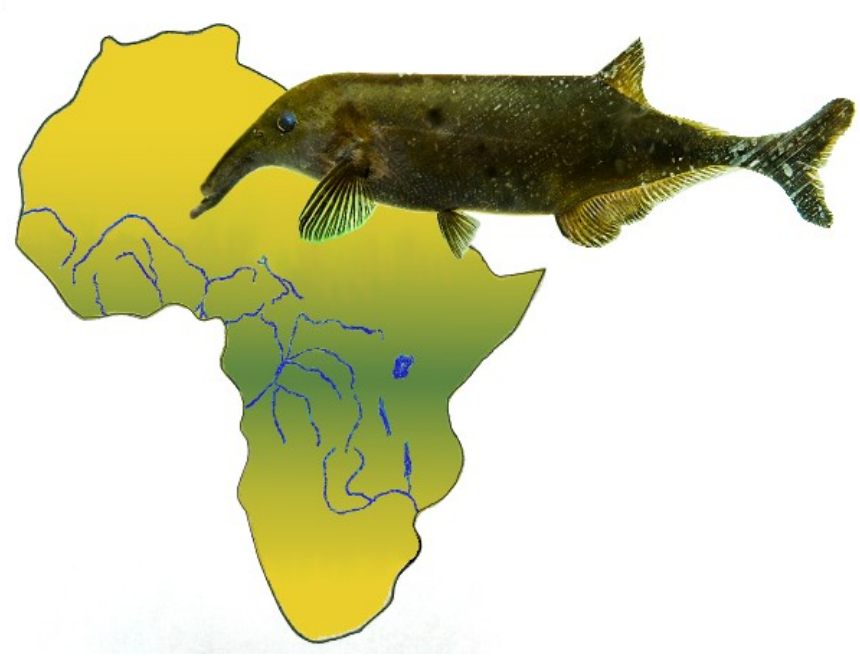
Institut für Biochemie und Biologie
Evolutionsbiologie / Spezielle Zoologie

Genome and karyotype evolution underlying speciation and diversification of electric organ discharges in African weakly electric fish

(*Campylomormyrus*, Mormyridae, Teleostei)



Dissertation
zur Erlangung des akademischen Grades
"doctor rerum naturalium"
(Dr. rer. nat.)
in der Wissenschaftsdisziplin Evolutionsbiologie

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
Institut für Biochemie und Biologie
der Universität Potsdam

von Julia Canitz

am 22. Mai 2019

Gutachter:

1. Prof. Dr. Ralph Tiedemann (Universität Potsdam)
2. Prof. Dr. Rüdiger Krahe (Humboldt Universität Berlin)
3. Prof. Dr. Walter Salzburger (Universität Basel)

Erklärung der Urheberschaft

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Arbeit ohne Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe.

Die Arbeit wurde bisher in gleicher oder ähnlicher Form in keiner anderen Prüfungsbehörde vorgelegt.

Ort, Datum Julia Canitz

For my family

Contents

List of Figures

List of Tables

1 Introduction

Teleost electric fish are characterized by the electric sense comprising generation (electric organ) and reception of electric signals (electroreception). They are very diverse and not restricted to a monophyletic group. They evolved independently multiple times within the teleosts and are represented by the families Uranoscopidae (Perciformes) and Malapteruridae and Mochokidae (Siluriformes), the superfamily Mormyroidei (Osteoglossifomes) consisting of two families (Gymnarchidae, Mormyridae) and the order Gymnotiformes which comprises five families (Gymnotidae, Rhamphichthyidae, Hypopomidae, Sternopygidae, Apteronotidae) (Table 1) (Nelson et al., 2016). Except for the marine species of Uranoscopidae, the fish occur in fresh water systems of South and Central America (Gymnotiformes) (Mago-Leccia, 1994) and Africa (Mormyroidei, Malapteruridae) (Poll, 1967; Hopkins, 1986; Moller, 1995) (Table 1). Teleost electric fish became a focus of evolutionary biology because of the uncommon use of their electric sense for hunting, communication and orientation (Bass, 1986a; Arnegard et al., 2006; Stoddard & Markham, 2008). Hereof, they developed the specialized organ, enabling them to generate an electric organ discharge (EOD) which in turn produces a surrounding electric field. Depending on the species, EODs can range from millivolts up to 800 volts. Those voltage differences are used to discriminate between weakly and strongly electric fish (Table 1). A strong EOD is represented in the marine Uranoscopidae and two freshwater families. One family is the neotropical Gymnotidae (Gymnotiformes) with the strong electric species *Electrophorus electricus*. The other one is the family of African electric catfish (Siluriformes, Malapteruridae) (Table 1). Strongly electric fish use their EOD mainly for predation and defense (Catania, 2015) while weakly electric fish use it for navigation, orientation, and communication (Stoddard & Markham, 2008; Worm et al., 2018). The weak EOD evolved in two monophyletic groups, the order Gymnotiformes, and the superfamily Mormyroidei (Lavoué et al., 2012). The fact that species of these groups share the trait of electrogenesis but are phylogenetically very distant, makes them an excellent example of convergent evolution (Lavoué et al., 2012).

Table 1: Overview of teleost electric fish sorted by continental distribution. Occurring combinations of main EOD characteristics in each taxonomic order are listed and the silhouette of a generic fish for the single feature combinations is given.

Distribution area	**Taxonomic order (family)**	**EOD strength**	**EOD type**	**Representative example**
South and Central America 	**Perciformes** (Uranoscopidae)	strong	pulse	*Astroscopus sp.*⋆ 
	Gymnotiformes (Gymnotidae, Rhamphichthyidae, Hypopomidae, Sternopygidae, Apteronotidae)	strong	pulse	*Electrophorus electricus* 
		weak	pulse	*Gymnotus sp.* 
		weak	wave	*Eigenmannia sp.* 
Africa 	**Siluriformes** (Malapteruridae, Mochokidae)	strong	pulse	*Malapterus sp.* 
	Osteoglossiformes (Mormyridae, Gymnarchidae)	weak	pulse	*Campylomormyrus sp.* 
		weak	wave	*Gymnarchus niloticus* 

⋆ marine species

1.1 African weakly electric fish

The superfamily Mormyroidei belongs to the basal teleost order Osteoglossiformes and consists of two families (Gymnarchidae and Mormyridae). Gymnarchidae contains only one species (*Gymnarchus niloticus*) while Mormyridae is the most species-rich family of its order ($\sim$ 220 described species) (Daget et al., 1984; Alves-Gomes & Hopkins, 1997). The species are endemic to Africa and are able to generate weak EODs. There are two EOD types among African weakly electric fish, the wave- and pulse-type EOD (Table 1). Wave-type EODs have a longer duration than pulse-type EODs. They are continuously emitted at equal time intervals and merge into a constant frequency wave (Figure 1). *Gymnarchus niloticus* is the only species of its superfamily with a wave-type EOD (Lissmann, 1951).

Pulse-type EODs are emitted in irregular sequences of pulse intervals whereby the pauses between two pulses are much longer and more variable than the pulse itself. They are as diverse as the mormyrid species and can be subdivided by polarity, number of phases, duration and overall shape (Figure 1). It is proposed that EODs are species-specific and highly stereotyped within individuals (Hopkins & Bass, 1980; McGregor & Westby, 1992). However, the EOD waveform can vary slightly among individuals of the same species, for example by gender, developmental stage, geographical origin or breeding season (Hopkins & Bass, 1980; Schwassmann et al., 2014; Gallant et al., 2017; Nguyen et al., 2017). Individuals of *Brienomyrus brachyistius* show variation in the EOD duration due to sexes and dominance among males (Carlson, 2002). Further, population-specific discharges based on the EOD duration and absence/presence of a so-called P0 phase were found among different *Paramormyrops kingsleyae* populations (Gallant et al., 2011).

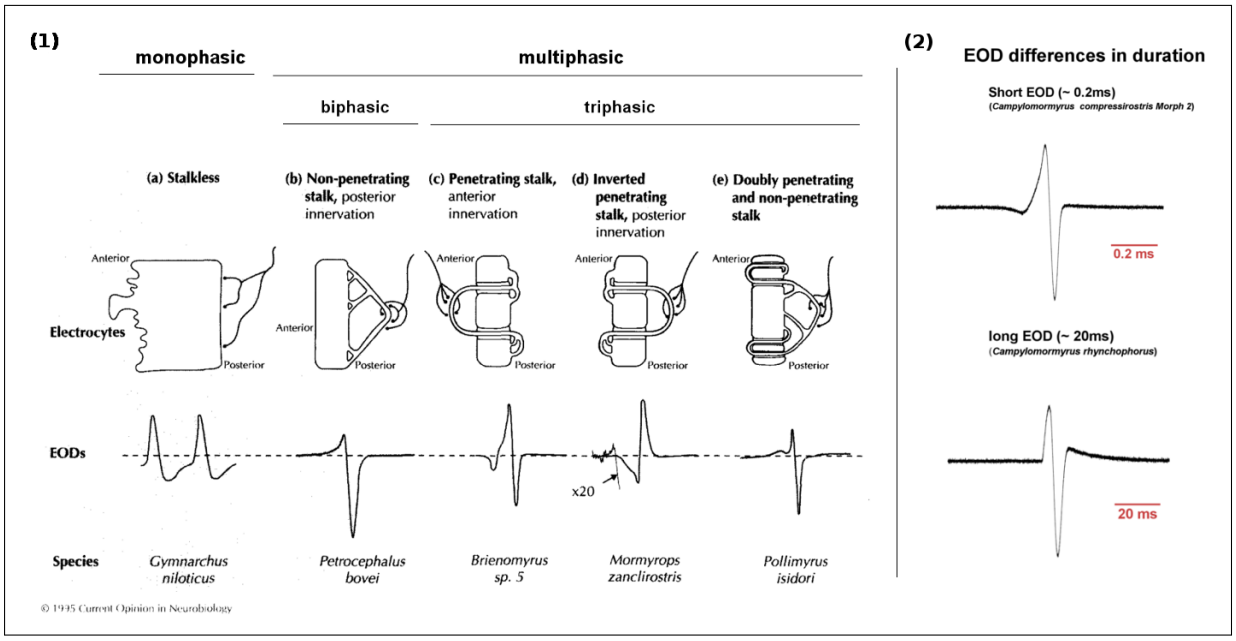


Figure 1: Different electrocyte geometries and the accompanied EODs in among species of the superfamily Mormyroidei. (1) Types of electrocytes (a-e) with innervation face and penetration are depicted. The appropriated EOD waveforms and mormyroid example are given below. (a) Stalkless electrocytes result in monophasic EODs. (b) A non-penetrating stalk is typical for a biphasic EOD. (c-e) Electrocytes with stalk penetrations lead to different triphasic EODs. Modified from Hopkins (1995). (2) EOD waveforms of two *Campylomormyrus* species with remarkably differences in EOD duration.

For evolutionary biology, the sympatric occurrence of African weakly electric fish is a further interesting fact. It raises the question whether their diversity is a consequence of sympatric speciation or allopatric speciation with secondary contact. In the case of sympatric speciation, certain evolutionary drivers are involved reaching reproductive

isolation. Temporal or ecological differences can cause reproductive isolation (micro-allopatry) if individuals are adapted to certain times of activity (nocturnal vs. diurnal) or habitats (e.g. ecological niches or seasonal migration) (Tsukamoto et al., 2002). In those cases, speciation depends on individuals not encountering each other despite occupying the same geographical range. Behavior is another important characteristic in the context of reproductive isolation. It plays a major role in mating processes and thus can drive speciation by factors like direct mate choice (Knight et al., 1998; Feulner et al., 2008) or male courtship (Kitano et al., 2009). These prezygotic reproductive isolation mechanisms can cause a mating pattern with preferences for individuals with conformities or disparities (assortative mating). Assortative mating, a form of sexual selection, can in turn act directly on signals involved in prezygotic reproductive isolation and thus can be a potent driver of speciation (Carlson et al., 2011). Either way, sexual selection is one of the main hypotheses in the evolution of sympatric species (Higashi et al., 1999; Boughman, 2001).

Speciation can also be a consequence of disruptive natural selection as shown in sympatric fish species where it acts on body size and shape resulting in ecological speciation (Schliewen et al., 1994). One uncertainty here is whether prezygotic isolation mechanisms emerged as a consequence of disruptive natural selection or whether disruptive natural selection acts on systems which already developed prezygotic isolation mechanisms (Schluter, 1996). In sympatrically occurring African weakly electric fish, the EOD is presumed to be a prezygotic isolation mechanism (Albert & Crampton, 2005; Feulner et al., 2009b) and the divergence of electric signals facilitated the diversification of electric fish under each of those evolutionary scenarios (Carlson et al., 2011). An alternative process which can cause sympatric speciation is postzygotic isolation by gametic incompatibility, zygotic mortality, hybrid sterility and inviability as well as hybrid breakdown. In addition, chromosomal changes have recently been identified as a postzygotic isolation mechanism in sympatric electric fish (Cardoso et al., 2018). In addition to these mechanisms, it should be noted that hybridization and gene flow between sympatric species may be further drivers of their divergence. For electric fish, this could confound their current species definition and classification, and prompt a revision of species descriptions with respect to the biological species concept.

1.2 Electrogenesis in African weakly electric fish (Mormyridae)

Over the last decades, histological and molecular properties of the electric organ have been studied in more detail among African weakly electric fish (Lissmann, 1958; Bennett, 1961; Bass, 1986a,b; Stoddard, 2006; Kirschbaum et al., 2016). The organ is derived from muscle tissue (myogenic organ) whereby its cells, called electrocytes, have lost the ability to contract, and became electrogenic. Histological investigations showed that electrocytes vary by structures including innervation faces, papillae, penetrations and stalk formations (Bass, 1986a; Hopkins, 1995; Paul et al., 2015). Variation in these electric organ structures are associated with particular EOD waveforms. For example, electric fish with stalk-penetrated electrocytes emit a triphasic EOD while no penetrations result in a biphasic EOD (Figure 1). However, the actual generation of a discharge is determined by cellular processes and molecular mechanisms of electrocytes. In mormyrids, an electric organ is composed of tens of electrocytes stacked in cylindrical columns which maximizes the voltage of an EOD. Four columns are arranged in parallel which increases the generated current. They are orientated rostro-caudally and are surrounded tightly by connective tissue, blood vessels and nerve fibers (Figure 2). The membrane of each electrocytes is excitable and equipped with voltage-gated ion channels, which serve as gates for positively charged sodium (Na^+) and potassium (K^+) ions (Stoddard & Markham, 2008). By ion flow through the membrane, the charge of the membrane and intra- and extracellular speace changes. These ions can flow through the membrane and change the charge of the intra- and extracellular space. Considering the whole individual, the EOD is similar to an oscillating dipole embracing the head and tail of the fish with changing head-positive and head-negative phases that cycle with the EOD repetition rate (Albert & Crampton, 2005).

Assuming a simple biphasic EOD with a first head-positive and a second head-negative phase, each electrocyte is innervated at the smooth posterior face by electromotor neurons via a stalk while the anterior face is grooved which increases the cell surface (Figure 2). The initial phase of an EOD is the resting potential where an excess of positively charged Na^+ ions exists in the extracellular space compared to the intracellular space. When the posterior face of electrocytes becomes activated, it initiates the first action potential and Na^+ ions enter the cell (depolarization) in anterior direction. A repolarization follows, meaning that the inward Na^+ ion current is countered by a K^+ ion outflow as potassium ion channels are opened at the anterior face of the electrocyte (Figure 2; red). During these milliseconds, the flow of positive ions is head-ward, charging the head of the fish

positively (head-positive phase). The second action potential is produced immediately after the first one (Figure 2; black). But this time, the de- and repolarization run vice versa leading to a current of positive ions from the head to the tail (head-negative phase). Afterwards, the resting potential is restored (Stoddard & Markham, 2008). The cellular principle is similar in all weakly electric fish but the diversity of EODs might be caused by multiple factors like electrocyte geometry, membrane excitability, or hormone levels (Bass, 1986a; Bass et al., 1986; Freedman et al., 1989).

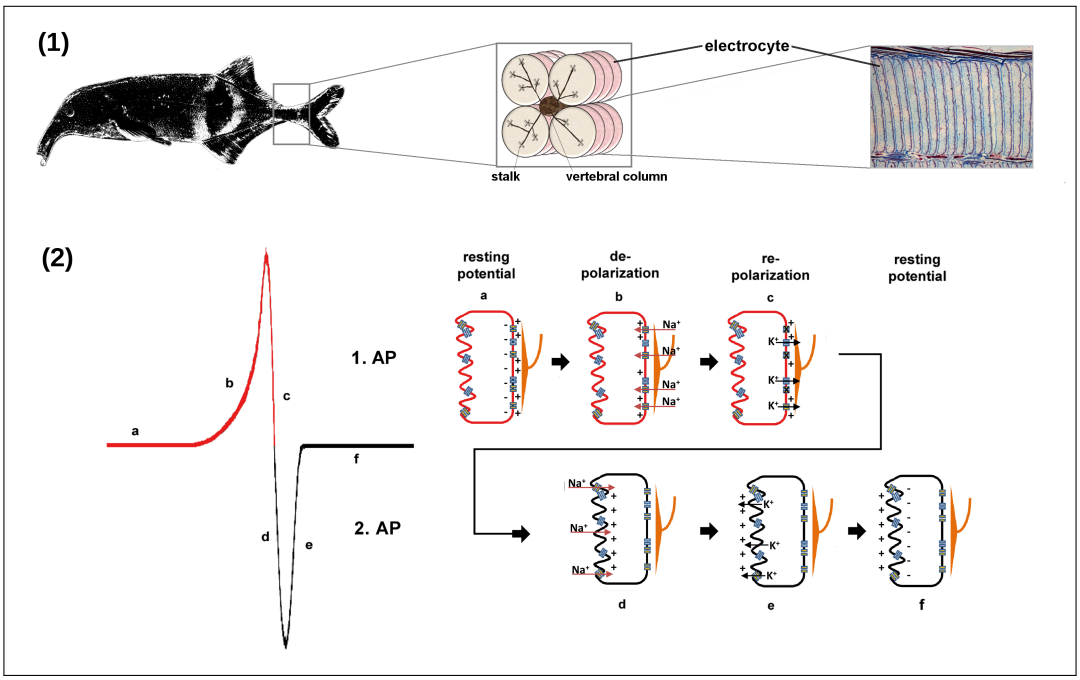


Figure 2: Morphological position, schematic illustration and histological picture of a mormyrid electrical organ, and the schematic sequence of ionic flows during the generation of an biphasic EOD. (1) The electric organ is positioned close to the tail and consists of electrocytes stacked in four columns. (2) On the left, a biphasic EOD with resting stages (a,f), depolarization (b,d) and re-polarization (d,e). On the right, a sketch of ion flows is illustrated. The first action potential is initiated by the motor neuron at the posterior site indicated by the first head-positive phase (red). The following head-negative phase is caused by a second action potential at the anterior face (black).

1.3 Genetics of African weakly electric fish (Mormyridae)

Today, one of the most common ways to classify taxonomic units and define species is the usage of genetic markers. Microsatellites and gene sequences serve as tools in topics of evolutionary biology like species phylogeny or population structure, but are also helpful in conservation biology. Several genetic markers (e.g. *cyt b* or *COI*) are used to identify taxonomic lineages. They have been utilized to confirm genetically the convergent evolution of weakly electric fish (Lavoué et al., 2000, 2012). Within mormyrids they were used

to reconstruct phylogenies and assign known species to monophyletic groups or distinct clades (Feulner et al., 2007, 2008; Lamanna et al., 2016). However, single markers have their restrictions when addressing additional evolutionary questions. Therefore, a wider range of genetic data is needed to avoid biased conclusions in areas like species admixture or when investigating the impact of selection on speciation.

The development of Next Generation Sequencing (NGS) allows researchers access to thousands of genetic markers in a cost-efficient way. It enables genome assemblies like the one of the mormyrid *Paramormyrops kingsleyae* (Gallant et al., 2017). This genome reconstruction has been a crucial step for future research of mormyrid evolution and can now serve as reference for gene annotations and expression analyses of closely related species. These NGS applications are essential to gain insights into the genetic variability of mormyrid fish. Knowledge about karyological features like chromosome number also supports a genome reconstruction as the number of scaffolds are easier to estimate. Unfortunately, chromosome numbers are only known for a few taxa in the species-rich Mormyridae (Uyeno, 1973; Ozouf-Costaz et al., 2015). Especially concerning the teleost-specific genome duplication, a known chromosome number would facilitate the estimate of genome sizes and rearrangements.

This genome duplication is assumed to have happened 300 to 350 million years ago (Meyer & Van De Peer, 2005; Inoue et al., 2015; Pasquier et al., 2016), and resulted in the presence of numerous paralogous genes which were able to gain sub- or neofunctionalizations (Ohno et al., 1967; Zhang, 2003; Ozernyuk & Myuge, 2013). Two gene families (sodium and potassium voltage-gated ion channel genes), which are involved in EOD generation, are paralogous and studies have focused on their evolution and molecular structure (Arnegard et al., 2010; Paul et al., 2016). It is presumed that their different genetic structure is one main factor leading to the EOD diversification, especially in terms of expression of paralogous genes in different tissues like the skeletal muscle and electric organ (Zakon et al., 2006; Lamanna et al., 2015; Nagel et al., 2017). Additionally, single mutations in DNA and amino acid sequences seem to be taxon-specific and can be allocated to distinct characteristics like the EOD duration (Paul et al., 2016; Swapna et al., 2018). In contrast to the examination of individual genes, entire transcriptome analyses can be applied by means of NGS. This has been already used for differential gene expression analyses across tissues, and across species with EOD variation (Gallant et al., 2012; Lamanna et al., 2015; Gallant et al., 2017).

Concerning EOD waveform variability of the electric fish, organ anatomy and electrocyte geometry have been well studied, but their underlying genetics have not yet been comprehensively compared. Consequently, genes which might be involved in the cell structure, development or differentiation, and thus in tissue formation could be promising candidate genes. In summary, gene mutation, expression, and regulation have been examined to address evolutionary questions in African weakly electric fish, but there is still a lack of knowledge about evolutionary and genetic backgrounds driving EOD diversification in mormyrids.

1.4 Genus *Campylomormyrus*

In 1874, the genus *Campylomormyrus* was defined for the first time (Bleeker, 1874), however subsequent specimens were listed under the genus name *Gnathonemus*, for which *Campylomormyrus* was synonymous (Boulenger, 1898; Poll, 1945). It was only in 1968 that Taverne (1968) reintroduced the genus name *Campylomormyrus* (Taverne, 1968; Poll et al., 1982). Since the first record, the number of valid *Campylomormyrus* species has varied. Due to different authors and their morphological analyses, specimens were denoted with different names (Table 1 in Feulner et al. (2007)). A recent example refers to the species *C. ibis* as the juvenile state of *C. numenius* (Paul et al., 2015), and should no longer be listed as a distinct species. Thus, 15 described species are currently accepted (Table 2).

Table 2: Overview of 15 accepted and two putative *Campylomormyrus* species. The list includes species name, a generic EOD and its time scale, the stalk form/innervation face, morphological example and the methods of classification. Fish body sizes are not relative to each other.

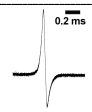
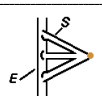
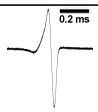
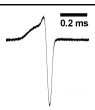
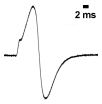
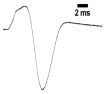
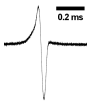
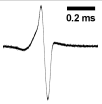
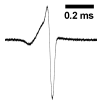
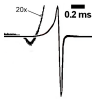
Species	**EOD**	**stalk form/ innervation face**	**morphological shape**	**classified by** ⋆
C. compressirostris Morph 1				genetics / morphometrics / holotypus / EOD
C. compressirostris Morph 2 ⋆⋆		?		genetics / morphometrics / EOD
C. curvirostris		?		genetics / morphometrics/ holotypus / EOD

Table 2: continued from previous page

Species	**EOD**	**stalk form/ innervation face**	**morphological shape**	**classified by** ⋆
C. tshokwe				genetics / morphometrics/ holotypus / EOD
C. numenius				genetics / morphometrics/ holotypus / EOD
C. rhynchophorus				genetics / morphometrics/ holotypus / EOD
C. elephas		?		genetics / morphometrics/ holotypus / EOD
C. alces		?		genetics / morphometrics/ holotypus / EOD
C. christyi		?		genetics / morphometrics/ holotypus
C. tamandua				genetics / morphometrics/ holotypus / EOD
C. n.sp. ⋆⋆		?		genetics / morphometrics
C. phantasticus ⋆⋆⋆		?		holotypus / EOD
C. bredori ⋆⋆⋆	?	?		holotypus
C. cassasicus ⋆⋆⋆	?	?		holotypus
C. luapulensis ⋆⋆⋆	?	?		holotypus
C. mirus ⋆⋆⋆	?	?		holotypus
C. orycteropus ⋆⋆⋆	?	?		morphology / holotypus

? = no data available, S = stalk, E = electrocyte

⋆ references: Taverne (1968); Poll et al. (1982); Feulner et al. (2008); Lamanna et al. (2016)

⋆⋆ not an accepted species but listed in the phylogeny of Lamanna et al. (2016)

⋆⋆⋆ pictures of morphological shape adapted from *http://mormyrids.myspecies.info/en*

Campylomormyrus occurs exclusively in the Congo Basin except for two species. *Campylomormyrus tamandua* has been also found in northern rivers like Volta, Chad/Shari and Niger (Daget et al., 1984) while *C. phantasticus* is endemic to the Sanaga River in Cameroon (Pellegrin, 1927). Since the distribution area of most of the species overlaps, they are considered as sympatric species (Feulner et al., 2008; Tiedemann et al., 2010). The natural habitat of *Campylomormyrus* species is rather diverse ranging from calm waters like Lake Moero to strong currents as it is found in some parts of the Lower Congo River (Gosse, 1984). They are nocturnal and feed on insect larvae, and small crustacean (Marrero & Winemiller, 1993). Their local name, the "poisson elephant", already indicates the distinct feature of the elongated trunk-like snout with the small mouth at the end (Table 2). Morphological characteristics of the head are used to classify species (Bleeker, 1874; Boulenger, 1898; Feulner et al., 2007; Lamanna et al., 2016), but the EOD is also appropriated for species differentiation as it is supposed to be highly stereotyped within species and shows no differences among sexes in *Campylomormyrus* (Feulner et al., 2009a; Paul et al., 2015). It has been shown that the EOD of some species changes dramatically during the ontogeny (Paul et al., 2015; Nguyen et al., 2017) but after maturity it remains constant in the individual and is a reliable discrimination trait (Feulner et al., 2007, 2009a).

Concerning the different EOD characteristics, the phase number and duration have been intensively investigated (Paul et al., 2015; Kirschbaum et al., 2016; Nguyen et al., 2017; Nagel et al., 2018a). The different EOD phases and their polarity are traced back to the anatomy of its electric organ and electrocytes. Duration differences were mostly studied histologically and genetically. Thereby, differences in electrocyte geometries could be stated (Paul et al., 2015), and single amino acid changes in the sodium voltage-gated ion channel were found being exclusively in species with elongated EODs (Paul et al., 2016). Nagel et al. (2017) detected different gene expression patterns of potassium voltage-gated ion channels among two species with different EOD waveform lengths. As in all mormyrids, *Campylomormyrus* uses the EOD for communication including species recognition and discrimination (Nagel et al., 2018a,b). Based on mate choice experiments, it is thought that the EOD facilitates assortative mating and functions as a prezygotic isolation mechanism which promotes natural divergent selection. Hence it is proposed to be a 'magic trait' (Feulner et al., 2009b; Tiedemann et al., 2010). Even if the EOD does not seem to be the only modality involved in species discrimination, it is hypothesized that discrimination by electric cues alone might be an apomorphic trait that evolved during a recent species

radiation in *Campylomormyrus* (Nagel et al., 2018a,b). Another hypothesis concerning their radiation is the morphological adaptation to ecological niches. Differences in their snout morphology leads to the assumption that species evolved by adaptation to different food sources promoting the divergence of *Campylomormyrus* (Feulner et al., 2009b). Those different traits might be involved in the divergence of sympatrically occurring *Campylomormyrus* species, and each scenario or even a combination seems likely. Nevertheless, it is still uncertain if its recent radiation is caused by ecological speciation (adaptation to food sources), sexual selection promoted by the different EODs (assortative mating), and/or natural divergent selection leading to the prezygotic reproductive isolation or caused by it. It should be considered that hybridization and gene flow can potentially occur among the sympatric *Campylomormyrus* species, especially as inter- and intra-genus F1 hybrids have been produced under laboratory conditions (Kirschbaum et al., 2016).

The recent species tree of *Campylomormyrus* was published by Lamanna et al. (2016). It includes eight morphologically described and one undescribed species, and is based on three genetic markers. Species delimitations are supported by microsatellite data, morphometric measurements and differences in EOD waveforms, but not in all cases the species affiliation is concordant. For example, a classification by morphometric measurements yields a lower species number than one by EOD waveform, as two EOD morphs of *C. compressirostris* are observed (Lamanna et al., 2016). However, in the inferred species tree the nine species are divided into two main clades (Figure 3). One of them is called the *alces*-complex (Figure 3) and comprises the three species *C. christyi, C. alces* and *C. elephas*. This clade is not fully resolved and EOD waveforms vary within species (unpublished data). The other main clade consists of the remaining species excluding *C. tamandua* (Clade 1; Figure 3), and splits into two sub-clades and a single lineage (undescribed species *C. n.sp.*). One sub-clade includes the species *C. compressirostris, C. curvirostris* and *C. tshokwe*. All species except for *C. compressirostris* Morph 2 have a biphasic EOD but with different duration. The other sub-clade consists of *C. numenius* and *C. rhynchophorus* showing elongated EODs (>10ms) but with differences in phase number (Table 2). *Campylomormyrus tamandua* is the most basal lineage and the sister group to both main clades . It possess a triphasic EOD which is consistent with the phase number of the genus' sister group *Gnathonemus* (Sullivan et al., 2000; Lamanna et al., 2016). Consequently, it is assumed that a triphasic EOD waveform is the ancestral state

in the genus. Regarding the EOD duration, a short pulse is likely ancestral, and thus, the elongation of discharges (>2ms) is considered a derived characteristic (Tiedemann et al., 2010). Comparing these features with the current phylogeny means that an EOD elongation evolved independently and is fixed in at least two phylogentic lineages of *Campylomormyrus* (sub-clade: *C. rhynchophorus* and *C. numenius*; species *C. tshokwe*).

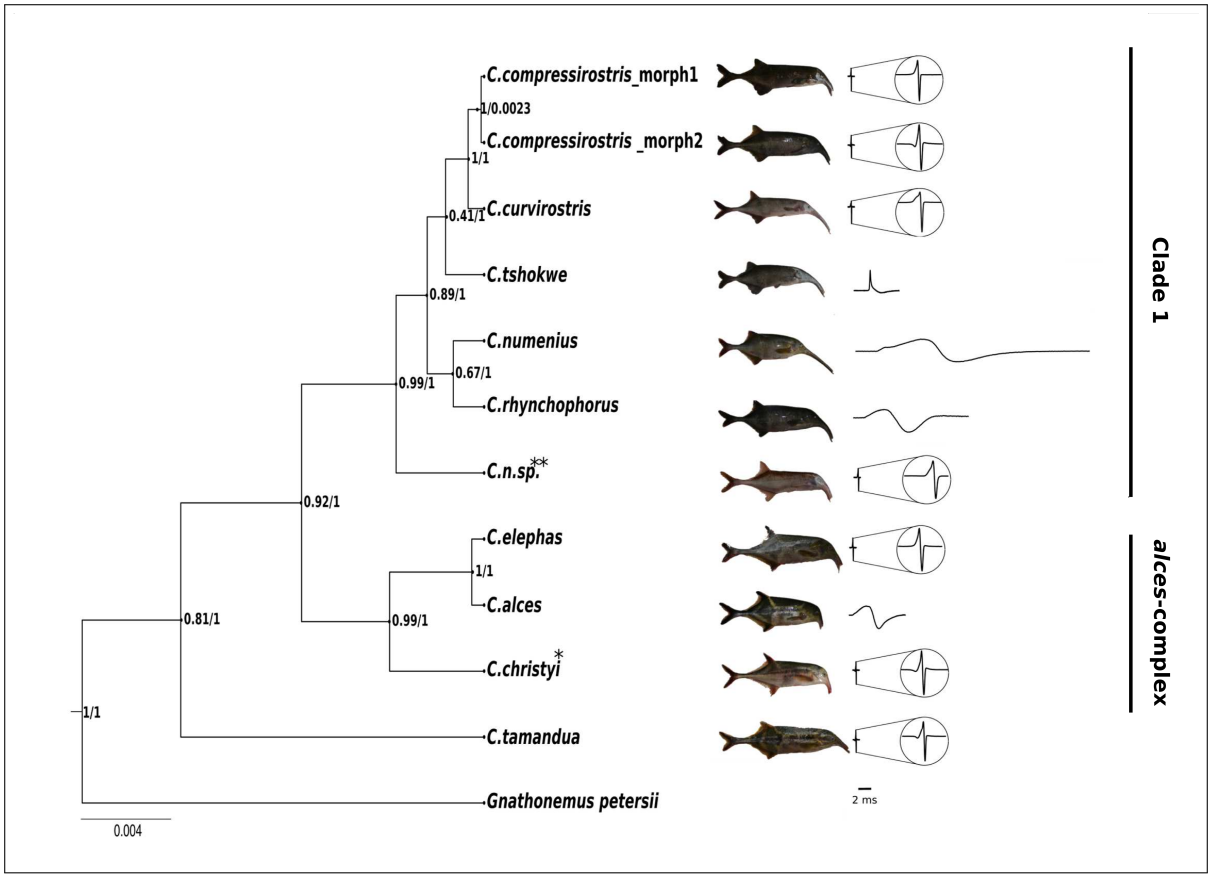


Figure 3: Current *Campylomormyrus* phylogeny. Eight species are morphologically described, one is still undescribed (*C. n.sp*) and 2 EOD morphs are represented for *C. compressirostris*. Clade 1 and the *alces*-complex are labels only used in this thesis. Modified from (Lamanna et al., 2016).

1.5 Aims of the study

The EOD as a species-specific trait in African weakly electric fish is proposed to promote prezygotic reproductive isolation by facilitating assortative mating (Feulner et al., 2009b). Hence, it plays a key role in the speciation of the sympatrically occurring fish. The overall aim of this study is a better understanding of the correlation between EOD characteristics and species diversification within the African weakly electric fish genus *Campylomormyrus*, and which evolutionary concepts are subject to this diversification. In more detail, I want to

investigate the genetic variation within *Campylomormyrus* underlying EOD diversification and species divergence to gain insights into evolutionary processes driving the probalble sympatric speciation in the genus. For this, I focus on chromosomal evolution, the genetic background of the EOD diversity, and the genomic differences among *Campylomormyrus* species. Therefore I followed three different approaches:

1) Karyological approach: Although some evolutionary events and processes can be identified by genomic structures, like chromosome number and shape, such a fundamental information is still lacking for several mormyrid species. Therefore, I aim to describe the first karyotype of the genus *Campylomormyrus* and to compare it with existing chromosome sets within Osteoglossiformes. With this insight, the genus can be classified in the context of chromosome evolution and chromosome rearrangements among osteoglossiform taxa can be inferred. Additionally, this information serves as basis for further investigations including genome reconstruction and gene localization.

2) Genomic approach: Even if the number of chromosomes is a basic information, differences in phenotypic traits (e.g. EOD) are encoded by single genes. So far, differential gene expression analyses were applied across *Campylomormyrus* species with different EODs, focusing on voltage-gated ion channel genes (Lamanna et al., 2015; Nagel et al., 2017). The second aim of this thesis is the identification of genetic differences in expressed genes among closely related *Campylomormyrus* species with different EOD duration. As RNA sequencing data are available, I set up a transcriptome-wide SNP analysis pipeline to discover candidate genes and cellular mechanisms with increased genetic variation which take part in the determination of EOD characteristics in a species with an elongated EOD (*C. tshokwe*).

3) Phylogenetic approach: *Campylomormyrus* fish occur in sympatry but most of them can be well discriminated morphologically, genetically, and by their species-specific EOD. In the current phylogeny, most of the clades can be assigned to previously described species but some are missing and uncertainties among species, like in the *alces*-complex, still exist (Lamanna et al., 2016). Thus, a further aim of this study is a comprehensive phylogenetic reconstruction based on an increased number of genetic markers and including two additional species. For this I used the ddRAD sequencing method to gain a higher amount of genetic markers (SNPs) which aid to resolve phylogenetic ambiguities, delimit species and provide more information about evolutionary mechanisms driving speciation.

2 Summary of Articles

2.1 Summary of Article I

Karyotype description of the African weakly electric fish *Campylomormyrus compressirostris* in the context of chromosome evolution in Osteoglossiformes.

Canitz, J., Kirschbaum, F., Tiedemann, R.
Journal of Physiology - Paris **110**, 273-280 (2016)

Despite several genetic studies about *Camyplomormyrus* species, a key information of this realm - its karyotype - is still lacking. Hence, I have described the first karyotype of *Campylomormyrus compressirostris* in this study. I have tested different conditions for chromosome preparation following the methods of Shao et al. (2010) and Karami et al. (2015) to gain the best chromosome spreads. To visualize the chromosomes I stained them with DAPI and Giemsa. Finally, I preserved the chromosomes with DPX and took microscopic pictures. The results lead to the conclusion that *C. compressirostris* has a diploid chromosome set of 2n=48 consisting of 4 submetacentric, 26 metacentric, and 18 acrocentric chromosomes (fundamental number=78). This diploid chromosome number is consistent with the chromosome number of its sister genus *Gnathonemus petersii* supporting their close relationship on the chromosome level. Furthermore, the results extend the current number of karyotypes in Osteoglossiformes to 21 species in 17 genera.

