INVITED REVIEWS AND META-ANALYSES

Speciation in fishes

GIACOMO BERNARDI
Department of Ecology and Evolutionary Biology, University of California Santa Cruz, 100 Shaffer Road, Santa Cruz, CA 95076, USA

Abstract

The field of speciation has seen much renewed interest in the past few years, with theoretical and empirical advances that have moved it from a descriptive field to a predictive and testable one. The goal of this review is to provide a general background on research on speciation as it pertains to fishes. Three major components to the question are first discussed: the spatial, ecological and sexual factors that influence speciation mechanisms. We then move to the latest developments in the field of speciation genomics. Affordable and rapidly available, massively parallel sequencing data allow speciation studies to converge into a single comprehensive line of investigation, where the focus has shifted to the search for speciation genes and genomic islands of speciation. We argue that fish present a very diverse array of scenarios, making them an ideal model to study speciation processes.

Keywords: ecological genetics, evolutionary theory, fish, speciation

Received 14 May 2013; revision received 8 August 2013; accepted 14 August 2013

In the past two decades the field of speciation has seen a renewed interest with major theoretical and empirical advances (Rice & Hostert 1993; Howard & Berlocher 1998), accompanied by the publication of seminal books that placed the field within a strong theoretical framework (Coyne & Orr 2004; Dieckmann et al. 2004; Gravill et al. 2004; Price 2007; Nosil 2012). Molecular approaches applied to aquatic organisms have traditionally lagged behind their terrestrial counterparts, and this also applies to the field of speciation. In a few cases, however, fishes have been at the forefront of speciation studies. Examples of parallel divergence of sympatric sticklebacks and runaway antagonistic sexual selection in cichlids, which will be discussed later, come to mind (Rundle & Schluter 2004; Farnell & Streelman 2013). The goal of this study is to review the state of knowledge of speciation in fishes and synthesize its idiosyncrasies and breakthroughs.

As there are numerous definitions, a treatment on species concepts goes beyond the scope of this review (De Queiroz 1998; Richards 2010; Hausdorf 2011; Carstens et al. 2013). Early species concepts mostly influenced by Mayr’s biological species concept (BSC) assumed some level of reproductive isolation (Mayr 1942). This concept, however, is clearly insufficient even for some simple cases, such as asexual and unisexual species, but also cannot accommodate for the fundamental and growing evidence of speciation in the absence of breeding barriers (Coyne & Orr 2004). Hausdorf (2013) argues that defining species as units that reach fitness maxima, and where any exchange would lower their fitness, encompass most situations and may tend towards a universal species concept. Importantly, many studies that will be mentioned in this review are more concerned by the speciation mechanisms rather than the end product of the species. So for all practical purposes, we will be using Hausdorf’s species concept (Hausdorf 2011) as it will suffice for our discussion.

Understanding how divergence emerges, which is the study of speciation, is a question that has been approached in fishes using three different angles. One approach considers the spatial and geographical component of speciation, another tack has focused on the ecological component of speciation, and a third has mostly considered the sexual part of the equation. These three approaches do intersect, at times rather significantly, but for clarity’s sake, we will distinguish them here. All three aspects of speciation are rooted in traditional studies and are entering the realm of genomics, making
a review of the material timely. Thus, this study, which
is constrained by its size, will limit itself by first provid-
ing a general background on fish, their taxonomy and
genomics, then will describe the history of the three
approaches to the study of speciation and their latest
developments, and will conclude by trying to bring
them together for a comprehensive view of the ques-
tion.

What are fishes?—general background
Fish are a paraphyletic group of vertebrates that have
evolved for the past 500 My and comprise approxi-
mately 30 000 species (Nelson 2006). Similarly to other
groups of organisms, fish species diversification is not
particularly related to the age of the group (Rabosky
et al. 2012). Fish include the agnathans (jawless fishes:
hagfish and lampreys, with about 70 species) and the
gnathostomes, which are jawed vertebrates that
comprise the chondrichthians (cartilaginous fishes: chi-
maeras, sharks and rays, with about 1000 species),
the actinopterygians (ray-finned fishes, with more than
25 000 species) and the sarcopterygians (lobe-finned
 fishes: lungfishes and coelacanths, 11 species) that are
closely related to tetrapods (Nelson 2006). While 2.8%
of the earth’s water is fresh and 2.7909% is not usable
by fish, approximately 35% of fish species are found in
freshwater. Therefore, freshwater habitats are about
10 000 times more speciose than marine ones, although
this may be more a reflection of an ancient marine
extinction rather than an actual mechanistic cause (Vega
& Wiens 2012). The centre of marine fish diversity is
the coral triangle, an area bound by the Philippines,
Indonesia and Papua New Guinea, where the highest
diversity of marine reef fishes is found, with approxi-
mately 2500–4000 species, mostly belonging to the order
Perciformes (Carpenter & Springer 2005; Allen 2008;
Allen & Erdmann 2012). The centre of freshwater fish
diversity is the Amazon Basin, with approximately 2600
species, mostly belonging to the superorder Ostario-
physyi (catfish, piranhas; Albert & Reis 2011; Chen et al.
2013). Most fishes are oviparous, and females produce
eggs, which are either released in the water (broadcast
spawners) or laid on a surface (benthic spawners; Helf-
man et al. 2009). In marine fishes, eggs (for broadcast
spawners) or larvae (for benthic spawners) move with
currents and disperse during a pelagic phase, which
typically lasts from a few days to several weeks (pelagic
larval duration or PLD; Leis 1991), as opposed to fresh-
water systems where most species do not have pelagic
eggs (Kunz-Ramsey 2008). Parental care, generally in
the form of male guarding, is common in fishes ( Gross
&Sargent 1985). A few species of marine fishes lack a
pelagic larval stage (apelagic fishes), by either brooding
their offspring (surperches, Embiotocidae), mouthbroo-
ding (Banggai cardinalfish, Pterapogon kauderni) or
guarding their babies (some damselfishes, Acanthochro-
mos polyscathus, Alchrithys azurelinatus, A. curatus;
Bernardi 2011). The genetic effects of such different life
histories also have a strong influence on the population
structure of these species (Doherty et al. 1995). Fishes
mature at very small sizes (<1 cm, Paecycypris progenti-
ca, Trimmatom nanus and Schindleria brevipinguis) or very
large ones (whale shark, Rhincodon typus, about 10 m),
and their lifespan varies between <1 year (annual killif-
ishes, e.g. Cynelebias, Notobranchius) to more than
150 years (e.g. rockfishes of the genus Sebastes). Taken
together, these characteristics result in variations in
mutation rates and are to be considered when estimat-
ing divergence and coalescence times.

