matrix-related genes are violet, bone-related genes are cyan, haem pathway genes are ochre and others black (as in Fig. III.3).

Transcription factor binding site analysis indicates regulatory basis of co-expression

A transcription factor binding site (TFBS) analysis was utilized to evaluate the regulatory basis of the identified patterns of co-expression amongst functional categories. Using the TFBS detection program Jaspar, a total of 10,741 TFBSs were identified in the promoter regions and 50-UTRs of our 19 candidate genes. Of these, 244 were binding sites of four key TFs that are putatively involved in establishing the observed plastic phenotypes. These include two mechanically responsive TFs, AP1 and CREB1, in addition to two bone-related TFs KLF4 and RUNX2B (Fig. III.6, Fig. III.7, Table S.III.7).

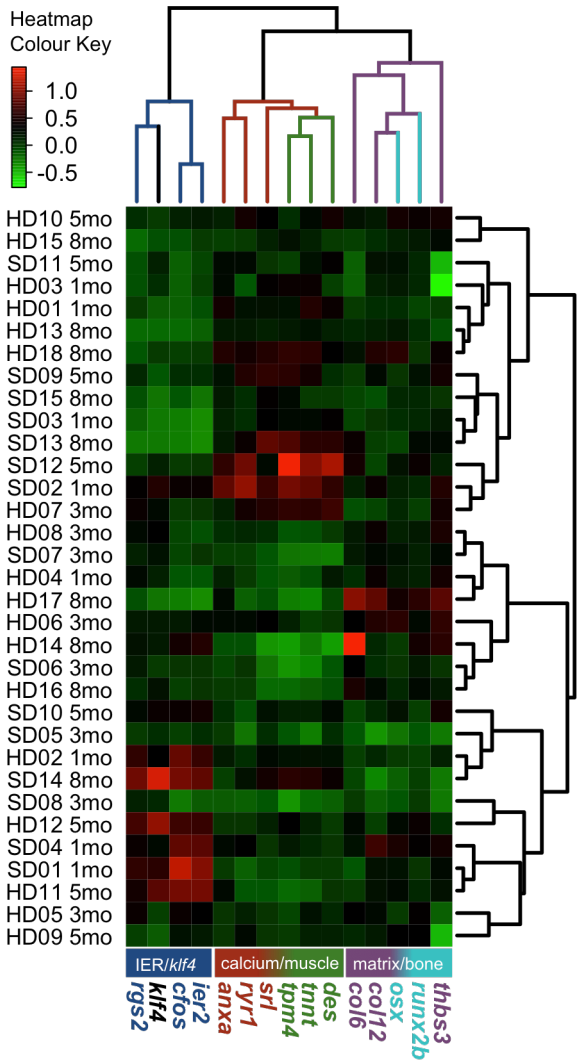


Fig. III.5 Hierarchical cluster analysis on candidate gene expression patterns indicates coregulation amongst functionally related genes. The analysis shows three main clusters: (i) *klf4* and the immediate early genes, (ii) bone-related genes and matrix-related genes and (iii) calcium pathway genes and muscle-related genes. HD = hard diet, mo = months of treatment, SD = soft diet. Coloration of the gene names refers to their respective functional class: immediate early genes are blue, calcium pathway genes are red, muscle-related genes are green, matrix-related genes are violet, bone-related genes are cyan and others black (as in Fig. III.3 &III.4).

We estimated the strength of regulatory relationships between the selected TFs and our candidate genes, based on the total number of predicted TFBSs in their promoters, and the predicted binding strength at each TFBS (based on the position weight matrix score) (Hallikas *et al.* 2006). Based on this, the strongest putative regulatory targets of AP1 were estimated to be the immediate early genes *cfos*, *rgs2* and *ier2*, both collagens and *runx2b* and des. Moreover, AP1 was estimated to have a more moderate regulatory influence on *abcb3*, *alas1*, *c1ql*, *gif*, *osx* and the remaining muscle-related and calcium pathway genes. We estimated the strongest putative regulatory target of CREB1 to be *runx2b*, while having a more moderate influence on *cfos*, *rgs2*, *klf4*, *srl* and *tnnt*. RUNX2B was estimated to have a moderate regulatory influence on *gif*, *runx2b* and *osx*. Finally, KLF4 was estimated to have the strongest regulatory influence on *alas1* and a more moderate influence on *c1ql*, *rgs2*, *klf4*, *osx*, *anxa6*, *ryr1*, *tnnt* and *des*. Of the 170 TFs for which binding sites were detected, twelve were shared amongst all candidate genes. Within functional categories, we detected characteristic TFBSs that may contribute to the observed co-expression. Amongst all immediate early genes and *klf4*, binding sites for

14 TFs were detected that were not shared by all members of any other functional group. These include the mechanically responsive TFs JUN, GATA4 and CEBPb (Sumpio *et al.* 1994; Swynghedauw 2006; Sen *et al.* 2009). For all matrix and bone-related genes, binding sites for three TFs were detected. For all muscle- related and calcium pathway genes, binding sites for two TFs were detected, including the muscle-specific TF MYOG that contributes to myogenic cell differentiation in mesenchymal stem cells (Wright *et al.* 1989). Binding sites for two TFs (MYCN and FOS) were detected in the calcium pathway genes *ryr1* and *anxa6*, but not in any other muscle-related gene or *srl*, potentially explaining their relatively distinct expression in comparison with the other genes within their co-expression cluster.

Interestingly, amongst TFs with the most abundant binding sites in the promoter regions of our candidate genes, there are two that are known to be involved in development and plasticity, which thus may have contributed to the observed plastic phenotypes (Table S.III.7). These include ARID3A, which is involved in the regulation of developmental plasticity in mouse and human cells (An *et al.* 2010) and PRRX2, which is involved in craniofacial bone development in mice (Lu *et al.* 1999) as well as matrix alterations of the vascular system and smooth muscles (Bergwerff *et al.* 1998). Binding sites for both TFs were found to be present in high numbers in the promoter of *osx*, a major regulator of bone development.

Discussion

Unraveling the molecular basis of phenotypic plasticity is a powerful first step in empirically evaluating the potential role of plasticity in evolution (West-Eberhard 2003, 2005; Bell & Aubin-Horth 2010; Beldade *et al.* 2011). Indeed, molecular tools are increasingly incorporated into studies of plasticity in phenotypic traits such as melanization in Daphnia

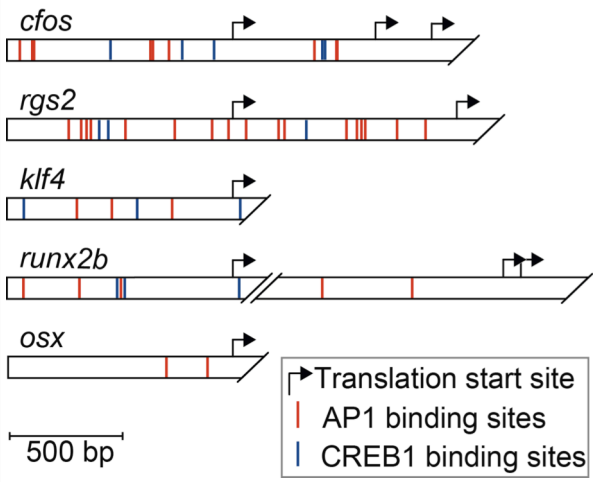


Fig. III.6 The regulatory regions of our candidate genes contain numerous predicted binding sites for mechanically responsive TFs. Colored lines indicate predicted binding sites for the mechanically responsive transcription factors AP1 and CREB1 (cAMP response element-binding protein 1) in the regulatory regions of *cfos*, *rgs2*, *klf4*, *runx2b* and *osx*. Immediate early genes *cfos* and *rgs2* include the highest density of mechanically responsive TFBSs.

(Scoville & Pfrender 2010), phase change in locusts (Wang & Kang 2014), hypoxia tolerance in killifish (Whitehead *et al.* 2011) and changing reproductive phenotypes in cichlids (Maruska *et al.* 2011). Our study examines gene expression across a developmental time-course, shedding light on the transcriptional network involved with the initiation of phenotypic plasticity in the LPJ of the cichlid *Astatoreochromis alluaudi*. Drawing upon our previous transcriptome study (Gunter *et al.* 2013), 19 candidate genes were selected, with putative functions that span various organizational levels within the molecular network that orchestrates LPJ development. By examining the expression of these candidate genes in LPJ samples taken during the early divergence of diet-induced phenotypes, we gained insight into the complex molecular network that drives the development of a famous example of adaptive phenotypic plasticity.

This study is, to our knowledge, the first to use controlled experimental conditions to examine the developmental onset of diet-induced plasticity in the LPJs of *A. alluaudi*, identifying the period between 3 and 5 months of diet treatment as a critical period for the first appearance of measurable plasticity. At this stage, the SLs of both treatment and control fish were ~55– 60 mm, consistent with previous observations on aquarium- raised *A. alluaudi* (Huysseune *et al.* 1994). This size at divergence is relatively delayed compared with the wild populations that inhabit Lake Victoria and its surrounding satellite lakes, for which LPJ divergence already occurs at an SL of ~40 mm (Hoogerhoud 1986c). Our observation suggests that the applied treatment was less intensive than that provided by natural conditions, where juvenile fish may ingest hard food items at an earlier age and where their diet is likely to be more diverse than what we have provided (Cosandey-Godin *et al.* 2008a). It is also plausible that selection may have acted upon the norm of reaction of this laboratory line, which has been bred in captivity for more than 25 years.

The candidate gene expression analyses indicated that major transcriptional changes coincide with the morphological transition period between 3 and 5 months of treatment. The first significant expression differences were detected between diet groups after 3 months of treatment, notably an upregulation of calcium pathway genes and immediate early genes amongst HD fish. Calcium channels are known to be rapidly activated in osteoblasts in response to mechanical strain (Walker *et al.* 2000), which in turn initiates the upregulation of immediate early genes (Chen *et al.* 2000). The combined upregulation of both calcium and immediate early genes after a treatment period of 3 months is consistent with our hypothesis that intense mechanical strain precedes the appearance of plastic phenotypes in *A. alluaudi*. A treatment period of 5 months was associated with a relative increase in expression of *osx*, a major regulator of osteoblast differentiation, in HD fish. Notably, 5 months of treatment coincides with the first period of detectable morphological divergence, most probably involving enhanced proliferation of osteoblasts. Finally, we

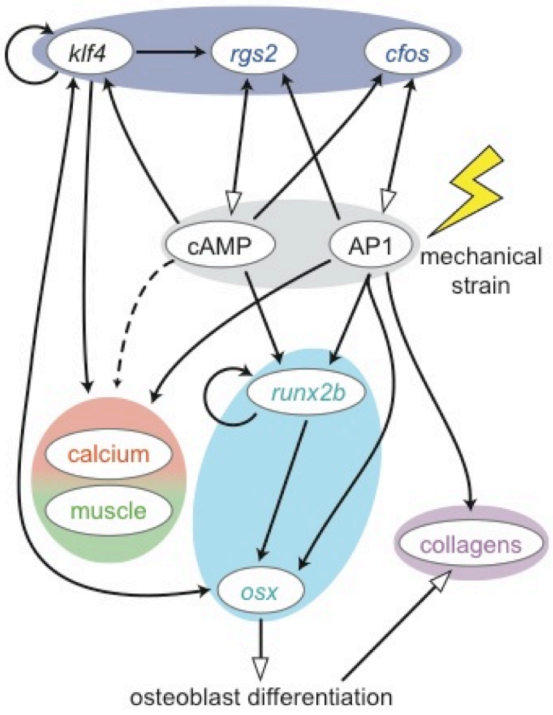


Fig. III.7 Major regulatory pathways of LPJ phenotypic plasticity as suggested by the TFBS analysis. Closed arrowheads indicate the direction of regulatory influence inferred from the analysis, while open heads indicate regulatory relationships described in previous publications (see text for details). The dashed arrow targeting 'calcium pathway' and 'muscle-related' genes indicates that only some, but not all of these genes exhibit cAMP (mediated by CREB1) TFBSs. The strain responsive transcription factor AP1 appears to be of central role in the upstream regulation of LPJ phenotypic plastic response to mechanical strain, regulating collagens, calcium and muscle-related genes, immediate early genes and *runx2b*. *runx2b* expression may also be regulated via the strain responsive second messenger cAMP (mediated by CREB1), which also influences *klf4*, *cfos*, calcium pathway and muscle-related genes and *rgs2* expression. RUNX2B appears as an intermediate TF that regulates expression of a key osteoblast regulator, *osx*, together with KLF4 and AP1.

identified *col12* to be significantly upregulated in HD fish after 8 months of treatment, which is likely to reflect alterations to the tensile properties of the extracellular matrix (Izu *et al.* 2011), concomitant with an overall increase in size of the LPJ. The majority of gene categories were relatively upregulated in HD at 3 months and beyond, albeit they infrequently attained statistical significance, likely due, in part, to our relatively low sample size. Nonetheless, it is interesting to note that statistically significant gene expression differences were observed in an order that reflects our hypothesis that mechanically responsive genes are initiated first, followed by osteoblast differentiation genes, then matrix-related genes (Gunter & Meyer 2014).

The cluster analysis indicated a high degree of co-expression amongst genes of related functional categories, suggesting that they form regulatory modules that act in concert, generating the observed phenotypic plasticity. Three major regulatory clusters were identified by our analyses, including one comprised of the immediate early genes and *klf4*, a second that includes matrix- and bone-related genes, and a third that includes muscle-related and calcium pathway genes. The identified clusters comprise genes of multiple

functional categories; however, it is likely that in this context, they are functionally related. For example, while *klf4* is a pleiotropic transcription factor (Dang *et al.* 2000), its clustering with immediate early genes suggests that *klf4* expression follows an immediate early pattern of induction in the *A. alluaudi* LPJ, similar to previous observations of its response to fluid shear stress (Peters *et al.* 2003). Moreover, the co-regulation of bone- and matrix-related genes, as well as muscle-related and calcium pathway genes could be due to the fact that osteoblasts secrete bone matrix, and calcium flux is integral to muscle function. Intriguingly, we observed differential expression of muscle genes in HD and SD LPJs, despite being cleaned of muscle prior to RNA extraction. This pattern is potentially due to the crosstalk between bone and muscle pathways (Mo *et al.* 2012; Bonewald *et al.* 2013). It has been demonstrated that muscle damage induces the differentiation of myoblastic precursors in the bone marrow cavities of mice, which migrate into the muscles and assist with their repair (Ferrari *et al.* 1998; LaBarge & Blau 2002). Thus, the observed upregulation of muscle genes in HD jaws at 3 months may be induced by damage to the LPJ adductor muscles during chewing. The detected co-expression of calcium and muscle genes further suggests that myocytes undergo differentiation in the LPJ, as calcium homeostasis is integral to myogenesis and muscle contractile function (Davies 1963; Hauser *et al.* 2008).

In order to further characterize the regulatory relationships amongst our candidate genes, the transcription factor binding sites in their promoter regions were analyzed. Specifically, we sought to identify genes whose expression is likely to respond to our diet treatments, first focusing on two mechanically responsive transcription factors, AP1 and CREB1 (cAMP response element-binding protein 1) (Davidovitch *et al.* 1984; Nomura & Takano-Yamamoto 2000), as mechanical strain is the environmental stimulus that induces plasticity in *A. alluaudi* (Gunter *et al.* 2013). AP1 binding sites were detected in the promoter regions of all 19 candidate genes, and CREB1 binding sites were detected in 12 of the 19 candidates. This abundance of TFBSs suggests that immediate early genes such as AP1 are major regulators of phenotypic plasticity in the cichlid LPJ, influencing gene expression at various levels of the regulatory cascade (Fig. III.7). Together, these observations and the detection of divergence in expression patterns prior to the onset of detectable morphological plasticity suggest that all genes in our analysis contribute to the establishment of the plastic phenotypes. However, constructing a structured regulatory hierarchy was beyond the scope of this project.

While we identified informative patterns of expression for most of our candidate genes, several genes displayed unexpected expression patterns, including some that were inverted in comparison with the results of our previous study (Gunter *et al.* 2013). Specifically, these include representatives from the immediate early, muscle-related and calcium pathway genes. We suggest that the dynamic expression of these genes reflects

heterologous cycles of bone remodeling and tooth replacement that occur in the LPJ. In response to microdamage caused by mechanical strain, bones are locally resorbed by osteoclasts, and new, stronger bone is subsequently secreted by osteoblasts, improving the mechanical robustness of the bone (Hadjidakis & Androulakis 2006). Indeed, many expression patterns induced in bones by mechanical strain vary considerably with time (Mantila Roosa *et al.* 2011). Remodeling is also an essential process in tooth replacement, which occurs approximately once per month in *A. alluaudi* (Huysseune 1995). Such dynamic gene expression patterns highlight the importance of including multiple developmental time points when trying to determine the molecular basis of phenotypic plasticity, as gene expression at a specific time point within a gene expression cycle may vary considerably. Moreover, we detected different expression patterns after 1 month of treatment, in comparison with 3, 5 or 8 months, with the majority of candidate genes being upregulated in SD (albeit not significantly). We hypothesize that this may be due to a stressful period of adjustment to the mechanically stimulating diet that caused HD individuals to receive a lower level of nutrition than SD. It is also possible that gene expression after 1 month of treatment reflects an early stage of the plastic response, characterized by an overall decrease in gene expression, potentially representing an innate cost of plasticity. However, further experiments that include a pretreatment sample and denser sampling earlier in the development of plasticity are required to investigate this phenomenon in more detail.

Our study demonstrates that environmental inputs can act on a developmental pathway at various hierarchical levels, generating an adaptive and functionally integrated trait. Nineteen candidate plasticity genes, which encode a combination of transcription factors, signaling and structural proteins, were found to contain putative binding sites for mechanically responsive TFs in their promoter regions. Moreover, trends towards differential expression of these genes were observed prior to the appearance of plastic phenotypes. In particular, muscle and calcium genes displayed tight co-expression, suggesting that mechanical strain induces the differentiation of myocytes in the LPJ, integrating the development of these two tissue types. Together, these results suggest that diet-induced mechanical strain directly influences gene expression across various pathways, which together result in the development of a functionally integrated phenotype, such as the PJA of cichlid fishes. This finding extends the hypothesis of Young (2013) that the molecular pathways underlying phenotypic plasticity involve environmental stimulation of 'upstream genes', which direct the acquisition of a plastic phenotype, further altering the expression of 'downstream genes'. Specifically, our results indicate that the mechanically responsive TFs AP1 and CREB1 can regulate the expression of a suite of LPJ development genes without signaling via intermediates. Similarly, in the sea urchin, thermal stress was shown to influence the expression of genes distributed throughout the endomesodermal

and ectomesodermal developmental network in a complex pattern that was not consistent with sequential gene regulation (Runcie *et al.* 2012).

The molariform phenotype of durophagous cichlids such as *A. alluaudi* is complex, involving coordinated alterations in the size of the pharyngeal jaws and their adductor muscles alongside an increase in tooth size and decrease in tooth number (Hoogerhoud 1986a; Huysseune *et al.* 1994; Huysseune 1995). Mismatch in any of these elements would render individuals less able to process hard food items, particularly in the light of the tight architectonic constraints within the oral cavity of teleosts (Meyer 1989; Smits *et al.* 1996; Chapman *et al.* 2001; Binning *et al.* 2010). Our previous investigation proposed that pleiotropic genes are likely to guide the development of both teeth and jaws (Gunter *et al.* 2013). Here, we propose a further molecular explanation for integrated development of muscles and bones of the PJA, namely through coordinated signaling between both tissues, which induces the differentiation of muscle precursor cells in the medullary cavity. Plasticity favors the development of complex, functionally integrated phenotypes without relying on multiple interdependent mutations, pleiotropy or linkage disequilibrium (West-Eberhard 2003, 2005). These investigations provide strong molecular evidence for the integration of different developmental modules during the establishment of adaptive phenotypes, which has previously been demonstrated on a morphological level (Badyaev *et al.* 2005).

In conclusion, through examining the expression of putative plasticity genes in a developmental context, we have identified that a strongly environmentally responsive network underlies adaptive plasticity in a key innovation of a cichlid fish. Our analysis identified the coordinated coexpression of functionally related genes, which we consider to represent regulatory modules. As mechanical strain directs the expression of each of these modules, it promotes the development of an integrated, complex phenotype, the molariform PJA, which enables the efficient exploitation of an alternative trophic niche in *A. alluaudi*. This study forms a platform to empirically assess the molecular trajectories that underlie adaptive phenotypes across a phylogenetic context in cichlid fishes. Ultimately, it will help to evaluate the importance of phenotypic plasticity in cichlid evolution, specifically in the light of evolutionary mechanisms such as genetic assimilation and genetic accommodation (Waddington 1953; West-Eberhard 2005).

Chapter IV

Molecular investigation of genetic assimilation during the rapid adaptive radiations of East African cichlid fishes

HELEN M. GUNTER, RALF F. SCHNEIDER, IMMANUEL KARNER, CHRISTIAN STURMBAUER & AXEL MEYER

Prepared for resubmission to MOLECULAR ECOLOGY, 2017

Abstract

Adaptive radiations are formed by adaptive diversification and speciation within a lineage that result in many ecologically specialized, phenotypically diverse species. It has been proposed that adaptive radiations originate from ancestral lineages with high developmental plasticity in adaptive traits, facilitating ecologically driven phenotypic diversification that later becomes genetically fixed through genetic assimilation. This study aimed to investigate whether phenotypic plasticity may underlie adaptive differences in the trophic apparatus of East African cichlid fishes, and if genetic assimilation might have fixed these differences during their adaptive radiation. A split brood experimental design compared adaptive plasticity in species from within and outside of adaptive radiations, whereby a plastic response was induced in the pharyngeal jaws through feeding individuals either a hard or soft diet. We find that non-radiating, basal lineages show higher levels of adaptive morphological plasticity than the derived, radiated lineages, suggesting that these differences have become partially genetically fixed during adaptive radiation. Two candidate genes that may have undergone genetic assimilation were identified. Taken together, our results suggest that genetic assimilation may have dampened the inducibility of plasticity related genes during the adaptive radiations of East African cichlids, canalizing their feeding phenotypes to suit progressively more narrow ecological niches.

Introduction

Adaptive radiations often involve either the colonization of novel habitats or movement into vacant niche space (i.e. ecological opportunity) that can arise after the eradication of existing diversity (Losos 2010). The formation of such species flocks is characterized by an explosive increase in species number within a lineage, typically via specialization to a multitude of ecological niches (Schluter 2000). While intense competition between species for limited resources often results in extinction (competitive displacement) (Sepkoski Jr 1996), lineages that undergo adaptive radiation respond, more often, through diversification via the *de novo* occupation of vacant ecological niches (ecological opportunity) (Schluter 2000). Thus, identifying the unique features of lineages that undergo adaptive radiations provides a first step in understanding the mechanisms that promote speciation (Brawand *et al.* 2014; Henning & Meyer 2014). In spite of the potential value of stem lineage characterization, few studies of adaptive radiation have focused on this aspect (Schluter 2000; West-Eberhard 2003).

It has been hypothesized that adaptive radiation may be initiated by lineages that have recently evolved key innovations (Simpson 1953), or display high levels of variation that is either genetic (Hedrick 2013), plastic (Pfennig *et al.* 2010) or a combination of the two (Gomez-Mestre & Jovani 2013). Subsequent inhabitation of new ecological niches causes these lineages to undergo diversification through processes such as ecological character displacement (Schluter 2000, but see (Stuart & Losos 2013)), where competitive interactions between sympatric species drive the evolution of exaggerated differences in phenotype and niche utilization.

Increasingly, evidence shows that adaptive phenotypic plasticity contributes to trait variability within lineages that are prone to adaptive radiation. Indeed, some initial phenotypic plasticity in key ecological traits is found in many of the model adaptive radiations (West-Eberhard 2003), such as sticklebacks (Wund *et al.* 2008; Wund *et al.* 2012; Lucek *et al.* 2014; Morris *et al.* 2014; Oke *et al.* 2015), Anolis lizards (Losos *et al.* 2000; Kolbe & Losos 2005), Darwin's finches (Tebbich *et al.* 2010), Hawaiian spiders (Yim *et al.* 2014) and cichlid fishes (Meyer 1987b; Muschick *et al.* 2011). Moreover, a recent study linked phenotypic plasticity to a major evolutionary transition – the origin of tetrapods (Standen *et al.* 2014). In spite of this, phenotypic plasticity is frequently omitted from formal discussions of adaptive radiation and evolution in general (West-Eberhard 2003). During the initial phases of adaptive radiation, phenotypic plasticity can provide an inherent competitive advantage during the colonization of new or heterogeneous habitats (Baldwin 1896; Richards *et al.* 2006; Ghalambor *et al.* 2007; Lande 2009; Morris *et al.* 2014). Phenotypically plastic species, which are commonly generalists (Van Tienderen 1997), are

also more likely to meet their resource requirements during the colonization of novel environments (Baldwin 1896). Additionally, phenotypic plasticity enhances trait variability within stem lineages, through generating multiple phenotypes from single genotypes (West-Eberhard 2003), through exaggerating trait differences between lineages (Schluter 1994) and through increasing the accumulation of standing genetic variation (Gomez-Mestre & Jovani 2013). Together, this would be expected to enhance the probability of rapid and parallel colonization of open niches, as is observed in adaptive radiations (Pfennig *et al.* 2010). Although phenotypic plasticity is likely to contribute to lineage diversification, the degree to which it contributes to this process remains unresolved (Hendry 2015), and is an ongoing topic of debate (De Jong 2005;Schneider & Meyer 2016).

Although it is broadly accepted that phenotypic plasticity can enhance survival in novel environments and trait variability, induced traits are very rarely passed onto the next generation as they seldom impact the germ line (see (Morgan *et al.* 1999) for an exception). So how might phenotypic plasticity impact speciation? A mechanistic link has been proposed, namely genetic assimilation (Crispo 2007; Pfennig *et al.* 2010; Schneider & Meyer 2016). In this scenario, alternative plastic phenotypes that are consistently induced over many generations through sustained exposure to particular environmental conditions are exposed to selection, and their environmental sensitivity subsequently declines (Waddington 1953, 1961). This may contribute to speciation when alternative adaptive phenotypes become reproductively isolated and natural selection acts either positively (DeWitt *et al.* 2000) or negatively (Ghalambor *et al.* 2015) to the direction of plastic change, also dependent on the cost of plasticity in a given ecological context (Van Kleunen & Fischer 2005; Auld *et al.* 2009; Murren *et al.* 2015). Moreover, selection can refine the phenotype beyond the values displayed by the original, plastic lineage, where an exaggerated phenotype provides a higher level of fitness. Although genetic assimilation has been demonstrated empirically in fruit flies by Conrad H. Waddington (1953), and a putative molecular mechanism has been proposed (Waddington 1961; Rutherford & Lindquist 1998), genetically assimilated loci have rarely been studied in natural populations, so their importance in natural evolutionary processes is difficult to evaluate. The mutations that contribute to genetic assimilation are thought to have arisen in the stem lineage through random processes, and specifically, their subsequent selection should influence on the degree of plasticity of ecologically relevant traits. For example genetically assimilated loci in a derived, specialist lineage may show a blunted expression response to a stimulus that exerts a plastic response in the stem lineage. In the case of adaptive radiations, then restriction of plasticity would make the trait increasingly heritable and thus allowing natural selection to fine-tune the phenotypes to the specific niche requirements. Thus, specialist lineages can outcompete more plastic, generalist lineages in such niches.

In order to test the role of phenotypic plasticity and subsequent genetic assimilation in adaptive radiations, we selected East African cichlid fishes as our models. As one of the most species-rich families of vertebrates, and with adaptive radiations that occurred independently on multiple continents, cichlid fishes represent a prime model for investigating speciation via adaptive radiation (Schluter 2000; Genner *et al.* 2007; Friedman *et al.* 2013). The adaptive radiations of the East African rift lakes are particularly impressive, with around 500 species in Lake Victoria alone, most of which evolved in less than 100 000 years (Meyer *et al.* 1990; Salzburger *et al.* 2005). Phenotypic plasticity may well have contributed to these dramatic radiations, as it plays a role in generating alternative feeding phenotypes that fine-tune the exploitation of specialized niches (Meyer 1987b, 1990; Wimberger 1991; Stauffer & van Snick Gray 2004; Van Dooren *et al.* 2010; Muschick *et al.* 2011; Machado‐Schiaffino *et al.* 2014). Arguably the best-studied cichlid example of adaptive phenotypic plasticity is *Astatoreochromis alluaudi* (Greenwood 1964; Gunter *et al.* 2013; Hoogerhoud 1986b; Hoogerhoud 1986c; Huysseune *et al.* 1994; Huysseune 1995; Smits *et al.* 1996; Smits 1996; Schneider *et al.* 2014). This species modulates the development of its pharyngeal jaws (a second pair of "crushing" jaws in the throat), in response to the robustness of the available diet. Ingesting a soft diet (SD; e.g. insect larvae) induces a slender 'papilliform' jaw with numerous fine teeth, while a hard diet (HD; hard-shelled snails) induces a hypertrophied 'molariform' jaw with larger, molar-like teeth, most notably in the Lower Pharyngeal Jaw (LPJ) (Fig. IV.1A). Thus, the resulting adults were suggested to more efficiently exploit the two alternative diet niches and avoid competition with other more specialized species (Greenwood 1959, 1964; Slootweg *et al.* 1994; Hulsey *et al.* 2005; Cosandey-Godin *et al.* 2008b). In addition to numerous morphological studies, recent molecular investigations of adaptive plasticity in the pharyngeal jaws of *A. alluaudi* have also shed light on its transcriptional underpinnings (Gunter *et al.* 2013; Schneider *et al.* 2014). Specifically, this research has identified a number of genes that display dynamic patterns of expression throughout development of molariform and papilliform jaws. These genes are putatively involved in the development of bone, tooth and muscle, as well as the response to mechanical strain, the putative plasticity stimulus.

It was the goal of this research to investigate whether diet-induced adaptive phenotypic plasticity is likely to have contributed to adaptive radiations in East African cichlid fishes. To examine this we conducted a comparative plasticity study that included two specialist species from two independent adaptive radiations in Lakes Victoria (*Haplochromis ishmaeli*) and Tanganyika (*Tropheus moorii*) and two riverine generalist species branching more basally - outside these adaptive radiations (*A. alluaudi*, and *Astatotilapia burtoni*), plus a member of a more basally branching ancestral lineage (Fig. IV.1A) (*Pseudocrenilabrus multicolor*: for phylogenetic relationships see Salzburger *et*

al. (2005) and Seehausen 2013. The relatively basal, non-radiating riverine lineages are hypothesized to fill a similar ecological niche to the common ancestor(s) of the radiations that colonized the lakes (Meyer 1993b; Clabaut *et al.* 2007), however it should be noted that this is an approximation of the ancestral state as a re-construction of the actual ancestor experimentally impossible (Shapiro 2016). Selected species from the adaptive radiations occupy more specialized trophic niches (Greenwood 1964; Yamaoka 1983). Broods were split and fed a diet of either hard-shelled snails (hard diet) or pulverized snails (soft diet) and their LPJs were sampled for analyses of morphology and gene expression. In line with the 'flexible stem' hypothesis (West-Eberhard 2003; Wund *et al.* 2008; Muschick *et al.* 2011), our research aims to determine whether the cichlid adaptive radiations are likely to have evolved from a phenotypically plastic ancestral lineage, which specialized and radiated to fill an array of narrow ecological (trophic) niches. This implies that the specialist species from within the adaptive radiations would be expected to show a lower degree of adaptive phenotypic plasticity as assessed by both morphological variation and gene expression than the more basal generalist species. Moreover, we aimed to identify genes that have putatively undergone genetic assimilation. We predict that such genes would show dampened differential expression between induced and baseline phenotypes in derived, non-plastic lineages than in more basal lineages (Fig. IV.1B).

Materials and Methods

Cichlid diet experiments

Five cichlid species were selected for diet manipulation experiments, based on previous evidence of diet-induced phenotypic plasticity, phylogenetic position and the diets of natural populations (Fig. IV.1A). These included two relatively basal species: *Astatoreochromis alluaudi* and *Pseudocrenilabrus multicolor. A. alluaudi* is known to have a highly plastic pharyngeal jaw that becomes larger and more robust in response to a hard diet such as hard-shelled snails. It is geographically widespread with populations that inhabit both riverine and lacustrine environments, exploiting both insects and algae (soft diet) and mollusks (hard diet) (Greenwood 1964; Witte 1980; Slootweg *et al.* 1994; Cosandey-Godin *et al.* 2008a). We used a laboratory population that was obtained from the Mwanza Gulf in Lake Victoria in 1984, and has been used in numerous plasticity experiments (Gunter *et al.* 2013; Hoogerhoud 1986b; Huysseune 1995;Huysseune *et al.* 1994; Smits *et al.* 1996; Smits 1996; Schneider *et al.* 2014). As a second outgroup we used the generalist riverine species *P. multicolor*, which branched basally to *A. alluaudi* (Salzburger *et al.* 2005; Binning & Chapman 2008) and has no previous record of plasticity in the pharyngeal jaws, albeit its gills (which are adjacent to the LPJ) are known to be plastic (Chapman *et al.* 2001; Crispo & Chapman 2008; Crispo & Chapman 2010). Additionally,

more phylogenetically derived species were included: the riverine species *Astatotilapia burtoni*, a generalist species that exploits mostly insects, but also plants and algae (Salzburger *et al.* 2005; Sturmbauer *et al.* 2003; Theis *et al.* 2014), *Haplochromis ishmaeli*, a snail-cracking specialist endemic to Lake Victoria (Greenwood 1964; Hoogerhoud 1986c; Slootweg 1987) and *Tropheus moorii*, an algae browsing specialist that is endemic to Lake Tanganyika (Yamaoka 1983; Sturmbauer *et al.* 1992; Sturmbauer & Meyer 1992; Sturmbauer *et al.* 2003). All species had been bred in captivity for multiple generations and were obtained from laboratory strains (*A. alluaudi*, *A. burtoni* and *T. moorii*) or the aquarium trade (*H. ishmaeli* and *P. multicolor*).

Diet manipulation experiments were conducted using methods modified from Gunter *et al.* 2013 and Meyer (1990). Briefly, 1-2 clutches of juvenile fish from each species (all bred at U Konstanz) were raised separately then split into two groups and fed diets that differed only in the mechanical strain required for processing, but not in their nutritional content. One group ingested a hard diet (HD) composed of hard-shelled snails that had to be cracked with their pharyngeal jaws, and the second group ingested a soft diet (SD) of an equivalent quantity of crushed snails. Both were also fed ad libidum with Tetramin flake food each morning. The experiment was terminated after approximately eight months, a period of time known to induce significant plasticity in *A. alluaudi* (Schneider *et al.* 2014). Our previous work showed that plasticity-driven size and shape differences develop gradually (Schneider *et al.* 2014), with statistically significant divergence after ~five months of treatment. For all species, HD individuals were observed to take the snails into their mouths during the experimental period (pers. obs.). The standard length for each individual was noted and the lower pharyngeal jaws (LPJ) were dissected and stored in RNA-later at -20°C for further processing. All statistical analyses were performed with R (Team 2013).

1. Analytical workflow

Three main datasets were analyzed in this study: LPJ linear morphometric measurements, LPJ geometric morphometric measurements and LPJ candidate gene expression measurements (details below). Each of the three raw datasets was processed, either for between-species comparisons or for within species diet-group comparisons (HD vs SD). For all six data subsets the following analyses were performed: (i) a Principle Component Analysis (PCA), in which all Principle Components (PCs) are considered until their cumulative explained variance exceeds 95% of the total variance; (ii) an Analysis Of Variance (ANOVA) or Multivariate ANOVA (MANOVA), testing for differences in PCs by species or diet (selection was depended on the number of considered PCs); (iii) pair-wise comparisons among species or between diet groups, performed separately for each considered PC (when species were compared, Tukey-HSD post-hoc correction was

performed). Additional dataset-specific analyses were performed as outlined below. Finally, we integrated gene expression and linear morphometric measurements using partial least squares regression (PLSR).

1.1 Linear morphometric analyses

LPJs were cleaned and photographed and linear and geometric (see below) morphometric measurements were made according to Gunter et al. (2013). Final sample sizes for linear morphometric measurements were (n=SD,HD): *A. alluaudi* n=9,11; *A. burtoni* n=15,11; *H. ishmaeli* n=8,9; *P. multicolor* n=5,9; and *T. moorii* n=7,6. Briefly, linear measurements included length, width and depth of the LPJ, width of the muscle attachment horns, various tooth measurements and LPJ weight (Fig. S.IV.1). For species comparisons, data were first standardized through dividing morphometric measurements by standard lengths and then centering each variable (through subtracting the mean value from each data point). For diet group comparisons, first a linear regression was fitted separately for each species and measurement using the morphometric variable as a dependent- and standard length as an independent variable to obtain size-corrected residuals. These fit residuals were then used subsequently for within-species diet group comparisons. In addition to the aforementioned PCA and downstream analyses, pair-wise comparisons of diets were conducted on all measurement variables using t-tests for unequal variances and False Discovery Rate (FDR) post-hoc p-value corrections.

1.2 Geometric morphometric analyses

The geometric morphometrics analyses used two landmarks and 14 semi-landmarks that outlined the LPJs (Fig. S.IV.1D). Standardization for species comparisons was done by aligning all species' landmarks together using the R geomorph package procrustes fit (Adams & Otárola‐Castillo 2013). For diet group comparisons, LPJ landmarks were aligned species-wise. Only the symmetric component of LPJ variation was considered, i.e. variations between the left and right side were removed. No allometric effect of standard length on shape variation could be detected within species. Final sample sizes were (n=SD,HD): *A. alluaudi* n=9,11; *A. burtoni* n=11,9; *H. ishmaeli* n=8,9; *P. multicolor* n=5,9; and *T. moorii* n=6,6.

1.3 Gene expression analyses

RNA was extracted from the LPJs using a modified protocol, suitable for extractions from bones (Gunter *et al.* 2013). Candidate genes were selected from our previous publications (Gunter *et al.* 2013; Schneider *et al.* 2014), based on having putative functions in driving plastic development in the LPJs. These fall into different categories, which are likely to

influence different aspects of the development of plasticity in the LPJ, including the immediate response to mechanical strain, bone modeling and re-modeling, and the development of larger teeth and muscles. Specifically, these include the "immediate early" genes *c-fos* and *rgs2*; genes that influence the osteoblast lineage *runx2b* and *osx*; extracellular matrix genes *col12* and *col6*; calcium pathway genes *ryr* and *srl*, and the muscle-related genes *tpm4* and *des*. Moreover, we included *gif*, *alas1* and *c1q-like*, genes with putative roles in the inflammatory response, which were observed to be repressed or induced in response to a strain-inducing diet in *A. alluaudi* (Gunter *et al.* 2013; Schneider *et al.* 2014). These genes showed a dynamic expression pattern during plastic development (Schneider *et al.* 2014) and the analysis of gene expression in the LPJs at ~8 months represents a snapshot of this pattern.

RNA extraction and gene expression analyses were performed only for *A. alluaudi*, *A. burtoni* and *H. ishmaeli* as morphometric measurements indicated no significant diet-induced divergence in *P. multicolor* and *T. moorii* (see Results section). Synthesis of cDNA and qRT-PCR were performed according to Schneider *et al.* (2014). Briefly, after confirming the quality of total extracted RNA (Bioanalyzer RIN values above 6.0 and A260:280 >1.8), cDNA was synthesized using Superscript III, primed by oligo dTs. Primers of candidate genes were designed to *A. alluaudi* based on RNA-seq contigs (Table S.IV.1; Gunter *et al.* (2013)), and their efficiencies were optimized for *A. alluaudi*, *A. burtoni* and *H. ishmaeli* through analyses of standard curves. Specifically, each primer pair was tested on a dilution series generated from pooled cDNAs from each species. In the cases where E<1.85 or >2.15 new primers were designed, based on genome traces for *A. burtoni* (*osx*, *rgs2* and *gif*)(Brawand *et al.* 2014), or alternative regions of the *A. alluaudi* gene sequence for *H. ishmaeli* (*actinR*, *twinfilin*, *osx, runx2* and *gif*). Final sample sizes were (n=SD,HD): *A. alluaudi* n=9,11; *A. burtoni* n=11,9; *H. ishmaeli* n=8,9.

qRT-PCR was used to analyze candidate gene expression in two technical replicates for each individual, where expression levels were averaged (Gunter *et al.* 2013). For among-species comparisons, gene expression levels were calculated using a specific formula that enabled between-species comparisons (see Supplementary information). Although our e-values were calculated using best practice methods, any biases in their calculation would more strongly affect the species comparisons than the between diet, within species comparisons. Across the three species, gene expression was then scaled and centered, i.e. the mean for each gene was subtracted from individual values. For diet group comparisons, we calculated standardized and normalized gene expression levels according to reference gene expression levels following Gunter *et al.* (2013). These gene expression values were then scaled and centered species-wise.

In addition to the aforementioned analyses, diet-induced differences in candidate gene expression were analyzed using pairwise t-tests (for unequal variances). To correct for multiple testing, false discovery rate corrections were applied to all p-values for each species. We conducted two-way ANOVAs on our gene expression dataset, using 'gene expression' as the dependent variable and 'diet-group' and 'species' as factors. To investigate changes in reaction norms between species, the interaction of the two factors was also included and all ANOVAs were performed as pairwise comparisons, each including only two of the three species. Notably, as gene expression was standardized gene-wise, gene effects are not informative so we focus our interpretation on the interaction term. To explore gene co-expression patterns, hierarchical clustering analyses were performed for each species (Haas *et al.* 2013). Finally, linear Discriminant Function Analyses (lDFA) were conducted for *H. ishmaeli*, *A. burtoni* and *A. alluaudi* to predict species and diet group

Table IV.1 Results of two-way ANOVA on single gene expression values. Note that species estimates are non-informative as gene expression values were standardized species-wise.

Gene	Species comparison	R square	Est. diet	Est. species	Est. interact
runx2b	alluaudi vs burtoni	0.277	0.095	0.528 **	-0.010
osx	alluaudi vs burtoni	0.577	-0.016	0.778 ***	-0.237
	alluaudi vs ishmaeli	0.172	-0.016	-0.062	0.247 *
	burtoni vs ishmaeli	0.529	-0.254 .	-0.840 ***	0.484 *
gif	alluaudi vs burtoni	0.406	0.691 **	0.184	0.086
	alluaudi vs ishmaeli	0.503	0.691 ***	0.438 **	-0.369 .
	burtoni vs ishmaeli	0.335	0.777 ***	0.254	-0.455 .
alas1	alluaudi vs burtoni	0.457	0.430 **	0.343 **	-0.014
	alluaudi vs ishmaeli	0.482	0.430 ***	0.518 ***	-0.349 *
	burtoni vs ishmaeli	0.298	0.416 ***	0.175 .	-0.334 *
col6	alluaudi vs ishmaeli	0.207	-0.412	-0.934 **	0.864 *
	burtoni vs ishmaeli	0.191	0.163	-0.266 .	0.288
col12	alluaudi vs burtoni	0.156	0.174	0.579 **	-0.477
	burtoni vs ishmaeli	0.565	-0.303 *	-0.858 ***	0.573 **
ryr1	burtoni vs ishmaeli	0.142	0.079	0.341 *	-0.089

memberships, respectively, using gene expression data (using the lda() function from the R "MASS" package). Accuracy and Leave One Out (LOO)-accuracy was calculated using the same dataset that was used to calculate the predictor to estimate predictor quality.

1.4 Integrating gene expression and morphology

To integrate gene expression measurements and morphometric measurements a PLSR was performed for *H. ishmaeli*, *A. burtoni* and *A. alluaudi* separately (using the plsreg1() function from the R "plsdepot" package; Sanchez & Sanchez 2012). The gene expression matrix was used as x matrix (the independent variable) and the PC1 of species-specific PCAs on linear morphometric measurements was used as y (the dependent variable). The PC1 of these analyses was chosen as it represents the main variation found in the linear morphometric

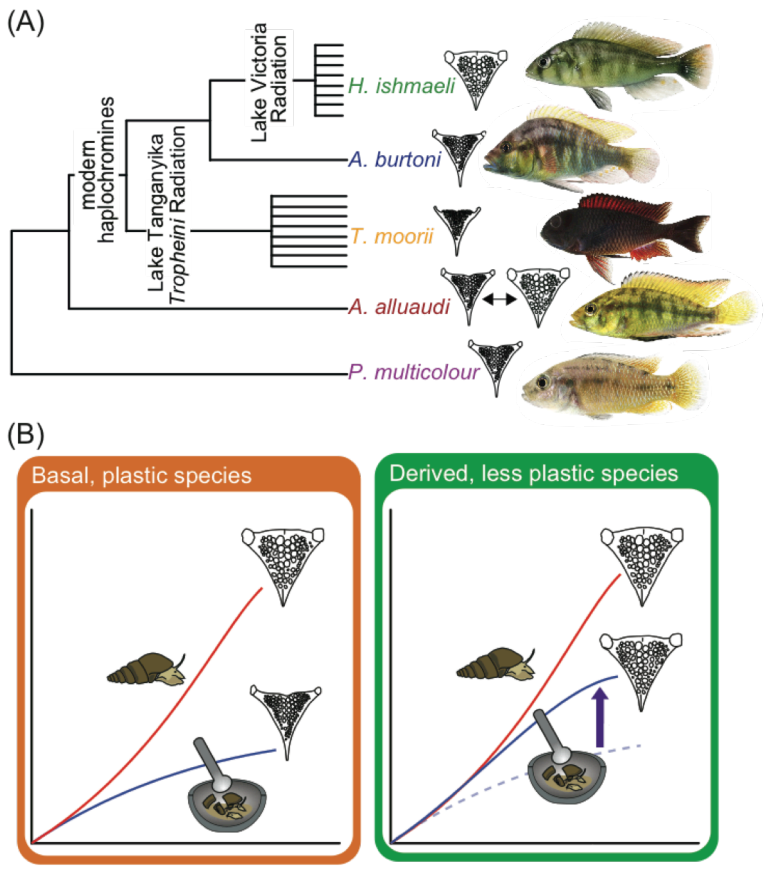


Fig. IV.1 Evolutionary relationships and lower pharyngeal jaw (LPJ) morphology sketches of cichlid species included in this study. (A) Representatives of lineages basal to adaptive radiations include the riverine generalist species *Astatoreochromis alluaudi*, *Pseudocrenilabrus multicolor* and *Astatotilapia burtoni.* More "derived" species from adaptive radiations include *Tropheus moorii*, a specialist representative of the Lake Tanganyika Tropheini radiation, and *Haplochromis ishmaeli*, a specialist representative of the Lake Victoria radiation. (B) Hypothesis for the direction of change in morphology and gene expression under the assumption of genetic assimilation. Here, the baseline condition (soft diet) of more derived species should approach the induced condition (hard diet) for plastic species, and there should be overall a more limited degree of plasticity.

measurements datasets and because diet groups are significantly differentiated on it for all three considered species.

Results

LPJs of the focal species have distinct sizes, shapes and gene expression patterns

After conducting controlled feeding trials on the five focal cichlid species, LPJs were dissected, enabling both linear and geometric morphometric analyses to be performed. Our among-species PCA analysis of linear morphometric measurements confirmed that each of the five investigated species has a unique LPJ morphology, putatively reflecting their ecological niches (Fig. IV.1, IV.2A, S.IV.2). The first PC, which was predominantly loaded with LPJ weight, explained >95% of the total variation (Fig. IV.2B,C). An ANOVA indicated that species differ significantly across PC1 (df=4, F=56.5, p<0.001). All pairwise comparisons between species on PC1 indicate significant differences, except for *A. alluaudi* vs. *H. ishmaeli* and *T. moorii* vs. *P. multicolor* (Fig. IV.2D).

Similarly, our PCA of geometric morphometric measurements indicates moderate separation between the focal species, albeit to a lesser degree than the linear morphometric analyses (compare Fig. IV.3 to IV.2A). Here, PC1 explains 51% of the variation, and PC2 explains 25% of the variation, with 3 further PCs required to explain 95% of the variation (Fig. S.IV.3). Jaw length and the shape of the posterior margin of the LPJ contribute most significantly to PC1 (Fig. IV.3), measures that are associated with relative molariformity (Muschick *et al.* 2011; Gunter *et al.* 2013). The different species clustered in a similar order to the linear morphometrics PCA, with *A. alluaudi* showing the most strongly molariform phenotype according to its position on PC1, and *T. moorii* showing the most papilliform phenotype. A MANOVA confirms that species differ according to the considered PCs (df=4, $Pillai$=1.6, F=10.2, p<0.001).

A PCA on gene expression was performed for the three species to visualize transcriptional differences that may underlie species-level differences in LPJ morphology (Fig. IV.4, S.IV.4). All species could be separated on the first two PCs, whereby *H. ishmaeli* was significantly separated from *A. burtoni* and *A. alluaudi* on PC1, and *A. burtoni* was significantly separated from *H. ishmaeli* and *A. alluaudi* on PC2 (Fig. IV.4A, D). PC1 explains 28% of the variation between individuals, and PC2 explains a further 25%. A further six PCs are required to explain 95% of the variation (Fig. IV.4C, S.IV.4). Different genes contribute to each of the PCs, with no single gene explaining a substantial portion of the variation (Fig. IV.4B). Our MANOVA confirms species-specific gene expression (df=2, $Pillai$=1.8, F=147.4, p<0.001).