Subsequently, I have used the software ChromEvol to predict the most likely haploid chromosome number of common ancestors within the order Osteoglossiformes. As the software takes into account phylogenetic information, a maximum likelihood approach has been used to generate a phylogenetic tree based on mitochondrial genomes of 16 osteoglossiform species. Based on the data, the software infers a most likely haploid chromosome number of n=27 for the most recent common ancestor of Osteoglossiformes. As other studies have suggested a haploid chromosome number of n=24, I compared two ChromEvol analyses with different fixed ancestral chromosome numbers (n=24 and n=27). The results of both runs deviate just slightly regarding the ancestral chromosome numbers.

Although the comparison indicates a haploid chromosome number of 27 as most likely, I suggest that the ancestral haploid chromosome number of Osteoglossiformes is n=24. This is consistent with the current hypothesis, supported by the most basal phylogenetic

branch, Pantodon (n=24), and it is assumed that the ancestral chromosome number of all teleostei is n=24. Furthermore, I suggest that the ancestral chromosome number of mormyrids is n=24-25 as the probabilities are the highest and second highest ones in both analyses. Besides all the differences, a higher number of chromosome losses than gains has been inferred in both runs. The total amount of chromosomal changes from the root to the tips is supported by the highly variable karyotype formula among the taxa of Osteoglossiformes. I assume that these changes are caused by Robertsonian fusion and pericentric inversion.

Authors contribution: I performed the laboratory experiments, took the microscopic pictures, did the data analyses, and drafted the manuscript. F. Kirschbaum provided the fish eggs and contributed to experimental design to hatch the eggs. R. Tiedemann conceptualized and supervised the study, and participated in the discussion and finalization of the manuscript. All authors read and revised the manuscript.

2.2 Summary of Article II

Transcriptome-wide single nucleotide polymorphisms related to electric organ discharge differentiation among African weakly electric fish species

Canitz, J., Kirschbaum, F., Tiedemann, R.
PlosONE, submitted

Pulse duration in African weakly electric fish is presumed to depend on electrocyte geometry/excitability or hormonal signaling. This study provides the first transcriptome-wide Single Nucleotide Polymorphism (SNP) analysis among species of the African weakly electric fish genus *Campylomormyrus*, as well as putative candidate genes and cellular mechanisms which are potentially involved in the determination of an elongated discharge in one of the *Campylomormyrus* species. To achieve those results, I compared orthologous transcripts of three *Campylomormyrus* species (*C. tamandua*, *C. compressirostris*, and *C. tshokwe*) and *Gnathonemus petersii* regarding species-specific non-synonymous SNPs for *C. tshokwe*, the only species in this group with an elongated EOD.

I identified 27 candidate genes with inferred positive selection and at least one non-synonymous *C. tshokwe*-specific SNP. These candidate genes have functions associated directly or indirectly with transcriptional regulation, cell proliferationa and differentiation. Additionally, I have found 27 GO terms and two KEGG categories for which *C. tshokwe* possesses a significantly higher frequency of expressed genes with species-specific substitutions than *C. compressirostris*. It is most likely that transcriptional regulation, cell proliferation and differentiation, which are pivotal for tissue morphogenesis, are involved in the determination of pulse duration. This supports the observed correlation between tissue structure of the electric organ and discharge duration.

Authors contribution: I set up the bioinformatical pipeline, analyzed the data, and drafted the manuscript. F. Kirschbaum participated in the manuscript drafting and supervision. R. Tiedemann conceptualized and supervised the study and contributed to manuscript finalization. All authors read and approved the manuscript.

2.3 Summary of Article III

Genome-wide SNPs provide new insights into the phylogeny of the African weakly electric mormyrid fish genus *Campylomormyrus* (Teleostei; Osteoglossiformes)

Canitz, J., Kirschbaum, F., Mamonekene, V., Tiedemann, R.
to be submitted to *Molecular Phylogenetics and Evolution*

The current species tree of *Campylomormyrus* is based on three genetic markers and contains ten species which occur sympatrically at the Lower Congo River (Lamanna et al., 2016). Since uncertainties in the phylogeny and species delimitation still exist, and six morphologically described species are not included, this study aims at a reduced genome approach with focus on individual phylogenetic relationships, species delimitations and evolutionary processes driving the assumed sympatric speciation. To extend the data set, I included two additional species (*C. phantasticus* and *C. cassaicus*) and to obtain an increased number of genetic markers I established the double digest restriction-site associated DNA (ddRAD) sequencing method in our laboratory. In total, libraries for 63 samples have been prepared covering twelve *Campylomormyrus* and one *Gnathonemus* species. Using a denovo assembly pipeline, provided by the analysis tool Stacks 2.2, I could assemble more than 150,000 loci. All samples were genotyped and the SNPs were filtered for different parameters like sequence quality, minor allele frequency and sample coverage. A Maximum Likelihood and Bayesian Inference analysis have been performed to reconstruct the phylogeny. In addition, more than 8,000 genotyped positions were used for a Principle Component Analysis (PCA) and a Discriminant Analysis of Principle Components (DAPC) to group the individuals by genetic similarity. For gaining a higher resolution among particular species I repeated the pipeline for the specimens of the so-called *alces*-complex.

The new phylogeny reveals a major change in the topology as the phylogenetic position of *C. tamandua* is not anymore the most basal lineage within the genus *Campylomormyrus*. It is now the sister group of the *alces*-complex. Moreover, the phylogenetic trees yield that the species *C. phantasticus* is the basal lineage in one of the main clades and *C. cassaicus* belongs to the *alces*-complex. The individual-based phylogeny reveals well-supported relationships within and among species (PP $>$ 90). However, the genetic similarity especially among individuals of the *alces*-complex is high and they cluster neither by species nor

by EOD. This in combination with its high EOD diversification could indicate incipient sympatric speciation for individuals of the *alces*-complex (Arnegard et al., 2005, 2006). Such an assumpition would also lead to a higher species number increasing the diversity of *Campylomormyrus*in the Lower Congo River. Thus, the ddRAD sequencing method is a useful tool to uncover species diversity as relationships even among individuals are solved and well supported.

Author's contributions: I did the laboratory work including protocol establishment and library preparation. Furthermore, I performed data processing and drafted the manuscript. V. Mamonekene supported the group during the field trips and contributed to a successful sampling. F. Kirschbaum participated in the field trips and contributed to species classification and manuscript drafting. Ralph Tiedemann also joined the field excursions, contributed to conceive the study and participated in manuscript drafting.

3 Article I

Karyotype description of the African weakly electric fish *Campylomormyrus compressirostris* in the context of chromosome evolution in Osteoglossiformes

Julia Canitz, Frank Kirschbaum, Ralph Tiedemann

Published in

Jornal of Physiology - Paris (2016)

Journal of Physiology - Paris 110 (2016) 273–280

Contents lists available at ScienceDirect

Journal of Physiology - Paris

journal homepage: www.elsevier.com/locate/jphysparis

Original Research Paper

Karyotype description of the African weakly electric fish *Campylomormyrus compressirostris* in the context of chromosome evolution in Osteoglossiformes

Julia Canitz [a], Frank Kirschbaum [b], Ralph Tiedemann [a,*]

[a] Unit of Evolutionary Biology/Systematic Zoology, Institute of Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Straße 24-25, 14476 Potsdam, Germany
[b] Faculty of Life Sciences, Biology and Ecology of Fishes, Humboldt University of Berlin, Philippstraße 13 (Haus 16), 10115 Berlin, Germany

ARTICLE INFO

Article history:
Received 29 June 2016
Received in revised form 21 December 2016
Accepted 14 January 2017
Available online 17 January 2017

Keywords:
Karyotype
Campylomormyrus
Chromosomal evolution
Osteoglossiformes
ChromEvol

ABSTRACT

Karyotyping is a basic method to investigate chromosomal evolution and genomic rearrangements. Sixteen genera within the basal teleost order Osteoglossiformes are currently described cytogenetically. Our study adds information to this chromosomal dataset by determining the karyotype of *Campylomormyrus compressirostris*, a genus of African weakly electric fish that has not been previously examined. Our results indicate a diploid chromosome number of 2n = 48 (4sm + 26m + 18a) with a fundamental number of FN = 72. This chromosome number is identical to the number documented for the sister taxon of the genus *Campylomormyrus*, *i.e.*, *Gnathonemus petersii* (2n = 48). These results support the close relationship of *Campylomormyrus* and *Gnathonemus*. However, the karyotype formula of *C. compressirostris* is different from *Gnathonemus petersii*, thereby confirming the high variability of karyotype formulae within the Mormyridae. We infer that the differences in chromosome number and formula of *Campylomormyrus* relative to other mormyrids may be caused by Robertsonian fusion and pericentric inversion. In addition to the karyotype description and classification of *Campylomormyrus*, a ChromEvol analysis was used to determine the ancestral haploid chromosome number of osteoglossiform taxa. Our results indicate a relatively conservative haploid chromosome number of n = 24 for the most recent common ancestor of Osteoglossiformes and for most of the internal nodes of osteoglossiform phylogeny. Hence, we presume that the high chromosome variability evolved recently on multiple independent occasions. Furthermore, we suggest that the most likely ancestral chromosome number of Mormyridae is either n = 24 or n = 25. To the best of our knowledge this is the first attempt to determine and classify the karyotype of the weakly electric fish genus *Campylomormyrus* and to analyze chromosomal evolution within the Osteoglossiformes based on Maximum Likelihood and Bayesian Inference analyses.

1. Introduction

Campylomormyrus, a genus of African weakly electric fish, has been intensively studied with regard to evolution, electric discharge, electric organ anatomy, and behavior (Feulner et al., 2007, 2009; Lamanna et al., 2015, 2016; Paul et al., 2015). The genus belongs to the ancient teleost order Osteoglossiformes. An understanding of phylogenetic relationships within Osteoglossiformes has been hampered by the evolutionary history of this order, its morphological heterogeneity, and its relatively low species diversity (in comparison to its sister taxon, the Elopocephala; Nelson, 1969; Bonde, 1996; Li et al., 1997; Hilton, 2003). Relationships within Osteoglossiformes were first inferred from anatomical and morphological traits (Greenwood et al., 1966; Nelson, 1969), while more recent studies have utilized genetic markers (e.g., Lavoué and Sullivan, 2004; Lavoué et al., 2011). Several phylogenetic hypotheses have been advanced for the order, suggesting different relationships within Osteoglossiformes (Bonde, 1996; Kumazawa and Nishida, 2000; Hilton, 2003). In the most recent phylogenetic analyses (Lavoué and Sullivan, 2004; Lavoué et al., 2011), there is strong support for the hypothesis that the family

Abbreviations: a, acrocentric; CI, centromeric index; CRND, Constant Rate with No Duplication; FN, Fundamental Number; m, metacentric; n, haploid chromosome number; p, short chromosome arm; q, long chromosome arm; PP, posterior probability; sm, submetacentric; 2n, diploid chromosome number.

* Corresponding author.
E-mail addresses: jcanitz@uni-potsdam.de (J. Canitz), frank.kirschbaum@staff.hu-berlin.de (F. Kirschbaum), tiedeman@uni-potsdam.de (R. Tiedemann).

http://dx.doi.org/10.1016/j.jphysparis.2017.01.002
0928-4257/

Pantodonidae is the sister group of the two osteoglossiform suborders Osteoglossoidei (Osteoglossidae) and Notopteroidei, the latter consisting of the families Notopteridae, Mormyridae, and Gymnarchidae. *Campylomormyrus*, a taxon that has recently diversified in an adaptive radiation (Feulner et al., 2007; Tiedemann et al., 2010), is part of the Mormyridae, the family with by far the highest species diversity (~180 species), while all other families are relatively species-poor (ranging from one species in Pantodonidae and Gymnarchidae to ten species in Notopteridae).

To date, 20 species from 16 genera of osteoglossiform fishes are described cytogenetically (Uyeno, 1973; Ozouf-Costaz et al., 2015; Table 1), but the karyotype of *Campylomormyrus* has not hitherto been described. Furthermore, previous studies of osteoglossiform karyotypes have mainly focused on chromosome morphology and karyotype formulae, but do not specifically address karyotype evolution (Suzuki et al., 1982; Marques et al., 2006; Krysanov and Golubtsov, 2014). Uyeno (1973) describes the chromosome formulae of nine osteoglossiform fishes, and hypothesizes a diploid chromosome number of 2n = 48 for the most recent common ancestor of Osteoglossiformes. This is supported by later studies in which the ancestral chromosome number of Teleostei is proposed to be 2n = 48–50 (Jaillon et al., 2004; Mank and Avise, 2006; Kohn et al., 2006; Nakatani et al., 2007). Here, we describe the first karyotype of a *Campylomormyrus* species (*C. compressirostris*) and compare it to the karyotype information from other Mormyrids as well as non-electrogenic osteoglossiforms. Additionally, we estimate the most likely ancestral chromosome number of Osteoglossiformes using Maximum Likelihood and Bayesian Inference methods to discuss patterns of chromosome number evolution. As this analysis requires a phylogenetic tree as input, we reconstructed a phylogeny of those Osteoglossiforms for which karyotypes are available, based on mitochondrial genomes.

2. Material and methods

2.1. Laboratory procedures

Chromosome preparations were obtained from embryonic tissue of *Campylomormyrus compressirostris*. The samples were provided by the department of Biology and Ecology of Fishes (Prof. Kirschbaum, HU Berlin) and the protocol was adapted from Shao et al. (2010) and Karami et al. (2015). First, freshly hatched fish embryos (less than 24 hours old) were exposed to colchicine solution. After incubation, they were put on ice until vital signals were no longer detected. The yolk sac and lipid membrane were removed. The remaining tissue was incubated in a hypotonic solution for 30 min. To yield the best chromosome spreads, the preparation parameters were altered as follows: Two colchicine concentrations (0.02% or 0.05%) were tested for two different incubation times (3 h or 5 h) with two alternative hypotonic solutions (0.4% KCL or ddH_2O). In total, this resulted in eight different preparations. Henceforth, each preparation was treated equally. First, an incubation in Carnoy's solution (ethanol:glacial acetic acid; 3:1) for 20 min was performed to ensure the fixation of cell components. This step was repeated twice. The fixed embryonic tissue was then ground with a pestle in 50% acetic acid, and 10 µl of the cell suspension was pipetted onto a clean slide. Subsequently, the steps as described in Karami et al. (2015) were performed. To visualize the chromosomes, each slide was stained either with Giemsa (G-Banding) or with DAPI. For G-Banding, the slides were first swivelled briefly in a mixture of 2.5%Trypsin and 0.9% NaCl solution, washed twice with 0.9% NaCl solution, and incubated in 6% Giemsa staining for 12 min. Afterwards, the slides were washed with Gurr's buffer and conserved using DPX. Regarding the DAPI staining, slides were covered with 0.001% DAPI solution and incubated for 15 min in the dark, followed by a washing step with phosphate buffer and the conservation with DPX. The slides were stored at 4 °C in the dark until they were viewed. All slides were scanned manually under 1000× magnification (oil immersion) using a Leica DM4000 B microscope. Chromosome spreads were photographed with a Leica DFC480 digital camera and saved as TIFF-files.

Among all treatments, the highest number of chromosome spreads per preparation was observed on slides treated with higher colchicine concentration (0.05%), an incubation of 5 h, and KCl as hypotonic solution. Chromosomes generally clustered together so that a complete chromosome set could be expected. In single cell preparations, we counted diploid chromosome numbers ranging from 40 to 48. For all other treatments, most cells were either not bloated so that chromosomes could not be distinguished or chromosome spreads were not visible. DAPI staining worked

Table 1
Available karyotypes in the teleost order Osteoglossiformes.

Family	Species	2n	Karyotype formula	Reference
Osteoglossidae	*Arapaima gigas*	56	4m, 12sm, 40a	Hinegardner and Rosen (1972)
			28m, 28a	Marques et al. (2006)
	Heterotis niloticus	40	26m, 10sm, 4a	Hirata and Urushido (2000)
	Osteoglossum bicirrhosum	56	1sm, 55a	Uyeno (1973)
			56a	Suzuki et al. (1982)
	Osteoglossum ferreirai	54	2m, 4sm, 48a	Suzuki et al. (1982)
	Scleropages formosus	50	4m, 46a	Hirata and Urushido (2000)
	Scleropages jardini	48	16m, 6sm, 26a	Hirata and Urushido (2000)
	Scleropages leichardti	44	16m, 8sm, 20a	Hirata and Urushido (2000)
Pantodontidae	*Pantodon buchholzi*	48	12m, 12sm, 24a	Uyeno (1973)
Mormyridae	*Gnathonemus petersii*	48	10m, 6sm, 32a	Uyeno (1973)
			18m, 2sm, 28a	Ozouf-Costaz et al. (2015)
	Marcusenius brachistius	48	1m, 4sm, 43a	Uyeno (1973)
	Marcusenius moori	50	4sm, 46a	Ozouf-Costaz et al. (2015)
	Ivindomyrus opdenboschi	50	10m, 2sm, 38a	Ozouf-Costaz et al. (2015)
	Brienomyrus sp.	50	2m, 6sm, 42a	Ozouf-Costaz et al. (2015)
	Stomatorhinus wallkeri	50	2sm, 48a	Ozouf-Costaz et al. (2015)
	Petrocephalus microphtalamus	50	2sm, 48a	Ozouf-Costaz et al. (2015)
	Pollimyrus cf. nigricans	40	2m, 38a	Krysanov and Golubtsov (2014)
	Campylomormyrus compressirostris	48	4sm, 26m, 18a	This study
Notopteridae	*Chitala chitala*	42	42a	Uyeno (1973)
	Notopterus notopterus	42	42a	Rishi and Singh (1983)
	Papycranus afer	34	4sm, 30a	Uyeno (1973)
	Xenomystus nigri	42	42a	Uyeno (1973)

successfully in all treatments. The Giemsa staining was partially successful, as it allowed for chromosome counting. It was however not possible to unambiguously reveal the G-Banding pattern, presumably because chromosomes were very small. Hence, homologous chromosomes could not be assigned.

2.2. Chromosome measurements

All photographed spreads were perused manually. If the magnification was insufficient or the coloration too dark, the filter tools of the software GIMP 2.8.10 (www.gimp.org, 1997–2014) were used to improve the quality. Among all images, the clearest chromosome spread (i.e., the spread where chromosomes appeared most separated) was chosen to extract individual chromosomes for a final karyotype alignment graphic. Chromosome sizes and arm length were also measured to calculate the centromeric index (CI) and the q/p quotient (q = long chromosome arm, p = short chromosome arm) using the image processing software ImageJ 1.50i. (Wayne Rasband, National Institute of Health, USA). Based on these indicators, chromosome number and karyotype formulae were specified according to the nomenclature of Levan et al. (1964). The final chromosome alignment graphic was composed using GIMP 2.8.10.

2.3. Phylogenetic reconstruction and inference of chromosomal evolution

To analyze patterns of chromosome number evolution, a phylogenetic tree was computed for 17 genera of Osteoglossiformes for which karyotype information was available, including *Campylomormyrus* studied here (Table 1). Mitochondrial genome sequences, downloaded from NCBI (Hrbek and Farias, 2008; Inoue et al., 2001, 2009; Lamanna et al., 2014; Lavoué et al., 2012; Yue et al., 2006; Supplementary Table 1), were used to generate the alignment for a Maximum Likelihood reconstruction (ML). The alignment was performed using the ClustalW algorithm, implemented in Geneious R8.0.5, setting the parameters gap opening and gap extension to 15 and 6, respectively. The alignment was cut to a total length of 15,744 bp. The program RAxML v8.2.X (Stamatakis, 2014) was applied for the construction of a phylogenetic ML tree. Prior to analysis, JModelTest 3.7 (Guindon and Gascuel, 2003; Darriba et al., 2012) was used to calculate the best fitting evolutionary model for the given data set (General Time Reversible with gamma distributed rate; GTR-Gamma), which was subsequently considered during the ML approach. The genus *Pantodon* served as the outgroup to Osteoglossidae + Notopteridae + Mormyridae, as its position as sister group to all remaining osteoglossiform lineages is phylogenetically well supported (Lavoué and Sullivan, 2004; Lavoué et al., 2011). The bootstrap parameter was set to 1000 replicates to evaluate the support of the inferred clades.

For determination of ancestral chromosome number, the software ChromEvol 2.0 (Mayrose et al., 2010; Glick and Mayrose, 2014) was applied, which is based on Maximum Likelihood and Bayesian phylogenetic inferences. ChromEvol provided eight models of chromosome-number changes, taking into account the following six parameters: polyploidization (chromosome number duplication) with rate ρ, demi-duplication (fusion of different ploidy gametes) with rate μ, and dysploidization (single chromosome increase (rate λ) or decrease (rate δ)). Two further parameters (λ_1 and δ_1) were rates of chromosome increase or decrease, assuming a linear relationship between rates and the actual chromosome number. The eight models were divided in two sets – one constant (using ρ, μ, λ, and δ only) and one linear (using ρ, μ, λ, δ, λ_1, and δ_1). Each set included three models with a different parameter combination and one null model that assumed no duplication events. First, all models were run to test for the model fitting best to our data. For the best fitting model, ChromEvol analysis was then performed with optimized parameters identified in the initial run and no fixed haploid chromosome number (n). The optimized model was repeated five times with different fixed basal chromosome numbers (n = 24; n = 25; n = 26; n = 27; n = 28) to compare the expected changes of chromosome transitions along each branch and the ancestral chromosome numbers at internal nodes under different conditions. Results of the runs were visualized using the ChromEvol-0.9.1 package in R (Dr. Natalie Cusimano; LMU München, 2013).

3. Results

3.1. Karyotyping

Thirty-three metaphases were photographed, and the diploid chromosome number for *Campylomormyrus compressirostris* has been determined to be 2n = 48. Chromosome numbers lower than 2 = 48, encountered in some preparations, have been attributed to losses during chromosome preparation and/or superimposition of chromosomes. Two images of chromosome spreads (Fig. 1A and 1B) have been utilized to establish the karyotype using the following nomenclature and abbreviations: m = metacentric, sm = submetacentric, and a = acrocentric (including the subtelocentric and telocentric chromosomes). Based on the calculation of chromosomal indicators (CI and q/p), the following karyotype formula has been identified: 4sm + 26 m + 18a chromosomes with a fundamental number (arm number) of FN = 78 (Fig. 1C). The average size of the four submetacentric chromosomes is 7.44 μm, ranging from 7.03 μm to 7.64 μm. For metacentric chromosomes, the mean total length is 5.05 μm (range from 3.46 μm to 6.69 μm) and for acrocentric chromosomes the mean is 4.39 μm (range from 3.71 μm to 5.61 μm) (Supplementary Table 4). Additionally, there are two conspicuous metacentric chromosomes which appear to have attached satellites (Fig. 1C, Chromosomes 29 and 30). Fig. 1A has been used to align the final karyotype (Fig. 1C).

3.2. Phylogenetic reconstruction

Our ML analysis resulted in a well-supported phylogenetic tree with bootstrap support of 95–100% for all nodes (Supplementary, Fig. 1). The topology of the families (Osteoglossidae, Notopteridae and Mormyridae) is consistent to the phylogeny of Lavoué and Sullivan (2004; Topology O). Within the Mormyridae, Lavoué and Sullivan (2004) include only two species, so we compare this portion of our tree to the phylogeny of Ozouf-Costaz et al. (2015). Both phylogenies were consistent, except for the position of *Stomatorhinus* and *Ivindomyrus*. This inconsistency could be caused by the different set of molecular markers used and/or the differences in sequence length (15,744 bp of the mitochondrial genome in our study vs. 1080 bp of the Cytochrome b gene (Ozouf-Costaz et al., 2015)).

Except for the genus *Pantodon* (used as outgroup), all osteoglossiform fishes are separated in two major clades corresponding to the suborders Osteoglossoidei and Notopteroidei (Fig. 2). Osteoglossoidei consists of two sister clades (*Osteoglossum + Scleropages; Heterotis + Arapaima*). The second major clade (Notopteroidei) comprises the families Notopteridae and Mormyridae (Fig. 2).

3.3. Patterns of chromosome number evolution

The analysis in ChromEvol 2.0 inferred that the best fitting model to our data was the Constant-Rate-With-No-Duplication

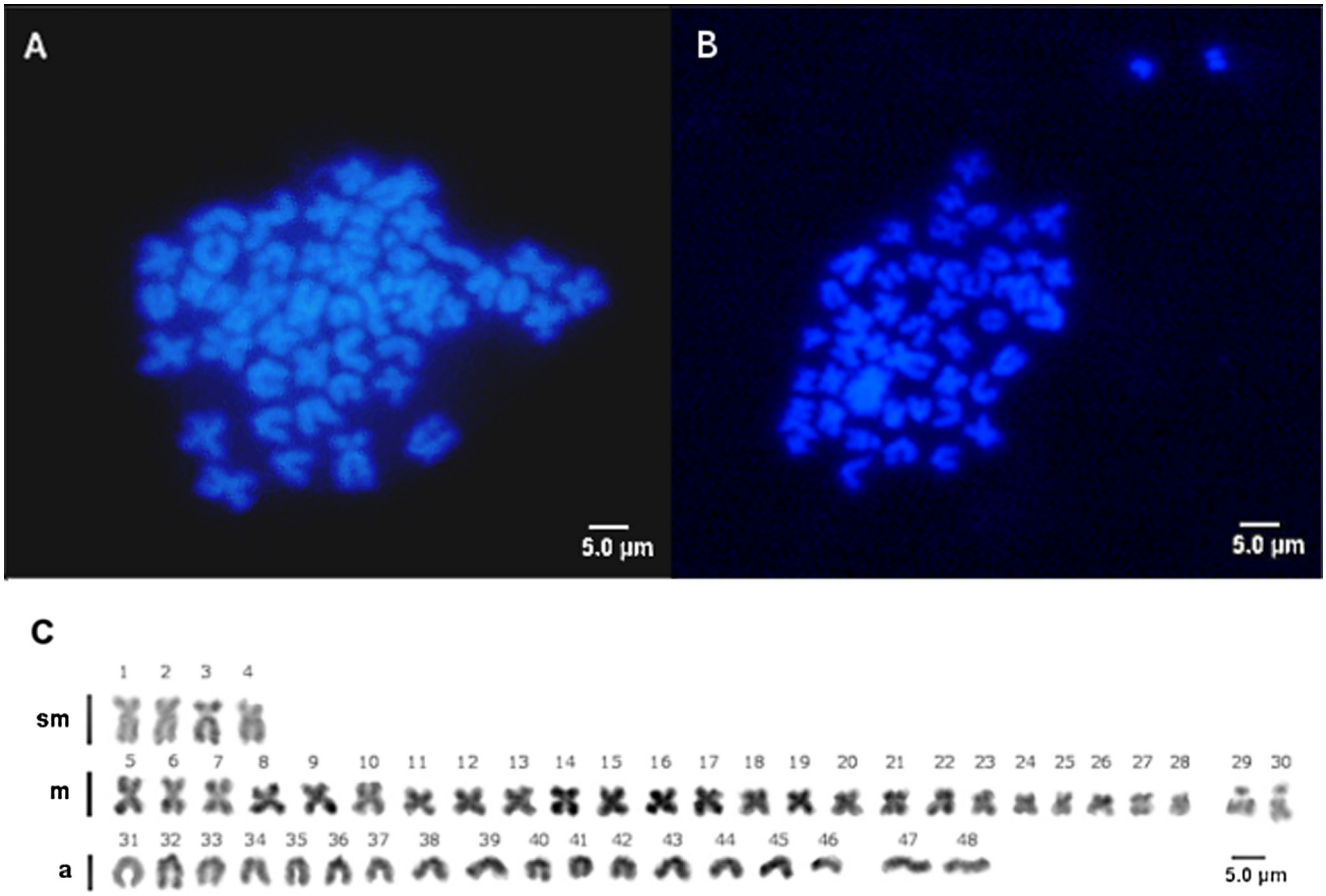


Fig. 1. Chromosomes of *Campylomormyrus compressirostris*. (A and B) two different chromosome spreads stained with DAPI; (A) This image was used to extract the single chromosomes and create the karyotype; (C) Karyotype – 48 chromosomes classified as submetacentric (sm), metacentric (m), and acrocentric (a) chromosomes, aligned by decreasing size and numbered.

(CRND), based on the Akaike information criterion (AIC; Supplementary Table 2). A second run with fixed CRND parameters inferred an ancestral haploid number of n = 27 with the highest posterior probability (PP) of 0.23. As the PP values of adjacent chromosome numbers (n = 24–26 and n = 28) are close to the probability of n = 27 (Table 2), the CRND model has been re-run independently for each haploid number (n = 24–28). The lowest computed AIC of the five runs is also found for n = 27 (AIC = 81.0429, bold in Table 2). As the haploid chromosome number of n = 24 is hypothesized to be the ancestral number of Teleostei, the results of two models (Run1: n = 24, the hypothetical ancestral number; and Run2: n = 27, the most likely number inferred from the data) were compared (Supplementary Table 3). Chromosome numbers of internal nodes and chromosome transitions along branches are visualized for both runs (main text Fig. 2 for n = 24 (Run1); Supplementary Fig. 2 for n = 27 (Run2)).

For Run1, a total of 88.2 chromosome transition events with an expectation above 0.5 have been inferred, including 39.5 chromosome gains and 48.7 chromosome losses. A similar number of total transition events is calculated for Run2 (86.2 events), but the difference between chromosome gains and losses is higher (gains: 27.7; losses: 58.5). No runs infer duplication or demi-duplication events. The rate of constant losses and gains in Run1 are $\delta = 20.430$ and $\lambda = 17.866$, whereas in Run2 the rates are $\delta = 33.47$ and $\lambda = 14.085$. Additionally, there are differences in inferred chromosome numbers among the two runs at all internal nodes. While the haploid chromosome number of the most recent common ancestor of Osteoglossoidei and Notopteroidei is inferred to be n = 24 in Run1, it is inferred as n = 26 (PP = 0.29) in Run2. Within the family Osteoglossidae, the chromosome number differs among the nodes in Run1 (n = 24 for N3 and N4; n = 25 for N5). This is not the case in Run2, where a chromosome number of n = 26 has been calculated at each node. The chromosome number of the most recent common ancestor of Notopteroidei is estimated as n = 23 (PP = 0.29) in Run1 and n = 25 (PP = 0.31) in Run2. In the family Notopteridae, both runs infer that the haploid chromosome number has decreased. The ancestral number of Notopteridae is predicted to be n = 22 (PP = 0.27) in Run1 and n = 24 (PP = 0.28) in Run2. For Mormyridae, Run1 inferred a haploid chromosome set of n = 24 for all internal nodes. Here, Run2 inferred a chromosome number of n = 25 except for the most recent clade including *Campylomormyrus* and *Gnathonemus* for which n = 24 (PP = 0.66) was inferred. The highest posterior probability at each node indicates the most likely haploid chromosome number of this node, but the differences between this and the second highest PP values are slight (Run1: $\Delta PP_{mean} = 0.10$; Run2: $\Delta PP_{mean} = 0.12$, Supplementary Table 3). For several nodes, one run infers the highest PP for a chromosome number which yields the second highest PP in the other run and vice versa (see "differences in n" in Supplementary Table 3).

4. Discussion

4.1. Karyotypes of Mormyridae

Here, we present the karyotype of the mormyrid fish *C. compressirostris* for the first time. We show that it has a diploid

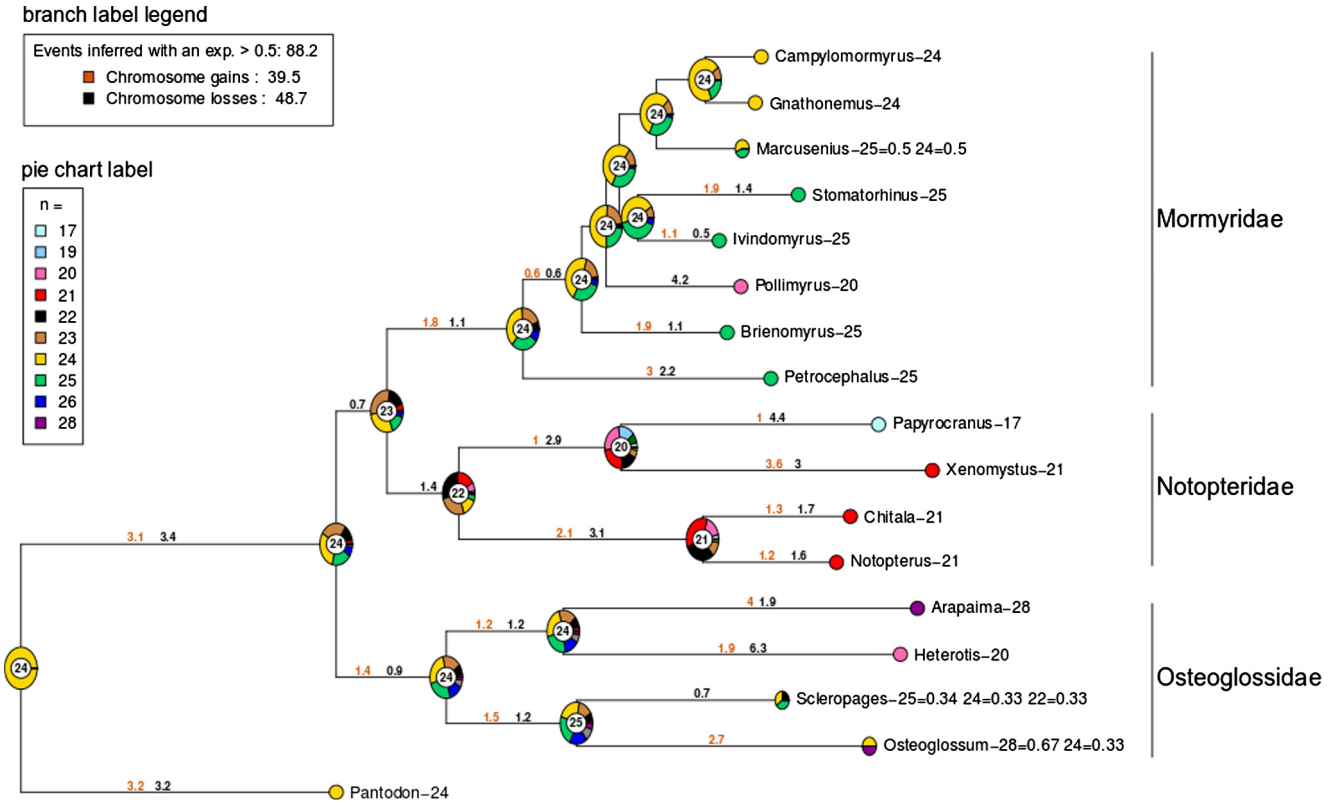


Fig. 2. Chromosome evolution along a phylogenetic tree of Osteoglossiformes based on the CRND model with ancestral haploid chromosome number n = 24. The phylogenetic tree is based on a ML analysis, and tree tips indicate the genera of Osteoglossiformes and their currently published haploid chromosome numbers (see Table 1; if karyotypes for more than one species of a genus are available, relative frequencies of different n-values are provided). Nodes are labelled by pie charts indicating the proportional probabilities of the different haploid chromosomes sets. Chromosome numbers with highest posterior probability (PP) at each node are centred in the pie chart. Color code for both is given in the pie chart legend. Branch labels indicate chromosome transitions (gains in brown, losses in black) along each branch. A total of 88.2 transition events with an expectation over 0.5 are predicted.

Table 2
AIC values of runs with different fixed haploid chromosome numbers (n) and their PP values of the initial CRND model (inferred most likely n in bold).

n	AIC	PP of initial optimized CRND run
24	81.2916	0.1028
25	81.1427	0.1695
26	81.0771	0.2172
27	**81.0429**	**0.2267**
28	81.0734	0.2136

chromosome number of 2n = 48 (Fig. 1); this inferred chromosome number is identical to the diploid number of 2n = 48 in the genus *Gnathonemus* (Uyeno, 1973; Ozouf-Costaz et al., 2015), its sister taxon. This is consistent with their close relatedness within the family Mormyridae, as evidenced by phylogenetic analyses (this study; Ozouf-Costaz et al., 2015).

Seven karyotypes of mormyrid genera have been previously described (Ozouf-Costaz et al., 2015; Krysanov and Golubtsov, 2014; Uyeno, 1973). Based on these data, our new *Campylomormyrus* karyotype, and our ChromEvol analysis, we have inferred an ancestral diploid chromosome number for Mormyridae of 2n = 48–50. Chromosome numbers in this range are indeed found in most of the extant mormyrids (Table 1). These diploid numbers are also considered ancestral in Teleostei (Arai, 2011; Mank and Avise, 2006; Sember et al., 2015). While chromosome numbers were inferred to be relatively constant in the evolution of most mormyrid lineages, karyotype formulae are highly variable across taxa. Specifically, a high variance in the number of metacentric and acrocentric chromosomes has been observed among the taxa of Mormyridae. Comparing karyotype formulae to the mormyrid phylogeny did not reveal any obvious congruence. Indeed, the changes in karyotype formulae of single taxa appear to evolve independently. In a previous study, Ozouf-Costaz et al. (2015) investigated chromosomal evolution in Mormyridae. They conclude that chromosome formulae have obviously diverged rapidly and through pericentric inversions. In this type of chromosome rearrangements, one break is on each of the two chromosome arms, such that the inverted chromosome segment includes the centromere. Another chromosomal rearrangement, called Robertsonian fusion or centric fusion, can also be responsible for changes in chromosome number. Here, a fragment of one chromosome is attached to a non-homologous chromosome. This type of rearrangement often leads to a higher number of sub-/metacentric and a lower number of acrocentric chromosomes.

Campylomormyrus and *Gnathonemus* (sm/m = 30; sm/m = 20) possess a higher number of sub-/metacentric chromosomes relative to their sister taxon, *Marcusenius* (sm/m = 4). This suggests that several pericentric inversions occurred, leading to the karyotype of *Campylomormyrus*.