Fish genomics
Vertebrate genomes have undergone an early duplica-
tion (1R) that was followed by a secondary duplica-
tion (2R; Ohno 1970). In ray-finned fishes, a third round of
duplication, termed ‘fish-specific genome duplication’
(FSGD, 3R), occurred (Ohno 1970; Taylor et al. 2003;
Meyer & Van de Peer 2005; Steinke et al. 2006; Kuraku &
Meyer 2010). Further duplication events may have
happened in additional lineages, most notably in salmo-
nids where a recent (50–100 Mya) duplication is well
documented (Allendorf & Thorgaard 1984; Koop &
Davidson 2008). In general, fish DNA base composition
is relatively homogeneous along the genome, resulting
in symmetrical peaks in denaturation curves and caes-
ium chloride (CsCl) ultracentrifugation gradients (Ber-
ardi & Bernardi 1990; Bucciarelli et al. 2002), which
also translates in weak banding patterns in fish chromo-
somes (Medrano et al. 1988). The size range of fish
genomes, a value known for more than 40 years (Hine-
gardner & Rosen 1972), varies widely from the smallest
size in some puffers (390 Mb, about one-eighth of the
human genome) to more than 300 times that size in
Protopterus aethiopicus (marbled lungfish, 130 000 Mb,
about 40 times the human genome; Metcalfe et al. 2012).
The difference in size is reflected by a large difference in
repeated sequences and transposable elements (Pizon
et al. 1984), which in turn prompted the early sequenc-
ing of the genome of puffers Takifugu rubripes and Tetra-
odon fluviatilis (Brenner et al. 1993; Jaillon et al. 2004).
A number of additional fish genomes were sequenced in
succession, including zebrafish, medaka, stickleback
and tilapia (Colosimo et al. 2005; Kasahara et al. 2007;
are currently being sequenced, with an additional
several hundreds being targeted in the near future (Ber-
nardi et al. 2012a,b). Besides the sequencing of entire

© 2013 John Wiley & Sons Ltd
SPECIATION IN FISHES

Spatial component

In its early days, the field of speciation was greatly influenced by the search for allopatric situations that would produce the genetic divergence necessary for reproductive isolation (Mayr 1942). Indeed, the description of trans-isthmian geminate fish species in Panama was a key step towards understanding allopatric speciation processes (Jordan 1908). To this day, geminate species are used to calibrate molecular clocks (Bermingham et al. 1997; Donaldson & Wilson 1999; Marko 2002; Domingues et al. 2005; Lessios 2008). Biogeographical data, where species distributions yield some insight into the speciation process, have been used since the allozyme era (Stepien et al. 1994), soon followed by studies based on mtDNA markers (Avise 1992; Grant & Bowen 1998; Bowen et al. 2001). Yet, biogeographical data based on species-level approaches can only give a very coarse idea of the speciation process. These approaches were, in turn, replaced by within-species-level studies, the realm of phylogeography (Avise 2000). An entire field of investigation started and is still active, where an explicit geographical component is combined with genetic data to identify the potential for abiotic factors to play a role in various biotic processes, including speciation (Rocha & Bowen 2008; Puritz et al. 2012; Andrew et al. 2013). For marine fishes, biogeographical boundaries were seen as the potential nexus of population genetic discontinuities (Dawson 2001; von der Heyden et al. 2011), be it the large expanse of open water between the western and eastern Pacific Ocean (Lessios & Robertson 2009), the somewhat permeable southern tip of the Baja California Peninsula (Present 1987; Terry et al. 2000; Huang & Bernardi 2001; Bernardi et al. 2003; Schinske et al. 2010), the stretched Hawaiian Archipelago (Ramón et al. 2008; Craig et al. 2010; Eble et al. 2011), the Mona Passage in the Caribbean Sea (Shulman & Bermingham 1995; Taylor & Hellberg 2006) or the complex Indo-West Pacific (Messmer et al. 2005; Drew & Barber 2012; Gaither & Rocha 2013). Physical boundaries that limit gene flow are mostly absent in marine systems, and an interplay between dispersal capabilities, via long PLDs, and oceanic currents was suggested as the main cause for incipient speciation and the subject of further research (Weersing & Toonen 2009; White et al. 2010; Faunby & Barber 2012). Several studies specifically investigated the relationship between PLD and gene flow with mixed results (Waples 1987; Doherty et al. 1995; Shulman & Bermingham 1995; Riginos & Victor 2001). However, the lack of a simple correlation between PLD and gene flow was attributed, at least in part, to theoretical and technical limitations (Weersing & Toonen 2009; Faunby & Barber 2012).