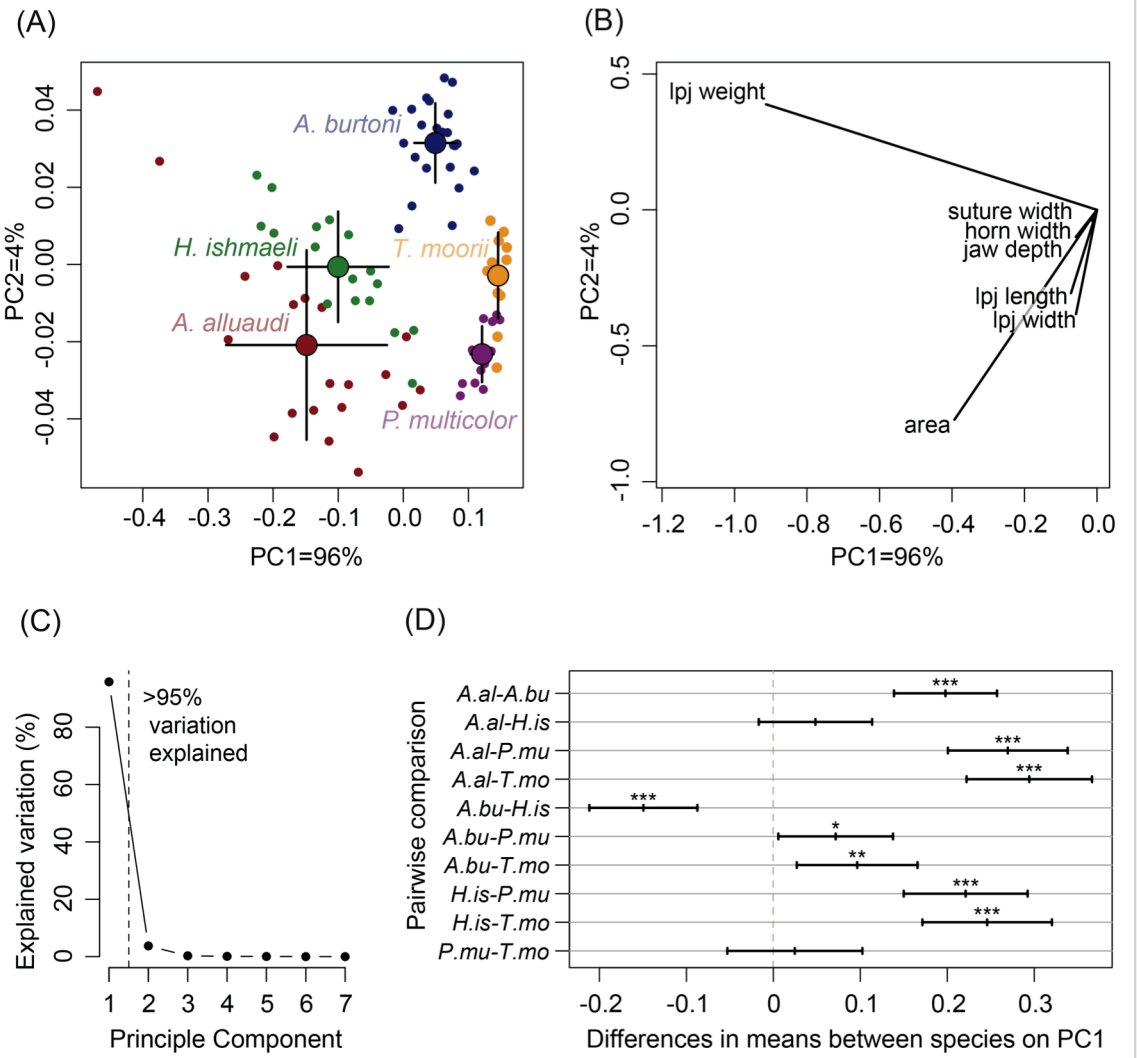


Fig. IV.2 The LPJs of 5 focal cichlid species occupy different morphospace, as per linear morphometric measurements. Principal component analyses were prepared for all linear LPJ measurements (absolute measurements divided by standard length). (A) The 5 focal cichlid species can be separated based on PC1 and PC2. The crosshairs show mean and standard deviations within each species. (B) PC loadings indicate that Jaw Weight is the most significant contributor to PC1. (C) PC1 explains >95% of the variance, with minor contributions from all other PCs. (D) Significant differences were detected for most pairwise comparisons between each species, as per our pairwise Tukey-HSD Test.

Specialist representatives of radiating lineages display lower adaptive phenotypic plasticity than generalist, non-radiating lineages

Similar to our previous experiments (Schneider *et al.* 2014), we confirmed that hard diet (HD) and soft diet (SD) treatments cause *A. alluaudi* to develop divergent LPJ morphologies after a feeding period of 8 months (Fig. S.IV.5). For this species, we observed strong differences in all linear morphometric measurements examined (Fig. S.IV.5, Table S.IV.2).

Interestingly, significant morphological plasticity was also detected in the more slender-jawed species, *A. burtoni,* where LPJ plasticity has not previously been demonstrated (Fig. S.IV.5; Table S.IV.2). For this species, the largest differences were observed in LPJ depth ($p<0.001$) and jaw area ($p<0.01$). Morphological divergence was also demonstrated for *H. ishmaeli*, albeit to a lesser extent than for *A. alluaudi* or *A. burtoni* (Figs S.IV.5, Table S.IV.2). Significant differences were observed in LPJ area and depth, centroid size, weight, horn width and suture width ($p<0.05$). No significant differences were observed in *T. moorii* or *P. multicolor* (Fig. S.IV.5, Table S.IV.2). However, within species morphological variability across species was comparable (Fig. S.IV.5). As morphological variation related to adaptive phenotypes appears to be mostly loaded on PC 1-2 (where most variation of the afore mentioned plastic species is found), it is likely that variation found in "non-plastic" species reflects non-adaptive plastic responses, mostly in *T. moori* and *P. multicolor*.

We conducted correlation-matrix based PCAs of linear morphometric measurements for each species to explore patterns of diet-induced size variation (Fig. IV.5). For *A. alluaudi*, PC1 explains 97% of the identified variation, and the two diet groups are clearly separable on this axis ($p<0.001$) (Fig. IV.5M-P). For PCAs generated for *A. burtoni* and *H. ishmaeli*, PC1 explains 90% and 98% of the total variation respectively, with both species showing significant divergence between the two diet treatments ($p<0.01$ for *A. burtoni* and $p<0.05$ for *H. ishmaeli*) (Fig. IV.5A-H). In contrast, only 64% and 51% of the variation is explained by PC1 in *P. multicolor* and *T. moorii* respectively, and diet-induced divergence is not statistically significant (Fig. IV.5I-L; Q-T).

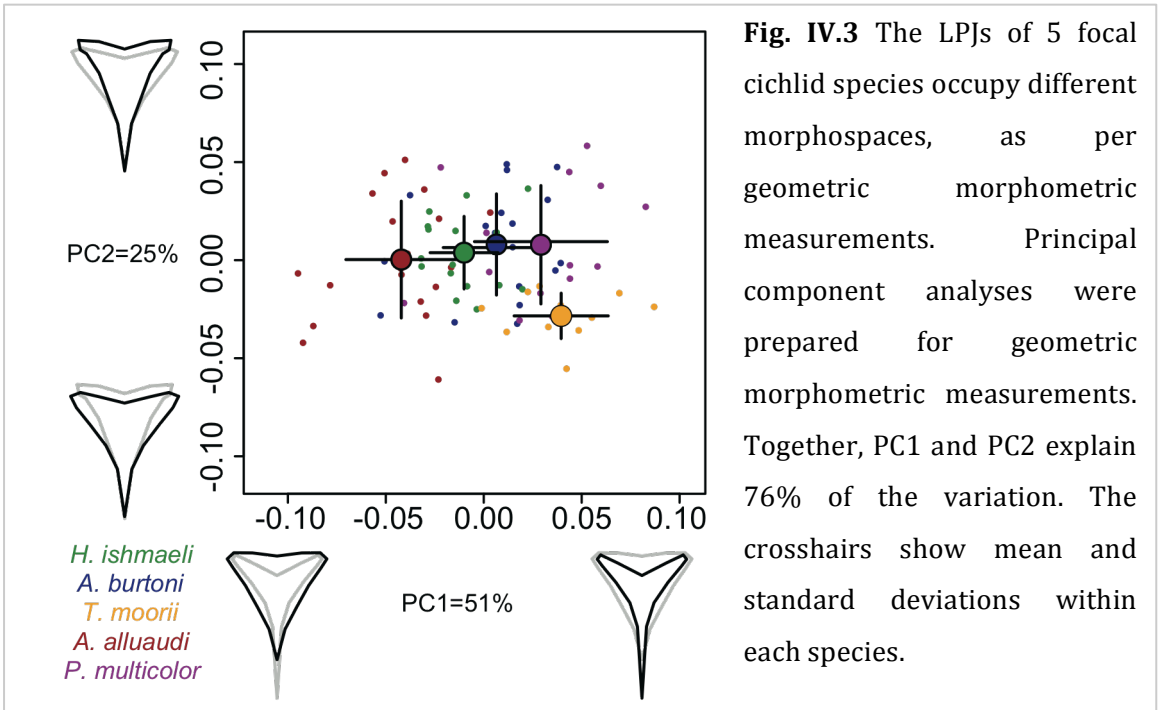


Fig. IV.3 The LPJs of 5 focal cichlid species occupy different morphospaces, as per geometric morphometric measurements. Principal component analyses were prepared for geometric morphometric measurements. Together, PC1 and PC2 explain 76% of the variation. The crosshairs show mean and standard deviations within each species.

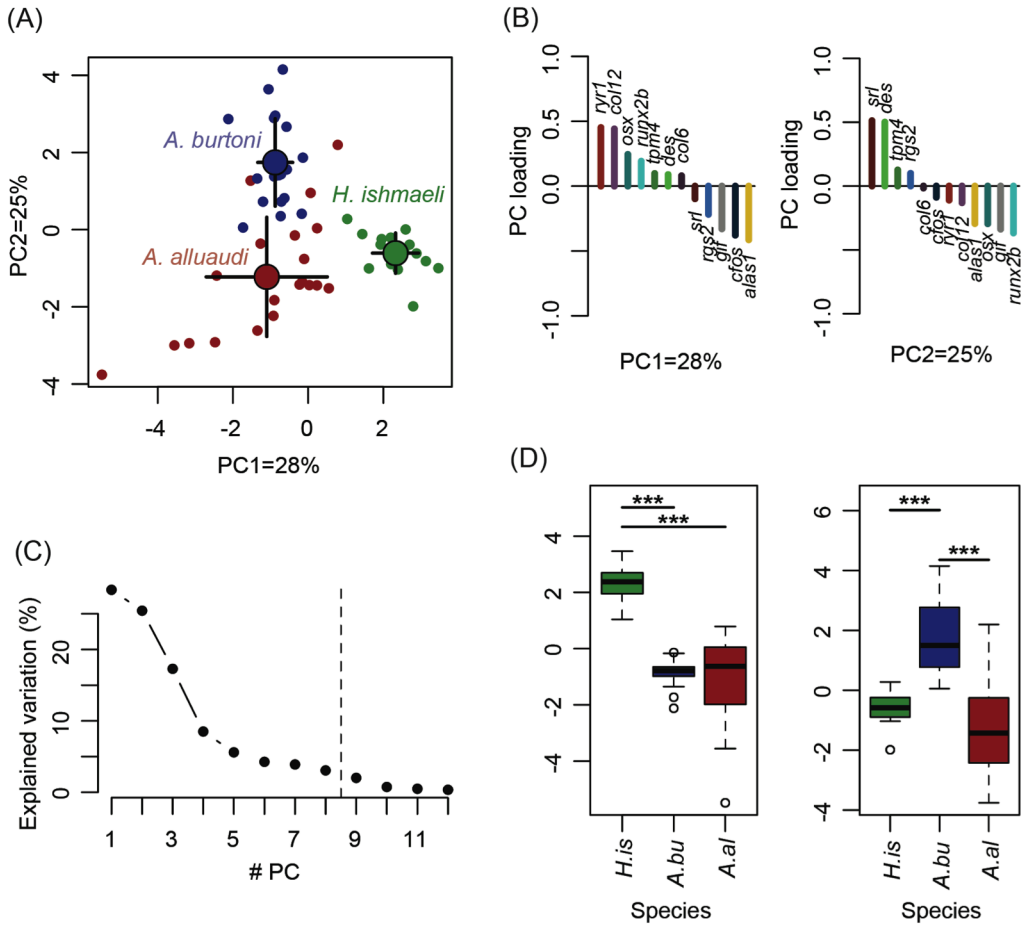
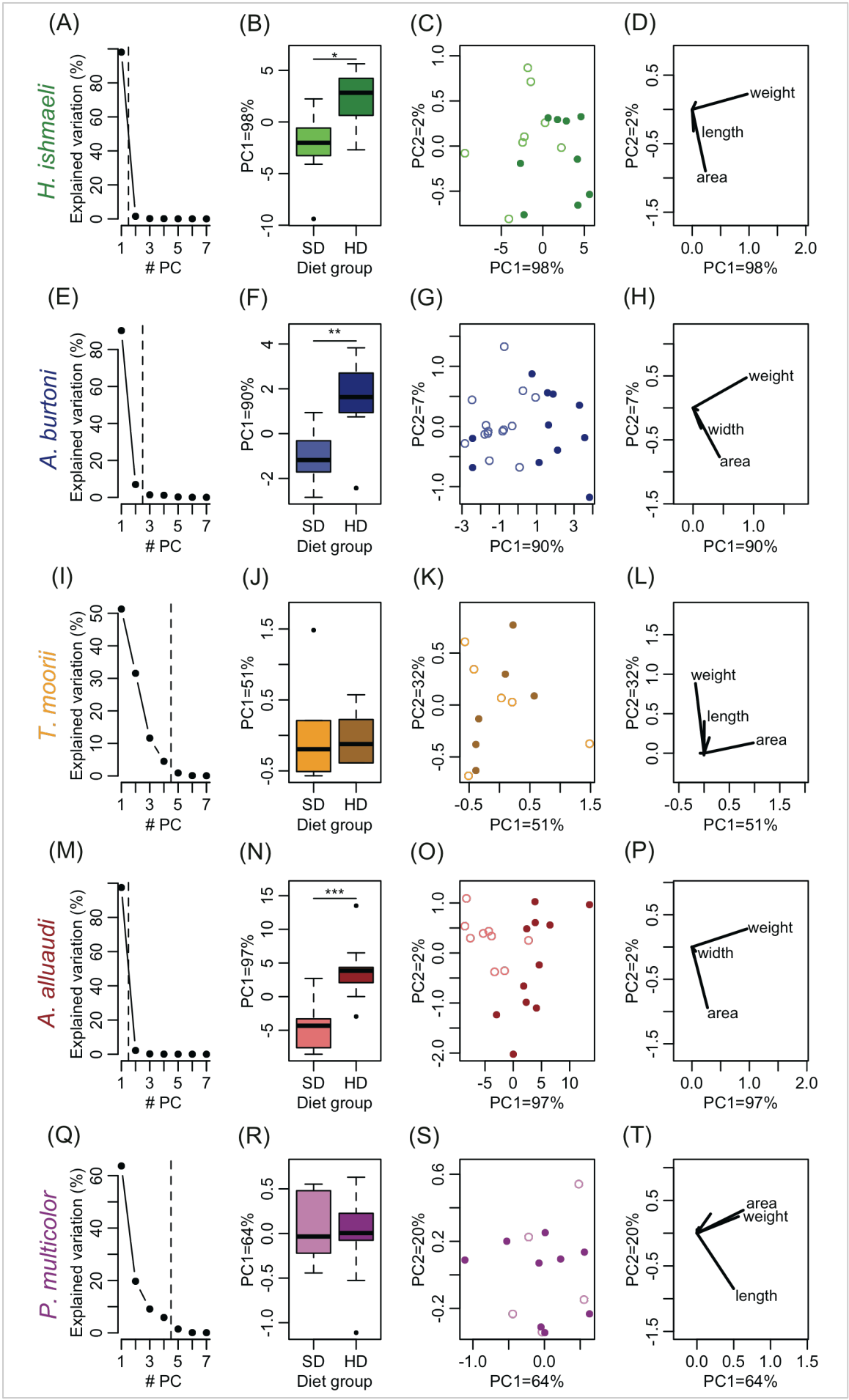


Fig. IV.4 The LPJs of 3 focal cichlid species display distinct gene expression patterns, based on candidate qRT-PCR. Principal component analyses were prepared for gene expression values (normalized to 2 housekeeping genes). (A) The 3 cichlid species used for gene expression analysis can be separated based on PC1 and PC2. The crosshairs show mean and standard deviations within each species. (B) Tukey-HSD tests indicated that each species could be significantly distinguished on either PC1 (left) or PC2 (right). (C) PC1 explains 28% of the variance, and together 9 PCs explain >95% of the variance. (D) The 12 candidate genes contributed variably to PC1 and PC2.

Next page: Fig. IV.5 Influence of diet on linear morphometric measurements in the LPJs of cichlid species from within and outside adaptive radiations. Principal component analyses were produced for all measurements, and (A, E, I, M, Q) percent explained variance for each principal component, (B, F, J, N, R) variation on PC1 for soft diet (SD) and hard diet (HD) treatments, (C, G, K, O, S) variation at PC1 and PC2 and (D, H, L, P, T) loadings of each measurement on PC1-2 are displayed. Significant differentiation at PC1 was detected between HD and SD treatments for (B) *H. ishmaeli* ($p<0.05$), (E) *A. burtoni* ($p<0.01$), (M) *A. alluaudi* ($p<0.001$), however significant differentiation was not detected for (I) *T. moorii* ($p>0.05$), or (Q) *P. multicolor* ($p>0.05$).


(A)
(B)
(C)
(D)
H. ishmaeli
Explained variation (%)
# PC
PC1=98%
Diet group
SD
HD
*
PC2=2%
PC1=98%
weight
length
area
(E)
(F)
(G)
(H)
A. burtoni
PC1=90%
**
PC2=7%
PC1=90%
weight
width
area
(I)
(J)
(K)
(L)
T. moorii
PC1=51%
PC2=32%
PC1=51%
weight
length
area
(M)
(N)
(O)
(P)
A. alluaudi
PC1=97%
***
PC2=2%
PC1=97%
weight
width
area
(Q)
(R)
(S)
(T)
P. multicolor
PC1=64%
PC2=20%
PC1=64%
area
weight
length

Further PCAs were conducted on geometric morphometric measurements for each species to explore patterns of diet-induced shape variation (Fig. S.IV.6, S.IV.7). Hard and soft diets induce significant differences on at least one PC for *A. alluaudi*, *A. burtoni* and *H. ishmaeli* (PCs 2, 1 and 3 respectively; Fig. S.IV.7). No significant shape differences were observed for *P. multicolor* or *T. moorii*, on any of the PCs that cumulatively explain >95% of the variation (Fig. S.IV.6, S.IV.7).

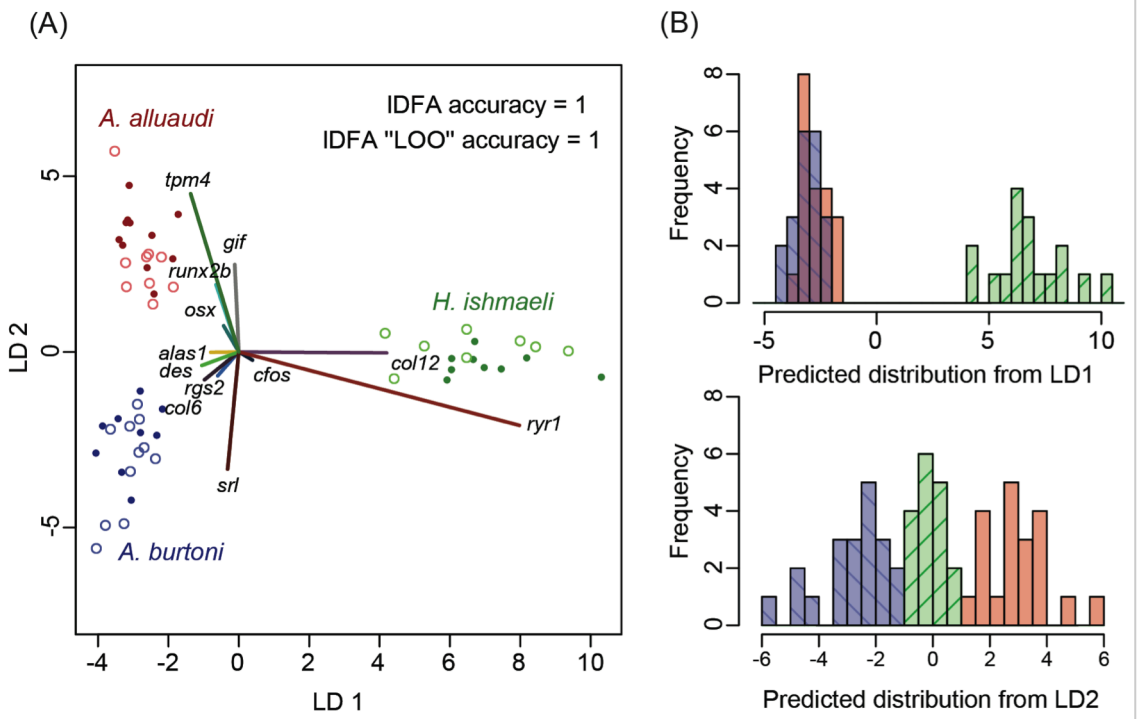


Fig. IV.6 Linear discriminant function analysis shows that gene expression can be used for a reliable predictor of diet for the 3 cichlid species analyzed. Here, diet (SD – HD) was the dependent variable and gene expression (absolute, normalized values) was the independent variable matrix. (A) The 3 focal species could be clearly differentiated and (B) standard as well as Leave One Out (LOO) cross-validation support the reliability of predictors.

Species-specific patterns of gene expression associated with adaptive plasticity in basal vs derived lineages

Expression of 13 candidate genes was examined in the LPJs of HD and SD individuals for *A. alluaudi*, *A. burtoni* and *H. ishmaeli*, using qRT-PCR (Table S.IV.3, Fig. S.IV.8). These were categorized according to functional annotations that include: immediate early genes, heme genes, matrix genes, bone genes, muscle genes and calcium genes (Gunter *et al.* 2013). Few genes showed statistically significant differences in gene expression between SD and HD groups. Among them were *gif* and *alas1* for *A. alluaudi*, and *osx* and *col6* for *H. ishmaeli*.

Moreover, an ANOVA predominantly supported these results, with significant interactions between species and diet for *osx*, *alas1*, *col6* and *col12* (Table IV.1), while marginal significance was detected for *gif*.

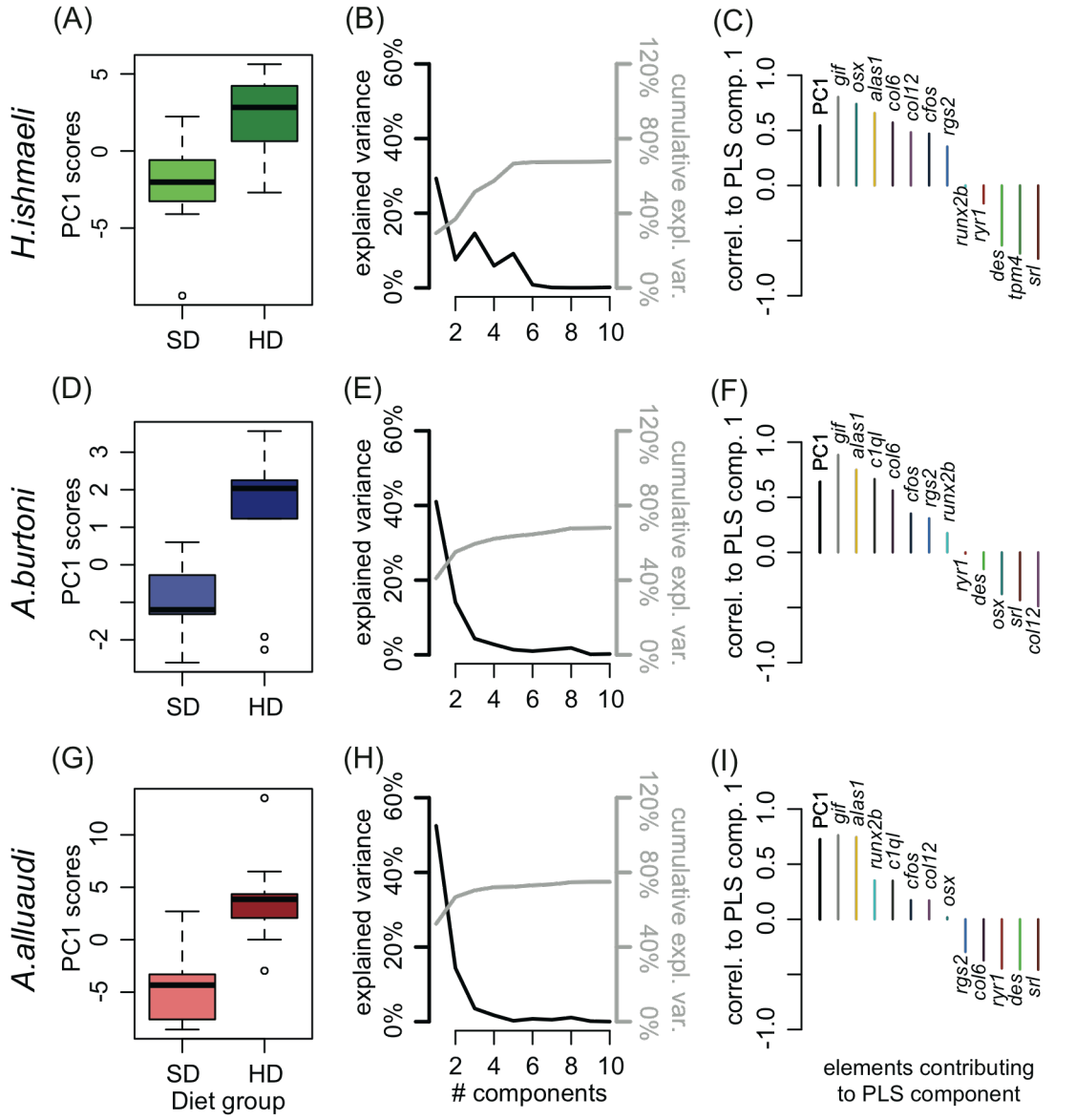


Fig. IV.7 Partial least squares regression shows that diet-induced morphology can be associated with gene expression. Here, the linear dependent variable was the PC1 scores of the PCA on linear morphometric measurements (Fig. IV.2A) and gene expression was used as independent matrix of variables (Fig. IV.8). Plots show (A, D, G) PC1 scores for the 3 species (NB, diet is not included in this analysis), (B, E, H) the amount of variation explained by each of the PLSR components (per variable: left; and cumulative: right) and (C, F, I) the correlation between the expression of each gene and the response with the first PLSR component.

Species-wise PCA on gene expression revealed that in *H. ishmaeli*, *A. burtoni* and *A. alluaudi* the scores of PC2, PC1 and PC3 & PC4 were significantly different between diet groups, respectively ($p<0.05$, Fig. S.IV.9). Combined with significant MANOVAs using the considered PCs as dependent and diet as independent variable (*df*=1,1,1 *Pillai*=0.8,0.71,0.8, *F*=5.14,4.22,8.86, *p*=0.013,0.014,0.0006; for *H. ishmaeli*, *A. burtoni* and *A. alluaudi*, respectively), our data suggests that gene expression patterns differ between diet groups, although other factors may affect gene coexpression, particularly in *A. alluaudi* and *H. ishmaeli*, as the scores of their most heavily loaded PCs do not differ significantly between diet groups.

Linear discriminant function analysis was used to evaluate whether a species membership could be predicted based on diet induced gene expression patterns (Fig. IV.6, S.IV.10). LD1 separated *H. ishmaeli* from *A. alluaudi* and *A. burtoni*, while LD2 separated all three species (Fig. IV.6A). Leave One Out (LOO) cross-validation indicated a high accuracy, showing that all 3 species could be separated on the basis of gene expression (Fig. IV.6B). Interestingly, genes relating to muscle structure and function on both LD 1 and LD 2 contributed strongest to the predictors (*ryr1*, *srl* and *tpm4*), suggesting that they may dictate the observed species-level differences in the diet-induced plastic response (Fig. S.IV.10).

Gene expression predicts morphology in 3 focal species

To investigate whether we could predict morphology using gene expression, we employed partial least squares regressions (PLSR) (Fig. IV.7). Specifically, this analysis relates the PC1 scores from the PCA on linear morphometric measurements to gene expression values, using within-species (diet comparison) datasets for the respective data types. Our analysis indicated that for *A. alluaudi*, the first component of the analysis explains ~52% of the variance of PC1 (Fig. IV.7A-C), while for *A. burtoni* and *H. ishmaeli* the explanatory power was lower, at 41% and 29% respectively (Fig. IV.7D-I). The genes with consistently high explanatory power for each species are *gif* and *alas1*.

Candidates for genetic assimilation identified in derived lineages

Our analysis of gene expression identified gene expression patterns that are suggestive of genetic assimilation, or more specifically, a reduction in plasticity in derived lineages. First, gene expression patterns were compared to our model of genetic assimilation (Fig. IV.1B; Fig. IV.8C-D). Expression of *gif* and *alas1* was consistent with our hypothesis of genetic assimilation, as both genes were significantly differentially expressed in *A. alluaudi*, and their normalized expression levels in *H. ishmaeli* SD individuals were more similar to HD than for the other two species (Fig. IV.8A). That is, the norm of reaction for these genes was

shallower for *H. ishmaeli* than for *A. alluaudi* or *A. burtoni*. Conversely, *osx* and *col6* showed expression patterns that were the opposite of genetic assimilation (Fig. IV.8B), as they were significantly differentially expressed in *H. ishmaeli*, however they were not differentially expressed in *A. alluaudi* or *A. burtoni*.

To investigate any putative alterations to gene expression networks in our focal species, we conducted hierarchical clustering analyses (Fig. IV.8E-G). As was observed in previous studies (Gunter *et al.* 2013; Schneider *et al.* 2014), genes from similar functional categories were co-expressed in *A. alluaudi*. For example, discrete clusters were generally observed for the matrix genes (*col6* and *col12*), the immediate early genes (*cfos* and *rgs2*) and the bone genes (*osx* and *runx2b*). Also similar to Schneider et al. (2014), muscle-related genes (*tpm4* and *des*) clustered with calcium-related genes (*ryr1* and *srl*). The majority of genes analyzed for *A. burtoni* and *H. ishmaeli* displayed patterns of co-expression that were similar to *A. alluaudi*, however they were not identical (Fig. IV.8E-G). Notably for *A. burtoni*, matrix genes *col6* and *col12* were not co-regulated, but rather *col6* clustered with the muscle/calcium genes, while *col12* clustered with the bone genes (Fig. IV.8F). Additionally, for *H. ishmaeli* the bone genes (*osx* and *runx2b*) clustered within the muscle/calcium genes, while the matrix genes (*col6* and *col12*) formed a discrete cluster (Fig. IV.8E). Moreover, three genes identified as being substantial for our lDFA predictor (*ryr1*, *srl* and *tpm4*) were also consistently co-expressed for all 3 species according to our hierarchical cluster analysis (Fig. IV.7, IV.8E-G).

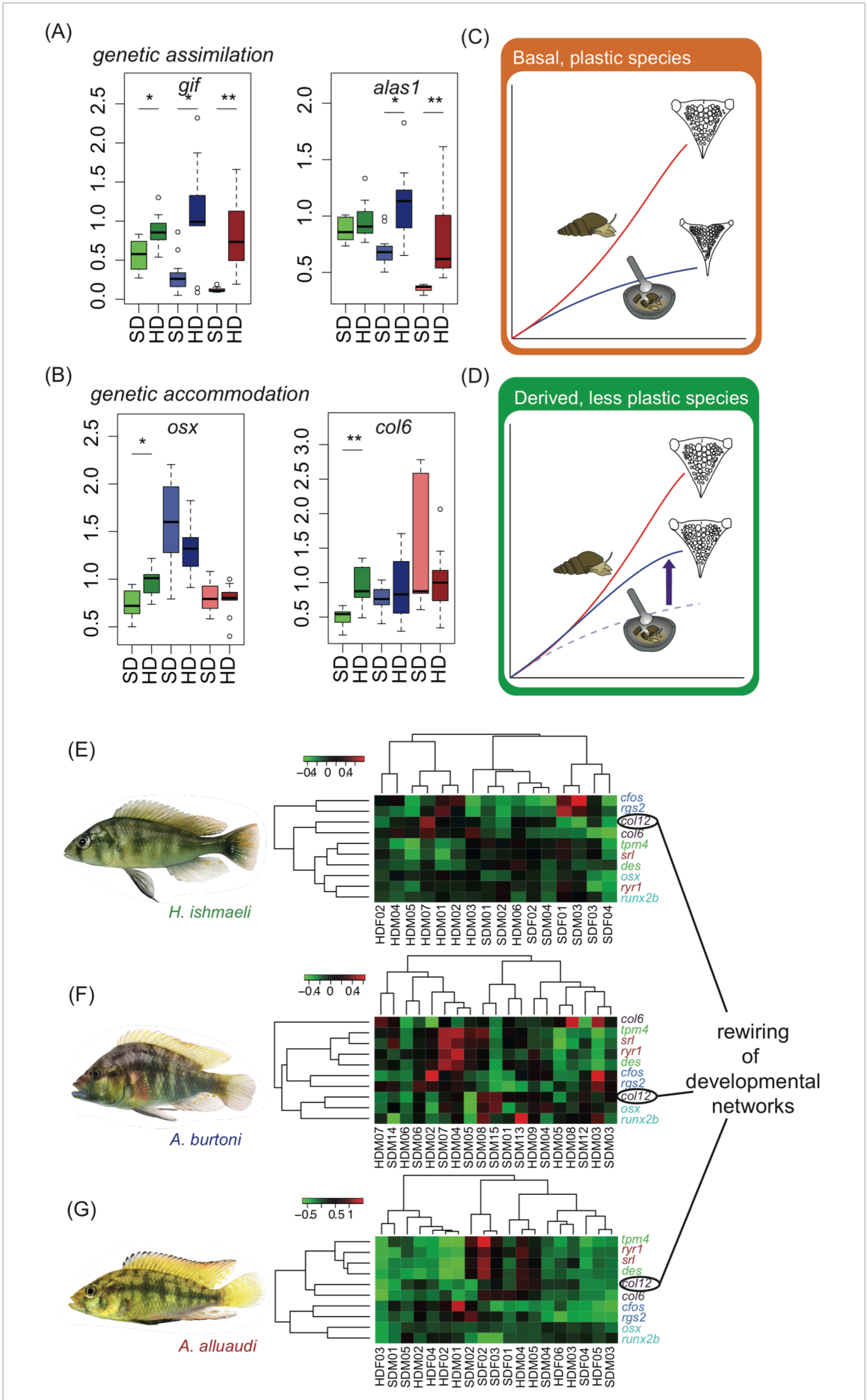
Discussion

Adaptive radiations occur when populations infiltrate empty niches, resulting in a proliferation of new phenotypes, each of which is uniquely adapted to a specific ecological niche. Adaptive radiations have been the subject of many research studies, however few studies have focused on the characteristics of lineages that tend to undergo adaptive radiation (Schluter 2000; West-Eberhard 2003). One proposal is that radiation-prone lineages harbor a high degree of phenotypic plasticity, driving the development of creative, integrated phenotypes that are tuned to the local environment (West-Eberhard 1989; West-Eberhard 2003, 2005). This hypothesis has gained empirical support in recent years through research that examined inducible plasticity in lineages, within a phylogenetic context (Meyer 1987b; Losos *et al.* 2000; Kolbe & Losos 2005; Wund *et al.* 2008; Tebbich *et al.* 2010; Muschick *et al.* 2012; Wund *et al.* 2012). Through comparing the impacts of diet manipulation (hard vs. soft diets) on morphology and gene expression in a range of East African cichlid species, we demonstrate that i) the jaws of each species occupy unique phenotype space ii) this phenotype space can be expanded through feeding on different

diets and iii) that phenotypic plasticity itself may have contributed to the East African cichlid radiations.

It has been proposed that adaptive radiations should, in part, be driven by competitive interactions, spurring on ecological character displacement (Schluter 2000, but see (Stuart & Losos 2013)). Indeed, the LPJs of the cichlid species included in our study predominantly occupy unique phenotype space, most likely underlain by their distinct gene expression patterns. Such differences in LPJ shape tend to arise convergently among cichlids and are correlated differences in trophic ecology (Muschick *et al.* 2012; Elmer *et al.* 2014; Kusche *et al.* 2014), most likely relating to differences in feeding efficiency. Interestingly, the size and shape of *A. alluaudi* and *H. ishmaeli* LPJs differ only subtly and they occupy very similar niches (both are pharyngeal mollusk crushers) in Lake Victoria (Witte 1980; Slootweg 1987). Similar observations of overlapping morphologies and niches have been made in other Lake Victorian cichlid species (Hoogerhoud *et al.* 1982). It is unknown whether these minor differences in LPJ shape imply subtle differences the trophic niches occupied or if they overlap completely. Ecological separation of these two species is plausible in the wild, as they may have a preference for different prey size classes (Slootweg 1987), or occupy slightly different spatial (Abila 2011) or bathymetric environments (Witte 1980). Alternatively, some cichlid species occupy almost indistinguishable trophic niches (Muschick *et al.* 2012), and with at least 13 molluskivorous species in the Mwanza Gulf alone, it is plausible that at least one species pair from this group would also occupy extremely similar trophic niches (Witte 1980).

Next page: Fig. IV.8 Analysis of gene expression and hierarchical cluster analysis on candidate gene expression suggest species specific gene expression by rewiring regulatory networks. Dampened norms of reaction in gene expression found in *gif* and *alas1* (A) suggest events of genetic assimilation in more specialized species (C,D) while steeper slopes in *osx* and *col6* in more specialized species (B) suggest rewiring of regulatory networks via genetic accommodation, potentially to induce altered new adaptive responses. In general, gene expression clusters by functional category for individuals fed a hard or soft diet for (E) *H. ishmaeli*, (F) *A. burtoni* and (G) *A. alluaudi*. HDM = hard diet male, HDF = hard diet female, SDM = soft diet female, SDF = soft diet female. Coloration of gene names refers to functional class: immediate early genes are blue, calcium pathway genes are red, muscle-related genes are green, matrix-related genes are violet, bone-related genes are cyan. Deviations in clustering among species (e.g. *col12*) may indicate a rewiring of the underlying gene regulatory network in more specialized species.


(A)
genetic assimilation
gif
alas1
SD HD SD HD SD HD
(B)
genetic accommodation
osx
col6
(C)
Basal, plastic species
(D)
Derived, less plastic species
(E)
H. ishmaeli
cfos
rgs2
col12
col6
tpm4
srl
des
osx
ryr1
runx2b
(F)
A. burtoni
(G)
A. alluaudi
rewiring
of
developmental
networks

In addition to the observed species-level differences between LPJ morphologies, phenotype space occupied by three of our five examined species was expanded through diet-induced phenotypic plasticity on PC1 and PC2, which reflect the adaptive response in *A. alluaudi*. Expanding the expressed phenotype to match available food resources is likely to confer a competitive advantage in terms of responding to seasonal and spatial variation in food availability, albeit this was not explicitly tested in our study. Indeed, similar studies have shown a putative advantage to adaptive phenotypic plasticity such as predator avoidance phenotypes in tadpoles (Pfennig & Murphy 2000; Pfennig *et al.* 2002), drought tolerance in plants (Sultan *et al.* 2009), and trophic polymorphisms in goodeid fish (Grudzien & Turner 1984), arctic charr (Nordeng 1983) and stickleback (Day *et al.* 1994; Wund *et al.* 2008; Svanbäck & Schluter 2012). Although it is clear that phenotypic plasticity may increase niche spread, the issue of whether adaptive plasticity promotes or hinders evolution remains a topic of debate. Few empirical studies have rigorously tested the influence of plasticity on evolution (Hendry 2015), and those that have, both support (Schaum & Collins 2014) and refute (Torres‐Dowdall *et al.* 2012) this hypothesis. It is likely that both are true under different ecological and evolutionary scenarios, which lead Hendry (2015) to stress that future research should rather tease apart the conditions under which adaptive plasticity promotes and constrains evolution.

Our dataset supports the hypothesis that ancestral plasticity may have been canalized in adaptive radiations. This is indicated by the relatively low inducible LPJ plasticity in *H. ishmaeli*, a snail cracking specialist from the Lake Victorian adaptive radiation (Greenwood 1964; Hoogerhoud 1986c; Slootweg 1987), compared to two relatively basal, generalist species, *A. alluaudi* and *A. burtoni* (Salzburger *et al.* 2005; Seehausen 2013). In terms of both size and shape, we observed the highest level of adaptive plasticity the LPJ of *A. alluaudi* (i.e. the highest level of statistical significance for the highest number of measured variables), an intermediate level in *A. burtoni*, a slightly lower level in *H. ishmaeli*, and no significant level for *T. moorii* or *P. multicolor*. Adaptive plasticity is well documented in *A. alluaudi* (Greenwood 1984; Hoogerhoud 1986b; Hoogerhoud 1986c; Huysseune *et al.* 1994; Huysseune 1995; Smits *et al.* 1996; Smits 1996), however diet-induced LPJ plasticity was not previously demonstrated in *A. burtoni* to the best of our knowledge. This observation implies that LPJ plasticity may be present among a considerably higher number riverine, generalist cichlid species than has previously been demonstrated, albeit confirmation of this would require the inclusion of further generalist species in future studies. The examination of additional basal representatives is of particular importance, as although our selected generalist species are believed to display traits similar to the common ancestor, it is not possible to test this hypothesis – as is the case for all studies that investigate plasticity in extant lineages. The observation of higher plasticity in the LPJ of *A. burtoni* than *H. ishmaeli*

is of particular interest, as *H. ishmaeli* has a considerably more molariform jaw than *A. burtoni* (Hoogerhoud 1986c), which is better suited to cracking snails and should provide a better starting point for the development of plasticity (Slootweg 1987).

The observed lack of significant inducible plasticity in the LPJ of a species from the Lake Tanganyika adaptive radiation, *T. moorii*, is in line with the flexible stem hypothesis as it displays a highly specialized feeding morphology and behavior as an algal scraper (Yamaoka 1983; Sturmbauer *et al.* 1992; Sturmbauer & Meyer 1992; Sturmbauer *et al.* 2003). Our data suggest that genetic assimilation is at an advanced state for this species – consistent with the age of the radiations as the Lake Tanganyika radiations considerably predate the re-filling of Lake Victoria and thus its major radiation. In spite of this we did observe HD individuals taking the snails into their mouths during the experimental period (per obs.). By experiencing the stimulus that would have led to an adaptive phenotype in a presumable plastic ancestor, rudiments of the original plasticity network that are still present in *T. moori* may still have responded and produced a now degenerated non-adaptive phenotype. These findings are in line with theoretical work recently published in which at intermediate stages of genetic assimilation, reduced adaptive plastic - but increased non-adaptive plastic responses are expected.

In a seemingly contradictory result, we also did not detect significant inducible adaptive phenotypic plasticity in the LPJ of *P. multicolor*, which is basal to *A. alluaudi* (Salzburger *et al.* 2005). However one factor that we had not previously considered is that *P. multicolor* is a dwarf species, which is considerably smaller than the other investigated species, with jaw muscles that are likely to be too small to efficiently crack snails. Indeed, its' mean SL at the termination of the experiment was 44.6 mm – considerably less than the SL at which *A. alluaudi* first displayed significant LPJ plasticity (55 – 60mm) in our previous developmental study on *A. alluaudi* (Schneider *et al.* 2014). Moreover, the ecological niche inhabited by *P. multicolor* may plausibly have selected for plasticity in other traits, such as the gills (Chapman *et al.* 2001; Crispo & Chapman 2008; Crispo & Chapman 2010), which are known to increase in surface area for this species, when raised in a hypoxic environment. This increase in gill size concomitantly reduces the space available to the LPJ, an adjacent structure (Binning *et al.* 2010).

Our observation of higher plasticity in the more basal lineages provides further support that the cichlid radiations have been initiated by flexible stems (Muschick *et al.* 2011; Parsons *et al.* 2016), as was suggested for other adaptive radiations, (West-Eberhard 2003; Wund *et al.* 2008; Tebbich *et al.* 2010). Although it is clear that selection on standing genetic variation is important in cichlid radiations (Brawand *et al.* 2014) - an important factor identified in leading models of adaptive radiation (Schluter 2000) - our research characterizes further sources of phenotypic flexibility in stem lineages (West-Eberhard

2003). This is of particular importance for East African cichlid radiations that involve the parallel evolution of multitudinous phenotypes (Fryer & Iles 1972; Liem 1974; Witte 1980), which would have been seeded by hyperflexible stem lineages (West-Eberhard 2003), rather than the parallel, binary flexibility displayed by the other well studied radiations such as stickleback (Rundle *et al.* 2000). It is likely that adaptation to these multitudinous, specialized trophic niches has involved concomitant changes in learning and behavior on top of or instead of morphology (Meyer 1986, 1987a), as behavioral changes have the potential to directly influence morphology through processes such as bone modeling and remodeling (Currey 2002). Although a behavioral component was not incorporated into our study design, we observed that species such as *A. alluaudi*, *A. burtoni* and *H. ishmaeli* received the offered diets more enthusiastically than the other species, particularly *T. moorii* (pers. obs.).

On top of interpreting our results in light of the flexible stem hypothesis, we are also able to make a broader comparison of generalist and specialist lineages. Our results show that generalist (basal) lineages (*A. alluaudi* and *A. burtoni*) display a higher level of plasticity than the more derived, specialist lineages (*H. ishmaeli* and *T. moorii*), which is in line with previous studies (Sultan *et al.* 2009; Svanbäck & Schluter 2012), and with the niche variation hypothesis (Van Valen 1965). Our results support the hypothesis that stem lineages leading to adaptive radiations contain generalist ancestral lineages, in line with previous studies (Nosil 2002) and models (Van Tienderen 1997; Stomp *et al.* 2008), however there are many notable exceptions, where specialists give rise to adaptive radiations (Schluter 2000). Based on models, phenotypic plasticity is likely to be lower in specialist species (Van Tienderen 1997), and therefore other factors such as standing genetic variation may play a relatively higher role in these cases. However without explicit testing the importance of phenotypic plasticity, it cannot be ruled out. This is especially true for radiations seeded by the colonization of novel, stressful environments such as islands, where induced plasticity could unlock cryptic genetic variation (Rutherford & Lindquist 1998; Queitsch *et al.* 2002), which would not otherwise have been exposed to natural selection (Schneider & Meyer 2016).

It is interesting to note that while *A. alluaudi* and *A. burtoni* can be broadly described as generalist species, there are nuanced differences between the diets of populations inhabiting different geographic locations. For example, *A. burtoni* populations show a lake-stream divergence in dietary preference, with lake fish showing a plant/algae and zooplankton-biased diet, and stream fish eating more snails, insects and plant seeds (Theis *et al.* 2014). The diets of *A. alluaudi* populations also differ significantly, with Lake Victorian populations being strictly molluskivorous (Bouton *et al.* 1997), and riverine and satellite lake species feeding on insects, algae, fish and mollusks, whose proportions vary spatially

(Mbabazi *et al.* 2004; Abila *et al.* 2008; Binning & Chapman 2008) and seasonally (Binning *et al.* 2009). This is particularly interesting when one considers that snails are not the preferred diet of *A. alluaudi* – where they have a choice they tend to select an insect-based diet (Slootweg 1987; Slootweg *et al.* 1994). Bouton *et al*. (1997) show that seasonal fluctuations in food availability in Lake Victoria can alter dietary composition for different species, and suggest that competition determines which species gain access to preferred food items. Given that *A. alluaudi* is restricted to non-preferred food items in Lake Victoria, it appears that they do not compete well with a proportion of Lake Victoria's endemic cichlid species, which may have gained a competitive advantage. In this case, LPJ plasticity is likely to benefit *A. alluaudi* by expanding its geographic distribution, but potentially comes at the expense of access to their preferred food items in regions of high competition. It should be noted that a single *A. alluaudi* population (from Lake Victoria) was used in this study, and we are unsure of how its plasticity compares to that of riverine and satellite lake populations. However it seems likely that the other populations are plastic to an equal or greater degree than our selected population, as riverine have been shown to ingest a more variable diet, so contemporary plasticity may be more beneficial.