It should be considered that only one species has been karyotyped for most genera (cf. Table 1). Interspecific and intraspecific variation was not evaluated in this study, but several previous studies have suggested that chromosome number variation occurs among closely related species and even among conspecific populations (Marques et al., 2006). In the genus *Marcusenius*, two species were found to differ in both chromosome number and karyotype formula (Uyeno, 1973; Ozouf-Costaz et al., 2015). In the species *Gnathonemus petersii*, two different karyotypes are described by different authors; both indicate a diploid chromosome number of 2n = 48, but they differ in their karyotype formulae

(2sm-18m-28a, Ozouf-Costaz et al., 2015 and 6sm-10m-32a, Uyeno, 1973). Unfortunately, the authors provide no detailed information about the geographic origin of their specimens, and no genotype information is available for the samples analyzed by Uyeno (1973). It is therefore not possible to conclude whether *Gnathonemus petersii* exhibits two karyotypes within a single interbreeding species or whether these different karyotypes may be indicative of cryptic species. Intraspecific variation in chromosome number has been also reported in the distantly related Gymnotiform weakly electric fish (e.g., in the genus *Gymnotus*). Here, fusion-fission and pericentric inversion are discussed as processes leading to a reduction in chromosome number (Milhomen et al., 2008).

Pericentric inversion has been considered an important post-zygotic reproduction isolation mechanism, and it has been suggested that variation of karyotypes could contribute to reproductive isolation and ultimately speciation (King, 1993; Milhomen et al., 2008). Indeed, cases of rapid speciation through chromosomal rearrangements have been reported in mammals (e.g., Malagasy lemurs; Ravaoarimanana et al., 2004). Given the recent radiation and sympatric evolution in the genus *Campylomormyrus* (Feulner et al., 2007; Tiedemann et al., 2010), it would be interesting to analyze karyotypes of further closely related species within this genus in the future.

In summary, the chromosome number of *C. compressirostris* is similar to that of several other species of Mormyridae. Although it is identical to the most closely related genus (*Gnathonemus*), the karyotype formula is different. Pericentric inversion and Robertsonian fusion are proposed to be the most common causes for changes in karyotypes among teleostean fishes (Silva et al., 2014; Sember et al., 2015). This may have also occurred during the evolution of Mormyridae and may have led to the high number of metacentric chromosomes in *Campylomormyrus* (cf. Table 1).

4.2. Chromosomal evolution in Osteoglossiformes

Our Maximum Likelihood inference indicates an ancestral haploid number of n = 27 chromosomes. This contrasts with a previous hypothesis of Uyeno (1973), who suggested a diploid number of 2n = 48 (n = 24). Our analysis has been applied for both a fixed ancestral chromosome number of n = 24 (the current hypothetical value; Run1) and n = 27 (the value with the highest likelihood; Run2). In both runs, the numbers of total events occurring from the root to the tips are similar, indicating that during the evolution of Osteoglossiformes more chromosome losses than gains occurred. However, there are also differences between the two analyses. While the results of the first run (fixed ancestral n = 24) shows a stable haploid number of n = 24 at almost all internal nodes, the run with a fixed ancestral number of n = 27 indicates a general decrease in chromosome number along the phylogeny. With regard to the number of necessary changes in chromosome number ($\Delta n = 7$ for n = 24; $\Delta n = 9$ for n = 27), the results of Run1 (n = 24) are more parsimonious than those of Run2 (n = 27). We therefore consider it more probable that the most recent common ancestor of Osteoglossiformes had a haploid chromosome number of n = 24. This is also supported by the chromosome number of the most basal taxon in osteoglossiform fishes presenting the same haploid number (*Pantodon*, n = 24). Furthermore, Osteoglossiformes is one of the earliest branching clades in the phylogeny of Teleostei, and the chromosome number of the teleostean ancestor is also assumed to be n = 24–25. This would indicate a relative constancy in chromosome number. The assumption of an ancestral chromosome number of n = 24 in osteoglossiform fishes and its relative stability during its evolution does, however, not apply to all of the taxa. Specifically within Notopteridae and Osteoglossidae, inferred gains and losses of chromosomes vary considerably relative to their respective common ancestor, such that chromosome numbers appear more dynamic in these groups. Gymnotiform weakly electric fish also show a high variation in chromosome numbers within and among families. There are families with a more conserved chromosome number (Rhamphichthyidae), while others are highly variable in this measure (Hypopomidae). Karyotype information can be beneficial to confirm an existing phylogenetic hypothesis or to resolve an uncertain relationship between taxa, as shown for the Gymnotiform tribe Steatogenini (Cardoso et al., 2011).

As a cautionary note, we have to point to the fact that our discussion of chromosome number evolution is mostly based on those chromosome numbers exhibiting the highest posterior probabilities. It must be stressed here that – for several internal nodes - the next higher or lower chromosome numbers were assigned almost as high posterior probabilities in the ChromEvol analysis (Supplementary, Table 3), rendering a conclusive inference of definite chromosome numbers difficult. At least within the family Mormyridae, it appears most likely that the ancestral chromosome number is either n = 24 or n = 25, as both runs reveal a high distance in posterior probabilities to the next highest values ($\Delta PP_{mean} = 0.3$).

From a methodological point of view, we argue that the Maximum Likelihood method applied here provides a framework for rigorous evaluation of chromosome number. Moreover, our analysis also reveals that the available karyotype data for Osteoglossiformes are still too scarce to infer distinct evolutionary scenarios with convincing probability. Hence, to verify our hypotheses about chromosome number evolution at the various taxonomic levels, karyotypes of more species are necessary. There are still genera which are not yet described cytogenetically (e.g., *Mormyrus* and *Paramormyrus*). A further inevitable limitation of our inference of ancestral character states is the small number of extant non-electrogenic osteoglossiforms. Given that osteoglossiforms comprise an ancient clade, a considerable part of the phylogenetic diversity may have been lost by extinction. Evidently, the inference of chromosomal evolution becomes more accurate if many extant related species can be assessed, as in more recent radiations (e.g., within the Mormyridae; cf. Suppl. Table 3, differences between runs 1 & 2).

To have a more complete picture of the chromosome evolution in Osteoglossiformes, such an extended analysis should include an assessment of karyotype variation both among closely related species and within species. Further assessments of chromosome numbers and morphology should also be complemented by analyses allowing for an exact identification of homologous chromosomes, such as Fluorescence in-situ hybridization (FISH) or C-banding. These methods would also provide a proof of hypothesized chromosomal rearrangements (as, e.g., in *Gymnotus carapo*; Milhomen et al., 2008).

In conclusion, the karyotype of the genus *Campylomormyrus* is similar to the already known chromosome sets of Mormyridae (cf. Table 1). For the most recent common ancestor of the family Mormyridae, a haploid chromosome number of n = 24–25 is inferred. We assume that pericentric inversion and Robertsonian fusion played a role in the chromosomal evolution of mormyrid taxa. Furthermore, our analyses is in congruence with the hypothesized ancestral chromosome number of n = 24 for the entire order Osteoglossiformes. Moreover, our study provides hypotheses about the most likely haploid chromosome numbers for ancestral nodes based on Maximum Likelihood and Bayesian Inference.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgments

We thank Khouloud El Arbarni and Linh Nguyen of the department of Biology and Ecology of Fishes (chair: Prof. Kirschbaum, HU Berlin) for feeding and reproducing the fish as well as taking care during the gestation period and providing embryonic tissue for chromosome preparation. We also thank Prof. Ralph Gräf of the Unit for Cell Biology at the University of Potsdam and his PhD student, Tatjana Peter, for their expertise and advices in microscopy and staining. We are very grateful to Dr. Alice Dennis for her native speaker's assessment of our manuscript. Furthermore, we acknowledge financial support by the University of Potsdam.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jphysparis.2017.01.002.

References

Arai, R., 2011. Fish Karyotypes - A Check List. Springer, Japan. http://dx.doi.org/10.1017/CBO9781107415324.004.

Bonde, N., 1996. Osteoglossids (Teleostei: Osteoglossomorpha) of the Mesozoic. Comments on their interrelationships. In: Arratia, G., Viohl, G. (Eds.), Mesozoic Fishes — Systematics and Paleoecology. Verlag Dr Friedrich Pfeil, München, Germany, pp. 273–284.

Cardoso, A.L., Pieczarka, J.C., Feldberg, E., Milhomen, S.S.R., Moreira-Almeida, T., dos Santos Silva, D., Corrêa da Silva, P., Nagamachi, C.Y., 2011. Chromosomal characterization of two species of the genus *Steatogenys* (Gymnotoformes: Rhamphichthyoidae: Steatogenini) from Amazon basin: sex chromosome and correlations with Gymnotifomres phylogeny. Rev. Fish Biol. Fisher. 21, 613–621.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. JModelTest 2: more models, new heuristics and parallel computing. Nat. Meth. 9, 772.

Feulner, P.G.D., Kirschbaum, F., Mamonekene, V., Ketmaier, V., Tiedemann, R., 2007. Adaptive radiation in African weakly electric fish (Teleostei: Mormyridae: *Campylomormyrus*): a combined molecular and morphological approach. J. Evol. Biol. 20, 403–414. http://dx.doi.org/10.1111/j.1420-9101.2006.01181.x.

Feulner, P.G.D., Plath, M., Engelmann, J., Kirschbaum, F., Tiedemann, R., 2009. Magic trait electric organ discharge (EOD) – dual function of electric signals promotes speciation in African electric fish. Commun. Integr. Biol. 2 (4), 329–331. http://dx.doi.org/10.4161/cib.2.4.8386.

Glick, L., Mayrose, I., 2014. ChromEvol: assessing the pattern of chromosome number evolution and the inference of polyploidy along a phylogeny. Mol. Biol. Evol. 31, 1914–1922. http://dx.doi.org/10.1093/molbev/msu122.

Greenwood, H.P., Rosen, D.E., Weitzman, S.H., Myers, G.S., 1966. Phyletic studies of teleostean fishes, with a provisional classification of living forms. Bull. Am. Museum Nat. History 131, 345–455.

Guindon, S., Gascuel, O., 2003. A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. Syst. Biol. 52, 696–704.

Hilton, E.J., 2003. Comparative osteology and phylogenetic systematics of fossil and living bony-tongue fishes (Actinopterygii, Teleostei, Osteoglossomorpha). Zool. J. Linnean Soc. 137, 1–100. http://dx.doi.org/10.1046/j.1096-3642.2003.00032.x.

Hinegardner, R., Rosen, D.E., 1972. Cellular DNA content and the evolution of teleostean fishes. Am. Natur. 106, 621–644.

Hirata, J., Urushido, T., 2000. Karyotypes and DNA contents in Osteoglossiformes fishes. Sci. Rep. Res. Inst. Evol. Biol. 9, 83–90 (In Japanese with English abstract).

Hrbek, T., Farias, I.P., 2008. The complete mitochondrial genome of the pirarucu (*Arapaima gigas*, Arapaimidae, Osteoglossiformes). Gen. Mol. Biol. 31, 293–302. http://dx.doi.org/10.1590/S1415-47572008000200024.

Inoue, J.G., Miya, M., Tsukamoto, K., Nishida, M., 2001. A mitogenomic perspective on the basal teleostean phylogeny: resolving higher-level relationships with longer DNA sequences. Mol. Phylogenet. Evol. 20, 275–285. http://dx.doi.org/10.1006/mpev.2001.0970.

Inoue, J.G., Kumazawa, Y., Miya, M., Nishida, M., 2009. The historical biogeography of the freshwater knifefishes using mitogenomic approaches: a Mesozoic origin of the Asian notopterids (Actinopterygii: Osteoglossomorpha). Mol. Phylogenet. Evol. 51, 486–499. http://dx.doi.org/10.1016/j.ympev.2009.01.020.

Jaillon, O., Aury, J.-M., Brunet, F., Petit, J.-L., Stange-Thomann, N., Mauceli, E., Bouneau, L., Fischer, C., Ozouf-Costaz, C., Bernot, A., Nicaud, S., Jaffe, D., Fisher, S., Lutfalla, G., Dossat, C., Segurens, B., Dasilva, C., Salanoubat, M., Levy, M., Boudet, N., Castellano, S., Anthouard, V., Jubin, C., Castelli, V., Katinka, M., Vacherie, B., Biémont, C., Skalli, Z., Cattolico, L., Poulain, J., De Berardinis, V., Cruaud, C., Duprat, S., Brottier, P., Coutanceau, J.-P., Gouzy, J., Parra, G., Lardier, G., Chapple, C., McKernan, K.J., McEwan, P., Bosak, S., Kellis, M., Volff, J.-N., Guigó, R., Zody, M. C., Mesirov, J., Lindblad-Toh, K., Birren, B., Nusbaum, C., Kahn, D., Robinson-Rechavi, M., Laudet, V., Schachter, V., Quétier, F., Saurin, W., Scarpelli, C., Wincker, P., Lander, E.S., Weissenbach, J., Roest Crollius, H., 2004. Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. Nature 431, 946–957. http://dx.doi.org/10.1038/nature03025.

Karami, A., Eghtesadi Araghi, P., Syed, M.A., Wilson, S.P., 2015. Chromosome preparation in fish: effects of fish species and larval age. Int. Aquat. Res. 7, 201–210. http://dx.doi.org/10.1007/s40071-015-0104-z.

King, M., 1993. Species Evolution: The Role of Chromosome Change. Cambridge University Press.

Kohn, M., Högel, J., Vogel, W., Minich, P., Kehrer-Sawatzki, H., Graves, J.A.M., Hameister, H., 2006. Reconstruction of a 450-My-old ancestral vertebrate protokaryotype. Trends Genet. 22, 203–210. http://dx.doi.org/10.1016/j.tig.2006.02.008.

Krysanov, E.Y., Golubtsov, A.S., 2014. Karyotypes of four fish species from the Nile and Omo-Turkana basins in Ethiopia. J. Ichthyol. 54, 889–892. http://dx.doi.org/10.1134/S0032945214100087.

Kumazawa, Y., Nishida, M., 2000. Molecular phylogeny of osteoglossoids: a new model for Gondwanian origin and plate tectonic transportation of the Asian arowana. Mol. Biol. Evol. 17, 1869–1878. http://dx.doi.org/10.1093/oxfordjournals.molbev.a026288.

Lamanna, F., Kirschbaum, F., Tiedemann, R., 2014. De novo assembly and characterization of the skeletal muscle and electric organ transcriptomes of the African weakly electric fish *Campylomormyrus compressirostris* (Mormyridae, Teleostei). Mol. Ecol. Resour. 14, 1222–1230. http://dx.doi.org/10.1111/1755-0998.12260.

Lamanna, F., Kirschbaum, F., Waurick, I., Dieterich, C., Tiedemann, R., 2015. Cross-tissue and cross-species analysis of gene expression in skeletal muscle and electric organ of African weakly electric fish (Teleostei; Mormyridae). BMC Genom. 16, 1–17.

Lamanna, F., Kirschbaum, F., Ernst, A.R.R., Feulner, P.G.D., Mamonekene, V., Paul, C., Tiedemann, R., 2016. Species delimination and phylogenetic relationships in a genus of African weakly-electric fishes (Osteoglossiformes, Mormyridae, *Campylomormyrus*). Mol. Phylogenet. Evol. 101, 8–18. http://dx.doi.org/10.1016/j.ympev.2016.04.035.

Lavoué, S., Sullivan, J.P., 2004. Simultaneous analysis of five molecular markers provides a well-supported phylogenetic hypothesis for the living bony-tongue fishes (Osteoglossomorpha: Teleostei). Mol. Phylogenet. Evol. 33, 171–185. http://dx.doi.org/10.1016/j.ympev.2004.04.021.

Lavoué, S., Miya, M., Arnegard, M.E., McIntyre, P.B., Mamonekene, V., Nishida, M., 2011. Remarkable morphological stasis in an extant vertebrate despite tens of millions of years of divergence. Proc. Roy. Soc. Lond. B: Biol. Sci. 278, 1003–1008. http://dx.doi.org/10.1098/rspb.2010.1639.

Lavoué, S., Miya, M., Arnegard, M.E., Sullivan, J.P., Hopkins, C.D., Nishida, M., 2012. Comparable ages for independent origins of electrogenesis in African and South American weakly electric fishes. PLoS ONE 7, 1–18. http://dx.doi.org/10.1371/journal.pone.0036287.

Levan, A., Fredga, K., Sandberg, A.A., 1964. Nomenclature for centromeric position at chromosomes. Hereditas, 201–220. http://dx.doi.org/10.1111/j.1601-5223.1964.tb01953.x.

Li, G.-Q., Wilson, M.V.H., Grande, L., 1997. Review of *Eohiodon* (Teleostei: Osteoglossomorpha) from western North America, with a phylogenetic reassessment of Hiodontidae. J. Paleontol. 71, 1109–1124.

Mank, J.E., Avise, J.C., 2006. Phylogenetic conservation of chromosome numbers in Actinopterygiian fishes. Genetica 127, 321–327. http://dx.doi.org/10.1007/s10709-005-5248-0.

Marques, D.K., Venere, P.C., Galetti, P.M., 2006. Chromosomal characterization of the bonytongue *Arapaima gigas* (Osteoglossiformes: Arapaimidae). Neotrop. Ichthyol. 4, 215–218. http://dx.doi.org/10.1590/S1679-62252006000200007.

Mayrose, I., Barker, M.S., Otto, S.P., 2010. Probabilistic models of chromosome number evolution and the inference of polyploidy. Syst. Biol. 59, 132–144. http://dx.doi.org/10.1093/sysbio/syp083.

Milhomen, S.S.R., Pieczarka, J.C., Crapton, W.G.R., Silva, D.S., De Souza, A.C.P., Carvalho Jr, J.R., Nagamachi, C.Y., 2008. Chromosomal evidence for a putative cryptic species in the *Gymnotus carapo* species-complex (Gymnotiformes, Gymnotidae). BMC Genet. 9, 1–10.

Nakatani, Y., Takeda, H., Kohara, Y., Morishita, S., 2007. Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates. Genome Res. 17, 1254–1265. http://dx.doi.org/10.1101/gr.6316407.

Nelson, G.J., 1969. Infraorbital bones and their bearing on the phylogeny and geography of osteoglossomorph fishes. Am. Museum Novit. 2394, 1–37.

Ozouf-Costaz, C., Coutanceau, J., Bonillo, C., Guidi-rontani, C., Paugy, D., 2015. First insights into karyotype evolution within the family Mormyridae. Cybium 39, 227–236.

Paul, C., Mamonekene, V., Vater, M., Feulner, P.G.D., Engelmann, J., Tiedemann, R., Kirschbaum, F., 2015. Comparative histology of the adult electric organ among four species of the genus *Campylomormyrus* (Teleostei: Mormyridae). J. Comp. Physiol. A 201, 357–374. http://dx.doi.org/10.1007/s00359-015-0995-6.

Ravaoarimanana, I.B., Tiedemann, R., Montagnon, D., Rumpler, Y., 2004. Molecular and cytogenetic evidence for cryptic speciation within a rare endemic Malagasy lemur, the Northern Sportive Lemur (*Lepilemur septentrionalis*). Mol. Phylogenet. Evol. 31, 440–448.

Rishi, K.K., Singh, J., 1983. Chromosomes of *Notopterus notopterus* (Pallas) (Notopteridae: Clupeiformes). Chrom. Inform. Serv. 34, 139–144.

Sember, A., Bohlen, J., Šlechtová, V., Altmanová, M., Symonová, R., Ráb, P., 2015. Karyotype differentiation in 19 species of river loach fishes (Nemacheilidae, Teleostei): extensive variability associated with rDNA and heterochromatin distribution and its phylogenetic and ecological interpretation. BMC Evol. Biol. 15, 251. http://dx.doi.org/10.1186/s12862-015-0532-9.

Shao, C.W., Wu, P.F., Wang, X.L., Tian, Y.S., Chen, S.L., 2010. Comparison of chromosome preparation methods for the different developmental stages of the half-smooth tongue sole, *Cynoglossus semilaevis*. Micron 41, 47–50. http://dx.doi.org/10.1016/j.micron.2009.08.002.

Silva, F.H.R., Pieczarka, J.C., Cardoso, A.L., da Silva, P.C., De Oliveira, J.A., Nagamachi, C.Y., 2014. Chromosomal diversity in three species of electric fish (Apteronotidae, Gymnotiformes) from the Amazon Basin. Genet. Mol. Biol. 37, 638–645. http://dx.doi.org/10.1590/S1415-47572014005000018.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 2014, 1–2. Open access link: <http://bioinformatics.oxfordjournals.org/content/early/2014/01/21/bioinformatics.btu033.abstract?keytype=ref&ijkey=VTEqgUJYCDcf0kP>.

Suzuki, A., Taki, Y., Urushido, T., 1982. Karyotypes of two Species of arowana, *Osteoglossum bicirrhosum* and *O. ferreirai*. Japanese J. Ichthyol. 29, 220–222.

Tiedemann, R., Feulner, P.G.D., Kirschbaum, F., 2010. Electric organ discharge divergence promotes ecological speciation in sympatrically occurring African weakly electric fish (*Campylomormyrus*). In: Glaubrecht, M. (Ed.), Evolution in Action – Adaptive Radiation, Speciation and the Origin of Biodiversity. Springer, Berlin, Heidelberg, pp. 307–321.

Uyeno, T., 1973. A comparative study of chromosomes in the teleostean fish order osteoglossiformes. Japanese J. Ichthyol. 20, 211–2017.

Yue, G.H., Liew, W.C., Orban, L., 2006. The complete mitochondrial genome of a basal teleost, the Asian arowana (*Scleropages formosus*, Osteoglossidae). BMC Genom. 7, 242. http://dx.doi.org/10.1186/1471-2164-7-242.

4 Article II

Transcriptome-wide single nucleotide polymorphisms related to electric organ discharge differentiation among African weakly electric fish species

Julia Canitz, Frank Kirschbaum, Ralph Tiedemann

Submitted to

PlosONE

PLOS ONE

Transcriptome-wide single nucleotide polymorphisms related to electric organ discharge differentiation among African weakly electric fish species

--Manuscript Draft--

Manuscript Number:	
Article Type:	Research Article
Full Title:	Transcriptome-wide single nucleotide polymorphisms related to electric organ discharge differentiation among African weakly electric fish species
Short Title:	Transcriptome-wide SNPs related to electric organ discharge differentiation in weakly electric fish
Corresponding Author:	Ralph Tiedemann University of Potsdam Potsdam, GERMANY
Keywords:	Campylomormyrus; SNP analysis; transcriptome; weakly-electric fish; electric organ discharge; candidate genes
Abstract:	African weakly electric fish of the mormyrid genus Campylomormyrus generate pulse-type electric organ discharges (EODs) for orientation and communication. Their pulse durations are species-specific and elongated EODs are a derived trait. Differential gene expression among tissue-specific transcriptomes across species with different pulses and point mutations in single ion channel genes indicate a relation of pulse duration and electrocyte geometry/excitability. However, a comprehensive assessment of expressed Single Nucleotide Polymorphisms (SNPs) throughout the entire transcriptome of African weakly electric fish, with the potential to identify further genes influencing EOD duration, is still lacking. This is of particular value, as discharge duration is likely based on multiple cellular mechanisms and various genes. Here we provide the first transcriptome-wide SNP analysis of African weakly electric fish species (genus Campylomormyrus) differing by EOD duration to identify candidate genes and cellular mechanisms potentially involved in the determination of an elongated discharge of C. tshokwe. Non-synonymous substitutions specific to C. tshokwe were found in 27 candidate genes with inferred positive selection among Campylomormyrus species. These candidate genes had mainly functions linked to transcriptional regulation, cell proliferation and cell differentiation. Further, by comparing gene annotations between C. compressirostris (ancestral short EOD) and C. tshokwe (derived elongated EOD), we identified 27 GO terms and 2 KEGG pathway categories for which C. tshokwe significantly more frequently exhibited a species-specific expressed substitution than C. compressirostris. The results indicate that transcriptional regulation as well cell proliferation and differentiation take part in the determination of elongated pulse durations in C. tshokwe. Those cellular processes are pivotal for tissue morphogenesis and might determine the shape of electric organs supporting the observed correlation between electrocyte geometry/tissue structure and discharge duration. The inferred expressed SNPs and their functional implications are a valuable resource for future investigations on EOD durations.
Order of Authors:	Julia Canitz
	Frank Kirschbaum
	Ralph Tiedemann
Opposed Reviewers:	Harold Zakon University of Texas at Austin Potential conflict of interest
	Matthew Arnegard Fred Hutchinson Cancer Research Center Potential conflict of interest
Additional Information:	

Transcriptome-wide single nucleotide polymorphisms related to electric organ discharge differentiation among African weakly electric fish species

Julia Canitz, Frank Kirschbaum and Ralph Tiedemann*

Julia Canitz
Unit of Evolutionary Biology/Systematic Zoology
Institute for Biochemistry and Biology
University of Potsdam
Karl-Liebknechtstr. 24-25, Haus 26
14476 Potsdam, Germany
julia.canitz@googlemail.com

Prof. Frank Kirschbaum
Faculty of Life Sciences, Chair: Biology and Ecology of Fishes
Humboldt University of Berlin
Philippstraße 13 (Haus 16)
10115 Berlin, Germany
frank.kirschbaum@staff.hu-berlin.de

Prof. Ralph Tiedemann (* corresponding author)
Unit of Evolutionary Biology/Systematic Zoology
Institute for Biochemistry and Biology
University of Potsdam
Karl-Liebknechtstr. 24-25, Haus 26
14476 Potsdam, Germany
tiedeman@uni-potsdam.de

Abstract

African weakly electric fish of the mormyrid genus *Campylomormyrus* generate pulse-type electric organ discharges (EODs) for orientation and communication. Their pulse durations are species-specific and elongated EODs are a derived trait. Differential gene expression among tissue-specific transcriptomes across species with different pulses and point mutations in single ion channel genes indicate a relation of pulse duration and electrocyte geometry/excitability. However, a comprehensive assessment of expressed Single Nucleotide Polymorphisms (SNPs) throughout the entire transcriptome of African weakly electric fish, with the potential to identify further genes influencing EOD duration, is still lacking. This is of particular value, as discharge duration is likely based on multiple cellular mechanisms and various genes. Here we provide the first transcriptome-wide SNP analysis of African weakly electric fish species (genus *Campylomormyrus*) differing by EOD duration to identify candidate genes and cellular mechanisms potentially involved in the determination of an elongated discharge of *C. tshokwe*. Non-synonymous substitutions specific to *C. tshokwe* were found in 27 candidate genes with inferred positive selection among *Campylomormyrus* species. These candidate genes had mainly functions linked to transcriptional regulation, cell proliferation and cell differentiation. Further, by comparing gene annotations between *C. compressirostris* (ancestral short EOD) and *C. tshokwe* (derived elongated EOD), we identified 27 GO terms and 2 KEGG pathway categories for which *C. tshokwe* significantly more frequently exhibited a species-specific expressed substitution than *C. compressirostris*. The results indicate that transcriptional regulation as well cell proliferation and differentiation take part in the determination of elongated pulse durations in *C. tshokwe*. Those cellular processes are pivotal for tissue morphogenesis and might determine the shape of electric organs supporting the observed correlation between electrocyte geometry/tissue structure and discharge duration. The inferred expressed SNPs and their functional implications are a valuable resource for future investigations on EOD durations.

Introduction

In closely related teleost fish, species-specific differences can be observed in morphology, behavior, reproduction, and communication [1-4]. These trait differences are often adaptive, especially if the species have evolved (or at least occur) in sympatry. African weakly electric fishes of the family Mormyridae (Osteoglossiformes; Teleostei) comprise numerous sympatric closely related species

differing in their electric sense. Mormyrids evolved an electric organ enabling them to produce electric organ discharges (EODs) for orientation and communication, i.e., mate recognition and species discrimination . The main structure of electric organs is formed by tens of specialized cells, called electrocytes, which have a disk-like shape and are stacked in cylindrical columns [15–21]. Either the anterior or the posterior face of electrocytes gives rise to several finger-like evaginations fusing in a stalk, serving as the interface to the electromotor neuron. Those stalks can penetrate the electrocyte and occur as single- or multiple stalk-systems . The electrocyte faces can be smooth or unevenly invaginated with papillae or folds increasing the membrane's surface [15, 16]. The membrane is excitable and packed with different ion channels, such that each single electrocyte is independently capable of generating an action potential [16]. The simultaneous release of all action potentials forms the weak pulse-type discharge which varies among closely related species in the number and orientation of phases as well as in its duration. In several African mormyrid weakly electric fish, slight differences in EOD characteristics occur between sexes and populations, while this was never detected in our focus genus *Campylomormyrus*, rendering their EOD a species-specific trait [14, 18]. Currently, 15 *Campylomormyrus* species are morphologically described and for 9 of them the EOD is known. Their waveforms can be bi- or triphasic with putatively ancestral pulse duration mostly shorter than 400µs (*C. compressirostris*; *C. tamandua*). A significantly longer EOD (> 4ms) occurs in *C. tshokwe*, *C. numenius* and *C. rhynchophorus* and is assumed to be the derived character state within the genus [22, 23].

In mormyrids, there is evidence that different patterns of electrocyte geometry, like electrocyte penetrations, cause different EOD waveforms. The electric organ of several mormyrid species has been histologically analyzed and a relation between electrocytes being stalk-penetrated or not and the EOD phase number could be observed [15–19]. Further, species with a multiple-stalk system are found to have a longer EOD than those with a single-stalk system [19]. Stalk size as well as the size of the electrocytes also seem to play a role in determining pulse duration. This is observed in *Brienomyrus* species where longer EODs emitted by males coincide with thicker electrocytes and larger stalks, relative to females [24, 25]. A comparable correlation can also be found among differentially discharging species of *Campylomormyrus* [18]. Moreover, a link between the increase of the electrocyte surface at the anterior face and EOD length has been found after hormone treatment (e.g., with testosterone) [24–27]. As these experiments are mainly performed in species with sex-specific EOD differences, it is not clear, whether hormones also contribute to inter-specific variation.

Further crucial components for an EOD generation are different types of ion channels integrated in the electrocyte membrane. They are responsible for generating the action potential, as they regulate

the in- and out-flux of sodium and potassium ions. Hence, variations in the abundance or molecular structure of ion channels may also contribute to different pulse durations. All these components are encoded in the genome. Sequence variation of genes coding for major cellular components like the ion channels has been assessed in relation to EOD duration in several mormyrid species [28–31], but could not fully explain the observed variation in EOD characteristics.

Mutations can affect a particular phenotypic trait in different ways. Single point mutations in the coding region of a particular gene can change the protein sequence and hence, its function, resulting in variation of a phenotypic trait [32–34]. Often, however, traits have a quantitative inheritance, i.e., they depend on several or many genes and multiple mutations in these genes may contribute to trait variability. Regarding the ion channels, SNPs have been detected between wave-type and pulse-type African weakly electric fish in one of the two paralogous genes of the voltage-gated potassium ion channel *kcna7a* [35]. Among *Campylomormyrus* species with shorter and longer EODs, SNPs are also found in one paralog of the voltage-gated sodium channel gene (*scn4aa*) [31]. In those genes, non-synonymous SNPs likely alter the protein function and thus, can contribute to a change of the EOD duration. Mutations can also affect gene expression by impairing promoter and transcription factor interactions, modifying the function of regulatory proteins like activators/repressors, or changing mRNA conformations (stability) [36–38]. Indeed, the elongated EOD in *C. tshokwe* is associated with significantly elevated expression in several ion channel genes [30] but the regulatory elements behind this up-regulation are not yet known. EOD characteristics are likely encoded by multiple genes being part of various cellular processes. Thus, an identification of genes and processes related to the EOD in weakly electric fish will provide further insights into its evolution and the mechanisms which cause the different pulse duration.

In this study we aim at identifying further genes potentially influencing the pulse duration by using RNA sequencing data of the electric organ of three *Campylomormyrus* and one *Gnathonemus* species with different EOD durations [40]. Specifically, we look for putative candidate genes with non-synonymous SNPs which are related to EOD duration. Such expressed single mutations may have an impact either on gene expression or protein functions. Furthermore, we put our data in the context of gene ontology to infer those biological processes which are likely involved in the determination of EOD duration. As EOD differentiation is likely due to divergent selection among closely related species, genes with species-specific mutations may be more common for those biological processes and mechanisms which are relevant for a derived elongation of pulse duration. Conceptually, we screen two closely related species with short and long EOD, respectively, for species-specific non-synonymous SNPs, having two further species with ancestral EOD characteristics as outgroup to discern putatively derived from ancestral SNP states.

Results

Transcriptome assemblies

In total, 66986804, 117994270, 41018678 and 99097520 high quality reads were used to assemble de novo the transcriptomes of *C. compressirostris*, *C. tshokwe*, *C. tamandua* and *Gnathonemus petersii*, respectively. The four Trinity assemblies produced between 141384 and 218372 contigs with N50 length of 1393 to 1881 bp (Table 1). According to the BUSCO analysis, the four transcriptomes matched an adequate proportion of 62.5% to 76.9% of the Actinopterygii core genes data set, depending on the species (Table 1). After isolating transcripts with the longest open reading frame (LORF) by following the TransDecoder pipeline, transcriptome size decreased by around 70%, leaving a remainder of 50241 transcripts for *C. compressirostris*, 48929 for *C. tamandua*, 65330 for *C. tshokwe*, and 52226 for *G. petersii* (Table 1). The identification of orthologous sequences among the four transcriptomes outputted a total number of 36285 orthogroups containing 16661 orthogroups present in all species of which 5284 were Single Copy Orthogroups (SCO). Subsequent alignment and filtering steps reduced the data set to a final number of 5071 SCO.

Approach I: Identification of candidate genes potentially related to EOD elongation in *C. tshokwe*

The goal of approach I was the identification of distinct putative candidate genes supposed to play a role in the EOD elongation in *C. tshokwe*. Candidate genes were identified by fulfilling two conditions: an inferred positive selection among the four mormyrid species ($\omega > 1$) and the occurrence of species-specific non-synonymous SNPs in *C. tshokwe*. The PAML/codeml analysis for inferring positive selection revealed 131 SCOs with $\omega > 1$. Among these SCOs, 39 sequences had at least one *C. tshokwe*-specific non-synonymous SNP. For 27 of them, only *C. tshokwe* was deviant, while *C. compressirostris* was identical to the outgroup taxa. These genes were considered as potential candidates related to the elongated EOD in *C. tshokwe* (see Methods for details). Twenty of these candidate genes had 1, six had 2 and one had 3 *C. tshokwe*-specific non-synonymous substitutions (Table 2).

The BLASTn analysis assigned 23 sequences to known coding genes, 3 sequences to uncharacterized proteins/regions and one sequence had no blast hit. Sixteen genes could be assigned to GO terms. Most of them were assigned to the terms 'binding' (75%) in the category Molecular Function (MF), 'cellular process' (25%), 'metabolic process' (25%) as well as 'response to stimuli'

(20%) in the category Biological Process (BP), and 'cell'/'cell part' (50%) and 'membrane' (15%) in the category Cellular Component (CC). Other GO terms were 'catalytic activity' (MF), 'localization' and 'biological regulation' (BP), 'protein-containing-complex' and 'organelle' (CC) (S1 Figure). The KEGG pathway annotation yielded 7 genes involved in 27 pathways, mainly belonging to the categories 'Human Diseases' (5 genes in 15 pathways), 'Environmental Information Processing' (3 genes in 5 pathways) and 'Organismal System' (3 genes in 4 pathways) (S1 Table). In three candidate genes, *C. tshokwe*-specific SNPs led to amino acid substitutions that likely impair/alter the protein function according to the MAPP analysis: SprT-like N-Terminal Domain (*sprtn),* EWS RNA binding protein 1a *(ewsr1),* and REV3 Like, DNA Directed Polymerase Zeta Catalytic Subunit *(rev3l*).

Approach II: Comparative annotation analyses of orthogroups between *C. compressirostri*s and *C. tshokwe*

Approach II was established to identify those biological mechanisms which are likely to take part in the EOD elongation of *C. tshokwe*. By SNP calling of multiple sequence alignments, 2203 and 1195 SCOs with species-specific SNPs were found for *C. tshokwe* and *C. compressirostris*, respectively. At least one non-synonymous SNP was detected in 824 SCOs for *C. tshokwe* and 825 for *C. compressirostris*. In 261 cases, the same SCO was affected (albeit at different non-synonymous sites), yielding 563 and 564 unique SCOs for *C. tshokwe* and *C. compressirostris*, respectively. After filtering for SCOs with inferred positive selection ($\omega > 1$), the data set of *C. tshokwe* contained 201 SCOs and that of *C. compressirostris* 192 SCOs. All the sequences of these SCOs matched the criteria of inferred positive selection and of carrying at least one non-synonymous SNP in only one of the two respective species (*C. tshokwe* or *C. compressirostris*) and named hereafter "candidate SCO-sequences".

For the identification of biological processes being relevant for the EOD duration, (i) a GO term annotation comparison and (ii) KEGG orthology annotation comparison of level A and B categories were performed among the two species. Regarding the first comparison, Blast2GO assigned 126 GO terms to 110 candidate SCO-sequences of *C. tshokwe*, while 116 candidate SCO-sequences were annotated to 114 GO terms for *C. compressirostris.* A combination of both annotations resulted in a total of 141 GO terms (S2 Table). For 42 GO-terms, there was a significant difference among the two species in the proportion of candidate SCO-sequences assigned to them. For 27 GO terms, this proportion was significantly higher in *C. tshokwe* than in *C. compressirostris* (Figure 1). These GO terms were mainly related to molecule binding, ion transport, signal transduction, cell communication, peptides, membrane, extracellular matrix and transcription factor complexes.

Conversely, the proportion of candidate SCO-sequences was significantly higher for 15 GO terms in *C. compressirostris* (Figure 1). These GO terms were associated with metal ion binding, cytoskeleton and phosphate related metabolic processes as well as cell organelle lumen and cytoskeleton components (S2 Table).

Figure 1: **Proportional ratio of candidate SCO-sequences between *C. tshokwe* and *C. compressirostris* for any annotated GO-term.** The figure illustrates the proportional ratio of candidate SCO-sequences between the two *Campylomormyrus* species (y-axis) relative to the total candidate SCO-sequence number for any annotated GO term (x-axis). The 95% confidence interval for an equal ratio (50:50) is depicted as the gray shaded area, rendering dots (i.e., GO terms) outside of the area significant (red circles).