The next step was to better assess and understand the role of larval movement in marine fishes. The unexpected discovery of high levels of self-recruitment in coral reef fishes in the late 1990s (Jones et al. 1999; Swearer et al. 1999) prompted the use of microsatellites and paternity analyses to identify the path of individual fish larvae (Jones et al. 2005; Almany et al. 2007; Saenz-Agudelo et al. 2009; Bernardi et al. 2012a,b; Berumen et al. 2012; Harrison et al. 2012). These studies showed, for the first time, that coral reef fish larvae may not travel far from their natal grounds, resulting in great potential for reduced gene flow, philopatry and local adaptation. These results ultimately reconciled the apparent paradox between the lack of physical barrier in the ocean, theoretically preventing allopatric speciation, and high species diversity in coral reefs. Reduced gene flow and assortative mating, evoked to explain the population structure of *Acanthochromis polyacanthus*, an apelagic damselfish, were also consistent with these results (Planes & Doherty 1997; Planes et al. 2001; Van Herwerden & Doherty 2006).

In general, freshwater bodies tend to be relatively small, thus providing limited possibility for spatial differentiation, diversity of ecological niches and ultimately speciation. Indeed, a quest to find allopatric isolation within a body of water was not as developed in freshwater systems for two main reasons (Puebla 2009). First, the larvae of freshwater fishes tend, in general, not to be transported over long distances by currents, because most freshwater species do not have dispersive larvae. Second, it is rare to find semi-permeable boundaries in freshwater systems. Nevertheless, when possible and relevant, this question has been investigated. For example, glacial cycles in North America created pockets of allopatry that produced large numbers of freshwater fish species and were characterized as ‘speciation pumps’ (April et al. 2013a,b). Another classic example is the rise and fall over geological times of the water level in Lake Tanganyika that resulted in almost complete physical separation of smaller bodies of water within the lake, eventually leading to allopatric genetic divergence for populations of cichlids trapped in those isolated pools. These were later rejoined when water levels rose again, the secondary contact completing the speciation process (Sturmbauer & Meyer 1992). This was shown in both rock-dwelling and open-water cichlids, including the spectacular case of the rock-dwelling *Tropheus* using amplified fragment...
length polymorphism (AFLP) data, where allopatric situations due to water level fluctuations resulted in more than 100 populations and sister species (Egger et al. 2007; Figs 1 and 2). Similarly, allopatric situations emerged due to the high flow of large rivers, such as the Amazon or the Congo, which are essentially impassable for most fishes, and are just as effective a barrier to gene flow as a hard physical boundary (Albert & Crampton 2010; Lundberg et al. 2010; Markert et al. 2010). Indeed, the Amazon outflow on the coast of Brazil is so powerful that it also greatly influences speciation in marine species (Rocha 2003). For freshwater fishes, where river drainages provide physical barriers, the focus shifted to the historical component of freshwater drainages that over time shifted their locations or resulted in captures (Burridge et al. 2006). In Africa, for example, plate tilting was shown to play an essential role in river captures resulting in fish speciation (Giddelo 2002; Hermann et al. 2011; Koblmüller et al. 2012; Musilová et al. 2013). In North America, the freshwater darters (genus *Percina*) are a case in point, where allopatric speciation resulted in more than 50 species among river systems found on the eastern continental divide of the United States (Near et al. 2011).

While geographical isolation, in itself, does eventually result in speciation, the mechanisms of reproductive isolation do not follow a single mode. The simplest one, at the mechanistic level, is drift, where in due time markers become fixed by chance alone (n or 4n generations in populations of n individuals, for mitochondrial and nuclear markers, respectively; Avise 2004). In this case, for accurate estimation of drift, markers must be neutral. Yet, more complex and realistic scenarios involve situations where assortative mating, reduced gene flow and local adaptation, which are in direct violation of Hardy–Weinberg equilibrium requirements, all play an important role in creating reproductive barriers (Puebla et al. 2012). For example, in livebearing fishes of the genus *Poecilia*, independent invasion of extreme habitats (sulphide springs) repeatedly resulted in local adaptation leading to speciation (Tobler et al. 2011; Kelley et al. 2012; Plath et al. 2013). In general, these situations are much more complex to assess genetically, and the power of few mitochondrial, or nuclear loci, is usually insufficient (thus, the prevalence of neutral studies in the literature). The use of thousands of single nucleotide polymorphisms (SNPs), once difficult to obtain, has recently become reachable with the advent of massively parallel sequencing (MPS) and has efficiently been used in both marine and freshwater systems (Hohenlohe et al. 2010a; Seeb et al. 2011; Puritz et al. 2012). In stick-

---

**Fig. 1** Map of Lake Tanganyika showing colour morphs diversity of cichlids in the genus *Tropheus*. Deep-water basins of the lake are represented with grey shading. The dots marking sampling sites were colour-coded to show the distribution of the four major AFLP clades. Note the genetic divergence as it relates to allopatric distribution at low-water levels. Redrawn from Egger et al. (2007).

**Fig. 2** Distribution of the yellow (circles), blue (diamonds) and white (triangles) forms of the reef goby *Elacatinus evelynae* across the Bahamas and Caribbean Sea. Green squares indicate localities where both blue and yellow forms have been reported. The 17 localities that were sampled by Taylor & Hellberg (2003) are indicated with orange lines. Redrawn from Taylor & Hellberg (2003).
lebacks, isolation of populations during the last glacio-

evidences of local adaptation have both been

shown using RAD sequencing (Hohenlohe et al. 2010a, 
b).