On top of demonstrating that basal, generalist lineages display higher plasticity than derived, specialist lineages, we have identified two candidate genes that show dampened inducibility in a derived lineage. These genes are excellent candidates for genetic assimilation, and may explain the lower plasticity observed in *H. ishmaeli*. These are *gif* and *alas1*, genes are that are involved in vitamin B12 sequestration and heme biosynthesis, respectively (Sadlon *et al.* 1999; Greibe *et al.* 2012) and have yet unknown functions in cichlid LPJ plasticity. Upstream regulators of these genes may have undergone regulatory evolution that reduced their inducibility, albeit this study did not include methods that could identify the regulatory bases of these expression differences. Our study also identified two genes that showed a pattern opposite to genetic assimilation (*osx* and *col6*), suggesting multidirectional patterns of regulatory evolution. It is difficult to ascertain how many genes would show the different patterns of regulatory evolution, as our study has analyzed 13 potential candidates that were non-randomly selected. Further genome-wide investigations such as RNA-seq would clarify the general trends in regulatory evolution, as has recently been reported for guppies (Ghalambor *et al.* 2015). Importantly, this would enable better analyses of pathways, which would contextualize the interpretation of relative up- and down-regulation of genes. It should be noted that although our study identified a relatively small number of loci that may have undergone genetic assimilation, this is similar to another recent study that identified a single locus (Parsons *et al.* 2016).

Moreover, we detected differences in gene co-expression within regulatory pathways, which may also have been subject to genetic assimilation. For example *col6* was co-

expressed with *col12* in *A. alluaudi* and *H. ishmaeli*, but it was co-expressed with the immediate early genes in *A. burtoni*. These differences in collagen regulation in the LPJ may plausibly have modified the stiffness of the extracellular matrix, leading to altered patterns of bone cell differentiation (McBeath *et al.* 2004; Engler *et al.* 2006) and/or altered mineralization (Wang *et al.* 2012), which may impact the mechanical properties of the bone (Nair *et al.* 2013). Interestingly, *col6*, *col12*, the IER genes are all predicted to contain binding sites for the mechanically-responsive transcription factor AP1 in their promoter regions (Schneider *et al.* 2014). Thus, promoter evolution for genes such as *col6* may have altered their mechanical responsiveness, leading to downstream changes in mineralization or cellular differentiation. Indeed, the role of regulatory evolution in mediating phenotypic plasticity is supported by theoretical (Schlichting & Pigliucci 1993; Espinosa-Soto *et al.* 2011), and empirical evidence (Li *et al.* 2006; Suzuki & Nijhout 2006; Ghalambor *et al.* 2015); but see (Sikkink *et al.* 2014).

Our putative identification of genetically assimilated loci within a cichlid adaptive radiation represents an important step in establishing a mechanism by which plasticity may have promoted their explosive speciation. This adds to other recent studies which have demonstrated that selection on plastic loci may have contributed to adaptive diversification in other species (Ghalambor *et al.* 2015). Our work will enable future investigations that pinpoint the loci associated with reduced plasticity in *H. ishmaeli* and will open the door to comparative studies investigating other genetically assimilated loci in other Lake Victorian cichlid species. As these species display vast differences in LPJ shape, it is likely that different loci would have been fixed through genetic assimilation. Comparisons between various different species with different LPJ sizes and shapes will allow us to differentiate the loci associated with LPJ architecture, vs those that modulate phenotypic plasticity itself. This research will add to previous studies that have identified important loci of adaptive radiations in general (Abzhanov *et al.* 2004; Colosimo *et al.* 2005; Albertson & Kocher 2006; Chan *et al.* 2010), and with the added benefit of providing correlative evidence that phenotypic plasticity may have played a role in the initial accumulation and selection of these mutations.

Cichlid adaptive radiations provide an extremely useful backdrop for studying both evolutionary parallelism of flexible stems, but also for investigating the different stages of genetic assimilation, as the radiations are of different ages. For example the Lake Victorian radiation is ~100 000 years old and the Tanganyikan cichlid radiations are 9-12M years old (Verheyen *et al.* 2003; Koblmüller *et al.* 2008; Elmer *et al.* 2009). Moreover, *T. moorii* of Lake Tanganyika displays a considerably more specialized feeding morphology and behavior than *H. ishmaeli* of Lake Victoria, which may, in part, be the product of the different ages of these radiations (Hoogerhoud 1986c; Sturmbauer *et al.* 1992; Sturmbauer & Meyer

1992; Sturmbauer *et al.* 2003). Thus it seems likely that genetic assimilation of any plastic traits may be complete, as no residual plasticity was detected for this species. The incorporation of lake-lake comparisons into future investigations will allow us to determine whether there is an increasing accumulation of mutations establishing phenotypic differences, associated with an ever-diminishing LPJ plasticity.

One potential objection to the hypothesis that cichlid adaptive radiation is guided by phenotypic plasticity, is that genetic assimilation would produce evolutionary dead ends, and that new plasticity would not arise. On the contrary, West-Eberhard (2003) argues that hypervariable flexible stems, such as those of cichlids have "the potential to rapidly evolve in any variety of new directions, should conditions change." This is backed up by the results of a range of experiments involving pond-based and natural populations, which indicate that adaptive traits fluctuate over extremely short timescales (year to year), with phenotypic plasticity providing the most plausible explanation (Kishe-Machumu *et al.* 2008; van Rijssel & Witte 2013; Rijssel *et al.* 2015). This flexibility may plausibly have assisted cichlids in dominating the East African lakes, which are characterized by rapid, dramatic changes including repeated desiccation and re-filling events, alongside major changes in salinity and food resources (Wimberger 1991). Indeed, rapid evolution of increased plasticity has been observed in lineages under strong selection pressure (Nussey *et al.* 2005).

Conclusion

Our investigation has added further evidence that phenotypic plasticity has the potential to expand niche space within adaptive radiations, which may have played a role in their formation. We demonstrate that specialist lineages from within adaptive radiations show reduced phenotypic plasticity in comparison to generalist lineages from outside of adaptive radiations. Furthermore, through comparative gene expression studies, we have identified putative candidates that may have undergone genetic assimilation, providing a potential genetic explanation for the dampened plasticity displayed by the derived lineages. This research has opened the door for future studies that examine the selection of genetically assimilated loci in natural populations, a powerful next step in establishing a role for phenotypic plasticity in adaptive radiation – a much debated hypothesis with limited genetic evidence to date.

Chapter V

Drivers of cichlid color evolution: no evidence for Fisher's run-away theorem on visual systems but further evidence for the importance of the ambient light regime and conspecific color in determining cichlid vision

RALF F. SCHNEIDER*, SINA J. ROMETSCH*, JULÍAN TORRES DOWDALL & AXEL MEYER

In preparation

Abstract

Cichlid fishes are famously known for their outstanding diversity in body coloration. This diversity is found across species, but also within species, e.g. across development and between sexes (sexual dimorphism in body coloration). Cichlids possess seven cone opsin genes and by expressing a subset of these, natural and sexual selection tuned their visual sensitivity in many cases correspondingly to their ecological requirements. It was proposed that sexually dimorphic body coloration patterns might be driven by Fisher's run-away sexual selection and thus potentially also reflected in sexually dimorphic visual systems. We examined the opsin expression patterns of all opsins in ten cichlid species, from the Afrotropics and Neotropics. No difference in opsin expression between sexes was detected, however rod opsin expression varied across species, potentially reflecting cichlids' different habitat light brightness levels. By integrating lens cutoffs and cone opsin absorption curves, effective retina sensitivities were calculated for African species and compared to estimated fish body reflectance patterns as they would be observable in their corresponding habitats. We find that sexually dimorphic species with divergent reflectance peaks (sexually 'dichromatic') typically express cone opsins covering a wider wavelength range than those species differing only in the height or width of a single reflective peak. Generally, we confirm that cone opsin expression patterns appear to be influenced by ambient light properties, but also conspecific reflectance patterns are crucially important. No evidence for run-away sexual selection having major influence on the evolution of cichlid coloration could be found in this study as no sexual dimorphism in the visual systems was detected. Nonetheless, run-away selection may still affect cichlid coloration and sexes' visual systems may just not have diverged, e.g. because a link of opsin expression and sex did not yet evolve in these fish.

Introduction

Cichlid fishes are evolutionary model organisms for studying the process of speciation, mostly due to their exceptionally high diversity and fast rate of phenotypic divergence (Kocher 2004; Genner & Turner 2005; Henning & Meyer 2014). Cichlids famously vary in morphological traits, such as body coloration and their trophic morphology, but also in behavioral traits such as parental care (Seehausen 1997; Wisenden *et al.* 2015). This remarkable diversity is likely to be the result of both, natural and sexual selection (Kocher 2004; Seehausen 2006). While natural selection is thought to have been the driving force for the evolution of a wide range of foraging mode associated morphologies (Keenleyside 1991), sexual selection was putatively instrumental during the evolution of an outstanding diversity of body coloration (Deutsch 1997; Allender *et al.* 2003). This color diversity is found across species, but also on an intraspecific level, as sexual dimorphism, with males showing different and often brighter colors and higher color intensities than females (Seehausen *et al.* 1999). As males typically exhibit more conspicuous body coloration patterns, female preference was suggested as the main driver of sexual dimorphism in cichlid fishes' body colorations (Seehausen *et al.* 1997).

The evolution of conspicuous coloration patterns in males in response to female preference is a frequently reported phenomenon in nature (Waitt *et al.* 2003; Sandkam *et al.* 2015). Examples include female preference for hormonally regulated reddening of facial skin in rhesus macaques (Waitt *et al.* 2003), dewlap color in *Anolis* lizards (Sigmund 1983), orange-red bill color in zebra finches (Burley & Coopersmith 1987), orange-reddish blotches in Trinidadian guppies (Kodric-Brown 1985) and red belly color in sticklebacks (Bakker & Mundwiler 1994). Several models aim to explain how sexual selection may facilitate the evolution of exaggerated or conspicuous traits, such as the 'handicap principle' (Zahavi & Zahavi 1999), the 'sensory bias' hypothesis or 'run-away' sexual selection (Fisher 1930; Endler 1992; Boughman 2002). The latter has been suggested to be involved in establishing the color diversity in cichlids, by facilitating the exaggeration of body color dimorphism (Sabbah *et al.* 2010). Fisher's 'run-away' model predicts that arbitrary characteristics in one sex (typically males) are preferentially chosen by the choosy sex (typically females). Males' characteristics then coevolved with female preference. Females with higher preference for exaggerated traits mate with males showing more extreme phenotypes. Thus male trait and females preference can become linked. Therefore, Fisherian run-away selection would lead to sexual dimorphism in the preferred characteristic (with the chosen sex having exaggerated phenotypes), but may also affect the sensory apparatus used to detect it (males). Considering the color diversity - but also the diversity found in visual systems (Carleton *et al.* 2016) across cichlids, it is plausible that

run-away selection may have contributed to this diversity by putting female visual properties in a different selective regime, not only driven by natural selection but also sexual selection, leading to divergent visual sensitivities between sexes (Kelber & Osorio 2010). Naturally, alternative drivers of body color and visual system diversity were suggested before, such as habitat type, feeding behavior and other forms of sexual selection (Seehausen *et al.* 2008; Smith *et al.* 2011). To find evidence for the potential importance of Fisherian run-away sexual selection during the evolution of cichlid diversity in body color and visual systems, it is of importance to also consider other central factors with presumed influence on body coloration and visual systems.

Organisms' visual systems are usually well adapted to the dominant ambient light conditions of that organism (Cronin *et al.* 2014; Cronin & Porter 2014). For instance, most birds have tetrachromatic vision presumably as they are descendants of diurnal active dinosaurs, which could utilize the full bright light spectrum of the sun (Finger & Burkhardt 1994). In contrast, most mammals evolved from nocturnal animals, which had to reduce color resolution in favor of higher light sensitivity in the very dim light conditions at night, and are thus still mostly dichromates (Jacobs 1993). For fish, not only the light brightness but also its composition can vary considerably (Lythgoe 1988). Light generally attenuates when transmitted through water and thus habitats get darker with increasing depth. In addition, clear water absorbs light of long wavelengths ('red' light) and very short wavelengths ('UV' light) particularly fast (Lythgoe 1988; Anthes *et al.* 2016). Thus, with increasing depth the ambient light spectrum becomes spectrally narrower and short to middle-long light wavelengths dominate (blue-green light). Furthermore, algae and particles suspended in the water change these transmission properties. Dissolved soil particles, for instance, absorb particularly short and middle wavelengths (blue-green light), leading to red-biased ambient light conditions for fish (Cronin *et al.* 2014).

Cichlid fishes live in a variety of water body types, such as small and large lakes, rivers or swamps, and thus their habitat-specific ambient light conditions can be very divergent. To investigate the visual system of cichlids it is crucial to have knowledge about their respective environmental light conditions, ideally by conducting spectral measurements in the field. If these are not feasible, natural light conditions can be estimated if the fish's preferred water depth is known, as well as the light transmission properties of the water. For example, Lake Malawi and Lake Tanganyika, which are inhabited by over 1,300 cichlid species, contain quite clear water (relatively few dissolved soil particles or algae). Thus sunlight is transmitted to great depth and with increasing depth, the ambient light becomes increasingly dominated by blue-green light (Dalton *et al.* 2010; Sabbah *et al.* 2011). In contrast, Lake Victoria with its around 600 cichlid species partially contains more murky water with a higher proportion of dissolved soil particles (depending on site and season).

With increasing water depth, it gets darker much faster than in the other two Great Lakes, and the ambient light is dominated by yellow and red light (Seehausen *et al.* 2008; Maan *et al.* 2010).

Light perception in vertebrates occurs in the eye's retina and, more precisely, in photoreceptors containing visual pigments. However, to reach the retina, light first has to pass through various media with specific transmission properties. Previous research showed that particularly eye lenses can modify incoming light spectra, usually by filtering out light of very short wavelengths (UV), which otherwise may damage the retina or compromise visual resolution as it scatters rapidly in the water column (Douglas & Marshall 1999). When light penetrates the photoreceptors of the retina it can excite visual pigments consisting of an opsin protein that is bound to a light-sensitive chromophore (Bowmaker 1990; Yokoyama 2000; Hofmann *et al.* 2009). There are two types of chromophores (A_1 and A_2) and each type has influence on the maximum absorption wavelength (λ_{max}) of the associated opsin (Wald 1961; Hárosi 1994; Carleton 2009). Once the chromophore is exited, a signal cascade is initiated that triggers a neural response.

In vertebrates, two classes of photoreceptor cells are distributed across the retina: rods and cones. Rods are involved in dim-light perception and have only one type of opsin (rhodopsin 1: *rho*). Cone photoreceptors are responsible for color vision, which is achieved by using multiple opsin proteins that are sensitive to different parts of the light spectrum (Wald 1968). Whereas humans have only three cone opsin genes, cichlids have a total of seven unique cone opsin genes (Hofmann *et al.* 2009). Each opsin gene can be characterized by the wavelength of its maximum absorption: *sws1* (UV), *sws2b* (violet), *sws2a* (blue), *rh2b* (green-blue), *rh2a-alpha* and *rh2a-beta* (green), and *lws* (red) (Carleton 2009). In most fishes, including cichlids, cones exist as single and double cones (Bowmaker *et al.* 1997). Single cones are smaller in size and contain visual pigments that are sensitive to rather short wavelengths (SWS opsins). Double cones consist of two cones joined together and are responsible for absorption at longer wavelengths (RH2 & LWS opsins) (Levine & MacNichol 1979). In spite of having seven different opsin genes, often only three opsin genes are expressed in cichlids at a time, resulting in a trichromatic visual system (Levine & MacNichol 1979; Fernald 1981, 1984; Carleton & Kocher 2001; Carleton 2009). As species differ in the set of opsin genes they express, their visual sensitivities can be highly variable (Carleton 2009). When opsin genes are not being expressed for considerable evolutionary periods, these may undergo pseudogenization, as was already reported for a number of cone opsins in different lineages (e.g. Weadick *et al.* 2012; Escobar‐Camacho *et al.* 2016).

The factors influencing visual properties of cichlid species have been the subject of numerous investigations (e.g. Levine & MacNichol 1979; Fernald 1981; Spady *et al.* 2006; Hofmann *et al.* 2009; Weadick *et al.* 2012). It is generally thought that each species

expresses an opsin composition particularly adaptive to their respective environment (Seehausen *et al.* 1997; Parry *et al.* 2005), and the food a species is specialized on (Seehausen *et al.* 2008; Hofmann *et al.* 2009; O'Quin *et al.* 2010; Sabbah *et al.* 2010), however both factors cannot fully explain measured opsin expression patterns (Jordan *et al.* 2006). It is likely that the perception of conspecific body coloration is of importance for cichlids' vision (e.g. for mate choice) and thus opsin expression may be tuned to perceive these colors particularly well (Parry *et al.* 2005; Maan & Seehausen 2010). Importantly, it has been hypothesized that sexual dimorphism, such as male nuptial coloration, is associated with differences in opsin gene expression between sexes: as nuptial coloration may be an indicator of male quality, evolution may tune female color vision to ensure optimal perception of mate coloration differences and facilitate mate selection (Sabbah *et al.* 2010). This may have led to a coevolution of male nuptial coloration and sexual dimorphism in color vision in response to sexual selection as predicted by Fisher's theory of run-away sexual selection (Fisher 1930; Sabbah *et al.* 2010; Bloch 2015).

This study aims to investigate whether conspecific males and females of selected cichlid species have dimorphic opsin expression patterns and whether this can be found more often in sexually dimorphic species compared to monomorphic species. Additionally, we explore further likely determinants of cichlid opsin expression by integrating species-specific environmental data, spectral body coloration measurements and opsin expression data of sexually dimorphic and monomorphic species of both sexes. Specifically, we investigate whether sexually dimorphic species use different opsins arrangements than sexually monomorphic species.

Materials and Methods

To investigate likely drivers of cichlid vision, comprehensive datasets comprising physiological and ecological data were collected for ten selected cichlid species, five from the Neotropics and five from the Afrotropics. More specifically, for each species we obtained sex-specific opsin expression levels. Furthermore, for the five investigated African cichlid species the following additional data were gathered, as their opsin expression proved to be more diverse than in the Neotropical species, either through measurements conducted in this study or by extracting the data from previously published papers: λ_{max} values for each opsin, lens absorption curves, sex-specific body-reflectance spectra, preferred habitat water properties and the water depth in which the species is most common. The following sections give details about the selected cichlid species, the data collection procedures and their analyses.

Species selection and maintenance

Cichlid species were selected to probe both the African as well as the Neotropical cichlid radiations. Generally, selected species include representatives with varying levels of sexual dimorphism in body coloration or brightness. Furthermore, selected cichlids represent different habitats, including clear vs. murky waters, shallow vs. deeper waters, temporally stabile vs. labile habitats. Using these criteria, the following species were chosen (Table V.1 for more details): *Hypsophrys nematopus, Hypsophrys nicaraguensis*, *Apistogramma agassizii*, *Apistogramma barlowi*, *Apistogramma ortegai*, *Pundamilia nyererei, Labidochromis caeruleus*, *Pseudotropheus lombardoi, Melanochromis auratus* and *Melanochromis johannii*. Used fish were captivity-bred and raised under artificial lightning until sexually mature (but typically not far beyond). Fish were euthanized by an overdose of Tricaine (MS222) in the early afternoon. For each fish both eyes were dissected and the lens was separated from the remaining eye. The lens-less eye was transferred to RNA*later* (Sigma) and stored at -20°C while the lens was transferred to PBS for short-term storage (<1h).

Table V.1 Details on examined species habitat.

Species	Conti-nent	Specific origin	Water Clarity*	Habitat depth*	Depth in analysis
H. nematopus	Neotropics	Nicara-gua & Costa Rica	Depending on specific habitat, relatively clear to very murky (Great Nicaraguan Lakes)	1-5m	NA
H. nicaraguensis				1-5m	
A. agassizii		Ama-zonas drainage rivers	Clear to very clear.	<1m	
A. barlowi				<0.5m	
A. ortegai					
P. nyererei	Afrotropics	Lake Victoria	Murky, brownish	3-8m	5m
L. caeruleus		Lake Malawi	Clear	10-30m, mostly 25m	25m
P. lombardoi				6-30m, mostly 10m	10m
M. auratus				3-8m	6m
M. johannii				3-8m	6m

* References: Konings 1990; Römer 1998; Maan *et al.* 2006; Römer & Hahn 2008; Britzke *et al.* 2014

1. Determination of opsin expression using qRT-PCR

RNA was extracted from whole lens-less eyes using an RNeasyMini Kit (QIAGEN) and concentration and purity were determined using a Colibri microvolume spectrometer (Titertek Berthold). Equal amounts of extracted RNA (750 ng) were used for first-strand cDNA synthesis (GoScript Reverse Transcription System, Promega). Quantitative PCR (qPCR) was performed on a CFX96 Real-Time System (Bio-Rad; using a GoTaq qPCR mix, Promega). For African and Neotropical cichlids separately, a single primer pair was designed for each opsin gene to amplify its target opsin across specimens of the whole clade (*rh2a* primers were designed to amplify both *rh2aα* and *rh2aβ*). Nonetheless, primer pairs were tested for each species separately and an efficiency value was calculated for each opsin and species combination (following Carleton & Kocher 2001). qPCR temperatures were adjusted for optimal efficiency, resulting in efficiency values between 0.762 and 1.085. For *P. nyererei rh2b* and *Apistogramma* species *sws1*, no reliable efficiency could be determined as the expression was too low (we assume an efficiency of 1.0 for downstream analyses), which is in accordance with previously published data (Escobar‐Camacho *et al.* 2016) (Table S.V.1). Primer sequences are given in Supplementary Table S.V.2.

The relative opsin gene expression was calculated following Fuller *et al.* (2004) and Carleton and Kocher (2001):

$$T_{i,g} = 1/\left(\left(1 + E_g\right)^{Ct_{i,g}}\right)$$

Where $T_{i,g}$ is the absolute expression value of the opsin gene g from individual i. E_g is the species-specific efficiency of g's primer and $Ct_{i,g}$ is the critical cycle number obtained from individual i using opsin primer g. For all six cone opsin primer pairs, individual relative expression values were then calculated by dividing each absolute cone opsin expression value by the sum of all cone opsin expression values. For the rod opsin rhodopsin 1 (*rho*), absolute *rho* expression was divided by the sum of all cone opsin absolute expression values. Averages across all individuals per species of relative opsin expression values were finally calculated and multiplied by 100, to obtain percentages. Pair-wise t-tests for unequal variances were used to test for differences in opsin expression between sexes. P-values were species-wise adjusted for multiple testing using the false-discovery-rate correction (Table S.V.3).

2. Meta-analysis of opsin pseudogene occurrence across cichlids

To investigate if sexual color-dimorphism may be related to opsin pseudogenization rates in cichlids, a meta-analysis was conducted on the published opsin sequences of African cichlids that were obtained from GeneBank. Furthermore, investigated species were scored

as being sexually dimorphic in body coloration or not (Table S.V.4). Species with ambiguous dimorphism were excluded from the analysis, leading to 38 considered species. Differences in the occurrence of pseudogenized opsin genes between the phenotype groups were evaluated using a χ^2- test.

3. Determination of visual properties and conspecific reflectance in African cichlids

To explore further likely determinates of cichlid vision, the properties of each species visual apparatus were determined in more detail, by including the transmission properties of the fishes' eye lenses and the opsins species-specific sensitivity peak wavelength. In addition the reflectance spectrum of each sex of each species in their respective habitat was modeled. As only African cichlids showed explicit inter-specific differences in opsin expression, these analyses were restrained to the five African cichlid species of this study (Fig. V.1).

3.1 Modeling African cichlids' visual sensitivity

Shortly after dissection, lenses from African species were mounted on an OceanOptics PX-2 light source and the transmitted light was measured using an Ocean Optics USB2000-UV-VIS-ES photospectrometer. Each relative lens transmittance spectrum was calculated by dividing the absolute transmittance spectrum by the spectrum of the light source, followed by standardization.

Default sensitivity peaks of visual pigments (λ_{max} values) for each cone opsin were obtained from previously published papers. In some cases, generic λ_{max} could be adjusted when species-specific λ_{max} or those of closely related species were available (see Table S.V.5 for details and references, and lens T50 values).

To model the visual sensitivity of each cichlid species, standardized cone opsin absorption curves were calculated according to Govardovskii *et al.* (2000) considering the species-specific λ_{max} of each opsin. Subsequently, these were weighted according to the respective opsins' relative expression levels as suggested by Hofmann *et al.* (2009). Retina sensitivity curves were then calculated by adding up the integrals of all cone opsin absorption curves. Finally, the sensitivity curves per species were corrected for the species-specific long-pass filter effects of the lens. Each individual lens spectrum was first corrected for potential skews in the 450nm to 800nm wavelength range (measurement artifacts, as mentioned in Hofmann *et al.* (2010), by fitting a linear regression to this range. Only the standardized residuals to the computed regression line across the full 280nm to 800nm range was then used for a sigmoid fit curve to describe the lens specific transmittance properties. Repeated measurements of the same lens showed that in this way sigmoid fits had a higher R^2. By averaging all sigmoid fits across each species, an average sigmoid fit line was calculated that described the species-specific lens transmittance properties (similar to

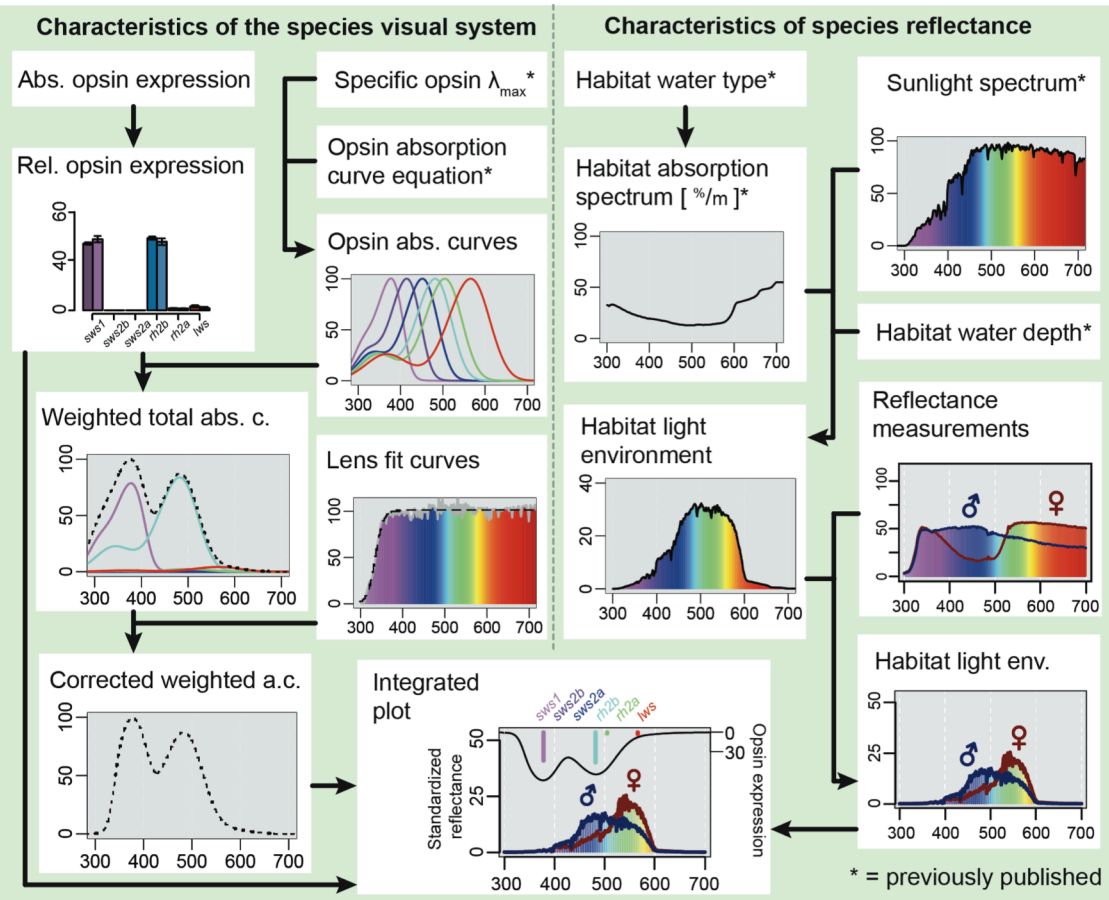


Fig. V.1 Pipeline to integrate opsin expression, body reflectance and natural habitat conditions. To characterize the visual system of investigated species, absolute opsin expression values were used to compute relative values. Opsin absorption peak wavelength and a general opsin absorption curve equation were used to calculate opsin absorption curves, which were weighted by the species-specific opsin expression profile and the transmission properties of the lens, using a sigmoid fit curve. For each species, the natural light environment was estimated by integrating the depth of the fish's natural habitat, a spectrum of the sunlight and the habitat water's specific absorption properties. The reflectance of each species and sex was characterized by measuring its relative reflectance, multiplying it with its habitat light spectrum and standardization. Visual capacities and reflective properties were then integrated for each species and measured spot in a common plot, showing the average modeled reflectance spectra for both sexes (left axis) and the weighted retina sensitivity curve and relative opsin expression levels upside down (right axis).

Hofmann *et al.* 2010). In contrast to previously published papers, however, we did not (only) extract the T50 value from these resulting average sigmoid fits, but the curves were multiplied with the total retina sensitivity curves of their respective species. Thus, a lens-corrected weighted retina sensitivity curve was estimated for each species.

3.2 Modeling African cichlids' body reflectance in their natural habitat

For the five African cichlid species, the reflective properties of the fish's bodies were measured at five to six spots on the fish, including cheek, back, belly, anal fin (anterior part), egg spots (if applicable) and an optional species-specific additional spot, representing a particular conspicuous or colorful patch (Fig. S.V.1 for more details). Euthanized fish tended to extract their melanophores, which let to generally darker fish compared to their normal coloration. However, the melanophores mostly altered the brightness of the fish and thus have only little effect on the relative spectral composition of reflected light, as melanin does not have an absorption peak in the visible light wavelength range (Zonios *et al.* 2008). The PX-2 pulsed xenon lamp was used to illuminate each measurement spot via a bifurcated fiber with a single probe that also allowed us to measure this spot at the same time and angle using the aforementioned spectrometer. Before each measurement, the respective spot was superficially dried with a tissue and then the probe was held at a 40-60° angle directly at the skin to measure the integrated reflective properties of an approximately 0.5cm^2 spot. Thus, particularly small structures (e.g. egg spots) may also contain some noise from unintentionally illuminated and measured surrounding skin. Comparisons to measurements conducted with in water submerged fish and probe showed no relevant differences from the surface water measurements. Reflectance measurements were recorded from 280 nm to 800nm light wavelength, but we will only report the visually relevant range from 300nm to 700nm. The relative reflectance spectrum per fish and spot (Fig. S.V.2) was calculated by dividing the absolute spectrum by the illumination spectrum, which was determined using a Spectralon diffuse white standard, and subsequent standardization.

To obtain estimates of species- and sex-specific body reflectance spectra of cichlids in their natural habitat, body light reflectance spectra have to be further processed. Due to potentially variable melanophore spread levels of sampled fish and somewhat variable tissue-probe-distances during measurements, raw absolute reflectance measurements are not necessarily comparable across individuals and spots. Melanin, the major pigment used by most animals to shade skin, does not have reflective peaks or valleys in the visual light wavelength range (Zonios *et al.* 2008). Instead melanin's reflectance is following roughly an exponentially decreasing curve, absorbing somewhat more light at short wavelengths and reflecting more at longer wavelengths.

To allow comparisons of spectral peaks nonetheless, relative reflectance spectra heights were corrected by standardizing them for equal integrals between sexes' average spectra, by adjusting the sex's spectrum with the lower integral to the sex's spectrum with the higher one. Thus, spectral peaks and valleys of reflectance are readily recognizable in

both sexes. However, it should be noted that all information on brightness (i.e. spectra heights) are non-informative. We provide photos of living fish of the investigated species to give some insights in how brightness levels may appear in live fish (Fig. V.3).

The habitat preference of each species was determined based on previously published literature (Table V.1 for details and references). To model the habitat ambient light environment, particularly the light absorption properties of the water of the respective habitat were of interest and two water-types were assumed for this study: 'Lake Malawi rock habitat' (very clear water; for all Malawi cichlids) and 'Lake Victoria Makobe' (somewhat murky water, for *P. nyererei*). We obtained the relative light absorption spectra per meter depth for the Lake Malawi habitats from Sabbah *et al.* (2011). The Lake Victoria habitat relative light absorption spectra per meter depth were calculated using data from Maan *et al.* (2010) (Fig. S.V.3). The water depth in which each species is most common was obtained through literature search and personal communication with Lake Malawi experts (all data and references in Table V.1).

Species reflectance patterns in bright sunlight above water (Fig. S.V.4) and in their respective habitats (Fig. S.V.5) were then modeled by multiplying the relative reflectance spectra with the sunlight spectrum (obtained from https://www.berthold.com/en/bio/how-do-i-convert-irradiance-photon-flux) and the putative available down-welling light spectrum of the fish's corresponding habitat, respectively (Fig. S.V.4). Sexes' spectra were then adjusted for equal integrals as outlined before. Furthermore, the percentage overlap of male and female reflectance spectra were calculated to estimate the degree of sexual dimorphism (SDi), as well as the overlap of both sexes' reflectance spectra with the spectra of the habitats' down-welling light to estimate general chromatic conspicuousness in the respective habitat (MC=male conspicuousness; FC=female conspicuousness; Table S.V.6).

3.3 Data analysis: integrating ecology and physiology in African cichlids

We integrate the visual sensitivity of African cichlid species and the modeled color patterns of males and females in their respective habitats by combining the weighted retina sensitivity curves, positions of expressed opsins and the reflectance patterns in single plots (Fig. S.V.6). Final sample sizes for all measurements are reported (Table S.V.7).

3.4 Simplifications and assumptions

The used pipeline integrates a wide range of data, however, several aspects of cichlid ecology and physiology had to be simplified as sufficient data was either not available or too complex. These simplifications include the usage of only two water types and the negligence of seasonality effects on those (Smith *et al.* 2012) and the fact that fish occur at a range of water depths and only one depth is used here. Also, our analysis neglects changes in water

absorption properties with increasing depth and only considers down-welling light (also for conspicuousness calculations, albeit side-welling light or rock reflectance may be more appropriate). Further simplifications include the negligence of chromophore usage (A1/A2 ratio, though its effect on visual systems is probably very small (Carleton & Kocher 2001)), cone structuring in single- and double cones and the potentially somewhat variable transmitting properties of lenses in the wavelength ranges beyond 410nm.

We further do not account for opsin expression variation during development (only sexually mature fish were used and, among those, no relationship of opsin expression and age could be found) and for opsin expression variations across the day (though most fish were killed in the early afternoon and the few that were killed at different times do not appear as outliers). Also, measuring physiological variables (opsin expression, lens transmission) can only approximate visual capabilities while actual visual abilities can only be determined using behavioral approaches (Kelber *et al.* 2003; Kalb *et al.* 2015). Finally, we only consider the measured spots and also neglect any body patterns due to differences in brightness.

Results

Cichlid fishes are famously known for their outstanding diversity in body coloration, which can even vary between sexes of the same species. We investigated opsin expression levels of males and females in ten cichlid species with varying degrees of sexual dimorphism to test whether sexual dimorphism in cichlid opsin expression readily can be found and whether it is correlated to sexual dimorphism in body coloration, as suggested by Fisher's run-away hypothesis.

No differences in opsin expression between sexes, irrespective of sexual dichromatism

When comparing the opsin expression levels between males and females across all species and opsin genes, we do not find any statistically significant differences after species-wise correction for multiple testing (Fig. V.2, Table S.V.4). Generally, all Neotropical cichlids examined express the same three cone opsin genes, *sws2a*, *rh2a* and *lws*, although some variation in the relative expression level can be found among species. *Apistogramma* species appear to have on average higher relative expression levels in the single cone opsin gene *sws2a* and a reduced expression of *rh2a*. African species were found to express much more diverse combinations of opsin genes. While all were found to express *lws* at moderate to high levels, no other cone opsin was expressed in these species. Expressed opsin numbers varied from two (*L. caeruleus*, *M. johannii*) to four (*P. lombardoi*). Rhodopsin 1 (*rho*) was found to be expressed in all species in considerable, but quite variable amounts: while in some species *rho* expression exceeds 90% of all expressed opsins (*H. nematopus*), in other

species it can be less than 20% (*A. ortegai*). This variation can be found in Neotropic as well as African cichlids.

Pseudogenization of the sws1 gene occurred more often in monomorphic cichlid species

Occurrences of opsin gene pseudogenization were investigated in 38 African cichlid species for which sequence information of opsin genes could be obtained and which could unambiguously be scored as sexually mono- or dimorphic in body coloration (Table S.V.4). Only in one opsin gene, s*ws1*, pseudogenization could be detected in several species: in seven of the 38 investigated cichlids *sws1* underwent pseudogenization. We determined that 13 monochromatic and 20 sexually dimorphic species had functional *sws1* genes, while all five pseudogenized *sws1* genes were found in monomorphic species. Thus, pseudogenization of the *sws1* opsin gene was thus found to occur significantly more often in monochromatic than in dichromatic species (χ^2= 6.397; *df*=1; *p* = 0.011).

Color spectra of investigated fishes relates to opsin expression patterns

We modeled body color reflectance spectra in African cichlids to investigate whether these can be related to opsin expression and, in particular, whether sexual dimorphisms in body coloration may affect the expressed opsin composition (Fig. V.3). All relative reflectance patterns of investigated fish show a reflection peak at around 340nm light wavelength (Fig. S.V.2). Besides this peak, reflection patterns of darker individuals are typically more flat (and would be lower, prior to standardization), while colored individuals have more or less pronounced reflectance peaks or shoulders. The sexually dimorphic species *P. lombardoi* and *M. johannii* have distinct reflective peaks for each sex (sexually dichromatic), while in sexually dimorphic species without chromatic differences between sexes (e.g. one sex is just 'darker' than the other) the darker sex typically has a flatter version of the reflective spectrum of the brighter sex (e.g. *L. caeruleus* and *M. auratus*).

Surface reflectance as approximated by our pipeline (Fig. V.3, Fig. S.V.3) shows that almost all species reflect sunlight at a wide range of wavelengths, although darker individuals may reflect overall less. Noteworthy is that all species included here do reflect light of very short wavelengths (UV) in considerable amounts. Furthermore, egg-spot color (in males only) appears to be very similar to the anterior part of the females' anal fins in *Melanochromis* species. In *P. lombardoi* egg-spot color was determined to be similar to those of *M. auratus*, however this does not coincide with female anal fin color, which is bluish, as the largest part of the female. Egg-spots in *P. nyererei* and *L. caeruleus* are ocher-brownish in our individuals. Generally, egg-spots do not contain 'unique' light wavelengths but conspicuousness is due to high contrast to the surrounding anal fin tissue, as exemplified by the anterior part of the anal fin.

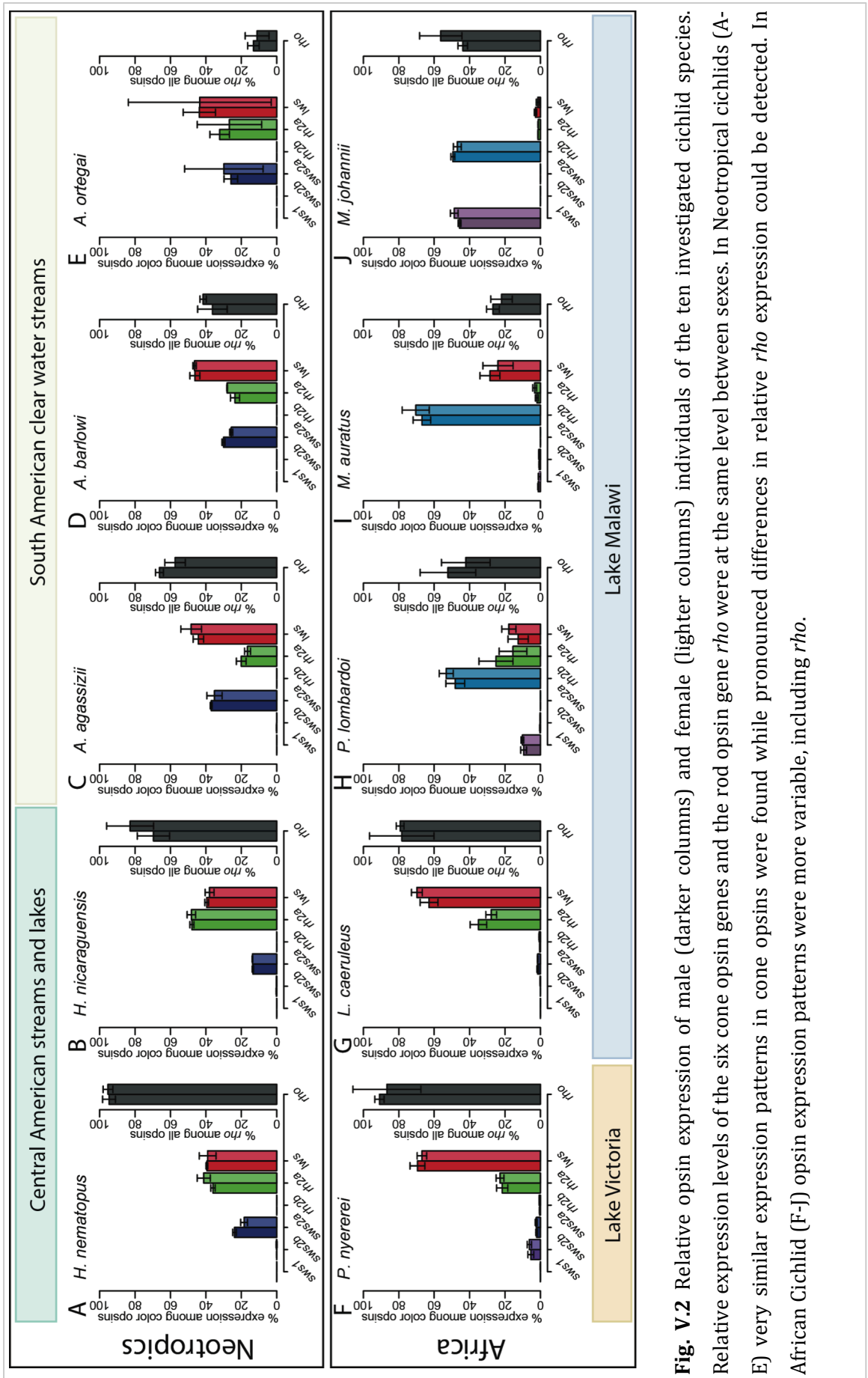


Fig. V.2 Relative opsin expression of male (darker columns) and female (lighter columns) individuals of the ten investigated cichlid species. Relative expression levels of the six cone opsin genes and the rod opsin gene *rho* were at the same level between sexes. In Neotropical cichlids (A-E) very similar expression patterns in cone opsins were found while pronounced differences in relative *rho* expression could be detected. In African Cichlid (F-J) opsin expression patterns were more variable, including *rho*.

Reflectance patterns change dramatically when the natural light environment is considered and reflectance (and thus body color) is predicted (Fig. V.3, Fig. S.V.4). Most notably, reflectance spectra become much more narrower, particularly in *P. nyererei* and *L. caeruleus*. The three Lake Malawi species inhabiting relatively shallow waters, still reflect considerable amounts of UV-light, while this is now virtually absent in *P. nyererei* and *L. caeruleus*. In dichromatic species (*P. lombardoi* & M. *johannii*), divergent reflectance peaks between sexes are still apparent at depth and chromatic reflectance differences extend clearly into short light wavelength around 400nm (Sexual-dichromatism values (SDi) were typically between 20% and 40%, i.e. 80% and 60% of sexes' reflectance curves overlap, respectively). In

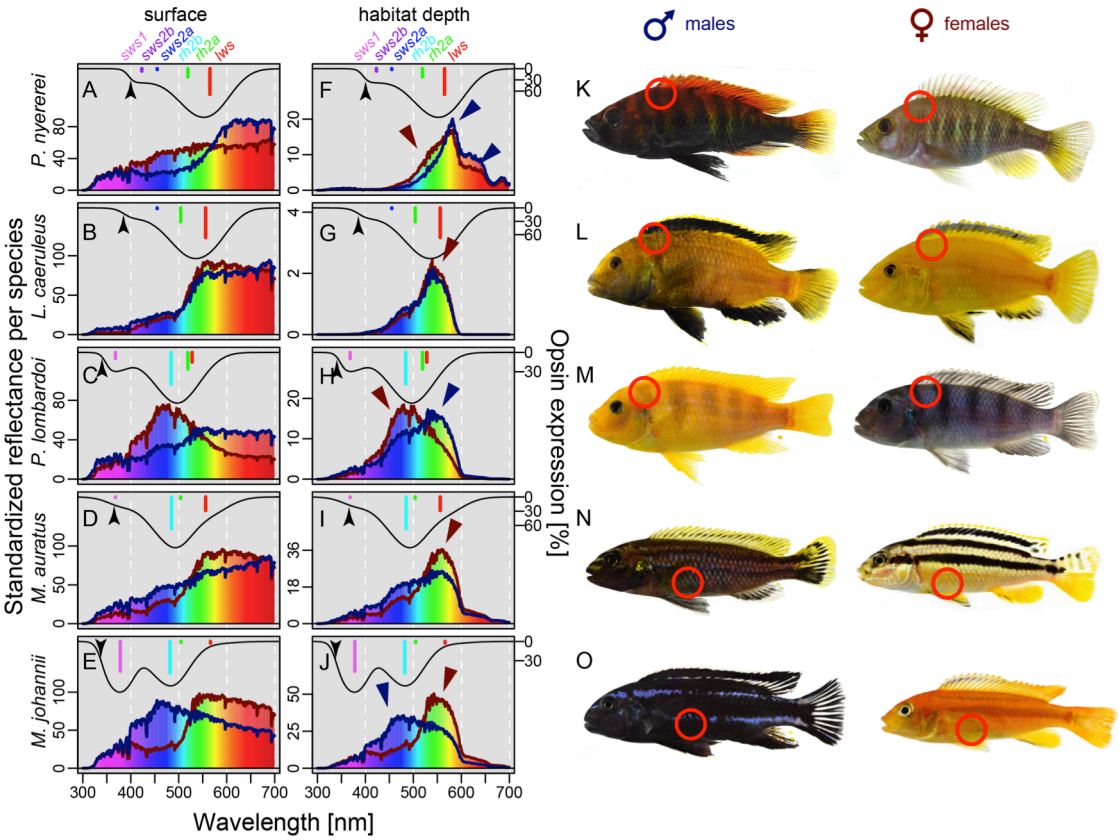


Fig. V.3 Modeled and standardized African cichlid body reflectance patterns above water (A-E) and at habitat depths (F-J) and indications where shown reflectance was measured (K-O). Additionally, the retina sensitivity curves and mean opsin expression levels are indicated bottom up on the right axis of each plot. Assumed habitat depths are 5m, 25m, 10m, 6m and 6m for F-J, respectively. Red and blue arrow-heads illustrate estimated reflectance peaks of fish in their natural habitat. Black arrow-heads indicate the T50-value of the lens. Species with broader reflectance spectra in their habitat appear to express a broader array of opsins. Species without UV reflectance in their natural habitat also do not express *sws1* (expression levels <1% not shown).

species with very narrow ambient light spectra (*P. nyererei* and *L. caeruleus*) only little divergence in peaks between the sexes is recognizable for most measured spots (SDi values barely exceeding 15%) and males appear to have the same chromatic reflectance as the female and are mostly just darker. There were three exceptions: one in *P. nyererei*, where males have a more narrow reflectance spectrum at the back resulting in a more reddish coloration than in females (SDi: 17%, MC: 19%, FC: 2%); one in the anal fin in *L. caeruleus*, with females having a much brighter fin (SDi: 17%, MC: 2%, FC: 16%); and egg-spots, which are only present in males. *L. caeruleus*, similarly to *P. nyererei*, has only one narrow reflectance peak, although in this species it does coincide much less with the peak wavelength of available light than in *P. nyererei*. This is reflected in the divergent conspicuousness values for the two species: while *P. nyererei* male and female reflectance almost matches the habitat light spectrum (almost all male (MC) and female (FC) conspicuousness values below 6%) the conspicuousness of *L. caeruleus* is much higher (MC and FC values repeatedly exceeding 20% and FC usually exceeding MC slightly). Males are divergent from females mostly by being much darker, particularly in the face, on the back and the anal fins, while reflectance matches almost perfectly between the sexes on the flanks (additional measuring spot in *L. caeruleus*). Although *M. auratus* lives in relatively shallow waters its reflectance coincides mostly with *L. caeruleus* in that males are dimorphic from females, but putatively by modulating reflectance intensity, but not so much hue. This makes FC values much higher than MC values, albeit this may be inversed for non-considered body regions, such as dorsal fins, which are brighter in males (Fig. V.3N).