The KEGG orthology analysis (KO) assigned 134 (*C. tshokwe*) and 130 *(C. compressirostris*) candidate SCO-sequences, respectively, to 145 and 188 KEGG pathways, of which 82 were found in both species, 63 only in *C. tshokwe* and 106 only in *C. compressirostris* (S3 Table).
For the KEGG level A categories, total candidate SCO-sequence counts were significantly different among the two species for 3 KEGG level A categories (Table 3). Candidate SCO-sequences were fewer in *C. tshokwe* for the categories 'Environmental Information Processing' (22 vs. 40 candidate SCO-sequences; $p = 0.034$), 'Cellular Processes' (25 vs. 45 candidate SCO-sequences; $p = 0.031$) and 'Organismal System' (26 vs. 73 candidate SCO-sequences; $p < 0.001$). Furthermore, we compared the distribution of candidate SCO-sequences in each KEGG level A category based on the next lower hierarchical KEGG level B (see Materials and Methods for details). The distribution of SCO-sequences to KEGG level B categories differed significantly among the two species within the A categories 'Genetic Information Processing' ($p = 0.009$), 'Cellular Process' ($p < 0.001$), 'Organismal Systems' ($p = 0.007$), and 'Human Diseases' ($p = 0.008$) (Table 3, S3 Table).
For a comparative analysis of the next lower hierarchal level (KEGG level B categories), the 251 pathways (82 shared, 63 and 106 unique pathways) were pooled to 45 KEGG level B categories (S3 Table). Among these, 11 categories exhibited significant differences in the percentage of candidate SCO-sequences among species (Figure 2). In *C. tshokwe* the 2 categories 'Transcription' and 'Cancer: specific types' were significantly overrepresented in terms of its proportion of candidate SCO-sequences, while in *C. compressirostris* 9 KEGG level B categories were overrepresented ('Signal transduction', 'Catabolism and transport', 'Endocrine system', 'Replication and repair', 'Cardiovascular diseases', 'Sensory system', 'Excretory system', 'Biosynthesis of other metabolites' and 'Drug resistance: antimicrobial') (Figure 2).

Figure 2: Proportional ratio of candidate SCO-sequences among *C. tshokwe* and *C. compressirostris* with annotated KEGG pathways. The figure represents the proportional ratio of candidate SCO-sequence counts among both species in a found KEGG level B category (y-axis), relative on the total number of candidate SCO-sequences in the respective category (x-axis). The 95% confidence interval for 50:50 ratio is depicted as gray shaded area, rendering dots (i.e., KEGG level B categories) outside of the area significant.

We also compared the annotations of the candidate SCO data set to those of the corresponding transcriptome (KEGG level A categories and GO terms) to evaluate whether certain annotation terms/categories are enriched among candidate SCO-sequences compared to all transcripts. In both species, no GO term was enriched at a FDR threshold of 0.05. Regarding the KEGG level A categories, a significant difference between the candidate SCO and transcriptome data set of *C. tshokwe* was observed in 5 categories (Figure 3). Here, the categories 'Metabolism', 'Genetic Information Processing' and 'Human Diseases' exhibited a significantly higher and 'Environmental Information Processing' and 'Organismal System' a significantly lower proportion in the candidate SCO data set. In *C. compressirostris*, only a single KEGG level A category, 'Human Diseases', yielded a significantly different (i.e., lower) proportion of sequence annotations in the candidate SCO data set compared to the entire transcriptome (Figure 3, S4 Table).

Figure 3: KEGG level A category assignments among the candidate SCO sequences, compared to the entire transcriptome. The bar chart shows the sequence percentage (y-axis) with an annotated KEGG Level A category (x-axis) in the SCO data set and entire transcriptomes of *C. tshokwe* (red) and *C. compressirostris* (blue). The error bars indicate the 95% confidence interval, taking the total absolute number of candidate SCO-sequences into account (confidence limits of proportions). Asterisks (*) depict significance at $p<0.05$.

Discussion

Within mormyrid species, several mechanisms are discussed to explain differences in EOD characteristics, i.e., electrocyte and stalk geometry, variation in hormone levels, and cell membrane excitability [18, 19, 43]. In this study, we identified 27 genes, which (1) are expressed in the electric organ, (2) are inferred to experience positive selection during the evolution of *Campylomormyrus*,

and (3) contain at least one non-synonymous mutation in the target species with the derived elongated EOD, i.e., *C. tshokwe*. These genes contain expressed genetic variation associated with EOD length differences and can hence be considered as potential candidates underlying this trait.

One of the most abundant functions among the candidate genes is transcriptional regulation. Ten of these genes regulate gene expression directly (e.g., zinc finger) or affect transcription factors and their pathways (e.g., Nuclear Factor kappa B (NFκB) and the Extracellular Signal-Regulated Kinase 1/2 (ERK1/ERK2) signaling cascade). The SLP Adaptor and CSK interacting membrane protein (*scimp*) positively regulates the ERK1/2 pathway which in turn modifies the activity of transcription factors, and hence changes gene expression levels of target genes that are important for the cell cycle progression and cell fate [44]. Additional three genes (*tsr2*, *mapkbp1* and *cd40*) play a role in the regulation of NFκB signaling pathway [45–47]. The protein complex NfκB is a transcription factor regulating the expression of several target genes being responsible for cell proliferation and differentiation as well as apoptosis. Neuroscientific studies proposed that the NFκB pathway mediates long-term changes in synaptic structures and neuronal plasticity via gene expression regulation [48, 49]. The co-repressor interacting with RBPJ1 (*cir1*) negatively regulates the NOTCH signaling pathway which is essential in cell-cell communication and cell differentiation processes at embryonic and adult stages [50–52]. It also plays a role in neuronal function and development, angiogenesis and cardiac valve homeostasis [53–56]. Furthermore, three genes code for zinc finger proteins (*znf32*, *znfx1* and *znf678*). This class of nucleic acid binding proteins has a zinc finger domain interacting with DNA and thus, acts as transcription factor. While the target genes of two of the zinc finger proteins (*znfx1* and *znf678*) are unknown in zebrafish, a knock-down of *znf32* suppresses the *SOX2* transcription which in turn enhances the regeneration of the nervous lateral line system [57]. In other model organisms, these three zinc finger proteins are associated with cancer pathways and epigenetic methylation [60–63]. Transcriptional regulation is also linked to the gene coding for Annexin A3 (*anxa3b*) and the EWS RNA binding protein 1 (*ewsr1*) [62, 63]. The relevance of this mechanism is corroborated by the GO terms 'cell communication', 'signal transduction' and 'transcription factor complexes' being significantly overrepresented among candidate SCO-sequences in *C. tshokwe* compared to *C. compressirostris* (S2 Table). Moreover, the KEGG orthology comparison of *C. tshokwe* yields the category ‘Genetic Information Processing’ (including processes of transcriptional regulation) to be significantly overrepresented among candidate SCO-sequences, compared to its occurrence in the entire transcriptome (Figure 3). Furthermore, the KEGG level B category ‘transcription’ is more abundant among candidate SCO-sequences of *C. tshokwe* than those of *C. compressirostris* (Figure 2). These categories refer to RNA polymerase, basal transcription factors and splicosome. Our results reveal that genes related to

processes and components of transcriptional regulation, exhibit an accelerated evolution in *C. tshokwe* with species-specific non-synonymous substitutions and inferred positive (divergent) selection. This emphasizes the importance of gene expression regulation for differences in EOD length among species. Indeed, previous studies already showed differential gene expression among *C. tshokwe* and *C. compressirostris* [30, 40], indicating the impact of transcriptional regulation on EOD waveforms. Transcriptional regulation would provide also a feasible explanation for the remarkable changes in EOD waveforms during the ontogeny of a *Campylomormyrus* fish [64], as expression levels may change during ontogeny [65, 66]. Experimental evidence on ontogenetic expression levels is however still lacking for our target taxa.

Our regulatory candidate genes are mainly associated with cell proliferation, cell differentiation and apoptosis, which holds true also for some non-regulatory candidate genes like the Vascular Cell Adhesion Molecule 1 (*vcam1*), DNA heat shock protein family member A1 (*dnaja1*) and REV3-like DNA directed polymerase subunit zeta (*rev3l*) (Table 2). The functions of their encoded proteins are related to cell expansion, cell survival and cell fate [67–70]. Furthermore, besides its regulatory function, Annexin A3 is mainly responsible for blood vessel and vascular cords formation [71, 72]. According to the MAPP results, the observed species-specific substitutions in the genes *ewsr1* and *rev3l* are likely to alter the protein function. The function of EWS RNA binding protein 1 (*ewsr1*) is shown to maintain miotic integrity and proneural cell survival in early developmental stages of zebrafish. A knock-down of this gene results in abnormalities of miotic spindles, followed by apoptosis and leading to a reduction of the proneural cell number and disorganization of neuronal networks during the early development stages [68]. Knock-down experiments of the second gene, *rev3l*, yield disorganized tissue with significantly reduced cell density [69]. A substitution at a functionally important site in these genes might lead to a functional loss or neofunctionalization and thus, to a modification of subsequent processes affecting cell fate. The association of cell proliferation processes to EOD duration is not only supported by many candidate genes (Approach I), but also by the KEGG categories ‘Human Diseases’ and ‘Cancer: Specific types’ which are significantly more abundant among candidate SCO-sequences of *C. tshokwe*. Indeed, most of the cancer pathways are linked to cell proliferation, cell differentiation and apoptosis, all crucial processes of tissue morphogenesis. Expressed genetic variation in these candidate genes may hence contribute to variation in electric organ tissue structures (e.g., multi- or single stalk systems) or cell morphs, supporting the hypothesis of an association between EOD duration and cell geometry [19, 24, 25].

Due to the teleost-specific whole genome duplication 350 mya [73], paralogous copies of essentially all genes emerged, and many of them were retained during teleost evolution. Some of

the paralogs may have retained their ancestral function or deteriorated into pseudogenes, but others underwent a neofunctionalization. One of our candidate genes, *anxa3*, is known to have two paralogs in zebrafish (*anxa3a*/ *anxa3b*) [74]. In *C. tshokwe* only one gene copy, *anax3b*, was found to be expressed in the electric organ. Annexin A3 has a similar function as some voltage gated ion channels shown to exhibit an electric organ-specific expression [30], i.e., the increase of membrane permeabilization activity and the influx regulation of calcium ions (Ca^{2+}) [75] which points out the known importance of ion activity and related proteins during EOD generation.

Conclusion

To our best knowledge, this is the first transcriptome-wide SNP analysis among African mormyrid weakly electric fish (genus *Campylomormyrus*). Our inferred 27 candidate genes and two molecular biological domains (transcriptional regulation and cell proliferation/cell fate) putatively support a link between tissue structures and EOD durations and provide new opportunities for molecular research regarding the EOD divergence in *Campylomormyrus* and other mormyrids. Genes affecting transcriptional regulation, and subsequent cell proliferation, cell differentiation and apoptosis seem likely to play a crucial role in determining pulse durations. Such processes are important for tissue morphogenesis and cell structures. They have hence the potential to contribute to different electric organ or electrocyte forms. Thus, our results are congruent with the hypothesis of the electric organ geometry not only to affect the shape of EOD pulses, but also their duration. Biochemical or physiological experiments via, e.g., knock-out trials were out of the scope of this study. Consequently, we can so far not proof a direct link between the inferred candidate genes and processes and the actual electric organ/electrocyte features. Thus, our study provides hypotheses about genes and processes relevant for EOD duration, which future research could build upon.

Methods

Transcriptome assemblies

We used RNA sequencing data from the electric organ (EO) of three *Campylomormyrus* species (*C. compressirostris*, *C. tshokwe* and *C. tamandua)* and the closely related species *G. petersii* to assemble tissue-specific transcriptomes (Figure 4A). We downloaded the Illumina raw reads from the Sequence Read Archive (SRA; http://www.ncbi.nlm.nih.gov/sra) with the accession number SRP050174 [40]. The processing of raw reads (quality filter, adapter trimming, etc.) was achieved

as described in Lamanna et al. (2015) [40]. Filtered paired-end reads of the four species were assembled de novo into separate transcriptomes using Trinity v. 2.2.0 with default parameters [76]. The four tissue-specific assemblies were tested for transcriptome completeness using BUSCO v3 [77]. For this purpose, transcriptomes were compared to the core gene set of Actinopterygii (state: 2018). The assemblies had been analyzed with the TransDecoder 3.0.1 pipeline to obtain the longest open reading frame for the transcripts [78]. The four transcriptomes served as input for the subsequent orthology analysis for which Orthofinder 1.1.10 was used (Figure 4) [79]. The Orthofinder tool is based on an all-versus-all blast of amino acid sequences, followed by a first sequence clustering taking into account the normalized bit score of the blast results. Afterwards, orthologous genes are selected and a final clustering by the Markov Cluster algorithm results in discrete orthogroups. Orthofinder distinguished between orthogroups with multiple sequences per species and Single Copy Orthogroups (SCO; one sequence per species). Further analyses were applied only to the SCOs to ensure analytical comparison among orthologous genes. The four nucleotide sequences of each SCO had been aligned codon-wise using PRANK v. 140110 (default parameters) [80] and trimmed to equal length by a customer bash script. To discard remaining paralogs, YASS 1.15 [81] was used. It compared all sequences pair-wisely and outputs the similarity for each pair in percentage. Single copy orthogroups with a similarity value below 90% were discarded from our data set. This procedure ensured that our retained genes are either clearly distinguishable from an ancient paralog (i.e., identified as separate SCOs) or do not have a paralog, at least not expressed in the electric organ. Furthermore, sequences shorter than 200bp were removed. Finally, a randomly chosen subset of SCOs was checked manually for correct filter criteria confirming the performance of the bioinformatical scripts and tools. The resulting data set of SCOs served as basis for our subsequent analyses (Figure 4).

Figure 4: Overview of the workflow of used data-analytical approaches. Shown are the major bioinformatical steps to create an input data set (A), steps for potential candidate gene identification (B), and the computational steps to create the candidate SCO data sets as well as their three annotation comparisons (C).

Approach I: Identification of candidate genes potentially related to EOD elongation in *C. tshokwe*

First, the ratio of non-synonymous (dN) to synonymous (dS) nucleotide substitution rates was calculated for all SCOs using the codeml package implemented in PAML v. 4.9 [82]. Therefore, the

sequence alignments of each SCOs were loaded to codeml separately and the site model M0 was chosen to compute the respective omega value ($\omega = dN/dS$). The ω value is an indicator of selective pressures on genes. A ratio significantly greater than 1 indicates positive selection. A ratio of 1 indicates neutral evolution at the protein level. A ratio less than 1 indicates selection to conserve the protein sequence (i.e., purifying selection). By using Geneious R 8.1.9, SCOs with $\omega > 1$ were screened manually for non-synonymous species-specific SNPs occurring only in the sequence of *C. tshokwe* (elongated EOD), when compared to the species with short EOD (*C. compressirostris*, *C. tamandua* and *G. petersii;* Figure 4B). As *C. tshokwe* and *C. compressirostris* are closely related, the direct comparison of these two species allowed us to focus on genes which experienced genetic changes since they diverged from their last common ancestor. Single Copy Orthogroups with $\omega > 1$ and non-synonymous, species-specific SNPs in *C. tshokwe* were considered as putative candidate genes for an elongated EOD (Figure 4B). Nucleotide sequences of these SCOs were blasted against the current nucleotide database of NCBI (nt database) using the BLASTn algorithm [83]. We also blasted sequences against the protein database of the Zebrafish Information Center (ZFIN) and the UniProtKB/Swiss-Prot database using the BLASTx algorithm. Putative candidate genes were assigned to GO terms using Blast2GO version 5.2.4 (java version) [84]. Gene ontology terms represent a controlled vocabulary of gene attributes which are organized hierarchically with three top categories: Biological Process (BP); Molecular Function (MF); Cellular Component (CC). In order to identify biological pathways in which the candidate genes occur, we uploaded the nucleotide sequences to the Kyoto Encyclopedia of Genes and Genomes Automatic Annotation Server (KAAS) [85] and blasted them against the gene data base of all available fish species as well as the human (*Homo sapiens*) and mouse (*Mus musculus*) data bases. Each run was performed with a bi-directional best hit algorithm (BBH). The Kyoto Encyclopedia of Genes and Genomes (KEGG) is a database of molecular functions which stores orthologs of experimentally characterized genes and proteins which are included in different biological processes (KEGG pathways). Each unit of these pathways is defined by a KO number. The KO numbers can be pooled by pathways which in turn can be grouped to level B categories. Several level B categories form a level A category. KEGG distinguishes six level A categories: 'Metabolism', 'Genetic Information Processing', 'Environmental Information Processing', 'Cellular Processes', 'Organismal Systems', and 'Human Diseases'. We applied a multivariate analysis of protein polymorphisms (MAPP) for the same gene set to detect impaired amino acids in the sequence. MAPP calculates a score to predict the impact of amino acid substitutions on protein function and structure. This impact score considers the properties of the amino acids and the phylogenetic relationship in the appropriated gene (gene tree) [86].

Approach II: Annotation comparisons of orthogroups between *C. compressirostri*s and *C. tshokwe*

The second approach aimed at identifying cellular and molecular mechanisms which play a role in the differentiation of EOD duration. Therefore, all SCO alignments were converted into a VCF file format using the tool SNP-sites to call SNPs from multiple sequence alignments [87]. Singe copy orthogroups with non-synonymous species-specific SNPs in *C. tshokwe* and *C. compressirostris* were isolated separately, creating two candidate SCO data sets. Subsequently, ω values for each remaining SCO were determined with the Ka/Ks Calculator 2.0 using the model selection according to the AICs (MS) [88]. Its underlying calculation deviates from the M0 site model (codeml), as it relies on a pairwise calculation across all input sequences and outputs a ω value for each possible combination. Only SCO with $\omega > 1$ for the pairing of *C. tshokwe* and *C. compressirostris* were retained in the candidate SCO data sets (Figure 4C). As we wanted to look for over-/underrepresentation of candidate SCO-sequences in certain biological processes, the candidate SCO data sets of both species were analyzed separately by a GO term annotation using Blast2GO version 5.2.4. Furthermore, KEGG orthology annotations were performed for both candidate SCO data sets as well as for both transcriptomes, considering for any gene only the transcript with the longest open reading frame (Figure 4C). We uploaded each data set separately to KAAS and used the same parameters and databases as in approach I. The GO term and KEGG pathway annotations were used for four annotation comparisons (Figure 4C).

The species-specific candidate SCO data sets were sorted by GO terms and the number of candidate SCO-sequences in each GO term was determined. Their proportion (in their respective candidate SCO data set) was compared among the two species. To account for different absolute sequence numbers underlying this comparison, we determined 95% confidence intervals of proportions at an equal ratio (50:50) for different total numbers of sequences ($2 \leq n \leq 200$) (S5 Table). Upper and lower confidence intervals were plotted using R version 3.4.4 (R Development Core Team, 2008). To identify the GO terms with a significant difference among *C. tshokwe* and *C. compressirostris*, we calculated the proportional ratio of candidate SCO-sequences between both species for each GO term and plotted them against the confidence interval, rendering GO terms outside the confidence interval significant. For the second analysis (KEGG annotation comparison), total numbers of candidate SCO-sequences in each KEGG level A category were compared among the two species and significant differences were tested with the Fisher Exact Test ($\alpha = 5\%$). In addition, for each KEGG level A category, it was tested with the Chi^2 test ($\alpha = 5\%$) whether candidate SCO-

sequences were assigned disproportionally to the level B categories (next lower hierarchical KEGG level) among the two species. For a closer look on the KEGG annotations, we grouped the pathways by KEGG level B categories and counted the candidate SCO-sequences in each KEGG level B category for both data sets. Here, KEGG categories were tested for significance among *C. tshokwe* and *C. compressirostris* analogous to the GO terms, i.e., identification of outliers to the 95% confidence interval of proportions. Data were visualized using R version 3.4.4. We further evaluated whether certain GO terms or KEGG Level A categories are enriched among the candidate SCO-sequences, compared to the respective transcriptome (Figure 4C). The enrichment analysis of Blast2GO (Fisher's Exact test; FDR=0.05) was used to identify over- or underrepresented GO terms in the candidate SCO data set. We applied this analysis to both species separately. To reveal KEGG level A categories with significant differences between the candidate SCO and transcriptome data set, the sequence number which matched each of the 6 KEGG level A categories were determined in each single data set separately (2x candidate SCO data sets and 2x transcriptome data sets). Their proportions (in their respective data set) were calculated and the respective 95% confidence intervals were determined using the online tool of the Allto Market Research web site [89]. The data were illustrated as bar chart with Microsoft Excel 2010.

Acknowledgements

The authors would like to acknowledge Prof. Michael Lenhard and Dr. Christian Kappel, University of Potsdam, for providing an access to their computational server to perform bioinformatic analyses. We also want to thank Dr. Ina-Maria Schedina and Dr. Sofia Paraskevopoulou for their support solving bioinformatic problems.

References

1. Smith-Vaniz WF, Kaufman LS, Glowacki J. Species-specific patterns of hyperostosis in marine teleost fishes. *Marine Biology*. 1995;121:573–80.
2. Fast MD, Sims DE, Burka JF, Mustafa A, Ross NW. Skin morphology and humoral non-specific defence parameters of mucus and plasma in rainbow trout, coho and Atlantic salmon. *Comp Biochem Physiol Part A*. 2002;132:645–57.
3. Assis CA. The lagenar otoliths of teleosts: their morphology and its application in species

identification, phylogeny and systematics. *J Fish Biol*. 2003;62:1268–95.

4. Liao C, Chen S, Guo Z, Ye S, Zhang T, Li Z, Murphy BR, Liu J. Species-specific variations in reproductive traits of three yellow catfish species (*Pelteobagrus spp*.) in relation to habitats in the Three Gorges Reservoir, China. *PLoS One*. 2018;13(7):e0199990. Available from: doi:10.1371/journal.pone.0199990.
5. Van Rijssel JC, Moser FN, Frei D, Seehausen O. Prevalence of disruptive selection predicts extent of species differentiation in Lake Victoria cichlids. *Proc R Soc Lond [Biol]*. 2018;285(1871). Available from: doi:10.1098/rspb.2017.2630.
6. Martin MD, Mendelson TC. Male behaviour predicts trait divergence and the evolution of reproductive isolation in darters (Percidae: Etheostoma). *Anim Behav*. 2016;112:179–86.
7. Carlson BA. Differences in electrosensory anatomy and social behavior in an area of sympatry between two species of mormyrid electric fishes. *J Exp Biol*. 2016;219:31–43.
8. Kocher TD. Adaptive evolution and explosive speciation: The cichlid fish model. *Nature Rev Genet*. 2004;5:288–98.
9. Hopkins CD, Bass AH. Temporal Coding of Species Recognition Signals in an Electric Fish. *Science*. 1981;212(4490):85-7. Available from: doi:10.1126/science.7209524.
10. Kramer B, Kuhn B. Species recognition by the sequence of discharge intervals in weakly electric fishes of the genus *Campylomormyrus* (Mormyridae, Teleostei). *Anim Behav*. 1994;48:435–45.
11. Moller P. 'Communication' in weakly electric fish, *Gnathonemus niger* (Mormyridae): I. Variation of electric organ discharge (EOD) frequency elicited by controlled electric stimuli. *Anim Behav*. 1970;18:768–86.
12. Nagel R, Kirschbaum F, Hofmann V, Engelmann J, Tiedemann R. Electric pulse characteristics can enable species recognition in African weakly electric fish species. *Sci Rep*. 2018;8(1):10799: Available from: doi:10.1038/s41598-018-29132-z.
13. Nagel R, Kirschbaum F, Engelmann J, Hofmann V, Pawelzik F, Tiedemann R. Male-mediated species recognition among African weakly electric fishes. *R Soc open sci*. 2018;5(2):170443. Available from: doi:10.1098/rsos.170443.
14. Feulner PGD, Plath M, Engelmann J, Kirschbaum F, Tiedemann R. Electrifying love: electric fish use species-specific discharge for mate recognition. *Biol Lett*. 2008;5:225–8.
15. Bass AH. Species Differences in Electric Organs of Mormyrids: Substrates for Species-Typical Electric Organ Discharge Waveforms. *J Comp Neurol*. 1986;244(3):313–30.
16. Bennett MVL, Grundfest H. Bioelectrogenesis. In: Chagas C, Carvalho AP. (eds.) *Studies on the morphology and electrophysiology of electric organs.* III. Electrophysiology of

542 electric organs in mormyrids. London: Elsevier; 1961. p. 113–35.
543 17. Gallant JR, Arnegard ME, Sullivan JP, Carlson BA, Hopkins CD. Signal variation and its
544 morphological correlates in *Paramormyrops kingsleyae* provide insight into the evolution
545 of electrogenic signal diversity in mormyrid electric fish. *J Comp Physiol [A]*.
546 2011;197(8):799–817. Available from: doi: 10.1007/s00359-011-0643-8
547 18. Paul C, Mamonekene V, Vater M, Feulner PGD, Engelmann J, Tiedemann R, Kirschbaum
548 F. Comparative histology of the adult electric organ among four species of the genus
549 *Campylomormyrus* (Teleostei: Mormyridae). *J Comp Physiol [A]*. 2015;201(4):357–74.
550 19. Bass AH. Electric Organs Revisited: Evolution of a Vertebrate Communication and
551 Orientation Organ. In: Bullock TH, Heiligenberg W. (eds.) *Electroreception*. New York:
552 Wiley; 1986. p. 13–70.
553 20. Denizot JP, Kirschbaum F, Max Westby GW, Tsuji S. The larval electric organ of the
554 weakly electric fish *Pollimyrus* (*Marcusenius*) *isidori* (Mormyridae, Teleostei). *J*
555 *Neurocytol*. 1978;7(2):165–81.
556 21. Bennett MVL. Electric Organs. In: Hoar WS, Randall DJ. (eds.) *Fish Physiology*. New
557 York: Academic Press. 1971. p. 347–491.
558 22. Tiedemann R, Feulner PGD, Kirschbaum F. Electric Organ Discharge Divergence
559 Promotes Ecological Speciation in Sympatrically Occurring African Weakly Electric Fish
560 (*Campylomormyrus*). In: Glaubrecht, M. (ed.) *Evolution in Action*. Berlin: Springer. 2010.
561 p. 307-21.
562 23. Lamanna F, Kirschbaum F, Ernst ARR, Feulner PGD, Mamonekene V, Paul C, Tiedemann
563 R. Species delimitation and phylogenetic relationships in a genus of African weakly-
564 electric fishes (Osteoglossiformes, Mormyridae, *Campylomormyrus*). *Mol Phylogenet Evol*.
565 2016;101:8–18. Available from: doi: 10.1016/j.ympev.2016.04.035
566 24. Freedman EG, Olyarchuk J, Marchaterre MA, Bass AH. A Temporal Analysis of
567 Testosterone-Induced Changes in Electric Organs and Electric Organ Discharges of
568 Mormyrid Fishes. *J Neurobiol*. 1989;20(7):619–34.
569 25. Bass AH, Denizot J, Marchaterre MA. Ultrastructural Features and Hormone- Dependent
570 Sex Differences of Mormyrid Electric Organs. *J Comp Neurol*. 1986;254(4):511–528.
571 26. Landsman RE, Moller P. Testosterone changes the electric organ discharge and external
572 morphology of the mormyrid fish, *Gnathonemus petersii* (Mormyriformes). *Experientia*.
573 1988;44(10):900–3.
574 27. Bass AH, Hopkins CD. Hormonal control of sex differences in the electric organ discharge
575 (EOD) of mormyrid fishes. *J Comp Physiol [A]*. 1985;156(5):587–604.

28. Zakon HH, Jost MC, Zwickl DJ, Lu Y, Hillis DM. Molecular evolution of Na^+ channels in teleost fishes. *Integr Zool*. 2009;4(1):64–74.
29. Zakon HH, Lu Y, Zwickl DJ, Hillis DM. Sodium channel genes and the evolution of diversity in communication signals of electric fishes: Convergent molecular evolution. *Proc Natl Acad Sci USA*. 2006;103(10):3675–80.
30. Nagel R, Kirschbaum F, Tiedemann R. Electric organ discharge diversification in mormyrid weakly electric fish is associated with differential expression of voltage-gated ion channel genes. *J Comp Physiol [A]*. 2017;203(3):183–95.
31. Paul C, Kirschbaum F, Mamonekene V, Tiedemann R. Evidence for Non-neutral Evolution in a Sodium Channel Gene in African Weakly Electric Fish (*Campylomormyrus*, Mormyridae). *J Mol Evol*. 2016;83(1-2):61–77. Available from: doi:10.1007/s00239-016-9754-8.
32. Gratten J, Beraldi D, Lowder BV, McRae AF, Visscher PM, Pemberton JM, Slate J. Compelling evidence that a single nucleotide substitution in TYRP1 is responsible for coat-colour polymorphism in a free-living population of Soay sheep. *Proc R Soc Lond [Biol]*. 2007;274(1610):619–26.
33. Hoekstra HE, Hirschmann RJ, Bundey RA, Insel PA, Crossland JP. A Single Amino Acid Mutation Contributes to Adaptive Beach Mouse Color Pattern. *Science*. 2006;313(5783):101–4.
34. Kamiya T, Kai W, Tasumi S, et al. A Trans-Species Missense SNP in *Amhr2* Is Associated with Sex Determination in the Tiger Pufferfish, *Takifugu rubripes* (Fugu). *PLoS Genet*. 2012;8(7):e1002798. Available from: doi:10.1371/journal.pgen.1002798.
35. Swapna I, Ghezzi A, Markham MR, Halling DB, Lu Y, Gallant JR, Zakon HH. Electrostatic Tuning of a Potassium Channel in Electric Fish. *Curr Biol*. 2018;28(13):2094–102. Available from: doi:10.1016/j.cub.2018.05.012.
36. Murgatroyd C. Impaired Repression at a Vasopressin Promoter Polymorphism Underlies Overexpression of Vasopressin in a Rat Model of Trait Anxiety. *J Neurosci*. 2004;24(35):7762–70.
37. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. *Nature Rev Genet*. 2009;10(3):155–9. Available from: doi:10.1038/nrg2521.
38. Shastry BS. SNPs: Impact on Gene Function and Phenotype. In: Komar A. (ed.) *Single Nucleotide Polymorphisms: Methods in Molecular Biology (Methods and Protocols)*. Totowa: Humana Press; 2009. p. 3-22.
39. Marguerat S, Bähler J. RNA-seq: from technology to biology. *Cell Mol Life Sci*.

2010;67:569–79. Available from: doi:10.1007/s00018-009-0180-6.

40. Lamanna F, Kirschbaum F, Waurick I, Dieterich C, Tiedemann R. Cross-tissue and cross-species analysis of gene expression in skeletal muscle and electric organ of African weakly-electric fish (Teleostei; Mormyridae). *BMC Genomics*. 2015;16:668. Available from: doi:10.1186/s12864-015-1858-9.

41. Gallant JR, Hopkins CD, Deitcher DL. Differential expression of genes and proteins between electric organ and skeletal muscle in the mormyrid electric fish *Brienomyrus brachyistius*. *J Exp Biol*. 2012;215(14):2479-94. Available from: doi:10.1242/jeb.063222.

42. Güth R, Pinch M, Samanta MP, Chaidez A, Unguez GA. *Sternopygus macrurus* electric organ transcriptome and cell size exhibit insensitivity to short-term electrical inactivity. *J Physiol (Paris)*. 2016;110(3):233–44.

43. Alves-Gomes JA, Hopkins CD. Molecular Insights into the Phylogeny of Momyriform Fishes and the Evolution of Their Electric Organs. *Brain Behav Evol*. 1997;49(6):324–50.

44. Roskoski Jr R. ERK1/2 MAP kinases: Structure, function, and regulation. *Pharmacol Res*. 2012;66(2):105–43. Available from: doi:10.1016/j.phrs.2012.04.005.

45. He H, Zhu D, Sun J, Pei R, Jia S. The Novel Protein TSR2 Inhibits the Transcriptional Activity of Nuclear Factor-κB and Induces Apoptosis. *Mol Biol*. 2011;45(3):451–7.

46. Elgueta R, Benson MJ, Vries de VC, Noelle RJ. Molecular mechanism and function of CD40⁄CD40L engagement in the immune system. *Immunol Rev*. 2009;229(1):152–72.

47. Lecat A, Di Valentin E, Somja J, et al. The JNK-binding protein (JNKBP1) acts as a negative regulator of NOD2 signaling by inhibiting its oligomerization process. *J Biol Chem*. 2012287(35):29213-26. Available from: doi:10.1074/jbc.M112.355545.

48. Mattson MP. NF-κB in the Survival and Plasticity of Neurons. *Neurochem Res*. 2005;30(6-7):883–93.

49. Mattson MP, Culmsee C, Yu Z, Camandola S. Roles of Nuclear Factor κB in Neuronal Survival and Plasticity. *J Neurochem*. 2000;74(2):443–56. Available from: doi:10.46/j.1471-4159.2000.740443.x.

50. Zhang J, Chen H, Weinmaster G, Hayward SD. Epstein-Barr Virus *Bam*Hi-A Rightward Transcript-Encoded RPMS Protein Interacts with the CBF1-Associated Corepressor CIR To Negatively Regulate the Activity of EBNA2 and NotchIC. *J Virol*. 2001;75(6):2946–56.

51. Contreras-Cornejo H, Saucedo-Correa G, Oviedo-Boyso J, Valdez-Alarcón JJ, Baizabal-Aguirre VM, Cajero-Juárez M, Bravo-Patiño A. The CSL proteins, versatile transcription factors and context dependent corepressors of the notch signaling pathway. *Cell Div*. 2016;11:12.

644 52. Hsieh JJ, Zhou S, Chen L, Young DB, Hayward SD. CIR, a corepressor linking the DNA
645 binding factor CBF1 to the histone deacetylase complex. *Proc Natl Acad Sci USA*.
646 1999;96(1):23–8.
647 53. Aguirre A, Rubio ME, Gallo V. Notch and EGFR pathway interaction regulates neural
648 stem cell number and self-renewal. Nature. 2011;467:323–7.
649 54. Gaiano N, Fishell G. The Role of Notch in Promoting Glial and Neural Stem Cell Fates.
650 *Annu Rev Neurosci*. 2002;25:471–90.
651 55. MacGrogan D, Luna-Zurita L, de la Pompa JL. Notch Signaling in Cardiac Valve
652 Development and Disease. *Birth Defects Res A*. 2011;91(6):449–59. Available from:
653 doi:10.1002/bdra.20815.
654 56. Grego-Bessa J, Luna-zurita L, Monte G, et al. Notch Signaling is Essential for Ventricular
655 Chamber Development. *Dev Cell*. 2007;12(3):415–29.
656 57. Wei Y, Li K, Yao S, et al. Loss of *ZNF*32 augments the regeneration of nervous lateral line
657 system through negative regulation of *SOX*2 transcription. *Oncotarget*. 2016;7(43):70420-
658 36.
659 58. Dominguez-Salas P, Moore SE, Baker MS, et al. Maternal nutrition at conception
660 modulates DNA methylation of human metastable epialleles. *Nat Commun*. 2014;5:3746.
661 Available from: doi:10.1038/ncomms4746.
662 59. Simeone P, Alberti S. Epigenetic heredity of human height. *Physiol Rep*. 2014;2(6):e12047.
663 Available from: doi:10.14814/phy2.12047.
664 60. Li Y, Zhang L, Li K, et al. ZNF32 inhibits autophagy through the mTOR pathway and
665 protects MCF-7 cells from stimulus-induced cell death. *Sci Rep*. 2015;5:9288. Available
666 from: doi:10.1038/srep09288.
667 61. Ishidate T, Ozturk AR, Durning DJ, Sharma R, Shen E, Chen H, Seth M, Shirayama M,
668 Mello CC. ZNFX-1 Functions within Perinuclear Nuage to Balance Epigenetic Signals.
669 *Mol Cell*. 2018;70(4):639–49. Available from: doi:10.1016/j.molcel.2018.04.009.
670 62. Park JE, Lee DH, Lee JA, Park SG, Kim NS, Park BC, Cho S. Annexin A3 is a potential
671 angiogenic mediator. *Biochem Biophy Res Co*. 2005;337(4):1283–7.
672 63. Luo Y, Blechingberg J, Fernandes AM, Li S, Fryland T, Børglum AD, Bolund L, Nielsen
673 AL. EWS and FUS bind a subset of transcribed genes encoding proteins enriched in RNA
674 regulatory functions. *BMC Genomics*. 2015;16:929. Available from: doi:10.1186/s12864-
675 015-2125-9.
676 64. Nguyen L, Paul C, Mamonekene V, Bartsch P, Tiedemann R, Kirschbaum F. Reproduction
677 and development in some species of the weakly electric genus *Campylomormyrus*

(Mormyridae, Teleostei). *Environ Biol Fish*. 2017;100(1):49–68.

65. Hassin S, Claire M, Holland H, Zohar Y. Ontogeny of Follicle-Stimulating Hormone and Luteinizing Hormone Gene Expression During Pubertal Development in the Female Striped Bass, *Morone saxatilis* (Teleostei). *Biol Reprod*. 2005;61(6):1608–15.

66. Rohlfing K, Stuhlmann F, Docker MF, Burmester T. Convergent evolution of hemoglobin switching in jawed and jawless vertebrates. *BMC Evol Biol*. 2016;16:30. Available from: doi:10.1186/s12862-016-0597-0.

67. Terada K, Yomogida K, Imai T, Kiyonari H, Takeda N, Kadomatsu T, Yano M, Aizawa S, Mori M. A type I DnaJ homolog, DjA1, regulates androgen receptor signaling and spermatogenesis. *EMBO J*. 2005;24(3):611–22.