Ecological component

Early work on the ecological component of speciation 
in fishes focused on feeding competition, resource parti-
tioning and character displacement (Fryer & Iles 1972). 
The ecological guild of marine surfperches (Embiotocidae) in California, for example, was described as a 
small flock with partitioned resources (Ebeling & Laur 
1986). In some cases, such as the separation of the clo-
sely related black surfperch, Embiotoca jacksoni and 
striped surfperch, E. lateralis, morphological differences 
in the feeding apparatus (the musculature of the pha-
ryngeal jaws) allowed winnowing (sorting food within 
the mouth; E. jacksoni), or not (E. lateralis), resulting in 
competition disequilibrium (Drucker & Jensen 1991). It 
is within this context that studies on feeding specialization 
and fish species flocks were set. Species flocks have 
been studied in few marine systems, such as the rock-
fishes in California (genus Sebastes), the notothens in the 
Antarctic, the South African clinids and the New Zea-
land triplefins (Johns & Avise 1998; Alesandrini & Ber-
2011; Janko et al. 2011). In contrast, literature on fresh-
water systems abounds. Extinct species flocks of sem-
ionotid fishes (related to the extantgars) have been 
described from ancient Mesozoic freshwater rift lakes 
within continental North America (McCune 2004), 
where both spatial and temporal fossil series are avail-
able (McCune et al. 1984), so the phenomenon of adap-
tive speciation within a lake is not new. Yet the study 
of current situations helped clarify some key questions 
as to the mechanisms of speciation within a confined 
environment. This started with the seminal study by 
Fryer and Iles on the cichlid species flocks of the Great 
Lakes of Africa (Fryer & Iles 1972), followed by the dis-
covery of other, more reduced, species flocks such as 
the Cyprinodon pupfishes in the Yucatan Peninsula 
(Humphries & Miller 1981), the barbs of lake Tana, 
Ethiopia (Nagelkerke et al. 1994), the sculpins of Lake 
Baikal (Kontula et al. 2003) and the sailfin silversides of 
the Mälli Lakes, Sulawesi (Herder et al. 2006). Niche 
partitioning and feeding mechanisms were a major 
focus, and indeed feeding specialization followed by 
reproductive isolation is central to ecological speciation 
in fish species flocks (Barlow 2000; Nosil 2012) and not 
unlike what is observed in Darwin’s finches (Grant & 
Grant 2006; Rand et al. 2013). Cichlid flocks that 
emerged independently in the Great African Lakes 
(Meyer et al. 1990; Verheyen et al. 2003; Sturmbauer 
et al. 2011; Loh et al. 2013) belong to the labroids, a 
group of fish with highly developed and specialized 
pharyngeal jaws (Wainwright et al. 2012). With more 
than 500 species found within a single lake (Lowe-
McConnell 1994), early morphological work on feeding 
specializations begat the search for developmental 
genes responsible for the origin of morphological differ-
ences in pharyngeal jaws, a proxy to understand the 
underpinnings of ecological speciation via feeding spe-
cialization in this group (Albertson et al. 2003). Yet, the 
mechanisms of speciation within a lake are the result of 
several interconnected factors, divergent selection on 
feeding types being just one of them. Indeed, besides 
the well-known cichlids, other, nonlabroid species, such 
as Mastacembelids eels and Synodontis catfish also show 
adaptive radiation in those lakes, albeit in lesser num-
bers (Day & Wilkinson 2006; Brown et al. 2010; Wright 
2011). To tease out the various factors involved, much 
simpler systems were sought.

The ecological component may be evaluated in its 
simplest form, when only two ecotypes are found. In 
that situation, ecological speciation occurs as a result of 
the adaptation to two different environments, later fol-
lowed by secondary contact and reproductive isolation 
(with reinforcement if hybrids exhibit lowered fitness; 
Rundle & Schluter 2004). A perfect case scenario was 
found in three-spined sticklebacks (Gasterosteus aculea-
tus), a species complex that is found both in freshwater 
and in coastal marine environments of the higher lati-
tudes in the northern hemisphere. In postglacial fresh-
water lakes in British Columbia, Canada, two sympatric 
forms are present, a limnetic form that specialized on 
plankton feeding and a benthic form that feeds primar-
ily on benthic invertebrates, the primary source of their 
ecological divergence (Schluter & McPhail 1992; Schluter 
1993). They consequently show differences in their 
shape and in particular in their gill rakers, which are 
comb-like structures used to trap food and most devel-
oped in planktonic feeding forms. This situation arose 
indeed independently several times, suggesting similar evolu-
tionary constraints and thus becoming a very clear and 
simple model for parallel evolution, adaptive radiation 
and ecological speciation (Nagel & Schluter 1998; Sch-
luter 2000; McKinnon & Rundle 2002; but see Ishikawa 
et al. 2013). With the sequencing of more than 20 stick-
leback genomes and population-level genome scans 
(RAD sequencing), the genomic regions under selection 
were recently identified (Hohenlohe, et al. 2010a; Jones 
et al. 2012). Importantly, the limnetic–benthic model 
goes beyond sticklebacks. It was found in other sys-
tems, mostly in North American glacial regions (Taylor 
1999), such as the lake whitefish (Coregonus clupeaformis; 
Gagnaire et al. 2013), but also in the Neotropical Midas 
cichlid species complex (Muschick et al. 2011).
The stickleback system fits the ecological speciation model, yet it does not require sympatry, at least not in its first stage, because in this model, adaptation to different environments may occur in allopatry. Indeed, it has been suggested that the limnetic form may be genetically more closely related to marine populations than to the benthic form, thus indicating that sympatric speciation within each lake may not be the most parsimonious explanation. Thus, the possibility of sympatric speciation was tantalizing but still contentious.