Lastly, we combined our approximations of fish reflectance with the modeled weighted and corrected retina sensitivity curves (Fig. V.3, Fig. S.V.6). Species with a single clearly defined reflectance peak show most pronounced opsin expression in close proximity to this peak (*P. nyererei*, *L. caeruleus*, *M. auratus*). However, in species with divergent reflectance peaks expressed opsins cover a wider range of wavelengths (*P. lombardoi*, *M. johannii*). Furthermore, species living in habitats in which light of short-wavelengths is still available, tend to express short-wavelength sensitive opsins only when there is also sexual dichromatism and a reflectance peak in short-wavelengths. *M. auratus* individuals, although living in a very similar habitat as *M. johannii*, did not express any sort-wavelength opsins in considerable amounts, although its lens transmission allows short wavelength light to reach the retina (T50 at 364nm). The weighted retina sensitivity curve thus reflects the reflectance curves of both sexes insofar as species with two distinct reflectance peaks differentiating the sexes also have two sensitivity peaks in their retina sensitivity curves. Lenses were found to have very variable transmission rates for light of very short wavelengths: while lenses of *P. nyererei* do not allow UV light to pass (T50 of 400nm), we find much lower T50 values in species with UV-sensitive retinas. Most notably, *P. lombardoi*

and *M. johannii* had average T50 values of 336nm and 339nm – some of the shortest ever-reported in fish (e.g. Hofmann *et al.* 2010).

Discussion

Cichlid fishes are famously known for their outstanding diversity in body colorations found across species, but also within species across development or between sexes (Seehausen & van Alphen 1998; Seehausen *et al.* 2008). Fisherian run-away selection was suggested to have contributed to the evolution of sexually dimorphic body colorations, associated with the evolution of sexually dimorphic visual sensitivities. Here, we investigated whether visual dimorphism can readily be identified across selected cichlid species. Furthermore, we investigated whether being monomorphic, dimorphic or even dichromatic is affecting opsin expression patterns. We also considered species-specific environmental variables to approximate a natural setting that allowed us to also investigate the potential influence of habitat ambient light and conspecific body coloration on cichlid opsin expression.

Opsin expression varies among species, but not sexes

We investigated the expression of all seven cichlid cone opsin genes (*rh2aα* and *rh2aβ* pooled) plus the rod opsin gene *rho*, but found no significant differences between sexes, independent of body coloration type and continent. Generally, we found much more pronounced differences in the repertoire of expressed opsins across African cichlids compared to Neotropical cichlids, confirming previously found patterns (e.g. Hofmann *et al.* 2010; Weadick *et al.* 2012; Escobar‐Camacho *et al.* 2016). Most opsin expression patterns determined in this study are in line with previously published expression patterns, but some deviations should be noted. Most relevant, we could not confirm sexual dimorphisms in opsin expression in *M. auratus*, which was found in a previous study (Sabbah *et al.* 2010). Instead, our findings coincide with a study showing that *M. auratus* sexes have the same color discrimination capabilities (Coniam 2014). Generally, *M. auratus* used for this study showed somewhat different opsin expression patterns than those used in some previous studies, as we could not detect considerable short-wavelength opsin expression, suggesting that individuals originating from different populations might have been used (Dalton *et al.* 2010; Hofmann *et al.* 2010; Sabbah *et al.* 2010). Similar expression patterns were found for *L. caeruleus* – a species in which opsin expression was not previously investigated, to our knowledge. For *M. johannii*, for which only another population (possibly species) has been investigated before (Hofmann *et al.* 2010), expression of almost exclusively two opsins (of which one is the UV-opsin *sws1*) could be detected and it does virtually not express *lws*, but *rh2b*. To our knowledge, it also has the lowest T50 measured in cichlids (336nm) thus far

and only very few coral fish probably have lower T50 (Siebeck & Marshall 2001). This emphasizes the potential importance of short-wavelength light for this species.

rho expression reflects habitats' ambient light environments

We found that the proportion of *rho* expressed relative to cone opsins is very variable among investigated species from both continents. Interestingly, *rho* expression seems to negatively correlate with the estimated ambient light brightness of the species habitats. African species living in dim habitats, such as *P. nyererei* and *L. caeruleus* express very high amounts of *rho*. On the other hand, species living in brighter environments show lower levels of *rho* expression (*P. lombardoi* and the *Melanochromis* species).

In the studied Neotropical cichlids *rho* expression follows similar patterns, but incorporates habitat heterogeneity, which is typically much higher than in cichlids from the African Great lakes (Escobar‐Camacho *et al.* 2016). *H. nematopus* was found to express the highest proportion of *rho* and it inhabits diverse habitats, including the very murky and dim great Nicaraguan lakes, Lake Nicaragua and Lake Managua, as well as some smaller lakes with different and sometimes variable light conditions (e.g. lake Xiloá) and some rivers (personal observations). Also, *H. nematopus* does not rely on chromatic body coloration, but breeding state is signalized by an inversion of a dark-and-light-grey pattern on the fish's flanks. *H. nicaraguensis*, another wide-spread species, partially co-occurring with *H. nematopus,* showed the second highest *rho* levels in examined Neotropical species. This species may express larger proportions of cone opsins compared to *H. nematopus* as it clearly uses body coloration (particularly green to red light) for intraspecific communication.

Less *rho* expression was detected in *Apistogramma* species. *A. agassizii*, one of the genus' most widely distributed species, shows highest proportions of relative *rho* expression in their retinas among investigated *Apistogrammas* (Römer 1998). This species is found in small shallow clear-water streams, but also somewhat larger stream with murkier water conditions and thus dimmer ambient light. The most clear and shallow habitats (with depths <50cm) are inhabited by *A. ortegai* and *A. barlowi*. These species also have very restricted distribution areas limited to one or few small streams (Römer & Hahn 2008; Britzke *et al.* 2014) and express among the lowest levels of *rho* in our study. Therefore, also in Neotropical species *rho* expression appears to be negatively correlated with natural habitat light conditions. Albeit *rho* sequence evolution has previously been identified to be under divergent selection across cichlids (e.g. Schott *et al.* 2014; Torres-Dowdall *et al.* 2015), *rho* expression was not yet described or related to cichlid ecology. Yet, the current study did not formally test whether variation in *rho* expression is indeed correlated to habitat brightness and we encourage future studies investigating this in more

detail. Finally, determining whether expression differences in *rho* are caused by varying numbers of rod cells in the retina, varying expression levels of *rho* per rod or a combination of both was beyond the scope of this study, but may also foster future research.

Pseudogenization of sws1 occurs more often in monomorphic cichlid species

We tested whether body dimorphism affects pseudogenization rates of cone opsin genes, and found that for *sws1*, the only gene found to underwent pseudogenization in multiple species, non-functional versions can be found significantly more often in monomorphic than dimorphic species. Besides a potential function in mate choice, expression of *sws1* was previously determined to be tightly linked to ontogenetic stage in tilapia and some other cichlids (but not all - Spady *et al.* 2006; Carleton *et al.* 2008). Whereas *sws1* is expressed in larvae and in early juveniles, it becomes increasingly replaced first by *sws2b* and later *sws2a*. Its function in early larvae was linked to feeding on zooplankton, which absorbs UV light particularly well (Loew & McFarland 1990; Sabbah *et al.* 2012). In Malawi cichlids it has been observed that some species have a more direct development with larvae more-or-less skipping the *sws1*-phase (Carleton *et al.* 2008). Such heterochronic shifts are a potential mechanism by which *sws1* is effectively removed from any stabilizing selection and thus may pseudogenize subsequently. We found that, besides for feeding in larvae cichlids, *sws1* is likely to be involved in mate recognition, as its expression is not only more often found in dichromatic species in our study, but also we find it to be pseudogenized exclusively in monomorphic species (among those considered). Albeit *sws1* thus appears to be pseudogenized relatively frequently its loss is typically observed relatively recently (Weadick *et al.* 2012; Escobar‐Camacho *et al.* 2016), casting some doubt on the long-term evolutionary success rate of cichlid lineages that lost *sws1*.

Body reflectance in natural habitats likely affects opsin expression

In an effort to determine the extent to which ecology and different sexes' body reflectance patterns may influence cone opsin expression patterns, species' light perception capabilities were modeled by calculating weighted retina sensitivity curves and body reflectance of individuals in their natural habitat was approximated. As little variation in opsin expression was observed among Neotropical cichlids, we focused on the more variable African cichlids. Examined African species were all considered sexually dimorphic when selected for this study. Additionally, we consider two species also being unambiguously sexually dichromatic (i.e. sexually dimorphism is not solely due to variations in melanin and thus brightness; *P. lombardoi* & *M. johannii*). In our study, species with divergent reflection peaks between sexes under natural habitat light (Fig. V.3F-J, Fig. S.V.6) (*P. lombardoi* and *M. johannii*) express opsins that cover a wider range of wavelengths.

In contrast, species with one reflectance peak, which may be more or less narrow (*P. nye*rerei) or more or less high (*M. auratus*, *L. caeruleus*), typically express opsins with peak sensitivities more closely to this reflectance peak. This is also reflected in the weighted retina sensitivity curves, which have two peaks in species with sexual dichromatism (i.e. two distinct sex-specific reflectance peaks), while the first sensitivity peak (<400nm wavelength) is not present in other species. This short-wavelength peak is a result of *sws1* expression, but also relies on lens-transmissions at short wavelengths.

We found that the occurrence of multiple distinct reflectance peaks might lead to a broader array of expressed opsins. To this point it remains untested whether reflectance peaks have to be the consequence of sexual dichromatism or whether they can also be due to distinct peaks on the same individual. Furthermore, the reflection of a specific light-wavelength range alone (without peak) is unlikely to influence opsin expression at the similar range, as exemplified by *M. auratus*: specimens were found to reflect UV-light at similar levels as *M. johannii*, but do not express *sws1*, while the latter does.

We did not identify considerable *sws1* expression in species whose habitat light is deprived of short-wavelength light (UV). Thus, our data confirms that habitat ambient light is a major determinant of cichlid opsin expression. However, the reverse of this rule is not true: *M. auratus*, for example, lives in water depths in which considerable amounts of UV light are still available but the fish do virtually not express *sws1*. Also, we noted that, while some cichlids mostly express cone opsins with peak sensitivities approximating the peak of the available light spectrum (*P. nyererei*, *P. lombardoi*, *M. auratus*), others do not (*L. caeruleus* & *M. johannii*). Therefore, these fish are likely to have reduced total brightness perception among cone opsins and may compensate this by either having increased *rho* expression (*L. caeruleus*) or by expressing fewer different cone opsins, effectively reducing perceived color resolution in favor of color brightness (*M. johannii*) (Kelber *et al.* 2003). Expressing opsins not coinciding with the ambient light spectrum may allow the fish to detect other more relevant color cues in their environment, such as conspecific body coloration, with the maximum possible sensitivity. This observation argues for the importance of conspecific body coloration as a factor influencing African cichlid opsin expression.

The apparent role of conspecific body coloration in cichlids contrasts our finding that sexes of investigated species do express opsins at the same levels. Our findings do not provide evidence for a major role of Fisherian run-away sexual selection in shaping dimorphic cichlid species, as suggested in previous hypotheses (Kelber & Osorio 2010; Sabbah *et al.* 2010). Naturally, we cannot exclude that other species may show expression differences or that differences are too small for detection in this study. This contrasts the finding in bird visual systems, where divergent opsin expression between sexes may

constitute a main driver of sexual selection (Bloch 2015). Also in non-vertebrates sex-specific visual sensitivities have been reported (Hilbrant *et al.* 2014). The lack of sexually dimorphic opsin expression is somewhat surprising, considering previous evidence, but it may be that dimorphic visual sensitivity did not evolve in investigated species because sex determination in cichlid fishes is a very labile process, presumably depending on testosterone hormone titer that is also affected by the biotic and abiotic environment (Francis & Barlow 1993; Baroiller *et al.* 1995; Baroiller & D'cotta 2001). Thus, linking opsin expression to a specific sex may be, evolutionarily, not as straightforward as in species with sex chromosomes. Still, this study confirms that both ecological parameters and conspecific body coloration, potentially with an exceptionally important role in sexually dichromatic species, are likely determinants of cichlid vision.

Concluding remarks

Freshwater visual environments are often much more complex and diverse than those of marine or terrestrial environments, due to higher variability in light-transmission properties of freshwater-ecosystems, compared to seawater of marine systems or air of terrestrial systems. Therefore, it is challenging to determine the importance of factors that may have affected fish vision. Our integrative approach allowed us to conclude that fish vision is adjusted to their respective habitats, possibly by adjusting rod opsin expression according to light quantity and cone opsin expression according to the habitat's prevalent light regime. However, we also found evidence that within the range of available light wavelength, variation in opsin expression is found and that cone opsin expression patterns are likely adjusted to optimize perception of conspecific coloration, at least in the sampled colorful species. Our conclusions are mostly in line with previously published data, however we provide additional knowledge to previous findings by excluding sexual visual dimorphism as a major factor in cichlid visual ecology and by highlighting the potential correlation of rod opsin expression with habitat light quantity. Compared to previous studies, our integrated approach allowed us a more precise approximation of how conspecifics are potentially seen by cichlids in their natural habitat (Dalton *et al.* 2010). More data on ecological parameters but also visual properties of individual cichlid species are necessary to draw a more complete picture of the diversity of visual ecologies in these colorful fish.

Chapter VI

General Discussion

Main Contributions

The chapters of this thesis investigated aspects of the origin of biological diversity and the evolutionary mechanisms driving it by promoting phenotypic diversification, lineage divergence and ultimately speciation. Using a full-genome scan approach, the seahorse proved to be an excellent model to investigate the genomic signatures of its many unique characteristics, such as male pregnancy, a full body-cover of bony plates or the loss of the caudal and pelvic fins. The study nicely illustrates how a variety of evolutionary mechanisms, such as gene duplication and -loss, as well as coding sequence evolution and alterations of regulatory elements can affect traits and lead to novel phenotypes. Also, it shows how effectively different methodological approaches can be integrated to extend or validate bioinformatically derived hypotheses, e.g. using transcriptomics or cutting-edge genome editing approaches, such as CRISPR-Cas9. The seahorse genome thus helps us to understand how novel characters can arise, however, it does not necessarily teach us about the future evolutionary and ecological importance of these characters down the evolutionary road.

Cichlid fishes are among the most famous models for investigations on rapid phenotypic diversification, ecological adaptation and speciation (Kocher 2004). The main focus of this thesis was put on studying one of the most diverse traits in these fishes and on understanding *how* this trait could generate such diversity: the pharyngeal jaws (PJ) (Greenwood 1964; Liem 1974; Hoogerhoud 1986a, b; Hoogerhoud 1986c; Muschick *et al.* 2011). The main goal of chapter II-IV was to understand how phenotypic plasticity, a property classically assumed to lower the rate of evolution (Sultan 2015), might have actually increased the diversifying potential of cichlids. In chapter II a comprehensive review on plasticity and its potential contribution to adaptive radiations is given, focusing on colonization events and subsequent lineage proliferation. The manuscript illustrates in more detail than previous manuscripts *why* and *how* plastic lineages might be be particularly successful colonizers and seeders of adaptive radiations. Next, molecular mechanisms are reviewed that plausibly depict how an originally plastic trait can become genetically fixed (via a process termed 'genetic assimilation'), even without the highly controversial 'cost of plasticity' (DeWitt *et al.* 1998; DeWitt & Scheiner 2004; Hendry 2015; Murren *et al.* 2015).

This review was also the first to link non-adaptive plastic responses to genetic assimilation and thus provides the first plausible hypothesis about the evolutionary origin of non-adaptive plasticity, albeit its evolutionary significance was noticed before (e.g. Ghalambor *et al.* 2007; Ghalambor *et al.* 2015). The manuscript describes the expected degree of adaptive and non-adaptive plasticity in lineages undergoing genetic assimilation, depending on how far genetic assimilation already proceeded. Thus, the paper provides a tool that assists in determining whether genetic assimilation occurs in a lineage and at which stage of genetic fixation a formerly plastic trait is. These patterns are discussed using previously published and recent data currently in preparation for publication and included in this thesis as chapter IV.

Besides providing a theoretical framework explaining how phenotypic plasticity potentially promotes evolution, this thesis contains two empirical studies that are among the first to investigate the transcriptional patterns that are associated with the PJ plastic response. Chapter III provides valuable insights in how morphological and transcriptional patterns diverge throughout the development of different diet-induced plastic PJ phenotypes. This allowed us to pinpoint a critical stage at which morphological divergence becomes observable and that is preceded by a marked change in expression patterns of candidate genes. Also, this study was the first to notice the modular nature of the transcriptional patterns associated with the PJ plastic response. Modularity is well-known to increase evolvability both on a morphological and transcriptional level (Ancel & Fontana 2000; Schlosser & Wagner 2004; Wagner *et al.* 2007). By conducting a transcription factor binding site analysis based on the genome of the closely related cichlid *Astatotilapia burtoni*, a preliminary regulatory network was synthesized that sketches some of the presumed regulatory pathways orchestrating the plastic response.

Chapter IV is the first study using diet-manipulation experiments in multiple cichlid species to compare their PJ plastic response. This study uses particularly comprehensive analyses to show that candidate species have distinct PJ morphologies and distinct transcriptional patterns. More importantly: the study shows that the degree of inducible adaptive plasticity is different among species, with generalists showing more pronounced adaptive plasticity than specialists. This is in line with patterns expected in a 'flexible-stem' system, where multiple specialized lineages emerge from a plastic ancestor via genetic assimilation (West-Eberhard 1989; West-Eberhard 2003). Two candidate genes are identified that may have undergone genetic assimilation, based on their expression reaction norms. Finally, and potentially most important for future studies: non-adaptive plasticity is indirectly recorded and, as predicted in chapter II, it is most pronounced in those lineages that presumably are at an intermediate stage of genetic assimilation. Thus, the study presented in chapter IV provides the first empirical data suggesting that non-adaptive

plasticity can be informative in identifying lineages undergoing genetic assimilation. Yet, to date available data is too limited to comprehensively compare the degree of adaptive and non-adaptive morphological and transcriptional plastic responses across lineages representing different stages of genetic assimilation and ongoing research will allow these analyses in the future.

Finally, the study presented in chapter V investigated another major aspect of phenotypic diversity among cichlids: body coloration and visual abilities. Albeit suggested by several previous studies (Kelber & Osorio 2010; Sabbah *et al.* 2010), no evidence for Fisher's run-away sexual selection to contribute to sexual dimorphism of body coloration diversity was found. Nonetheless, a link between sexual color dimorphism and the visual system was identified: the short-wavelength cone opsin gene *sws1* pseudogenized significantly more often in monomorphic than dimorphic species. Interestingly, it seems that all species with non-functional *sws1* have lost this opsin only since relatively short evolutionary time-spans (Escobar‐Camacho *et al.* 2016). This raises the question whether lineages without functional *sws1* are evolutionarily less successful. Considering that no sexually dimorphic species investigated had lost *sws1*, being sexually dimorphic in body coloration may offer the benefit of keeping *sws1* under the influence of natural selection and thus reducing the chance of its pseudogenization.

This study is also the first to identify a potential link between the rod opsin *rho* expression and habitat light quantity in cichlids, although this is not formerly tested (but see Bowmaker *et al.* (1994)). Anyhow, considerable variation across species in the relative expression of *rho* was identified. Future studies are encouraged to determine whether variations in *rho* is due to variations in rod numbers, variations in *rho* expression levels per rod or a combination of both. By integrating ecological and transcriptional data the study finds that sexually dimorphic species with divergent reflectance peaks between sexes typically also express opsins that cover a wider range of wavelengths compared to species with a single reflectance peak. Therefore, dimorphic species are more likely to express *sws1*, which may explain previously identified biases in *sws1* pseudogenization patterns. This study illustrates some of the many aspects of cichlid visual ecology and, by integrating available data, demonstrates how important integrative approaches are when complex traits, such as color vision in freshwater fish, are examined.

Advancing the understanding of the evolutionary significance of phenotypic plasticity

Phenotypic plasticity has long been recognized as being a central ecological trait (or trait property, depending on the point of view) as it promotes population persistence in heterogeneous environments. In contrast, its evolutionary consequences remained a topic of controversial debate for the last decades. Only in the last years empirical studies

identified cases in which ancestral plasticity likely has contributed to diversification and increased evolutionary rates (e.g. Moczek & Kijimoto 2014; Standen *et al.* 2014; Parsons *et al.* 2016) and generally the evolutionary significance of plasticity becomes more broadly accepted and acknowledged. Our understanding of the environmental properties that facilitate (or hamper) diversification based on plastic traits is reasonably advanced (Pfennig *et al.* 2002; West-Eberhard 2003; Pfennig *et al.* 2010; Ehrenreich & Pfennig 2015; Sultan 2015; Levis & Pfennig 2016). Still, one of the major weaknesses of a plasticity-promotes-diversification framework is our deficient understanding of how molecular mechanisms do fix a plastic trait genetically – and when. It should be the focus of future theoretical and empirical studies to shed light on this 'black-box of genetic assimilation' to promote evolutionary hypotheses incorporating plasticity.

The central contribution of this thesis is thus that it highlighted the potential evolutionary importance of phenotypic plasticity and to advance the field by reviewing some plausible but still very basic molecular mechanisms that could genetically fix plastic phenotypes. Also, it was the intention of the manuscript presented in chapter II to emphasize that genetic assimilation can (theoretically) occur without any need for a 'cost of plasticity' (DeWitt *et al.* 1998). This is a relevant contribution as 'costs of plasticity' are rarely shown and probably negligible (Hendry 2015) and it has been a major criticism that genetic assimilation would necessarily require these costs. Future studies are therefore encouraged to elaborate theoretically the molecular mechanisms potentially involved in genetic assimilation, but, most importantly, to also empirically test for ongoing genetic assimilation in a broad range of taxa. Ultimately, descriptions of the molecular processes that drove the genetic fixation of formerly plastic phenotypes in investigated taxa can complete our picture of plasticity's effects on evolution and diversification.

References

Abila R (2011) Preliminary gut content and dentition analysis reveal subtle resource partitioning and feeding adaptations within a haplochromine cichlid community of Lake Victoria satellite lake. *African Journal of Environmental Science and Technology* **5**, 457-463.

Abila R, Salzburger W, Ndonga MF*, et al.* (2008) The role of the Yala swamp lakes in the conservation of Lake Victoria region haplochromine cichlids: evidence from genetic and trophic ecology studies. *Lakes & Reservoirs: Research & Management* **13**, 95-104.

Abouheif E, Wray GA (2002) Evolution of the gene network underlying wing polyphenism in ants. *Science* **297**, 249-252.

Abzhanov A, Protas M, Grant BR, Grant PR, Tabin CJ (2004) Bmp4 and morphological variation of beaks in Darwin's finches. *Science* **305**, 1462-1465.

Adams DC, Otárola‐Castillo E (2013) geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods in Ecology and Evolution* **4**, 393-399.

Adler FR, Harvell DC (1990) Inducible defenses, phenotypic variability and biotic environments. *Trends in Ecology & Evolution* **5**, 407-410.

Albertson R, Kocher T (2006) Genetic and developmental basis of cichlid trophic diversity. *Heredity* **97**, 211-221.

Allender CJ, Seehausen O, Knight ME, Turner GF, Maclean N (2003) Divergent selection during speciation of Lake Malawi cichlid fishes inferred from parallel radiations in nuptial coloration. *Proceedings of the National Academy of Sciences* **100**, 14074-14079.

Alonso CR, Wilkins AS (2005) The molecular elements that underlie developmental evolution. *Nature Reviews Genetics* **6**, 709-715.

An G, Miner CA, Nixon JC*, et al.* (2010) Loss of Bright/ARID3a function promotes developmental plasticity. *Stem Cells* **28**, 1560-1567.

Ancel LW, Fontana W (2000) Plasticity, evolvability, and modularity in RNA. *Journal of Experimental Zoology* **288**, 242-283.

Anthes N, Theobald J, Gerlach T, Meadows MG, Michiels NK (2016) Diversity and Ecological Correlates of Red Fluorescence in Marine Fishes. *Frontiers in Ecology and Evolution* **4**, 126.

Arora R, Metzger RJ, Papaioannou VE (2012) Multiple roles and interactions of Tbx4 and Tbx5 in development of the respiratory system. *PLoS Genet* **8**, e1002866.

Attanasio C, Nord AS, Zhu Y*, et al.* (2013) Fine tuning of craniofacial morphology by distant-acting enhancers. *Science* **342**, 1241006.

Aubin-Horth N, Letcher BH, Hofmann HA (2009) Gene-expression signatures of Atlantic salmon's plastic life cycle. *General and comparative endocrinology* **163**, 278-284.

Aubin-Horth N, Renn SC (2009) Genomic reaction norms: using integrative biology to understand molecular mechanisms of phenotypic plasticity. *Molecular Ecology* **18**, 3763-3780.

Auld JR, Agrawal AA, Relyea RA (2009) Re-evaluating the costs and limits of adaptive phenotypic plasticity. *Proceedings of the Royal Society of London B: Biological Sciences*, rspb20091355.

Badyaev AV, Foresman KR, Young RL (2005) Evolution of morphological integration: developmental accommodation of stress‐induced variation. *The American Naturalist* **166**, 382-395.

Bailly X, Leroy R, Carney S*, et al.* (2003) The loss of the hemoglobin H2S-binding function in annelids from sulfide-free habitats reveals molecular adaptation driven by Darwinian positive selection. *Proceedings of the National Academy of Sciences* **100**, 5885-5890.

Bakker TC, Mundwiler B (1994) Female mate choice and male red coloration in a natural three-spined stickleback (Gasterosteus aculeatus) population. *Behavioral Ecology* **5**, 74-80.

Baldwin JM (1896) A new factor in evolution (Continued). *American Naturalist*, 536-553.

Barluenga M, Meyer A (2010) Phylogeography, colonization and population history of the Midas cichlid species complex (Amphilophus spp.) in the Nicaraguan crater lakes. *BMC Evolutionary Biology* **10**, 326.

Barluenga M, Stölting KN, Salzburger W, Muschick M, Meyer A (2006) Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature* **439**, 719-723.

Baroiller J, D'cotta H (2001) Environment and sex determination in farmed fish. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* **130**, 399-409.

Baroiller JF, Chourrout D, Fostier A, Jalabert B (1995) Temperature and sex chromosomes govern sex ratios of the mouthbrooding cichlid fish Oreochromis niloticus. *Journal of Experimental Zoology* **273**, 216-223.

Barrett RD, Schluter D (2008) Adaptation from standing genetic variation. *Trends in Ecology & Evolution* **23**, 38-44.

Beaty LE, Wormington JD, Kensinger BJ*, et al.* (2016) Shaped by the past, acting in the present: transgenerational plasticity of anti‐predatory traits. *Oikos*.

Bejerano G, Pheasant M, Makunin I*, et al.* (2004) Ultraconserved elements in the human genome. *Science* **304**, 1321-1325.

Beldade P, Mateus ARA, Keller RA (2011) Evolution and molecular mechanisms of adaptive developmental plasticity. *Molecular Ecology* **20**, 1347-1363.

Bell AM, Aubin-Horth N (2010) What can whole genome expression data tell us about the ecology and evolution of personality? *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 4001-4012.

Bento G, Ogawa A, Sommer RJ (2010) Co-option of the hormone-signalling module dafachronic acid-DAF-12 in nematode evolution. *Nature* **466**, 494-497.

Bergwerff M, Groot GD, Adriana C*, et al.* (1998) Patterns of paired‐related homeobox genes PRX1 and PRX2 suggest involvement in matrix modulation in the developing chick vascular system. *Developmental Dynamics* **213**, 59-70.

Bian C, Hu Y, Ravi V*, et al.* (2016) The Asian arowana (Scleropages formosus) genome provides new insights into the evolution of an early lineage of teleosts. *Scientific reports* **6**.

Binning S, Chapman L, Cosandey‐Godin A (2009) Specialized morphology for a generalist diet: evidence for Liem's paradox in a cichlid fish. *Journal of Fish Biology* **75**, 1683-1699.

Binning S, Chapman L, Dumont J (2010) Feeding and breathing: trait correlations in an African cichlid fish. *Journal of Zoology* **282**, 140-149.

Binning SA, Chapman LJ (2008) Feeding ecology and diet overlap in riverine cichlids from western Uganda. *Internationale Vereinigung fur Theoretische und Angewandte Limnologie Verhandlungen* **30**, 283-286.

Birney E, Clamp M, Durbin R (2004) GeneWise and genomewise. *Genome research* **14**, 988-995.

Blanchette M, Kent WJ, Riemer C*, et al.* (2004) Aligning multiple genomic sequences with the threaded blockset aligner. *Genome research* **14**, 708-715.

Bloch NI (2015) Evolution of opsin expression in birds driven by sexual selection and habitat. *Proceedings of the Royal Society of London B: Biological Sciences* **282**, 20142321.

Bonewald L, Kiel D, Clemens T*, et al.* (2013) Forum on bone and skeletal muscle interactions: Summary of the proceedings of an ASBMR workshop. *Journal of Bone and Mineral Research*.

Boughman JW (2002) How sensory drive can promote speciation. *Trends in Ecology & Evolution* **17**, 571-577.

Bouton N, Seehausen O, Alphen Jv (1997) Resource partitioning among rock‐dwelling haplochromines (Pisces: Cichlidae) from Lake Victoria. *Ecology of Freshwater Fish* **6**, 225-240.

Bowmaker J, Govardovskii V, Shukolyukov S*, et al.* (1994) Visual pigments and the photic environment: the cottoid fish of Lake Baikal. *Vision research* **34**, 591-605.

Bowmaker JK (1990) Visual pigments of fishes. In: *The visual system of fish*, pp. 81-107. Springer.

Bowmaker JK, Heath L, Wilkie S, Hunt D (1997) Visual pigments and oil droplets from six classes of photoreceptor in the retinas of birds. *Vision research* **37**, 2183-2194.

Bradshaw AD (1965) Evolutionary significance of phenotypic plasticity in plants. *Advances in genetics* **13**, 115-155.

Brawand D, Wagner CE, Li YI*, et al.* (2014) The genomic substrate for adaptive radiation in African cichlid fish. *Nature* **513**, 375–381.

Brewer MS, Carter RA, Croucher PJ, Gillespie RG (2015) Shifting habitats, morphology, and selective pressures: Developmental polyphenism in an adaptive radiation of Hawaiian spiders. *Evolution* **69**, 162-178.

Britten RJ, Davidson EH (1971) Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty. *Quarterly Review of Biology*, 111-138.

Britzke R, Oliveira C, Kullander SO (2014) Apistogramma ortegai (Teleostei: Cichlidae), a new species of cichlid fish from the Ampyiacu River in the Peruvian Amazon basin. *Zootaxa* **3869**, 409-419.

Burley N, Coopersmith CB (1987) Bill color preferences of zebra finches. *Ethology* **76**, 133-151.

Cahan SH, Julian GE, Rissing SW*, et al.* (2004) Loss of phenotypic plasticity generates genotype-caste association in harvester ants. *Current Biology* **14**, 2277-2282.

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972-1973.

Carleton K (2009) Cichlid fish visual systems: mechanisms of spectral tuning. *Integrative Zoology* **4**, 75-86.

Carleton KL, Dalton BE, Escobar‐Camacho D, Nandamuri SP (2016) Proximate and ultimate causes of variable visual sensitivities: insights from cichlid fish radiations. *genesis*.

Carleton KL, Kocher TD (2001) Cone opsin genes of African cichlid fishes: tuning spectral sensitivity by differential gene expression. *Molecular biology and evolution* **18**, 1540-1550.

Carleton KL, Parry JW, Bowmaker JK, Hunt DM, Seehausen O (2005) Colour vision and speciation in Lake Victoria cichlids of the genus Pundamilia. *Molecular Ecology* **14**, 4341-4353.

Carleton KL, Spady TC, Streelman JT*, et al.* (2008) Visual sensitivities tuned by heterochronic shifts in opsin gene expression. *BMC biology* **6**, 1.

Carroll SB (2008) Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**, 25-36.

Chan YF, Marks ME, Jones FC*, et al.* (2010) Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science* **327**, 302-305.

Chapman L, Galis F, Shinn J (2001) Phenotypic plasticity and the possible role of genetic assimilation: Hypoxia‚Äêinduced trade‚Äêoffs in the morphological traits of an African cichlid. *Ecology Letters* **3**, 387-393.

Chen NX, Ryder KD, Pavalko FM*, et al.* (2000) Ca2+ regulates fluid shear-induced cytoskeletal reorganization and gene expression in osteoblasts. *American Journal of Physiology-Cell Physiology* **278**, C989-C997.

Chevin L-M, Lande R, Mace GM (2010) Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. *PLoS Biol* **8**, e1000357.

Clabaut C, Bunje PM, Salzburger W, Meyer A (2007) Geometric morphometric analyses provide evidence for the adaptive character of the Tanganyikan cichlid fish radiations. *Evolution* **61**, 560-578.

Clausen J, Keck DD, Hiesey WM (1941) Regional differentiation in plant species. *American Naturalist* **75**, 231-250.

Colombo M, Diepeveen ET, Muschick M*, et al.* (2013) The ecological and genetic basis of convergent thick‐lipped phenotypes in cichlid fishes. *Molecular Ecology* **22**, 670-684.

Colosimo PF, Hosemann KE, Balabhadra S*, et al.* (2005) Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* **307**, 1928-1933.

Coniam M (2014) *Sex Differences in Visual Perception in Melanochromis auratus*, Bowling Green State University.

Cosandey-Godin A, Binning SA, Chapman LJ (2008a) Specialized morphology for a non-specialized diet: Liem's paradox in an African cichlid fish. *McGill Science Undergraduate Research Journal* **3**.

Cosandey-Godin A, Binning SA, Chapman LJ (2008b) Specialized morphology for a non-specialized diet: Liem's paradox in an African cichlid fish. *McGill Science Undergraduate Research Journal* **3**, 19-23.

Cotman CW, Berchtold NC (2002) Exercise: a behavioral intervention to enhance brain health and plasticity. *Trends in neurosciences* **25**, 295-301.

Crispo E (2007) The Baldwin effect and genetic assimilation: revisiting two mechanisms of evolutionary change mediated by phenotypic plasticity. *Evolution* **61**, 2469-2479.

Crispo E, Chapman L (2010) Geographic variation in phenotypic plasticity in response to dissolved oxygen in an African cichlid fish. *Journal of Evolutionary Biology* **23**, 2091-2103.

Crispo E, Chapman LJ (2008) Population genetic structure across dissolved oxygen regimes in an African cichlid fish. *Molecular Ecology* **17**, 2134-2148.

Cronin TW, Johnsen S, Marshall NJ, Warrant EJ (2014) *Visual ecology* Princeton University Press.

Cronin TW, Porter ML (2014) The evolution of invertebrate photopigments and photoreceptors. In: *Evolution of visual and non-visual pigments*, pp. 105-135. Springer.

Currey JD (2002) *Bones: structure and mechanics* Princeton University Press.

Dalton BE, Cronin TW, Marshall NJ, Carleton KL (2010) The fish eye view: are cichlids conspicuous? *The Journal of experimental biology* **213**, 2243-2255.

Dang DT, Pevsner J, Yang VW (2000) The biology of the mammalian Krüppel-like family of transcription factors. *The international journal of biochemistry & cell biology* **32**, 1103-1121.

Davidovitch Z, Shanfeld J, Montgomery P*, et al.* (1984) Biochemical mediators of the effects of mechanical forces and electric currents on mineralized tissues. *Calcified tissue international* **36**, S86-S97.

Davidson AM, Jennions M, Nicotra AB (2011) Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta‐analysis. *Ecology Letters* **14**, 419-431.

Davidson EH, Erwin DH (2006) Gene regulatory networks and the evolution of animal body plans. *Science* **311**, 796-800.

Davies R (1963) A molecular theory of muscle contraction: calcium-dependent contractions with hydrogen bond formation plus ATP-dependent extensions of part of the myosin-actin cross-bridges. *Nature* **199**, 1068-1074.

Day T, Pritchard J, Schluter D (1994) A comparison of two sticklebacks. *Evolution*, 1723-1734.

De Jong G (2005) Evolution of phenotypic plasticity: patterns of plasticity and the emergence of ecotypes. *New Phytologist* **166**, 101-118.

Deméré TA, McGowen MR, Berta A, Gatesy J (2008) Morphological and molecular evidence for a stepwise evolutionary transition from teeth to baleen in mysticete whales. *Systematic Biology* **57**, 15-37.

Denver RJ (1997) Environmental stress as a developmental cue: corticotropin-releasing hormone is a proximate mediator of adaptive phenotypic plasticity in amphibian metamorphosis. *Hormones and behavior* **31**, 169-179.

Deutsch JC (1997) Colour diversification in Malawi cichlids: evidence for adaptation, reinforcement or sexual selection? *Biological Journal of the Linnean Society* **62**, 1-14.

DeWitt TJ, Robinson BW, Wilson DS (2000) Functional diversity among predators of a freshwater snail imposes an adaptive trade-off for shell morphology. *Evolutionary Ecology Research* **2**, 129-148.

DeWitt TJ, Scheiner SM (2004) *Phenotypic plasticity: functional and conceptual approaches* Oxford University Press Oxford.

DeWitt TJ, Sih A, Wilson DS (1998) Costs and limits of phenotypic plasticity. *Trends in Ecology & Evolution* **13**, 77-81.

Dieckmann U, Doebeli M (1999) On the origin of species by sympatric speciation. *Nature* **400**, 354-357.

Dodson S (1988) The ecological role of chemical stimuli for the zooplankton: Predator‐avoidance behavior in Daphnia. *Limnology and Oceanography* **33**, 1431-1439.

Doebeli M, Dieckmann U (2003) Speciation along environmental gradients. *Nature* **421**, 259-264.

Don EK, de Jong-Curtain TA, Doggett K*, et al.* (2016) Genetic basis of hindlimb loss in a naturally occurring vertebrate model. *Biology open*, bio. 016295.

Douglas R, Marshall N (1999) A review of vertebrate and invertebrate ocular filters. In: *Adaptive mechanisms in the ecology of vision*, pp. 95-162. Springer.

Draghi JA, Whitlock MC (2012) Phenotypic plasticity facilitates mutational variance, genetic variance, and evolvability along the major axis of environmental variation. *Evolution* **66**, 2891-2902.

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research* **32**, 1792-1797.

Edidin M (1983) MHC antigens and non-immune functions. *Immunology today* **4**, 269-270.

Ehrenreich IM, Pfennig DW (2015) Genetic assimilation: a review of its potential proximate causes and evolutionary consequences. *Annals of botany* **115**, 769-779.

Elmer KR, Fan S, Kusche H*, et al.* (2014) Parallel evolution of Nicaraguan crater lake cichlid fishes via non-parallel routes. *Nature communications* **5**.

Elmer KR, Kusche H, Lehtonen TK, Meyer A (2010) Local variation and parallel evolution: morphological and genetic diversity across a species complex of neotropical crater lake cichlid fishes. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 1763-1782.

Elmer KR, Reggio C, Wirth T*, et al.* (2009) Pleistocene desiccation in East Africa bottlenecked but did not extirpate the adaptive radiation of Lake Victoria haplochromine cichlid fishes. *Proceedings of the National Academy of Sciences* **106**, 13404-13409.

Emlen DJ (1994) Environmental control of horn length dimorphism in the beetle Onthophagus acuminatus (Coleoptera: Scarabaeidae). *Proceedings of the Royal Society of London. Series B: Biological Sciences* **256**, 131-136.

Endler JA (1992) Signals, signal conditions, and the direction of evolution. *American Naturalist*, S125-S153.

Engler AJ, Sen S, Sweeney HL, Discher DE (2006) Matrix elasticity directs stem cell lineage specification. *Cell* **126**, 677-689.

Escobar‐Camacho D, Ramos E, Martins C, Carleton KL (2016) The Opsin Genes of Amazonian Cichlids. *Molecular ecology*.

Espinosa-Soto C, Martin OC, Wagner A (2011) Phenotypic plasticity can facilitate adaptive evolution in gene regulatory circuits. *BMC Evolutionary Biology* **11**, 5.

Fernald RD (1981) Chromatic organization of a cichlid fish retina. *Vision research* **21**, 1749-1753.

Fernald RD (1984) Vision and Behavior in an African Cichlid fish: Combining behavioral and physiological analyses reveals how good vision is maintained during rapid growth of the eyes. *American Scientist* **72**, 58-65.

Ferrari G, Angelis D, Coletta M*, et al.* (1998) Muscle regeneration by bone marrow-derived myogenic progenitors. *Science* **279**, 1528-1530.

Fierst J (2011) A history of phenotypic plasticity accelerates adaptation to a new environment. *Journal of Evolutionary Biology* **24**, 1992-2001.

Finger E, Burkhardt D (1994) Biological aspects of bird colouration and avian colour vision including ultraviolet range. *Vision research* **34**, 1509-1514.

Fisher RA (1930) *The genetical theory of natural selection: a complete variorum edition* Oxford University Press, Oxford.

Foster S, Wund M, Graham M*, et al.* (2015) Iterative development and the scope for plasticity: contrasts among trait categories in an adaptive radiation. *Heredity* **115**, 335-348.

Francis RC, Barlow GW (1993) Social control of primary sex differentiation in the Midas cichlid. *Proceedings of the National Academy of Sciences* **90**, 10673-10675.

Friedman M, Keck BP, Dornburg A*, et al.* (2013) Molecular and fossil evidence place the origin of cichlid fishes long after Gondwanan rifting **280**, 20131733.

Fryer G, Iles TD (1972) The Cichlid fishes of the Great Lakes of Africa: their biology and evolution. *Oliver & Boyd, Croythorn House, 23 Ravelston Terrace, Edinburgh EH 4 3 TJ, UK. 656*, 1972.

Fuller R, Carleton K, Fadool J, Spady T, Travis J (2004) Population variation in opsin expression in the bluefin killifish, Lucania goodei: a real-time PCR study. *Journal of Comparative Physiology A* **190**, 147-154.

Galis F (2001) Key innovations and radiations.

Gavrilets S (2003) Perspective: models of speciation: what have we learned in 40 years? *Evolution* **57**, 2197-2215.

Gelse K, Pöschl E, Aigner T (2003) Collagens—structure, function, and biosynthesis. *Advanced drug delivery reviews* **55**, 1531-1546.

Genner MJ, Seehausen O, Lunt DH*, et al.* (2007) Age of cichlids: new dates for ancient lake fish radiations. *Molecular biology and evolution* **24**, 1269-1282.

Genner MJ, Turner GF (2005) The mbuna cichlids of Lake Malawi: a model for rapid speciation and adaptive radiation. *Fish and Fisheries* **6**, 1-34.

Ghalambor CK, Hoke KL, Ruell EW*, et al.* (2015) Non-adaptive plasticity potentiates rapid adaptive evolution of gene expression in nature. *Nature*.

Ghalambor CK, McKay JK, Carroll SP, Reznick DN (2007) Adaptive versus non‐adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional ecology* **21**, 394-407.

Gibson G (2008) The environmental contribution to gene expression profiles. *Nature Reviews Genetics* **9**, 575-581.

Gilbert SF (2001) Ecological developmental biology: developmental biology meets the real world. *Developmental biology* **233**, 1-12.

Gilbert SF (2005) Mechanisms for the environmental regulation of gene expression: ecological aspects of animal development. *Journal of biosciences* **30**, 65-74.

Gilbert SF, Bosch TC, Ledón-Rettig C (2015) Eco-Evo-Devo: developmental symbiosis and developmental plasticity as evolutionary agents. *Nature Reviews Genetics* **16**, 611-622.

Gilbert SF, Epel D (2009) *Ecological developmental biology* Sinauer Associates Sunderland.

Gilbert SF, Epel D (2015) *Ecological developmental biology: the environmental regulation of development, health, and evolution* Sinauer Associates, Incorporated.

Gittenberger E (1991) What about non‐adaptive radiation? *Biological Journal of the Linnean Society* **43**, 263-272.

Gomez-Mestre I, Jovani R (2013) A heuristic model on the role of plasticity in adaptive evolution: plasticity increases adaptation, population viability and genetic variation **280**, 20131869.

Gosline WA (1980) The evolution of some structural systems with reference to the interrelationships of modern lower teleostean fish groups. *Japanese Journal of Ichthyology*.

Govardovskii VI, Fyhrquist N, Reuter T, Kuzmin DG, Donner K (2000) In search of the visual pigment template. *Visual neuroscience* **17**, 509-528.

Grabherr MG, Haas BJ, Yassour M*, et al.* (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology* **29**, 644-652.

Greenwood P (1959) The monotypic genera of cichlid fishes in Lake Victoria, Part II. *Bulletin of the British Museum (Natural History), Zoology* **5**, 165-177.

Greenwood P (1964) Environmental effects on the pharyngeal mill of a cichlid fish, *Astatoreochromis alluaudi*, and their taxonomic implications. *Proceedings of the Linnean Society of London* **176**, 1-10.

Greenwood P (1984) African cichlids and evolutionary theories.

Greibe E, Fedosov S, Nexo E (2012) The cobalamin-binding protein in zebrafish is an intermediate between the three cobalamin-binding proteins in human. *PLOS one* **7**, e35660.

Grudzien TA, Turner BJ (1984) Direct evidence that the Ilyodon morphs are a single biological species. *Evolution*, 402-407.

Gunter HM, Fan S, Xiong F*, et al.* (2013) Shaping development through mechanical strain: the transcriptional basis of diet-induced phenotypic plasticity in a cichlid fish. *Molecular Ecology* **22**, 4516–4531.

Gunter HM, Meyer A (2014) Molecular investigation of mechanical strain‐induced phenotypic plasticity in the ecologically important pharyngeal jaws of cichlid fish. *Journal of Applied Ichthyology* **30**, 630-635.

Gutteling E, Riksen J, Bakker J, Kammenga J (2007) Mapping phenotypic plasticity and genotype–environment interactions affecting life-history traits in Caenorhabditis elegans. *Heredity* **98**, 28-37.

Haas BJ, Papanicolaou A, Yassour M*, et al.* (2013) De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature protocols* **8**, 1494-1512.

Hadjidakis DJ, Androulakis II (2006) Bone remodeling. *Annals of the New York Academy of Sciences* **1092**, 385-396.

Hallikas O, Palin K, Sinjushina N*, et al.* (2006) Genome-wide prediction of mammalian enhancers based on analysis of transcription-factor binding affinity. *Cell* **124**, 47-59.

Harlin-Cognato A, Hoffman EA, Jones AG (2006) Gene cooption without duplication during the evolution of a male-pregnancy gene in pipefish. *Proceedings of the National Academy of Sciences* **103**, 19407-19412.

Hárosi FI (1994) An analysis of two spectral properties of vertebrate visual pigments. *Vision research* **34**, 1359-1367.

Harris J (1936) The role of the fins in the equilibrium of the swimming fish. *I.-J. exper. Biol* **13**, 476-493.

Harris R (2007) Improved pairwise alignment of genomic DNA [PhD thesis].[University Park (PA)]: The Pennsylvania State University.

Hauser J, Saarikettu J, Grundström T (2008) Calcium regulation of myogenesis by differential calmodulin inhibition of basic helix-loop-helix transcription factors. *Molecular biology of the cell* **19**, 2509-2519.

Hayden EJ, Ferrada E, Wagner A (2011) Cryptic genetic variation promotes rapid evolutionary adaptation in an RNA enzyme. *Nature* **474**, 92-95.

Hedrick PW (2013) Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Molecular Ecology* **22**, 4606-4618.

Hendry AP (2015) Key questions on the role of phenotypic plasticity in eco-evolutionary dynamics. *Journal of Heredity*, esv060.

Henning F, Meyer A (2014) The evolutionary genomics of cichlid fishes: explosive speciation and adaptation in the postgenomic era. *Annual review of genomics and human genetics* **15**, 417-441.

Hilbrant M, Almudi I, Leite DJ*, et al.* (2014) Sexual dimorphism and natural variation within and among species in the Drosophila retinal mosaic. *BMC Evolutionary Biology* **14**, 240.

Hinman VF, Yankura KA, McCauley BS (2009) Evolution of gene regulatory network architectures: examples of subcircuit conservation and plasticity between classes of echinoderms. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* **1789**, 326-332.

Hoekstra HE, Coyne JA (2007) The locus of evolution: evo devo and the genetics of adaptation. *Evolution* **61**, 995-1016.

Hofmann CM, O'Quin KE, Marshall NJ*, et al.* (2009) The eyes have it: regulatory and structural changes both underlie cichlid visual pigment diversity. *PLoS Biol* **7**, e1000266.

Hofmann CM, O'Quin KE, Marshall NJ, Carleton KL (2010) The relationship between lens transmission and opsin gene expression in cichlids from Lake Malawi. *Vision research* **50**, 357-363.

Hölldobler B (1990) *The ants* Harvard University Press.

Hoogerhoud R (1984) A taxonomic reconsideration of the haplochromine genera Gaurochromis Greenwood, 1980 and Labrochromis Regan, 1920 (Pisces, Cichlidae). *Netherlands Journal of Zoology* **34**, 539-565.

Hoogerhoud R (1986a) The Adverse Effects of Shell Ingestion for Molluscivorous Cichlids, a Constructional Morphological Approach. *Netherlands Journal of Zoology* **37**, 3-4.

Hoogerhoud R (1986b) Taxonomic and ecological aspects of morphological plasticity in molluscivorous haplochromines (Pisces, Cichlidae). *Annalen-Koninklijk Museum voor Midden Afrika. Zoologische Wetenschappen*.