68. Azuma M, Embree LJ, Sabaawy H, Hickstein DD. Ewing Sarcoma Protein Ewsr1 Maintains Mitotic Integrity and Proneural Cell Survival in the Zebrafish Embryo. *PLoS One*. 2007;2(10):e979.

69. Wittschieben J, Shivji MKK, Lalani E, Jacobs MA, Marini F, Gearhart P., Rosewell I, Stamp G, Wood RD. Disruption of the developmentally regulated *Rev3l* gene causes embryonic lethality. *Curr Biol*. 2010;10(19):1217–20.

70. Rosen GD, Sanes JR, LaChance R, Cunningham JM, Roman J, Dean DC. Roles for the Integrin VLA-4 and Its Counter Receptor VCAM-1 in Myogenesis. *Cell*. 1992;69(7):1107-19.

71. Meadows SM, Cleaver O. Annexin A3 Regulates Early Blood Vessel Formation. *PLoS One*. 2015;10(7):e0132580. Available from: doi:10.1371/journal.pone.0132580.

72. Farber SA, De Rose RA, Olson ES, Halpern ME. The Zebrafish Annexin Gene Family. *Genome Res*. 2003;13(6A):1082–96.

73. Pasquier J, Cabau C, Nguyen T, et al. Gene evolution and gene expression after whole genome duplication in fish: The PhyloFish database. *BMC Genomics*. 2016;17:368. Available from: doi:10.1186/s12864-016-2709-z.

74. Maqueen DJ, Johnson IA. Evolution of follistatin in teleost revealed through phylogenetic, genomic and expression analysis. *Dev Genes Evol*. 2008;218(1). Available from: doi:10.1007/s00427-007-0194-8.

75. Hofmann A, Raguénès-Nicol C, Favier-Perron B, Mesonero J, Huber R, Russo-Marie F, Lewit-Bentley A. The Annexin A3 - Membrane Interaction Is Modulated by an N-Terminal Tryptophan. *Biochemistry*. 2000;39(26):7712–21.

76. Grabherr MG, Haas B, Yassour M, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 2011;29(7):644–52. Available from:

712 doi:10.1038/nbt.1883.
713 77. Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO:
714 assessing genome assembly and annotation completeness with single-copy orthologs.
715 *Bioinformatics*. 2015;31(19):3210–2. Available from: doi:10.1093/bioinformatics/btv351
716 78. Haas BJ, Papanicolaou A, Yassour M, et al. De novo transcript sequence reconstruction
717 from RNA-Seq using the Trinity platform for reference generation and analysis. *Nat Protoc*.
718 2013;8:1494-512. Available from: doi:10.1038/nprot.2013.084.De.
719 79. Emms DM, Kelly S. OrthoFinder: solving fundamental biases in whole genome
720 comparisons dramatically improves orthogroup inference accuracy. *Genome Biol*. 2015;
721 16:157. Available from: doi:10.1186/s13059-015-0721-2.
722 80. Goldman N, Löytynoja A. An algorithm for progressive multiple alignment of sequences
723 with insertions. *Proc Nat Acad Sci USA*. 2005;102(30):10557-62.
724 81. Noé L, Kucherov G. YASS: enhancing the sensitivity of DNA similarity search. *Nucleic*
725 *Acids Res*. 2005;33:540–3.
726 82. Yang Z. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Mol Biol Evol*.
727 2007;24(8):1586–91.
728 83. Zhang Z, Schwartz S, Wagner L, Miller W. A Greedy Algorithm for aligning DNA
729 sequences. *J Comput Biol*. 2000;7(1-2):203–14.
730 84. Conesa A, Götz S, García-gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal
731 tool for annotation, visualization and analysis in functional genomics research.
732 *Bioinformatics Application Note*. 2005;21(18):3674–6.
733 85. Moriya Y, Itoh M, Okuda S, Yoshizawa A, Kanehisa M. *KAAS KEGG Automatic*
734 *Annotation Server*. Available from: www.genome.ip/kaas-bin/kaas_main [Accessed 10th
735 December 2018].
736 86. Stone EA, Sidow A. Physicochemical constraint violation by missense substitutions
737 mediates impairment of protein function and disease severity. *Genome Res*.
738 2005;15(7):978-86.
739 87. Keane JA, Page AJ, Delaney AJ, Taylor B, Seemann T, Harris SR, Soares J. SNP-sites:
740 rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb Genom*.
741 2016;2(4):e000056. Available from: doi: 10.1099/mgen.0.000056.
742 88. Zhang Z, Li J, Zhao X, Wang J, Wong GK, Yu J. KaKs_Calculator : Calculating Ka and Ks
743 Through Model Selection and Model Averaging. *Genom Proteom Bioinf*. 2006;4(4):259–
744 63.
745 89. McCallum D, Layton K. *Allto Market Research*. https://www.allto.co.uk/tools/statistic-

calculators/confidence interval-for-proportions-calculator/ [accessed 15th December 2018].

Supporting information

S1 Figure: GO term annotations regarding the 27 candidate genes of *C. tshokwe*. Pie charts depict the proportion of sequences annotated to the main GO categories Biological Process (A), Cellular Component (B) and Molecular Function (C).

S1 Table: KEGG orthology annotation of 27 candidate genes of *C. tshokwe*. The number of pathway annotations per gene in each KEGG level A category is listed.

S2 Table: List of GO terms occurring in the candidate SCO data sets of *C. tshokwe* and *C. compressirostris*. 141 GO term names and IDs, corresponding absolute number of candidate SCO-sequences and relative proportions of both species and for each GO term are given.

S3 Table: List of inferred KEGG pathways of both candidate SCO data sets for *C. tshokwe* and *C. compressirostris*. Raw data are separated by sheets. Shared and species-specific KEGG pathways, their level B categories and corresponding number of candidate SCO-sequences for both species are given.

S4 Table: Confidence interval of proportion among KEGG results of the candidate SCO data sets and the respective transcriptome. Sequence counts [%] matching the 6 KEGG level A categories and their respective confidence intervals are listed at a 5% level are listed. Data are available for the candidate SCO data sets and the transcriptomes (LORF) in *C. tshokwe* and *C. compressirostris*.

S5 Table: Upper and Lower confidence intervals of proportion for total numbers (Approach II).

Table 1: Assembly statistics for the four transcriptomes

	C. compressirostris	***C. tshokwe***	***C. tamandua***	***G. petersii***
# of processed reads	66 986 804	117 994 270	41 018 678	99 097 520
# of contigs (Trinity assembly)	160 665	218 372	141 384	176 155
N50 (Trinity assembly)	1393	1873	1881	1643
# of transcripts with LORF (TransDecoder pipeline)	50 241	65 330	48 929	52 226
BUSCO completeness (Actinopterygii core gene set)	62.5 %	76.9 %	64.4 %	68.0 %

778

779

Table 2: List of 27 candidate genes potentially related to EOD elongation in *C. tshokwe*.

Gene	**Gene Name**	**SNP position ****	**Amino acid change**	**General function**
trs2	TSR2, Ribosome Maturation Factor	A 432 G	K 138 E	- transcriptional regulation - apoptosis
znf32	Zinc Finger Protein 32	T 110 C T 452 C	V 37 A V 151 A	- transcriptional regulation
mapkbp1	Mitogen-Activated Protein Kinase Binding Protein 1-like	G 122 A	S 40 N	- immune system response - regulatory function
cd40	Tumor Necrosis Factor (TNF) Receptor superfamily member 5	T 330 G	D 110 E	- immune system response
sprtn	SprT-like N-Terminal Domain (Spartan)	T 919 C	S 307 P *	- DNA damage response
trim56	Tripartite Motif Containing 56	G 271 A G 413 A	V 91 I G 138 E	- immune system response
nfu1	NFU1 iron-sulfur cluster scaffold homolog	A 454 G	I 152 V	- iron-sulfur cluster biogenesis
wdyhv1	WDYHV Motif Containing 1	G 305 C	R 102 P	- cellular protein modification process
cir1	Corepressor Interacting With RBPJ, 1	T 1000 C A 1238 G	F 334 L D 413 G	- transcriptional regulation - signal transduction
dnaja1	DnaJ Heat Shock Protein Family (Hsp40) Member A1	T 838 A T 1034 C	S 280 T V345 A	- protein folding - regulation of androgene receptor activity

cstl1	Cystatin-like	G 92 A	G 31 E	- regulatory function
ewsr1	EWS RNA-binding protein 1	A 598 G	T 200 A*	- neuron development - transcriptional regulation
anxa3b	Annexin A3b	A 544 G	N 182 D/E	- cell morphogenesis - membrane permeability
rev3I	REV3 Like, DNA Directed Polymerase Zeta Catalytic Subunit	A 400 G	T 134 A*	- DNA repair - cell proliferation
ap4b1	Adaptor Related Protein Complex 4 Subunit Beta 1	G 404 A C 806 T	G 135 D A 269 V	- localization
vcam1	Vascular Cell Adhesion Molecule 1	G 77 A	A 26 N	- cell-cell recognition
znfx1	Zinc finger, NFX1-type containing 1	T 173 C	V 58 A	- DNA-binding - transcription factor activity
hbba1	Hemoglobin, beta adult 1	T 57 G	F 19 L	- oxygen transport
Scimp	SLP Adaptor and CSK Interacting Membrane Protein	T 256 C	S 87 P	- immune synapse formation - signal transduction
znf678	Zinc Finger Protein 678	T 491 C	V 164 T/A	- transcriptional regulation
igdcc4	Immunoglobulin superfamily DCC subclass member 4	A 428 G	N 143 S	- binding
atp5mf	ATP Synthase Membrane Subunit F	A 275 G	D 92 G/S	- ATP production
Rgmb	Repulsive Guidance Molecule BMP co-receptor b (3'UTR)	G 8 C	W 3 S	- development of nervous system
cunh2orf42	Chromosome unknown C2orf42 homolog	A 52 G T 97 C	K 18 E S 33 P	- integral component of membrane
-	uncharacterized LOC111853234 transcript variant X2	G 209 A	G 70 E	-
-	uncharacterized protein LOC109871595	A 268 G	T 90 A	
-	unknown gene	T 65 C C 150 G T 166 C	V 22 A D 50 E W 56 R	-

780 * amino acid substitutions predicted to impair/alter protein function; ** SNP position refers to the SCO alignments.

781

782

783

784

785

Table 3: Comparison of proportional assignment of candidate SCO data among *C. tshokwe* and *C. compressirostris* for KEGG level A categories.

KEGG level A category	**Species-wise comparison of the total number of candidate SCO-sequences assigned to the respective level A category (Fisher Exact Test; p-values)**	**Species-wise comparison of candidate SCO-sequence distributions in a KEGG level A category according to the KEGG level B assignment (Chi^2 Test; p-values)**
Metabolism	0.691	0.090
Genetic Information Processing	0.877	0.009
Environmental Information Processing	0.034	0.578
Cellular Processes	0.031	< 0.001
Organismal Systems	< 0.001	0.007
Human Diseases	0.691	0.008

786

Figure 1

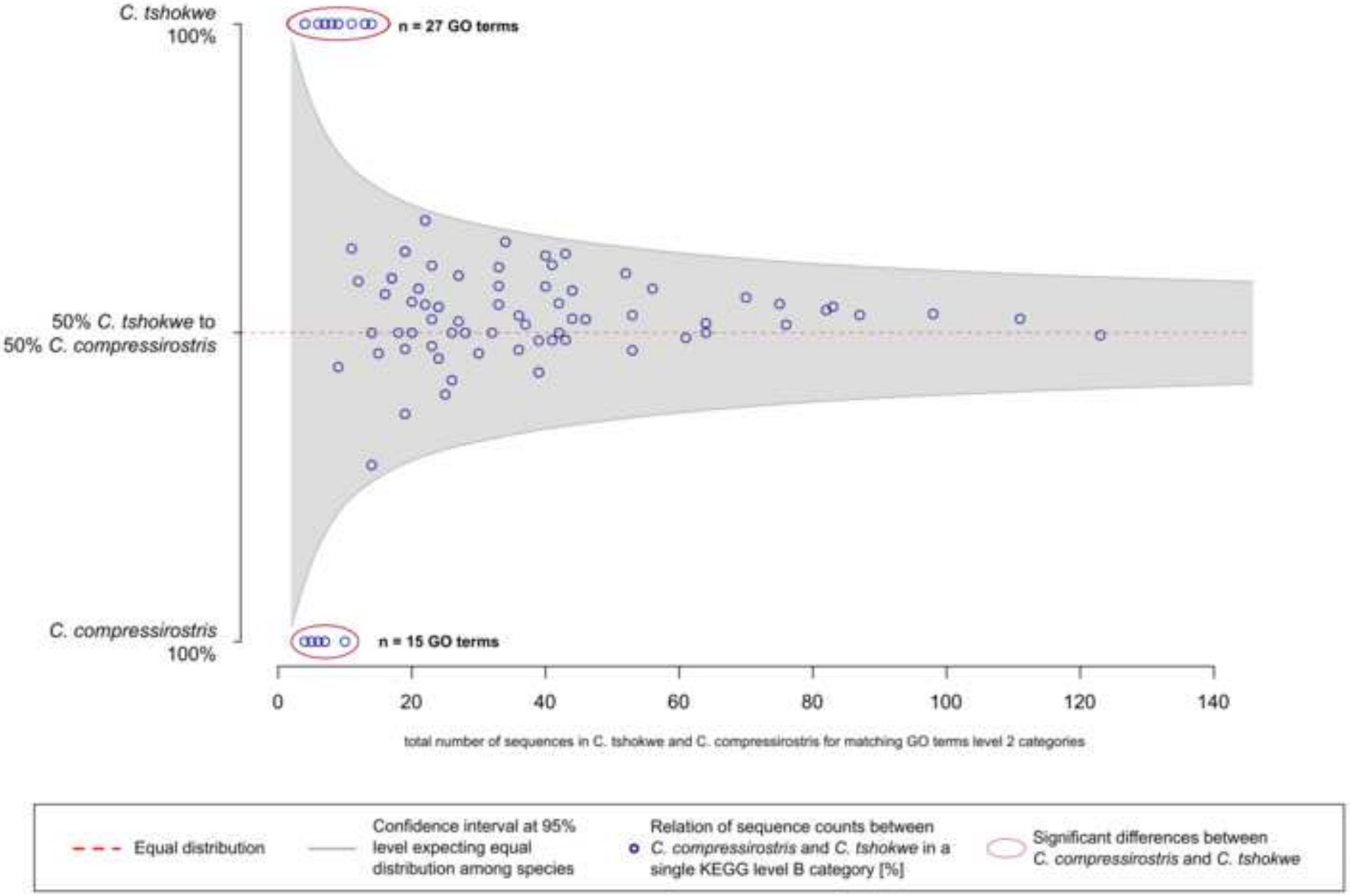


Figure 2

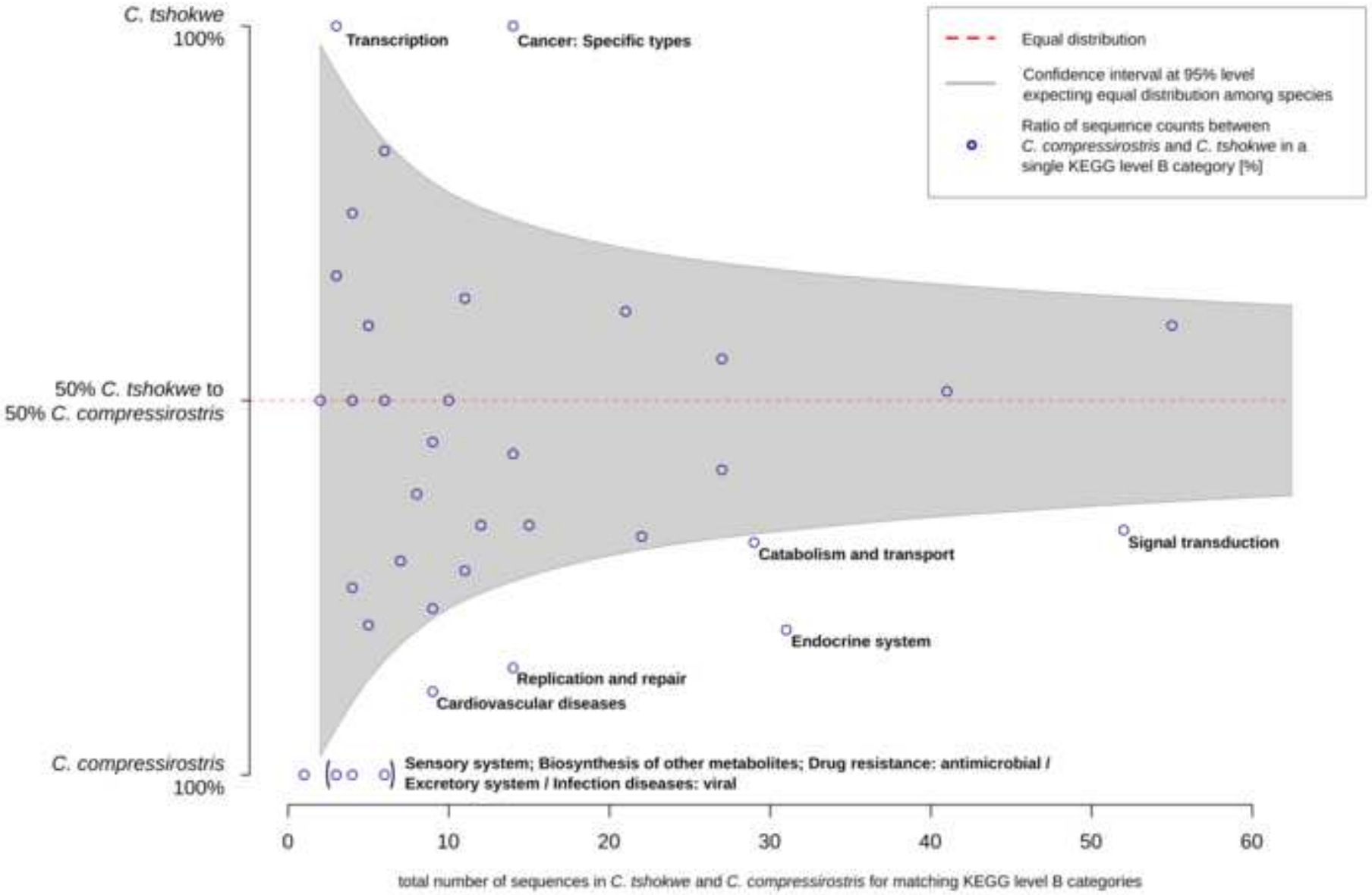


Figure 3

Click here to access/download;Figure;Canitz-PlosOne-Figure3.png

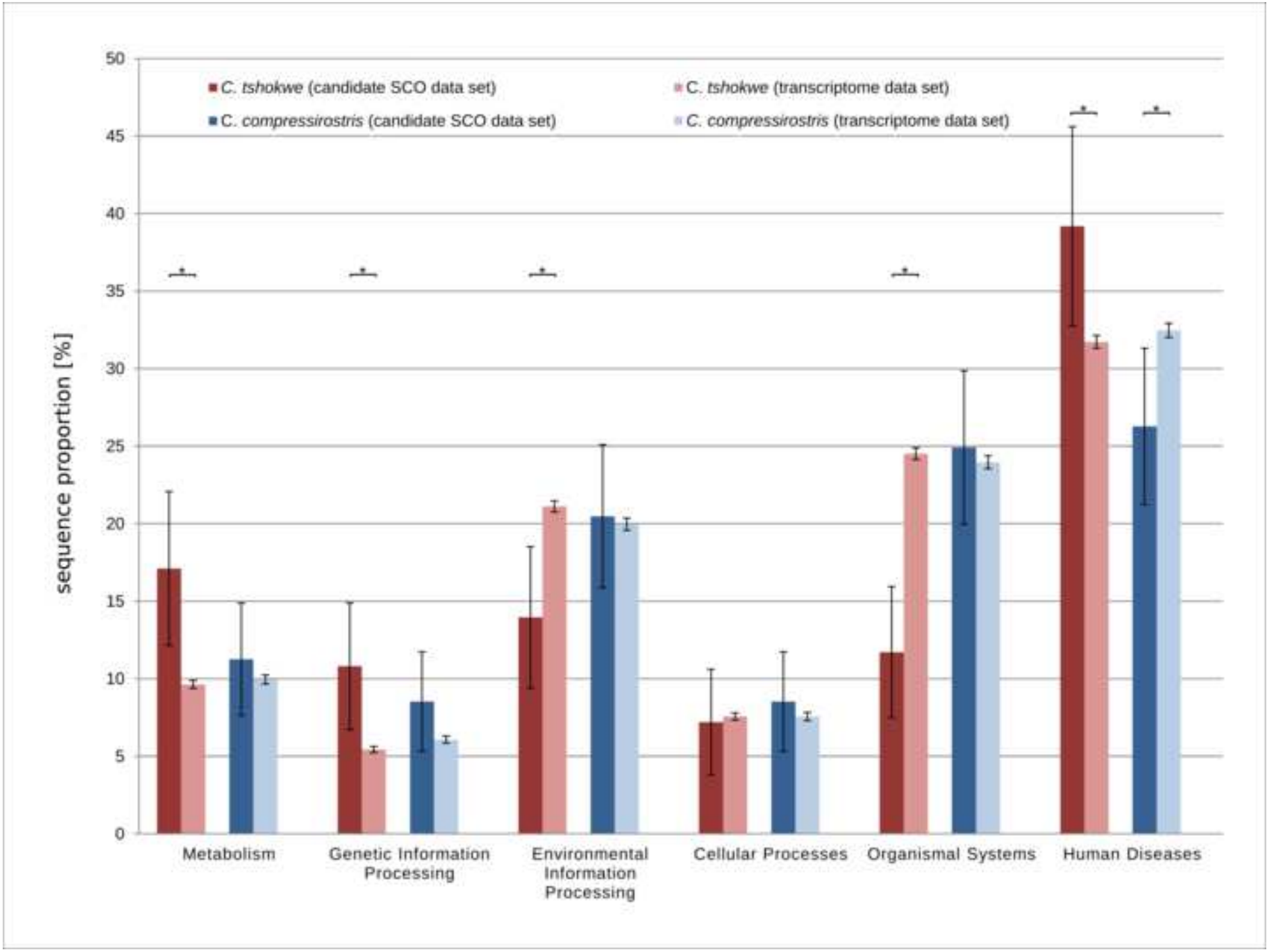
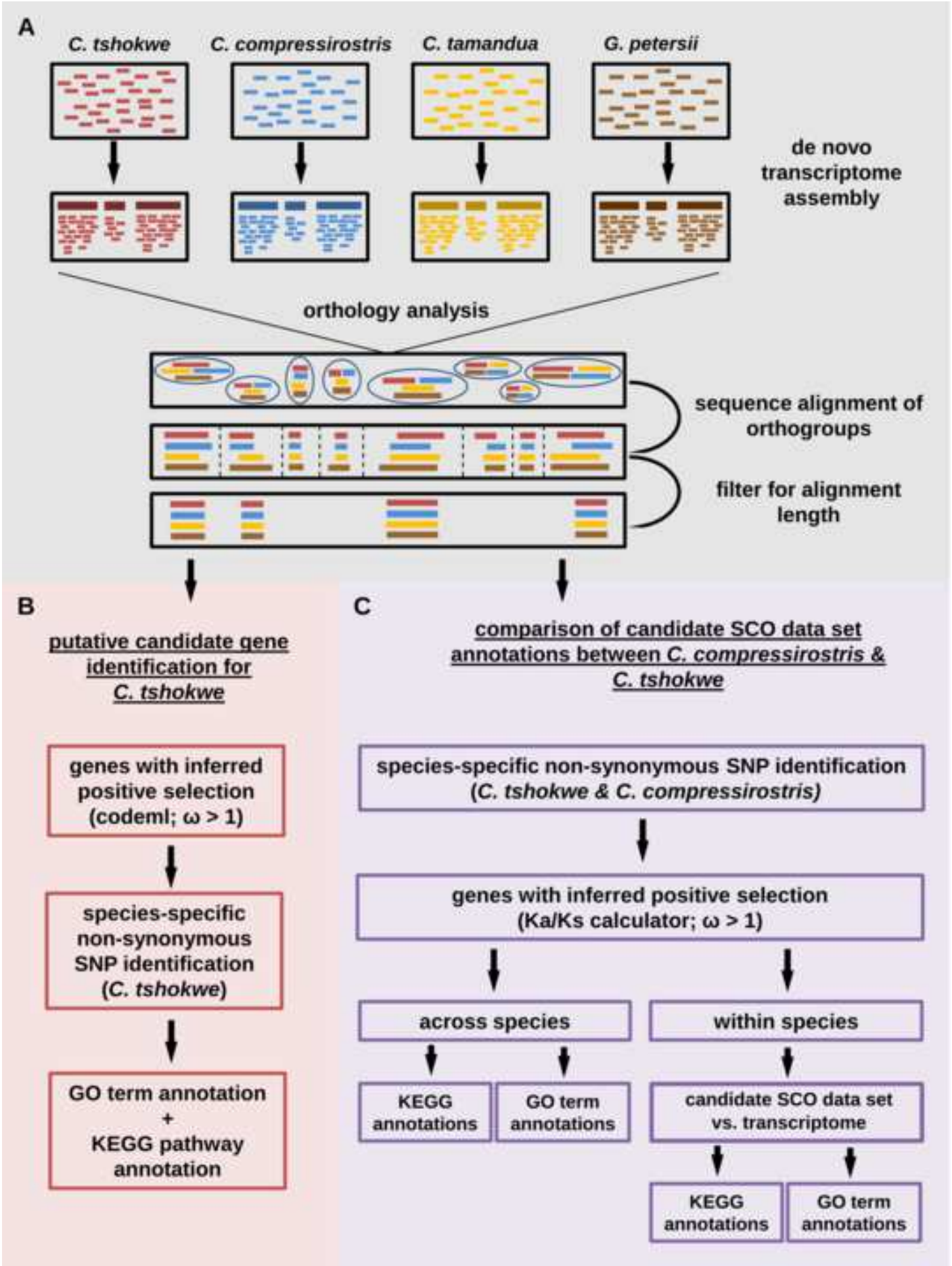


Figure 4

Click here to access/download;Figure;Canitz-PlosOne-Figure4.png



5 Article III

Genome-wide SNPs provide new insights into the phylogeny of the African weakly electric mormyrid fish genus *Campylomormyrus* (Teleostei; Osteoglossiformes)

Julia Canitz, Frank Kirschbaum, Viktor Mamonekene, Ralph Tiedemann

to be submitted to

Molecular Phylogenetics and Evolution

Genome-wide SNPs provide new insights into the phylogeny of the African weakly electric mormyrid fish genus *Campylomormyrus* (Osteoglossiformes, Teleostei)

Julia Canitz[1], Frank Kirschbaum[2], Viktor Mamonekene[3], Ralph Tiedemann[4*]

[1] Julia Canitz
Institute of Biochemistry and Biology/ Unit of Evolutionary Biology/ Systematic Zoology
University of Potsdam
Karl-Liebknecht Straße 24-25, 14776 Potsdam
jcanitz@uni-potsdam.de

[2] Prof. Frank Kirschbaum
Faculty of Life Sciences, Chair: Biology and Ecology of Fishes
Humboldt University of Berlin
Philippstraße 13 (Haus 16)
frank.kirschbaum@staff.hu-berlin.de

[3] Dr. Victor Mamonekene
Institut de Développement Rural, Université Marien Ngouabi,
B.P. 69 Brazzaville, Republic of Congo
vito.mamonekene@gmail.com

[4] Prof. Ralph Tiedemann ([*]corresponding author)
Institute of Biochemistry and Biology/ Unit of Evolutionary Biology/ Systematic Zoology
University of Potsdam
Karl-Liebknecht Straße 24-25, 14776 Potsdam
tiedeman@uni-potsdam.de

Abstract

The African weakly electric fish genus *Campylomormyrus* belongs to the species-rich family which is capable of electrogenesis. In the past, the exact species number of *Campylomormyrus* was highly discussed ranging two to 16 with currently 15 accepted species. The most recent phylogeny is only based on three genetic markers and just 10 species are included which possess a species-specific electric organ discharge (EOD). We aim in this study a comprehensive analysis based on the ddRAD sequencing method to obtain genetic markers (Single Nucleotide Polymorphisms) that cover a large part of the genome. Furthermore, two species (*Campylomormyrus phantasticus* and *Campylomormyrus cassaicus*) not yet part of the phylogeny are now included. The phylogenetic reconstruction reveals two main clades. *C. phantasticus* is positioned as sister group to one of the main clades and *C. cassaicus* belongs to the *alces*-complex. This species complex is not fully resolved by former studies and its specimens possess different EODs. In applying an individual-based approach we obtain well-supported relationships among individuals within the *alces*-complex (bootstrap>90 and posterior probabilities=1) although they are genetically similar. As those specimens cluster neither by species nor by EOD waveforms, incipient sympatric speciation is hypothesized to explain this pattern. Moreover, one single specimen (K71), originally identified as *C. elephas*, shows an outstanding phylogenetic position within the *alces*-complex and might better classified as unknown species. The main difference to previous phylogenies is the position of *C. tamandua* which is now the sister group to the *alces*-complex and not the most basal taxon within *Campylomormyrus*.

In conclusion, the phylogenetic analysis inferred by genome-wide distributed differences aids at resolving relationship uncertainties among single individuals. They further support current species classification, but unravel a discordance among different classification methods (e.g. by genetics and EOD waveforms). A taxonomic revision of the genus *Campylomormyrus* is therefore necessary, as new information have been continuously revealed during the last decade, in particular through genetic analysis.

Keywords: ddRAD sequencing, *Campylomormyrus*, phylogeny

1. Introduction

The African weakly electric fish family Mormyridae is the most diverse lineage within the basal teleost order Osteoglossiformes. The family is monophyletic, comprises more than 200 species in 20 genera, and occurs exclusively in the freshwater systems of Africa (Daget et al., 1984; Lavoué and Sullivan, 2004; Poll, 1945). A specialized electric organ enables mormyrid fish to generate electric organ discharges (EODs) which are used for orientation and communication (Arnegard et al., 2006; Bass, 1986a; Moller, 1995; Stoddard and Markham, 2008). Among mormyrid species, EOD waveforms vary in duration, phase number, and orientation (Bass, 1986b; Hopkins, 1995). Thus, it is proposed to be a species-specific trait and assists in species classification (Friedman and Hopkins, 1996; Hopkins, 1980). In some mormyrids, sex dimorphism and population differences of EOD waveforms are reported (Bass and Hopkins, 1985; Gallant et al., 2017b) but among *Campylomormyrus* species, the target genus of this study, those variations have not yet been observed (Feulner et al., 2008).

Campylomormyrus occurs mainly in the Congo Basin and is characterized by an elongated snout whose differences belong to the major morphological distinctions. In the Lower Congo River, most of its species occur in sympatry but are supposed to be reproductive isolated. The EOD is used for species discrimination and mate recognition and serves likely as prezygotic reproductive isolation mechanism which promotes assortative mating (Feulner et al., 2008; Nagel et al., 2018; Tiedemann et al., 2010).

Campylomormyrus species were first described morphologically at the end of the 19th century (Bleeker, 1874). Since then, the number of species varied between two and 16 (Boulenger, 1898; Poll, 1945; Taverne, 1968). Today, 15 species are accepted and can be discriminated by morphology, EOD waveform and/or genetics. In first phylogenetic analyses monophyletic groups correspond to known species but for several individuals the species affiliation is ambiguous (Feulner et al., 2007; Feulner et al., 2008). However, the current well supported species tree represents the main relationships of *Campylomormyrus* species (Lamanna et al., 2016). It includes nine accepted and one yet undescribed species (Figure 1) and is based on three genetic markers (*Cyt b*, *rps7* and *scn4aa*). Species delimitations are supported by microsatellite data, morphometric measurements, and EOD waveform differences. The species tree states *C. tamandua* as the most basal taxon. Due to

its phylogenetic position, its short triphasic EOD waveform is proposed to be the ancestral state implying that a bi-phasic waveform evolved after the divergence of *C. tamandua*.

The other nine species are divided in two main clades (Figure 1). One clade comprises the species *C. alces*, *C. christyi* and *C. elephas* (*alces*-complex) while the other one includes the six remaining species (Figure 1). Within species of the *alces*-complex, individuals are morphologically similar but show several differences in EOD characteristics like phase number and duration (unpublished data), and the species cannot be distinguished due to microsatellite data (Lamanna et al., 2016). In contrast, species delimitation in the second main clade (Clade 1) is consistent among the different approaches except for the two EOD morphs of *C. compressirostris* (EOD morph 1 is bi-phasic while EOD morph 2 is tri-phasic) (Figure 1). Clade 1 is further separated in two sub-clades and the undescribed species *C. n.sp.*. *Campylomormyrus rhynchophorus* and *C. numenius* (both with an elongated EOD) form one sub-clade, while *C. tshokwe* (intermediate long EOD), *C. curvirostris* (short EOD) and *C. compressirostris* (short EOD) are part of the second sub-clade. All of these species possess a bi-phasic EOD (except *C. compressirostris* EOD morph 2). Although the species tree represents well supported phylogenetic relationships, there are still uncertainties between genetic, morphological, and EOD classifications (Feulner et al., 2008; Lamanna et al., 2016; Tiedemann et al., 2010). This aspect raises the question whether those differences reflect the existence of cryptic species or certain species are not as reproductively isolated as expected.

Figure 1: Current species tree of nine *Campylomormyrus* species. The tree depicts the phylogenetic relationships among nine *Campylomormyrus* species. A morphological type example and a generic EOD record are depicted for each species. Modified from Lamanna et al. (2016).

Since the development of Next Generation Sequencing (NGS) methods, phylogenetic analyses are no longer restricted to single genes representing only a nominal part of the genome. A specialized NGS method is the double digest restriction enzyme associated DNA sequencing (ddRADseq) (Peterson et al., 2012). This approach allows the sequencing of thousands of genetic markers distributed throughout the genome. It is based on the DNA digest by two restriction enzymes and enables the comparison of DNA fragments among

multiple individuals of closely related species. The method was initially used to infer population structures but its application on the species level is increasing (Alter et al., 2017; Chattopadhyay et al., 2016; Rancilhac et al., 2019; Xu et al., 2017).

From an evolutionary perspective, those genomic approaches have changed the view on speciation as it turned out that gene flow/hybridization among closely related species occurs more often than expected during and after speciation (Campbell et al., 2018; Foote, 2018; Han et al., 2017; Nosil, 2008). There are different scenarios explaining sympatric speciation by either no, neutral or aided (secondary) gene flow. In the latter case, gene flow may trigger sympatric divergence or promotes previous initial divergence in sympatry. Furthermore, secondary contact after allopatric occurrence can lead to coexistence or reinforcement if species do not merge to an admixed population (reviewed in preprint Richards et al., 2018). Among cichlid fish, there are some evidences of introgression in the past even in long diverged species (Martin, 2012; Meier et al., 2017; Wagner and McCune, 2009). In addition, genes contributing to reproductive isolation are rare and distributed over the entire genome. Hence, investigations of genome architecture and its link to species divergence is necessary for a better understanding of sympatric speciation in the face of gene flow, introgression, and/or hybridization (Campbell et al., 2018; Harrison and Larson, 2014; Payseur and Rieseberg, 2016).

This study aims at the resolution of uncertain relationships among *Campylomormyrus* species with different EOD waveforms and the ascertainment of the phylogenetic position of two additional species which have not yet been included in phylogenetic reconstructions (*C. phantasticus* and *C. cassaicus*). By performing ddRAD sequencing, we presume that the usage of a highly increased number of genetic markers will clarify the relationships among species. Furthermore, we expect that a genome-wide approach provides more insights into evolutionary mechanisms driving speciation in the genus *Campylomormyrus*.

2. Material and methods

2.1 Sampling and laboratory work

Tissue samples of *Campylomormyrus* and *Gnathonemus* specimens were collected from two samples sites at the Lower Congo River near Brazzaville/Kinshasa during three field trips in

2004, 2006, and 2012 (Supplementary Table 1). DNA of 109 individuals was extracted using the Qiagen Dneasy Blood & Tissue Kit following the manufacturer's instructions. With a Quibit 3 Fluorometer (Invitrogen/ ThermoFisher) and the 4200 TapeStation System (Agilent) DNA concentrations and fragment length distributions were measured. 65 samples with high quality DNA were used for ddRAD sequencing covering the twelve *Campylomormyrus* and one *Gnathonemus* species (Supplementary Table 1). Library preparation was adapted to the protocol of Peterson et al. (2012) and Illumina adapters were modified by an insertion of four randomly chosen nucleotides in front of the inline barcode (5-7bp length). Library preparation is described briefly: 500ng DNA per sample were digested by the restriction enzymes EcoRI and MspI (New England Biolabs); modified adapters were ligated to the digested DNA fragments; samples were pooled equimolar by Illumina index sequence; size-selection of DNA fragments (300-450bp) was conducted by gel extraction (Macherey-Nagel PCR Cleanup Kit); Illumina adapters were completed via PCR attaching different Illumina indexes to DNA fragments of the respective sample pool. Library molarity was calculated by DNA concentration (Qubit 3 Fluorometer) and average fragment length (4200 TapeStation). Sequencing was performed using the Illumina NextSeq 500/550 machine at the Adaptive Genomics group (University Potsdam) to generate paired-end reads of 150 bp.