Where situations are more complex than the simple limnetic-benthic dichotomy, more than two species are observed (Østbye et al. 2006), and ultimately when the size of the body of water and the diversity of potential ecological niches are great, then species flocks may emerge, as is the case for Great African Lake cichlids (Sturmbauer et al. 2011). At the macroscopic scale, where phylogenies may be seen as evidence of speciation past (Barrachough & Nee 2001), early work suggested that the cichlid species flocks found in the Great African Lakes had evolved within each lake, thus sympatric speciation in some groups was a likely explanation (Meyer et al. 1990; Verheyen et al. 2003; but see Loh et al. 2013). In simpler situations, cichlids offered better situations to test for sympatric speciation, namely in Cameroon’s and Nicaragua’s crater lakes (Schliewen et al. 1994, 2001; Barluenga et al. 2006a). In the latter case, a combined morphological, ecological, phylogeographical and genetic approach provided convincing evidence that sympatric speciation had occurred in the Midas cichlid (Amphilophus sp.) species complex in Lake Apoyo, Nicaragua (Barluenga et al. 2006a), a claim that resulted in some debate (Barluenga et al. 2006b; Schliewen et al. 2006). In marine systems, evidence for sympatric speciation is much more difficult to assess. In the Neotropical reef gobies in the genus Elacatinus, where about 30 species are found around the Isthmus of Panama, no sister relationships were found across the Isthmus (geminante species); instead, all sister relationships were found within an ocean basin (Taylor & Hellberg 2005). Genetic differences instead matched ecological and coloration patterns and were suggestive of sympatric speciation (Taylor & Hellberg 2003, 2005, 2006).

Coyne & Orr (2004) defined four criteria needing to be fulfilled for sympatric speciation to be inferred: a sympatric distribution, a monophyletic sister species relationship, an ecological setting where allopatric speciation is unlikely, such as a crater lake of recent origin as described above, and complete reproductive isolation. The latter condition, however, needs to be qualified because more recent speciation concepts specifically do not require reproductive isolation (Hausdorf 2011). For greenling from Japan (genus Hexagrammos), where these conditions are met (except for reproductive isolation), a combination of experiments and fitness levels of artificial crosses were tested, thus providing a solid case for sympatric speciation in marine fishes (Crow et al. 2010).

**Sexual component**

Studies that focus on the sexual component of speciation have steadily been increasing in the past few years (Coyne & Orr 2004). In fishes, male courtship translates primarily in behavioural displays and coloration signals. Where water is naturally turbid, male display may take unusual forms. One spectacular example is found in West Africa’s Mormyrid elephant fishes that use weak electric fields generated by muscular contractions to assess their environment and feed. There, an entire species flock is the result of sexual selection on male electric displays (Sullivan et al. 2002; Arnegard et al. 2010). In clear water, the most prominent aspect of male display involves reflection and coloration (Maan et al. 2010). Indeed, when clear water becomes turbid due to unnatural alterations, as was the case in Lake Victoria following the introduction of the Nile perch (which required smoking as opposed to the traditional drying of smaller cichlid fishes, thus leading to deforestation and the subsequent soil runoff), break down of assortative mating resulted in rampant hybridization and its associated loss of diversity in local cichlids (Goldschmidt 1996; Seehausen 1997).

Divergent selection on sensory systems may cause speciation through sensory drive. This is observed when female sensory bias plays a role in initiating the process of a given display as was shown in swordtails (genus Xiphophorus), where females select large colourful swords (Meyer 1997; Rosenthal & Evans 1998). For mouthbrooding haplochromine cichlids, high speciation rates were presumed to be driven by sensory bias where females are attracted to egg spots (egg-shaped markings in the males’ anal fin), but experimental work determined that this was unlikely (Henning & Meyer 2012). In contrast, female sensory bias for particular male nuptial colours was theoretically predicted, modelled and observed in Lake Victoria (Maan et al. 2006; Kawata et al. 2007; Seehausen et al. 2008). A correlation was found between the colour of Pundamilia cichlids (red or blue) and their visual pigments (opsin genes; Carleton et al. 2005). There, opsins are first adapted to different shallow- and deep-water light environments with their respective H and P alleles. Allele H is most associated with red male nuptial display, while the P allele is mostly associated with the blue phenotype. Males specifically display in wavelengths that are best seen by females, resulting in assortative mating pairs, and eventually speciation is the outcome of sensory
drive. The predicted replicated blue/red species pairs were observed several times in the *Pundamilia* species complex (Seehausen et al. 2008; Terai & Okada 2011; Fig. 3). Expectedly, mate choice based on male coloration is a way to rapidly create reproductive isolation (Wagner et al. 2012). The resulting assortative mating has been widely observed directly or indirectly (using genetic signatures), in both marine and freshwater fishes (Knight & Turner 2004; Puebla et al. 2007, 2012; Blais et al. 2009; Leray et al. 2010; Smadja & Butlin 2011). For example, at least nine colour morphs of the Caribbean hamlets (genus *Hypoplectrus*) have been observed mating assortatively, as suggested prior to incipient speciation (Fischer 1980; Domeier 1994; Puebla et al. 2007, 2012), yet nearly no genetic divergence has yet been found between *Hypoplectrus* morphospecies when using a number of different molecular assays (McCartney et al. 2003; Ramon et al. 2003; Puebla et al. 2007, 2012; Barreto & McCartney 2008).

To fully appreciate the visual environment of fishes, work has been done either at the ecological level, as for the classic case of guppies (Endler 1980; Kodric-Brown 1985), looking at visual pigments (Cummings & Partridge 2001), or by sequencing opsin genes (visual pigments) to indirectly elucidate visual potential driven by sexual selection, as mentioned above and in other cases as well (Carleton et al. 2005; Terai et al. 2006; Seehausen et al. 2008; Miyagi et al. 2012). The more difficult task of identifying the genes responsible for display coloration was achieved by looking at differential expression in transcriptomes of divergent colour morphs of the Midas cichlid (Henning et al. 2013).