Hoogerhoud R, Witte F, Barel C (1982) The ecological differentiation of two closely resembling Haplochromis species from Lake Victoria (H. iris and H. hiatus; Pisces, Cichlidae). *Netherlands Journal of Zoology* **33**, 283-305.

Hoogerhoud RJC (1986c) *Ecological morphology of some cichlid fishes*, RJC Hoogerhoud.

Hubisz MJ, Pollard KS, Siepel A (2010) PHAST and RPHAST: phylogenetic analysis with space/time models. *Briefings in bioinformatics*, bbq072.

Hulsey CD (2006) Function of a key morphological innovation: fusion of the cichlid pharyngeal jaw. *Proceedings of the Royal Society B: Biological Sciences* **273**, 669-675.

Hulsey CD, Hendrickson DA, De León FJG (2005) Trophic morphology, feeding performance and prey use in the polymorphic fish Herichthys minckleyi. *Evolutionary Ecology Research* **7**, 303-324.

Hunt BG, Ometto L, Wurm Y*, et al.* (2011) Relaxed selection is a precursor to the evolution of phenotypic plasticity. *Proceedings of the National Academy of Sciences* **108**, 15936-15941.

Hunter JP (1998) Key innovations and the ecology of macroevolution. *Trends in Ecology & Evolution* **13**, 31-36.

Huysseune A (1995) Phenotypic plasticity in the lower pharyngeal jaw dentition of *Astatoreochromis alluaudi* (Teleostei: cichlidae). *Archives of Oral Biology* **40**, 1005-1014.

Huysseune A, Sire J, Meunier F (1994) Comparative study of lower pharyngeal jaw structure in two phenotypes of *Astatoreochromis alluaudi* (Teleostei: Cichlidae). *Journal of Morphology* **221**, 25-43.

Indjeian VB, Kingman GA, Jones FC*, et al.* (2016) Evolving new skeletal traits by cis-regulatory changes in bone morphogenetic proteins. *Cell* **164**, 45-56.

Iwasaki WM, Tsuda ME, Kawata M (2013) Genetic and environmental factors affecting cryptic variations in gene regulatory networks. *BMC Evolutionary Biology* **13**, 1.

Izu Y, Sun M, Zwolanek D*, et al.* (2011) Type XII collagen regulates osteoblast polarity and communication during bone formation. *The Journal of cell biology* **193**, 1115-1130.

Jablonka E, Lamb MJ (2007) The expanded evolutionary synthesis—a response to Godfrey-Smith, Haig, and West-Eberhard. *Biology & Philosophy* **22**, 453-472.

Jablonka E, Raz G (2009) Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *The Quarterly Review of Biology* **84**, 131-176.

Jacobs GH (1993) The distribution and nature of colour vision among the mammals. *Biological Reviews* **68**, 413-471.

Janzen FJ, Paukstis GL (1991) Environmental sex determination in reptiles: ecology, evolution, and experimental design. *Quarterly Review of Biology*, 149-179.

Jarvela AMC, Hinman VF (2015) Evolution of transcription factor function as a mechanism for changing metazoan developmental gene regulatory networks. *EvoDevo* **6**, 3.

Jordan R, Kellogg K, Howe D*, et al.* (2006) Photopigment spectral absorbance of Lake Malaŵi cichlids. *Journal of Fish Biology* **68**, 1291-1299.

Kalb N, Schneider RF, Sprenger D, Michiels NK (2015) The Red‐Fluorescing Marine Fish Tripterygion delaisi can Perceive its Own Red Fluorescent Colour. *Ethology* **121**, 566-576.

Karttunen JT, Lehner PJ, Gupta SS, Hewitt EW, Cresswell P (2001) Distinct functions and cooperative interaction of the subunits of the transporter associated with antigen processing (TAP). *Proceedings of the National Academy of Sciences* **98**, 7431-7436.

Kawaguchi M, Tomita K, Sano K, Kaneko T (2015) Molecular events in adaptive evolution of the hatching strategy of ovoviviparous fishes. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* **324**, 41-50.

Kawaguchi M, Yasumasu S, Hiroi J*, et al.* (2006) Evolution of teleostean hatching enzyme genes and their paralogous genes. *Development Genes and Evolution* **216**, 769-784.

Kawasaki K (2011) The SCPP gene family and the complexity of hard tissues in vertebrates. *Cells Tissues Organs* **194**, 108-112.

Kawasaki K, Amemiya CT (2014) SCPP genes in the coelacanth: tissue mineralization genes shared by sarcopterygians. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* **322**, 390-402.

Kawecki TJ (1997) Sympatric speciation via habitat specialization driven by deleterious mutations. *Evolution*, 1751-1763.

Keenleyside MH (1991) *Cichlid fishes: behaviour, ecology and evolution* Springer Science & Business Media.

Kelber A, Osorio D (2010) From spectral information to animal colour vision: experiments and concepts. *Proceedings of the Royal Society of London B: Biological Sciences* **277**, 1617-1625.

Kelber A, Vorobyev M, Osorio D (2003) Animal colour vision–behavioural tests and physiological concepts. *Biological Reviews* **78**, 81-118.

Kent WJ (2002) BLAT—the BLAST-like alignment tool. *Genome research* **12**, 656-664.

Kerschbaumer M, Postl L, Koch M, Wiedl T, Sturmbauer C (2011) Morphological distinctness despite large-scale phenotypic plasticity—analysis of wild and pond-bred juveniles of allopatric populations of Tropheus moorii. *Naturwissenschaften* **98**, 125-134.

Kijimoto T, Costello J, Tang Z, Moczek AP, Andrews J (2009) EST and microarray analysis of horn development in Onthophagus beetles. *BMC genomics* **10**, 1.

Kijimoto T, Moczek AP, Andrews J (2012) Diversification of doublesex function underlies morph-, sex-, and species-specific development of beetle horns. *Proceedings of the National Academy of Sciences* **109**, 20526-20531.

King M-C, Wilson AC (1975) *Evolution at two levels in humans and chimpanzees* na.

Kishe-Machumu M, Witte F, Wanink JH (2008) Dietary shift in benthivorous cichlids after the ecological changes in Lake Victoria. *Animal Biology* **58**, 401-417.

Koblmüller S, Sefc KM, Sturmbauer C (2008) The Lake Tanganyika cichlid species assemblage: recent advances in molecular phylogenetics. In: *Patterns and Processes of Speciation in Ancient Lakes*, pp. 5-20. Springer.

Kocher TD (2004) Adaptive evolution and explosive speciation: the cichlid fish model. *Nature Reviews Genetics* **5**, 288-298.

Kocher TD, Conroy JA, McKaye KR, Stauffer JR (1993) Similar morphologies of cichlid fish in Lakes Tanganyika and Malawi are due to convergence. *Molecular phylogenetics and evolution* **2**, 158-165.

Kodric-Brown A (1985) Female preference and sexual selection for male coloration in the guppy (Poecilia reticulata). *Behavioral Ecology and Sociobiology* **17**, 199-205.

Kolbe JJ, Losos JB (2005) Hind-limb length plasticity in Anolis carolinensis. *Journal of Herpetology* **39**, 674-678.

Kondrashov AS, Kondrashov FA (1999) Interactions among quantitative traits in the course of sympatric speciation. *Nature* **400**, 351-354.

Konings A (1990) Konings's book of cichlids and all the other fishes of Lake Malawi.

Kuiter RH (2000) *Seahorses, pipefishes and their relatives: a comprehensive guide to Syngnathiformes* Twayne Publishers.

Kusche H, Recknagel H, Elmer KR, Meyer A (2014) Crater lake cichlids individually specialize along the benthic–limnetic axis. *Ecology and Evolution* **4**, 1127-1139.

Kwon AT, Arenillas DJ, Hunt RW, Wasserman WW (2012) oPOSSUM-3: advanced analysis of regulatory motif over-representation across genes or ChIP-Seq datasets. *G3: Genes| Genomes| Genetics* **2**, 987-1002.

LaBarge MA, Blau HM (2002) Biological progression from adult bone marrow to mononucleate muscle stem cell to multinucleate muscle fiber in response to injury. *Cell* **111**, 589-601.

Lahti DC, Johnson NA, Ajie BC*, et al.* (2009) Relaxed selection in the wild. *Trends in Ecology & Evolution* **24**, 487-496.

Lande R (2009) Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *Journal of Evolutionary Biology* **22**, 1435-1446.

Lande R, Arnold SJ (1983) The measurement of selection on correlated characters. *Evolution*, 1210-1226.

Landry CR, Aubin-Horth N (2014) Recent Advances in Ecological Genomics: From Phenotypic Plasticity to Convergent and Adaptive Evolution and Speciation. In: *Ecological Genomics*, pp. 1-5. Springer.

Le Vinh TJ, Vandenbrooks J, Angilletta Jr M (2016) Developmental plasticity evolved according to specialist-generalist trade-offs in experimental populations of Drosophila melanogaster. *Biology letters* **12**.

Leberer E, Charuk J, Green NM, Maclennan DH (1989) Molecular cloning and expression of cDNA encoding a lumenal calcium binding glycoprotein from sarcoplasmic reticulum. *Proceedings of the National Academy of Sciences* **86**, 6047-6051.

Lee PN, Callaerts P, de Couet HG, Martindale MQ (2003) Cephalopod Hox genes and the origin of morphological novelties. *Nature* **424**, 1061-1065.

Leichty AR, Pfennig DW, Jones CD, Pfennig KS (2012) Relaxed genetic constraint is ancestral to the evolution of phenotypic plasticity. *Integrative and Comparative Biology* **52**, 16-30.

Levine JS, MacNichol EF (1979) Visual pigments in teleost fishes: effects of habitat, microhabitat, and behavior on visual system evolution. *Sensory processes*.

Levis NA, Pfennig DW (2016) Evaluating 'plasticity-first'evolution in nature: key criteria and empirical approaches. *Trends in Ecology & Evolution*.

Lewontin RC (1974) *The genetic basis of evolutionary change* Columbia University Press New York.

Leysen H, Christiaens J, De Kegel B*, et al.* (2011) Musculoskeletal structure of the feeding system and implications of snout elongation in Hippocampus reidi and Dunckerocampus dactyliophorus. *Journal of Fish Biology* **78**, 1799-1823.

Li H, Handsaker B, Wysoker A*, et al.* (2009a) The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078-2079.

Li N, Felber K, Elks P, Croucher P, Roehl HH (2009b) Tracking gene expression during zebrafish osteoblast differentiation. *Developmental Dynamics* **238**, 459-466.

Li Y, Álvarez OA, Gutteling EW*, et al.* (2006) Mapping determinants of gene expression plasticity by genetical genomics in C. elegans. *PLoS Genet* **2**, e222.

Liem KF (1974) Evolutionary strategies and morphological innovations: cichlid pharyngeal jaws. *Systematic Biology* **22**, 425-441.

Liem KF (1986) The pharyngeal jaw apparatus of the Embiotocidae (Teleostei): a functional and evolutionary perspective. *Copeia*, 311-323.

Lin Q, Fan S, Zhang Y*, et al.* (2016) The seahorse genome and the evolution of its specialized morphology. *Nature* **540**, 395-399.

Lindblad-Toh K, Garber M, Zuk O*, et al.* (2011) A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* **478**, 476-482.

Loew ER, McFarland WN (1990) The underwater visual environment. In: *The visual system of fish*, pp. 1-43. Springer.

Lopez-Fernández H, Honeycutt RL, Winemiller KO (2005) Molecular phylogeny and evidence for an adaptive radiation of geophagine cichlids from South America (Perciformes: Labroidei). *Molecular phylogenetics and evolution* **34**, 227-244.

López-Fernández H, Winemiller KO, Montaña C, Honeycutt RL (2012) Diet-morphology correlations in the radiation of South American geophagine cichlids (Perciformes: Cichlidae: Cichlinae). *PLOS one* **7**, e33997.

Losos JB (2010) Adaptive radiation, ecological opportunity, and evolutionary determinism. *The American Naturalist* **175**, 623-639.

Losos JB, Creer DA, Glossip D*, et al.* (2000) Evolutionary implications of phenotypic plasticity in the hindlimb of the lizard Anolis sagrei. *Evolution* **54**, 301-305.

Louchart A, Viriot L (2011) From snout to beak: the loss of teeth in birds. *Trends in Ecology & Evolution* **26**, 663-673.

Lu M-F, Cheng H-T, Kern MJ*, et al.* (1999) prx-1 functions cooperatively with another paired-related homeobox gene, prx-2, to maintain cell fates within the craniofacial mesenchyme. *Development* **126**, 495-504.

Lucek K, Sivasundar A, Seehausen O (2014) Disentangling the role of phenotypic plasticity and genetic divergence in contemporary ecotype formation during a biological invasion. *Evolution* **68**, 2619-2632.

Lundsgaard‐Hansen B, Matthews B, Vonlanthen P, Taverna A, Seehausen O (2013) Adaptive plasticity and genetic divergence in feeding efficiency during parallel adaptive radiation of whitefish (Coregonus spp.). *Journal of Evolutionary Biology* **26**, 483-498.

Luo R, Liu B, Xie Y*, et al.* (2012) SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience* **1**, 1.

Lythgoe JN (1988) Light and vision in the aquatic environment. In: *Sensory biology of aquatic animals*, pp. 57-82. Springer.

Maan ME, Hofker KD, van Alphen JJ, Seehausen O (2006) Sensory drive in cichlid speciation. *The American Naturalist* **167**, 947-954.

Maan ME, Seehausen O (2010) Mechanisms of species divergence through visual adaptation and sexual selection: Perspectives from a cichlid model sys-tem.

Maan ME, Seehausen O, Van Alphen JJ (2010) Female mating preferences and male coloration covary with water transparency in a Lake Victoria cichlid fish. *Biological Journal of the Linnean Society* **99**, 398-406.

MacArthur DG, Seto JT, Raftery JM*, et al.* (2007) Loss of ACTN3 gene function alters mouse muscle metabolism and shows evidence of positive selection in humans. *Nature genetics* **39**, 1261-1265.

Machado‐Schiaffino G, Henning F, Meyer A (2014) Species‐specific differences in adaptive phenotypic plasticity in an ecologically relevant trophic trait: hypertrophic lips in Midas cichlid fishes. *Evolution* **68**, 2086-2091.

Manousaki T, Hull PM, Kusche H*, et al.* (2013) Parsing parallel evolution: ecological divergence and differential gene expression in the adaptive radiations of thick‐lipped Midas cichlid fishes from Nicaragua. *Molecular Ecology* **22**, 650-669.

Mantila Roosa SM, Liu Y, Turner CH (2011) Gene expression patterns in bone following mechanical loading. *Journal of Bone and Mineral Research* **26**, 100-112.

Maruska KP, Levavi-Sivan B, Biran J, Fernald RD (2011) Plasticity of the reproductive axis caused by social status change in an African cichlid fish: I. pituitary gonadotropins. *Endocrinology* **152**, 281-290.

Masel J, King OD, Maughan H (2007) The loss of adaptive plasticity during long periods of environmental stasis. *The American Naturalist* **169**, 38.

Mayo AE, Setty Y, Shavit S, Zaslaver A, Alon U (2006) Plasticity of the cis-regulatory input function of a gene. *PLoS biology* **4**, 555.

Mayr E (1963) *Animal species and evolution* Belknap Press.

Mazzarella AB, Voje KL, Hansson TH, Taugbøl A, Fischer B (2015) Strong and parallel salinity‐induced phenotypic plasticity in one generation of threespine stickleback. *Journal of Evolutionary Biology*.

Mbabazi D, Ogutu‐Ohwayo R, Wandera S, Kiziito Y (2004) Fish species and trophic diversity of haplochromine cichlids in the Kyoga satellite lakes (Uganda). *African Journal of Ecology* **42**, 59-68.

McBeath R, Pirone DM, Nelson CM, Bhadriraju K, Chen CS (2004) Cell shape, cytoskeletal tension, and RhoA regulate stem cell lineage commitment. *Developmental cell* **6**, 483-495.

McLean CY, Bristor D, Hiller M*, et al.* (2010) GREAT improves functional interpretation of cis-regulatory regions. *Nature biotechnology* **28**, 495-501.

McLean CY, Reno PL, Pollen AA*, et al.* (2011) Human-specific loss of regulatory DNA and the evolution of human-specific traits. *Nature* **471**, 216-219.

Mei J, Zhang Q-Y, Li Z, Lin S, Gui J-F (2008) C1q-like inhibits p53-mediated apoptosis and controls normal hematopoiesis during zebrafish embryogenesis. *Developmental biology* **319**, 273-284.

Meissner G (1994) Ryanodine receptor/Ca2+ release channels and their regulation by endogenous effectors. *Annual Review of Physiology* **56**, 485-508.

Meredith RW, Zhang G, Gilbert MTP, Jarvis ED, Springer MS (2014) Evidence for a single loss of mineralized teeth in the common avian ancestor. *Science* **346**, 1254390.

Meyer A (1986) Changes in behavior with increasing experience with a novel prey in fry of the Central American cichlid, Cichlasoma managuense (Teleostei: Cichlidae). *Behaviour* **98**, 145-167.

Meyer A (1987a) First feeding success with two types of prey by the Central American cichlid fish, Cichlasoma managuense (Pisces, Cichlidae): morphology versus behavior. *Environmental Biology of Fishes* **18**, 127-134.

Meyer A (1987b) Phenotypic plasticity and heterochrony in Cichlasoma managuense (Pisces, Chichlidae) and their implications for speciation in Cichlid fishes. *Evolution* **41**, 1357-1369.

Meyer A (1989) Cost of morphological specialization: feeding performance of the two morphs in the trophically polymorphic cichlid fish, Cichlasoma citrinellum. *Oecologia* **80**, 431-436.

Meyer A (1990) Ecological and evolutionary consequences of the trophic polymorphism in Cichlasoma citrinellum (Pisces: Cichlidae). *Biological Journal of the Linnean Society* **39**, 279-299.

Meyer A (1993a) Phylogenetic relationships and evolutionary processes in East African cichlid fishes. *Trends in Ecology & Evolution* **8**, 279-284.

Meyer A (1993b) Trophic polymorphisms in cichlid fish: do they represent intermediate steps during sympatric speciation and explain their rapid adaptive radiation. *Trends in Ichthyology* **7**, 257-266.

Meyer A, Kocher TD, Basasibwaki P, Wilson AC (1990) Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. *Nature* **347**, 550-553.

Michikami I, Fukushi T, Tanaka M*, et al.* (2012) Krüppel-like factor 4 regulates membranous and endochondral ossification. *Experimental Cell Research* **318**, 311-325.

Miller AH (1949) Some ecologic and morphologic considerations in the evolution of higher taxonomic categories. *Ornithologie als biologische Wissenschaft*, 84-88.

Mirouze M, Paszkowski J (2011) Epigenetic contribution to stress adaptation in plants. *Current opinion in plant biology* **14**, 267-274.

Mo C, Romero-Suarez S, Bonewald L, Johnson M, Brotto M (2012) Prostaglandin E2: from clinical applications to its potential role in bone-muscle crosstalk and myogenic differentiation. *Recent patents on biotechnology* **6**, 223.

Moczek A, Emlen D (1999) Proximate determination of male horn dimorphism in the beetle Ontophagus taurus (Coleoptera: Scarabaeidae). *Journal of Evolutionary Biology* **12**, 27-37.

Moczek AP, Kijimoto T (2014) Development and evolution of insect polyphenisms: novel insights through the study of sex determination mechanisms. *Current Opinion in Insect Science* **1**, 52-58.

Morgan HD, Sutherland HG, Martin DI, Whitelaw E (1999) Epigenetic inheritance at the agouti locus in the mouse. *Nature genetics* **23**, 314-318.

Morgan L (1947) A variable phenotype associated with the fourth chromosome of Drosophila melanogaster and affected by heterochromatin. *Genetics* **32**, 200.

Morris MR, Richard R, Leder EH*, et al.* (2014) Gene expression plasticity evolves in response to colonization of freshwater lakes in threespine stickleback. *Molecular Ecology* **23**, 3226-3240.

Murren CJ, Auld JR, Callahan H*, et al.* (2015) Constraints on the evolution of phenotypic plasticity: limits and costs of phenotype and plasticity. *Heredity* **115**, 293-301.

Muschick M, Barluenga M, Salzburger W, Meyer A (2011) Adaptive phenotypic plasticity in the Midas cichlid fish pharyngeal jaw and its relevance in adaptive radiation. *BMC Evolutionary Biology* **11**, 116.

Muschick M, Indermaur A, Salzburger W (2012) Convergent evolution within an adaptive radiation of cichlid fishes. *Current Biology* **22**, 2362–2368.

Naiche L, Papaioannou VE (2003) Loss of Tbx4 blocks hindlimb development and affects vascularization and fusion of the allantois. *Development* **130**, 2681-2693.

Nair AK, Gautieri A, Chang S-W, Buehler MJ (2013) Molecular mechanics of mineralized collagen fibrils in bone. *Nature communications* **4**, 1724.

Nakashima K, Zhou X, Kunkel G*, et al.* (2002) The novel zinc finger-containing transcription factor osterix is required for osteoblast differentiation and bone formation. *Cell* **108**, 17-29.

Navratilova P, Fredman D, Hawkins TA*, et al.* (2009) Systematic human/zebrafish comparative identification of cis-regulatory activity around vertebrate developmental transcription factor genes. *Developmental biology* **327**, 526-540.

Near TJ, Dornburg A, Eytan RI*, et al.* (2013) Phylogeny and tempo of diversification in the superradiation of spiny-rayed fishes. *Proceedings of the National Academy of Sciences* **110**, 12738-12743.

Nei M, Kumar S (2000) *Molecular evolution and phylogenetics* Oxford university press.

Nei M, Maruyama T, Wu C-I (1983) Models of evolution of reproductive isolation. *Genetics* **103**, 557-579.

Newman RA (1992) Adaptive plasticity in amphibian metamorphosis. *Bioscience*, 671-678.

Nicotra AB, Atkin OK, Bonser SP*, et al.* (2010) Plant phenotypic plasticity in a changing climate. *Trends in plant science* **15**, 684-692.

Nijhout HF (2003) Development and evolution of adaptive polyphenisms. *Evolution & Development* **5**, 9-18.

Nijhout HF (2015) A Developmental-Physiological Perspective on the Development and Evolution of Phenotypic Plasticity. In: *Conceptual Change in Biology*, pp. 147-173. Springer.

Nomura S, Takano-Yamamoto T (2000) Molecular events caused by mechanical stress in bone. *Matrix Biology* **19**, 91-96.

Nonaka MI, Nonaka M (2010) Evolutionary analysis of two classical MHC class I loci of the medaka fish, Oryzias latipes: haplotype-specific genomic diversity, locus-specific polymorphisms, and interlocus homogenization. *Immunogenetics* **62**, 319-332.

Nordeng H (1983) Solution to the" char problem" based on Arctic char (Salvelinus alpinus) in Norway. *Canadian Journal of Fisheries and Aquatic Sciences* **40**, 1372-1387.

Nose K, Shibanuma M (1994) Induction of early response genes by hypergravity in cultured mouse osteoblastic cells (MC3T3-E1). *Experimental Cell Research* **211**, 168-170.

Nosil P (2002) Transition rates between specialization and generalization in phytophagous insects. *Evolution* **56**, 1701-1706.

Nussey DH, Postma E, Gienapp P, Visser ME (2005) Selection on heritable phenotypic plasticity in a wild bird population. *Science* **310**, 304-306.

Nylin S, Gotthard K (1998) Plasticity in life-history traits. *Annual review of entomology* **43**, 63-83.

Nylin S, Wahlberg N (2008) Does plasticity drive speciation? Host-plant shifts and diversification in nymphaline butterflies (Lepidoptera: Nymphalidae) during the tertiary. *Biological Journal of the Linnean Society* **94**, 115-130.

O'Quin KE, Hofmann CM, Hofmann HA, Carleton KL (2010) Parallel evolution of opsin gene expression in African cichlid fishes. *Molecular biology and evolution* **27**, 2839-2854.

Odling-Smee FJ, Laland KN, Feldman MW (1996) Niche construction. *The American Naturalist* **147**, 641-648.

Ohno S (1970) The enormous diversity in genome sizes of fish as a reflection of nature's extensive experiments with gene duplication. *Transactions of the American Fisheries Society* **99**, 120-130.

Ohno S (2013) *Evolution by gene duplication* Springer Science & Business Media.

Oke KB, Bukhari M, Kaeuffer R*, et al.* (2015) Does plasticity enhance or dampen phenotypic parallelism? A test with three lake–stream stickleback pairs. *Journal of Evolutionary Biology*.

Orizaola G, Laurila A (2016) Developmental plasticity increases at the northern range margin in a warm‐dependent amphibian. *Evolutionary applications* **9**, 471-478.

Ostwald W (1904) Experimentelle Untersuchungen über den Saisonpolymorphismus bei Daphniden. *Development Genes and Evolution* **18**, 415-451.

Ott CE, Bauer S, Manke T*, et al.* (2009) Promiscuous and Depolarization‐Induced Immediate‐Early Response Genes Are Induced by Mechanical Strain of Osteoblasts. *Journal of Bone and Mineral Research* **24**, 1247-1262.

Paaby AB, Rockman MV (2014) Cryptic genetic variation: evolution's hidden substrate. *Nature Reviews Genetics* **15**, 247-258.

Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289-290.

Parry JW, Carleton KL, Spady T*, et al.* (2005) Mix and match color vision: tuning spectral sensitivity by differential opsin gene expression in Lake Malawi cichlids. *Current Biology* **15**, 1734-1739.

Parsons KJ, Concannon M, Navon D*, et al.* (2016) Foraging environment determines the genetic architecture and evolutionary potential of trophic morphology in cichlid fishes. *Molecular Ecology*.

Perry S (1998) Troponin T: genetics, properties and function. *Journal of Muscle Research & Cell Motility* **19**, 575-602.

Peter IS, Davidson EH (2011) Evolution of gene regulatory networks controlling body plan development. *Cell* **144**, 970-985.

Peters DG, Zhang X-C, Benos PV, Heidrich-O Hare E, Ferrell RE (2003) Genomic analysis of immediate/early response to shear stress in human coronary artery endothelial cells. *Physiological genomics* **12**, 25-34.

Peterson KJ, Davidson EH (2000) Regulatory evolution and the origin of the bilaterians. *Proceedings of the National Academy of Sciences* **97**, 4430-4433.

Pfennig DW, Ehrenreich IM (2014) Towards a gene regulatory network perspective on phenotypic plasticity, genetic accommodation and genetic assimilation. *Molecular Ecology* **23**, 4438-4440.

Pfennig DW, McGee M (2010) Resource polyphenism increases species richness: a test of the hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 577-591.

Pfennig DW, Murphy PJ (2000) Character displacement in polyphenic tadpoles. *Evolution* **54**, 1738-1749.

Pfennig DW, Murphy PJ, Wiens J (2002) How fluctuating competition and phenotypic plasticity mediate species divergence. *Evolution* **56**, 1217-1228.

Pfennig DW, Pfennig KS (2012) *Evolution's wedge: competition and the origins of diversity* Univ of California Press.

Pfennig DW, Reeve HK, Sherman PW (1993) Kin recognition and cannibalism in spadefoot toad tadpoles. *Animal Behaviour* **46**, 87-94.

Pfennig DW, Wund MA, Snell-Rood EC*, et al.* (2010) Phenotypic plasticity's impacts on diversification and speciation. *Trends in Ecology & Evolution* **25**, 459-467.

Piálek L, Říčan O, Casciotta J, Almirón A, Zrzavý J (2012) Multilocus phylogeny of Crenicichla (Teleostei: Cichlidae), with biogeography of the C. lacustris group: species flocks as a model for sympatric speciation in rivers. *Molecular phylogenetics and evolution* **62**, 46-61.

Pigliucci M (2007) Do we need an extended evolutionary synthesis? *Evolution* **61**, 2743-2749.

Pigliucci M, Müller GB (2010) Elements of an extended evolutionary synthesis. *Evolution: The extended synthesis*, 3-17.

Pigliucci M, Murren CJ (2003) Perspective: Genetic assimilation and a possible evolutionary paradox: can macroevolution sometimes be so fast as to pass us by? *Evolution* **57**, 1455-1464.

Pigliucci M, Murren CJ, Schlichting CD (2006) Phenotypic plasticity and evolution by genetic assimilation. *Journal of Experimental Biology* **209**, 2362-2367.

Price TD, Qvarnström A, Irwin DE (2003) The role of phenotypic plasticity in driving genetic evolution. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **270**, 1433-1440.

Prud'homme B, Gompel N, Carroll SB (2007) Emerging principles of regulatory evolution. *Proceedings of the National Academy of Sciences* **104**, 8605-8612.

Queitsch C, Sangster TA, Lindquist S (2002) Hsp90 as a capacitor of phenotypic variation. *Nature* **417**, 618-624.

Ragsdale EJ, Müller MR, Rödelsperger C, Sommer RJ (2013) A developmental switch coupled to the evolution of plasticity acts through a sulfatase. *Cell* **155**, 922-933.

Renn SCP, Schumer ME (2013) Genetic accommodation and behavioural evolution: insights from genomic studies. *Animal Behaviour* **85**, 1012-1022.

Richards CL, Bossdorf O, Muth NZ, Gurevitch J, Pigliucci M (2006) Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. *Ecology Letters* **9**, 981-993.

Rijssel JC, Hoogwater ES, Kishe‐Machumu MA*, et al.* (2015) Fast adaptive responses in the oral jaw of Lake Victoria cichlids. *Evolution* **69**, 179-189.

Robinson BW, Parsons KJ (2002) Changing times, spaces, and faces: tests and implications of adaptive morphological plasticity in the fishes of northern postglacial lakes. *Canadian Journal of Fisheries and Aquatic Sciences* **59**, 1819-1833.

Rodriguez-Esteban C, Tsukui T, Yonei S*, et al.* (1999) The T-box genes Tbx4 and Tbx5 regulate limb outgrowth and identity. *Nature* **398**, 814-818.

Römer U (1998) *Cichliden-Atlas* Mergus, Verlag für Natur-und Heimtierkunde Baensch.

Römer U, Hahn I (2008) Apistogramma barlowi sp. N.: description of a new faculative mouth-breeding cichlid species (Teleostei: Perciformes: Geophaginae) from Northern Peru. *Vertebrate Zoology* **58**, 49-66.

Rüber L, Adams D (2001) Evolutionary convergence of body shape and trophic morphology in cichlids from Lake Tanganyika. *Journal of Evolutionary Biology* **14**, 325-332.

Runcie DE, Garfield DA, Babbitt CC*, et al.* (2012) Genetics of gene expression responses to temperature stress in a sea urchin gene network. *Molecular Ecology* **21**, 4547-4562.

Rundell RJ, Price TD (2009) Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *Trends in Ecology & Evolution* **24**, 394-399.

Rundle HD, Nagel L, Boughman JW, Schluter D (2000) Natural selection and parallel speciation in sympatric sticklebacks. *Science* **287**, 306-308.

Rutherford SL, Lindquist S (1998) Hsp90 as a capacitor for morphological evolution. *Nature* **396**, 336-342.

Sabbah S, Gray SM, Boss ES*, et al.* (2011) The underwater photic environment of Cape Maclear, Lake Malawi: comparison between rock-and sand-bottom habitats and implications for cichlid fish vision. *The Journal of experimental biology* **214**, 487-500.

Sabbah S, Hui J, Hauser FE, Nelson WA, Hawryshyn CW (2012) Ontogeny in the visual system of Nile tilapia. *Journal of Experimental Biology* **215**, 2684-2695.

Sabbah S, Laria RL, Gray SM, Hawryshyn CW (2010) Functional diversity in the color vision of cichlid fishes. *BMC biology* **8**, 1.

Sabherwal N, Bangs F, Röth R*, et al.* (2007) Long-range conserved non-coding SHOX sequences regulate expression in developing chicken limb and are associated with short stature phenotypes in human patients. *Human molecular genetics* **16**, 210-222.

Sadlon TJ, Dell'Oso T, Surinya KH, May BK (1999) Regulation of erythroid 5-aminolevulinate synthase expression during erythropoiesis. *The international journal of biochemistry & cell biology* **31**, 1153-1167.

Salzburger W, Mack T, Verheyen E, Meyer A (2005) Out of Tanganyika: genesis, explosive speciation, key-innovations and phylogeography of the haplochromine cichlid fishes. *BMC Evolutionary Biology* **5**, 17.

Sanchez G, Sanchez MG (2012) Package 'plsdepot'. *Partial Least Squares (PLS) Data Analysis Methods, v. 0.1* **17**.

Sandkam B, Young CM, Breden F (2015) Beauty in the eyes of the beholders: colour vision is tuned to mate preference in the Trinidadian guppy (Poecilia reticulata). *Molecular Ecology* **24**, 596-609.

Schaum CE, Collins S (2014) Plasticity predicts evolution in a marine alga. *Proceedings of the Royal Society of London B: Biological Sciences* **281**, 20141486.

Scheiner SM (1993) Genetics and evolution of phenotypic plasticity. *Annual review of ecology and systematics*, 35-68.

Scheiner SM, Holt RD (2012) The genetics of phenotypic plasticity. X. Variation versus uncertainty. *Ecology and Evolution* **2**, 751-767.

Schlichting C (2004) The role of phenotypic plasticity in diversification. *Phenotypic Plasticity: functional and conceptual approaches*, 191-200.

Schlichting CD (1986) The evolution of phenotypic plasticity in plants. *Annual review of ecology and systematics*, 667-693.

Schlichting CD, Pigliucci M (1993) Control of phenotypic plasticity via regulatory genes. *American Naturalist*, 366-370.

Schlichting CD, Pigliucci M (1998) *Phenotypic evolution: a reaction norm perspective* Sinauer Associates Incorporated.

Schlichting CD, Wund MA (2014) Phenotypic plasticity and epigenetic marking: an assessment of evidence for genetic accommodation. *Evolution* **68**, 656-672.

Schlosser G, Wagner GP (2004) *Modularity in development and evolution* University of Chicago Press.

Schluter D (1993) Adaptive radiation in sticklebacks: size, shape, and habitat use efficiency. *Ecology*, 699-709.

Schluter D (1994) Experimental evidence that competition promotes divergence in adaptive radiation. *Science*, 798-798.

Schluter D (2000) *The ecology of adaptive radiation* OUP Oxford.

Schneider RF, Li Y, Meyer A, Gunter HM (2014) Regulatory gene networks that shape the development of adaptive phenotypic plasticity in a cichlid fish. *Molecular Ecology* **23**, 4511-4526.

Schneider RF, Meyer A (2016) How plasticity, genetic assimilation and cryptic genetic variation may contribute to adaptive radiations. *Molecular Ecology*.

Schön I, Martens K (2004) Adaptive, pre-adaptive and non-adaptive components of radiations in ancient lakes: a review. *Organisms Diversity & Evolution* **4**, 137-156.

Schott RK, Refvik SP, Hauser FE, López-Fernández H, Chang BS (2014) Divergent positive selection in rhodopsin from lake and riverine cichlid fishes. *Molecular biology and evolution* **31**, 1149-1165.

Schwander T, Leimar O (2011) Genes as leaders and followers in evolution. *Trends in Ecology & Evolution* **26**, 143-151.

Scoville AG, Pfrender ME (2010) Phenotypic plasticity facilitates recurrent rapid adaptation to introduced predators. *Proceedings of the National Academy of Sciences* **107**, 4260-4263.

Seehausen O (1997) Distribution of and reproductive isolation among color morphs of a rock‐dwelling Lake Victoria cichlid (Haplochromis nyererei). *Ecology of Freshwater Fish* **6**, 59-66.

Seehausen O (2006) African cichlid fish: a model system in adaptive radiation research. *Proceedings of the Royal Society B: Biological Sciences* **273**, 1987-1998.

Seehausen O (2013) Conditions when hybridization might predispose populations for adaptive radiation. *Journal of Evolutionary Biology* **26**, 279-281.

Seehausen O, Mayhew P, Van Alphen J (1999) Evolution of colour patterns in East African cichlid fish. *Journal of Evolutionary Biology* **12**, 514-534.

Seehausen O, Terai Y, Magalhaes IS*, et al.* (2008) Speciation through sensory drive in cichlid fish. *Nature* **455**, 620-626.

Seehausen O, van Alphen JJ (1998) The effect of male coloration on female mate choice in closely related Lake Victoria cichlids (Haplochromis nyererei complex). *Behavioral Ecology and Sociobiology* **42**, 1-8.

Seehausen O, Van Alphen JJ, Witte F (1997) Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science* **277**, 1808-1811.

Sen B, Styner M, Xie Z*, et al.* (2009) Mechanical Loading Regulates NFATc1 and β-Catenin Signaling through a GSK3β Control Node. *Journal of Biological Chemistry* **284**, 34607-34617.

Sepkoski Jr JJ (1996) Competition in macroevolution: the double wedge revisited. *Evolutionary paleobiology. University of Chicago Press, Chicago*, 211-255.

Serobyan V, Ragsdale EJ, Müller MR, Sommer RJ (2013) Feeding plasticity in the nematode Pristionchus pacificus is influenced by sex and social context and is linked to developmental speed. *Evolution & Development* **15**, 161-170.

Serobyan V, Xiao H, Namdeo S*, et al.* (2016) Chromatin remodelling and antisense-mediated up-regulation of the developmental switch gene eud-1 control predatory feeding plasticity. *Nature communications* **7**.

Shapiro B (2016) Pathways to de‐extinction: how close can we get to resurrection of an extinct species? *Functional ecology*.

Shiga M, Takigawa I, Mamitsuka H (2007) Annotating gene function by combining expression data with a modular gene network. *Bioinformatics* **23**, i468-i478.

Shubin N, Tabin C, Carroll S (2009) Deep homology and the origins of evolutionary novelty. *Nature* **457**, 818-823.

Siebeck UE, Marshall NJ (2001) Ocular media transmission of coral reef fish—can coral reef fish see ultraviolet light? *Vision research* **41**, 133-149.

Sigmund WR (1983) Female preference for Anolis carolinensis males as a function of dewlap color and background coloration. *Journal of Herpetology*, 137-143.

Sikkink KL, Reynolds RM, Ituarte CM, Cresko WA, Phillips PC (2014) Rapid evolution of phenotypic plasticity and shifting thresholds of genetic assimilation in the nematode Caenorhabditis remanei. *G3: Genes| Genomes| Genetics* **4**, 1103-1112.

Simpson G (1953) *The major features of evolution* Simon & Schuster, New York.
Simpson SJ, McCaffery A, Hägele BF (1999) A behavioural analysis of phase change in the desert locust. *Biological Reviews* **74**, 461-480.
Slootweg R (1987) Prey selection by molluscivorous cichlids foraging on a schistosomiasis vector snail, Biomphalaria glabrata. *Oecologia* **74**, 193-202.
Slootweg R, Malek E, McCullough F (1994) The biological control of snail intermediate hosts of schistosomiasis by fish. *Reviews in Fish Biology and Fisheries* **4**, 67-90.
Smith AR, D'Annunzio L, Smith AE*, et al.* (2011) Intraspecific cone opsin expression variation in the cichlids of Lake Malawi. *Molecular Ecology* **20**, 299-310.
Smith AR, van Staaden MJ, Carleton KL (2012) An evaluation of the role of sensory drive in the evolution of lake Malawi cichlid fishes. *International journal of evolutionary biology* **2012**.
Smith JM (1989) *Evolutionary genetics* Oxford University Press, Oxford.
Smits J, Witte F, Veen F (1996) Functional changes in the anatomy of the pharyngeal jaw apparatus of Astatoreochromis alluaudi (Pisces, Cichlidae), and their effects on adjacent structures. *Biological Journal of the Linnean Society* **59**, 389-409.
Smits JD (1996) Trophic flexibility through spatial reallocation of anatomical structures in the cichlid fish Astatoreochromis alluaudi. *DPhil, University of Leiden, Leiden, The Netherlands*.
Snell‐Rood EC, Van Dyken JD, Cruickshank T, Wade MJ, Moczek AP (2010) Toward a population genetic framework of developmental evolution: the costs, limits, and consequences of phenotypic plasticity. *BioEssays* **32**, 71-81.
Sommer RJ, Ogawa A (2011) Hormone signaling and phenotypic plasticity in nematode development and evolution. *Current Biology* **21**, R758-R766.
Song G, Harding SE, Duchen MR*, et al.* (2002) Altered mechanical properties and intracellular calcium signaling in cardiomyocytes from annexin 6 null-mutant mice. *The FASEB Journal* **16**, 622-624.
Spady TC, Parry JW, Robinson PR*, et al.* (2006) Evolution of the cichlid visual palette through ontogenetic subfunctionalization of the opsin gene arrays. *Molecular biology and evolution* **23**, 1538-1547.
Spight TM, Emlen J (1976) Clutch sizes of two marine snails with a changing food supply. *Ecology*, 1162-1178.
Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688-2690.
Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web servers. *Systematic Biology* **57**, 758-771.
Standen E (2008) Pelvic fin locomotor function in fishes: three-dimensional kinematics in rainbow trout (Oncorhynchus mykiss). *Journal of Experimental Biology* **211**, 2931-2942.
Standen EM, Du TY, Larsson HC (2014) Developmental plasticity and the origin of tetrapods. *Nature* **513**, 54-58.
Stauffer JR, van Snick Gray E (2004) Phenotypic plasticity: its role in trophic radiation and explosive speciation in cichlids (Teleostei: Cichlidae). *Animal Biology* **54**, 137-158.
Stebbins Jr C (1950) Variation and evolution in plants. *Variation and evolution in plants.*
Stelkens RB, Schmid C, Selz O, Seehausen O (2009) Phenotypic novelty in experimental hybrids is predicted by the genetic distance between species of cichlid fish. *BMC Evolutionary Biology* **9**, 283.
Stiassny ML, Meyer A (1999) Cichlids of the rift lakes. *Scientific American - American Edition* **280**, 64-69.
Stölting KN, Wilson AB (2007) Male pregnancy in seahorses and pipefish: beyond the mammalian model. *BioEssays* **29**, 884-896.
Stomp M, van Dijk MA, van Overzee HM*, et al.* (2008) The timescale of phenotypic plasticity and its impact on competition in fluctuating environments. *The American Naturalist* **172**, E169-E185.

Stuart YE, Losos JB (2013) Ecological character displacement: glass half full or half empty? *Trends in Ecology & Evolution* **28**, 402-408.

Sturmbauer C, Hainz U, Baric S, Verheyen E, Salzburger W (2003) Evolution of the tribe Tropheini from Lake Tanganyika: synchronized explosive speciation producing multiple evolutionary parallelism. *Hydrobiologia* **500**, 51-64.

Sturmbauer C, Koblmüller S, Sefc KM, Duftner N (2005) Phylogeographic history of the genus Tropheus, a lineage of rock-dwelling cichlid fishes endemic to Lake Tanganyika. In: *Aquatic Biodiversity II*, pp. 335-366. Springer.

Sturmbauer C, Mark W, Dallinger R (1992) Ecophysiology of aufwuchs-eating cichlids in Lake Tanganyika: niche separation by trophic specialization. *Environmental Biology of Fishes* **35**, 283-290.

Sturmbauer C, Meyer A (1992) Genetic divergence, speciation and morphological stasis in a lineage of African cichlid fishes. *Nature* **358**, 578-581.

Sultan S (2015) *Organism & Environment*, 1 edn. Oxford University Press, Oxford University Press, New York, USA.

Sultan SE, Barton K, Wilczek AM (2009) Contrasting patterns of transgenerational plasticity in ecologically distinct congeners. *Ecology* **90**, 1831-1839.

Sumpio BE, Du W, Xu W-J (1994) Exposure of endothelial cells to cyclic strain induces c-fos, fosB and c-jun but not jun B or jun D and increases the transcription factor AP-1. *Endothelium* **2**, 149-156.

Superti—Furga A, Scambler PI, Winter RM (1998) Mutation and deletion of the pseudoautosomal gene SHOX cause Leri-Weill dyschondrosteosis. *Nature genetics* **19**, 71.

Susoy V, Ragsdale EJ, Kanzaki N, Sommer RJ (2015) Rapid diversification associated with a macroevolutionary pulse of developmental plasticity. *Elife* **4**, e05463.

Suzuki Y, Nijhout HF (2006) Evolution of a polyphenism by genetic accommodation. *Science* **311**, 650-652.

Svanbäck R, Schluter D (2012) Niche specialization influences adaptive phenotypic plasticity in the threespine stickleback. *The American Naturalist* **180**, 50-59.

Swynghedauw B (2006) Phenotypic plasticity of adult myocardium: molecular mechanisms. *Journal of Experimental Biology* **209**, 2320-2327.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585-595.

Tamura K, Yonei-Tamura S, Belmonte JCI (1999) Differential expression of Tbx4 and Tbx5 in Zebrafish fin buds. *Mechanisms of development* **87**, 181-184.

Tanaka M, Hale LA, Amores A*, et al.* (2005) Developmental genetic basis for the evolution of pelvic fin loss in the pufferfish Takifugu rubripes. *Developmental biology* **281**, 227-239.

Taylor JS, Braasch I, Frickey T, Meyer A, Van de Peer Y (2003) Genome duplication, a trait shared by 22,000 species of ray-finned fish. *Genome research* **13**, 382-390.

Team RC (2013) R: A language and environment for statistical computing.

Tebbich S, Sterelny K, Teschke I (2010) The tale of the finch: adaptive radiation and behavioural flexibility. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **365**, 1099-1109.

Teske PR, Cherry MI, Matthee CA (2004) The evolutionary history of seahorses (Syngnathidae: Hippocampus): molecular data suggest a West Pacific origin and two invasions of the Atlantic Ocean. *Molecular phylogenetics and evolution* **30**, 273-286.

Theis A, Ronco F, Indermaur A, Salzburger W, Egger B (2014) Adaptive divergence between lake and stream populations of an East African cichlid fish. *Molecular Ecology* **23**, 5304-5322.

Thibert-Plante X, Hendry A (2011) The consequences of phenotypic plasticity for ecological speciation. *Journal of Evolutionary Biology* **24**, 326-342.

Torres-Dowdall J, Henning F, Elmer KR, Meyer A (2015) Ecological and Lineage-Specific Factors Drive the Molecular Evolution of Rhodopsin in Cichlid Fishes. *Molecular biology and evolution* **32**, 2876-2882.

Torres‐Dowdall J, Handelsman CA, Reznick DN, Ghalambor CK (2012) Local adaptation and the evolution of phenotypic plasticity in Trinidadian guppies (Poecilia reticulata). *Evolution* **66**, 3432-3443.

Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105-1111.

Treves S, Anderson AA, Ducreux S*, et al.* (2005) Ryanodine receptor 1 mutations, dysregulation of calcium homeostasis and neuromuscular disorders. *Neuromuscular Disorders* **15**, 577-587.

Trewavas E, Green J, Corbet SA (1972) Ecological studies on crater lakes in West Cameroon fishes of Barombi Mbo. *Journal of Zoology* **167**, 41-95.

Tucker R, Hagios C, Chiquet‐Ehrismann R, Lawler J (1997) In situ localization of thrombospondin‐1 and thrombospondin‐3 transcripts in the avian embryo. *Developmental Dynamics* **208**, 326-337.

Turcotte MM, Levine JM (2016) Phenotypic Plasticity and Species Coexistence. *Trends in Ecology & Evolution*.

Van Dooren TJ, Van Goor HA, Van Putten M (2010) Handedness and asymmetry in scale‐eating cichlids: antisymmetries of different strength. *Evolution* **64**, 2159-2165.

Van Kleunen M, Fischer M (2005) Constraints on the evolution of adaptive phenotypic plasticity in plants. *New Phytologist* **166**, 49-60.

Van Laere A-S, Nguyen M, Braunschweig M*, et al.* (2003) A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature* **425**, 832-836.

van Rijssel JC, Witte F (2013) Adaptive responses in resurgent Lake Victoria cichlids over the past 30 years. *Evolutionary Ecology* **27**, 253-267.

Van Tienderen PH (1997) Generalists, specialists, and the evolution of phenotypic plasticity in sympatric populations of distinct species. *Evolution*, 1372-1380.

Van Valen L (1965) Morphological variation and width of ecological niche. *American Naturalist*, 377-390.

Vandesompele J, De Preter K, Pattyn F*, et al.* (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology* **3**, research0034.

Venkatesh B, Kirkness EF, Loh Y-H*, et al.* (2006) Ancient noncoding elements conserved in the human genome. *Science* **314**, 1892-1892.

Venkatesh B, Lee AP, Ravi V*, et al.* (2014) Elephant shark genome provides unique insights into gnathostome evolution. *Nature* **505**, 174-179.

Verheyen E, Salzburger W, Snoeks J, Meyer A (2003) Origin of the superflock of cichlid fishes from Lake Victoria, East Africa. *Science* **300**, 325-329.

Versele M, de Winde JH, Thevelein JM (1999) A novel regulator of G protein signalling in yeast, Rgs2, downregulates glucose-activation of the cAMP pathway through direct inhibition of Gpa2. *The EMBO journal* **18**, 5577-5591.

Visel A, Prabhakar S, Akiyama JA*, et al.* (2008) Ultraconservation identifies a small subset of extremely constrained developmental enhancers. *Nature genetics* **40**, 158-160.