2.2 Sequence clean-up and read assembly

Raw reads were adapter-clipped using Cutadapt (Martin, 2011) and reads shorter than 32bp were removed from the datasets. The four bases which had been inserted before the inline barcode were trimmed by using a customer bash script. FastX Toolkit 0.0.13 (Hannon, 2010) was performed for quality and artifact filtering. Finally, reads were de-multiplexed by inline barcode using the package *process_radtags* implemented in Stacks 2.2 with default settings (Catchen et al., 2013). Reads shorter than 128 bp were removed. The remaining de-multiplexed reads were assembled following the reference-aligned pipeline of Stacks 2.2 (ref_map.pl). Bowtie2-2.2.9 (Langmead and Salzberg, 2012) mapped reads of each sample separately to the reference genome of *Paramormyrops kingsleyae* (Gallant et al., 2017a) with default parameters. Mapping results were controlled with Samtools 1.3.1. Two samples of *C. tamandua* (C08 and C105) were excluded from the data set as read numbers and mapping qualities were low. The package *gstacks* (Stacks 2.2) was applied to build a catalog

of RAD tags which were represented in the entire data set. Afterwards, *gstacks* mapped the reads of each sample back to the RAD tag catalog with a minimum mapping quality of 8 (--min-mapq 8) and genotyped them. Subsequently, the *populations* script (Stacks2.2) was executed to filter for loci shared among a given number of species. Based on this principle, we created two data sets with a different species number. In both data sets, all parameters were set to default except for minimum allele frequency (--min-maf 0.05) and the minimum proportion of individuals in a species to process a locus (-r 0.5). For an individual approach, the program was told to treat each of the 63 individuals as different 'species' and filter for loci shared by a minimum number of 60 'species' (-p 60; individual-based data set). For the other data set, individuals were assigned to the 13 known *Campylomormyrus* species (considering both EOD morphs of *C. compressirostris* as separate species) and one *Gnathonemus* species, and the *populations* script was executed with a minimum species number covering a locus of 13 (-p 13; species-based data set). In both data sets, the *populations* script called only one SNP per locus (--write-single-snp), produced outputs in multiple formats (i.e. structure, vcf and phylip), and generated several statistic summaries of population genetics like observed heterozygocity or nucleotide diversity.

2.3. Clustering analysis regarding all species

In order to identify species clusters within *Campylomormyrus*, we conducted a Principal Component Analysis (PCA), followed by a K-means clustering approach with an increasing number of clusters which was subsequently implemented in a Discriminant Analysis of Principle Components (DAPC). Those analyses were executed by functions implemented in the R package adegenet 2.1.1 (Jombart, 2008; Jombart and Ahmed, 2011). It ran with default parameters and 300 axis were retained. K-means algorithm calculated the best fitting cluster number for the given data set by computing the Bayesian Information Criterion (BIC). K-means clustering was executed with 10000 iterations, a maximum cluster number of 20 and 100 randomly choses starting centroids. As the computed clustering number (k = 8) deviated from our expectations (k = 14) we obtained the species-cluster matrix for both k values.

231 2.3 Phylogenetic analyses
232
233 The *populations* script outputted concatenated SNPs fixed within a population as well as
234 concatenated biallelic SNPs encoded by the IUPAC nucleotide code, as PHYLIP format.
235 Both variants of the individual- and species-based data sets were used to compute
236 phylogenetic relationships, as they contained different number of informative sites. RaxML
237 7.2.8 (Stamatakis, 2006) implemented in Geneious R8.1.9 was used to reconstruct the
238 phylogeny by maximum likelihood (ML). The substitution model was set to GTR+CAT and
239 *Gnathonemus petersii* served as outgroup. We chose the algorithm of rapid bootstrapping
240 with searching for best-scoring ML tree. The program ran with 1000 bootstrap replicates.
241 Furthermore, a bayesian inference (BI) approach was performed on the same data sets with
242 MrBayes (Geneious R8.1.9, Hülsenbeck and Ronquist, 2001) using the GTR+GAMMA
243 substitution model, 1,000,000 generations running with four heated chains and a burn-in of
244 100,000 generations. Again *G. petersii* was used as outgroup.
245
246 2.4 Analyses of the main clades
247
248 We performed the same Stacks2.2 pipeline for the two sub-sets comprising one of the main
249 clades inferred by the phylogenetic analyses, respectively (Clade 1 and *alces*-complex), to
250 gain insights into the species structure of each clade. The *populations* script was executed
251 with a minimum number of species covering a locus of 4 (-p 4; *alces*-complex) and 6 (-p 6;
252 Clade 1). The clades were analyzed by PCA, K-mean clustering, and DAPC using the R
253 package adegenet.
254
255
256 **3. Results**
257
258 3.1 ddRAD sequencing and SNP calling
259
260 The sequencing of 63 *Campylomormyrus* individuals resulted in a total of 222.1M raw
261 reads. After adapter clipping, quality and artifact filter, and discarding short sequences, the
262 read number has been decreased to 154.99M reads (69.78%). After the de-multiplexing by
263 inline barcode, the average number of reads per individual ranges from 194K to 3.91M with

a mean of 1.87M (± 881K SD). On average, 61.97% ± 13.9% of reads per sample map to the references genome of *P. kingsleyae* with a mean mapping quality of 32.7 ± 1.01 (samtools phred score) and a mean coverage depth of 9.43 ± 3.55.

Table 1: SNP calling statistics of the Stacks reference-aligned pipeline for the four different data sets

	individual-based data set	**species-based data set**	***alces*-complex**	**Clade 1**
number of shared loci in at least 2 species/individuals	125,833	106,307	82,807	110,858
number of loci passed population filter (--min-maf; -p; -r)	33,269	50,050	65,688	68,979
# of fixed SNPs	1,794	3,203	311	1,976
# of variable sites	4,329	12,109	20,959	20,669
missing data [%]	13.55	2.89	21.0	18.5

Among the 77.67M mapped read records, the individual-based data set detect 125,833 and the species-based 106,307 RAD loci which are shared by at least two species (Table 1). Due to the parameter setting, the *populations* scripts filtered 33,269 and 50,050 RAD loci for the individual-based and species-based data set, respectively. The shared RAD loci of the individual-based data set contained 1794 fixed SNPs within each individual, but different between at least two of them, 4329 variable sites (biallelic and fixed SNPs) and include 13.55% missing data. In contrast, the number of fixed SNPs and variable sites in the species-based data set is 3203 and 12,109, respectively, and only 2.89% of the data are missing (Table 1). Genomic statistics inferred by Stacks2.2 for the species-based data set reveal a similar number of informative sites, observed heterozygocity (H_{obs}), and nucleotide diversity (π) among species. For all species the analysis computes a nominal negative inbreeding coefficient (F_{IS}) (Table 2).

Table 2: Summary of genetic diversity statistics for the species-based data set calculated by Stacks2.2

Taxon	**N**	**sites**	**% poly**	**H_{obs}**	**π**	**F_{IS}**
C. compressirostris Morph 1	4.85	3,075	0.051	0.0002	0.0002	-0.00001
C. compressirostris Morph 2	4.89	4,874	0.081	0.0003	0.00029	-0.00004

C. curvirostris	5.63	5,182	0.086	0.00028	0.00027	-0.00002
C. tshokwe	5.71	3,197	0.053	0.0002	0.00018	-0.00005
C. numenius	10.41	5,640	0.094	0.00028	0.00026	-0.00006
C. rhynchophorus	5.75	3,460	0.057	0.00027	0.00023	-0.00009
C. n.sp.	1.92	3,225	0.054	0.00047	0.0004	-0.00011
C. phantasticus	1.95	2,137	0.035	0.00022	0.00021	-0.00001
C. alces	3.76	2,754	0.046	0.00023	0.00022	-0.00002
C. elephas	3.82	2,888	0.048	0.00025	0.00023	-0.00003
C. christyi	3.82	4,161	0.069	0.00039	0.00033	-0.0001
C. cassaicus	1	2,053	0.036	0.00036	0.00036	-
C. tamandua	5.74	4,190	0.070	0.00045	0.00032	-0.00023
G. petersii	1	1,278	0.012	0.00023	0.00023	-

N = mean number of individuals genotyped at each locus; sites = number of polymorphic sites across the data set; % poly = proportion of polymorphic loci; H_{obs} = average observed heterozygocity per species; π = average nucleotide diversity; F_{IS} = average inbreeding coefficient

3.2 Phylogenetic analyses

The phylogenetic reconstruction of the individual-based data sets (1,794 fixed SNPs and 4,329 variable sites) are concordant among analysis approaches (ML versus BI), but the topology differs slightly among data sets. As several splits in the phylogenetic tree of the fixed SNP data set are not supported (BS < 90) and some relationships are polytomous (Supplementary Figure 1), the tree topology inferred from 4,329 variable sites of the individual-based data set is described here (Figure 2). The tree is separated in two main clades. One clade includes the species *C. alces*, *C. cassaicus*, *C. elephas*, and *C. christyi* (*alces*-complex), in which a distinct split of *C. christyi* is shown. Individuals associated with *C. alces* and *C. elephas* are mixed and *C. cassaicus* is positioned among them. Its sister group *C. tamandua* is separated by a long branch representing a high number of substitutions since divergence. All major splits in this main clade are well supported by bootstrap values (BS > 90) and posterior probabilities (PP > 0.9; Figure 2). Within the second main clade (Clade 1), individuals being classified as one species are clustered in monophyletic groups (BS > 90, PP > 0.9; Figure 2). Clade 1 is divided in two sub-clades and

C. phantasticus which splits first from the other species. One sub-clade comprises *C. numenius* and *C. rhynchophorus* and the other is formed by *C. n.sp.*, *C. curvirostris, C. tshokwe,* and *C. compressirostris*. Individuals of *C. compressirostris* possessing difference EOD morphs are mixed within its monophyletic group (Figure 2).

Figure 2: Phylogenetic reconstruction based on the 4,329 variable sites of the individual-based data set. Branch labels indicate bootstrap values (BS > 90) and posterior probabilities (PP > 0.9) of the major clades. Scale represents substitutions per site.

The main topology of the species-based phylogeny is consistent among data sets and analysis approaches (Supplementary Figure 2). Only the position of *C. tshokwe* changes in the ML phylogeny inferred from the fixed SNPs but this relationship is not supported (BS < 50). In comparison to the inferred topology of the individual-based data sets, the position of *C. n.sp.* is different and segregates after the split of *C. phantasticus*. The topology is well supported at all nodes of the species-based tree, except for the separation of *C. curvirostris* and *C. compressirostris* (Supplementary Figure 2). The branch length of *C. compressirostris* (both EOD morphs), *C. curvirostris*, *C. alces*, and *C. elephas* are remarkable short compared to other branches which lead to extant species.

3.3 Clustering analyses regarding all species

The PCA analysis yields a similar pattern as shown by monophyletic groups in the phylogeny (Figure 3A). The first two principle components (PC) explain more than 50% of variation among species separating the samples in four clusters which represent one to seven taxa. A segregation of the individuals by the first and third component (~ 44.0%) structures the data set in six clusters with an overlap of individuals of the *alces*-complex as well as an overlap of both *C. compressirostris* morphs with *C. curvirostris* individuals (Figure 3A).
The K-means clustering analysis inferred k = 8 as best fitting cluster number (lowest BIC value; Supplementary Figure 3A). Individuals of both *C. compressirostris* morphs and *C. curvirostris* share one cluster, but are also represented in a second one together with *C. n.sp.* (Figure 3B). A similar pattern occurs for species of the *alces*-complex. The four species are assembled to one cluster by the K-means algorithm (Figure 3B). *C. phantasticus* and *G. petersii* share a cluster which is consistent with the PCA analysis. To compare the best

fitting cluster number with the expected one (k = 14), the cluster-species matrix for 14 clusters is also plotted (Figure 3C). In this case, the cluster corresponds approximately to the described species, but both morphs of *C. compressirostris* share one cluster, individuals of *C. numenius* are split in two clusters, and the individuals of the *alces*-complex are mixed and divided into two clusters (Figure 3C).

Figure 3: Principle Component Analysis (PCA) and species-cluster matrix of 13 *Campylomormyrus* and one *Gnathonemus* species. (A) The first two components of PCA are plotted. Dots represent individuals and are colored by species (see legend). (B) Species assigned to the best fitting cluster number inferred by K-means algorithm (k=8) and (C) to the expected number of clusters (k=14). Sizes of squares represent number of individuals (see legend).

The DAPC results in a similar pattern in which the individuals of *C. tamandua*, *C. tshokwe*, *C. rhynchophorus*, and *C. numenius* form well separated groups. Resolutions within the *alces*-complex as well as the separation of *C. compressirostris*, *C. curvirostris*, and *C. n.sp.*, cannot be deduced by this data set (Supplementary Figure 4).

3.4 Analyses of separate clades

Running *gstacks* for the separated clades reveals 82,807 and 110,858 RAD loci shared by at least two species for the *alces*-complex and Clade 1, respectively. After executing the *populations* script 65,688 (*alces*-complex) and 68,979 (Clade 1) RAD loci passed the filter parameters. Among loci shared by species of the *alces*-complex, 311 fixed SNPs and 20,669 variable sites with 21% missing data are found. 1,976 fixed SNPs and 20,959 variable sites have been called in Clade 1 including 18.5% missing data (Table 1).

Within the data set of the *alces*-complex, K-means clustering result in a best fitting k-number of 2 (lowest BIC values; Supplementary Figure 3B). The two clusters separate *C. christyi* from the other species (*C. alces, C. elephas, C. cassaicus*). Setting the cluster number to the expected species number (k=4) yields a single cluster for *C. cassaicus* and *C. christyi*, but individuals of *C. elephas* and *C. alces* are merged except for the specimen K71 which represents the fourth cluster (Figure 4A). PCA and DAPC analyses separate the species in three groups whereby *C. cassaicus* and *C. christyi* are well segregated and

individuals classified as *C. alces* and *C. elephas* overlap (Figure 4A, Supplementary Figure 5).

Figure 4: Results of PCA and cluster analysis regarding Clade 1 and the *alces*-complex. (A) The first two principle components of the PCA are plotted above, explaining 34.9% of variation in the *alces*-complex and the species-cluster matrix for k=2 (best fitting k) and k=4 (expected k) is shown below. (B) Results of the PCA analyses for Clade 1 are plotted above in which the first two axis represent 27.8% of variation. Below, the species-cluster matrix for k=6 (best fitting k) and k=8 (expected k) are shown. Square sizes in the matrices represent the individual number of each species in the particular cluster.

For Clade 1, k=6 fits best to the data according to K-means clustering algorithm (lowest BIC value; Supplementary Figure 3C). The species *C. tshokwe, C. numenius, C. rhynchophorus*, and *C. phantasticus* form single clusters, respectively. The individuals of *C. compressirostris* (morph 1) and *C. curvirostris* are split in two clusters sharing it once with *C. compressirostris* morph 2 and once with *C. n.sp.* (Figure 4B). Setting the cluster number to the expected number of species (k=8) results in a species-wise segregation except for both *C. compressirostris* morphs which share a cluster, while *C. numenius* is split into two (Figure 4B). The PCA and DAPC indicate a similar pattern with an overlap of both *C. compressirostris* morphs and *C. curvirostris*, and a distinct separation of the other species (Figure 4B, Supplementary Figure 6).

4. Discussion

It is the first time that a phylogenetic reconstruction of *Campylomormyrus* includes the species *C. phantasticus* and *C. cassaicus*. Our results demonstrate that *C. phantasticus* is positioned as basal lineage within one of the main clades representing the sister taxon to most of *Campylomormyrus* species (Figure 2). Although type specimens of *C. phantasticus* are described and EOD records as well as its distribution area is known (Daget et al., 1991; Pellegrin, 1927; Stiassny et al., 2007), an identification of our individuals by a morphological comparison to the type specimens cannot be provided as neither pictures nor EOD records exist. However, we emphasize that the clear separation of two individuals from

all others in the phylogenetic and cluster analysis, means the presence of another distinctive *Campylomormyrus* species in the Lower Congo River being most likely *C. phantasticus*.

The inclusion of new species and the increase of genetic markers infer a new phylogenetic position for *C. tamandua* as sister taxon of the 'alces' group which is supported by tree reconstructions of different data sets (Supplementary Figures 1 and 2). In all phylogenetic trees *C. tamandua* has a remarkably long branch compared to the other species which indicates that the species is genetically highly different compared to others. The early divergence but low diversification of the *C. tamandua* lineage might be due to an under-representation of the species. It is the only species which has apart from the Congo Basin additional distribution areas like parts of the Chad and Niger River (Daget et al., 1984). An incomplete taxon sampling can have those effects in the phylogeny as closely related taxa which would intersect a long branch, are missing (Cummings and Meyer, 2005; Heath et al., 2008).

The first classification of our *C. cassaicus* sample (K56) was obtained via visual comparison to the type specimen described by Poll (1967) (Supplementary Figure 7A). In the phylogeny, it forms a clade with two *C. elephas* individuals that represents the sister group to a clade containing two *C. alces* and one *C. elephas* specimens (Figure 2). These relationships are well supported in the Bayesian analysis (PP=1). However, genetic differences among those individuals are low according to branch lengths. PCA and DAPC clearly separate *C. cassaicus* from other species of the *alces*-complex (Figure 4) and considering a cluster number of 4 supports this segregation. As a sample set of one is less sufficient for distinct inferences of phylogenetic relationships our results can just confirm that *C. cassaicus* is part of the *alces*-complex. The same sample is also included in a previous phylogeny which is based on two genetic markers (Feulner et al., 2008). There it is part of the separated clade E assigned to the species *C. elephas* which is consistent with our results. Based on the genetic data, it can therefore be assumed that *C. cassaicus* belongs to *C. elephas*, which raises the question of whether it is a distinct species. Hence, it can be presumed that *C. cassaicus* is just a specimen of *C. elephas* raising the question if it can be considered as separated species.

From a phylogenetic perspective and regardless of the species affiliation, the 'alces' group is divided into three lineages whereby one is represented by four specimen classified as *C. christyi*, one by eight individuals which are previously assigned to the species *C. alces*, *C. elephas* and *C. cassaicus* and the one containing only the sample K71. The separation of *C.*

christyi is supported by the K-means clustering analysis, as k=2 is suggested to be the best fitting number of clusters and all individuals of *C. christyi* belong to one of the two clusters. Those results support that *C. christyi* is a distinct species (Boulenger, 1898; Lamanna et al., 2016).

The pooling of the eight species-mixed individuals is reasonable, as in comparison to other monophyletic clades representing a species (e.g. *C. rhynchophorus*) the genetic differences among specimen are similar according to branch lengths (Figure 2). Furthermore, clustering analysis with k=2 merges those individuals to one cluster. Even a calculation with the expected number of k=4 isolates only the *C. cassaicus* sample from the group and the PCA and DAPC analyses demonstrate an overlap of *C. elephas* and *C. alces* specimens (Figure 4, Supplementary Figure 5). Hence, due to our results, the samples of *C. elephas* and *C. alces* used in this study do not represent two separate species casting doubts whether those species can be discriminated. Low genetic difference but high EOD diversity within a sympatrically occurring group is also shown for the species *magnostipes*-complex of *Brienomyrus* (Arnegard et al., 2006, 2005). There are several hypotheses including behavioral and ecological selection as driver for EOD diversification. Only the assumption of incipient sympatric speciation takes into account the high genetic similarity among individuals with different EODs (Arnegard et al., 2006). As our genome-wide approach indicates high genetic similarity among individuals of the *alces*-complex with different EOD waveforms, incipient sympatric speciation might be a sufficient explanation as EODs are thought to serve as prezygotic isolation mechanism (Feulner et al., 2008). Furthermore, this would mean that our RAD loci cover no genes which are responsible for EOD waveform differences.

In the individual-based phylogeny a special position is inferred for the sample K71 which diverged even before the split of *C. christyi*. This outstanding position and its segregation by PCA and cluster analysis (k=4) as well as the special coloration (Supplementary Figure 7B) suggests that this individual may represent an unknown species. Among mormyrid fish, species diversity and especially the identification of new and cryptic species are increasing so that even an entire genus has been recently described (Rich et al., 2017).

The usage of a highly increased number of genetic markers for phylogenetic analysis confirms the monophyly of the species *C. rhynchophorus*, *C. numenius*, *C. tshokwe*, *C. compressirostris*, *C. curvirostris,* and *C. n.sp*. which are stated by previous studies (Feulner et al., 2008, 2007; Lamanna et al., 2016), but there are discordances of their phylogenetic

relationships. Additionally, even an increase of genetic markers does not indicate a separation by EOD morph among *C. compressirostris* specimens which can be explained by the same hypotheses mentioned for EOD variation within the *alces*-complex.

Moreover, we want to point out that the usage of ddRAD sequencing and the gathering of thousands of genetic markers lead to the resolution of polytomous relationships among single individuals in a problematic clade like the *alces*-complex (individual splits have a PP=1 and BS>90). It is also important to emphasize that our data show discrepancies among species- and individual-based phylogenies which are a result of species synonymization or miss-assignment. Especially in taxonomic groups where specimens are hard to discriminate by morphology or other features are missing (e.g. EOD records for weakly electric fish), species tree inferences may miss diversity and recent divergence within a species if the affiliation is not clear. Hence, before relationships between species are analyzed, the specimens need to be classified uniquely otherwise it should be kept separately.

In conclusion, our results show that main phylogenetic relationships among *Campylomormyrus* species are consistent with previous studies (Feulner et al., 2008; Lamanna et al., 2016) and that even a good resolution on the individual level is discordant with species classification. Especially the species of the *alces*-complex do not show a clear separation by species which raises the question of their evolutionary history. The remaining and reappearing uncertainties within the phylogeny of *Campylomormyrus* species despite an immense increase of genetic marker indicate that for an informative phylogeny (especially of the *alces*-complex) the sample set needs to be increased, different sampling areas should be considered and if possible, genetic data of the type species should be included. Furthermore, an a priori revision of the *Campylomormyrus* taxonomy would be advantageous as the knowledge about this genus has increased immensely during the last decade.

Author's contributions

JC established the ddRAD sequencing method in the laboratory, conceived and conducted the bioinformatical analyses, and drafted the manuscript. VM supported the field expeditions to Brazzaville, Republic of the Congo. FK participated in sample collection, manuscript

drafting, and supervision. RT participated in sample collection, conceived and supervised the study, and contributed to manuscript drafting. All authors read and drafted manuscript.

Acknowledgements

We thank Francesco Lamanna, Philine Feulner and Christiane Paul for collecting samples during the excursions and our technician Anja R.R. Ernst for her support in the laboratory. We appreciate the comments and advices of Axel Barlow from the Adaptive Genomics group (University of Potsdam) during the establishment of the method. Financial support is acknowledged from the University of Potsdam.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Alter, S.E., Munshi-South, J., Stiassny, M.L.J., 2017. Genomewide SNP data reveal cryptic phylogeographic structure and microallopatric divergence in a rapids-adapted clade of cichlids from the Congo River. Mol. Ecol. 26, 1401–1419.

Arnegard, M.E., Bogdanowicz, S.M., Hopkins, C.D., 2005. Multiple Cases of Striking Genetic Similarity between Alternate Electric Fish Signal Morphs in Sympatry. Evolution (N. Y). 59, 324–343.

Arnegard, M.E., Jackson, S.A., Hopkins, C.D., 2006. Time-domain signal divergence and discrimination without receptor modification in sympatric morphs of electric fishes. J. Exp. Biol. 209, 2182–2198.

Bass, A.H., 1986a. Electric Organs Revisited: Evolution of a Vertebrate Communication and Orientation Organ, in: Bullock, T.H., Heiligenberg, W. (Eds.), Electroreception. John Wiley and Sons, New York, pp. 13–70.

534 Bass, A.H., 1986b. Species differences in electric organs of mormyrids: Substrates for
535 species-typical electric organ discharge waveforms. J. Comp. Neurol. 244, 313–330.

536 Bass, A.H., Hopkins, C.D., 1985. Hormonal control of sex differences in the electric organ
537 discharge (EOD) of mormyrid fishes. J. Comp. Physiol. A 156, 587–604.

538 Bleeker, P., 1874. Description of the Genus Campylomormyrus type sp. Mormyrus
539 tamandua GTHR, VIII. ed. Verslag en Med. K. Ak., Amsterdam.

540 Boulenger, G.A., 1898. Matériaux pour la faune du Congo: Poissons Nouveaux, Annales du
541 Musee du Congo. Bruxelles.

542 Campbell, C.R., Poelstra, J.W., Yoder, A.D., 2018. What is Speciation Genomics? The roles
543 of ecology, gene flow, and genomic architecture in the formation of species. Biol. J.
544 Linn. Soc. 124, 561–583.

545 Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., Cresko, W.A., 2013. Stacks: an
546 analysis tool set for population genomics. Mol. Ecol. 22, 3124–40.

547 Chattopadhyay, B., Garg, K.M., Kumar, A.K.V., Doss, D.P.S., Rheindt, F.E., Kandula, S.,
548 Ramakrishnan, U., 2016. Genome-wide data reveal cryptic diversity and genetic
549 introgression in an Oriental cynopterine fruit bat radiation. BMC Evol. Biol. 16, 1–15.

550 Cummings, M.P., Meyer, A., 2005. Magic bullets and golden rules: Data sampling in
551 molecular phylogenetics. Zoology 108, 329–336.

552 Daget, J., Gosse, J.P., Thys van den Audenaerde, D.F.E., 1984. Check-List of the Freshwater
553 Fishes of Africa. OSTROM, MRAC, Paris, Bruxelles.

554 Feulner, P.G.D., Kirschbaum, F., Mamonekene, V., Ketmaier, V., Tiedemann, R., 2007.
555 Adaptive radiation in African weakly electric fish (Teleostei: Mormyridae:
556 Campylomormyrus): A combined molecular and morphological approach. J. Evol. Biol.
557 20, 403–414.

558 Feulner, P.G.D., Kirschbaum, F., Tiedemann, R., 2008. Adaptive radiation in the Congo
559 River: An ecological speciation scenario for African weakly electric fish (Teleostei;
560 Mormyridae; Campylomormyrus). J. Physiol. Paris 102, 340–346.

561 Foote, A.D., 2018. Sympatric Speciation in the Genomic Era. Trends Ecol. Evol. 33, 85–95.

562 Friedman, M.A., Hopkins, C.D., 1996. Tracking individual mormyrid electric fish in the
563 field using electric organ discharge waveforms. Anim. Behav. 51, 391–407.

564 Gallant, J.R., Losilla, M., Tomlinson, C., Warren, W.C., 2017a. The genome and adult
565 somatic transcriptome of the mormyrid electric fish Paramormyrops kingsleyae.
566 Genome Biol. Evol. 9, 3525–3530.

567 Gallant, J.R., Sperling, J., Cheng, C., Arnegard, M.E., Carlson, B.A., Hopkins, C.D., 2017b.
568 Microevolutionary processes underlying macroevolutionary patterns of electric signal
569 diversity in mormyrid fish. bioRxiv 154047.

570 Han, F., Lamichhaney, S., Grant, B.R., Grant, P.R., Andersson, L., Webster, M.T., 2017.
571 Gene flow, ancient polymorphism, and ecological adaptation shape the genomic
572 landscape of divergence among Darwin's finches. Genome Res. 27, 1004–1015.

573 Hannon, G.J., 2010. FastX Toolkit. http://hannonlab.cshl.edu/fastx_toolkit.

574 Harrison, R.G., Larson, E.L., 2014. Hybridization, introgression, and the nature of species
575 boundaries. J. Hered. 105, 795–809.

576 Heath, T.A., Hedtke, S.M., Hillis, D.M., 2008. Taxon sampling and accuracy of
577 phylogenetic analyses. J. Syst. Evol. 46, 239–257.

578 Hopkins, C.D., 1995. Convergent designs for electrogenesis and electroreception. Curr.
579 Opin. Neurobiol. 5, 769–777.

580 Hopkins, C.D., 1980. Evolution of electric communication channels of mormyrids. Behav.
581 Ecol. Sociobiol. 7, 1–13

582 Hülsenbeck, J.P., Ronquist, F., 2001. MrBayes: Bayesian inference of phylogenetic trees.
583 Bioinforma. Appl. Note 17, 754–755.

584 Jombart, T., 2008. Adegenet: A R package for the multivariate analysis of genetic markers.
585 Bioinformatics 24, 1403–1405.

586 Jombart, T., Ahmed, I., 2011. adegenet 1.3-1: New tools for the analysis of genome-wide
587 SNP data. Bioinformatics 27, 3070–3071.

588 Lamanna, F., Kirschbaum, F., Ernst, A.R.R., Feulner, P.G.D., Mamonekene, V., Paul, C.,
589 Tiedemann, R., 2016. Species delimitation and phylogenetic relationships in a genus of

590 African weakly-electric fishes (Osteoglossiformes, Mormyridae, Campylomormyrus).
591 Mol. Phylogenet. Evol. 101, 8–18.

592 Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat.
593 Methods 9, 357–359.

594 Lavoué, S., Sullivan, J.P., 2004. Simultaneous analysis of five molecular markers provides a
595 well-supported phylogenetic hypothesis for the living bony-tongue fishes
596 (Osteoglossomorpha: Teleostei). Mol. Phylogenet. Evol. 33, 171–185.

597 Martin, C.H., 2012. Weak Disruptive Selection and Incomplete Phenotypic Divergence in
598 Two Classic Examples of Sympatric Speciation: Cameroon Crater Lake Cichlids. Am.
599 Nat. 180, E90–E109.

600 Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing
601 reads. EMBnet.journal 17, 10–12.

602 Meier, J.I., Marques, D.A., Mwaiko, S., Wagner, C.E., Excoffier, L., Seehausen, O., 2017.
603 Ancient hybridization fuels rapid cichlid fish adaptive radiations. Nat. Commun. 8, 1–
604 11.

605 Moller, P., 1995. Electrish fishes: History and Behavior, 17th ed. Chapman & Hall, London.

606 Nagel, R., Kirschbaum, F., Engelmann, J., Hofmann, V., Pawelzik, F., Tiedemann, R., 2018.
607 Male-mediated species recognition among African weakly electric fishes. R. Soc. Open
608 Sci. 5, 170443.

609 Nosil, P., 2008. Speciation with gene flow could be common. Mol. Ecol. 17, 2103–2106.

610 Payseur, B.A., Rieseberg, L.H., 2016. A genomic perspective on hybridization and
611 speciation. Mol. Ecol. 25, 2337–60.

612 Pellegrin, J., 1927. Mormyridés du Cameroun recueillis par M. Th. Monod. Description d'un
613 genre, de quatre espèces et d'une variété. Bull. la société R. Zool. 52, 294–300.

614 Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E., 2012. Double digest
615 RADseq: an inexpensive method for de novo SNP discovery and genotyping in model
616 and non-model species. PLoS One 7, e37135.

617 Poll, M., 1967. Contribution à la faune ichthyologique de l'Angola. Diamang Publ. Cult. 75,
618 381.

619 Poll, M., 1945. Descriptions de Mormyridae · et de Characidae nouveaux du Congo bel avec
620 une étude du genre Stotnator hinus et des genres de Characidae nains africains. Rev.
621 Zool. Bot. Africaines 39, 36–77.

622 Rancilhac, L., Goudarzi, F., Gehara, M., Hemami, M.R., Elmer, K.R., Vences, M., Steinfarz,
623 S., 2019. Phylogeny and species delimitation of near Eastern Neurergus newts
624 (Salamandridae) based on genome-wide RADseq data analysis. Mol. Phylogenet. Evol.
625 133, 189–197.

626 Rich, M., Sullivan, J.P., Hopkins, C.D., 2017. Rediscovery and description of
627 Paramormyrops sphekodes (Sauvage, 1879) and a new cryptic Paramormyrops
628 (Mormyridae: Osteoglossiformes) from the Ogooué River of Gabon using
629 morphometrics, DNA sequencing and electrophysiology. Zool. J. Linn. Soc. 180, 613–
630 646.

631 Richards, E.J., Servedio, M., Martin, C.H., 2018. New criteria for sympatric speciation in the
632 genomic era. bioRxiv 367623.

633 Stamatakis, A., 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses
634 with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.

635 Stiassny, M.L.J., Teugels, G.G., Hopkins, C.D., 2007. The Fresh and Brackish Water Fishes
636 of Lower Guinea, West-Central Africa, Vol. 42. ed. IRD Editions.

637 Stoddard, P.K., Markham, M.R., 2008. Signal Cloaking by Electric Fish. Bioscience 58,
638 415–425.

639 Taverne, L., 1968. Ostéologie du genre Campylomormyrus Bleeker (Pisces
640 Mormyriformes). Bull. la Société R. Zool. Belge 98, 1–41.

641 Tiedemann, R., Feulner, P.G.D., Kirschbaum, F., 2010. Electric organ discharge divergence
642 promotes ecological speciation in sympatrically occurring African weakly electric fish
643 (Campylomormyrus), in: Glaubrecht, M. (Ed.), Evolution in Action: Case Studies in
644 Adaptive Radiation, Speciation and the Origin of Biodiversity. Springer-Verlag Berlin
645 Heidelberg, Berlin, Heidelberg, pp. 307–321.

646 Wagner, C.E., McCune, A.R., 2009. Contrasting patterns of spatial genetic structure in
647 sympatric rock-dwelling cichlid fishes. Evolution (N. Y). 63, 1312–1326.

Xu, T., Sun, J., Lv, J., Watanabe, H.K., Li, T., Zou, W., Rouse, G.W., Wang, S., Qian, P.-Y., Bao, Z., Qiu, J.-W., 2017. Genome-wide discovery of single nucleotide polymorphisms (SNPs) and single nucleotide variants (SNVs) in deep-sea mussels: Potential use in population genomics and cross-species application. Deep. Res. Part II 137, 318–326.

Supplementary

Tables:

Supplementary Table 1: Overview of the 65 samples used for the ddRAD sequencing analysis. Sample ID, classified species, EOD phase number and approximate duration and collection site are listed.

Figures:

Supplementary Figure 1: Phylogenetic analysis based on the fixed and biallelic SNPs of the individual-based data set. Branch labels indicate either bootstrap values or posterior probabilities.

Supplementary Figure 2: Phylogenetic analysis based on the fixed and biallelic SNPs of the species-based data set. Branch labels indicate either bootstrap values or posterior probabilities.

Supplementary Figure 3: Graphical depiction of the K-means clustering analyses. Bayesian information criteria (BIC) are shown for a range of potential clusters. (A) Graph of complete data set; Graphs of the separate analyses regarding (B) *alces*-complex; and (C) Clade 1

Supplementary Figure 4: DAPC plot of the complete data set. The two graphs show the variation based on DA eigenvalues 1 and 2, and 1 and 3, respectively. Dots represent each individual and are colored by species (see legend).

Supplementary Figure 5: DAPC plot of the first two eigenvalues for the separated analysis of the *alces*-complex. Dots represent single individuals and are colored by species (see legend).

Supplementary Figure 6: DAPC plot of the first two eigenvalues for the separated analysis of Clade 1. Dots represent single individuals and are colored by species (see legend).

Supplementary Figure 7: (A) Picture of the specimen K71 collected from the market in Kintele (Brazzaville, Republic of the Congo). (B) Above the type specimen of *C. cassaicus* obtained from http://mormyrids.myspecies.info/en and below the sample K56 classified as *C. cassaicus* and collected from Les Rapides (Brazzaville, Republic of the Congo).

Figure 1

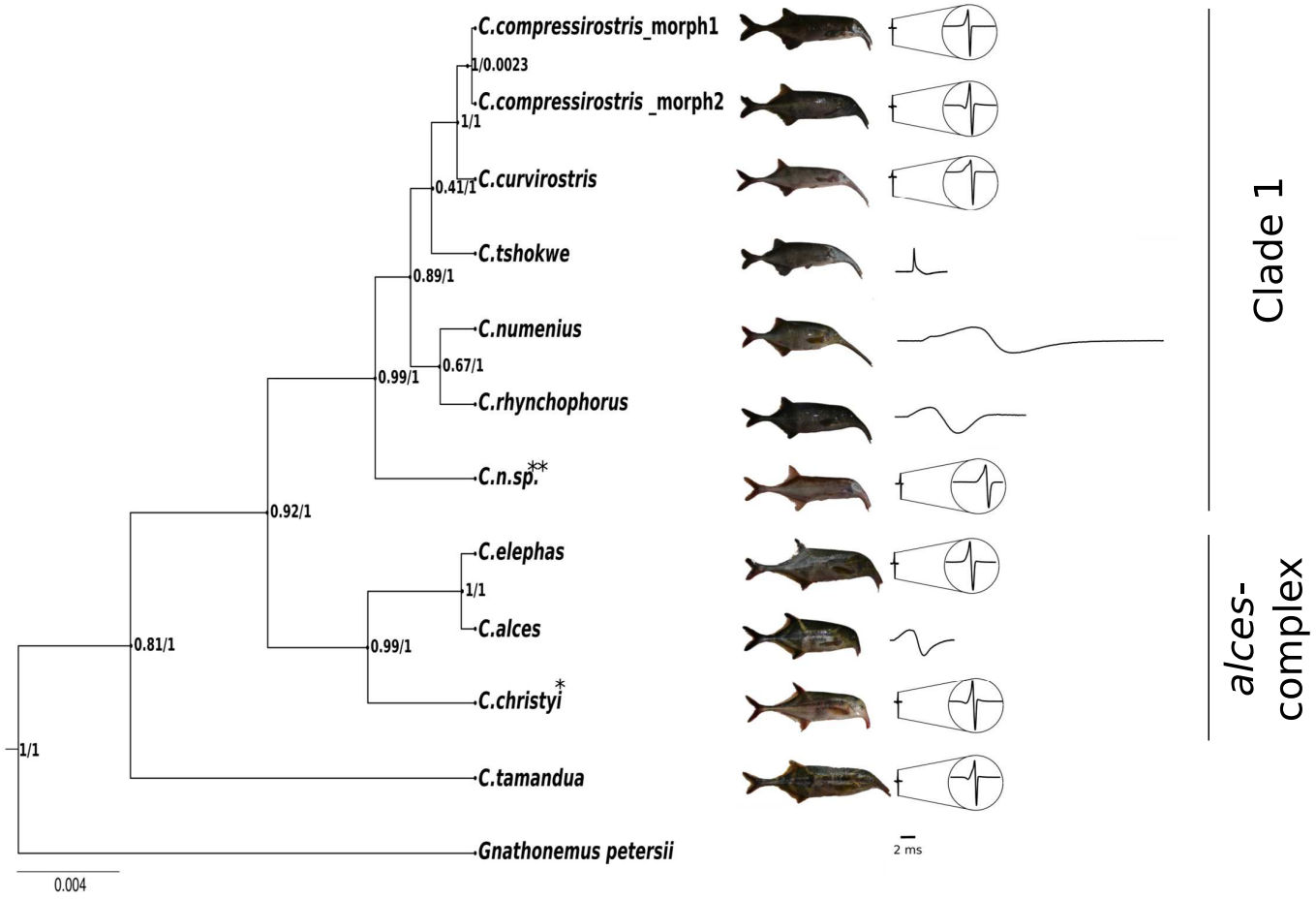

C.compressirostris_morph1
1/0.0023
C.compressirostris _morph2
1/1
C.curvirostris
0.41/1
C.tshokwe
0.89/1
C.numenius
0.67/1
C.rhynchophorus
0.99/1
C.n.sp.**
0.92/1
C.elephas
1/1
C.alces
0.81/1
0.99/1
C.christyi*
1/1
C.tamandua
Gnathonemus petersii
0.004
2 ms
Clade 1
alces-complex

Figure 2

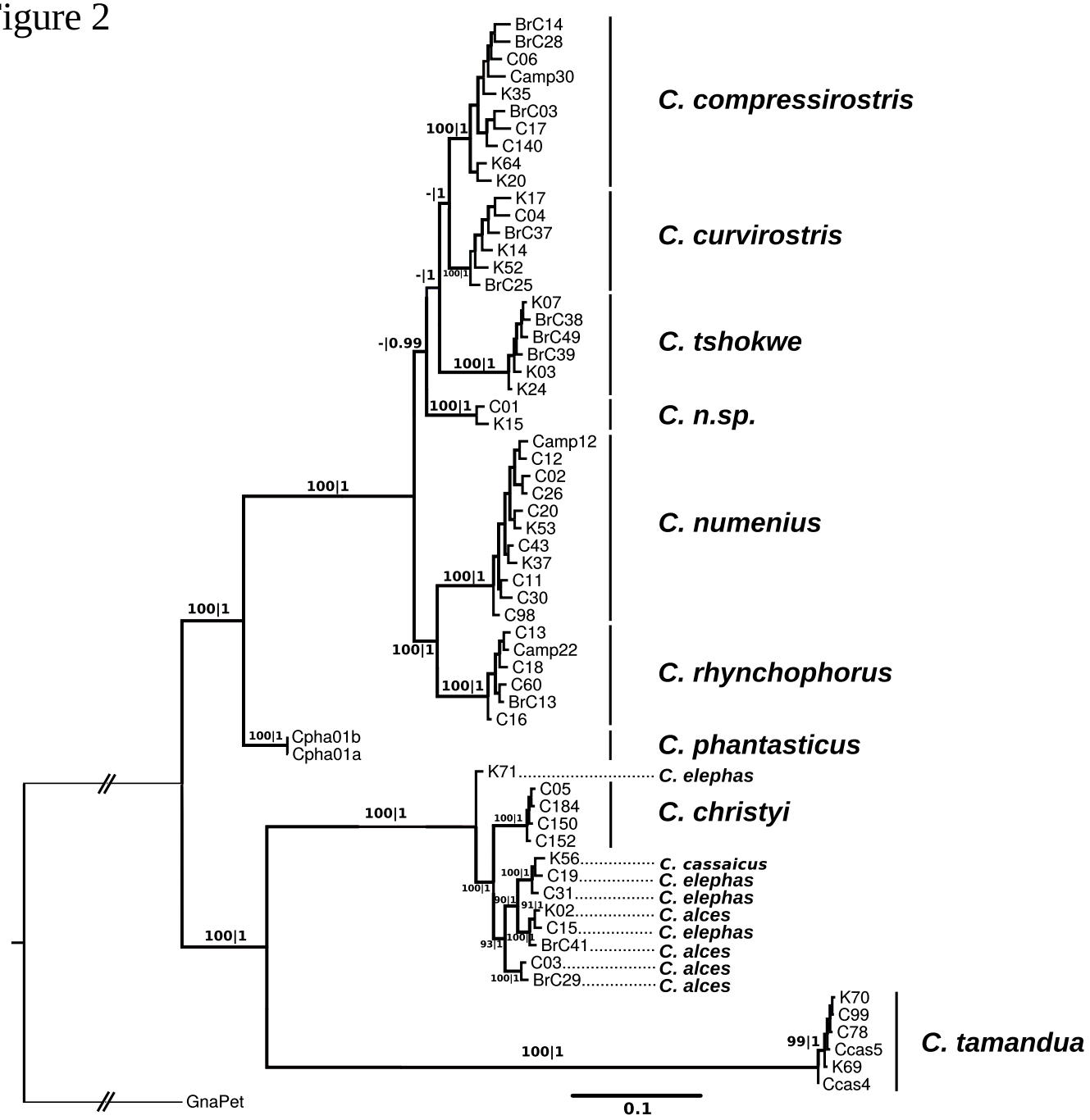

BrC14
BrC28
C06
Camp30
K35
BrC03
C17
C140
K64
K20
C. compressirostris
K17
C04
BrC37
K14
K52
BrC25
C. curvirostris
K07
BrC38
BrC49
BrC39
K03
K24
C. tshokwe
C01
K15
C. n.sp.
Camp12
C12
C02
C26
C20
K53
C43
K37
C11
C30
C98
C. numenius
C13
Camp22
C18
C60
BrC13
C16
C. rhynchophorus
Cpha01b
Cpha01a
C. phantasticus
K71
C. elephas
C05
C184
C150
C152
C. christyi
K56
C. cassaicus
C19
C. elephas
C31
C. elephas
K02
C. alces
C15
C. elephas
BrC41
C. alces
C03
C. alces
BrC29
C. alces
K70
C99
C78
Ccas5
K69
Ccas4
C. tamandua
GnaPet
0.1

Figure 3

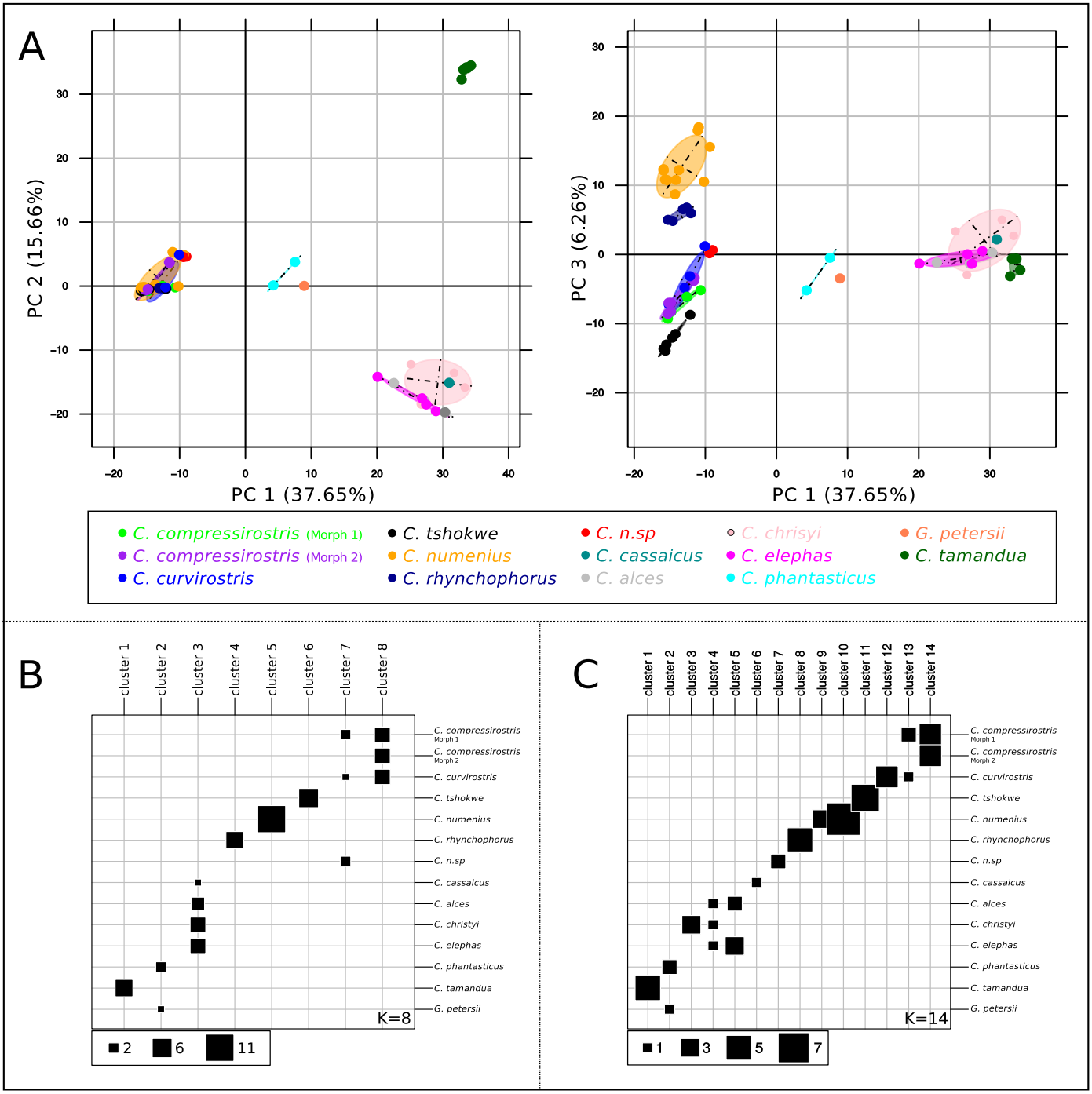
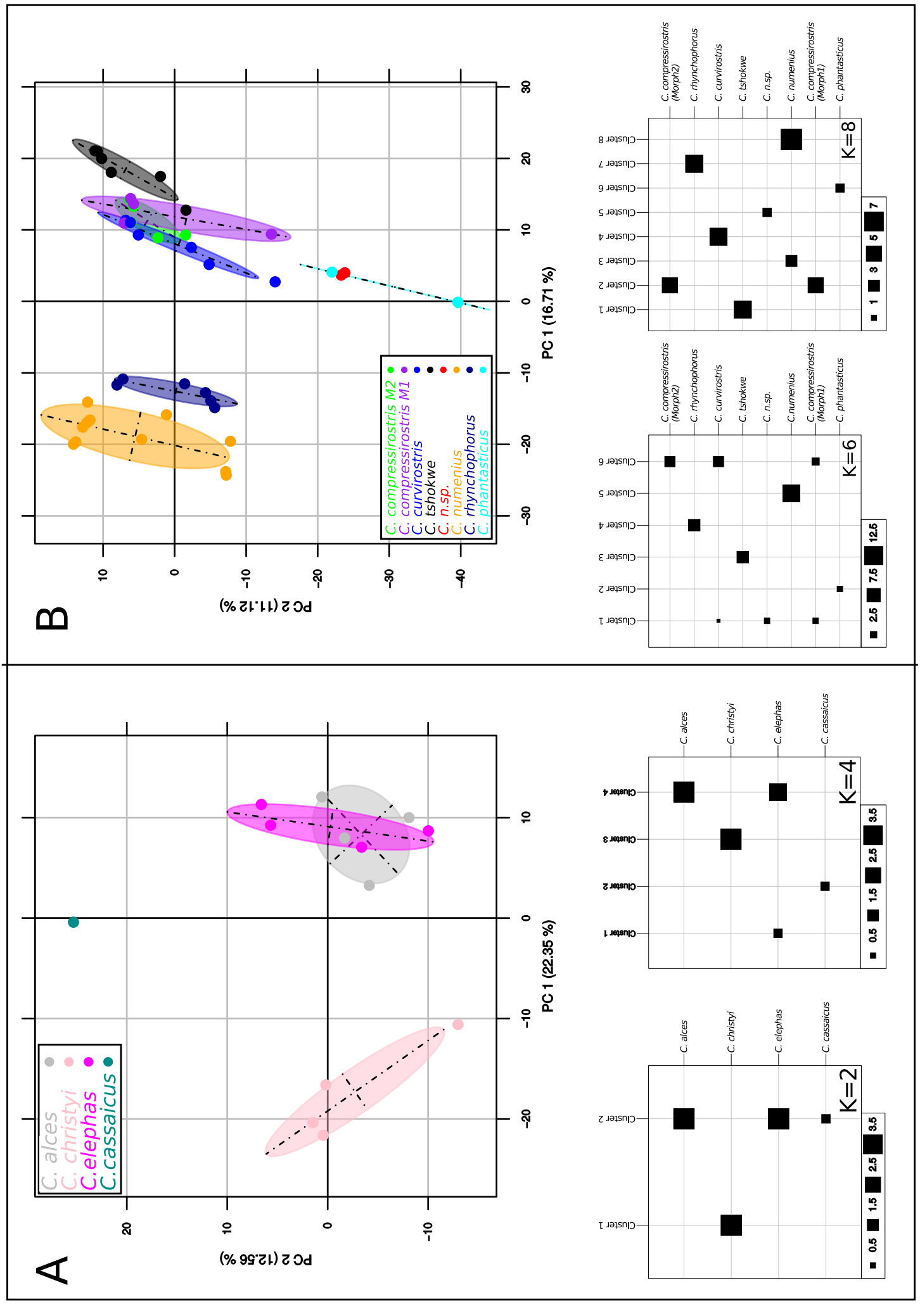

A
PC 2 (15.66%)
PC 1 (37.65%)
PC 3 (6.26%)
PC 1 (37.65%)
C. compressirostris (Morph 1)
C. compressirostris (Morph 2)
C. curvirostris
C. tshokwe
C. numenius
C. rhynchophorus
C. n.sp
C. cassaicus
C. alces
C. chrisyi
C. elephas
C. phantasticus
G. petersii
C. tamandua
B
cluster 1
cluster 2
cluster 3
cluster 4
cluster 5
cluster 6
cluster 7
cluster 8
C. compressirostris Morph 1
C. compressirostris Morph 2
C. curvirostris
C. tshokwe
C. numenius
C. rhynchophorus
C. n.sp
C. cassaicus
C. alces
C. christyi
C. elephas
C. phantasticus
C. tamandua
G. petersii
K=8
2
6
11
C
cluster 1
cluster 2
cluster 3
cluster 4
cluster 5
cluster 6
cluster 7
cluster 8
cluster 9
cluster 10
cluster 11
cluster 12
cluster 13
cluster 14
C. compressirostris Morph 1
C. compressirostris Morph 2
C. curvirostris
C. tshokwe
C. numenius
C. rhynchophorus
C. n.sp
C. cassaicus
C. alces
C. christyi
C. elephas
C. phantasticus
C. tamandua
G. petersii
K=14
1
3
5
7

Figure 4:


A
B
PC 1 (22.35 %)
PC 2 (12.56 %)
C. alces
C. christyi
C.elephas
C.cassaicus
PC 1 (16.71 %)
PC 2 (11.12 %)
C. compressirostris M2
C. compressirostris M1
C. curvirostris
C. tshokwe
C. n.sp.
C. numenius
C. rhynchophorus
C. phantasticus
K=2
K=4
K=6
K=8
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
C. compressirostris (Morph2)
C. rhynchophorus
C. curvirostris
C. tshokwe
C. n.sp.
C. numenius
C. compressirostris (Morph1)
C. phantasticus

6 Discussion

The ability of bioelectrogenesis, the sympatric occurrence, and the question of their underlying mechanism render the African weak electric fish genus *Campylomormyrus* an interesting research organism in evolutionary biology. Several studies cover histological, behavioral, and genetic investigations on their EOD diversification and its evolutionary effects (Kramer & Kuhn, 1994; Feulner et al., 2008; Paul et al., 2015, 2016; Lamanna et al., 2015, 2016; Kirschbaum et al., 2016; Nagel et al., 2018a,b). However, far too little is known to draw accurate conclusions. Karyological and genomic approaches have been pursued in this study to expand this knowledge and emphasize the importance of conventional as well as modern methods. In this study, I describe the first karyotype of *Campylomormyrus compressirostris* by means of chromosome preparation from free embryos and two staining methods (Giemsa and DAPI). The results are analyzed in the context of chromosomal evolution among Osteoglossiformes fish. Furthermore, I set up a SNP analysis pipeline which is based on transcriptomic data, and have identified potential candidate genes and cellular processes that are most likely involved in the determination of the elongated EOD in *C. tshokwe*. The genes and processes are linked to biological functions including cell proliferation and differentiation which support the existing hypothesis of associations between electrocyte geometry and EOD characteristics. Moreover, I have established the ddRAD sequencing method in the laboratory of the Evolutionary Biology/Zoological Systematics group at the University of Potsdam, to gain insights into the phylogeny and evolution of *Campylomormyrus* considering genome-wide distributed genetic markers. In the following, the obtained results will be discussed in the context of current research.

6.1 Karyological evolution

Cytogenetic investigations provide pivotal information about the genome architecture and its evolution (Potter et al., 2017). In the order Osteoglossiformes and in the family Mormyridae in particular, the karyotype is assumed to be conserved which is shown by the narrow range of mormyrid diploid chromosome numbers (2n=48-50) (Ráb et al., 2016; Canitz et al., 2016; Barby et al., 2019). However, in the species-rich Mormyridae, cytogenetic information are missing for a high number of species. As *Campylomormyrus* is among them, the first karyotype of *C. compressirostris* has been described (Canitz

et al., 2016). Its diploid chromosome number (2n=48) is accordant with those of other mormyrids and supports its high conservation within the family. The fundamental number (FN=78) indicates that *C. compressirostris* has a relatively high number of bi-armed chromosomes (metacentric and sub-metacentric) which is unusual for mormyrid species (Ráb et al., 2016). A slightly higher number of bi-armed chromosomes is also described for *G. petersii* (Uyeno, 1973; Ozouf-Costaz et al., 2015) the phylogenetic sister group of *Campylomormyrus* (Lamanna et al., 2016). Hence, an increased number of bi-armed chromosomes within the monophyletic group, comprising both genera, indicates that events of chromosome rearrangements or gains by pericentric inversion and Robertsonian fusion probably already happened before they diverged. However, to underline such a hypothesis, karyological examinations for further species of both genera as well as a larger sample set are required.

Recently, the karyotype of *Gymnarchus niloticus*, which is the only species within Gymnarchidae has been described with a karyotype formula of 2n=34 consisting only bi-armed chromosomes (Hatanaka et al., 2018). In contrast, a second karyotype of the same species was published by Jegede et al. (2018) with a highly deviating chromosome number and karyotype formula (2n=54; 26m+14sm+14a) (Jegede et al., 2018). As both karyotypes originate from individuals of different localities, it is assumed that such a high intra-specific variation might be attributed to two different species of *Gymnarchus* (Jegede et al., 2018). Even if the argumentation of a new species based on the karyotype information is slightly overrated, intra-specific karyological diversity is observed, also within other mormyrid species including *G. petersii* and *Pantodon buchholzi* (Uyeno, 1973; Ozouf-Costaz et al., 2015; Ráb et al., 2016). Thus, variation in chromosome number and karyotype formula highlights the needs of further cytogenetic investigations among and in particular within species. In the genomic era, karyotyping is a basic tool on which further analysis are built upon. Especially regarding fish genomes which experienced the teleost-specific whole genome duplication (Pasquier et al., 2016), the identification of homologous chromosomes, and chromosome rearrangements by for example fluorescence in-situ hybridization aids to understand the complexity of genome structures and how they arose (Potter et al., 2017; Barby et al., 2019).

Karyological results are also helpful to infer speciation scenarios (Bakloushinskaya, 2017; Romanenko et al., 2018; Supiwong et al., 2019). In terms of sympatric speciation, reproductive incompatibility by chromosome rearrangements is suggested to play a primary

role for the speciation of a spider-complex (Rezác et al., 2018). Such a scenario might be also a potential postzygotic isolation mechanism among *Campylomormyrus* species. Moreover, considering inter-specific hybridization, karyological analyses of hybrids can provide information about possible speciation mechanisms. In butterflies, for example, inter-specific hybrids become reproductive isolated by the fixation of novel chromosome combinations through fusion and/or fission but without a change in the chromosome number (homoploid hybrid speciation) (Lukhtanov et al., 2015). As inter-specific hybridization is shown for some *Campylomormyrus* species at least under laboratory conditions (Kirschbaum et al., 2016), cytogenetic analysis of these hybrids might provide new information about evolutionary processes leading to sympatric speciation. Although these scenarios may be feasible for *Campylomormyrus*, extensive and thorough studies including approaches with fluorescence in-situ hybridization or comparative genomic hybridization, and whole chromosome painting are needed.

6.2 Diversification of electric organ discharges

Within *Campylomormyrus* EOD pulses are characterized by their phase number, discharge polarity, and pulse duration. An elongated EOD pulse ($\sim$ 40ms) is shown in the species *C. rhynchophorus* and *C. numenius*, which are phylogenetically closely related and form a monophyletic group (Feulner et al., 2008). The longer EOD occurs also in some specimens of the *alces*-complex, which includes the species *C. alces, C. elephas, C. cassaicus* and *C. christyi*. *Campylomormyrus tshokwe* possesses an intermediate long EOD ($\sim$ 4ms). It is the sister group of *C. curvirostris* and *C. compressirostris* which emit a short EOD ($\sim$ 0.4ms) like the remaining species (Feulner et al., 2008; Lamanna et al., 2016) (Figure 1).

In order to understand the EOD diversification among *Campylomormyrus* species, histological and genetic examinations among species with different EOD characteristics have been conducted (Lamanna et al., 2014, 2015; Paul et al., 2015, 2016; Nguyen et al., 2017; Nagel et al., 2017). Differential gene expression analyses reveal that genes associated with electrical and metabolic activity as well as transcriptional regulation are up-regulated in the electric organ compared to the skeleton muscle (Lamanna et al., 2015). Especially, several ion channel genes are up-regulated in the electric organ of *C. tshokwe* compared to those of *C. compressirostris* (Nagel et al., 2017). Furthermore, non-synonymous nucleotide substitutions are observed in the sodium voltage-gated ion channel gene, *scn4aa*, which

are exclusive to two species with a long EOD (*C. rhynchophorus* and *C. numenius*) (Paul et al., 2016). These results support the hypothesis that ion channels play a major role in the diversification of EOD waveforms in mormyrid fish (Zakon et al., 2008). However, they could not fully explain the observed variation in EOD characteristics.

Histological examinations of the electric organs indicate an association between pulse duration and particular electrocyte and stalk structures. In several mormyrid fish, the electrocytes of species with longer EODs show increased cell surfaces by the development of papillae and have multi-stalk systems (Bass, 1986b; Paul et al., 2015). This hypothesis is supported by the candidate genes and biological processes I obtained from the transcriptome-wide SNP analysis considering to two main criteria: (1) Positive selection is inferred for the gene ($\omega > 1$) and (2) it carries a species-specific non-synonymous SNP in the species with elongated pulse (*C. tshokwe*). These genes have functions related to cell proliferation and differentiation as well as apoptosis which are crucial for tissue morphogenesis. Non-synonymous nucleotide substitutions within those genes may lead to a functional change or loss of the corresponding protein and thus likely contribute to the shaping of electric organs. Apart from the gene functions related to cell proliferation and differentiation, transcriptional regulation is a further process which is suggested to be relevant for determining the EOD duration. The importance of this process is already emphasized by the differential gene expression patterns among *C. tshokwe* and *C. compressirostris* (Lamanna et al., 2014, 2015; Nagel et al., 2017). Variation in genes coding for transcriptional regulation can cause changes in expression patterns of their target gene. Thus, non-synonymous SNPs might alter their protein confirmations leading to an impaired interaction between transcription factor and promoter region or modifying the function of transcriptional activators or repressors (Murgatroyd, 2004; Shastry, 2009). The results obtained by the transcriptome-wide SNP analysis study indicate an impact of both processes, transcriptional regulation and cell proliferation, on the elongation of EODs at least in the species *C. tshokwe*.

In other mormyrid weakly electric fish, a further aspect, the hormone level, is considered to mediate the EOD duration (Bass & Hopkins, 1982; Bass et al., 1986; Carlson et al., 2000). Experiments with gonad steroids reveal that after hormonal administration electrocyte sizes and membrane surfaces are increased and that emitted EODs are elongated, regardless of sex (Bass et al., 1986). As those experiments are restricted to species with sex dimorphism in EOD duration raises the question if hormone levels are also involved in inter-specific

determination of this characteristic. Furthermore, hormone effects are studied in the gymnotiform fish *Sternopygus* for which it is proposed to impact membrane integrated voltage-gated ion channels in terms of regulation or property change (Few & Zakon, 2006; Ferrari & McAnelly, 1995). Although genetic indications of processes related to, for example, hormone metabolism or segregation have not been reported for *Campylomormyrus*, their impacts on the EOD duration should not be ignored.

Combining the results of all studies provides a general overview of the genetic complexity which underlies the determination of EOD characteristics. At the same time, it also shows that the various factors identified, such as transcriptional regulation, cell proliferation and ion channel activity, are interrelated. In this case a potential scenario could be that an enlarged cell surface provides space for a higher number of membrane-integrated ion channels which are controlled by transcriptional regulation, and each factor can be affected by genetic mutations. An untoward aspect is that these investigations were only carried out for a few species. Therefore it cannot be ruled out whether the results found can be attributed to the actual characteristic including EOD duration, or to the respective diverging lineage. Hence, to continue this work a large-scale transcriptome analysis is needed including differential gene expression of the obtained candidate genes, and further species with different EOD duration.

Electric organ discharges represent a communication system within mormyrid fish and are used for species discrimination and mate recognition (Hopkins & Bass, 1980). It is a species-specific trait (Hopkins & Bass, 1980) which diversified along the evolution of these fish and, at least for *Campylomormyrus* it is thought to be a prezygotic isolation mechanism that is either the cause or effect of selection (Tiedemann et al., 2010). Hence knowing the underlying genes which contributes to the EOD diversification may help to understand to which extent EODs are involved in evolutionary processes and speciation.

It should be point out that transcriptomic data provide a huge potential to identify genes and biological processes which are involved in the determination of particular phenotypic traits, especially as those genes are expressed and their genetic variation affects directly the phenotype. Hence, regarding the identification of candidate genes coding for certain traits, transcriptomic analysis regardless of whether differential gene expression or SNP detection, should be continued and improved.

6.3 Phylogenetic classification of *Campylomormyrus*

The species of *Campylomormyrus* have been sampled during three field trips in 2004, 2006, and 2012. Those specimens originated from the Lower Congo River where they occur in sympatry. Among these individuals, several species have been be identified by phylogenetic and morphometric analyses (Feulner et al., 2006, 2008; Lamanna et al., 2016). Six species show clear genetic differences (*C. compressirostris*, *C. curvirostris*, *C. tshokwe*, *C. rhynchophorus*, *C. numenius*, and *C. tamandua*) while three species, *C. elephas, C. alces*, and *C. christyi*, are genetically not well distinguishable (Lamanna et al., 2016). For this reason, they are still regarded as species-complex and are combined to the *alces*-complex. However, these nine species are accepted and it is thought that they are reproductively isolated (Feulner et al., 2008). Phylogenetic relationships among these *Campylomormyrus* species were inferred by a comprehensive analysis of coalescence-based species tree estimation with well-supported posterior probabilities at each split (Lamanna et al., 2016) but there are still uncertainties among relationships of single individuals (Feulner et al., 2007, 2008). Although each species of the *alces*-complex is accepted, genetic differences among them are low based on the used genetic markers including microsatellite data and the genes *cytb*, *rps7*, and *scn4aa* (Feulner et al., 2007, 2008; Lamanna et al., 2016).

The genome-wide distributed loci, which are obtained in this study, provide informative phylogenetic relationships among individuals, lead to a change of the main species tree topology and might uncover cryptic species. Previously inferred monophyletic groups of the species *C. numenius, C. rhynchophorus, C. tshokwe, C. n.sp., C. compressirostris, C. curvirostris*, and *C. tamandua* are supported by the genome-wide genetic structure and except for *C. compressirostris*, all of them possess a species-specific EOD. In contrast, the specimens of the *alces*-complex cluster neither by species nor by EOD. Short branch lengths as well as species overlaps in the cluster analyses indicate their high genetic similarity. Among mormyrids, a sympatrically occurring species complex of the genus *Brienomyrus* (*magnostipes*-complex) presents a similar pattern of high diversity in EOD waveforms but low genetic differences (Arnegard et al., 2005, 2006). It is hypothesized that this pattern is representing incipient sympatric speciation meaning that the species of this complex just recently started to diverge and that EOD diversification arises before genome-wide genetic differences accumulate between sister species (Arnegard et al., 2006). This might be also the case for the *alces*-complex of *Campylomormyrus*.

In this study, two additional species, *C. cassaicus* and *C. phantasticus*, are added to the analysis which have not been included in the species tree. *Campylomormyrus phantasticus* is positioned as basal taxa within one of the inferred main clades including most of the *Campylomormyrus* species. As the identity of the used specimens is uncertain because of missing pictures and EOD records, I can just claim that a further genetically distinct *Campylomormyrus* species is distributed in the Lower Congo River. If they are representing *C. phantasticus*, it would mean that the species is not endemic to the Sanaga River (Cameroon) as assumed by Daget et al. (1984) and Stiassny et al. (2007), and has a much wider distribution area than other *Campylomormyrus* species (except for *C. tamandua*).

The second newly added species is *C. cassaicus*. This species is represented by a single individual and clusters among the specimens of the *alces*-complex. The unique color pattern assumes that the specimen belongs to the *C. cassaicus*, but the genome-wide analysis indicates low genetic difference to other individuals of the complex. Color traits can be involved in sympatric evolution due to mate choice. For example, among sympatrically occurring cichlid species, color patterns are linked to sexual selection (Seehausen et al., 1997; Seehausen, 2000). In *Campylomormyrus* visual signals are unlikely to affect mate choices. The usage of the EOD for species discrimination and mate recognition (Tiedemann et al., 2010; Nagel et al., 2018a) underlines that the vision might be less pronounced. Hence, the distinct color pattern of *C. cassaicus* is probably an insufficient character to distinguish the species. In contrast, as the specimen is part of the *alces*-complex, and incipient speciation is assumed for this group, its different color pattern would be another trait, which evolved at early stages of the species history and is uniquely for this lineage. However, the species is represented by a single individual, which poorly supports these hypotheses. A larger sample set of *C. cassaicus* and further information on ecology and distribution would help to corroborate or reject them.

The high species diversity within the Mormyridae family (Poll, 1945, 1967; Gosse, 1984) would support the theory that species are just diverged. Especially in the recent years, new species were identified and even an entire new genus has been described (Feulner et al., 2008; Sullivan et al., 2016; Rich et al., 2017). An increased diversity might be also found within *Campylomormyrus*. The individual-based approach, conducted by this study, reveals an outstanding phylogenetic position of a single individual, labeled as K71. It diverged even before the monophyletic group of *C. christyi* splits from the remaining

specimens and first observations of its morphology and color show differences to other specimens. In comparison to *C. cassaicus* which has also a deviating color pattern, its genetic difference to other specimens of the *alces*-complex is higher and indicates the occurrence of cryptic species. The *alces*-complex, in general, is obviously an interesting group of specimens, which should be investigated in more detail.

The inclusion of two additional species and the higher amount of genetic markers have changed one major aspect in the topology of the species tree. Regardless whether the tree is based on genetic differences of single individuals or on a previously merging of those to described species, the results show that *C. tamandua* is not anymore inferred as basal lineage being the sister group to all other *Campylomormyrus* species. Instead it is positioned as sister taxon of the *alces*-complex. Within *Campylomormyrus*, *C. tamandua* is the only distinct species with a clear triphasic EOD and a description of stalk-penetrating electrocytes (Paul et al., 2015). These are considered to be ancestral according to its former phylogenetically basal position, and the suggestion of an early evolution of the same characteristics in the family Mormyridae (Hopkins, 1995; Sullivan et al., 2000). Consequently, non-penetrated electrocytes and the accompanied biphasic EOD evolved after *C. tamandua* diverged, and are derived. Such a reversal to non-penetrating electrocytes and the biphasic EOD occurs in multiple phylogenetically independent lineages and clades (Sullivan et al., 2000, 2004). Hence, the new phylogenetic position of *C. tamandua* raises alternative hypotheses about the ancestral state of electric organ structures and its corresponding EOD waveform within *Campylomormyrus*. These theories depend further on the electric organ structure and EOD of the species thought to be *C. phantasticus*. If *C. phantasticus* and thus a triphasic EOD waveform is represented by the two individuals examined here, an ancestral triphasic EOD waveform is more likely. According to the parsimony principle, this scenario needs only two instead of three evolutionary events to explain current EOD characteristics within *Campylomormyrus* (Figure 4.1). The multiple reversals of a biphasic EOD in mormyrid linages increases the probability that the common ancestor of *Campylomormyrus* fish had a biphasic EOD. Otherwise, if the individuals belong to a species with biphasic EOD, both scenarios, an ancestral biphasic and an ancestral triphasic EOD, are likely according to the parsimony principle (Figure 4.2). In both cases, two evolutionary events are required to explain the pattern of extant species. However, a triphasic EOD waveform in the common ancestor of *Campylomormyrus* is supported by the presence of the same EOD waveform in its sister clade *Gnathonemus*.

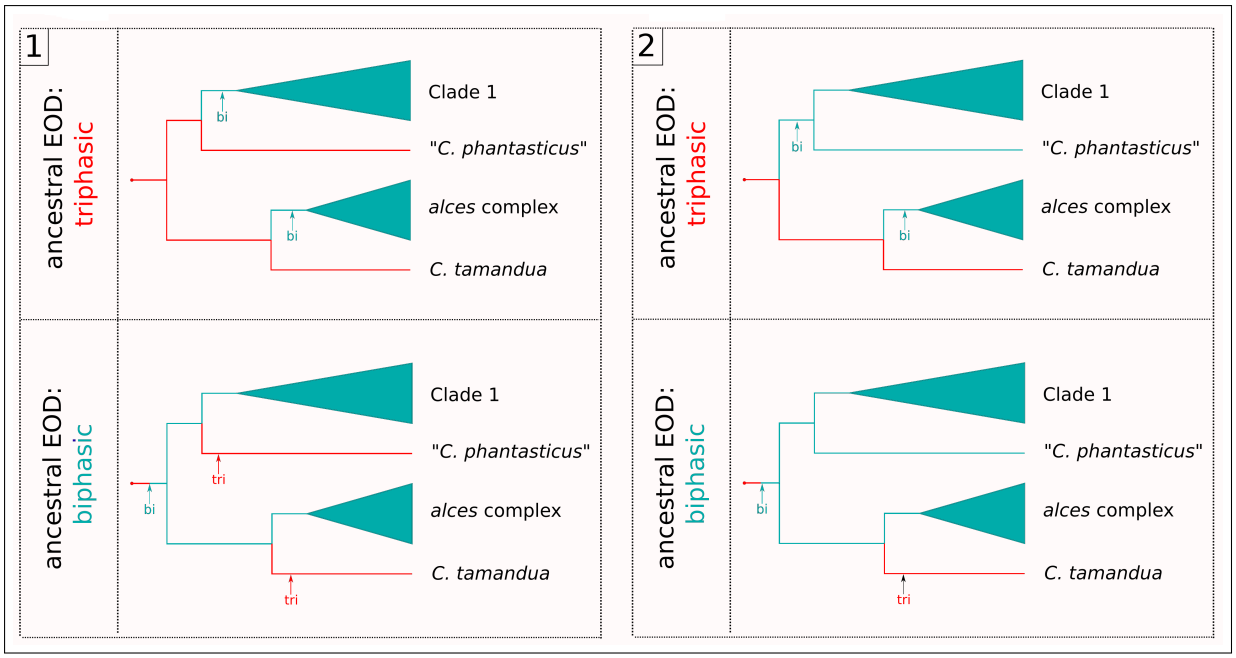


Figure 4: Schematic display of two different scenarios of EOD evolution within *Campylomormyrus* assuming (1) a triphasic or (2) a biphasic EOD for *"C. phantasticus"*. The cladograms represent the alternate tree topology obtained by this study and taxa sharing the same waveform are collapsed. The two scenarios assume either a tri- or a biphasic ancestral EOD and arrows indicate the number and positions of evolutionary events regarding the EOD waveform. Colors code always either for biphasic (bi - turquoise) or for triphasic (tri - red) EODs.

6.4 Conclusion and future perspective

The different approaches conducted by this study, regardless of being conventional (karyology) or modern (genomics), provide new information about the evolutionary history and EOD diversification of *Campylomormyrus*. In the context of sympatric speciation, which is the evolutionary divergence of reproductive isolated species in the absence of any barriers to gene flow, the EOD is considered as prezygotic isolation mechanism among *Campylomormyrus* species. The transcriptome-wide SNP analysis identified new candidate genes and biological processes those functions have not yet been considered for the determination of EOD waveform characteristics including its duration. Hence, the identification of genetic differences in expressed genes is a crucial step to understand the EOD diversification of mormyrids. Future genetic analyses, regardless of gene expression or sequence variation, should focus on genes which are involved in transcription and cell formation.

Due to the continuously increase of genomic approaches, conventional methods such as karyotyping, are getting more attention again. Comparative analyses including cytogenetic and molecular markers can provide a better understanding of genome organization and can demonstrate evolutionary dynamics and hybridization events as it is shown in the teleost genus *Clarias* (Maneechot et al., 2016). For this reason, the karyotype description

of *C. compressirostris* is a first step to obtain a complete picture of cytogenetic processes driving speciation in *Campylomormyrus*. Further karyotype descriptions are needed to draw conclusions on chromosome rearrangements among and within species. Moreover, the karyotype is a key aspect of the genome and can be linked to a whole genome sequence to generate a so-called chromosome-level assembly. Hence, the karyotype of *Campylomormyrus* would support the genome sequencing of the species which is demanded for future genomic studies.

Genomic approaches have changed the perspective on sympatric speciation, because hybridization and gene flow are assumed to occur more frequently than expected before and after speciation (Nosil, 2008; Han et al., 2017; Foote, 2018; Campbell et al., 2018). For example, there are some evidences for gene introgression among long diverged cichlid species (Wagner & McCune, 2009; Martin, 2012; Meier et al., 2017). The genome architecture plays an important role to comprehend sympatric speciation in the face of signals of gene flow, introgression, and/or hybridization (Harrison & Larson, 2014; Payseur & Rieseberg, 2016; Campbell et al., 2018). Considering those aspects, the genomic data of *Campylomormyrus* which have been obtained by the ddRAD sequencing method, are an useful basis for investigating the impact of these processes on the evolution of *Campylomormyrus*. Especially as hybridization under laboratory conditions has been shown among some of its species (Kirschbaum et al., 2016), those processes should be considered as potential driver of speciation and EOD diversification. Furthermore, this method is an important tool to improved the inference of phylogenetic relationships among single individuals as shown for the *alces*-complex. With a higher amount of genetic markers it is also helpful to uncover species diversity.

For future research, I suggest to continue the application of these methods for the investigation of *Campylomormyrus*' evolutionary history. Phylogenetic analysis based on genomic data should consider more specimens and a wider sampling area. In addition, a revision of the genus *Campylomormyrus* is requested as the knowledge about the species and, in particular its genetics, is continuously increasing. Finally, it is suggested to apply the established genomic approaches to other mormyrids, as Carlson & Gallant (2013) already underlines the importance of genomic resources to understand the evolutionary processes underlying mormyrid diversification in terms of communication and behavior.

7 Abstract

The African weakly electric fish genus *Campylomormyrus* is a well-investigated fish group of the species-rich family Mormyridae. They are able to generate species-specific electric organ discharges (EODs) which vary in their waveform characteristics including polarity, phase number and duration. In mormyrid species EODs are used for communication, species discrimination and mate recognition, and it is thought that they serve as pre-zygotic isolation mechanism driving sympatric speciation by promoting assortative mating. The EOD diversification, its evolutionary effects and the link to species divergence have been examined histologically, behaviorally, and genetically. Molecular analyses are a major tool to identify species and their phenotypic traits by studying the underlying genes. The genetic variability between species further provides information from which evolutionary processes, such as speciation, can be deduced.

Hence, the ultimate aim of this study is the investigation of genetic variability within the African weakly electric fish genus *Campylomormyrus* to better understand their sympatric speciation and comprehend their evolutionary drivers. In order to extend the current knowledge and gain more insights into its species history, karyological and genomic approaches are being pursued considering species differences.

Previous studies have shown that species with different EOD duration have specific gene expression patterns and single nucleotide polymorphisms (SNPs). As EODs play a crucial role during the evolution of *Campylomormyrus* species, the identification of its underlying genes may suggest how the EOD diversity evolved and whether this trait is based on a complex network of genetic processes or is regulated by only a few genes. The results obtained in this study suggest that genes with non-synonymous SNPs, which are exclusive to *C. tshokwe* with an elongated EOD, have frequent functions associated with tissue morphogenesis and transcriptional regulation. Therefore, it is proposed that these processes likely co-determine EOD characteristics of *Campylomormyrus* species. Furthermore, genome-wide analyses confirm the genetic difference among most *Campylomormyrus* species. In contrast, the same analyses reveal genetic similarity among individuals of the *alces*-complex showing different EOD waveforms. It is therefore hypothesized that the low genetic variability and high EOD diversity represents incipient sympatric speciation. The karyological description of a *Campylomormyrus* species provides crucial information about chromosome number and shapes. Its diploid chromosome number of 2n=48 supports the

conservation of this trait within Mormyridae. Differences have been detected in the number of bi-armed chromosomes which is unusually high compared to other mormyrid species. This high amount can be due to chromosome rearrangements which could cause genetic incompatibility and reproductive isolation. Hence an alternative hypothesis regarding processes which cause sympatric speciation is that chromosome differences are involved in the speciation process of *Campylomormyrus* by acting as postzygotic isolation mechanism.

In summary, the karyological and genomic investigations conducted in this study contributed to the increase of knowledge about *Campylomormyrus* species, to the solution of some existing ambiguities like phylogenetic relationships and to the raising of new hypothesis explaining the sympatric speciation of those African weakly electric fish. This study provides a basis for future genomic research to obtain a complete picture for causes and results of evolutionary processes in *Campylomormyrus*.

8 German Abstract

Die Gattung *Campylomormyrus* gehört zur Familie der afrikanisch schwach elektrischer Fische Mormyridae (Osteoglossiformes, Teleostei). Die Arten können elektrische Organentladungen (engl. Electric Organ Discharge; EOD) generieren, die bezüglich der Wellenform und -dauer artspezifisch sind und somit der Kommunikation und Arterkennung dienen. Durch Verhaltensexperimente wurde festgestellt, dass bei *Campylomormyrus* Arten das EOD auch die Partnerwahl beeinflusst, was zur Annahme führt, dass EODs als ein prezygotischer Isolationsmechanismus fungieren. Dies fördert die assortative Paarung und kann somit zur sympatrischen Artbildung führen. Die Vielfalt der EOD Wellenform in *Campylomormyrus* und ihr Einfluss auf die evolutionären Geschichte der Arten, sind bereits vielfach Gegenstand histologischer, verhaltensbedingter und genetischer Studien gewesen. Derzeit sind vorallem genetische Untersuchungen ein wichtiges Instrument, um Arten zu unterscheiden oder genetische Variabiliät in Genen zu finden, die den verschiedenen Phänoypen zu Grunde liegen. Genetische Unterschiede zwischen Arten können auch Informationen liefern, von denen evolutionäre Prozesse, wie die sympatrische Artentstehung, abgeleitet werden können.

Das Ziel dieser Arbeit ist es deshalb, genetische Unterschiede zwischen den Arten der Gattung *Campylomormyrus* zu identifizieren und Rückschlüsse auf mögliche Mechanismen zu ziehen, die zur ihrer sympatrischen Artbildung geführt haben. Zur Erweiterung des aktuellen Wissens über die Gattung *Campylomormyrus*, werden genomische und karylogische Ansätze unter Berücksichtigung von Artunterschieden verfolgt. Bisherige Studien haben gezeigt, dass Arten mit unterschiedlichen EOD Längen auch verschiedene Genexpressionenmuster aufweisen und spezifische nicht-synonyme Einzelnukleotid-Polymorphismus (engl. Single Nucleotide Polymorphism; SNP) besitzen. Da die EOD Diversifizierung eine wichtige Rolle in der Evolution der *Campylomormyrus* Arten spielt, ist davon auszugehen, dass die Identifizierung der zu Grunde liegenden Gene ein Fortschritt bezüglich des Verständnis der Evolution der EOD Viellfalt darstellt. Die identifizierten Kandidaten-Gene können Hinweise liefern, ob dieses Merkmal auf einem komplexen Netzwerk von genetischen Prozessen basiert oder nur von einigen wenigen Gene reguliert wird. Die Ergebnisse dieser Studie haben ergeben, dass bei *Campylomormyrus tshokwe*, eine Art mit verlängertem EOD, nicht-synonymen SNPs vor allem in Genen, die im Zusammenhang mit Gewebebildung oder transkriptioneller Regulation stehen auftreten. Das deutet darauf hin, dass diese

biologischen Prozesse die Bestimmung der EOD Eigenschaften bei *Campylomormyrus* Arten wharscheinlich mitbestimmen. Au"serdem bestätigen die genom-weiten Analysen den genetischen Unterschied der meisten *Campylomormyrus* Arten. Dennoch zeigen sie auch, dass die Individuen des *alces*-Komplexes sich genetische sehr ähnlich sind, obwohl sie unterschiedliche EOD Wellenformen emmittieren. Somit kann hypothetisiert werden, dass die geringe genetische Variabilit und hohe EOD Diversität eine beginnende sympatrische Artbildung darstellt.

Die karyologische Beschreibung einer Art liefert wichtige Informationen über die Anzahl und Form der Chromosomen. Die für *Campylomormyrus compressirostris* beschriebene diploide Chromosomenzahl von 2n=48 ist übereinstimmend mit der konservierten Chromosomenanzahl innerhalb der Mormyridae (2n=48-50). Unterschiede gibt es jedoch in der Anzahl der zweiarmigen Chromosomen, die im Vergleich zu den meisten mormyriden Arten ungewöhnlich hoch ist. Diese Variation kann auf Chromosomenumstellungen oder -neuordnungen zurückzuführen sein, die zu genetischer Unverträglichkeit und reproduktiver Isolation führen können. Eine alternative Hypothese zum Ursprung von sympatrischer Artbildung ist daher, dass die Chromosomenunterschiede bei der Artbildung von *Campylomormyrus* als postzygotischer Isolationsmechanismus beteiligt sind. Zusammenfassend haben die Ergebnisse dieser Arbeit dazu beigetragen, das Wissen über *Campylomormyrus* Arten zu erweitern, einige bestehende Unklarheiten, wie phylogenetische Beziehungen, zu lösen, sowie neue Hypothesen aufzustellen, die ihre sympatrische Artentstehung erklären.

9 References

Albert, J. S., & Crampton, W. G. R. (2005). Electroreception and Electrogenesis. In D. Evans (Ed.) *The Physiology of Fishes*, chap. 12, (pp. 431–472). New York: CRC, 3 ed.

Alves-Gomes, J. A., & Hopkins, C. D. (1997). Molecular Insights into the Phylogeny of Mormyriform Fishes and the Evolution of Their Electric Organs. *Brain, Behavior and Evolutionn*, *49*, 324–351.

Arnegard, M. E., Bogdanowicz, S. M., & Hopkins, C. D. (2005). Multiple Cases of Striking Genetic Similarity between Alternate Electric Fish Signal Morphs in Sympatry. *Evolution*, *59*(2), 324–343.

Arnegard, M. E., Jackson, S. A., & Hopkins, C. D. (2006). Time-domain signal divergence and discrimination without receptor modification in sympatric morphs of electric fishes. *Journal of Experimental Biology*, *209*(11), 2182–2198.

Arnegard, M. E., Zwickl, D. J., Lu, Y., & Zakon, H. H. (2010). Old gene duplication facilitates origin and diversification of an innovative communication system–twice. *Proceedings of the National Academy of Sciences*, *107*(51), 22172–22177.

Bakloushinskaya, I. Y. (2017). Chromosomal rearrangements, genome reorganization, and speciation. *Biology Bulletin*, *43*(8), 759–775.

Barby, F. F., Bertollo, L. A. C., de Oliveira, E. A., Yano, C. F., Hatanaka, T., Ráb, P., Sember, A., Ezaz, T., Artoni, R. F., Liehr, T., Al-Rikabi, A. B., Trifonov, V., de Oliveira, E. H., Molina, W. F., Jegede, O. I., Tanomtong, A., & de Bello Cioffi, M. (2019). Emerging patterns of genome organization in Notopteridae species (Teleostei, Osteoglossiformes) as revealed by Zoo-FISH and Comparative Genomic Hybridization (CGH). *Scientific Reports*, *9*(1), 1–11.

Bass, A. H. (1986a). A hormone-sensitive communication system in an electric fish. *Journal of Neurobiology*, *17*(3), 131–155.

Bass, A. H. (1986b). Electric Organs Revisited: Evolution of a Vertebrate Communication and Orientation Organ. In T. H. Bullock, & W. Heiligenberg (Eds.) *Electroreception*, chap. 2, (pp. 13–70). New York: John Wiley and Sons.

Bass, A. H., Denizot, J., & Marchaterre, M. A. (1986). Ultrastructural features and hormone dependent sex differences of mormyrid electric organs. *Journal of Comparative Neurology*, *254*(4), 511–528.

Bass, A. H., & Hopkins, C. D. (1982). Hormonal Control of Sexual Differentiation: Changes in Electric Organ Discharge Waveform. *Science*, *220*, 971–974.

Bennett, M. V. L. (1961). Modes of Operation of Electric Organs. *Annals of the New York Academy of Sciences*, *94*(2), 458–509.

Bleeker, P. (1874). *Description of the Genus Campylomormyrus type sp. Mormyrus tamandua GTHR*, vol. 2. Amsterdam: Verslag en Med. K. Ak., 8 ed.

Boughman, J. W. (2001). Divergent sexual selection enhances reproductive isolation in sticklebacks. *Nature*, *411*(June), 944–946.

Boulenger, G. A. (1898). Matériaux pour la faune du Congo: Poissons Nouveaux. Tech. rep., État Independant du Congo, Bruxelles.

Campbell, C. R., Poelstra, J. W., & Yoder, A. D. (2018). What is Speciation Genomics? The roles of ecology, gene flow, and genomic architecture in the formation of species. *Biological Journal of the Linnean Society*, *124*(4), 561–583.

Canitz, J., Kirschbaum, F., & Tiedemann, R. (2016). Karyotype description of the African weakly electric fish Campylomormyrus compressirostris in the context of chromosome evolution in Osteoglossiformes. *Journal of Physiology Paris*, *110*(3), 273–280.

Cardoso, A. L., Pieczarka, J. C., Crampton, W. G., Ready, J. S., Ready, W. M. F., Waddell, J. C., de Oliveira, J. A., & Nagamachi, C. Y. (2018). Karyotypic diversity and evolution in a sympatric assemblage of Neotropical electric knifefish. *Frontiers in Genetics*, *9*(MAR), 1–12.

Carlson, B. A. (2002). Electric signaling behavior and the mechanisms of electric organ discharge production in mormyrid fish. *Journal of Physiology Paris*, *96*(5-6), 405–419.

Carlson, B. A., & Gallant, J. R. (2013). From sequence to spike to spark: Evo-devo-neuroethology of electric communication in mormyrid fishes. *Journal of Neurogenetics*, *27*(3), 106–129.

Carlson, B. A., Hasan, S. M., Hollmann, M., Miller, D. B., Harmon, L. J., & Arnegard, M. E. (2011). Diversification of Electric Fishes. *Science*, *583*(6029), 583–586.

Carlson, B. A., Hopkins, C. D., & Thomas, P. (2000). Androgen correlates of socially induced changes in the electric organ discharge waveform of a mormyrid fish. *Hormones and Behavior*, *38*(3), 177–186.

Catania, K. C. (2015). Electric Eels Concentrate Their Electric Field to Induce Involuntary Fatigue in Struggling Prey. *Current Biology*, *25*(22), 2889–2898.

Daget, J., Gosse, J., & Thys van den Audenaerde, D. (1984). *Check-List of the Freshwater Fishes of Africa*. Paris: OSTROM, MRAC.

Ferrari, M., & McAnelly, M. (1995). Individual variation in and androgen-modulation of the sodium current in electric organ. *The Journal of*, *15*(May), 4023–4032.

Feulner, P. G. D., Kirschbaum, F., Mamonekene, V., Ketmaier, V., & Tiedemann, R. (2007). Adaptive radiation in African weakly electric fish (Teleostei: Mormyridae: Campylomormyrus): a combined molecular and morphological approach. *Journal of Evolutionary Biology*, *20*(1), 403–414.

Feulner, P. G. D., Kirschbaum, F., Schugardt, C., Ketmaier, V., & Tiedemann, R. (2006). Electrophysiological and molecular genetic evidence for sympatrically occuring cryptic species in African weakly electric fishes (Teleostei: Mormyridae: Campylomormyrus). *Molecular Phylogenetics and Evolution*, *39*(1), 198–208.

Feulner, P. G. D., Kirschbaum, F., & Tiedemann, R. (2008). Adaptive radiation in the Congo River: An ecological speciation scenario for African weakly electric fish (Teleostei; Mormyridae; Campylomormyrus). *Journal of Physiology Paris*, *102*(4-6), 340–346.

Feulner, P. G. D., Plath, M., Engelmann, J., Kirschbaum, F., & Tiedemann, R. (2009a). Electrifying love: Electric fish use species-specific discharge for mate recognition. *Biology Letters*, *5*(2), 225–228.

Feulner, P. G. D., Plath, M., Engelmann, J., Kirschbaum, F., & Tiedemann, R. (2009b). Magic trait Electric Organ Discharge (EOD). *Communicative & Integrative Biology*, *2*(4), 329–331.

Few, W. P., & Zakon, H. H. (2006). Sex Differences in and Hormonal Regulation of Kv1 Potassium Channel Gene Expression in the Electric Organ: Molecular Control of a Social Signal. *Developmental neurobiology*, *67*(5), 535–549.

Foote, A. D. (2018). Sympatric Speciation in the Genomic Era. *Trends in Ecology and Evolution*, *33*(2), 85–95.

Freedman, E. G., Olyarchuk, J., Marchaterre, M. A., & Bass, A. H. (1989). A Temporal Analysis of Testosterone-Induced Changes in Electric Organs and Electric Organ Discharges of Mormyrid Fishes. *Journal of Neurobiology*, *20*(7), 619–634.

Gallant, J. R., Arnegard, M. E., Sullivan, J. P., Carlson, B. A., & Hopkins, C. D. (2011). Signal variation and its morphological correlates in Paramormyrops kingsleyae provide insight into the evolution of electrogenic signal diversity in mormyrid electric fish. *Journal of comparative physiology. A, Neuroethology, sensory, neural, and behavioral physiology*, *197*(8), 799–817.

Gallant, J. R., Hopkins, C. D., & Deitcher, D. L. (2012). Differential expression of genes and proteins between electric organ and skeletal muscle in the mormyrid electric fish Brienomyrus brachyistius. *Journal of Experimental Biology*, *215*(14), 2479–2494.

Gallant, J. R., Losilla, M., Tomlinson, C., & Warren, W. C. (2017). The Genome and Adult Somatic Transcriptome of the Mormyrid Electric Fish Paramormyrops kingsleyae. *Genome Biology and Evolution*, *9*(12), 3525–3530.

Gosse, J.-P. (1984). Mormyriformes. In J. Daget, J.-P. Gosse, & D. F. E. van den Audenaerde (Eds.) *Check-list of freshwater fishes in Africa*, (pp. 63–124). Paris, Tervuren: Ostrom/MRAC.

Han, F., Lamichhaney, S., Grant, B. R., Grant, P. R., Andersson, L., & Webster, M. T. (2017). Gene flow, ancient polymorphism, and ecological adaptation shape the genomic landscape of divergence among Darwin's finches. *Genome Research*, *27*(6), 1004–1015.

Harrison, R. G., & Larson, E. L. (2014). Hybridization, introgression, and the nature of species boundaries. *Journal of Heredity*, *105*(S1), 795–809.

Hatanaka, T., De Oliveira, E. A., Ráb, P., Yano, C. F., Bertollo, L. A., Ezaz, T., Jegede, O. O. I., Liehr, T., Olaleye, V. F., & Cioffi, M. D. B. (2018). First chromosomal analysis in Gymnarchus niloticus (Gymnarchidae: Osteoglossiformes): Insights into the karyotype evolution of this ancient fish order. *Biological Journal of the Linnean Society*, *125*(1), 83–92.

Higashi, M., Takimoto, G., & Yamamura, N. (1999). Sympatric speciation by sexual selection. *Nature*, *402*, 523–526.

Hopkins, C. D. (1986). Behavior of Mormyridae. In T. H. Bullock, & W. Heiligenberg (Eds.) *Electroreception*, (pp. 527–576). New York: John Wiley and Sons.

Hopkins, C. D. (1995). Convergent designs for electrogenesis and electroreception. *Current Opinion in Neurobiology*, *5*(6), 769–777.

Hopkins, C. D., & Bass, A. H. (1980). Temporal Coding of Species Recogition Signals in an Electric Fish. *Science*, *212*(3), 85–87.

Inoue, J., Sato, Y., Sinclair, R., Tsukamoto, K., & Nishida, M. (2015). Rapid genome reshaping by multiple-gene loss after whole-genome duplication in teleost fish suggested by mathematical modeling. *Proceedings of the National Academy of Sciences*, *112*(48), 14918–14923.

Jegede, O. I., Akintoye, M. A., & Awopetu, J. I. (2018). Karyotype of the African Weakly Electric Fish, Gymnachus niloticus (Osteoglossiformes: Gymnarchidae) from Oluwa River, Nigeria. *Ife Journal of Science*, *20*(3), 539–545.

Kirschbaum, F., Nguyen, L., Baumgartner, S., Chi, H. W. L., Wolfart, R., Elarbani, K., Eppenstein, H., Korniienko, Y., Guido-Böhm, L., Mamonekene, V., Vater, M., & Tiedemann, R. (2016). Intragenus (Campylomormyrus) and intergenus hybrids in mormyrid fish: Physiological and histological investigations of the electric organ ontogeny. *Journal of Physiology Paris*, *110*(3), 281–301.

Kitano, J., Ross, J. A., Mori, S., Kume, M., Jones, F. C., Chan, Y. F., Absher, D. M., Grimwood, J., Schmutz, J., Myers, R. M., Kingsley, D. M., & Peichel, C. L. (2009). A role for a neo-sex chromosome in stickleback speciation. *Nature*, *461*(7267), 1079–1083.

Knight, M. E., Turner, G. F., Rico, C., von Oppen, M. J. H., & Hewitt, G. M. (1998). Microsatellite paternity analysis on captive Lake Malawi cichlids supports reproductive isolation by direct mate choice. *Molecular Ecology*, *7*, 1605–1610.

Kramer, B., & Kuhn, B. (1994). Species recognition by the sequence of discharge intervals in weakly electric fishes of the genus Campylomormyrus (Mormyridae, Teleostei). *Animal Behaviour*, *48*(2), 435–445.

Lamanna, F., Kirschbaum, F., Ernst, A. R. R., Feulner, P. G. D., Mamonekene, V., Paul, C., & Tiedemann, R. (2016). Species delimitation and phylogenetic relationships in a genus of African weakly-electric fishes (Osteoglossiformes, Mormyridae, Campylomormyrus). *Molecular Phylogenetics and Evolution*, *101*, 8–18.

Lamanna, F., Kirschbaum, F., & Tiedemann, R. (2014). De novo assembly and characterization of the skeletal muscle and electric organ transcriptomes of the African weakly electric fish Campylomormyrus compressirostris (Mormyridae, Teleostei). *Molecular Ecology Resources*, *14*(6), 1222–1230.

Lamanna, F., Kirschbaum, F., Waurick, I., Dieterich, C., & Tiedemann, R. (2015). Cross-tissue and cross-species analysis of gene expression in skeletal muscle and electric organ of African weakly-electric fish (Teleostei; Mormyridae). *BMC Genomics*, *16*(1), 1–17.

Lavoué, S., Bigorne, R., Lecointre, G., & Agnèse, J.-F. (2000). Phylogenetic Relationships of Mormyrid Electric Fishes (Mormyridae; Teleostei) Inferred from Cytochrome b Sequences. *Molecular Phylogenetics and Evolution*, *14*(1), 1–10.

Lavoué, S., Miya, M., Arnegard, M. E., Sullivan, J. P., Hopkins, C. D., & Nishida, M. (2012). Comparable ages for the independent origins of electrogenesis in African and South American weakly electric fishes. *PLoS ONE*, *7*(5), 1–18.

Lissmann, H. W. (1951). Continuous Electrical Signal from the Tail of a Fish, Gymnarchus niloticus Cuv. *Nature*, *167*(3), 201–202.

Lissmann, H. W. (1958). On the Function and Evolution of Electric Organs in Fish. *Journal of Experimental Biology*, *35*(1), 156–191.

Lukhtanov, V. A., Shapoval, N. A., Anokhin, B. A., Saifitdinova, A. F., & Kuznetsova, V. G. (2015). Homoploid hybrid speciation and genome evolution via chromosome sorting. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1807).

Mago-Leccia, F. (1994). Electric fishes of the continental waters of America. *Biblioteca de la Academia de Ciencias Fisicas Matematicas, y Naturals*, *29*, 1–225.

Maneechot, N., Yano, C. F., Bertollo, L. A. C., Getlekha, N., Molina, W. F., Ditcharoen, S., Tengjaroenkul, B., Supiwong, W., Tanomtong, A., & De Bello Cioffi, M. (2016). Genomic organization of repetitive DNAs highlights chromosomal evolution in the genus Clarias (Clariidae, Siluriformes). *Molecular Cytogenetics*, *9*(1), 1–10.

Marrero, C., & Winemiller, K. O. (1993). Tube-snouted gymnotiform and mormyriform fishes: convergence of a specialized foraging mode in teleosts. *Environmental Biology of Fishes*, *38*(4), 299–309.

Martin, C. H. (2012). Weak Disruptive Selection and Incomplete Phenotypic Divergence in Two Classic Examples of Sympatric Speciation: Cameroon Crater Lake Cichlids. *The American Naturalist*, *180*(4), E90–E109.

McGregor, P. K., & Westby, G. W. M. (1992). Discrimination of individually characteristic electric organ discharges by a weakly electric fish. *Animal Behaviour*, *43*(6), 977–986.

Meier, J. I., Marques, D. A., Mwaiko, S., Wagner, C. E., Excoffier, L., & Seehausen, O. (2017). Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nature Communications*, *8*(May 2016), 1–11.

Meyer, A., & Van De Peer, Y. (2005). From 2R to 3R: Evidence for a fish-specific genome duplication (FSGD). *BioEssays*, *27*(9), 937–945.

Moller, P. (1995). *Electrish fishes: History and Behavior*. London: Chapman & Hall, 17 ed.

Murgatroyd, C. (2004). Impaired Repression at a Vasopressin Promoter Polymorphism Underlies Overexpression of Vasopressin in a Rat Model of Trait Anxiety. *Journal of Neuroscience*, *24*(35), 7762–7770.

Nagel, R., Kirschbaum, F., Engelmann, J., Hofmann, V., Pawelzik, F., & Tiedemann, R. (2018a). Male-mediated species recognition among African weakly electric fishes. *Royal Society Open Science*, *5*(2), 170443.

Nagel, R., Kirschbaum, F., Hofmann, V., Engelmann, J., & Tiedemann, R. (2018b). Electric pulse characteristics can enable species recognition in African weakly electric fish species. *Scientific Reports*, *8*(1), 1–12.

Nagel, R., Kirschbaum, F., & Tiedemann, R. (2017). Electric organ discharge diversification in mormyrid weakly electric fish is associated with differential expression of voltage-gated ion channel genes. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, *203*(3), 183–195.

Nelson, J. S., Grande, T. C., & Wilson, M. V. (2016). *Fishes of the World*. John Wiley and Sons.

Nguyen, L., Paul, C., Mamonekene, V., Bartsch, P., Tiedemann, R., & Kirschbaum, F. (2017). Reproduction and development in some species of the weakly electric genus Campylomormyrus (Mormyridae, Teleostei). *Environmental Biology of Fishes*, *100*(1), 49–68.

Nosil, P. (2008). Speciation with gene flow could be common. *Molecular Ecologycology*, *17*(9), 2103–2106.

Ohno, S., Wolf, U., & Aktin, N. B. (1967). Evolution From Fish To Mammal by Gene Duplication. *Hereditas*, *59*(1), 169–187.

Ozernyuk, N. D., & Myuge, N. S. (2013). Large-scale genome duplications and paralog divergence in fish. *Russian Journal of Genetics*, *49*(1), 63–69.

Ozouf-Costaz, C., Coutanceau, J.-P., Bonillo, C., Belkadi, L., Fermon, Y., Agnese, J.-F., Guidi-Rontani, C., & Paugy, D. (2015). First insights into karyotype evolution within the family Mormyridae. *Cybium: International Journal of Ichthyology*, *39*(3), 227–236.

Pasquier, J., Cabau, C., Nguyen, T., Jouanno, E., Severac, D., Braasch, I., Journot, L., Pontarotti, P., Klopp, C., Postlethwait, J. H., Guiguen, Y., & Bobe, J. (2016). Gene evolution and gene expression after whole genome duplication in fish: The PhyloFish database. *BMC Genomics*, *17*(368).

Paul, C., Kirschbaum, F., Mamonekene, V., & Tiedemann, R. (2016). Evidence for Non-neutral Evolution in a Sodium Channel Gene in African Weakly Electric Fish (Campylomormyrus, Mormyridae). *Journal of Molecular Evolution*, *83*(1-2), 61–77.

Paul, C., Mamonekene, V., Vater, M., Feulner, P. G. D., Engelmann, J., Tiedemann, R., & Kirschbaum, F. (2015). Comparative histology of the adult electric organ among four species of the genus Campylomormyrus (Teleostei: Mormyridae). *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, *201*(4), 357–374.

Payseur, B. A., & Rieseberg, L. H. (2016). A genomic perspective on hybridization and speciation. *Molecular ecology*, *25*(11), 2337–60.

Pellegrin, J. (1927). Mormyridés du Cameroun recueillis par M. Th. Monod. Description d'un genre, de quatre espèces et d'une variété. *Bulletin de la société royale de zoologie*, *52*, 294–300.

Poll, M. (1945). Descriptions de Mormyridae et de Characidae nouveaux du Congo bel avec une étude du genre Stotnator hinus et des genres de Characidae nains africains. *Revue de Zoologie et de Botanique Africaines*, *39*(1), 36–77.

Poll, M. (1967). Contribution à la faune ichthyologique de l'Angola. *Diamang Publishing Culture*, *75*, 381.

Poll, M., Gosse, J.-P., & Orts, S. (1982). Le genre Campylomormyrus Bleeker, 1874, étude systematique et description d'une espece nouvelle (Pisces, Mormyridae). *Bulletin de l'Institut Royal des Sciences de Belgique Biologie*, *54*(5), 1–34.

Potter, S., Bragg, J. G., Blom, M. P., Deakin, J. E., Kirkpatrick, M., Eldridge, M. D., & Moritz, C. (2017). Chromosomal speciation in the genomics era: Disentangling phylogenetic evolution of rock-wallabies. *Frontiers in Genetics*, *8*(FEB), 1–18.

Ráb, P., Yano, C. F., Lavoué, S., Jegede, O. I., Bertollo, L. A., Ezaz, T., Majtánová, Z., De Oliveira, E. A., & Cioffi, M. B. (2016). Karyotype and Mapping of Repetitive DNAs in the African Butterfly Fish Pantodon buchholzi, the Sole Species of the Family Pantodontidae. *Cytogenetic and Genome Research*, *149*(4), 312–320.

Rezác, M., Arnedo, M. A., Opatova, V., Musilová, J., Rezácová, V., & Král, J. (2018). Taxonomic revision and insights into the speciation mode of the spider Dysdera erythrina species-complex (Araneae: Dysderidae): sibling species with sympatric distributions. *Invertebrate Systematics*, *32*(1), 10.

Rich, M., Sullivan, J. P., & Hopkins, C. D. (2017). Rediscovery and description of Paramormyrops sphekodes (Sauvage, 1879) and a new cryptic Paramormyrops (Mormyridae: Osteoglossiformes) from the Ogooué River of Gabon using morphometrics, DNA sequencing and electrophysiology. *Zoological Journal of the Linnean Society*, *180*(3), 613–646.

Romanenko, S. A., Serdyukova, N. A., Perelman, P. L., Trifonov, V. A., Golenishchev, F. N., Bulatova, N. S., Stanyon, R., & Graphodatsky, A. S. (2018). Multiple intrasyntenic rearrangements and rapid speciation in voles. *Scientific Reports*, *8*(1), 2–10.

Schliewen, U. K., Tautz, D., & Pääbo, S. (1994). Sympatric speciation suggested by monophyly of crater lake cichlids. *Nature*, *368*(6472), 629–632.

Schluter, D. (1996). Ecological speciation in postglacial fishes. *Phil. Trans. Royal Soc. London Series B*, *351*, 807–814.

Schwassmann, H. O., Assunção, M. I. S., & Kirschbaum, F. (2014). Ontogeny of the electric organs in the electric eel, Electrophorus electricus: Physiological, histological, and fine structural investigations. *Brain, Behavior and Evolution*, *84*(4), 288–302.

Seehausen, O. (2000). Explosive speciation rates and unusual species richness in haplochromine cichlid fishes: Effects of sexual selection. *Advances in Ecology Research*, *31*, 237–274.

Seehausen, O., Van Alphen, J. J., & Witte, F. (1997). Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science*, *277*(5333), 1808–1811.

Shastry, B. S. (2009). SNPs: Impact on Gene Function and Phenotype. In A. Komar (Ed.) *Single Nucleotide Polymorphisms: Methods in Molecular Biology (Methods and Protocols)*, vol. 578. Totowa, New York: Humana Press.

Stiassny, M. L. J., Teugels, G. G., & Hopkins, C. D. (2007). *The Fresh and Brackish Water Fishes of Lower Guinea, West-Central Africa*. IRD Editions, vol. 42 ed.

Stoddard, P. K. (2006). Plasticity of Electric Organ Discharge Waveform: Contexts, Mechanisms, and Implications for Electrocommunication. In F. Ladich, S. P. Collin, P. Moller, & B. G. Kapoor (Eds.) *Communication in Fish*, vol. 2, chap. 22, (pp. 623–646). Science Publishers.

Stoddard, P. K., & Markham, M. R. (2008). Signal Cloaking by Electric Fish. *BioScience*, *58*(5), 415–425.

Sullivan, J. P., Lavoué, S., Arnegard, M. E., & Hopkins, C. D. (2004). AFLPs resolve phylogeny and reveal mitochondrial introgression within a species flock of African electric fish (Mormyroidea: Teleostei). *Evolution*, *58*(4), 825–841.

Sullivan, J. P., Lavoué, S., & Hopkins, C. D. (2000). Molecular systematics of the African electric fishes (Mormyroidea: teleostei) and a model for the evolution of their electric organs. *The Journal of experimental biology*, *203*(Pt 4), 665–83.

Sullivan, J. P., Lavoué, S., & Hopkins, C. D. (2016). Cryptomyrus: A new genus of mormyridae (teleostei, osteoglossomorpha) with two new species from gabon, West-Central Africa. *ZooKeys*, *561*, 117–150.

Supiwong, W., Pinthong, K., Seetapan, K., Saenjundaeng, P., Bertollo, L. A., De Oliveira, E. A., Yano, C. F., Liehr, T., Phimphan, S., Tanomtong, A., & Cioffi, M. B. (2019). Karyotype diversity and evolutionary trends in the Asian swamp eel Monopterus albus (Synbranchiformes, Synbranchidae): A case of chromosomal speciation? *BMC Evolutionary Biology*, *19*(1), 1–9.

Swapna, I., Ghezzi, A., York, J. M., Markham, M. R., Halling, D. B., Lu, Y., Gallant, J. R., & Zakon, H. H. (2018). Electrostatic Tuning of a Potassium Channel in Electric Fish. *Current Biology*, *28*(13), 2094–2102.e5.

Taverne, L. (1968). Ostéologie du genre Campylomormyrus Bleeker (Pisces Mormyriformes). *Bulletin de la Société Royale de Zoologie Belge*, *98*, 1–41.

Tiedemann, R., Feulner, P. G. D., & Kirschbaum, F. (2010). Electric organ discharge divergence promotes ecological speciation in sympatrically occurring African weakly electric fish (Campylomormyrus). In M. Glaubrecht (Ed.) *Evolution in Action: Case studies in Adaptive Radiation, Speciation and the Origin of Biodiversity*, (pp. 307–321). Berlin: Springer-Verlag Berlin Heidelberg.

Tsukamoto, K., Aoyama, J., & Miller, M. J. (2002). Migration, speciation, and the evolution of diadromy in anguillid eels. *Canadian Journal of Fisheries and Aquatic Sciences*, *59*(12), 1989–1998.

Uyeno, T. (1973). Teruya Uyneo. *Japanese Journal of Ichthyology*, *20*(4), 211–217.

Wagner, C. E., & McCune, A. R. (2009). Contrasting patterns of spatial genetic structure in sympatric rock-dwelling cichlid fishes. *Evolution*, *63*(5), 1312–1326.

Worm, M., Kirschbaum, F., & von der Emde, G. (2018). Disembodying the invisible: Electrocommunication and social interactions by passive reception of a moving playback signal. *The Journal of Experimental Biology*, *221*(5), jeb172890.

Zakon, H. H., Lu, Y., Zwickl, D. J., & Hillis, D. M. (2006). Sodium channel genes and the evolution of diversity in communication signals of electric fishes: Convergent molecular evolution. *Proceedings of the National Academy of Sciences*, *103*(10), 3675–3680.

Zakon, H. H., Zwickl, D. J., Lu, Y., & Hillis, D. M. (2008). Molecular evolution of communication signals in electric fish. *Journal of Experimental Biology*, *211*(11), 1814–1818.

Zhang, J. (2003). Evolution by gene duplication: An update. *Trends in Ecology and Evolution*, *18*(6), 292–298.

10 Acknowledgements

First of all, I would like to thank my supervisor Prof. Dr. Ralph Tiedemann for the opportunity to do my PhD at his group at the University of Potsdam. I appreciated his support, the discussions about concepts and approaches as well as his constructive criticism. Thank you for the chances I had implementing the project and for your confidence in my work.

Furthermore, I thank Prof. Dr. Michael Hofreiter for acting as second supervisor and member of my defense committee.

I am grateful for the help of Prof. Dr. Frank Kirschbaum. His knowledge about the fish is incredible and I acknowledge his help in some confusing situations. Thank you for acting as member of my defense committee.

I also would like to express my sincere gratitude to Dr. Ina Maria Schedina, who helped me to understand the world of transcriptomes, and to Dr. Alice Dennis for discussing new ideas. Furthermore, I thank Dr. Rebecca Nagel for being the best office mate ever and for highly qualitative discussions about the *Campylomormyrus* project.

I would like to thank all those whose assistance proved to be a milestone in the accomplishment of my final goal. In particular, Anja Ernst, Katja Havenstein, Tonio Pieterek who not just lend me their support in the laboratory, but who were also very encouraging. But the rest of the group Evolutionary Biology/Systematic Zoology also deserves this thanks. I appreciated the help of Sofia Paraskevopoulou and Marijke Autenrieth in computer matters and discussing related problems.

I would like to thank the rest of my thesis committee: Prof. Rüdiger Krahe and Prof. Walter Salzburger, and Prof. Jörns Fickel. I am grateful for financial and logistic support from the University of Potsdam throughout my PhD studies.

Finally, thanks to my family and friends. I would not have reached my goals without the support of my mother, father, sister, and their partners. You are amazing and my tower of strength, not only during the recent years. Thank you for your help in any way, your patience, your sympathy and words of encouragement in the more stressful periods of my PhD. You have no idea how much you supported me.