An intriguing case of potential runaway selection was uncovered in species of Lake Malawi mbuna (rock-dwelling) cichlids. Ecological and molecular mapping work was done on a colour morph, orange blotch (OB), which is predominantly found in females in both Lake Malawi and Lake Victoria (Streelman et al. 2003; Ser et al. 2010). Recently, the situation became more complex because the interest in coloration and mating shifted to the issue of sexual conflicts. In this context, speciation may occur through the interplay of assortative mating, sexual conflict and colour patterns (Kocher 2004; Parnell & Streelman 2013). What was unexpected was that the OB colour was found to be linked with the ZW female sex-determining locus (found on chromosome 5), and the blue nuptial colour pattern was linked with two XY male sex-determining loci found on two different chromosomes (Parnell & Streelman 2013). Thus, this system is a perfect setup for a role of ‘polygenic sex-determining systems in rapid evolutionary diversification’ (Parnell & Streelman 2013). This system may explain both the high diversification of Great African Lake cichlid flocks and its unusual fast rate.

As discussed above, prezygotic isolation via assortative mating is well documented in fish. The fertilization stage has also been an important topic of investigation as it relates to potential for reproductive barriers. The study of the fast-evolving gamete recognition proteins is key to understanding speciation in marine invertebrates (Vacquier 1998; Palumbi 2009); yet, little is known about their counterparts in vertebrates in general and in fishes in particular. Zona pellucida, ZP,
glycoproteins, which are represented by at least seven genes in fish, have been sequenced and are good candidates for the egg portion of the equation, but no experimental work has been done to relate genotype, gamete incompatibilities and speciation in fishes (Sun et al. 2010). Instead, an indirect measure of sexual conflict and sperm competition has been gathered in a very large number of studies through the evaluation of levels of multiple pregnancy or multiple maternity (in reversed gender role pipefish and seahorses) in fishes (Jones & Avise 1994; Reisser et al. 2009; Coleman & Jones 2011; Liu & Avise 2011). High levels of multiple paternity (intraspecific) and hybridization (interspecific) clearly show that prezygotic isolation at the fertilization stage is not the primary means of reproductive isolation in fishes. In fact, in some cases, hybridization has generated diversity resulting in speciation (Dwelling & DeMerais 1993; Keller et al. 2012). For example, in cases where new ecological niches can quickly be invaded, fitness advantages may be gained by populations where hybridization is common (Seehausen 2004). Following this model, empirical data have been accumulating, showing that this phenomenon, once considered extremely rare, may be more common than once thought (Salzburger et al. 2002; Koblmüller et al. 2007; Genner & Turner 2012; Ward et al. 2012; Cui et al. 2013). The general (and likely more common) rule, however, that post-zygotic isolation occurs via lowered hybrid fitness (known as the Dobzhansky–Muller model of hybrid incompatibility) has been show in several examples (Rogers & Bernatchez 2006; Crow et al. 2007, 2010; Gagnaire et al. 2012a,b; Montanari et al. 2012).

Conclusion—towards a unifying genomic approach

Speciation in different systems tend to result from some predominant factor; yet, artificially separating the different factors that play a role in speciation, as was done above, is to caricature a complex process that has multiple facets. Luckily, it is now possible to take a more comprehensive view of the entire process (Maan & Seehausen 2011). As was shown extensively above, fish present a very diverse array of scenarios, making them an ideal model to study speciation processes. That said, speciation in fishes has constraints that are peculiar to this group. Mating processes that rely on communication have to accommodate the aquatic medium, be it turbid, clear, with or without colour, ultraviolet or polarized light. This in turn will influence female choice and male display, which will drive the strength of assortative mating (Langerhans & Makowiec 2013; Langerhans & Riesch 2013). External fertilization has also resulted in strong potential for prezygotic isolation. Once fertilized, eggs and larvae generally move freely in the water column and thus influence dispersal, gene flow and the likelihood of breeding incompatibilities via neutral processes.

All these factors have been studied at length, as we have discussed, but mostly in isolation. For technical reasons, trying to understand speciation at the genetic level used to be restricted to the search of a single or few genes that would be the speciation smoking gun. Following the Dobzhansky–Muller (DM) model of hybrid incompatibility, where a gene would bring overall fitness to zero in a hybrid, the melanoma-causing gene Xmrk-2 in Xiphophorus swordtail hybrids, which has also deleterious effects in other fish species, such as medaka, was considered a good speciation gene candidate (Wu & Ting 2004). The recent developments of massively parallel sequencing are now allowing us to have a much more global view of the speciation process. They are bringing together, for the first time, genomic approaches that are unifying the language of speciation (Butlin et al. 2012; Gagnaire et al. 2013). All the factors influencing speciation leave strong genomic signatures and time is ripe to identify the fundamental molecular mechanisms responsible for these effects (Andrew et al. 2013).

In the early phase of speciation, gene flow occurs between two incipient species, which is theoretically an impediment to the genetic diversification necessary to produce a breeding barrier (Kulathinal et al. 2009). At the genomic level, however, chromosomal rearrangements and differential recombination will produce differential gene flow levels along the genome (Rieseberg 2001). In the best studied case of recent speciation, humans and chimps, genomic regions that are syntenic (where gene order is maintained in a colinear fashion) show much higher levels of gene flow than regions that are not (Farré et al. 2013). These regions are visualized as a local increase in Fst levels and create an opportunity for genetic disruption resulting in speciation, and these regions are often referred to as speciation islands (Nielsen et al. 2009; Nosil & Feder 2012; Hemmer-Hansen et al. 2013; Fig. 4). Therefore, high recombination would theoretically increase speciation rates. Genomic regions of high recombination were identified either as GC-rich regions or as fragile regions where GC content rapidly changes along the chromosomes (Fullerton et al. 2001; Marsolier-Kergoat & Yeramian 2009; Watanabe & Maekawa 2013). In general, fishes tend to be AT rich and homogeneous (meaning with few compositional jumps in their genome; Bucciarelli et al. 2002, 2009), suggesting that at the macroscopic scale, speciation rates should be relatively low in fishes.