Visser ME, Holleman LJ, Caro SP (2009) Temperature has a causal effect on avian timing of reproduction. *Proceedings of the Royal Society B: Biological Sciences* **276**, 2323-2331.

Waddington CH (1942) Canalization of development and the inheritance of acquired characters. *Nature* **150**, 563-565.

Waddington CH (1953) Genetic assimilation of an acquired character. *Evolution* **7**, 118-126.

Waddington CH (1961) Genetic assimilation. *Advances in genetics* **10**, 257-293.

Wagner A (2008) Robustness and evolvability: a paradox resolved. *Proceedings of the Royal Society of London B: Biological Sciences* **275**, 91-100.

Wagner GP, Pavlicev M, Cheverud JM (2007) The road to modularity. *Nature Reviews Genetics* **8**, 921-931.

Wainwright PC, Smith WL, Price SA*, et al.* (2012) The Evolution of Pharyngognathy: A Phylogenetic and Functional Appraisal of the Pharyngeal Jaw Key Innovation in Labroid fishes and Beyond. *Systematic Biology*.

Waitt C, Little AC, Wolfensohn S*, et al.* (2003) Evidence from rhesus macaques suggests that male coloration plays a role in female primate mate choice. *Proceedings of the Royal Society of London B: Biological Sciences* **270**, S144-S146.

Wald G (1961) The visual function of the vitamins A. *Vitamins & Hormones* **18**, 417-430.

Wald G (1968) The molecular basis of visual excitation. *Nature* **219**, 800-807.

Walker L, Publicover S, Preston M, Said Ahmed M, El Haj A (2000) Calcium‐channel activation and matrix protein upregulation in bone cells in response to mechanical strain. *Journal of cellular biochemistry* **79**, 648-661.

Wang X, Kang L (2014) Molecular mechanisms of phase change in locusts. *Annual review of entomology* **59**, 225-244.

Wang Y, Azaïs T, Robin M*, et al.* (2012) The predominant role of collagen in the nucleation, growth, structure and orientation of bone apatite. *Nature materials* **11**, 724-733.

Warner RR, Swearer SE (1991) Social control of sex change in the bluehead wrasse, Thalassoma bifasciatum (Pisces: Labridae). *The Biological Bulletin* **181**, 199-204.

Weadick CJ, Loew ER, Rodd FH, Chang BS (2012) Visual pigment molecular evolution in the Trinidadian pike cichlid (Crenicichla frenata): a less colorful world for Neotropical cichlids? *Molecular biology and evolution* **29**, 3045-3060.

Wellborn GA, Langerhans RB (2015) Ecological opportunity and the adaptive diversification of lineages. *Ecology and Evolution* **5**, 176-195.

West-Eberhard MJ (1989) Phenotypic plasticity and the origins of diversity. *Annual review of ecology and systematics*, 249-278.

West-Eberhard MJ (2003) *Developmental plasticity and evolution* Oxford University Press, USA.

West-Eberhard MJ (2005) Developmental plasticity and the origin of species differences. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 6543-6549.

Whitehead A, Roach JL, Zhang S, Galvez F (2011) Genomic mechanisms of evolved physiological plasticity in killifish distributed along an environmental salinity gradient. *Proceedings of the National Academy of Sciences* **108**, 6193-6198.

Whitelaw NC, Whitelaw E (2006) How lifetimes shape epigenotype within and across generations. *Human molecular genetics* **15**, R131-R137.

Whittington CM, Griffith OW, Qi W, Thompson MB, Wilson AB (2015) Seahorse brood pouch transcriptome reveals common genes associated with vertebrate pregnancy. *Molecular biology and evolution* **32**, 3114-3131.

Williams GC (1966) *Adaptation and natural selection: a critique of some current evolutionary thought* Princeton University Press, Princton New Jersey.

Wilson AB, Vincent A, Ahnesjö I, Meyer A (2001) Male pregnancy in seahorses and pipefishes (family Syngnathidae): rapid diversification of paternal brood pouch morphology inferred from a molecular phylogeny. *Journal of Heredity* **92**, 159-166.

Wimberger PH (1991) Plasticity of jaw and skull morphology in the neotropical cichlids Geophagus brasiliensis and G. steindachneri. *Evolution*, 1545-1563.

Wimberger PH (1994) Trophic polymorphisms, plasticity, and speciation in vertebrates. *Advances in Fish Foraging Theory and Ecology* **18**, 19-43.

Wisenden BD, Stumbo AD, Self PA*, et al.* (2015) Co-evolution of offspring antipredator competence and parental brood defense in convict cichlids. *Hydrobiologia* **748**, 259-272.

Witte F (1980) Initial results of the ecological survey of the haplochromine cichlid fishes from the Mwanza Gulf of Lake Victoria (Tanzania): breeding patterns, trophic and species distribution. *Netherlands Journal of Zoology* **31**, 175-202.

Woltereck R (1913) Beitrag zur Analyse der „Vererbung erworbener Eigenschaften": Transmutation und Präinduktion beiDaphnia. *Molecular and General Genetics MGG* **9**, 146-155.

Wray GA (2007) The evolutionary significance of cis-regulatory mutations. *Nature Reviews Genetics* **8**, 206-216.

Wright WE, Sassoon DA, Lin VK (1989) Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD. *Cell* **56**, 607-617.

Wund MA, Baker JA, Clancy B, Golub JL, Foster SA (2008) A test of the flexible stem model of evolution: ancestral plasticity, genetic accommodation, and morphological divergence in the threespine stickleback radiation. *The American Naturalist* **172**, 449-462.

Wund MA, Valena S, Wood S, Baker JA (2012) Ancestral plasticity and allometry in threespine stickleback reveal phenotypes associated with derived, freshwater ecotypes. *Biological Journal of the Linnean Society* **105**, 573-583.

Yamanoue Y, Setiamarga D, Matsuura K (2010) Pelvic fins in teleosts: structure, function and evolution. *Journal of Fish Biology* **77**, 1173-1208.

Yamaoka K (1983) Feeding behaviour and dental morphology of algae scraping cichlids (Pisces: Teleostei) in Lake Tanganyika.

Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Molecular biology and evolution* **24**, 1586-1591.

Yeh PJ, Price TD (2004) Adaptive phenotypic plasticity and the successful colonization of a novel environment. *The American Naturalist* **164**, 531-542.

Yim KM, Brewer MS, Miller CT, Gillespie RG (2014) Comparative transcriptomics of maturity-associated color change in Hawaiian spiders. *Journal of Heredity* **105**, 771-781.

Yokoyama S (2000) Molecular evolution of vertebrate visual pigments. *Progress in retinal and eye research* **19**, 385-419.

Young KA, Snoeks J, Seehausen O (2009) Morphological diversity and the roles of contingency, chance and determinism in African cichlid radiations. *PLOS one* **4**, e4740.

Young RL (2013) Linking conceptual mechanisms and transcriptomic evidence of plasticity-driven diversification. *Molecular Ecology* **22**, 4363-4365.

Yu X, Lin J, Zack DJ, Qian J (2006) Computational analysis of tissue-specific combinatorial gene regulation: predicting interaction between transcription factors in human tissues. *Nucleic acids research* **34**, 4925-4936.

Zahavi A, Zahavi A (1999) *The handicap principle: A missing piece of Darwin's puzzle* Oxford University Press.

Zhang G, Li C, Li Q*, et al.* (2014) Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* **346**, 1311-1320.

Zhang XS (2005) Evolution and maintenance of the environmental component of the phenotypic variance: benefit of plastic traits under changing environments. *The American Naturalist* **166**, 569-580.

Zimova M, Mills LS, Nowak JJ (2016) High fitness costs of climate change‐induced camouflage mismatch. *Ecology Letters* **19**, 299-307.

Zonios G, Dimou A, Bassukas I*, et al.* (2008) Melanin absorption spectroscopy: new method for noninvasive skin investigation and melanoma detection. *Journal of biomedical optics* **13**, 014017-014017-014018.

Acknowledgements

General acknowledgements

Firstly, I am thankful to Prof. Axel Meyer for giving me the opportunity to conduct my PhD research in his lab and to be supportive in research matters as well as aspects of career development. Also, I am particular thankful to him as he gave me the opportunity to develop my own research ideas while keeping a barely manageable variety of fish species. Also, his grants paid my salary and I thank him and the DFG and the University of Konstanz Young Scholar Fund for it.

Prof. Manfred Schartl agreed to referee my thesis and I am very thankful for this as well as for valuable comments on my projects during the PAC meeting. Also I want to thank the whole IMPRS for the pleasant and very helpful courses, retreats and other events. I am very grateful to Maggi Hieber-Ruiz and Evi Fricke for all the support.

I am thankful for the supervision and collaborations with Helen Gunter and Joost Woltering. Both researchers taught me a lot and helped me with many issues related to research and career development. I want to thank all my collaborators and colleagues for the productive and pleasant past four years.

I like to further specifically thank the people whose organizational talents and helpfulness keep the Meyer-Lab running: Ingrid Bader, Christiane Weber, Gabi Legant, all of our TAs and the whole TFA stuff.

Corinna Waider was of invaluable help in all kinds of matters, including research question, organizational issues, motivation and many more throughout my whole PhD (and before) and specifically while I was compiling my thesis. I genuinely thank her for all of it. I also like to thank Finn for having been a source of joy during a stressful period.

Finally, I cordially thank my parents for supporting me throughout my whole life and for fostering my interest in all living things.

Chapter specific acknowledgements

Chapter I

We thank the Laboratory Animal Center of Sun Yat-Sen University for providing experimental facilities and the Agency for Science, Technology and Research (A* STAR) Computational Resource Centre for use of its high-performance computing facilities. This research was supported by the National Science Fund for Excellent Young Scholars (41322038), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA13020103), the Youth Foundation of National High Technology Research and

Development Program (863 Program) (2015AA020909), the Outstanding Youth Foundation in Guangdong Province (S2013050014802), the National Natural Science Foundation of China (41576145), the National Key Basic Research Program of China (2015CB452904), the Special Project on the Integration of Industry, Education and Research of Guangdong Province (no. 2013B090800017) and the Biomedical Research Council of A* STAR, Singapore.

Chapter II

We are grateful for funding from the Deutsche Forschungsgemeinschaft (SPP1819) and the University of Konstanz (A.M., R.S.). AM acknowledges support from the European Research Council through ERC advanced grant GenAdap (297300). We thank Helen M. Gunter, Claudius Kratochwil, Maggie Sefton, Pooja Singh, Julian Torres-Dowdall, Adam Wilkins and Joost Woltering for valuable comments and discussions.

Chapter III

We gratefully acknowledge financial support from the Deutsche Forschungsgemeinschaft (GU 1160/2-1) (HG), the University of Konstanz (AM), by the Zukunftskolleg at the University of Konstanz for HG and the China Scholarship Council for YL. We like to thank Xiong Fan for support in the laboratory, Carmelo Fruciano for assistance with morphometric analyses and Patrick Niemantsverdriet and the late Dr. Frans Witte from the University of Leiden for generously providing their strain of *A. alluaudi*.

Chapter IV

We wish to acknowledge the funding agencies that generously provided funding for this work, including the Deutsche Forschungsgemeinschaft (to HG and AM), the Zukunftskolleg at the University of Konstanz (HG), The University of Konstanz (HG and AM), Frospects (IK), Austrian Science Fund grant P22737 (CS) and the International Max Planck Research School (RS). Hans Hoffman, Frans Witte, Erwin Schraml and Thomas Uecker and Marco Welss generously supplied fish used in this study.

Chapter V

We thank the International Max Planck Reasearch School Organismal Biology (IMPRS-OB) for funding for this project. We are also grateful for funding from the Zukunftskolleg Uni Konstanz and from the DFG to cover salaries of RFS and JTD.

Author contributions

Chapter I

Author Contributions Q.L., A.M. and B.V. designed the scientific objectives. Q.L., B.V., A.M., Q.S. and Y.Z. oversaw the project. H.Z., G.Q., B.V. and Y.Z. collected samples for sequencing DNA and RNA. M.X., Y.Y., J.M. and Q.G. performed genome sequencing, assembly and annotation. S.F., J.M. and Y.Y. performed phylogenomic analysis and molecular evolutionary rate analysis. S.F. and M.X. characterized repetitive sequences and GC content. S.F., R.F.S., P.X., V.R. and B.V. annotated and analysed Hox clusters. V.R. and B.V. annotated and analysed SCPP genes. A.P.L., V.R., Z.W.L. and B.V. performed CNE analysis and functional assay of zebrafish CNEs. Y.Z., M.X., C.Z. and D.S. assembled and annotated RNA-seq data. H.M.G. and S.F. interpreted RNA-seq results and designed the qRT–PCR experiment. Y.Z., H.Z. and X.W. performed qRT–PCR to validate the expression levels of transcripts. Y.Z., M.X., H.Z. and V.R. analysed the patristacin gene family. Y.Z., Q.L., J.M.W., R.F.S. and A.M. performed tbx4 knockout analysis. Y.Q., J.B., C.B., Y.S. and X.Z. were involved in data analysis. L.H., G.L., W.L., Z.G., K.W. and H.Q. participated in the discussions related to data analysis. Q.L., Y.Z., S.F., H.M.G., A.M. and B.V. wrote the manuscript with input from all other authors.

Chapter II

R.F.S. wrote the draft of the manuscript. A.M. revised the draft manuscript.

Chapter III

H.G. and A.M. conceived and designed the experiments; H.G. and R.S. conducted the experiments; R.S. and Y.L. performed the data analysis; R.S. and H.G. wrote the first draft of manuscript and all authors contributed to and approved the final version.

Chapter IV

H.M.G., A.M. and C.S. designed the study, and experimental work was carried out by I.K. and H.M.G. Data analysis was performed by R.F.S. and H.M.G. The initial draft of this manuscript was prepared by H.M.G. and all authors assisted with additional drafts and reviewed the final version.

Chapter V

J.T.D., R.F.S. and S.J.R. conceptualized the project. S.J.R. and R.F.S. conducted experimental work and analyzed the data. R.F.S. and S.J.R. wrote the first draft of the paper and all authors revised the manuscript in later stages.

Supplementary Material

Chapter I

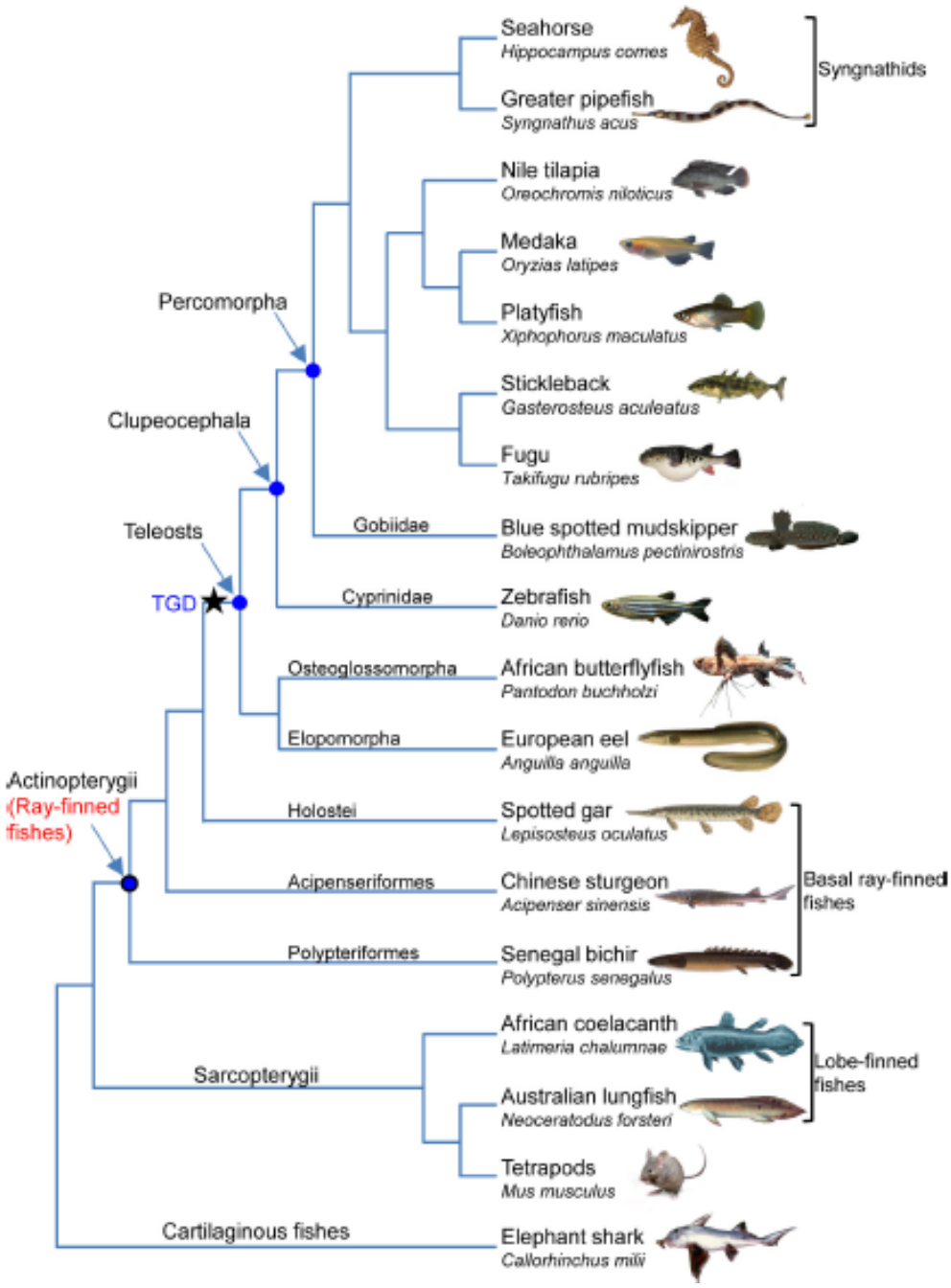


Fig. S.I.1 Phylogenetic relationships of ray-finned fishes discussed in this study. Phylogenetic relationships of ray-finned fishes depicted here are based on the current study and Bian *et al.* 2016. Ray-finned fishes (Actinopterygii) are divided into basal ray-finned fishes (Polypteriformes, Acipenseriformes and Holostei) and teleosts. The latter comprise ~ 99% of the extant ray-finned fishes. The star represents the teleost-specific genome duplication event that occurred in the common ancestor of all teleost fishes. Syngnathids (seahorse and pipefish) display the unique phenomenon of 'male pregnancy'.

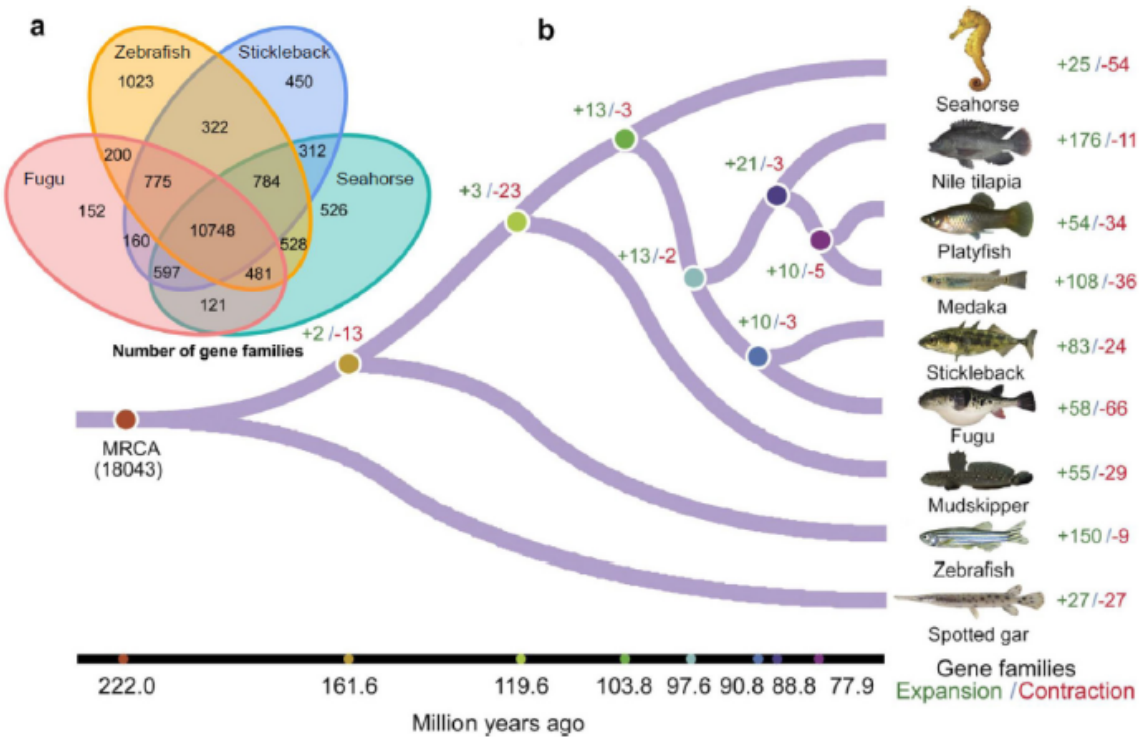


Fig. S.I.2 Number of gene families in various teleosts and the spotted gar. a, Venn diagram of shared orthologous gene families in seahorse (*H. comes*), fugu, zebrafish and stickleback. b, The phylogeny and divergence times of seahorse and other teleost fishes based on analysis of genome-wide one-to-one orthologous protein sequences. The numbers at nodes indicate the number of gene families expanded and contracted at different evolutionary time points.

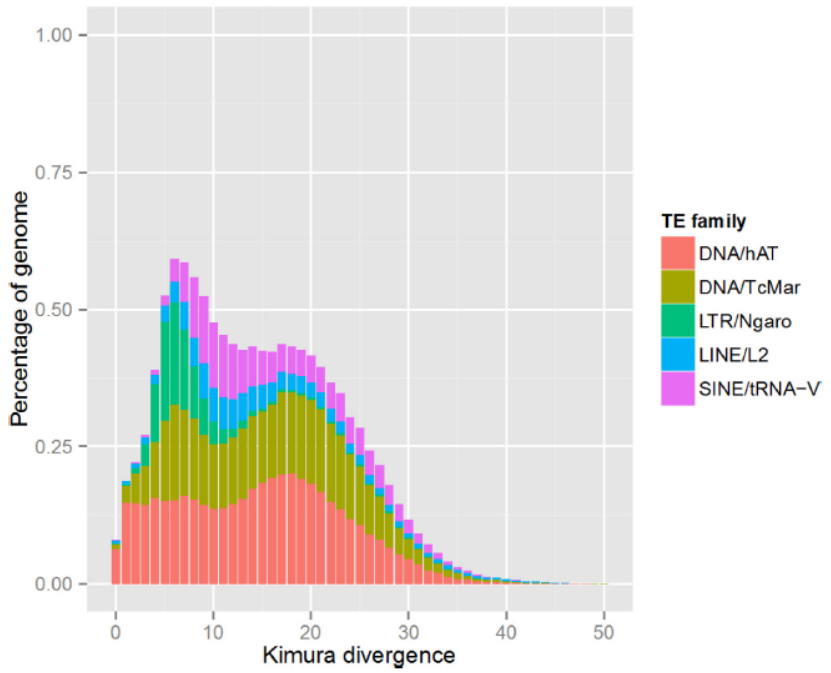


Fig. S.I.3 Divergence distribution of transposable elements compared to consensus in the transposable element library. The divergence rate was calculated between the identified transposable elements (TEs) in the *H. comes* genome and the consensus sequence in the transposable element library.

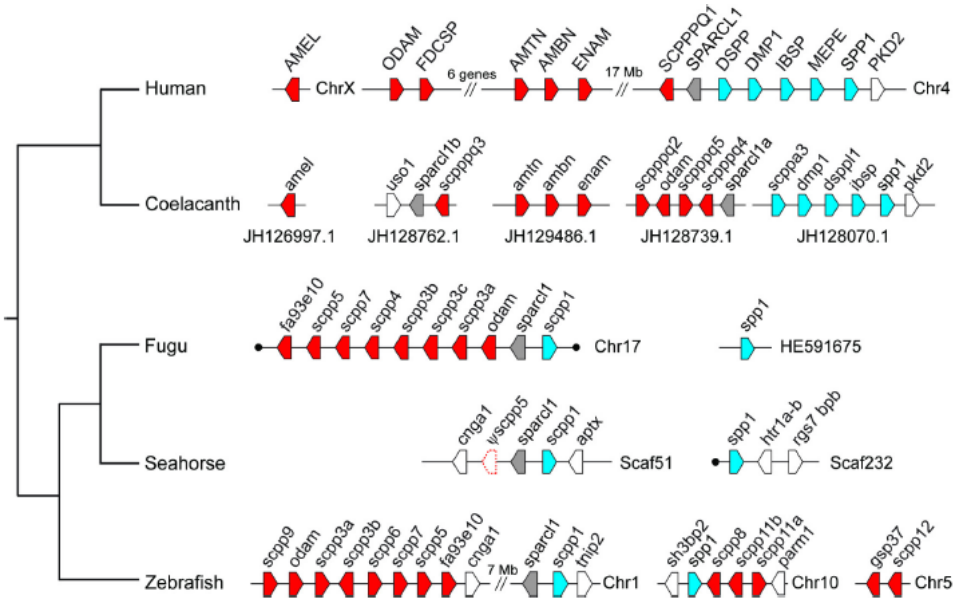


Fig. S.I.4 SCPP genes in *H. comes* and other jawed vertebrates. Gene loci for elephant shark, human, coelacanth and zebrafish were adapted from other publications (Kawasaki & Amemiya 2014; Venkatesh *et al.* 2014). *sparcl1*, which is the ancestral gene that gave rise to SCPP genes is shown in grey; P/Q-rich SCPP genes are shown in red; acidic SCPP genes are shown in blue. In seahorse, *scpp5* is a pseudogene and is denoted by ψ . Owing to space constraints, the P/Q-rich SCPP genes encoding milk casein and salivary proteins in human have been omitted. Black circles mark the ends of scaffolds.

Fig. S.I.5 Maximum-likelihood phylogenetic tree of OR genes in *H. comes* and other ray-finned fishes.

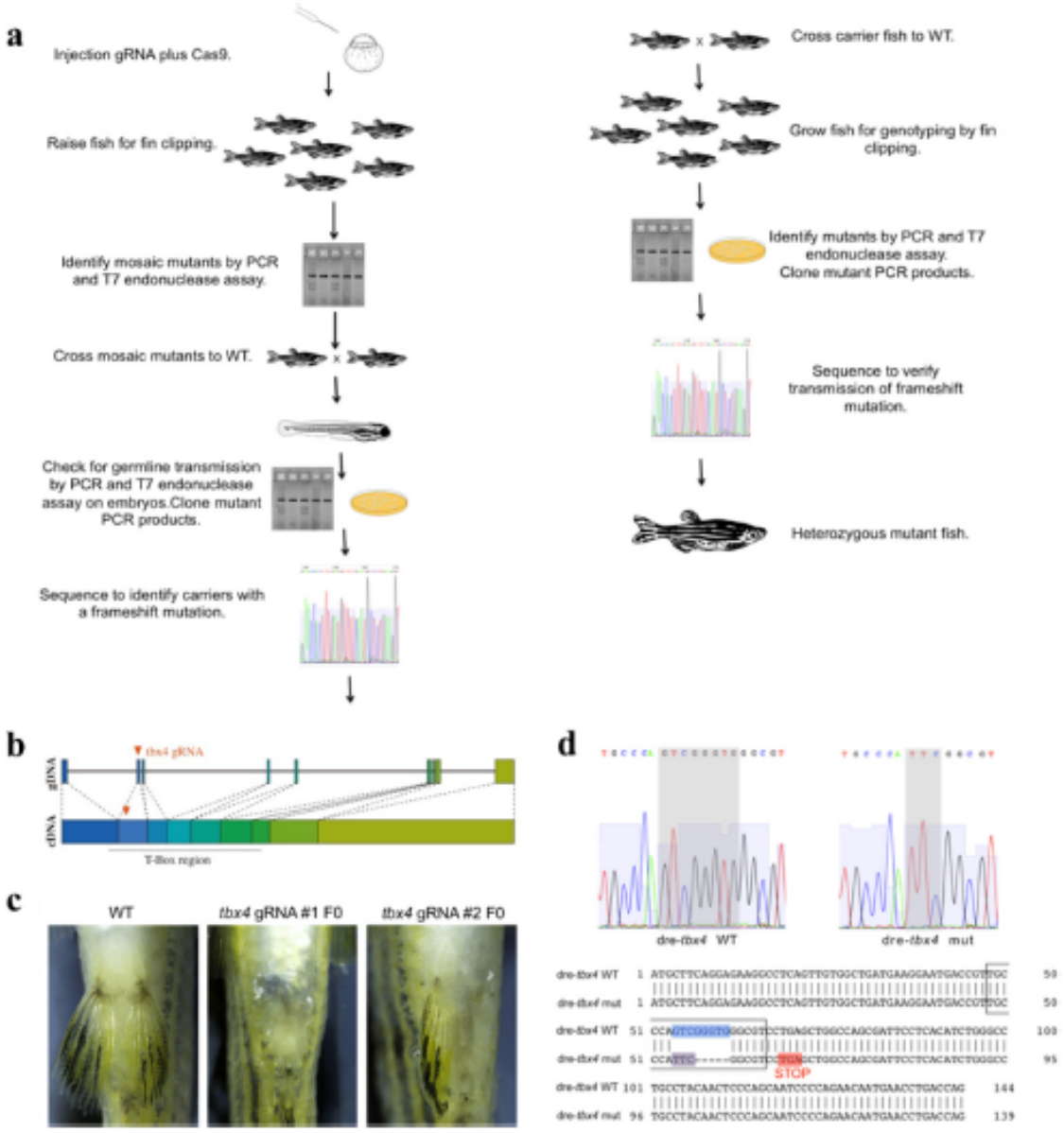


Fig. S.I.6 CRISPR–Cas9 mediated knockdown of *tbx4* in zebrafish. a, CRISPR–Cas9 mutagenesis strategy. b, CRISPR–Cas9 sites targeted in zebrafish *tbx4* gene. c, Loss of function *tbx4* phenotypes in F0 mosaic mutants. Pelvic fin loss was observed with low frequency in F0 mosaic mutant fish. Frequency of animals with either single- or double-sided loss of pelvic fins was 3/42 for gRNA#1 and 1/34 for gRNA#2. d, Identification of zebrafish *tbx4* mutant line. Top shows sequencing chromatograms of wild-type (left) and mutant (right) alleles. Bottom shows alignment of *tbx4* exon 2 from wild-type and mutant. The region for which the chromatograms are shown is indicated with a box. In the mutant a deletion (indicated in blue in the wild-type sequence)/substitution (indicated in lilac in the mutant sequence) was identified. The deletion/substitution area is indicated with a grey box in the chromatograms.

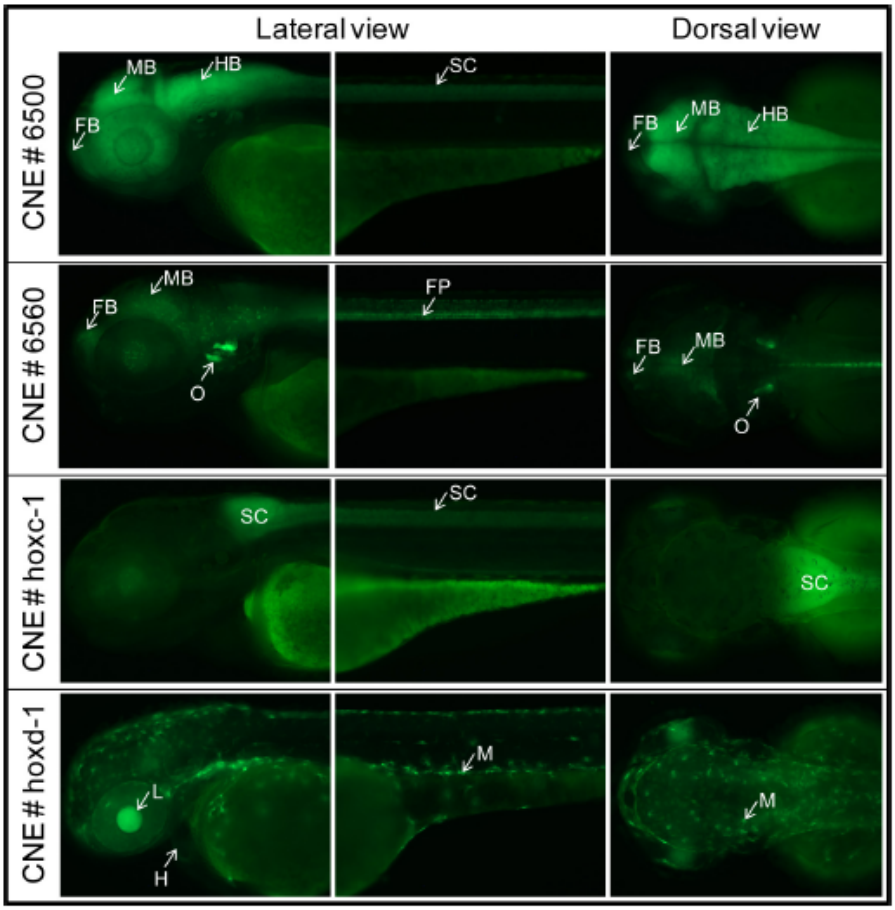


Fig. S.I.7 Reporter gene expression pattern driven by zebrafish CNEs that are lost in *H. comes*. Lateral and dorsal views of 72 h post-fertilization F1 transgenic zebrafish embryos. The lost CNEs (#6500, #6560, #hoxc-1 and #hoxd-1) were assayed for their reporter gene expression potential in transgenic zebrafish. FB, forebrain; FP, floor plate; H, heart; HB, hindbrain; L, lens; M, melanocytes; MB, midbrain; O, otic vesicle; SC, spinal cord.

Online Supplementary Data

http://www.nature.com/nature/journal/v540/n7633/extref/nature20595-s1.pdf

Chapter II

-None-

Chapter III

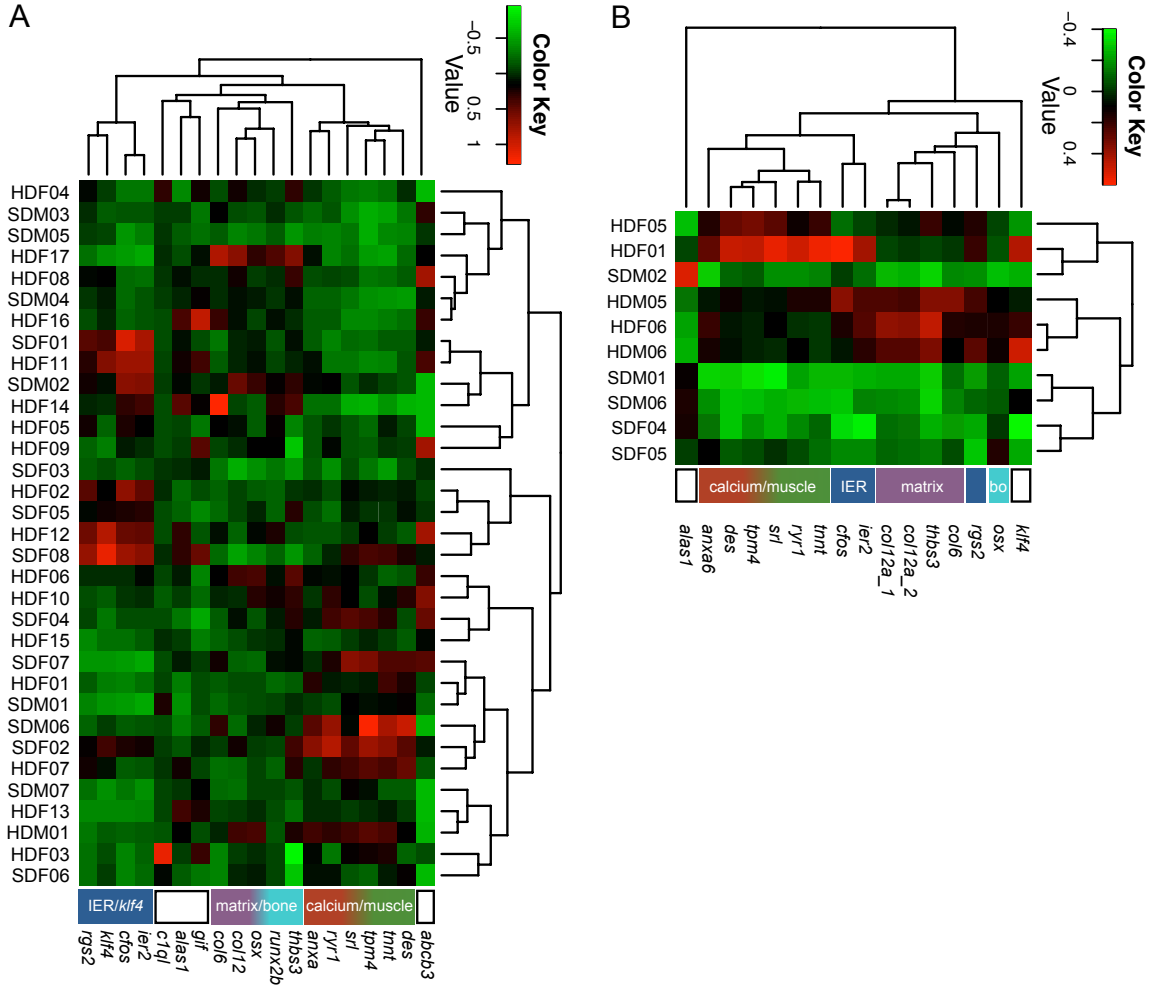


Fig. S.III. 1 Hierarchical gene clustering for all 19 candidate gene expression levels of this study across time-points (A) and for 16 out of these 19 genes, for which expression was detected in the RNAseq study of Gunter et al. (2013) after 18 months of diet treatment (B). Please find information on the analyses in the Material & Methods section of the main document. Genes mostly clustered according to their function and clusters coincide between studies. Muscle and calcium genes are always tightly clustered as matrix and bone genes are. Immediate early genes (IER) are also co-expressed, except for rgs2 (B). The identification of such similar co-expression patterns between independent studies confirms the reliability of the determined gene expression levels and virtually excludes significant tank effects.

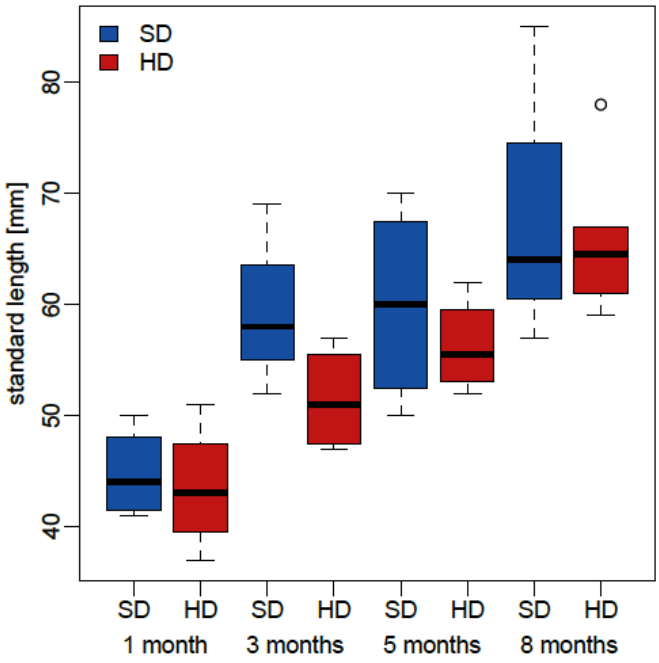


Fig. S.III.2 The mean standard length of fish between diet groups did not differ for any of the four time-points of the study (Wilcoxen Signed Rank tests, all p>0.05)

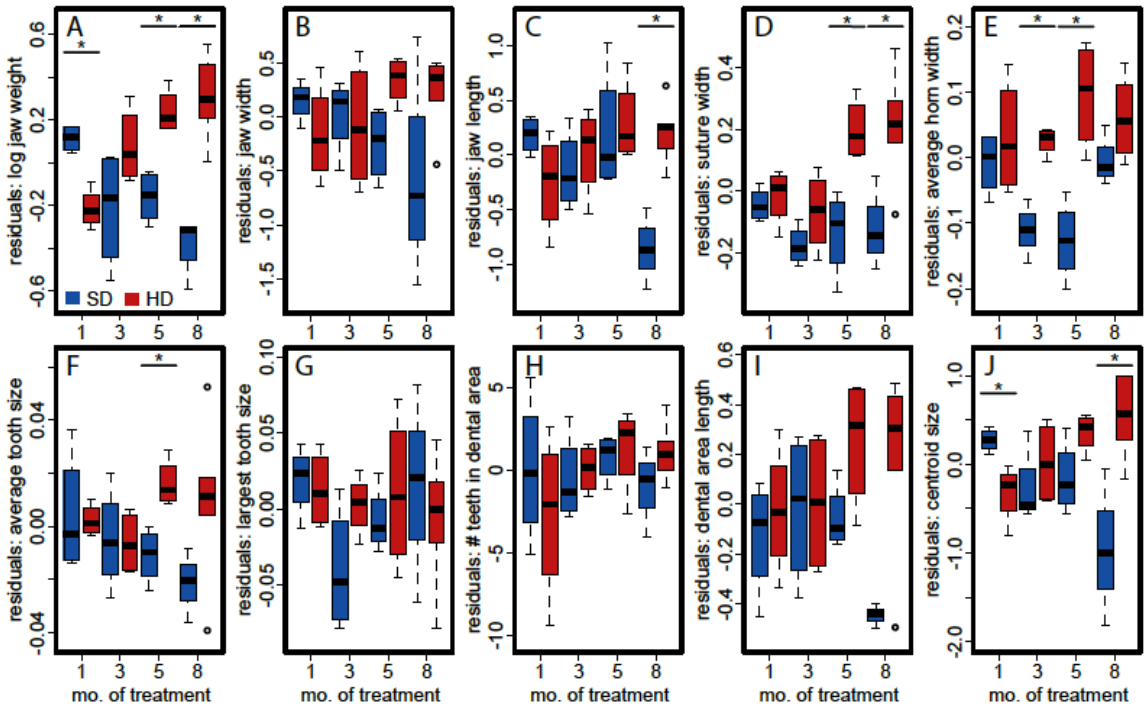


Fig. S.III.3 Residuals of the ten linear morphometric measurements of the LPJ across diet groups and time-points. Boxplots show the median, the 1st and 3rd quartiles as hinges and upper and lower whiskers.

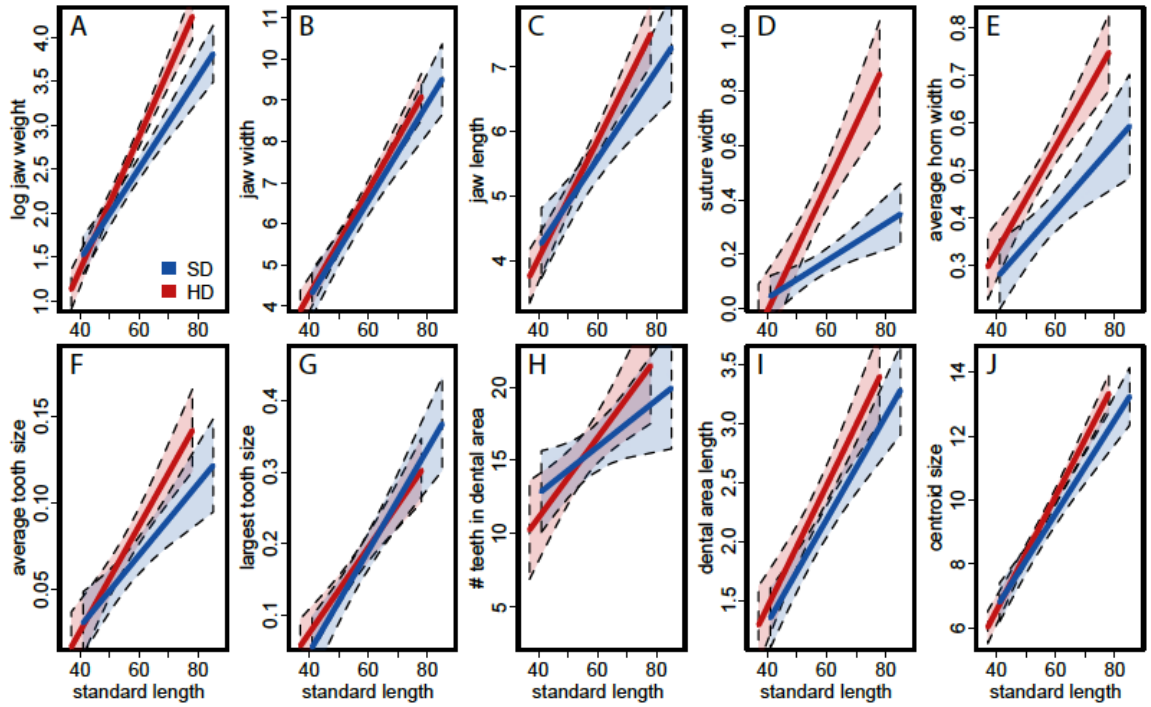


Fig. S.III.4 Plots of the linear morphometric measurements of the LPJ across standard length for the two diet groups. Marked areas reflect 95% CIs.

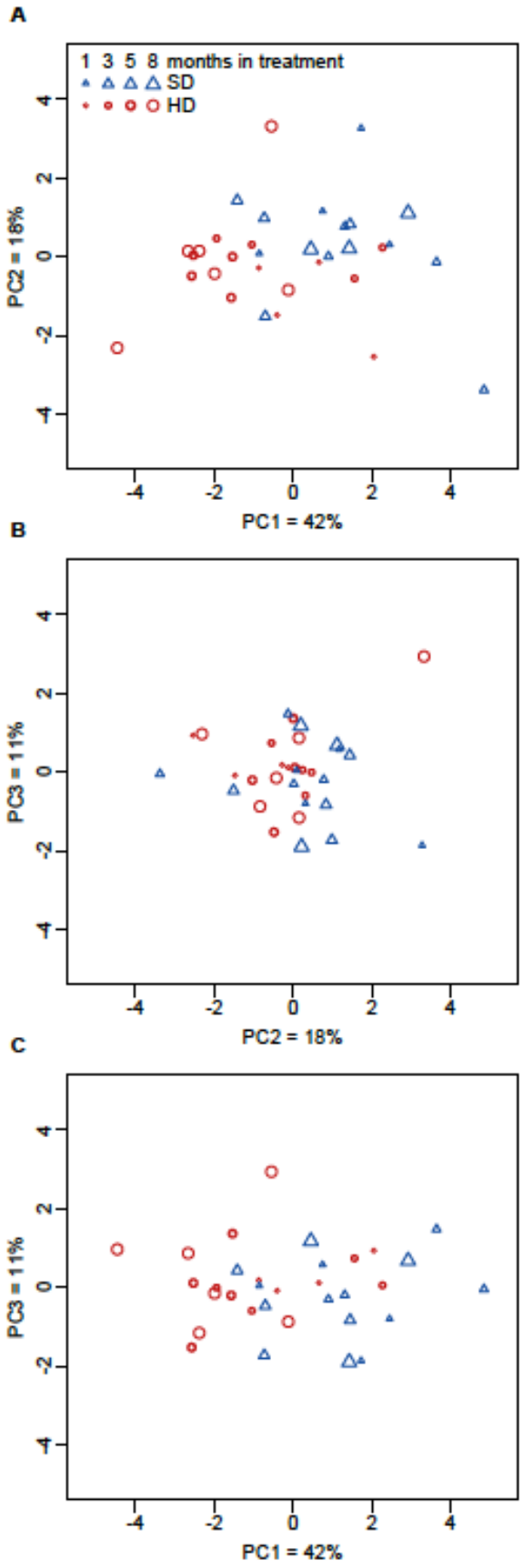


Fig. S.III.5 PCA on linear morphometric measurements.

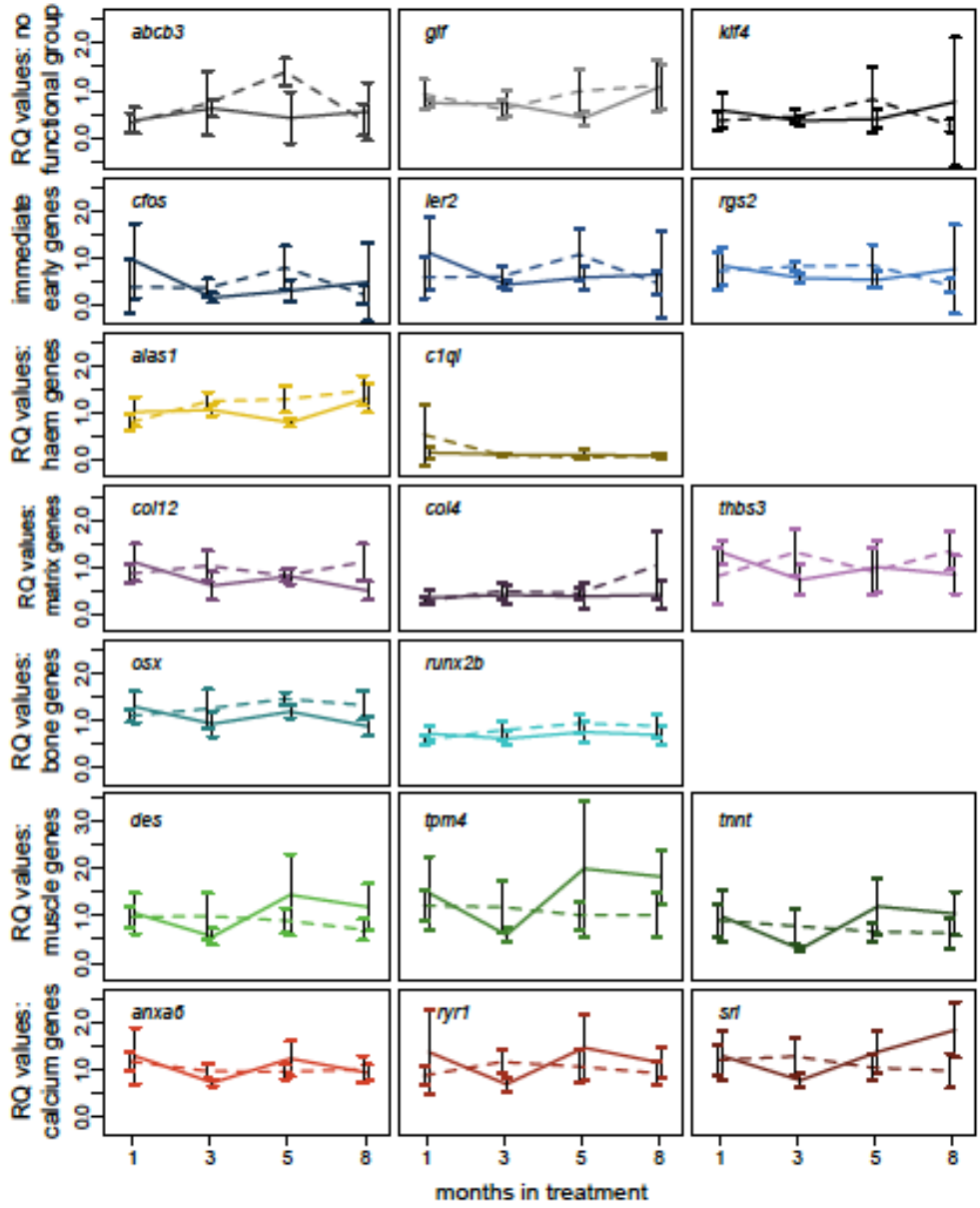


Fig. S.III.6 Gene expression levels between the diet groups (SD = solid line, HD = dashed line) across the four time-points and according to their functional categories. Error-bars indicate 95% CIs.

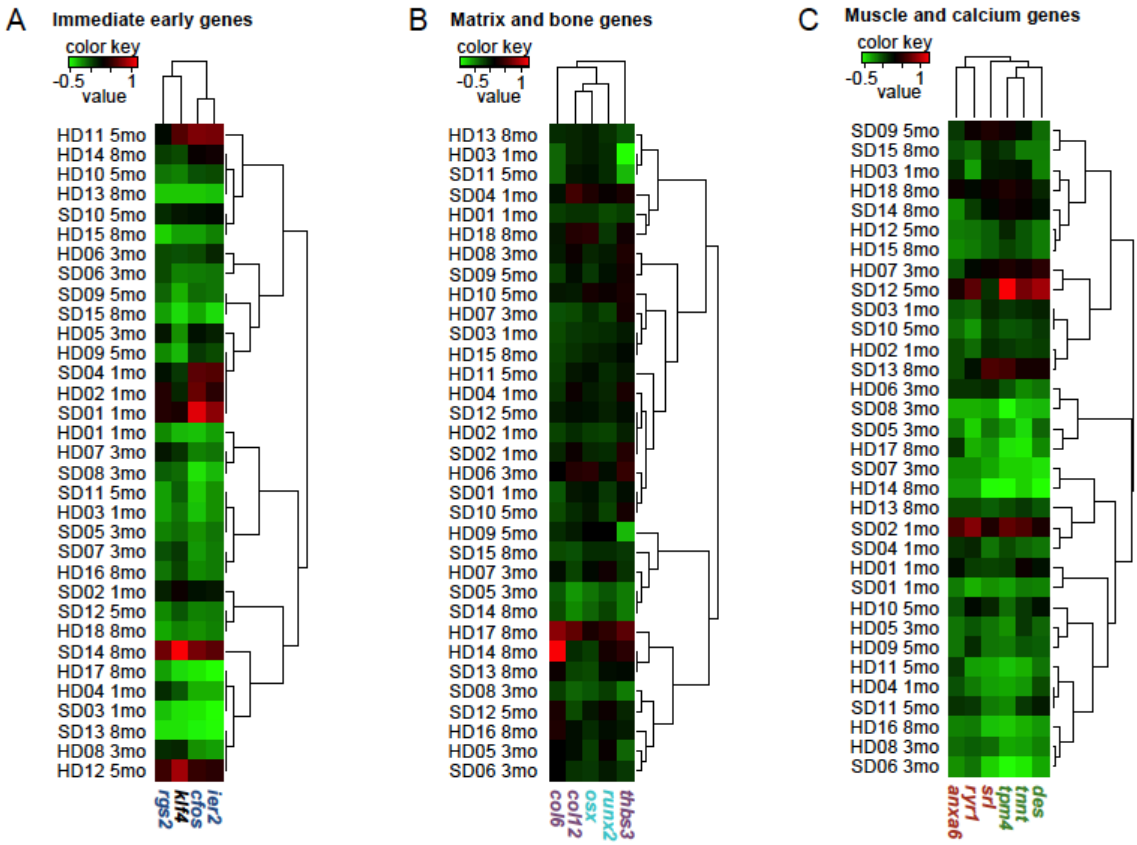


Fig. S.III.7 Individual cluster analyses for the three identified main clusters of gene expression.

Table S.III.1 List of gene abbreviations, the full names, sequences of the used primer pairs and assigned functional categories

Gene abbr.	Full name	Primer sequence	Functional category
gif	gastric intrinsic factor-like	Fwd: gacctccacctggacacca	no class
		Rev: ttgccactctaataaaccc	
abcb3	ATP-binding cassette-transporter subfamily-B type 3	Fwd: taagcggtggcaaagttaa	
		Rev: gcacagactcggtggtttc	
klf4	Kruppel-like factor 4	Fwd: gacttggatggtcggtgga	
		Rev: ttgaaggcaaatcgtggtt	
rgs2	Regulator of G-protein signalling 2	Fwd: tccaccgtattcattccag	immediate early genes
		Rev: caatcattcccgcttcctc	
ier2	immediate early response 2	Fwd: gcggaaagtttgaggaaga	
		Rev: gcggcagtcacttgtttag	
cfos	c-fos protein	Fwd: ggaagcagcagcaatgagc	
		Rev: atagccctgtgatcggcac	
alas1	5-aminolevulinate synthase	Fwd: tttgttgccaatgactccac	haem pathway genes
		Rev: gccgctgttcctgatacct	
c1ql	complement component 1, q subcomponent-like	Fwd: gcagaaactggacccttta	
		Rev: ctggtgcagtgaaaatacc	
thbs3	thrombospondin 3	Fwd: aaacaccggctggatggacaaa	matrix related genes
		Rev: gcgtgtcccttcaaacagctttac	
col12	collagen 12A1	Fwd: aggggagattttgcatcccagaac	
		Rev: tggtccaggagtgttgctacgata	
col6	collagen 6A3	Fwd: gacctacccaaagtgcaacagcaa	
		Rev: aaatcacgcatggcaggaaagc	
runx2b	Runt-related transcription factor 2b	Fwd: gtatccgtcctacctgagtccc	bone related genes
		Rev: cactgatggcaggcagac	
osx	osterix	Fwd: acaagtaaacatgctcgtg	
		Rev: ctacttttccttttg	
des	desmin	Fwd: ctttccagcatgtgctacgtgtct	muscle related genes
		Rev: cgcatgctgtctgaggaatgctaa	
tpm4	tropomyosin alpha-4 chain-like isoform 1	Fwd: gccctcaacgatatgactacactg	
		Rev: agtagcactgtggagtggaggaaa	
tnnt	troponin T	Fwd: ctgcagaaacacagcaagaagg	
		Rev: tgttgctgtggatgtcatcg	
anxa6	annexin a6 isoform 1	Fwd: ctgtgatggttcacgtcagcatct	calcium pathway genes
		Rev: atcaagtgacctatgacctcgcag	
ryr1	ryanodine receptor 1	Fwd: gtgtggaagatgtaccaggaacga	
		Rev: gcagtatgaggaccagcttggata	
srl	sarcalumenin-like	Fwd: attcagtgcgatctgctgtgctgt	
		Rev: aggagcaccattccctcttccatt	
actinr	ActinR	Fwd: tcaagagggtctatgggaact	*housekeep. genes*
		Rev: caaacactgaggcaaagca	
twinfilin	Twinfilin	Fwd: tcccaatgagaatccctaa	
		Rev: tgtaccctggcatagaataaa	

Table S.III.2 Comparisons of linear morphometric measures between diet groups across the developmental time-course

residuals of	M	*W*	*n*	*p*	residuals of	M	*W*	*n*	*p*
jaw weight	1	16	8	0.029	average tooth size	1	7	8	0.886
	3	4	8	0.343		3	8	8	1.000
	5	0	8	0.029		5	0	8	0.029
	8	0	9	0.024		8	3	9	0.167
jaw width	1	11	8	0.486	largest tooth size	1	8	8	1.000
	3	8	8	1.000		3	3	8	0.200
	5	1	8	0.057		5	6	8	0.686
	8	6	9	0.548		8	12	9	0.5476
jaw length	1	14	8	0.114	# teeth in dental area	1	10	8	0.686
	3	6	8	0.686		3	6	8	0.686
	5	6	8	0.686		5	4	8	0.343
	8	0	9	0.024		8	4	9	0.262
suture width	1	5	8	0.486	dental area length	1	6	8	0.686
	3	4	8	0.343		3	7	8	0.886
	5	0	8	0.029		5	2	8	0.114
	8	1	9	0.048		8	2	9	0.095
average horn width	1	6	8	0.686	centroid size	1	16	8	0.029
	3	0	8	0.029		3	3	8	0.200
	5	0	8	0.029		5	2	8	0.114
	8	3	9	0.167		8	1	9	0.048

Table S.III.3 ANCOVAs on linear morphometric measures considering diet groups, time-point and SL of the fish

	Est.	Std.Err.	*t* value	*p*		Est.	Std.Err.	*t* value	*p*
log jaw weight					**average tooth area**				
(Interc.)	0.465	0.645	0.721	0.476	(Interc.)	-0.012	0.056	-0.209	0.836
Diet	-1,070	0.411	-2,604	0.014	Diet	-0.042	0.035	-1,185	0.246
SL	0.028	0.011	2,540	0.017	SL	0.001	0.001	1,125	0.270
Diet:SL	0.024	0.007	3,335	0.002	Diet:SL	0.001	0.001	1,592	0.122
jaw width					**largest tooth area**				
(Interc.)	-0.278	1,541	-0.181	0.858	(Interc.)	-0.308	0.120	-2,576	0.015
Diet	-0.250	0.981	-0.254	0.801	Diet	0.072	0.076	0.939	0.356
SL	0.110	0.027	4,137	<0.001	SL	0.008	0.002	3,995	<0.001
Diet:SL	0.008	0.017	0.491	0.627	Diet:SL	-0.001	0.001	-0.839	0.408
jaw length					**number teeth in dental area**				
(Interc.)	2,534	1,403	1,806	0.081	(Interc.)	12,498	8,989	1,390	0.175
Diet	-1,078	0.894	-1,206	0.238	Diet	-6,166	5,725	-1,077	0.290
SL	0.046	0.024	1,905	0.067	SL	0.047	0.155	0.303	0.764
Diet:SL	0.023	0.016	1,458	0.156	Diet:SL	0.113	0.100	1,135	0.266
suture width					**dental area length**				
(Interc.)	0.456	0.378	1,205	0.238	(Interc.)	-0.282	0.859	-0.328	0.746
Diet	-0.687	0.241	-2,851	0.008	Diet	-0.165	0.547	-0.302	0.765
SL	-0.009	0.007	-1,411	0.169	SL	0.036	0.015	2,455	0.020
Diet:SL	0.016	0.004	3,814	<0.001	Diet:SL	0.008	0.010	0.793	0.434
average horn width					**centroid size**				
(Interc.)	0.091	0.206	0.443	0.661	(Interc.)	2,182	1,624	1,343	0.190
Diet	-0.101	0.131	-0.767	0.449	Diet	-1,384	1,034	-1,338	0.191
SL	0.003	0.004	0.892	0.380	SL	0.114	0.028	4,087	<0.001
Diet:SL	0.004	0.002	1,714	0.097	Diet:SL	0.032	0.018	1,785	0.085

Table S.III.4 Comparisons of relative gene expression (RQ) between diet groups across the developmental time-course using Wilcoxen Signed Rank tests.

gene	M	*W*	*n*	*p*	gene	M	*W*	*n*	*p*
rgs2	1	10	8	0.686	*col6*	1	10	8	0.686
	3	0	8	0.029		3	5	8	0.486
	5	4	8	0.343		5	4	8	0.343
	8	9	9	1.000		8	3	9	0.167
cfos	1	10	8	0.686	*tpm4*	1	9	8	0.886
	3	3	8	0.200		3	1	8	0.057
	5	2	8	0.114		5	12	8	0.343
	8	8	9	0.905		8	16	9	0.095
c1ql	1	4	8	0.343	*ryr1*	1	11	8	0.486
	3	9	8	0.886		3	0	8	0.029
	5	11	8	0.486		5	11	8	0.486
	8	9	9	1.000		8	14	9	0.262
abcb3	1	10	8	0.686	*tnnt*	1	8	8	1.000
	3	8	8	1.000		3	0	8	0.029
	5	1	8	0.057		5	14	8	0.114
	8	9	9	1.000		8	14	9	0.262
alas1	1	11	8	0.486	*thbs3*	1	14	8	0.114
	3	3	8	0.200		3	2	8	0.114
	5	0	8	0.029		5	8	8	1.000
	8	7	9	0.714		8	4	9	0.262
ier2	1	11	8	0.486	*des*	1	9	8	0.886
	3	4	8	0.343		3	2	8	0.114
	5	2	8	0.114		5	12	8	0.343
	8	8	9	0.905		8	16	9	0.095
klf4	1	11	8	0.486	*col12*	1	10	8	0.686
	3	5	8	0.486		3	2	8	0.114
	5	7	8	0.886		5	8	8	1.000
	8	6	9	0.548		8	0	9	0.024
gif	1	6	8	0.686	*osx*	1	11	8	0.486
	3	11	8	0.486		3	4	8	0.343
	5	1	8	0.057		5	0	8	0.029
	8	10	9	0.905		8	2	9	0.095
anxa6	1	6	8	0.686	*runx2b*	1	12	8	0.343
	3	0	8	0.029		3	3	8	0.200
	5	13	8	0.200		5	4	8	0.343
	8	8	9	0.905		8	6	9	0.548
srl	1	9	8	0.886					
	3	1	8	0.057					
	5	11	8	0.486					
	8	16	9	0.095					

Table S.III.5 Loadings and importance of components for the PCA on SL corrected linear morphometric measurements (Fig. S5). rjwe = jaw weight; rjwi = jaw width; rjle = jaw length; rswi = suture width; rhav = average horn width; ravt = average tooth size; rlth = largest tooth size; rnth = # teeth in dental area; rdal = dental area length; rces = centroid size.

Correlation matrix

ID	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5	Comp.6	Comp.7	Comp.8	Comp.9	Comp.10
HD01	-0.408	-1.477	-0.083	2.155	-2.089	0.205	0.503	1.044	-0.115	-0.163
HD02	2.049	-2.541	0.938	-0.356	-0.473	-1.014	-0.533	-0.145	0.355	-0.087
HD03	-0.872	-0.280	0.184	-1.874	0.177	1.109	0.096	0.000	0.359	0.059
HD04	0.656	-0.133	0.119	0.149	-0.061	0.178	0.753	-0.054	0.288	-0.165
HD05	2.268	0.238	0.053	0.421	-0.422	-0.931	0.384	-0.010	0.354	0.057
HD06	-1.945	0.470	-0.005	-0.036	0.110	0.203	-0.538	0.270	0.058	0.005
HD07	-1.049	0.312	-0.592	-0.053	-0.815	-0.573	-0.281	0.559	0.242	0.036
HD08	1.558	-0.548	0.741	-0.349	0.854	-0.490	-0.129	0.256	-0.133	-0.235
HD09	-1.572	-1.035	-0.202	-0.090	1.346	-0.088	0.087	-0.055	-0.624	0.470
HD10	-1.539	0.005	1.368	0.301	0.723	-0.815	-1.163	-0.519	-0.190	0.112
HD11	-2.532	0.044	0.111	2.673	-0.199	-0.294	-0.482	0.263	0.052	0.704
HD12	-2.574	-0.484	-1.524	-1.292	0.610	-0.166	-0.114	0.153	0.019	0.188
HD13	-0.123	-0.841	-0.871	1.129	1.535	0.496	0.244	0.323	0.002	-0.379
HD14	-0.551	3.322	2.943	0.473	0.026	-0.575	0.427	-0.034	0.404	-0.039
HD15	-2.384	0.151	-1.157	0.259	0.553	-0.247	-0.196	0.032	-0.397	-0.757
HD16	-2.665	0.149	0.870	0.257	0.331	-0.495	0.155	-0.225	-0.054	-0.658
HD17	-1.994	-0.425	-0.147	0.754	0.647	0.700	-0.349	0.334	-0.493	-0.005
HD18	-4.465	-2.314	0.970	-0.152	-0.224	0.651	1.283	-0.688	0.378	0.280
SD01	0.746	1.165	0.583	-0.066	-0.916	-0.224	0.898	-0.405	-0.666	-0.077
SD02	-0.859	0.088	0.045	-0.824	-0.254	0.256	0.449	-0.006	0.047	0.055
SD03	2.448	0.314	-0.797	0.034	0.293	0.795	0.416	-0.405	-0.097	-0.177
SD04	1.726	3.274	-1.860	0.189	-0.446	-1.060	0.630	-0.213	-0.398	0.286
SD05	0.897	0.016	-0.302	0.079	-0.260	0.385	-0.332	-0.276	-0.191	-0.075
SD06	1.310	0.777	-0.193	-0.614	0.641	1.223	0.256	0.045	-0.052	0.364
SD07	4.846	-3.381	-0.045	-0.187	0.461	-0.879	0.053	0.105	0.179	0.182
SD08	3.642	-0.131	1.479	0.394	1.799	0.336	0.152	0.001	-0.184	0.001
SD09	-1.419	1.438	0.432	-1.144	0.751	-0.568	-0.077	0.580	0.442	0.081
SD10	-0.707	-1.514	-0.467	-0.263	-2.055	-0.278	-0.299	-1.070	-0.369	-0.046
SD11	1.446	0.834	-0.828	1.053	-0.660	1.229	-0.495	-0.154	1.110	-0.141
SD12	-0.733	0.984	-1.728	-0.706	0.063	-0.407	-0.759	-0.608	0.552	-0.043
SD13	2.921	1.117	0.679	0.957	-0.545	1.387	-0.966	-0.333	-0.431	0.075
SD14	1.425	0.214	-1.895	-0.415	0.298	-0.511	0.414	0.360	-0.003	0.094
SD15	0.450	0.194	1.180	-2.855	-1.799	0.463	-0.486	0.874	-0.443	-0.001

Loadings:

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5	Comp.6	Comp.7	Comp.8	Comp.9	Comp.10
rjwe	-0.377		0.416	-0.203	-0.225	0.255	0.522		0.445	-0.221
rjwi	-0.37	0.251	0.206	-0.123	-0.245	0.468	-0.573		-0.125	0.345
rjle	-0.308	0.427	-0.236	0.248	-0.195	-0.367	0.359			0.551
rswi	-0.335	-0.281	0.151	0.253	0.492	0.28	0.322		-0.521	0.155
rhav	-0.245	-0.359	0.508	0.223		-0.568	-0.329	0.121	0.206	0.108
ravt	-0.345	-0.39	-0.317					-0.762	0.183	
rlth	-0.191	-0.351	-0.232	-0.69	-0.18	-0.245		0.343	-0.29	0.112
rnth	-0.139	0.385		-0.461	0.723	-0.138		-0.12	0.229	
rdal	-0.341	-0.1	-0.541	0.27	0.173	0.159	-0.186	0.497	0.36	-0.204
rces	-0.404	0.336			-0.119	-0.274		-0.124	-0.414	-0.659

Importance of components:

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5	Comp.6	Comp.7	Comp.8	Comp.9	Comp.10
Standard deviation	2.048	1.344	1.027	1.004	0.890	0.680	0.533	0.427	0.376	0.273
Proportion of Variance	0.419	0.181	0.106	0.101	0.079	0.046	0.028	0.018	0.014	0.007
Cumulative Proportion	0.419	0.600	0.705	0.806	0.886	0.932	0.960	0.978	0.993	1.000

Table S.III.6 Loadings and importance of components for the PCA on gene expression (Fig. III.5).

Loadings:

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5	Comp.6	Comp.7	Comp.8	Comp.9	Comp.10	Comp.11	Comp.12	Comp.13	Comp.1	4 Comp.1	Comp.16	Comp.17	Comp.18	Comp.19
rgs2		-0.385						-0.179	0.230		-0.240	0.182	0.369	0.125	-0.208	0.330	0.519	0.130	0.248
cfos		-0.527			0.182			0.179	-0.181	-0.243	0.295							0.105	-0.652
ier2		-0.531		-0.107	0.155		-0.109	0.196	-0.160	-0.110	0.147		-0.201		0.166	-0.117	-0.186	-0.174	0.636
alas1		-0.152			-0.413		0.261	0.156	-0.367	0.500			-0.144	0.264	-0.380	0.173	-0.125	0.125	
c1ql			-0.136			-0.256		0.362	0.557	-0.145			-0.220	0.593		0.125			-0.103
ryr1	0.367		0.178				-0.201	-0.137	-0.393	-0.167	-0.470		0.147	0.415	0.325			-0.155	-0.142
anxa6	0.218		0.141			-0.325				-0.194	-0.262	-0.229	-0.498	-0.342	-0.137	0.304		0.401	
srl	0.309			0.262			0.622	0.355	-0.122	-0.123		0.353		-0.254	0.256	0.175			
tnnt	0.375							0.108			-0.126	0.288			-0.327	-0.731	0.199	0.189	
tmp4	0.630	-0.119			-0.169			0.157	0.214	0.170	0.278	-0.510	0.257					-0.141	
des	0.372						-0.221	-0.406		0.101	0.325	0.552	-0.257			0.282	-0.209		
col6			0.450	-0.287	-0.400	0.410	-0.162	0.274	0.153	-0.186	-0.143	0.170	-0.161	-0.180			0.182	-0.257	
col12			0.401		0.144	-0.443			0.105		-0.111	0.148	0.326	-0.155	-0.438		-0.343	-0.357	
thbs3			0.634	0.104	0.124	0.148	0.433	-0.357	0.123		0.222	-0.183	-0.126	0.246		-0.169		0.158	
osx			0.286		0.120	-0.425	-0.160	0.146		0.542	0.151	0.166			0.339		0.453		
runx2b			0.212		-0.136		-0.275	0.193	0.124			0.146	0.381		0.199		-0.362	0.672	0.110
gif				-0.146	-0.641	-0.473	0.143	-0.271		-0.310	0.259		0.124		0.161		0.103		
abcb3		-0.104		0.870	-0.266		-0.262			-0.101			-0.147		-0.168				
klf4		-0.472			-0.129		0.107	-0.235	0.390	0.287	-0.396			-0.247	0.234	-0.201	-0.280		-0.213

Importance of components:

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5	Comp.6	Comp.7	Comp.8	Comp.9	Comp.10	Comp.11	Comp.12	Comp.13	Comp.14	Comp.15	Comp.16	Comp.17	Comp.18	Comp.19
Standard deviation	1.131	0.841	0.666	0.519	0.493	0.395	0.327	0.281	0.263	0.242	0.222	0.169	0.161	0.137	0.110	0.090	0.083	0.061	0.044
Proportion of Variance	0.358	0.199	0.124	0.075	0.068	0.044	0.030	0.022	0.019	0.016	0.014	0.008	0.007	0.005	0.003	0.002	0.002	0.001	0.001
Cumulative Proportion	0.358	0.557	0.681	0.757	0.825	0.869	0.898	0.921	0.940	0.956	0.970	0.978	0.986	0.991	0.994	0.996	0.998	0.999	1.000

Table S.III.7 Detected number of transcription factor binding sites (TFBSs) in the promoters of candidate genes for focal transcription factors (TFs).

candidate gene	Transcription factor binding sites					
	AP1	CREB1	RUNX2B	KLF4	PRRX2	ARID3A
abcb3	0 / 0 / 5	0 / 0 / 1	0 / 0 / 2	0 / 0 / 1	1 / 0 / 3	6 / 2 / 4
alas1	0 / 2 / 8	0 / 0 / 2	0 / 0 / 1	1 / 0 / 1	3 / 0 / 6	5 / 3 / 9
c1ql	0 / 1 / 4	0 / 0 / 0	0 / 0 / 1	0 / 1 / 2	7 / 0 / 1	3 / 3 / 8
gif	0 / 1 / 3	0 / 0 / 1	0 / 2 / 0	0 / 0 / 1	4 / 0 / 3	2 / 1 / 4
cfos	2 / 2 / 7	0 / 4 / 5	0 / 0 / 2	0 / 0 / 0	7 / 0 / 1	9 / 1 / 9
rgs2	2 / 7 / 10	0 / 2 / 2	0 / 0 / 1	0 / 1 / 1	5 / 0 / 2	4 / 4 / 5
ier2	2 / 3 / 6	0 / 0 / 3	0 / 0 / 0	0 / 0 / 1	6 / 0 / 0	2 / 2 / 3
klf4	0 / 0 / 3	0 / 1 / 3	0 / 0 / 0	0 / 2 / 2	3 / 0 / 2	4 / 2 / 5
col12	2 / 1 / 10	0 / 0 / 0	0 / 0 / 2	0 / 0 / 1	4 / 0 / 1	4 / 4 / 9
col6	1 / 1 / 5	0 / 0 / 0	0 / 0 / 1	0 / 0 / 0	8 / 0 / 3	6 / 4 / 1
thbs3	0 / 0 / 4	0 / 0 / 0	0 / 0 / 0	0 / 0 / 0	6 / 0 / 1	0 / 0 / 3
runx2b	1 / 1 / 4	2 / 0 / 3	0 / 2 / 4	0 / 0 / 4	5 / 0 / 4	3 / 1 / 8
osx	0 / 1 / 1	0 / 0 / 0	0 / 1 / 0	0 / 1 / 2	13 / 0 / 1	10 / 2 / 12
anxa6	0 / 2 / 8	0 / 0 / 2	0 / 0 / 1	0 / 1 / 0	2 / 0 / 3	1 / 0 / 3
ryr1	0 / 2 / 6	0 / 0 / 0	0 / 0 / 1	0 / 1 / 2	3 / 0 / 1	3 / 4 / 2
srl	0 / 1 / 5	0 / 1 / 1	0 / 0 / 3	0 / 0 / 1	0 / 0 / 3	2 / 0 / 4
tnnt	0 / 1 / 4	0 / 1 / 1	0 / 0 / 1	0 / 1 / 1	6 / 0 / 1	2 / 2 / 2
tpm4	0 / 0 / 7	0 / 0 / 0	0 / 0 / 1	0 / 0 / 1	7 / 0 / 3	4 / 6 / 3
des	1 / 2 / 5	0 / 0 / 1	0 / 0 / 0	0 / 1 / 7	11 / 0 / 8	6 / 3 / 5

Chapter IV

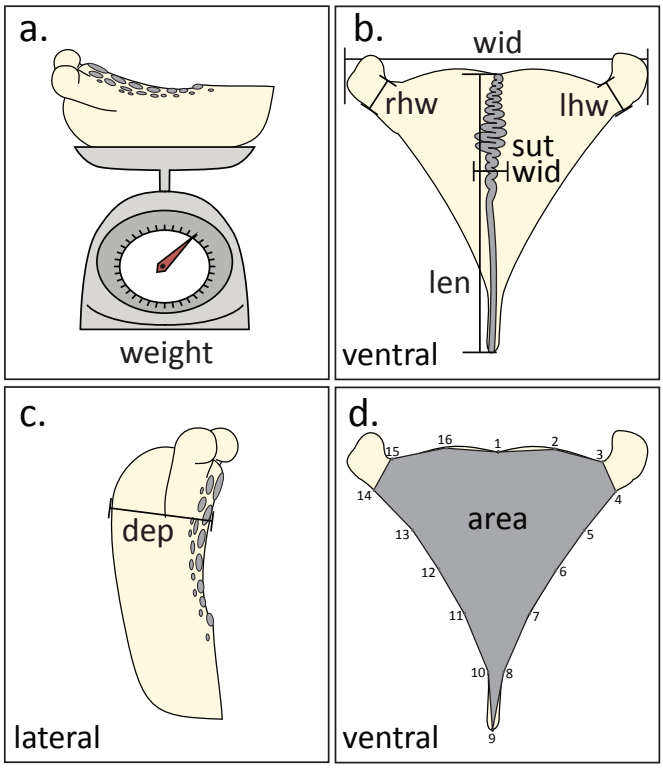


Fig. S.IV.1 Measurements made for linear and geometric morphometric analyses.

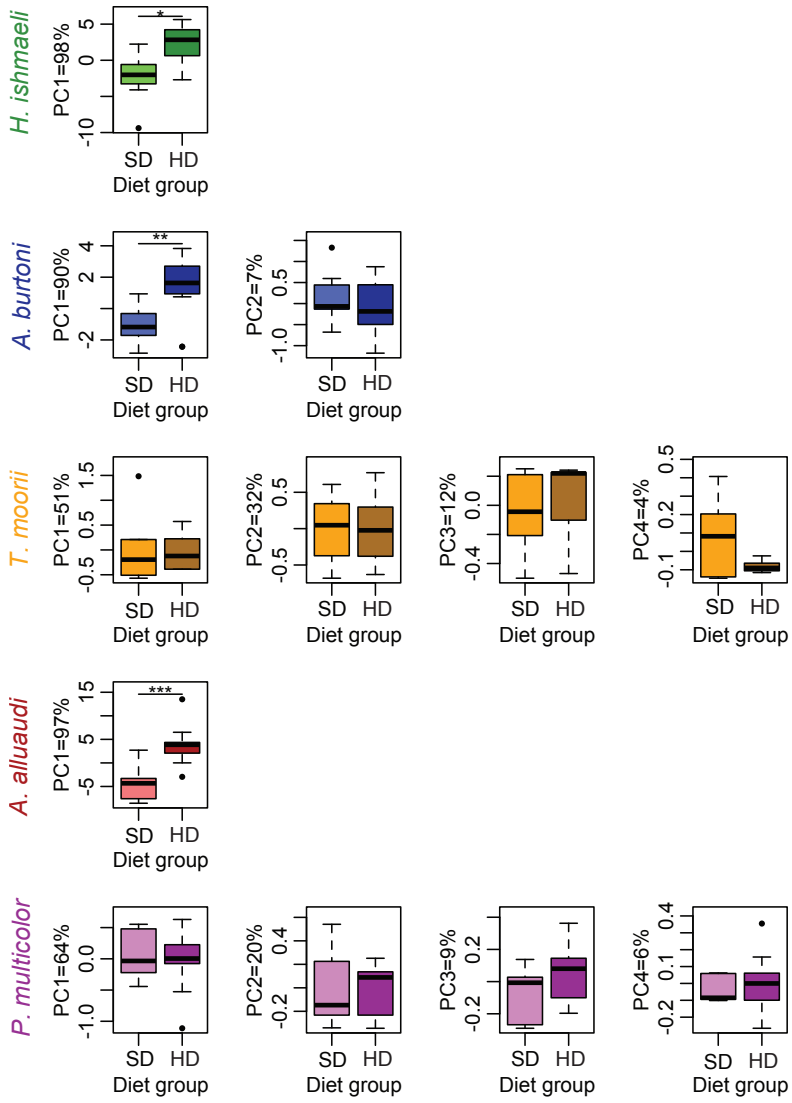


Fig. S.IV.2 Boxplots of PC scores for all considered PCs of the species-wise PCAs on linear morphometric measurements.

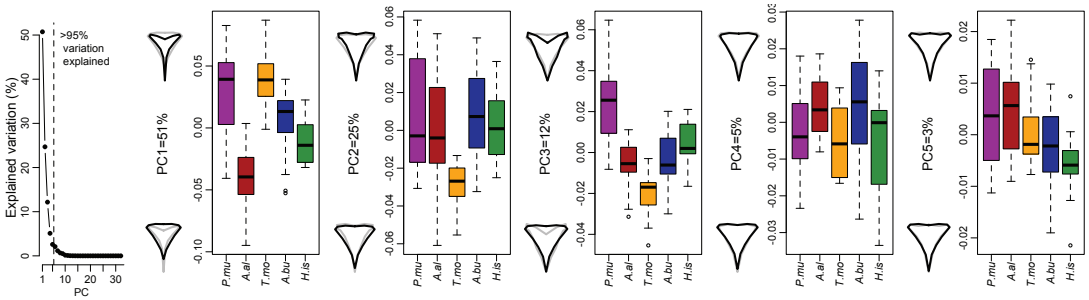


Fig. S.IV.3 Scree-plot for all PCs and boxplots of PC scores for considered PCs 1-5 of the among-species PCA on geometric morphometric measurements.

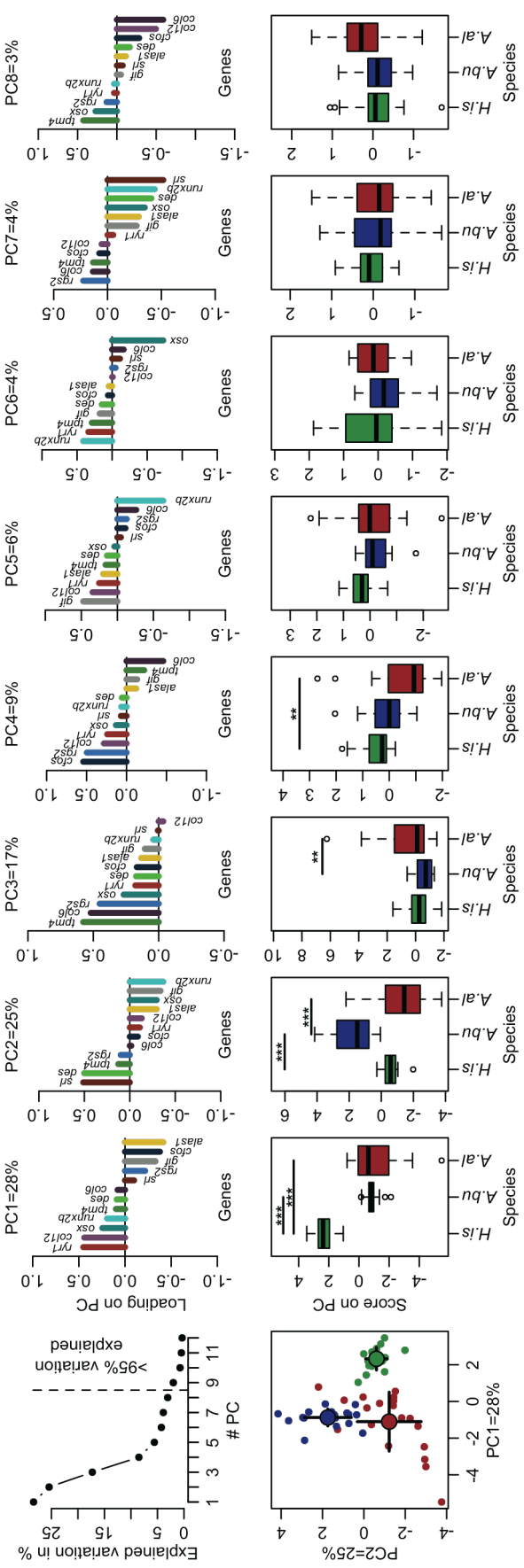


Fig. S.IV.4 Scree-plot, scatterplot of PC1 and PC2, PC loadings and PC scores (as boxplots) of the among-species PCA on gene expression measurements.

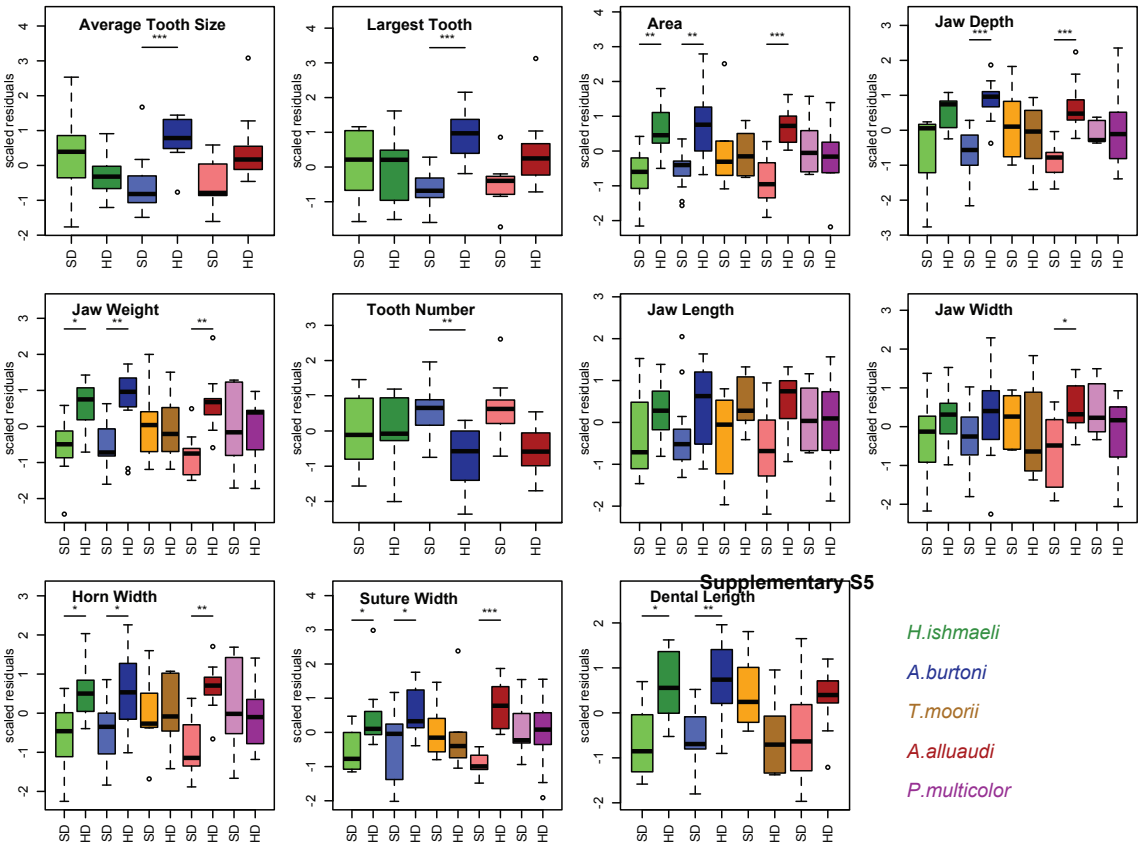


Fig. S.IV.5 Boxplots for all linear morphometric measurements and all species. Linear measurements have been standardized species-wise prior to plotting.

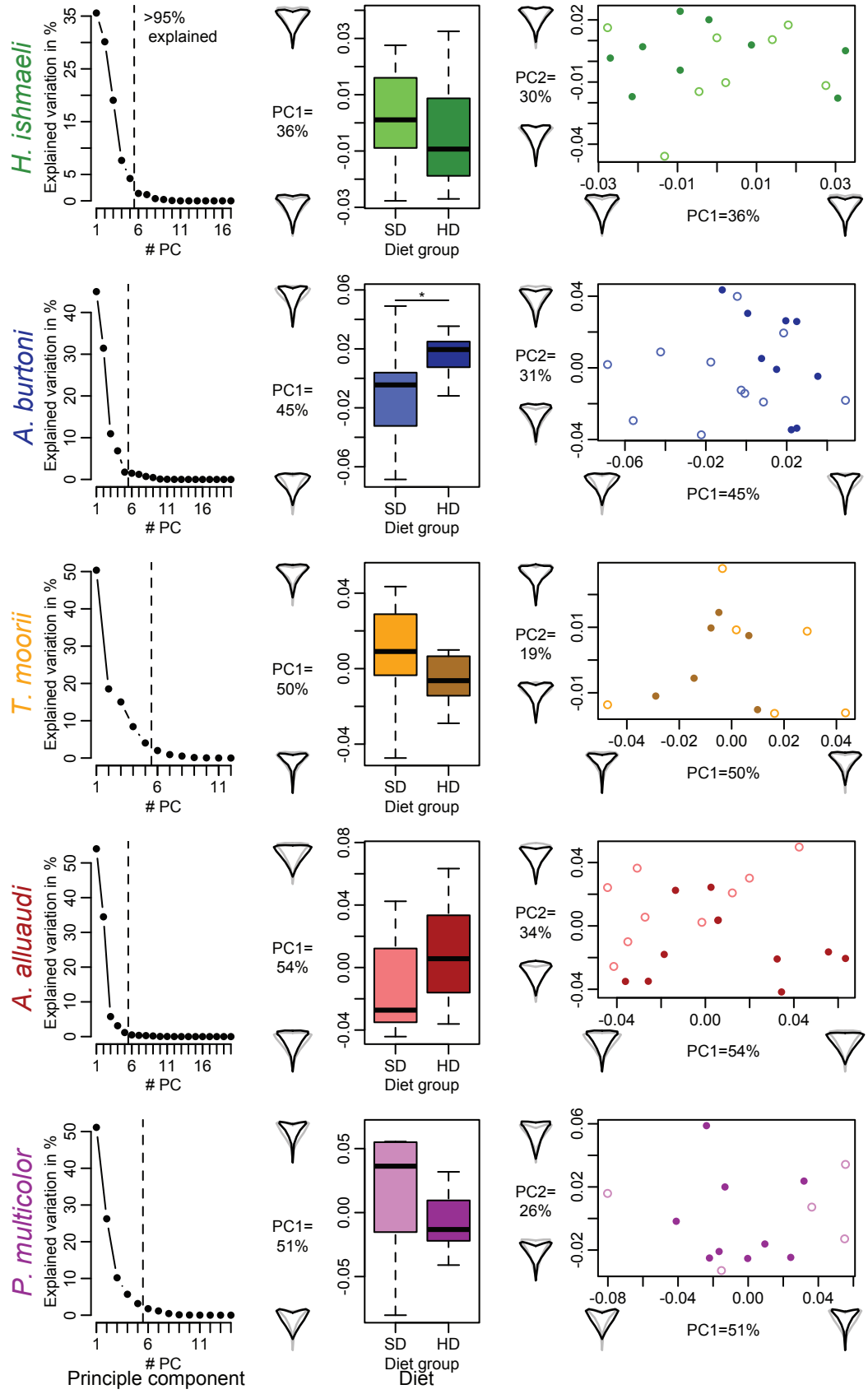


Fig. S.IV.6 Species-specific PCAs on geometric morphometric measurements. Scree-plots indicating explanatory value of each PC, boxplots showing the scores on PC1 and a scatterplot showing the scores on PC1 and PC2.

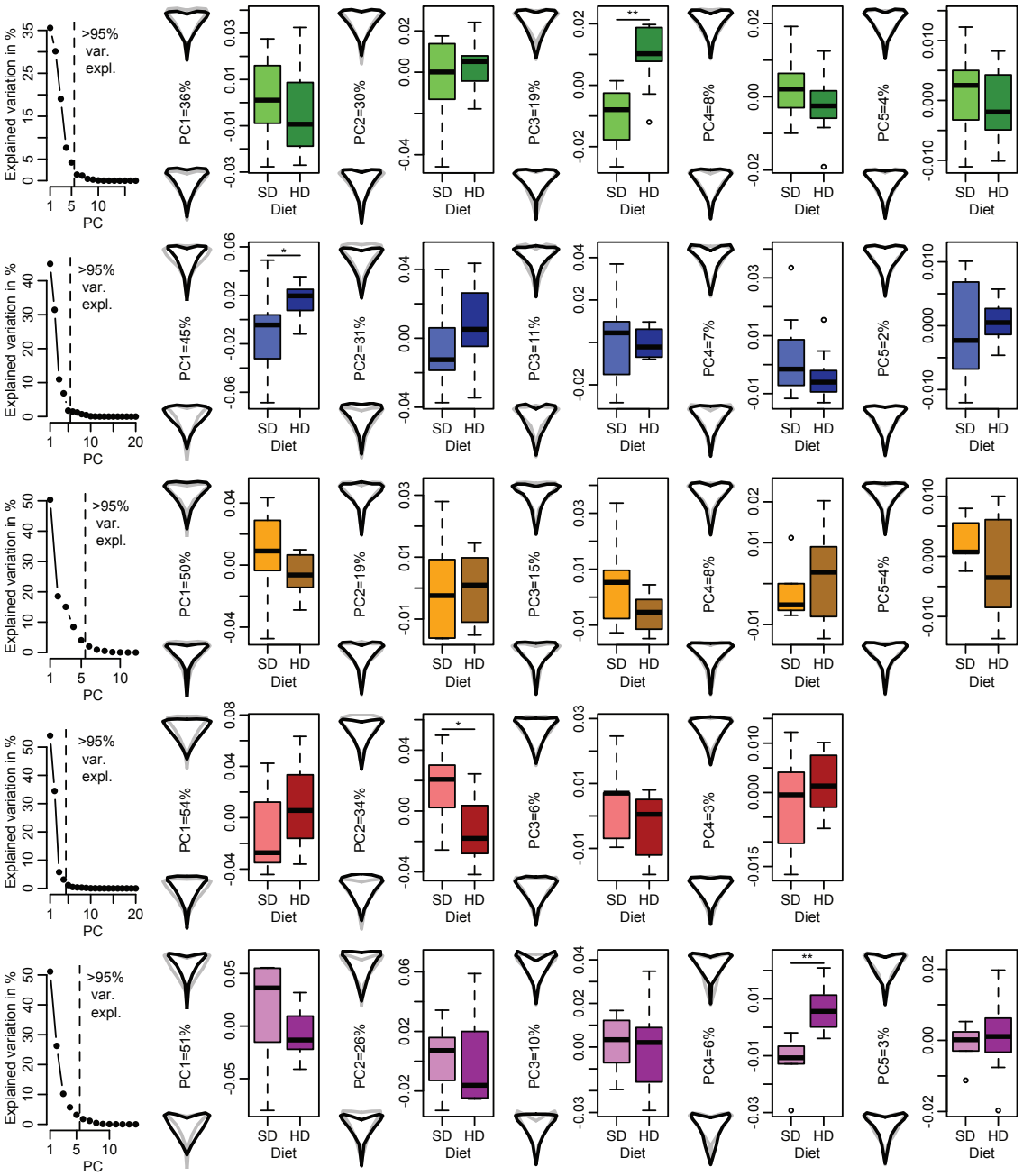


Fig. S.IV.7 Species-specific PCAs on geometric morphometric measurements. Scree-plots indicating explanatory value of each PC, boxplots showing the scores on all considered PCs per species.

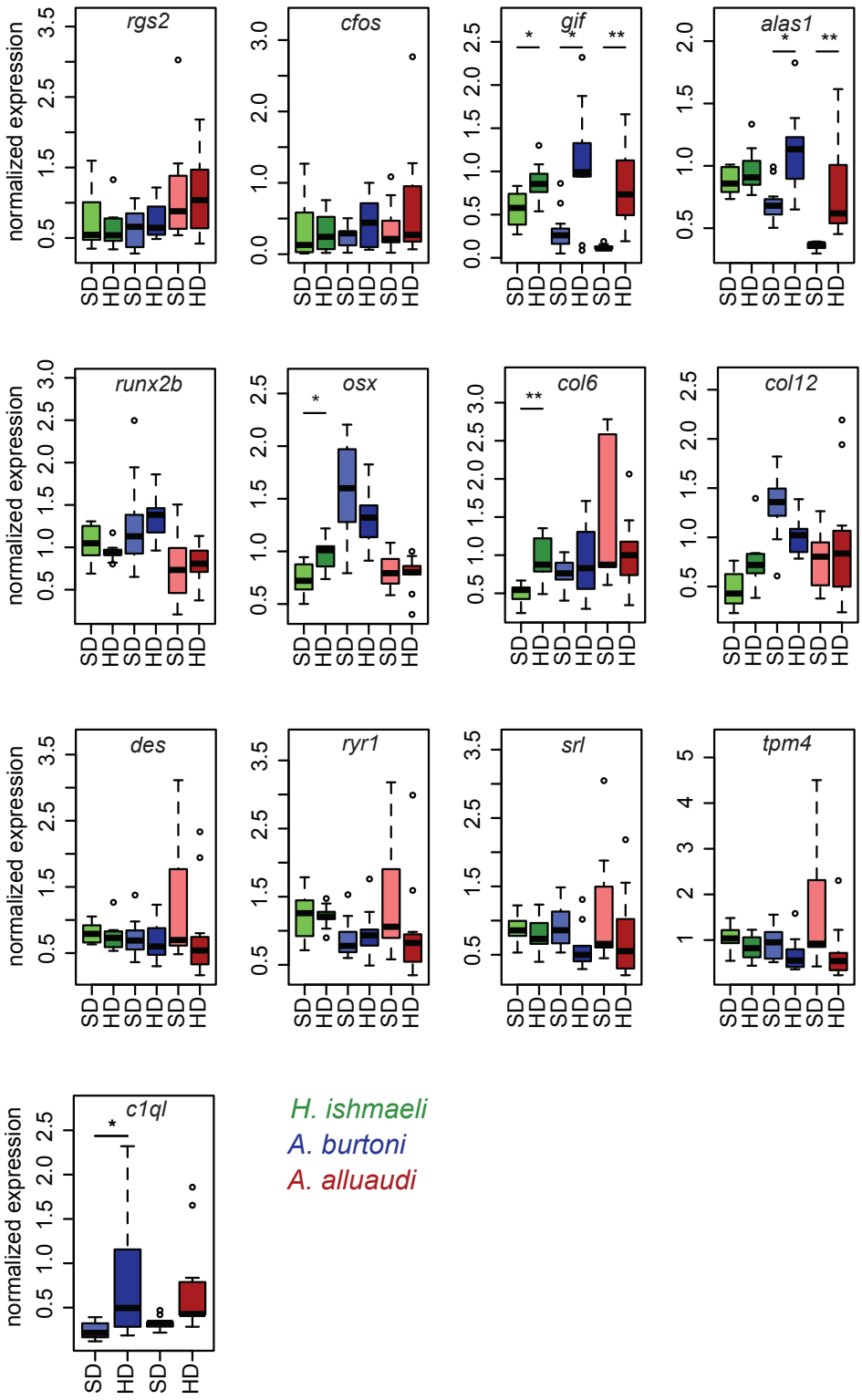


Fig. S.IV.8 Species-wise standardized gene expression values for all candidate genes in the three considered species.

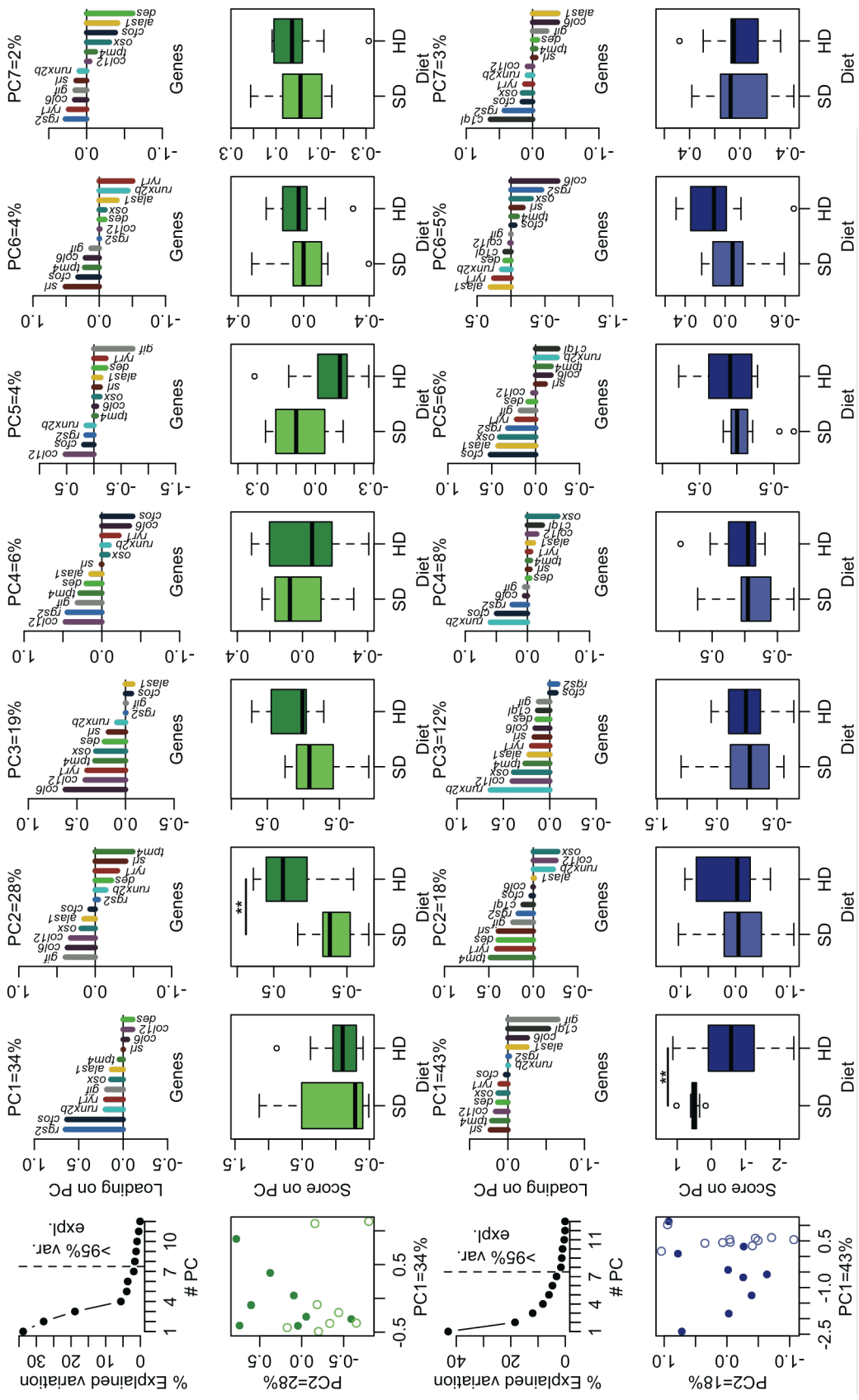


Fig. S.IV.9a Scree-plot, scatterplot of PC1 and PC2, PC loadings and PC scores (as boxplots) of species-wise PCAs on gene expression measurements (*H. ishmaeli* & *A. burtoni*).

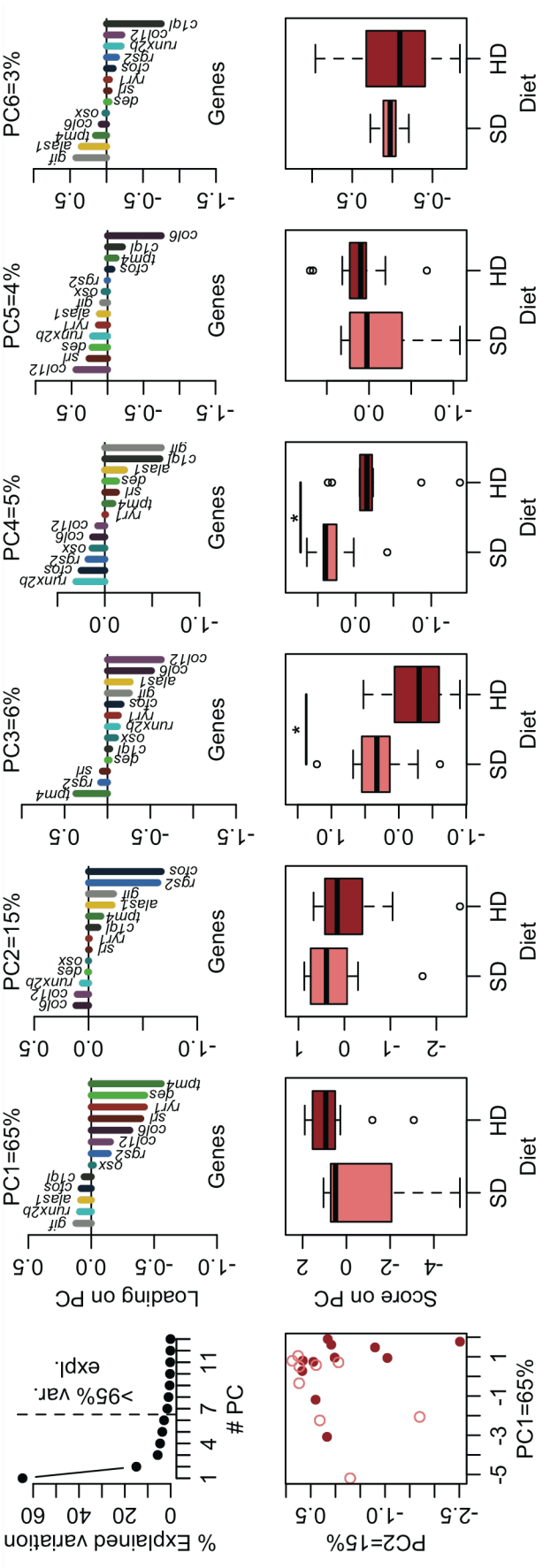


Fig. S.IV.9b Scree-plot, scatterplot of PC1 and PC2, PC loadings and PC scores (as boxplots) of species-wise PCAs on gene expression measurements (*A. alluaudi*).

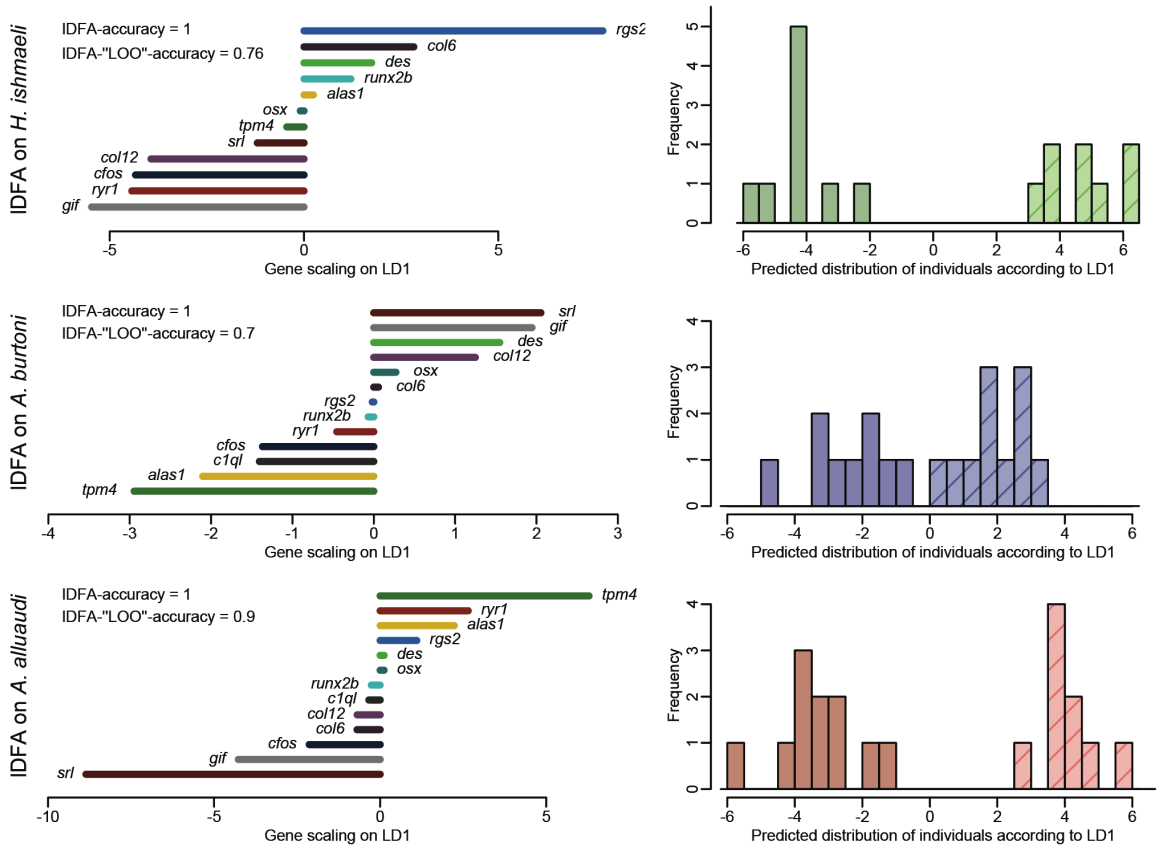


Fig. S.IV.10 Contributions of genes to lDFA diet-predictors in the three considered species and histograms of predicted diet-group memberships of specimens.

Table S.IV.1 qRT-PCR primers used in this study

A. alluaudi

gene	name	sequence
actinR	fx_actinR_f2	TCAAGAGGGTCTATGGGAACT
	fx_actinR_r2	CAAACACTGAGGCAAAGCA
twinfilin	fx_twinfilin_f2	TCCCAATGAGAATCCCTAA
	fx_twinfilin_r2	TGTACCCTGGCATAGAATAAA
runx2b	xf-RT-Runx2b-F1	GTATCCGTCCTACCTGAGTCCC
	xf-RT-Runx2b-R1	CACTGATGGCAGGCAGACC
osx	XF-RT-osx–f2	ACAAGTAAACATGCTCGTG
	XF-RT-osx–r2	CTACTTTTCCTTTTGTGGC
c-fos	fx_e7fkf1_f1	GGAAGCAGCAGCAATGAGC
	fx_e7fkf1_r1	ATAGCCCTGTGATCGGCAC
rgs2	fx_rgs2_f1	TCCACCGTATTCATTCCAG
	fx_rgs2_r1	CAATCATTCCCGCTTCCTC
gif	fx_gif_f2	GACCTCCACCTGGACACCA
	fx_gif_r2	TTGCCACTCTAATAAACCC
alas1	fx_alas1_f1	TTTGTTGCCAATGACTCCAC
	fx_alas1_r1	GCCGCTGTTCCTGATACCT
c1q-like	fx_c1ql_f3	GCAGAAACTGGACCCTTTA
	fx_c1ql_r3	CTGGTGCAGTGAAAATACC

A. burtoni

gene	name	sequence
actinR	fx_actinR_f2	TCAAGAGGGTCTATGGGAACT
	fx_actinR_r2	CAAACACTGAGGCAAAGCA
twinfilin	fx_twinfilin_f2	TCCCAATGAGAATCCCTAA
	fx_twinfilin_r2	TGTACCCTGGCATAGAATAAA
runx2b	xf-RT-Runx2b-F1	GTATCCGTCCTACCTGAGTCCC
	xf-RT-Runx2b-R1	CACTGATGGCAGGCAGACC
osx	IK-Abuosx-F1	AAAGAGCTCTGAAGCCGCAAAAGG
	IK-Abuosx-R1	CTCTAGGCATGCAATGCTTCATCG
c-fos	fx_e7fkf1_f1	GGAAGCAGCAGCAATGAGC
	fx_e7fkf1_r1	ATAGCCCTGTGATCGGCAC
rgs2	IK-Aburgs2-F1	GTCATTTGCTTACCCCTGTGCTGT
	IK-Aburgs2-R1	CTGGCCAAATCATTCCTGCTTCCT
gif	IK-Abugif-F1	CTTGCGAGTCAAAAGGCTTCTGAG
	IK-Abugif-R1	ACCCCTGTAGCAGCCCTGAAA
alas1	fx_alas1_f1	TTTGTTGCCAATGACTCCAC
	fx_alas1_r1	GCCGCTGTTCCTGATACCT
c1q-like	fx_c1ql_f3	GCAGAAACTGGACCCTTTA
	fx_c1ql_r3	CTGGTGCAGTGAAAATACC

H. ishmaeli

gene	name	sequence
actinR	RT_His_ActinR_f2	GGTGGACAACCTCATCTTTGTTGG
	RT_His_ActinR_r2	TGACGAGTTGCTCAAGAGGGTCTA
twinfilin	RT_His_twinfilin_f1	CGCCCTCATGGGAATGTTTGTAGA
	RT_His_twinfilin_r1	TGAACAGGAGCTCATTCGGTTGTG
runx2b	HG_HisRunx2b-F1/2	ACATCCCTGGTGAACCCTAACCT
	HG_HisRunx2b-R	CACTGATGGCAGGCAGACC
osx	RT_His_osx_f1	GCACAAATATGAGCTACCGAGACA
	RT_His_osx_r1	TTCCGAGCAGACGCACTTTTG
c-fos	fx_e7fkf1_f1	GGAAGCAGCAGCAATGAGC
	fx_e7fkf1_r1	ATAGCCCTGTGATCGGCAC
rgs2	IK-Aburgs2-F1	GTCATTTGCTTACCCCTGTGCTGT
	IK-Aburgs2-R1	CTGGCCAAATCATTCCTGCTTCCT
gif	RT_His_gif_f1	GCTTGCGAGTCAAAAGGCTTCT
	RT_His_gif_r1	GTAGCAGCCCTGAAAAGCTCAA
alas1	fx_alas1_f1	TTTGTTGCCAATGACTCCAC
	fx_alas1_r1	GCCGCTGTTCCTGATACCT

Table S.IV.2 Summary statistics for morphological analysis

	A. alluaudi				*A. burtoni*				*H. ishmaeli*			
	d. f.	**t**	**p-value**	**fdr**	**d. f.**	**t**	**p-value**	**fdr**	**d. f.**	**t**	**p-value**	**fdr**
Jaw area	14.095	-5.600	0.000	0.000	14.078	-3.743	0.002	0.009	14.681	-3.448	0.004	0.029
Centroid size	10.311	-4.404	0.001	0.002	14.006	-2.965	0.010	0.016	13.034	-3.104	0.008	0.033
Jaw depth	16.699	-6.031	0.000	0.000	22.950	-5.855	0.000	0.000	9.298	-2.423	0.038	0.050
Jaw weight	17.000	-5.044	0.000	0.000	13.862	-3.345	0.005	0.013	14.095	-2.776	0.015	0.035
Jaw length	11.072	-3.420	0.006	0.006	19.824	-2.314	0.031	0.036	12.553	-1.268	0.228	0.260
Jaw width	13.383	-3.157	0.007	0.007	16.357	-1.145	0.269	0.269	13.617	-1.139	0.274	0.274
Horn width	12.343	-5.428	0.000	0.000	17.051	-2.621	0.018	0.024	13.755	-2.699	0.017	0.035
Suture width	13.773	-7.132	0.000	0.000	22.956	-2.962	0.007	0.014	13.565	-2.507	0.026	0.041

	P. multicolor				*T. moori*			
	d. f.	**t**	**p-value**	**fdr**	**d. f.**	**t**	**p-value**	**fdr**
Jaw area	8.238	0.316	0.760	0.823	7.741	-0.234	0.821	0.861
Centroid size	8.835	0.573	0.581	0.823	7.630	-1.070	0.317	0.635
Jaw depth	10.429	-0.301	0.769	0.823	7.650	0.414	0.690	0.861
Jaw weight	4.069	-0.238	0.823	0.823	7.951	0.379	0.715	0.861
Jaw length	8.047	0.516	0.620	0.823	7.867	-1.085	0.310	0.635
Jaw width	6.868	1.412	0.202	0.823	7.577	2.149	0.066	0.526
Horn width	3.734	0.260	0.809	0.823	7.953	1.078	0.312	0.635
Suture width	5.986	0.271	0.796	0.823	7.061	-0.182	0.861	0.861

Table S.IV.3 Summary statistics for gene expression

	A. alluaudi				*A. burtoni*				*H. ishmaeli*			
	d. f.	**t**	**p-value**	**fdr**	**d. f.**	**t**	**p-value**	**fdr**	**d. f.**	**t**	**p-value**	**fdr**
runx2b	11.456	-0.610	0.554	0.654	15.400	-0.451	0.658	0.658	9.859	1.275	0.232	0.463
osx	17.136	0.221	0.828	0.828	16.373	1.567	0.136	0.284	14.796	-3.104	0.007	0.031
rgs2	14.374	0.260	0.799	0.828	17.492	-1.074	0.297	0.429	12.251	0.655	0.524	0.699
cfos	14.325	-1.063	0.305	0.496	10.117	-1.487	0.168	0.284	11.356	0.088	0.932	0.946
gif	10.122	-5.059	0.000	0.006	9.490	-3.102	0.012	0.077	14.888	-3.079	0.008	0.031
alas1	10.183	-3.836	0.003	0.021	10.390	-3.246	0.008	0.077	13.390	-1.138	0.275	0.471
col6	11.310	1.212	0.250	0.496	9.878	-0.937	0.371	0.482	11.896	-4.148	0.001	0.017
col12	14.705	-0.804	0.434	0.565	17.411	2.348	0.031	0.134	13.638	-2.281	0.039	0.118
des	15.339	1.149	0.268	0.496	16.555	0.505	0.620	0.658	14.337	0.487	0.634	0.761
ryr1	16.298	1.210	0.243	0.496	14.334	-0.515	0.614	0.658	9.796	0.069	0.946	0.946
srl	14.204	0.953	0.356	0.515	16.359	2.028	0.059	0.154	14.534	0.866	0.400	0.600
tpm4	10.680	1.795	0.101	0.328	16.51	1.418	0.175	0.284	14.54	1.634	0.124	0.297

Chapter V

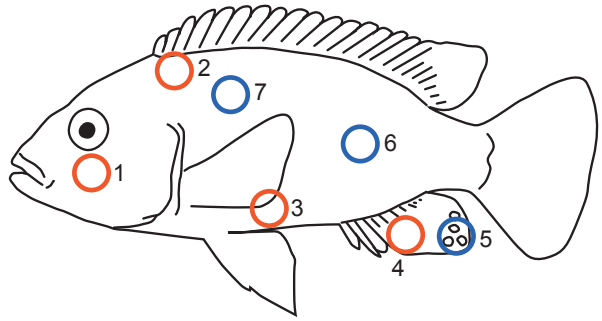


Fig. S.V.1 Measured spots on African cichlid species. Red circles (1-4) indicate spots measured on all individuals. Spot number five was only measured in males, spot number six only in *L. caeruleus* (this was the most intensely looking yellow spot on both sexes) and spot number seven in *M. auratus* and *M. johannii* (this was focused on the position with the male blue stripe).

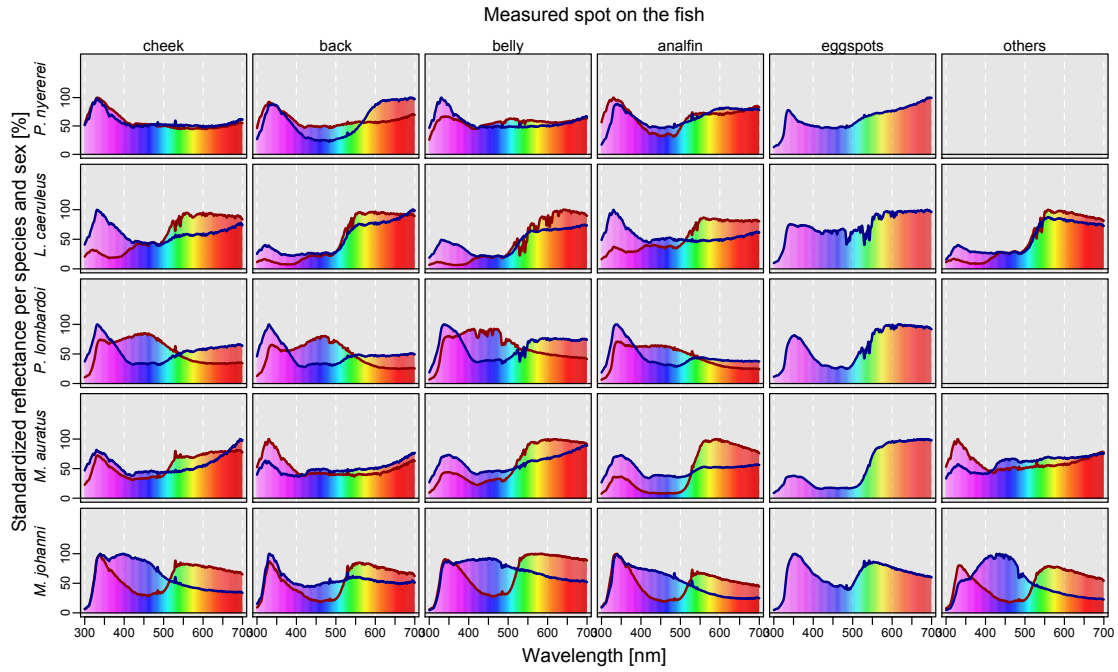


Fig. S.V.2 Average relative reflectance of measured spots on African cichlid bodies. Reflectance was standardized (highest reflectance value was set to 100%) and male and female reflectance was equalized according to the curves integral, where applicable. Male and female spectra are indicated using blue and red lines, respectively.

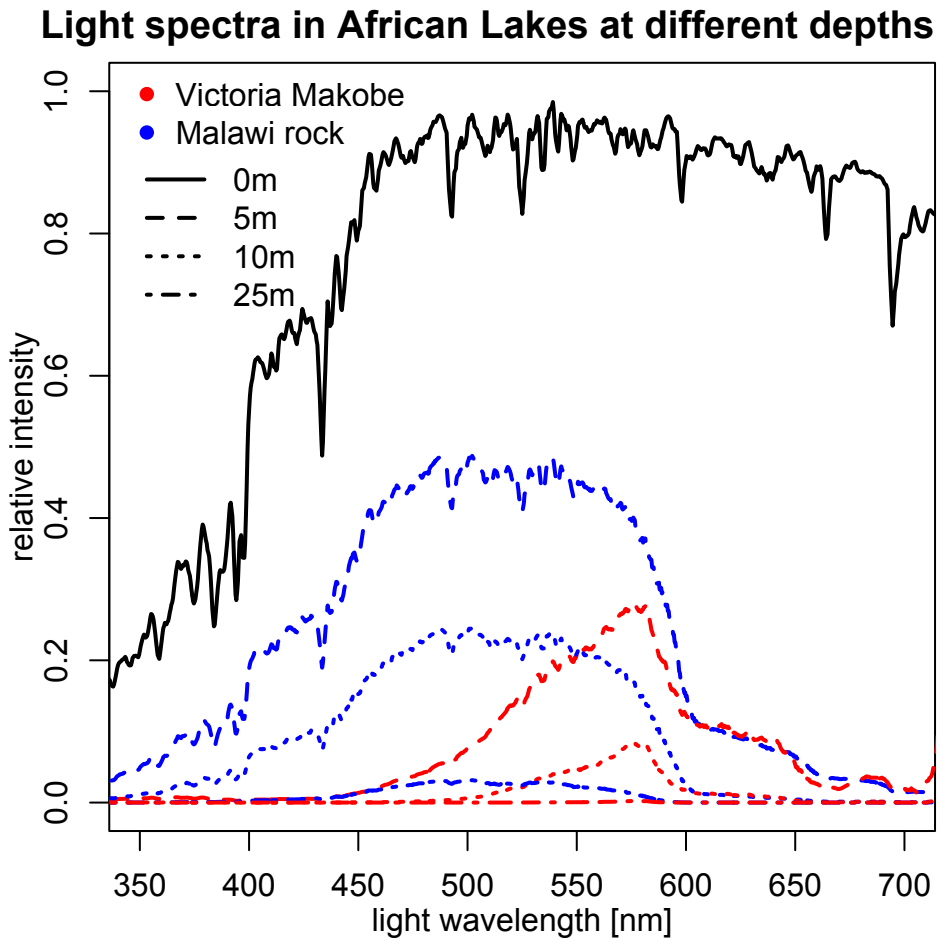


Fig. S.V.3 Light spectra assumed for habitat depths used in this study for Lake Malawi (rock habitat) and Lake Victoria (Makobe island). Maximum sunlight intensity at surface level was set to be 1 (black line) and blue and red lines indicate available light spectra at different depths in Lake Malawi and Victoria, respectively.

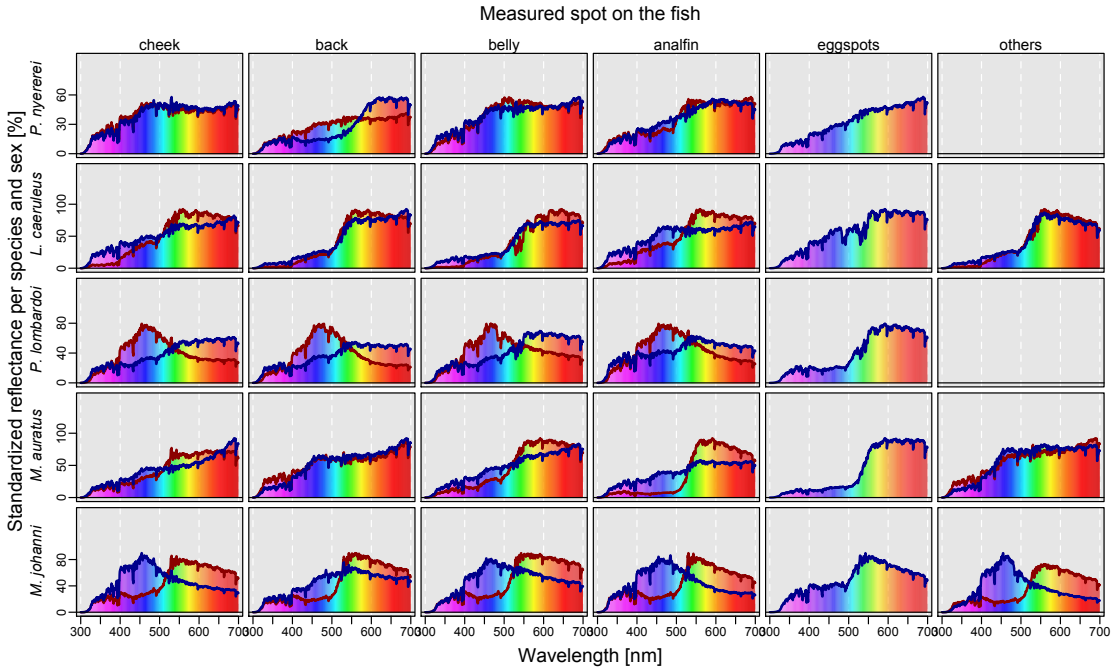


Fig. S.V.4 Assumed reflectance spectra for measured spots on investigated cichlids. Spectra were derived by multiplying the sunlight spectrum with the relative reflectance spectra per spot and species. Male and female spectra are indicated using blue and red lines, respectively. Per plot, both spectra were standardized the height of the spectrum of the sex with the lower integral was adjusted to equalize integrals.

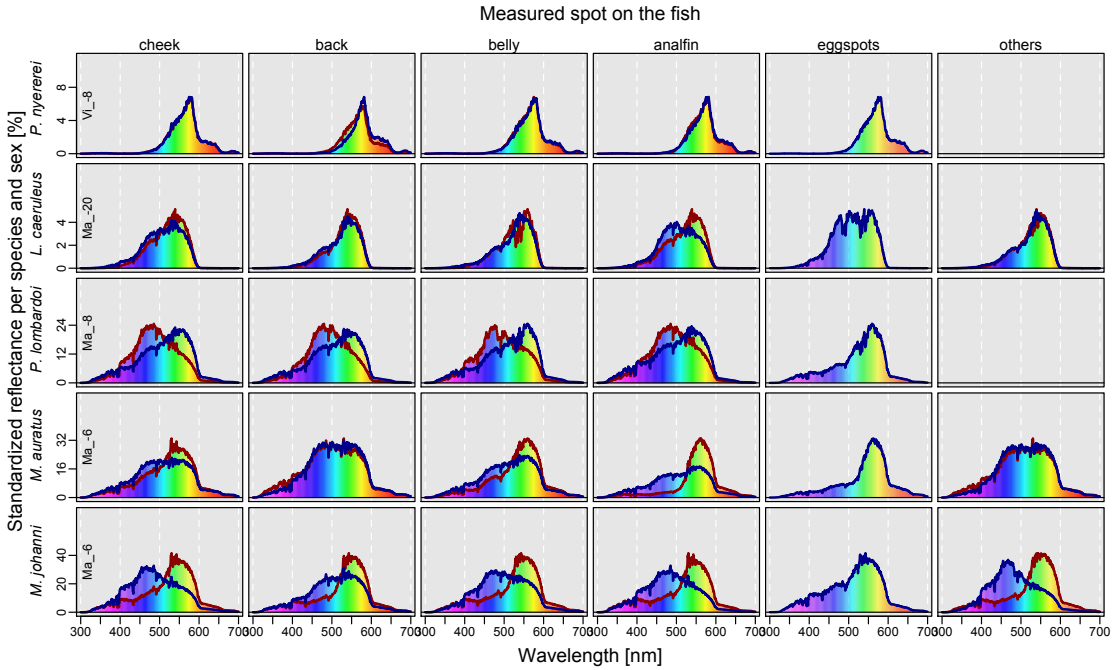


Fig. S.V.5 Assumed reflectance spectra for measured spots on investigated cichlids at habitat depth. Spectra were derived by multiplying the respective habitat light spectrum with the relative reflectance spectra per spot and species. Male and female spectra are indicated using blue and red lines, respectively. Per plot, the height of the spectrum of the sex with the lower integral was adjusted to equalize integrals.

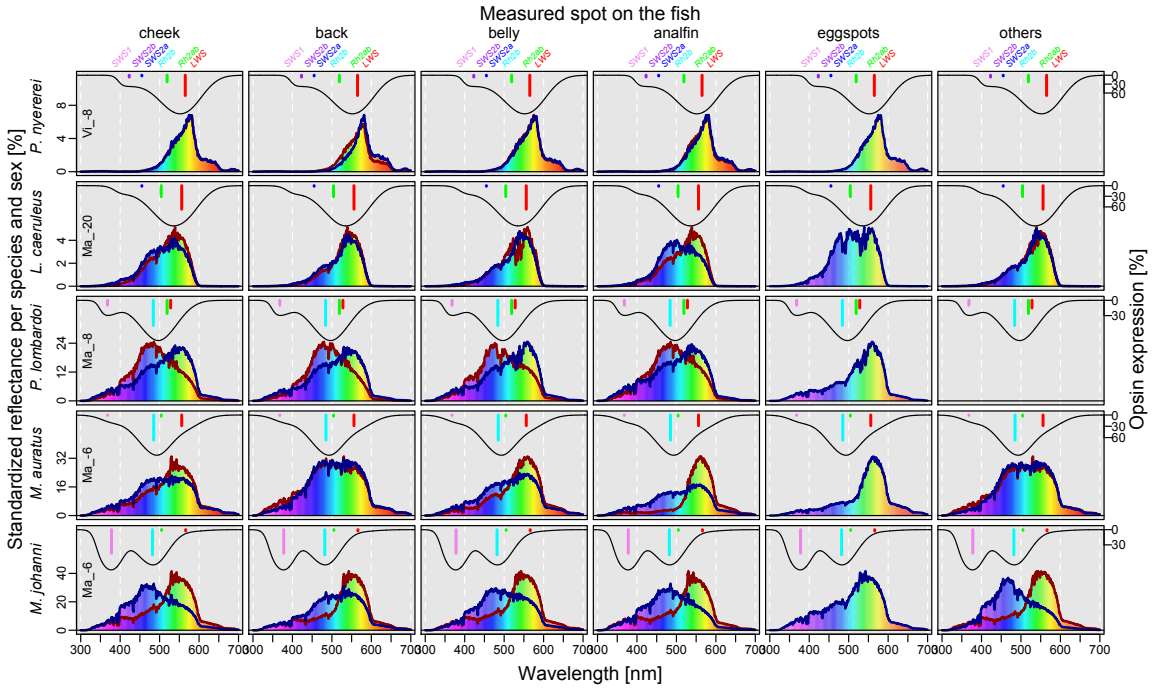


Fig. S.V.6 Assumed reflectance spectra for measured spots on investigated cichlids at habitat depth with modeled retina sensitivities and opsin expression bottom-up on the right axis of each plot. Male and female spectra are indicated using blue and red lines, respectively. Per plot, the height of the spectrum of the sex with the lower integral was adjusted to equalize integrals.

Table S.V.1 Efficiencies and R^2 values for the different opsin primers in the ten species, R^2 are shown in parenthesis

Species	*sws1*	*sws2b*	*sws2a*	*rh2b*	*rh2a*	*lws*	*rho*
H. ne	0.987 (0.98)	1.077 (1)	1.055 (1)	0.986 (0.96)	1.08 (0.98)	1.04.9 (1)	1.049 (1)
H. ni	0.979 (0.98)	0.869 (1)	0.977 (1)	0.932 (0.89)	0.962 (1)	1.006 (1)	1.069 (1)
A. ag	NA	0.983 (0.99)	0.835 (0.99)	1.00 (0.75)	0.978 (1)	0.978 (0.99)	0.921 (1)
A. ba	NA	0.882 (0.99)	0.86 (1)	1.00 (0.97)	0.923 (1)	0.961 (1)	0.929 (0.99)
A. or	NA	1.035 (0.99)	0.908 (1)	1.002 (0.9)	0.995 (1)	0.965 (1)	0.983 (1)
L. ca	0.871 (0.95)	1.004 (1)	0.968 (1)	0.934 (0.99)	0.879 (1)	0.975 (1)	0.974 (1)
P. lo	0.92 (1)	0.894 (0.99)	0.958 (1)	0.948 (1)	0.93 (1)	0.921 (1)	1.01 (1)
M. au	0.844 (1)	1.055 (1)	1.011 (0.98)	1.026 (1)	1.085 (1)	1.028 (1)	1.016 (1)
M. jo	0.762 (0.98)	0.92 (0.99)	0.952 (0.94)	0.934 (1)	1.077 (0.98)	0.955 (1)	1056 (0.99)
P. ny	0.858 (0.96)	0.967 (1)	0.948 (1)	NA	0.963 (1)	0.988 (1)	0.982 (0.99)

Table S.V.2 Primer sequences code-number for all eight opsin genes among all ten investigated species and code-primer-pair-sequence list

	sws1	*sws2b*	*sws2a*	*rh2b*	*rh2a alpha*	*lws*	*rho*
A.agassizii	1	8	3	4	5	12	13
A.barlowi	1	8	3	4	5	12	13
A.ortegai	1	2	3	4	5	6	13
H.nematopus	7	8	9	10	11	12	13
H.nicaraguensis	7	8	9	10	11	12	13
L.caeruleus	7	8	9	10	11	12	13
M.auratus	7	8	9	10	11	12	13
M.johannii	7	8	9	10	11	12	13
P.lombardoi	7	8	9	10	11	12	13
P.nyererei	7	8	9	10	11	12	13

1:	AH_SWS1_F2	GATATACCTGCTCCAGCCAAAG
	AH_SWS1_R2	GGTCATCTGTAAACCATTTGGAG
2:	AH_SWS2B_F1	CACGAACGGAACTGCTTATTG
	AH_SWS2B_R1	GCCATTTGACCTGAGACTGG
3:	JTD_SWS2a_F1.1	GAAACTTCGATCCCACCTCA
	JTD_SWS2a_R4	GGCTTACCATACCACCGAGT
4:	AH_RH2B_F1	GCTTGCAGCAGCCTTCAC
	AH_RH2B_R1	CCTGTGGACCGGATTACTACAC
5:	JTD_RH2Abeta_F3	TGCTGGAGTACTCTTCACATGG
	JTD_RH2A_R3	CTCATTGTTGAAGCCTGGAG
6:	JTD_LWS_F1.1	GGCGGTACCATGAAGATACAAC
	JTD_LWS_R4.1	ACCACGAAAAGCATCCAGAG
7:	JTD_SWS1.AF_F426	TTGGCTCAAATCACGCTCTG
	JTD_SWS1.Af_R519	AGGGATGTACCTGCTCCAG
8:	JTD_SWS2bAf_F370	ACTCTTGGTGGTATGGTTAGCC
	JTD_SWS2b_R456	AAAGTTCCCAAGTGGCTTGC
9:	JTD_SWS2a_F369	AACACTCGGTGGTATGGTAAGC
	JTD_SWS2aAF_R469	TGCAAATCACAAGCCATCGC
10:	JTD_RH2BAf_F543	TGGCTGGTCAAGGTACATTCC
	JTD_RH2BAf_R661	AGAAGTGGCAGGTGAACATG
11:	JTD_RH2AAf_F708	GACAGTCAAAGCTGCAGCAG
	JTD_RH2AAf_R848	ATCCAACCAGCAAAGCTAGC
12:	JTD_LWSAf_F844	ACCTTCTTTGCCTGCTTTGC
	JTD_LWSAf_R992	ATGATGCATGTGCGGAACTG
13:	RS_RH0_F4	AACTTCCGCTTTGGGGAGAA
	RS_RH0_R4	CCTCAGGGATGTAACGAGACC

Table S.V.3 Test-statistics of pair-wise comparisons of opsin expression between male and female. P-values were adjusted for multiple testing within species.

t

	Hne	Hni	Aag	Aba	Aor	Lca	Plo	Mau	Mjo	Pny
SWS1	2.936	1.132	-0.894	0.906	-0.312	-0.56	-0.927	-0.096	-0.925	-0.528
SWS2b	-0.833	1.093	1	-1.166	0.1	-1.855	1.025	-0.647	0.849	-0.69
SWS2a	2.357	-1.548	0.307	3.617	-0.048	-0.536	1.525	0.322	-0.142	0.296
RH2b	0.848	-0.932	0.679	0.91	-0.512	-1.778	-0.639	-0.466	0.556	-0.228
RH2ab	-0.901	-0.062	1.911	-2.487	0.399	1.129	0.636	-0.531	-0.051	-0.853
LWS	-0.309	0.204	-1.152	0.698	-0.158	-0.94	-0.308	0.511	0.635	0.915
RHO	-0.619	-0.426	1.539	-0.774	0.697	-0.83	-0.217	0.583	-0.879	1.063

df

	Hne	Hni	Aag	Aba	Aor	Lca	Plo	Mau	Mjo	Pny
SWS1	3.552	4.819	2.292	2.083	1.006	5.966	4.322	3.78	2.578	5.952
SWS2b	3.13	4.62	2.295	4.074	1.967	5.92	3.17	4.837	4.768	5.622
SWS2a	4.646	4.999	2.067	4.805	1.06	3.98	3.057	6.419	2.36	5.772
RH2b	5.843	1.001	3.893	2.137	1.065	3.554	5.559	5.476	2.975	5.66
RH2ab	3.67	1.498	3.353	2.016	1.19	5.184	5.769	6.436	3.998	5.378
LWS	3.061	1.386	2.962	2.496	1.103	4.949	5.399	5.551	4.1	5.107
RHO	4.974	2.063	2.572	2.18	1.491	3.086	5.895	5.089	2.209	3.113

p

	Hne	Hni	Aag	Aba	Aor	Lca	Plo	Mau	Mjo	Pny
SWS1	0.049	0.311	0.455	0.457	0.807	0.596	0.403	0.928	0.433	0.617
SWS2b	0.464	0.328	0.411	0.307	0.93	0.114	0.377	0.547	0.436	0.518
SWS2a	0.069	0.182	0.787	0.016	0.969	0.621	0.223	0.758	0.898	0.777
RH2b	0.43	0.522	0.535	0.453	0.695	0.159	0.548	0.659	0.617	0.828
RH2ab	0.423	0.958	0.142	0.13	0.749	0.308	0.549	0.613	0.962	0.43
LWS	0.777	0.864	0.334	0.544	0.899	0.391	0.77	0.629	0.559	0.401
RHO	0.563	0.71	0.236	0.514	0.578	0.466	0.836	0.584	0.464	0.363

p.adjusted

	Hne	Hni	Aag	Aba	Aor	Lca	Plo	Mau	Mjo	Pny
SWS1	0.242	0.765	0.624	0.544	0.969	0.621	0.769	0.928	0.864	0.828
SWS2b	0.65	0.765	0.624	0.544	0.969	0.556	0.769	0.884	0.864	0.828
SWS2a	0.242	0.765	0.787	0.112	0.969	0.621	0.769	0.884	0.962	0.828
RH2b	0.65	0.914	0.624	0.544	0.969	0.556	0.769	0.884	0.864	0.828
RH2ab	0.65	0.958	0.624	0.455	0.969	0.621	0.769	0.884	0.962	0.828
LWS	0.777	0.958	0.624	0.544	0.969	0.621	0.836	0.884	0.864	0.828
RHO	0.657	0.958	0.624	0.544	0.969	0.621	0.836	0.884	0.864	0.828

Table S.V.4 Details on the species used for the meta-study

species	sexual dimorphism	functional *sws1*
Aulonocara baenschi	yes	yes
Aulonocara hueseri	yes	yes
Copadichromis borleyi	yes	yes
Haplochromis fischeri	yes	yes
Haplochromis sp. 'macula'	yes	yes
Julidochromis regani	no	yes
Labidochromis caeruleus	no	yes
Labidochromis chisumulae	yes	yes
Labidochromis flavigulus	no	yes
Labidochromis ianthinus	yes	yes
Labidochromis vellicans	yes	yes
Lethrinops parvidens	yes	yes
Maylandia barlowi	yes	yes
Maylandia benetos	yes	yes
Maylandia callainos	yes	yes
Maylandia lombardoi	yes	yes
Maylandia mbenji	yes	yes
Maylandia phaeos	yes	yes
Maylandia pyrsonotos	yes	yes
Melanochromis auratus	yes	yes
Mylochromis lateristriga	yes	yes
Neolamprologus brichardi	no	no
Neolamprologus furcifer	no	yes
Neolamprologus mondabu	no	yes
Neolamprologus mondabu	no	no
Neolamprologus tretocephalus	no	yes
Oreochromis niloticus	no	yes
Perissodus microlepis	no	yes
Stigmatochromis modestus	yes	yes
Tramitichromis intermedius	no	yes
Tyrannochromis maculatus	yes	yes
Xenotilapia bathyphila	no	yes
Xenotilapia boulengeri	no	no
Xenotilapia flavipinnis	no	no
Xenotilapia leptura	no	yes
Xenotilapia ochrogenys	no	no
Xenotilapia spiloptera	no	yes

Table S.V.5 Details on λ_{max} per species and corresponding references. Also Lens T50 values are shown.

	Default*	*P.ny*	*L.ca*	*P.lo*	*M.au*	*M.jo*
SWS1	381	368	381	**378**	376	**377**
SWS2b	427	**427**	427	415	424	415
SWS2a	460	455	460	452	457	452
RH2b	490	490	490	**482**	**485**	**485**
RH2a(α&β)	535	**528**	**535**	**534**	525	525
LWS	566	**565**	**566**	**566**	**561**	556
T50	-	400	387	339	364	336
REF	Sabbah *et al.* (2010): *M. zebra* all, but LWS Parry *et al.* (2005): *P. acei* - LWS	Seehausen *et al.* (2008): all but SWS1, SWS2b & RH2b *Default: SWS1, SWS2b & RH2b	*Default: all	Parry *et al.* (2005): *P. acei*	Sabbah *et al.* (2010): *M. auratus*	Parry *et al.* (2005): *M. vermivorus* - all but SWS2a, SWS1 & RH2a Sabbah *et al.* (2010): *M. auratus* - SWS1, SWS2a & RH2a

Table S.V.6 Divergence between color spectra after standardization in % between sexes (SDi), between males (MC) and the habitat light spectrum, and the females and the habitat light spectrum (FC).

	cheek			back			belly			anal fin			egg spots	other		
	SDi	MC	FC	SDi	MC	FC	SDi	MC	FC	SDi	MC	FC	- NA -	SDi	MC	FC
P.ny	2	2	3	17	19	2	3	2	2	4	6	5	-			
L.ca	10	6	16	7	21	27	10	21	22	17	2	16	-	4	23	24
P.lo	22	10	12	20	9	11	20	13	9	15	8	9	-			
M.au	13	5	15	4	3	4	17	8	23	15	8	41	-	4	4	3
M.jo	31	14	18	21	4	25	27	6	22	32	13	21	-	41	17	25

Table S.V.7 Final sample sizes per opsin gene expression, lens cutoff and reflectance measurement.

	sex	*sws1*	*sws2b*	*sws2a*	*rh2b*	*rh2a*	*lws*	LC	reflect
H. ne	M	4	4	4	4	4	4	-	-
	F	4	4	4	4	4	4		-
H. ni	M	5	5	5	5	5	5	-	-
	F	2	2	2	2	2	2		-
A. ag	M	3	3	3	3	3	3	-	-
	F	3	3	3	3	3	3		-
A. ba	M	3	3	3	3	3	3	-	-
	F	4	4	4	4	4	4		-
A. or	M	3	3	3	3	3	3	-	-
	F	2	2	2	2	2	2		-
P. ny	M	4	4	4	4	4	4	1	4
	F	4	4	4	4	4	4		3
L. ca	M	4	4	4	4	4	4	1	4
	F	4	4	4	4	4	4		5
P. lo	M	4	4	4	4	4	4	2	4
	F	4	4	4	4	4	4		3
M. au	M	6	6	6	6	6	6	1	4
	F	4	4	4	4	4	4		4
M. jo	M	4	4	4	4	4	4	6	1
	F	3	3	3	3	3	3		1