Within the islands of speciation, it is unclear what may produce breeding barriers. The rapid divergence of
regions that experience low levels of gene flow may, in itself, prevent effective reproduction, not unlike the predictions of the DM model of hybrid incompatibility. In that respect, the islands of speciation, which persist in the face of high gene flow, were proposed to contain the genes responsible for reproductive isolation (Turner et al. 2005), which would be identified using Est outlier methods for example (Storz 2005); however, this is clearly not always the case (Macaya-Sanz et al. 2011). Instead, reproductive barriers may simply be generated by genetic incompatibilities due to a lack of syntenic regions (Stukenbrock et al. 2010), and there should be little difference if the regions are coding or noncoding, yet this remains controversial (Woolston 2013). Admittedly, the presence in those regions of key genes involved in reproduction predicts an increase in speciation rates. Because genes tend to be at higher concentration in GC-rich regions (Bernardi 2000) and those are subjected to higher recombination levels, there may be a link between a higher than expected concentration of genes and the location of speciation islands.

These are exciting times for the field of speciation, because we have new tools to quickly identify large numbers of SNPs and their significance to genetic divergence and speciation (Bernatchez et al. 2010; Hohenlohe et al. 2012; Krück et al. 2013). Vast amount of sequencing is affordable (e.g. Amemiya et al. 2013), and analytical tools are available to assess the demographic history of a species using the genomic sequence of very few or even a single individual, thus allowing for probing speciation events from the recent past (Li & Durbin 2011; Miller et al. 2012). Yet, as genomic tools become incredibly powerful and affordable, major biological questions remain, and these can only be fully understood in the light of careful examination and understanding of the actual organism in its natural environment. It is ironic that the mechanisms of speciation are becoming best understood in an era where the environment that promotes speciation is being relentlessly degraded.

Acknowledgements

I would like to thank my colleagues of the Department of Ecology and Evolutionary Biology at the University of California Santa Cruz for having broadened my biological knowledge over the years, in particular Pete Raimondi, Mark Carr, Grant Posgren, Bruce Lyon, Barry Sinervo and John Thompson. I would like to thank my father, Giorgio Bernardi, for discussing genome evolution for the better part of the last 30 years. Finally, I would also like to thank the members of my laboratory for their patience, good food and good humour, which makes my daily work most enjoyable.

References


Albert JS, Reis R (2011) Historical Biogeography of Neotropical Freshwater Fishes. University of California Press, ????.


© 2013 John Wiley & Sons Ltd
unification and terminological recommendations. In: Endless Forms Species and Speciation (eds Howard DJ, Berlocher SH), pp. 57–75. Oxford University Press, ????.


Fullerton SM, Carvalho AB, Clark AG (2001) Local rates of recombination are positively correlated with GC content in the human genome. Molecular Biology and Evolution, 18, 1139–1142.


Gagneaire P-A, Normandean E, Pavey SA, Bernatchez L (2012b) Mapping phenotypic, expression and transmission ratio distortion QTL using RAD markers in the Lake Whitefish (Coregonus clupeaformis). Molecular Ecology, ????, ????.


von der Heyden S, Bowie RCK, Prochazka K et al. (2011) Phylogeographic patterns and cryptic speciation across


Markert JA, Schelly RC, Stiasny ML (2010) Genetic isolation and morphological divergence mediated by high-energy rapiids in two cichlid genera from the lower Congo rapids. BMC Evolutionary Biology, 10, 149.


of the pharyngeal jaw key innovation in labroid fishes and beyond. *Systematic Biology*, 0, 1–27.


© 2013 John Wiley & Sons Ltd
Dear Author,

During the copy-editing of your paper, the following queries arose. Please respond to these by marking up your proofs with the necessary changes/additions. Please write your answers on the query sheet if there is insufficient space on the page proofs. Please write clearly and follow the conventions shown on the attached corrections sheet. If returning the proof by fax do not write too close to the paper’s edge. Please remember that illegible mark-ups may delay publication.

Many thanks for your assistance.

<table>
<thead>
<tr>
<th>Query reference</th>
<th>Query</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AUTHOR: A running head short title was not supplied; please check if this one is suitable and, if not, please supply a short title of up to 40 characters that can be used instead.</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>AUTHOR: Keywords are used form the author PDF version. Please check.</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>AUTHOR: Hausdorf (2013) has not been included in the Reference List, please supply full publication details.</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>AUTHOR: Bernardi et al. 2012 has been changed to Bernardi et al. 2012a, 2012b so that this citation matches the Reference List. Please confirm that this is correct.</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>AUTHOR: Bernardi, Beldade, et al. 2012 has been changed to Bernardi et al. 2012a, 2012b so that this citation matches the Reference List. Please confirm that this is correct.</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>AUTHOR: Please check and approve the edit made in the sentence “These results ultimately reconciled … species diversity in coral reefs.”</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>AUTHOR: April, Hanner, Dion-Côté, et al. 2013; April, Hanner, Mayden, et al. 2013 has been changed to April, Hanner, Dion-Côté, and Bernatchez 2013a, 2013b so that this citation matches the Reference List. Please confirm that this is correct.</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>AUTHOR: Figure 2 was not cited in the text. An attempt has been made to insert the figure into a relevant point in the text—please check that this is OK. If not, please provide clear guidance on where it should be cited in the text.</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>AUTHOR: Partridge 2001 has been changed to Cummings and Partridge 2001 so that this citation matches the Reference List. Please confirm that this is correct.</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>AUTHOR: Gagnaire et al. 2012 has been changed to Gagnaire et al. 2012a, 2012b so that this citation matches the Reference List. Please confirm that this is correct.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AUTHOR: Please provide the publisher location for reference Albert and Reis (2011).</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>----------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>AUTHOR: Please provide the page range for reference Albertson et al. (2003).</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>AUTHOR: Please provide the volume number, page range for reference Carstens et al. (2013).</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>AUTHOR: Please provide the volume number for reference Cummings and Partridge (2001).</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>AUTHOR: Please provide the publisher location for reference De Queiroz (1998).</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>AUTHOR: Please provide the publisher location for reference Dieckman et al. (2004).</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>AUTHOR: Please provide the page range for reference Fryer et al. (1972).</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>AUTHOR: Please provide the volume number, page range for reference Gagnaire et al. (2012b).</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>AUTHOR: Please provide the volume number, page range for reference Gagnaire et al. (2013).</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>AUTHOR: Please provide the volume number, page range for reference Gaiter and Rocha (2013).</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>AUTHOR: Please provide the publisher location for reference Helfman et al. (2009).</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>AUTHOR: Please check the details for reference Hohenlohe et al. (2012).</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>AUTHOR: Please provide the publisher location, page range for reference Howard and Berlocher (1998).</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>AUTHOR: Please provide the volume number, page range for reference Howe et al. (2013).</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>AUTHOR: Please provide the volume number, page range for reference Keller et al. (2012).</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>AUTHOR: Please provide the publisher location for reference Koop and Davidson (2008).</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>AUTHOR: Please provide the volume number, page range for reference Krück et al. (2013).</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>AUTHOR: Please provide the publisher location for reference Kunz-Ramsey (2008).</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>AUTHOR: Please provide the editor name(s), publisher location for reference Kuraku and Meyer (2010).</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>AUTHOR: Please provide the page range for reference Langerhans and Makowicz (2013).</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>AUTHOR: Please provide the volume number, page range for reference Lessios (2008).</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>AUTHOR: Please provide the volume number for reference Lowe-McConnell (1994).</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>AUTHOR: Please provide the publisher location for reference Mayr (1942).</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>AUTHOR: Please provide the publisher location for reference McCune (2004).</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>AUTHOR: Please provide the publisher location for reference Ohno (1970).</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>AUTHOR: Please provide the volume number for reference Planes and Doherty (1997).</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>AUTHOR: Please provide the volume number, page range for reference Plath et al. (2013).</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>AUTHOR: Please provide the volume number for reference Schinske et al. (2010).</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>AUTHOR: Please provide the volume number, page range for reference Vega and Wiens (2012).</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>AUTHOR: Please supply a brief Author Contributions Box describing the specific contributions of the authors to the published work. This is mandatory.</td>
<td></td>
</tr>
</tbody>
</table>
Using e-Annotation Tools for Electronic Proof Correction

Required software to e-Annotate PDFs: Adobe Acrobat Professional or Adobe Reader (version 8.0 or above). (Note that this document uses screenshots from Adobe Reader X.) The latest version of Acrobat Reader can be downloaded for free at: http://get.adobe.com/reader/

Once you have Acrobat Reader open on your computer, click on the Comment tab at the right of the toolbar:

This will open up a panel down the right side of the document. The majority of tools you will use for annotating your proof will be in the Annotations section, pictured opposite. We’ve picked out some of these tools below:

1. **Replace (Ins) Tool** – for replacing text.

   ![Replace Tool](image)

   **How to use it**
   - Highlight a word or sentence.
   - Click on the Replace (Ins) icon in the Annotations section.
   - Type the replacement text into the blue box that appears.

2. **Strikethrough (Del) Tool** – for deleting text.

   ![Strikethrough Tool](image)

   **How to use it**
   - Highlight a word or sentence.
   - Click on the Strikethrough (Del) icon in the Annotations section.

3. **Add note to text Tool** – for highlighting a section to be changed to bold or italic.

   ![Add Note Tool](image)

   **How to use it**
   - Highlight the relevant section of text.
   - Click on the Add note to text icon in the Annotations section.
   - Type instruction on what should be changed regarding the text into the yellow box that appears.

4. **Add sticky note Tool** – for making notes at specific points in the text.

   ![Add Sticky Note Tool](image)

   **How to use it**
   - Click on the Add sticky note icon in the Annotations section.
   - Click at the point in the proof where the comment should be inserted.
   - Type the comment into the yellow box that appears.
5. **Attach File Tool** – for inserting large amounts of text or replacement figures.

*Inserts an icon linking to the attached file in the appropriate pace in the text.*

**How to use it**
- Click on the Attach File icon in the Annotations section.
- Click on the proof to where you'd like the attached file to be linked.
- Select the file to be attached from your computer or network.
- Select the colour and type of icon that will appear in the proof. Click OK.

6. **Add stamp Tool** – for approving a proof if no corrections are required.

*Inserts a selected stamp onto an appropriate place in the proof.*

**How to use it**
- Click on the Add stamp icon in the Annotations section.
- Select the stamp you want to use. (The Approved stamp is usually available directly in the menu that appears).
- Click on the proof where you'd like the stamp to appear. (Where a proof is to be approved as it is, this would normally be on the first page).

7. **Drawing Markups Tools** – for drawing shapes, lines and freeform annotations on proofs and commenting on these marks.

*Allows shapes, lines and freeform annotations to be drawn on proofs and for comment to be made on these marks.*

**How to use it**
- Click on one of the shapes in the Drawing Markups section.
- Click on the proof at the relevant point and draw the selected shape with the cursor.
- To add a comment to the drawn shape, move the cursor over the shape until an arrowhead appears.
- Double click on the shape and type any text in the red box that appears.

For further information on how to annotate proofs, click on the Help menu to reveal a list of further options: