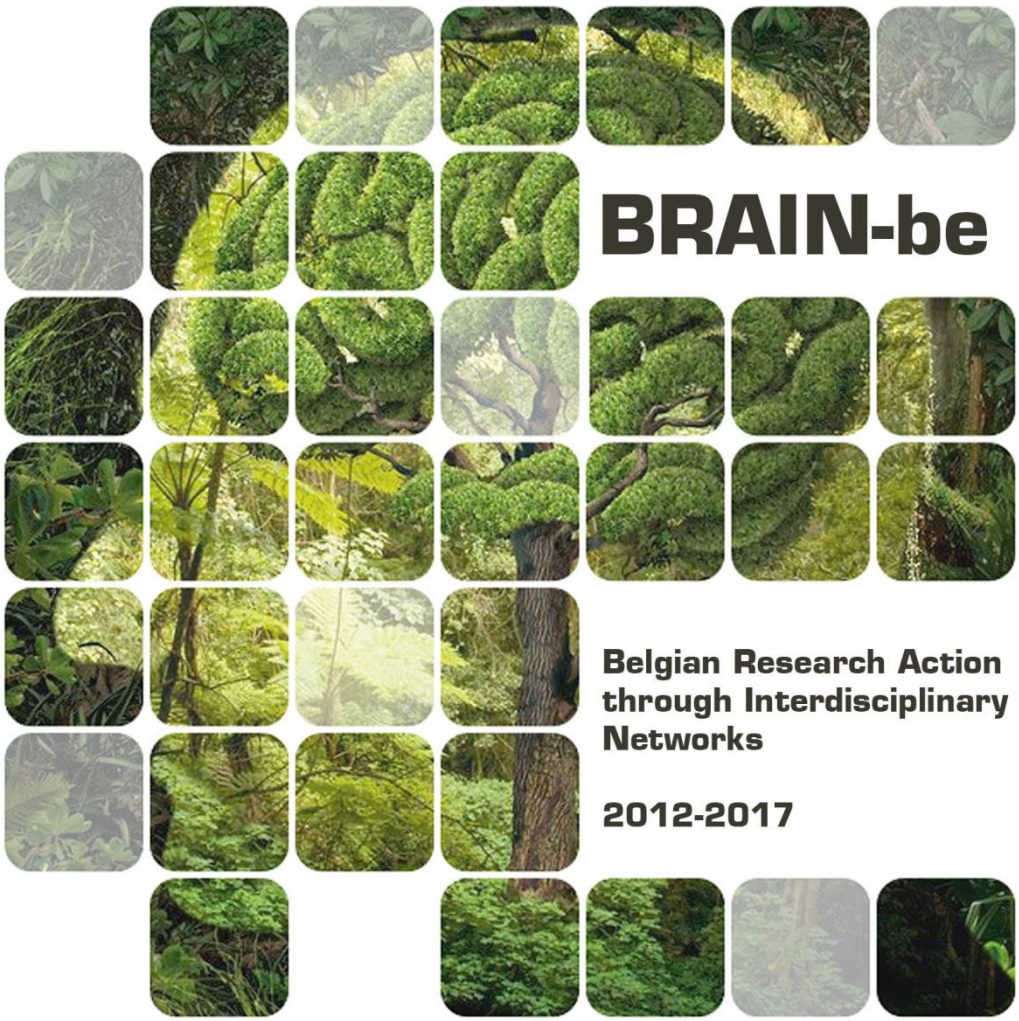

GENBAS

Genomic and Behavioural Aspects of Cichlid Speciation

Maarten Van Steenberge (RBINS), Jos Snoeks (RMCA), Gregory Maes (KU Leuven), Pascal Poncin (ULiège), Sofie Derycke (RBINS), Loïc Kéver (ULiège), Eric Parmentier (ULiège), Noémie Jublier (ULiège), Álvaro Cortés Calabuig (KU Leuven), Sophie Gresham (RBINS), Jeroen Van Houdt (KU Leuven), Erik Verheyen (RBINS).

Axis 1: Ecosystems, biodiversity and evolution

belspo .be

NETWORK PROJECT

GENBAS

Genomic and Behavioural Aspects of Cichlid Speciation

Contract - BR/132/A1/GENBAS

FINAL REPORT

PROMOTORS: JOS SNOEKS (RMCA)
GREGORY MAES (KU Leuven)
PASCAL PONCIN (ULiège)
ERIK VERHEYEN (RBINS)

AUTHORS: MAARTEN VAN STEENBERGE (RBINS)
JOS SNOEKS (RMCA)
GREGORY MAES (KU Leuven
PASCAL PONCIN (ULiège)
SOPHIE DERYCKE
LOÏC KEVER (ULiège)
ERIC PARMENTIER (ULiège)
NOEMIE JUBLIER (ULiège)
ALVARO CORTES CALABUIG (KU Leuven)
SOPHIE GRESHAM (RBINS)
JEROEN VAN HOUDT (KU Leuven)
ERIK VERHEYEN (RBINS)

Published in 2021 by the Belgian Science Policy Office
WTCIII
Simon Bolivarlaan 30 Boulevard Simon Bolivar
B-1000 Brussels
Belgium
Tel: +32 (0)2 238 34 11 - Fax: +32 (0)2 230 59 12
http://www.belspo.be
http://www.belspo.be/brain-be

Contact person: Koen Lefever (koen.lefever@belspo.be)
Tel: +32 (0)2 238 35 51

Maarten van Steenberge, Jos Snoeks, Gregory Maes, Pascal Poncin, Sofie Derycke, Loïc Kéver, Eric Parmentier, Noémie Jublier, Álvaro Cortés Calabuig, Sophie Gresham, Jeroen Van Houdt, Erik Verheyen. Genomic and Behavioural Aspects of Cichlid Speciation. Final Report. Brussels : Belgian Science Policy Office 2021 – 49 p. (BRAIN-be - (Belgian Research Action through Interdisciplinary Networks))

TABLE OF CONTENTS

ABSTRACT

Adaptive radiations, such as those find in cichlid fish, often harbour closely related species with overlapping niches and distribution ranges. Such species are known to sometimes hybridize in nature, which raises the question how they coexist without merging into one hybrid species. Both behavioural and molecular mechanisms have been invoked to explain these patterns of co-existence, however, very few studies have looked at both mechanisms simultaneously on a same study system. Often, the high number of potentially hybridising species, and the uncertainties in species delimitation, limits such studies. In GENBAS (GENomic BAsis of Speciation), we used *Ophthalmotilapia* cichlids from Lake Tanganyika as a model system. This small genus of ecologically and behaviourally similar species was shown to hybridise under natural conditions, although the species could still be easily distinguished. By combining ethological and genomic approaches, we were able to infer the extent of hybridization and introgression among species within Ophthalmotilapia; to provide a starting point for candidate genes involved in the origin of a pair of sister species (speciation) and to infer the phylogeny of the four Ophthalmotilapia species and outgroups from the Ectodini to ascertain their sister species status.

Keywords: Speciation, Genomics, Gene Expression, Mate Pairing, Acoustic Communication

1. INTRODUCTION

"Why are there so many species on earth", "How are species formed?", "How are species barriers maintained?" Such questions lie at the centre of evolutionary biology. Hence, these questions have been addressed using a multitude of model systems and using several approaches including behavioural, ecological and molecular methods. With the advent of genomic sequencing techniques, a new toolbox opened itself to this field. However, every approach has its limitations and by relying solely on one toolbox, input from other field of research might be missed.

In the GENBAS project, we examined how species diversity can be maintained by looking at a single model system: *Ophthalmotilapia* from Lake Tanganyika combining behavioural, transcriptomic, genomic and genetic data. More specifically, our experimental design involves the combination of four studies: (1) to explore behavioral aspects of mate pairing during con- and hetero-specific interactions among species; (2) to investigate differential gene expression in the female brain associated with pairing behaviour during con- and hetero-specific interactions prior and immediately after mating; (3) to elucidate genetic differences within and among relatively recently diverged and hybridizing species at a genome-wide level; and (4) to examine whether the part of the genome that is expressed during mate recognition in the brains of the female cichlids correspond with one or more of the genomic areas with increased differentiation within or among these species.

2. STATE OF THE ART AND OBJECTIVES

Goals

The objective of this project is to investigate the genetic basis of the speciation process itself through the use of an integrated behavioural and genomic approach on the Tanganyika cichlid genus *Ophthalmotilapia* that we have selected as a suitable model. The general objective of our proposal is to characterize the genomic differentiation that drives the speciation process and to verify whether/and to what extent the same genomic changes are involved in to maintain the 'integrity' of the gene pools of the resulting sister species.

We have attempted to address this question through a dual approach: one that compares the genome wide differentiation within and between sister species and one that explores the genetic basis of the behavioural aspects (display and acoustic signalling during courtship) that allow female fishes to differentiate between con-specific and hetero-specific males. More specifically, our study design was designed in order to (1) experimentally explore behaviour during of mate pairing during con and hetero-specific interactions prior and after mating (2) to investigate differential gene expression in the female brain associated with pairing behaviour during con- and hetero-specific interactions; (3) to attempt to elucidate genetic differences within and among relatively recently diverged and hybridizing species at a genome-wide level; and (4) examine whether the part of the genome that is expressed during mate recognition in the brains of the female cichlids correspond with one or more of the genomic areas with increased differentiation within or among these species.

Model system

Cichlid radiations in the East African Great Lakes

The cichlid faunas from the East African Great Lakes: Tanganyika, Malawi and Victoria, contain the most species-rich and phenotypically-diverse adaptive radiations in vertebrates. They are also among the best-documented examples of one of the least understood processes in speciation: the sympatric evolution of reproductive isolation (Seehausen & van Alphen 1999). Several properties of cichlids, such as their unique jaw morphology, have been suggested to explain their explosive evolution. However, as a large proportion of these closely related species live in sympatry without apparent eco-morphological differences the question remains how such species maintain the integrity of their gene pool. It has been suggested that reproductive isolation in East African cichlids was mostly maintained through prezygotic isolation (Turner et al. 2001).

In GENBAS, we will use the Lake Tanganyika cichlid genus *Ophthalmotilapia* Pellegrin, 1904 as a model to study the genomic and behavioural basis of pre-zygotic isolation. *Ophthalmotilapia* species are maternal mouth brooders that occur at the rocky and intermediate (rocky patches separated by sand) shores of Lake Tanganyika. The genus contains four currently valid species: *O. ventralis* (Boulenger 1898), *O. boops* (Boulenger 1901), *O. heterodonta* (Poll & Maréchal 1962) and *O. nasuta* (Poll & Maréchal 1962, Hanssens et al. 1999). They are sexually dimorphic, maternal mouthbrooders in which males are territorial and protect a spawning site. *Ophthalmotilapia* males are unique among Great Lake cichlids by the presence of egg-shaped lappets at the distal ends of their greatly enlarged pelvic fins. The four species have very different, but partially overlapping, distribution ranges (Fig. 1). In one part of the Lake, up to three species of *Ophthalmotilapia* can be found in sympatry (Hanssens et al. 1999).

Even though specimens of *Ophthalmotilapia* can be unambiguously assigned to one of the valid species (Hanssens et al. 1999), incomplete reproductive isolation was discovered between the species. Nevado et al. (2011) observed that specimens of *O. nasuta* often carried mitochondrial DNA typical of the other species (Fig. 1), whereas the opposite was much less often the case. They argued that this asymmetrical pattern could be explained by asymmetry in reproductive behaviour. This would entail that females of all species would mate with *O. nasuta* males, whereas *O. nasuta* females would be more discriminatory towards heterospecific males. By including samples from several sites along the Lake, Nevado et al. (2011) also found that hybridisation events mostly occurred in the South-East of the Lake, i.e. the region were up to three species of *Ophthalmotilapia* occur in sympatry.

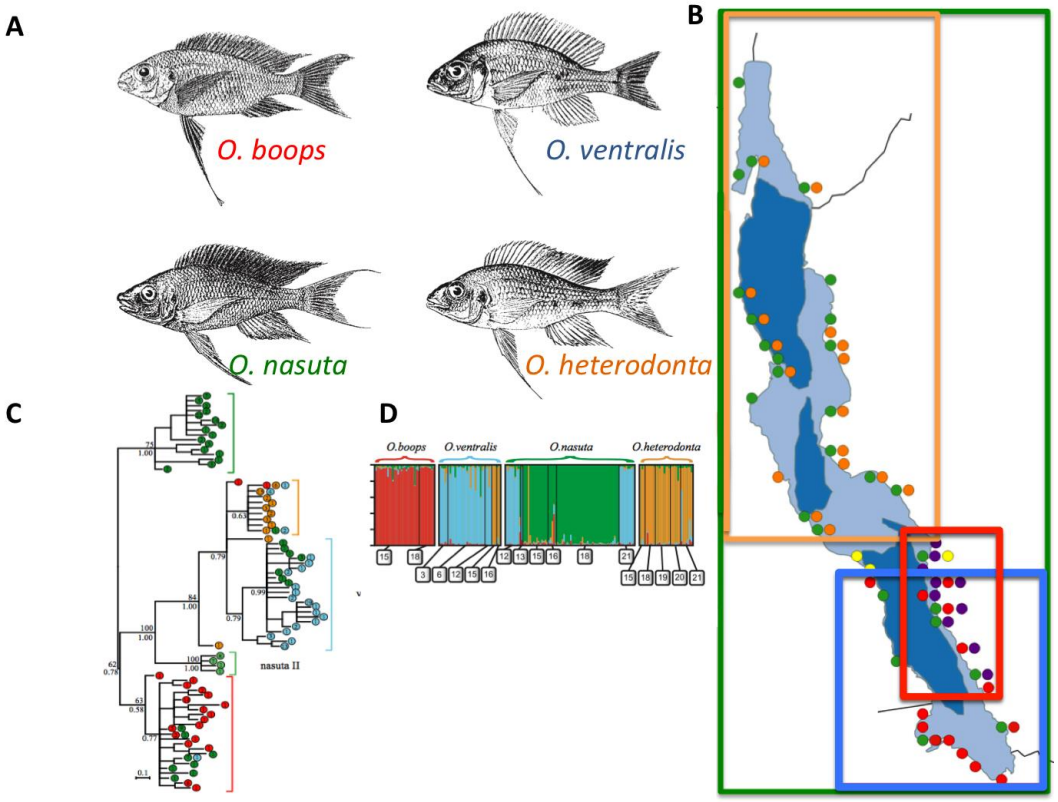

Fig. 1: Overview of the model system, with (A) the four species in the genus *Ophthalmotilapia* (male specimens, figure from Hanssens et al. 1999) and (B) their distribution across the Lake (colours as in A). In a previous study, Nevado et al. (2011) found a mismatch between mitochondrial DNA (C: ML tree derived from the mitochondrial DLoop, colours at nodes indicating species identification), whereas nuclear DNA corresponded with morphological identification (D: output from structure analysis using on MS data).

Aims of the project

The asymmetry in reproductive behaviour in *Ophthalmotilapia* suggests that the barriers to gene flow between the species in the genus have different strengths. This provides us with a natural experiment to identify the behavioural, neurological and genomic mechanisms that allow maintaining species co-existence. An additional advantage of this model system is that it only contains four species that are phylogenetically distinct from other Lake Tanganyika cichlids. This in

contrast with other cichlid taxa from these lakes, in which hybridisation occurs across tens or even hundreds of species (Meier et al. 2017).

Although the study of Nevado et al. (2011) allowed us to discover the patterns of this hybridisation, it fell short of explaining what mechanisms caused this as only a limited number of genetic markers were used. Therefore, we re-examined the population structure of the four species in the genus using a genomic approach. Taking advantage of the annotated cichlid genome (Brawand et al. 2015), this allows us to not only to identify the differentiation, but also to scan the genome for regions that drive differentiation between species and population.

Maintaining species boundaries through pre-zygotic isolation ultimately comes down to a decision process, hence triggering neural processes. The development of RNA sequencing techniques further allowed us to identify 'speciation genes' as it allowed us to sequence the neural transcriptomes of fishes in different stages of the mating process. By comparing these transcriptomes obtained from specimens during the mating process with genomic patterns of differentiation between the species, we aimed to identify the genes responsible maintaining the species barriers.

Approach

We aimed to identify the speciation genes in *Ophthalmotilapia* by following two different disciplines on the same samples and datasets. These can be illustrated by three independent pillars: A) pairing experiments, B) gene expression and C) population genomics (Fig. 2). Although these pillars could be considered stand-alone research projects, additional insights were gained by performing them in an interconnected way. (See: 3 Methodology)

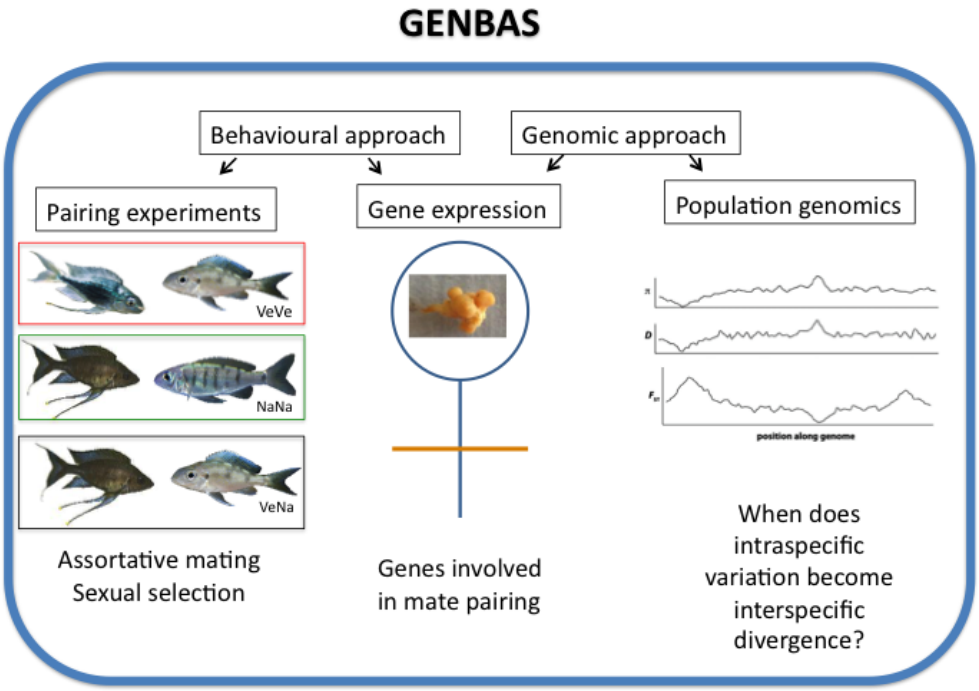

Fig 2: Outline of the project. Using insights obtained from two different disciplines: Animal behaviour and Genomics, the research could be divided into three, partially overlapping pillars: pairing experiments, gene expression studies and population genomics. The combination of these three pillars allowed us to identify speciation genes.

Situation of the project within an international context

Until recently, most studies attempted to clarify to which extent speciating populations need to be geographically isolated (allopatric, peripatric, parapatric, and sympatric), whether sexual selection could be identified as a driving factor, or whether speciation rates are relatively constant over time or occurs only over relatively brief intervals, often assumed to be triggered by an environmental event. All these questions have been extensively addressed in the cichlid fishes from the Great African lakes (e.g. see publications of coordinator and partners). These cichlids are often cited as one of the ultimate animal models to investigate adaptive radiation and speciation in vertebrates (Kocher 2004). The recent advent of novel techniques that allow to explore the genomic aspects concerning the evolutionary processes resulting in adaptive radiation and speciation in these organisms have for the most part been focused on linking genes with phenotypes 'a priori' defined genetic differences and to investigate the genetic basis for interspecific differences, and the repeated evolution of the same phenotypic difference during independent episodes of speciation (Allender et al. 2003; Elmer et al. 2010; Hulsey et al. 2010; Loh et al. 2013).

The project fills a gap in cichlid research by being the first to attempt to investigate the genetic basis of the speciation process itself through the use of an integrated behavioural and genomic approach on the Tanganyika cichlid genus *Ophthalmotilapia that we have selected* as a suitable model. The general objective of our proposal is to characterize the genomic differentiation that drives the speciation process and to verify whether/and to what extent the same genomic changes are involved in to maintain the 'integrity' of the gene pools of the resulting sister species.

Relevance to society, link with the federal competences

Speciation has become a topic of major interest, and thanks to the availability of novel molecular techniques many important advances have been made. We argue that the fundamental insight into the speciation process is of importance for at least two major issues that our society faces: First, the management of the declining biodiversity under the pressure of human induced environmental changes – a federal competence - secondly, the growing debate on the modalities of evolution and more specifically the impact of creationist thinking on many of the younger citizens in our otherwise highly educated society.

3. METHODOLOGY

3.A Pairing experiments

In this pillar, we examined reproductive barriers by investigating the behavioural responses towards potential mates. This was done by recording the behavioural responses of males and females towards potential mates of the four *Ophthalmotilapia* species under con- and heterospecific conditions. However, mating can be considered an end point of a complex decision making process. During this process, several potential mates can be evaluated and certain behavioural steps might be required. In order to encompass both the very first (first encounter and mate recognition) and the last step (oviposition and fertilisation), two different sets of experiments were set up. This is a deviation from the original project proposal were only the latter set of experiments were planned.

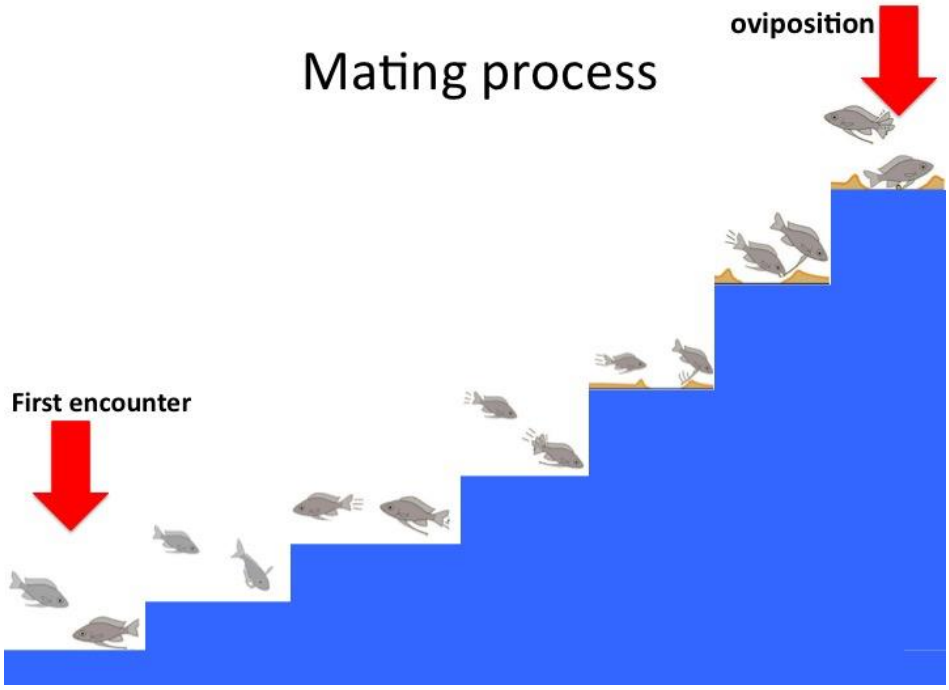

Fig. 3 Natural mating repertoire of *Ophthalmotilapia* species. In nature, a mating event starts takes a few discrete steps (see 4. results), each of which could be repeated, lead to a next step, or lead to a termination of the mating. However, as the first (encounter) and the last (oviposition and fertilisation) steps of this process are fixed, we studied them in a set of experiments.

In both sets of experiments, the behavioural responses of males and females of the four *Ophthalmotilapia* species was compared to both con- and heterospecific mates. For this, specimens of selected populations (see pillar C) of *O. boops*, *O. ventralis* and *O. nasuta* were obtained from a commercial supplier (cichlidenstadl, Germany). Specimens were all wild-caught and stem from populations were the three species co-occur naturally. As the specimens were acquired were either juveniles or subadults, they were kept to reach maturity in the aquarium rooms of the ULg and the RBINS. Here, fishes were also housed in between experiments. They were kept in monospecific tanks (photoperiod: 12:12 h L:D; water temperature: 26 ± 1 °C; carbonate hardness: > 8 dKH) and were fed once a day with 'Tropical Spirulina forte' mini-granules ad libitum. Water was changed (2/3 replacement) and tanks were cleaned every week. Prior to the establishment of the aquaria, an expert (S. Loose) was invited to provided us with advice on how to best keep the fishes. Since some

issues were encountered with the housing of *O. boops*, it was decided that the focus of the experiments was mostly on *O. ventralis* and *O. nasuta*.

Specimens used in the pairing experiments were also used for the study on gene expression (see pillar B). Hence, after the experiments, specimens were euthanized following a protocol approved by the ethical committee of the University of Liège. This also allowed us to examine the stage of gonadal development.

3.A.1: Mating study

In a first study, male and female specimens of the tree different species were housed in different tanks and constantly monitored until mating took place. Besides serving for pillar 2 (see further); this allowed us to describe the mating system of the species. We studied both conspecifc and heterospecific encounters by constructing monospecific (males and females belong to the same species) and heterospecific (males and females belong to different species).

At the onset of the experiments, three males and four or five females, all adults, were introduced simultaneously in monospecific tanks with a single water circulation system and a similar layer of sand on the bottom. Hiding places and flat rocks were also provided (Fig. 4). Surprisingly, this setup worked well only for *O. nasuta*. For this species, courting behaviours were observed within weeks. In the tanks of *O. ventralis* and *O. boops*, courting behaviours were also observed within weeks but the constant aggressiveness of the males towards the females seemed to prevent the latter from mating. The content of the monospecific tanks for *O. ventralis* and *O. boops* were changed with a single male and three to four females per tank. Again, hiding places, sand and flat rocks were provided. In these setup, it took several weeks before a first reproduction was observed. After successful conspecific reproductions were recorded, three females of *O. ventralis* were kept with one male of *O. nasuta* in a heterospecific tank. Because studying the mating behaviour of O. boops in the lab was shown to be difficult, we choose not to focus on this species in the remainder of the experiments.

During the cause of these experiments, the light conditions were altered to allow an experimenter to be present in the room when lights were turned on. Matings were recorded using a camera directed at the tank. A hydrophone [HTI Min-96, −164.4 dB re. 1 V µPa−1; bandwidth 2 Hz and 30 kHz (Long Beach, MS, USA), or a Brüel and Kjær 8101 hydrophone connected to a calibrated Brüel and Kjær 2610 am- plifier, sensitivity −184 dB re. 1 V µPa−1; bandwidth 0.1 Hz to 200 kHz (Nærum, Denmark)] was placed in each tank, at half the height of the water column, over or just next to the bower. We connected this hydrophone to a Tascam DR-05 recorder (TEAC, Wiesbaden, Germany) at a sampling rate of 44.1 kHz.

Fig. 4 Typical tank used for the mating experiment. The tanks contains both sand and flat stones that allow for bower building (necessary for the mating behaviour of the species) and hiding places were females and subordinate males can find shelter from the aggression of the dominant male.

The recordings of reproduction events were analysed starting from 20min before the first egg was laid to 10min after the last egg was laid. In these recordings, nineteen behaviours could be identified (Table II). These were encoded in point and state events and analysed using the Boris v 2.72 open source software (Friard and Gamba, 2016). Sounds emitted by *Ophthalmotilapia* were shown to be very weak and could therefore only be analysed when produced close to the hydrophone. A successful mating can consist of several spawning bouts (separate cycles of oviposition and fertilisation, Fig 2.). After the last spawning bout, we waited for at least 10 minutes. After this period, we assumed the mating event was over and the female was isolated, euthanized and dissected (see 3.B). The weight and size of both partners used in a reproduction event were collected. Euthanized fishes were stored in ethanol for further reference.

Table II: Behaviours used for the quantitative analysis of *Ophthalmotilapia* reproductions. PE: point event; SE: state event. Modifiers: different modifiers were used when a behaviour could be performed in different contexts.

PE	Bite	Fish bites another fish
PE	Chase	Fish swims very fast towards another fish
SE	Clean	Fish removes dirt/sand to clean/build the pit
PE	Follow	Fish swims behind another fish
PE	Frontal display	Fish spreads his pectorals, swells is throat, head high
SE	Inside	Fish is inside the nest (less than one body length from the pit floor)
PE	Invite	Fish performs a fast and unilateral contraction of its body musculature, generally head down, close to another fish
PE	Lateral display	Fish presents his flank, swells is throat and leans slightly
PE	Lay	Fish lies an egg
PE	Lead	Fish swims at the front of another fish and guides it to the nest
SE	Pelvics flickering	Fish alternatively moves its pelvic fins up and down (fast movements)
PE	Present	Fish presents his pelvic fins to the mouth of a second fish, generally egg dummies are on the pit floor
PE	Present over the back	Fish abduct the pelvics slightly swimming just over the female. The latter being inside the nest
PE	Show	Fish stops at the entrance of the nest and tail waggle.
PE	Sound	Fish emits an acoustic signal of interest

SE	Tail wagging/ swimming on the spot	Fish shakes the caudal part of its body (caudal fin and caudal part of the dorsal fin) with head slightly up or down + Fish performs alternate forward movements of the pectoral.
PE	Take	Fish takes or tries to take something in mouth (modifiers: eggs or egg dummies)
PE	Tilt	Fish leans on one side
PE	Uncommon	Any uncommon behavior that was recorded in rare occasions

The study of the mating behaviour of *O. ventralis*, *O. boops* and *O. nasuta*, revealed that the tree species had a very similar mating ritual (Fig. 5). This allowed us to define behavioural variables that to score in a subsequent set of experiments (see 3.A.2). The similarity of the courtship behaviour and the mating ritual could explain why hybridisation was encountered in nature. This hypothesis was confirmed by the fact that a normal mating behaviour was also observed in matings between *O. ventralis* females and *O. nasuta* males. However, in such matings fewer eggs were laid and we only managed to obtain hybrid offspring once (out of three heterospecific matings). One typical behaviour, called 'invite' was, however, displayed by males of *O. boops* and *O. ventralis,* but not by males of *O. nasuta*. However, the acceptance of male *O. nasuta* as mates by members of the other species (both in the lab and in the field), suggests that this behaviour is not necessary to obtain successful matings.

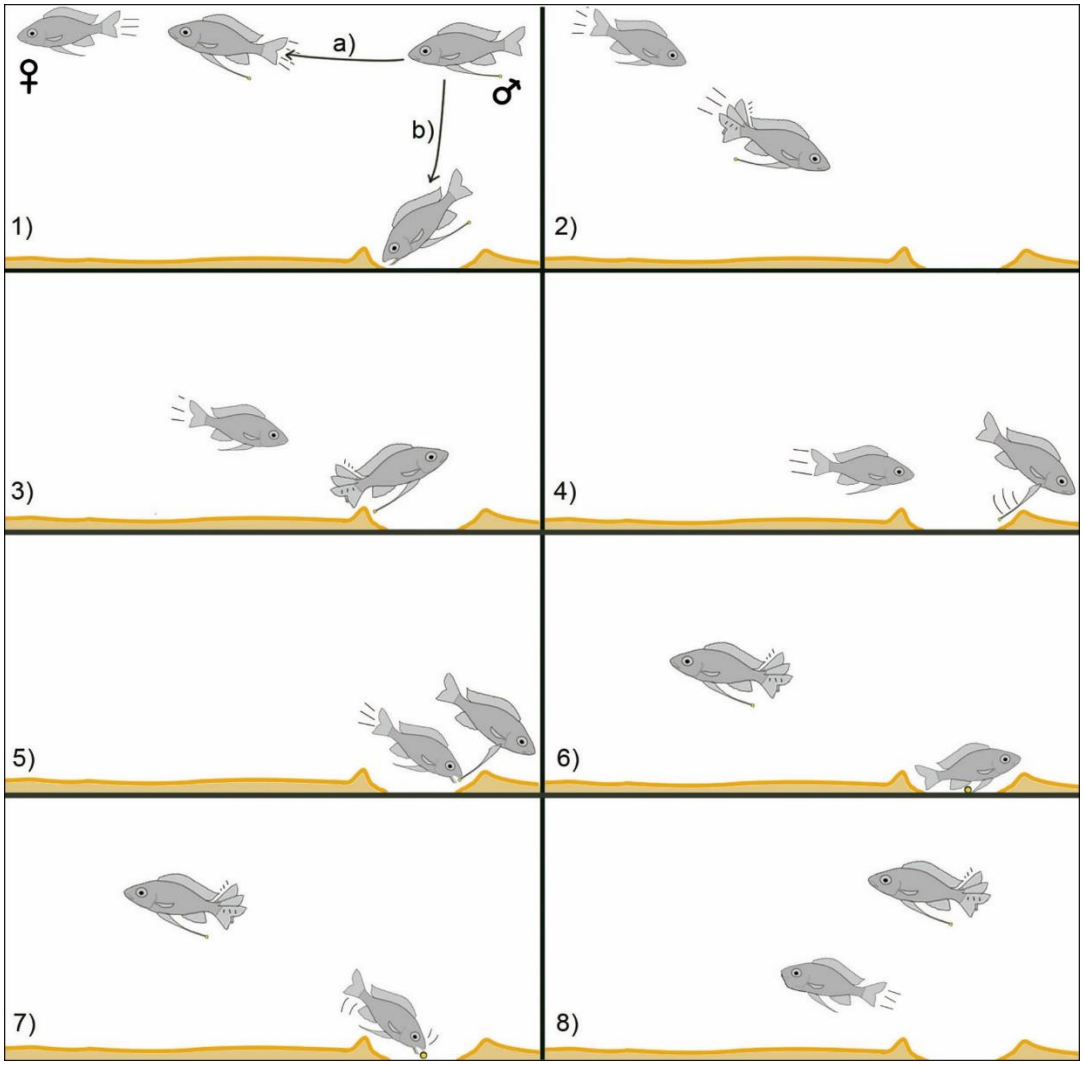

Fig. 5: The typical spawning bout of Ophthalmotilapia ventralis, O. nasuta, and O. boops. (1) The male chases the female and prepares the bower. (2) The female follows the male to his bower. (3) The male shows the bower to the female. (4) The male presents its pelvic fins. (5) The female enters the bower, swims towards the egg dummies, and opens the mouth. (6) The female lays an egg in the bower. (7) The female takes the egg in mouth. (8) The female leaves the bower. Stages 6 and 7 are generally repeated before 8.

3.A.2: Presentation study

A potential flaw in using only data from the experiments described above in detecting speciation genes through RNA sequencing (see 3.B) is that we do measure readiness to mate with a con- or heterospecific male, but that we fail to measure to what degree a female recognises a male as being con- or heterospecific. Additionally, the mating experiments also fail to collect data of females that do not mate with a certain mail, which could introduce a bias in our results. For this aim, we also performed presentation experiments in which we simply presented a female to a con- or a heterospecific male. Although the mail goal of this study was to detect differences in neural gene expression between treatments, it also allowed us to detect a behavioural measure for recognition, defined as ‘‘a measurable difference in behavioural response toward conspecifics as compared to heterospecifics’’ (Mendelson and Shaw 2012).

We performed two independent sets of experiments using females and males of *O. ventralis* and *O. nasuta*. The first focused on the behaviour of focal *O. nasuta* females (ON experiments), the second

on that of focal *O. ventralis* females (OV experiments). We also performed some experiments with *O. boops* males, but given the difficulties we encountered with keeping this fish in the lab (see also 3.A.1, 3.A.3), we stopped doing so after two trials. Specimens stem from the same populations as those examined in 3.A.1. Prior to the onset of the experiment, the sex of the specimens was checked by visually inspecting their genital papillae. Female specimens were kept isolated from males and heterospecific specimens in a separate tank for at least two weeks. During that period, males were kept in monospecific tanks in which they were visually isolated from each other using opaque partitions.

We used three identical experimental aquaria (88cm*50cm*40cm with water level ca. 40cm), which we divided into two equal parts by a perforated transparent partition (separation wall), through which fishes could not pass, and by an opaque wall (visual barrier) that could be removed (Fig. 6). A flower pot was placed on each side of the separation to allow the fish to take refuge. We kept the fishes in these aquaria for at least twelve hours before they were used in the experiments. During the experiments, the visual barrier was removed.

We recorded the behaviour of focal specimens (*O. nasuta* females or *O. ventralis* females) in four different experimental conditions. They were either exposed to (i) no other specimen, (ii) a conspecific female, (iii) a conspecific male or (iv) a heterospecific male. For each set of experiments (ON and OV) and for each condition, we conducted a minimum of five replicates. We filmed (using a CANON Legria HF R606) the entire aquarium (i.e. focal and non-focal fishes) during one hour: from 15 minutes before to 45 minutes after the visual barrier was removed. Experimenters were only briefly present in the room to remove the visual barrier. As *Ophthalmotilapia* males are known to produce weak-pulsed sounds during the inviting behaviour (see 3.A.1), we record sounds during the whole experiment. The hydrophone was positioned near the separation wall, at half the height of the water column, on the side of the non-focal specimen. At the start of each experiment, we switched off the aeration of the tank so that sounds could be recorded.

After each experiment, both the focal and the non-focal specimens were weighed. Focal specimens were euthanized, measured, dissected (3.B) and the stage of gonad development was scored following Panfili et al. (2006). As we expect behaviour of female fishes to be influenced by the reproductive cycle and by the size difference between the partner, we made sure that the specimens used in different comparisons did not differ in gonadal development or size.

In total, we performed 28 ON and 21 OV experiments, with a maximum of three experiments per day (Table III). However, after the dissections (see below), we observed that six focal *O. nasuta* females from the first set of experiments possessed male or ambiguous gonads. These specimens were referred to as floater males (see 3.A.3) and the recordings for these experiments were not analysed. As we suspected that these specimens had changed sex, we photographed the genital papillae of the focal females that were to be used in the subsequent OV experiments two weeks before the onset of the experiment. A comparison between papillae of the same individuals after two weeks confirmed that a sex change did indeed took place in several specimens. These specimens were not included in the experiments. After each experiment, the aquarium was cleaned and the water fully renewed.

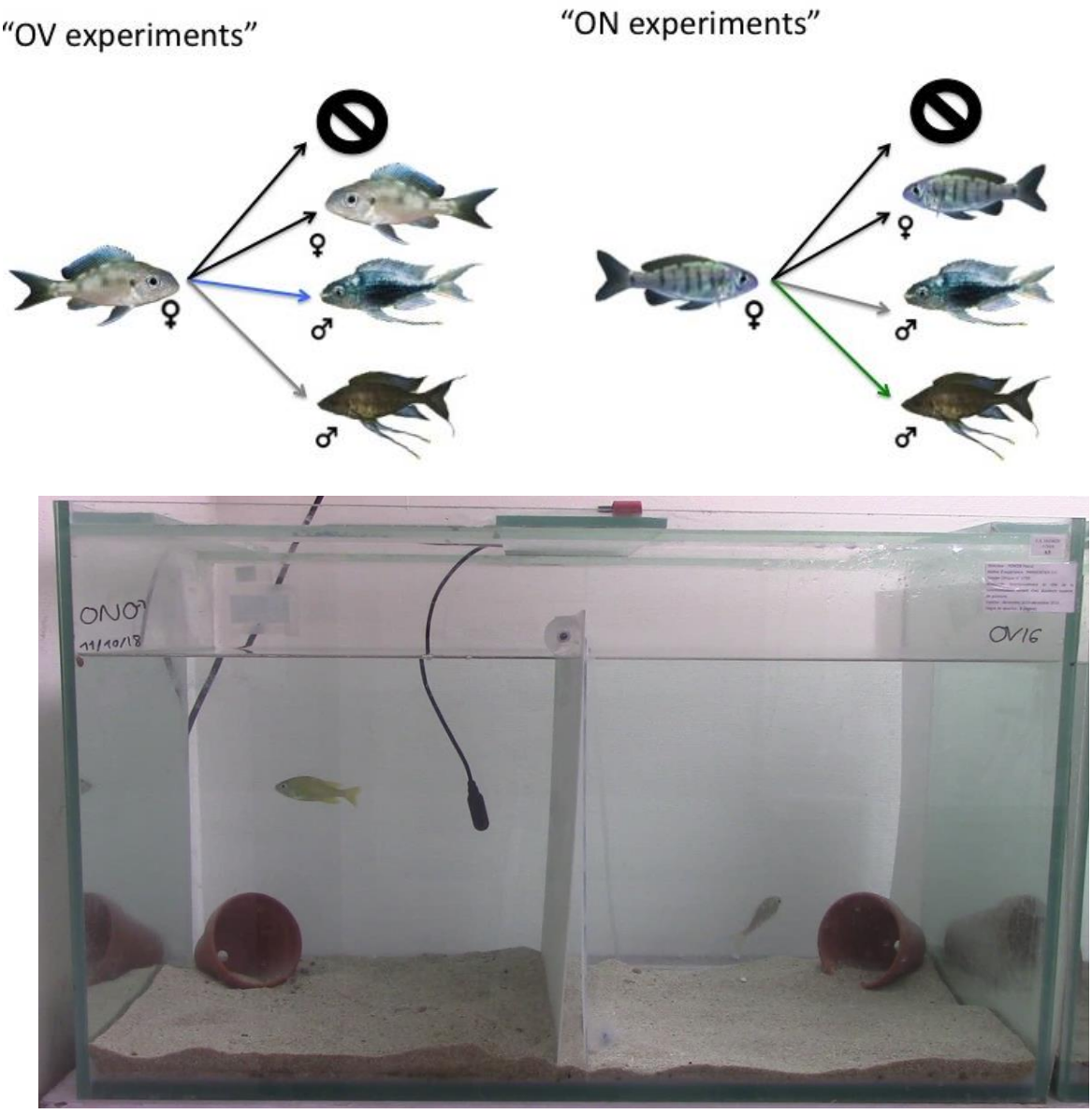

Fig. 6 Experimental setup of the presentation experiments. Above: set-up of the experiments: A focal female of either *O. nasuta* or *O. ventralis* was placed in one half of the experimental tank whereas no fish, a conspecific female or a hetero- or conspecific male was placed in the other half. Below: tank used for the experiments. The tank was divided in two by a transparent wall and a white visual barrier. A microphone (E) was placed on the side of the non-focal specimen and an empty flowering pot was placed in both halves of the tank, allowing fishes to take refuge. Video and audio recordings were made 15 minutes prior and 45 minutes after the visual barriers were removed.

Table III: Summary of the presentation experiments. Specimens indicated with* displayed courtship behaviour and those recordings were therefore not included in the analyses. The same holds for experiments that included 'floater males'. These are specimens for which the external morphology indicated that they were female but which possessed testes or intermediate gonads.

Specimen	Date	Tank	Focal specimen	Side focal	Weight focal	Gonadal stage	Non-focal specimen	Weight non-focal	Comment
ON21	09 Feb 2016	1	O. nasuta floater male	Left	12.7	testes	O. nasuta male	20.7	Not used
ON22	09 Feb 2016	2	O. nasuta female	Right	16	4	O. nasuta female		
ON23	09 Feb 2016	3	O. nasuta female	Right	19.8	3	O. ventralis male	10.28	
ON24	10 Feb 2016	1	O. nasuta female	Left	19.5	5	No fish	/	
ON26	11 Feb 2016	1	O. nasuta floater male	left	19.4	testes	O. ventralis male	15.3	Not used
ON27	11 Feb 2016	2	O. nasuta female	Right	15.4	5	O. nasuta male	12.7	

ON28	11 Feb 2016	3	O. nasuta female	Right	11	3	O. nasuta female	6.16	
ON30	12 Feb 2016	3	O. nasuta female	Right	12.92	5	No fish	/	
ON31	13 Feb 2016	1	O. nasuta female	Left	11	4	O. nasuta female		
ON32	13 Feb 2016	2	O. nasuta floater male	Right	10.5	testes	O. ventralis male	10.01	Not used
ON33	13 Feb 2016	3	O. nasuta floater male	Right	9.55	testes	O. nasuta male	20.1	Not used
ON34	14 Feb 2016	1	O. nasuta female	left	9.88	3	O. ventralis male	15.4	
ON35	14 Feb 2016	2	O. nasuta floater male	Right	7.8		No fish	/	Not used
ON36	14 Feb 2016	3	O. nasuta female	Right	6.7	5	O. nasuta male	20.7	
ON37	15 Feb 2016	1	O. nasuta female	Left	8.7	5	O. nasuta male	20.7	
ON38	15 Feb 2016	2	O. nasuta female	Right	6.76	3	O. ventralis male*	10.1	Not used
ON39	15 Feb 2016	3	O. nasuta female	Right	6.16	3	No fish	/	
ON41	16 Feb 2016	1	O. nasuta female	left	11.8	4	O. nasuta male	12.7	
ON43	16 Feb 2016	3	O. nasuta floater male	Right	12.6	testes	O. ventralis male	10.8	Not used
ON44	28 Aug 2018	1	O. nasuta female	Right	19	4	O. ventralis male	19	
ON45	28 Aug 2018	2	O. nasuta female	Left	11	4	No fish	/	
ON46	28 Aug 2018	3	O. nasuta female	Left	30	3	O. nasuta male	19	
ON47	29 Aug 2018	1	O. nasuta female	Right	11	5	No fish	/	
ON48	29 Aug 2018	2	O. nasuta female	Left	17	5	O. ventralis male	12	
ON49	29 Aug 2018	3	O. nasuta female	Left	13	5	O. nasuta female	11	
ON50	30 Aug 2018	1	O. nasuta female	Right	7	4	O. nasuta female	15	
ON51	30 Aug 2018	2	O. nasuta female	Left	17	4	O. ventralis male*	19	Not used
OV13	10 Oct 2018	1	O. ventralis female	Left	14	2	O. ventralis female	10	
OV14	10 Oct 2018	2	O. ventralis female	Right	16	2	O. ventralis male	16	
OV15	10 Oct 2018	3	O. ventralis female	Left	12	5	O. nasuta male	17	
OV16	11 Oct 2018	1	O. ventralis female	Right	10	2	O. nasuta male	16	
OV17	11 Oct 2018	2	O. ventralis female	Left	12	3	No fish		
OV18	11 Oct 2018	3	O. ventralis female	Right	10	4	O. ventralis male	17	
OV19	12 Oct 2018	1	O. ventralis female	Left	11	1	O. ventralis male	16	
OV20	12 Oct 2018	2	O. ventralis female	Right	8	3	O. ventralis female	10	
OV21	12 Oct 2018	3	O. ventralis female	Right	13	4	No fish	/	
OV22	15 Oct 2018	1	O. ventralis female	Left	9	4	No fish	/	
OV23	15 Oct 2018	2	O. ventralis female	Left	10	4	O. nasuta male	42	
OV24	15 Oct 2018	3	O. ventralis female	Right	8	5	O. ventralis male	17	
OV25	16 Oct 2018	1	O. ventralis female	Right	9	5	O. ventralis male	17	
OV26	16 Oct 2018	2	O. ventralis female	Right	10	3	No fish	/	
OV27	16 Oct 2018	3	O. ventralis female	Left	14	5	O. ventralis female	10	
OV28	17 Oct 2018	1	O. ventralis female	Right	12	4	O. ventralis female	9	
OV29	17 Oct 2018	2	O. ventralis female	Right	9	5	O. nasuta male	18	
OV30	17 Oct 2018	3	O. ventralis female	Left	9	5	No fish	/	
OV31	18 Oct 2018	1	O. ventralis female	Left	9	4	O. nasuta male	22	
OV32	18 Oct 2018	2	O. ventralis female	Left	6	3	O. ventralis female	8	
OV33	18 Oct 2018	3	O. ventralis female	Right	10	4	O. nasuta male	38	

The video footage of the experiments was used to collect two sets of behavioural parameters: point events and tracking data (Table IV). The point events were based on those identified as important in 3.A.1 and were collected using the same software. Tracking data was collected by scoring the position of each fish, at every second, for all of the video recordings. For this video files were converted into JPG images using Adapter v2.1.6, capturing one frame per second and coordinates were collected using ImageJ v1.49 (Schneider et al. 2012). Using these coordinates, six tracking parameters could be calculated that describe the position and swimming behaviour of the fish. Point events and tracking data were analysed jointly and compared between focal fishes of the same species that subjected to a different experiment (different non-focal individual). Following the definition of species recognition discussed above, we concluded that a specimens behaved differently in a certain treatment whenever a significant difference in behaviour was discovered. Data was collected both for the recordings done prior to the onset of the experiment (removal of separation wall), and after this wall was removed. This allowed us to verify whether the behavioural differences observed were induced by the presentation to a second fish.

Table IV. Tracking parameters and point events recorded during the presentation experiments. Data was collected for both the focal and non-focal individuals, for point events, interpretation of the behavior was added.

Tracking parameters	Description
Dist.wall	Distance to the visual barrier (% of length of compartment).
Dist.fish	Distance to the fish on the other side of the transparent wall (cm).
Sp	Average speed (cm/s).
SpX	Horizontal speed (cm/s).
SpY	Vertical speed (cm/s).
Height	Mean height (% of height of water column).
Point events	**Description**
Chase	The fish suddenly swims very fast towards the other fish and rams (or almost rams) the separation wall. (Contextual: agonistic and male courtship behaviour)
Flee	The fish suddenly swims away from the other fish. (Contextual: agonistic and female courtship behaviour)
Lateral	Lateral display: The fish positions itself perpendicular to the other fish, keeping its head slightly downwards, erects its fins and bends its body. (Signal movement, agonistic behaviour)
Frontal	Frontal display: The fish faces the other fish head up and erects its fins. (Signal movement, agonistic behaviour)
Bite	Biting the wall: The fish bites the separation wall (possibly trying to bite the other fish). (Signal movement, agonistic behaviour)
Ram	Ram into the wall: The fish tries to enter the other part of the aquarium and rams (not very fast) the separation wall.
Sand	Sand picking: The fish takes sand in its mouth. (Courtship behaviour when linked to construction of bower, signal movement when nipping off a substrate)
Spasm	A quick, strong, and unilateral contraction of the trunk musculature that results in a displacement of the head and the caudal fin in the same direction. (Contextual: comfort behaviour or signal movement. Observed in courtship behaviour or inter-territorial fights depending on the genus)
Tail	Tail-wagging: Exaggerated movements of the caudal fin (+ caudal part of the dorsal fin). At its zenith, the movement of the caudal fin is completely counterbalanced by backpedalling. (Signal movement, courtship behaviour)
Flicker	Pelvics flickering: The fish quickly and alternatively moves its right and left pelvic fins. (Comfort behaviour)

3.A.3 Difficulties experienced during the mating experiment

Originally, it was planned to study the acoustic behaviour of the species. This because acoustic communication remained underexplored in cichlid fishes, although accounts were made of its importance. Additionally, the lab in Liège is a centre of excellence in acoustic communication in fishes. However, in spite of an 'invite' sound made by males during the mating process (see 4.A.1) of *O. boops* and *O. ventralis*; no sounds were recorded for this species. Hence, the focus of the behavioural part of the project was shifted towards studying the role of visual communication.

Originally, we planned to investigate the reproductive barriers between *O. boops, O. ventralis* and *O. nasuta*. However, keeping *O. boops* in our aquaria prove to be difficult and several specimens died before experiments could be conducted, or even during the experiments. Hence, we gradually abandoned this species and focussed on a comparison between *O. ventralis* and *O. nasuta* instead. This was unfortunate since the majority of hybridisation events recorded in nature were between *O. boops* and *O. nasuta*. However, this change of plans did not alter the outline of the project too much since indications for unidirectional hybridisation into *O. nasuta* was found for *O. ventralis* as well.

The difficulty in keeping *O. boops* was unexpected since we were told that *O. nasuta* was the most aggressive and demanding species to keep (S. Loose, pers. comm.).

Prior to the onset of the presentation experiment (3.A.2), the sex of the specimens was checked by visually inspecting their genital papillae. Female specimens were kept isolated from males and heterospecific specimens in a separate tank for at least two weeks. However, when performing the experiments, we observed that some of the presumed females possessed male or unambiguous gonads (Fig. 7). Since we suspected that sex-changed had occurred, we photographed the genital region of all specimens that were to be used in a second experiment. After two weeks, some of the females had indeed changed into males, which was (somewhat) visible at the level of the papillae and at the length of the pelvic fins (a sexually dimorphic trait in *Ophthalmotilapia*). None of these specimens, however, developed a male colour pattern and they also did not deviate in behaviour from the remaining females. Although sporadic accounts of sex-change in cichlids have been reported by aquarists, this has not yet been mentioned in the literature. Hence, if we would re-run these experiments and remove all possible effects of males on female behaviour, specimens would need to be isolated completely.

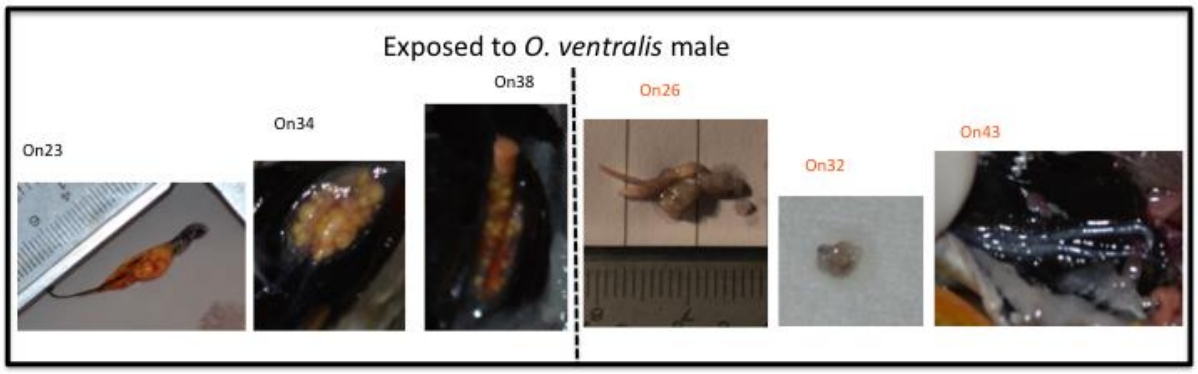

Fig. 7. Dissection of the gonads of focal specimens. Focal specimens of *O. nasuta* that were presented with an *O. ventralis* male in the presentation experiments. All of these specimens were identified as female based on external morphology, but some of them proved to have male (ON32, 43, or ambiguous (ON26)) gonads. Specimen codes as in Table III.

3.B Differential gene expression

3.B.1 Optimisation of RNA sequencing protocol

In this pillar, we aimed to identify changes in the neural transcriptome that could be linked to species recognition and mate choice in *Ophthalmotilapia*. For this, we sequenced RNA from the brain of female specimens that were used in the experiments explained above (3.A.1; 3.A.2). However, since we had no prior experience with RNA sequencing of fish tissue in the lab, we first optimised a protocol on Nile tilapia (*Oreochromis niloticus*). This because specimens of this species were larger and easier to keep than those of *Ophthalmotilapia*. Nile tilapia were made available by the Tilapia breeding facility of the University of Liège and raised till maturity at the facilities of the RBINS. We used these specimens to optimise a standardised dissection protocol (Fig. 8). Because RNA degrades rapidly, we made sure that this dissection could be finalised within a reasonable time (<5min). Additionally, a standardised protocol also allowed us to remove and store the different brain parts: Cerebellum (CE); Optic Tectum (OT); Diencephalon (DI); Brain Stem (BS); Olfactory Bulbs (OB); Telencephalon (TE) (Fig. 7). Tissue samples were stored in RNAlater and kept at -20°C. We

dissected several specimens over a course of three months and compared extracts of fresh (1 day old) and old (up to three months old) samples for RNA quality. RNA extraction was performed using the RNeasy Lipid Tissue Mini Kit from Qiagen following the manufacturer's protocol. Brain tissues were homogenised using pestles and a cordless motor (Sigma Aldrich). Comparing fresh and old RNA extracts revealed that the quantity and quality of RNA was very similar and yielded RNA with RIN values as high as 10. RNA profiles on the Bioanalyzer also showed no signs of degradation, as two sharp peaks were observed for all samples. RNA extracts were of sufficient quantity and quality for RNA sequencing (see 3.B.2). Consequently, we decided that the storage of the brain parts in RNA later at -20°C does not affect RNA. This was relevant as it allowed us with some flexibility in the planning of our experiments.

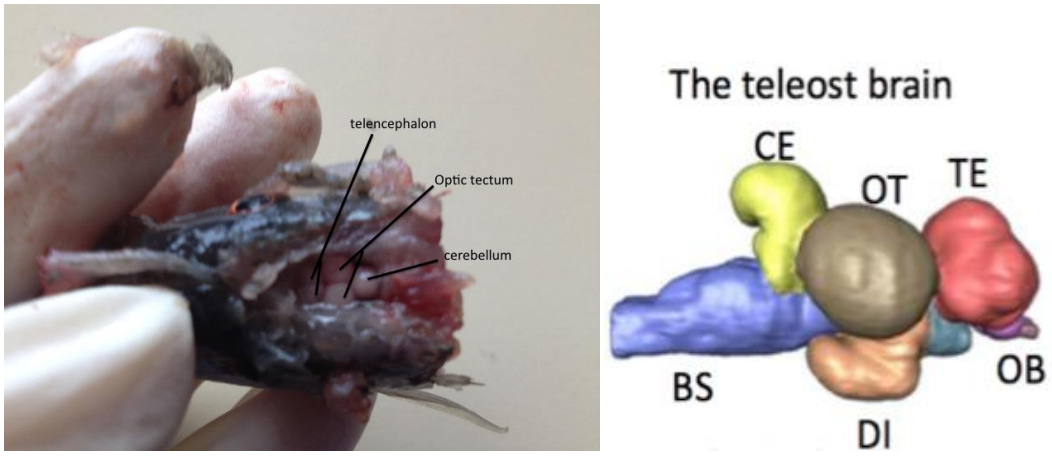

Fig. 8. Dissection of a fish brain. Left: head of a Nile Tilapia dissected for during the optimalisation of the protocol, right: schematic description of the teleost brain with BS: Brain Stem; CE: Cerebellum; OT: Optic Tectum; DI: Diencephalon; TE: Telencephalon; OB: Olfactory Bulbs. The posterior side of the fish is situated at the right side of the brain.

3.B.2 Gene expression in the female brain in a control setting

Although they provided us with new insights in the behaviour of *Ophthalmotilapia* species, the experiments explained above (3.A.1; 3.A.2) were ultimately set up to discover changes in the neural transcriptome that could be linked to species recognition and mate choice. Because of the difficulties encountered with *O. boops*, these experiments were only performed for *O. ventralis* and *O. nasuta*. However, before these could be interpreted, we decided that we needed a baseline for the neural transcriptome of both of the species in the absence of mating. This would allow us to tell to which degree the background transcriptome differed between these closely-related species. This baseline study would also allow us to verify whether different parts of the brain needed to be treated as separately, or whether transcriptomes could be obtained from a pooled brain sample.

Our controls consist of females that have never been in contact with males. These female fish were kept in separate aquaria (one for each species), which were located in the same room, and received the same maintenance throughout a 51 day period. After 51 days, ten females (five for each species) were captured and euthanized. The brain was dissected following the protocol outlined above (3.B.1). The time between catching the female and storing the brain parts in RNA later varied between 14 and 21 minutes. RNA extraction was done in batches of 12 samples that were randomly picked from the 60 samples. Bioanalyzer profiles indicated that the obtained RNA was of high quality

for all samples, but the RNA quantity of the olfactory bulbs proved to be borderline for adequate sequencing. This deviated from the results obtained for the Nile tilapia. Nevertheless, all samples (including the olfactory bulbs) were included for library preparation and sequencing.

The 60 RNA samples (6 brain parts x 5 fishes x2 species) were prepared using the QuantSeq 3' mRNA-Seq Library Prep Kit for Illumina (Lexogen). This method generates sequences close to the 3' end of polyadenylated RNA. Library preparation involved reverse transcription with oligodT primers, followed by removal of RNA and second strand cDNA synthesis with random primers. The resulting fragments were amplified using PCR with primers that also contain the Illumina adaptors and sample specific barcodes. All 60 samples were pooled on one Illumina Hiseq lane and sequenced single end for 50bp. Raw reads were trimmed and mapped to the annotated *Oreochromis niloticus* genome version ASM185804v2. Of the raw reads, between 52 and 70% mapped against the *O. niloticus* genome. Read counts were obtained using HTSeq-count 0.6.1p1 (Anders et al. 2015).

After filtering and normalisation, 11 577 genes were kept in the dataset, the bulk of which were expressed in all 59 samples (7 224 genes, 62.4 %). Transcriptomes, i.e. read counts of fragments that could be mapped, were compared across species and brain parts using a stage wise statistical testing procedure to adequately identify differential gene expression between six brain parts of two cichlid species. One sample was removed from the dataset because of a low read count. The total number of genes expressed in each sample was very similar and ranged between 9 565 and 11 499 (mean 11 201.8). Not a single gene was uniquely expressed in one brain part.

Statistical analyses revealed that the six brain regions profoundly differed in gene expression (Fig. 9). Of the 11577 expressed genes, we found 8748 differentially expressed genes between at least two of the brain parts. These differentially expressed genes belonged to 140 biological processes. Importantly, the expression pattern between both species was very similar since we only detected 32 differentially expressed genes between brain parts across the two species.

3.B.3 Gene expression in the female brain during the experiments

Here, we will investigate the transcriptome for the specimens used in the mating and in the presentation experiments (3.A.1, 3.A.2). The results of the control experiment indicated the importance of comparing the transcriptomes per brain part instead of for the entire brain. Hence, also for the specimens used in the mating and presentation experiments (3.A.1, 3.A.2), the six brain parts were sequenced separately. Additionally, because the control experiment showed the similarity in the background transcriptome of *O. ventralis* and *O. nasuta*, we know we will be able to attribute any differences seen during mating and presentation to different responses, and not to inherent characteristics of the species.

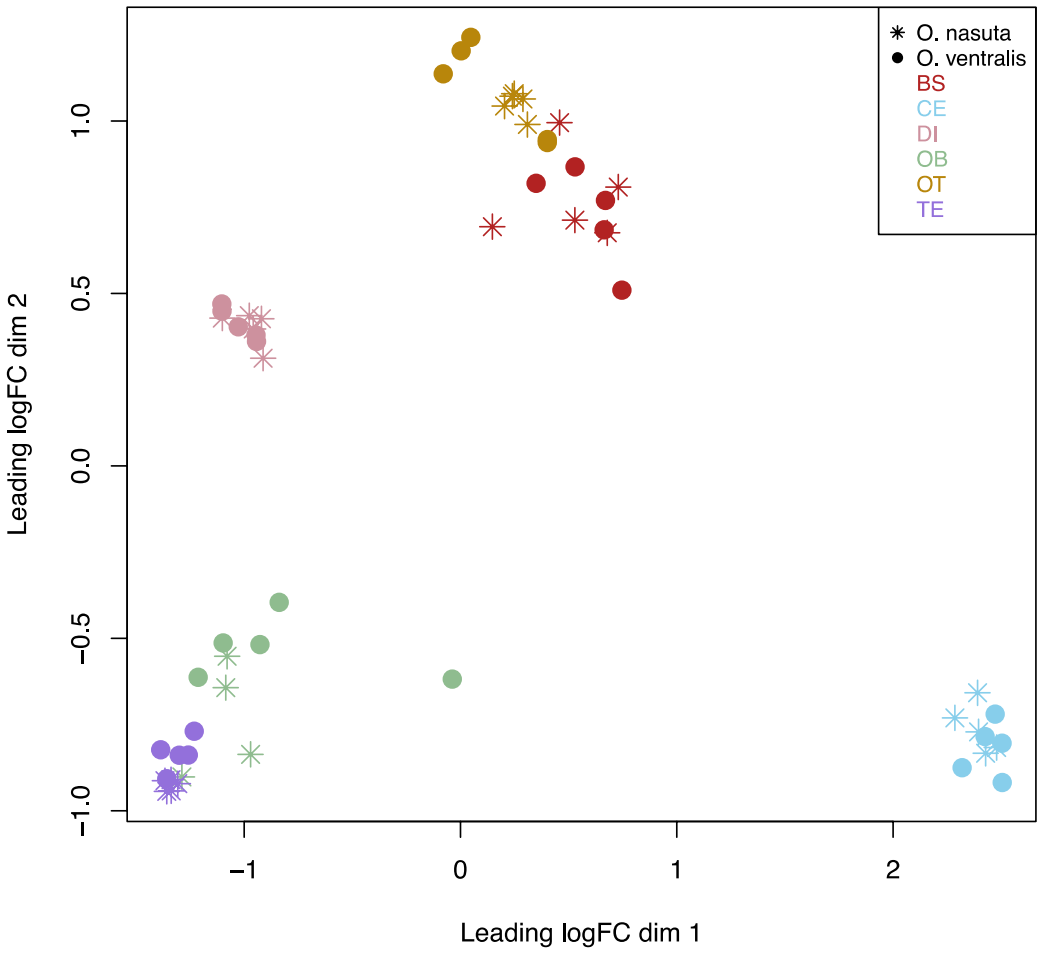

Fig. 9. MDS plot of 59 samples obtained from the control experiments. Colours indicate the different brain parts (BS: Brain Stem; CE: Cerebellum; OT: Optic Tectum; DI: Diencephalon; TE: Telencephalon; OB: Olfactory Bulbs), symbols denote the different species.

After the experiments (see 3.A.1, 3.A.2), focal female specimens were euthanized and the brains were dissected. Preservation of samples, and extraction of RNA was performed as in 3.B.1. Sequencing and analyses of the samples obtained before 2018 were sequenced in the same way. However, these do not include the last set of experiments of the presentation study (OV experiments, see 3.A.1). At this point, a novel sequencing platform was available at the genomics core that proved to be more cost effective. Hence, the new samples were run on this platform. For the sake of compatibility, all previous samples (save the olfactory bulbs) were resequenced. Hence, for 6 brain parts of 36 specimens (216 samples), transcriptomes were sequenced in a first batch (results presented here). For 29 specimens (used in the more recent presentation experiments), 5 brain parts were sequenced, together with 5 brain parts of 21 samples from the previous run (250 samples). The reads of the latter run have not yet been analysed at the time of writing.

3.B.4 Difficulties experienced with RNA sequencing

The olfactory bulbs proved to be problematic for both the control, as for the mating and presentation experiments. Often, these did not yield enough RNA to meet the minimum

requirements for sequencing. Additionally, even those that did meet these requirements often behaved differently in the analyses. We believe this is because we optimised the protocol for RNA extraction in Nile Tilapia specimens. These were significantly larger specimens and hence the brain parts were also larger. In the smaller *Ophthalmotilapia* specimens, olfactory bulbs were much smaller. Additionally, they were often difficult to distinguish from the telencephalon, to which they were attached. This might also explain the sometimes-high similarity seen between transcriptomes of olfactory bulbs and telencephalons. For this reason, we chose not to include the olfactory bulb tissues in the resequencing of the samples. (3.B.3)

Difficulties with the changing platform. Re-run of samples.

The statistical analyses on the sequencing read data proved to be difficult since we had an atypical design. Our aim to test between treatments was hindered by the fact that that individuals were fully nested within treatments (an individual can only be in one treatment, so we have to correct for the individual effect) and that we are interested in the interaction between treatments and brainparts. Lieven Clement and Koen Van Den Berghe (UGent) have developed a custom-made, two stage testing procedure to boosts the power of detecting gene-wise False Discovery Rate (FDR).

3.C Population genomics

3.C.1 Population genetics

The GENBAS project was inspired by the study of Nevado et al. (2011), who used mitochondrial Dloop sequences to identify patterns of hybridisation between *Ophthalmotilapia* species. However, Dloops were also shown to be informative of population structure in several Lake Tanganyika cichlids. Hence, we expanded the dataset of Nevado et al. (2011) by sequencing all available *Ophthalmotilapia* specimens available in the collections. We decided to also include the genus *Cyathopharynx* in this dataset, as previous studies showed that *Cyathopharynx* rendered *Ophthalmotilapia* paraphyletic (Koblmüller et al., 2004). In parallel, the entire collection of *Ophthalmotilapia* (and *Cyathopharynx*) was re-examined morphologically. For this, we especially tried to identify potential hybrid specimens.

We obtained a dataset of, in total, 515 sequences of *Ophthalmotilapia* and *Cyathopharynx* specimens belonging to 115 populations and originating from 75 locations. Of these, 328 were downloaded from Genbank whereas 187 new sequences were newly generated. The dataset also included all specimens used in the mating and presentation experiments. New samples were extracted with the Nucleospin Tissue kit (Macherey-Nagel, Düren, Germany). The mitochondrial control region (Dloop) was amplified following Nevado et al. (2009).

Biogeographic, genetic and demographic analyses were performed on the dataset of mitochondrial Dloops. We specifically compared classifications obtained by this mitochondrial marker with alternative classifications of the genus *Ophthalmotilapia* as proposed in aquarist literature. Dloops were further used to verify the origin of the specimens used in the experiments. Finally, the population structure shown by the DLoop dataset aided us in selecting samples for the GBS analysis.

3.C.2 Population genomics and Genotyping By Sequencing (GBS)

A total of 500 DNA extracts were sequenced using the genotype-by-sequencing (GBS) protocol for SNP discovery (Elshire et al. 2011). Following this protocol, DNA samples were digested with the PstI enzyme, ligated with adaptors allowing for the identification of specimens during demultiplexing and then sequenced on Illumina Hiseq. Library preparation and sequencing was carried out in three pools (Fig. 9). The resultant reads were quality checked using FastQC (version and citation), demultiplexed using GBSX.

The sample consisted of 60 *O. boops*, 119 *O. ventralis*, 166 *O. nasuta*, 51 *O. heterodonta*, 11 *O*. cf. *ventralis* (phenotypically intermediate between *O. ventralis* and *O. heterodonta*) and 4 lab-bread hybrids between *O. ventralis* and *O. nasuta*. Additionally, 24 samples of *Cyathopharynx* were included. The remaining specimens included representatives of the other genera of the Ectodini (the tribe of Lake Tanganyika cichlids to which *Ophthalmotilapia* belongs). Samples divided in five illumina libraries and run in three different Pools, containing 192, 189 and 121 samples, respectively (Fig. 10). The latter pool also contained samples that did not obtain enough reads in the first two pools.

Reads were demultiplexed, barcodes and adapters where automatically removed and overlapping reads were merged to fragments using FLASH. Bowtie2 was used in end-to-end mode to map the data to the *Oreochromis niloticus* genome version SM185804v2 using GBSXgenome. Variant calling was performed using FreeBayes, with mapping quality and base quality cut-offs of 15.

The resulting SNP dataset was filtered using VCFtools v.0.1.14 (Danecek et al. 2011). In view of the difficulties encountered with the structure of our data (see 3.C.3) constructed a pipeline for filtering that consisted of the following steps. We demanded: (1) minimum quality (QUAL) of 20 per SNP i.e. removal of SNPs with less than 99% base call accuracy, (2) removal of individuals with a frequency of missing data >90% (F_MISS >0.9), (3) removal of multiallelic SNPs keeping biallelic SNPs only (--max-alleles 2 --min-alleles 2), (4) removal of indels (--remove-indels), (5) removal of homozygotes for the reference allele using a minor allele frequency minimum (--MAF) of 0.01, (6) removal of homozygotes for the alternate allele using a minor allele frequency maximum (--max-MAF) of 0.99, (7) minimum reading depth of 3 (--minDP) to remove mapping and sequencing errors, (8) removal of SNPS with a frequency of missing data across all individuals (--max-missing) >90% and >70% (the dataset was split into two, stringent and less-stringent from this point, see 3.C.3), (9) removal of individuals with a frequency of missing data >70% after filtering (F_MISS >0.7), and (10) removal of SNPs which had over 0.5 observed heterozygosity per species. After filtering, the stringent 0.9 genotyped dataset consisted of 1584 SNPs and the less-stringent 0.7 genotyped dataset consisted of 10416 SNPs. For certain subsequent analyses, additional filtering by pruning was required. Pruning was performed using PLINK v1.90b5.4 (Chang et al. 2015); loci pairs with strong linkage disequilibrium (R2 > 0.1), within a sliding 50 SNP window at a step of 5 SNPs, were randomly removed of one SNP from the pair. After pruning, the stringent 0.9 genotyped dataset had 675 SNPs and the less-stringent 0.7 genotyped dataset had 3243 SNPs.

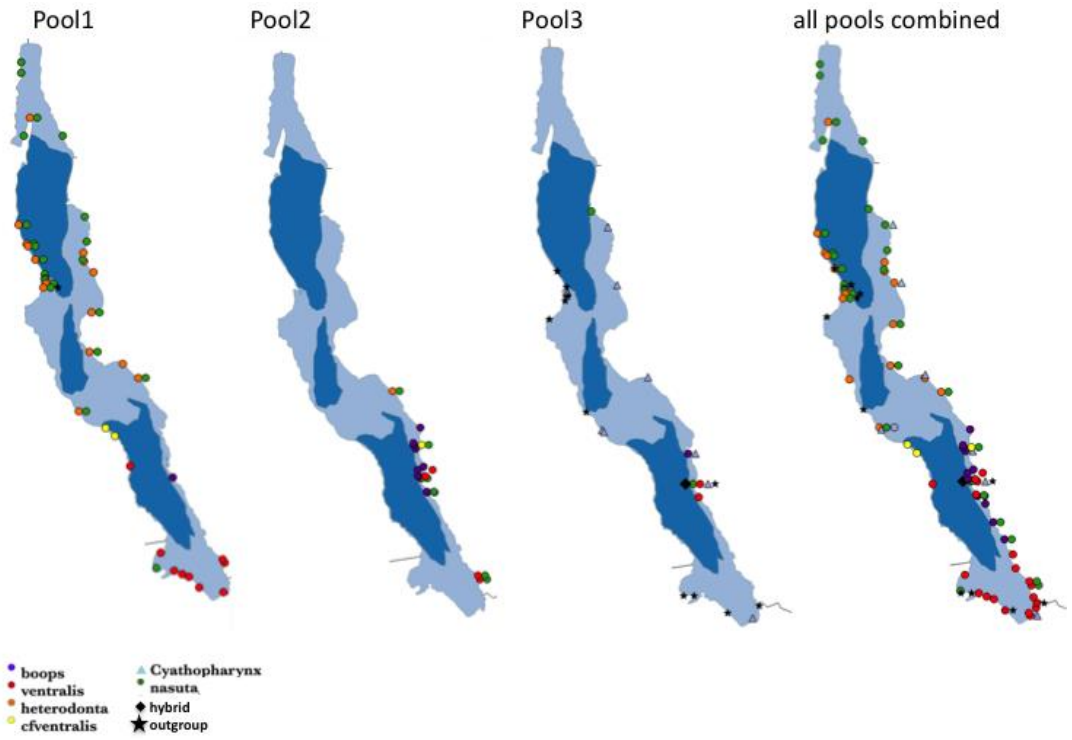

Fig. 10. Origin and identify of the samples used in the three different pools for GBS along the shores of Lake Tanganyika. Colours and symbols denote the different species. Outgroups (star) all belong to the tribus Ectodini. Hybrids were lab bred, but they were indicated as originating from their parents' catch locality on the map.

We specifically investigated the dataset for traces of hybridisation. For this, we used, amongst others, ADMIXTURE, ABBA BABA, and Bayes factor delimitation.

3.C.3 Difficulties experienced with population genomics

Technical issues were explored at the onset of the project. We discovered that, when using the standard Nucleospin extraction kit, RNA extracts showed a smear when being loaded on a gel, indicating partial degradation. This especially held for older (1992, 1995) samples. To test whether this was due to the extraction protocol, we extracted five tissue samples with the Nucleospin protocol and again with the Paulino protocol and load them in a testrun. As the latter protocol does not use a spin filter, and produces a clear intact band of high molecular weight DNA, we expected to obtain more reads from the tissues extracted with the Paulino protocol. Three of the five samples yielded comparable read numbers between both extraction protocols. For the two other samples, the nucleospin protocol clearly outperformed the Paulino protocol in terms of read numbers. Hence, we decided that this potential degradation was not an issue and continued to use the Nucleospin extraction kits. We further also discussed which cutting enzyme to use for the genotyping by sequencing (GBS) approach: ApeKI or PstI. We also performed a test run, which revealed a good coverage and good sequencing depth for PstI.

At the onset of the project, mapping of GBS and RNA sequencing data was performed to the genome of *Metriaclina zebra*. This was reference was chosen as it was the most closely related species for which a reference genome was available. However, in the course of the project we decided to use the genome of *Oreochromis niloticus* instead (SM185804v2). Although this is a more distantly related species, we choose this approach since it is a more detailed annotation and since the annotation of the available *M. zebra* genome was based on that of the *O. niloticus* genome anyway.

Although data from the first pool(s) could be analysed successfully, we run into some issues when analysing the data from all three pools jointly. When using the same methodology and the same pipelines that successfully worked on the data from the first and the second pools, the dataset contained a lot of missing data, regardless of the filtering and pruning parameters used. At first, we hypothesised that the reason why the combined dataset was difficult to analyse was that pool 3 contained less samples (121) than pool 1 and pool 2 (192). Hence, pool 3 contained 30% more data for each sample. However, this could not explain that, even with very relaxed filtering, very few shared SNPs could be obtained. Eventually, we found out that, for the first pool, a different size selection was performed when preparing the libraries. Taking this into account, we could design a custom-made pipeline for filtering that still gave us sufficient SNPs for the analyses. However, for some of the analyses, this difference between the pools forced us to use two different datasets, one obtained with a more stringent, and one with a less stringent approach (see 3.C.2).

4. SCIENTIFIC RESULTS AND RECOMMENDATIONS

4.A Pairing experiments

4.A.1 presentation experiments

In these experiments, we predicted a difference in species recognition between females of *O. ventralis* and *O. nasuta*. Specifically, we hypothesize that *O. nasuta* females would discriminate between con- and heterospecific males at the initial stages of an encounter. For *O. ventralis* females, however, we expect that this discriminatory behaviour would be less pronounced or absent. In order to test this, we compared the behaviour of both focal and non-focal specimens in five different comparisons.

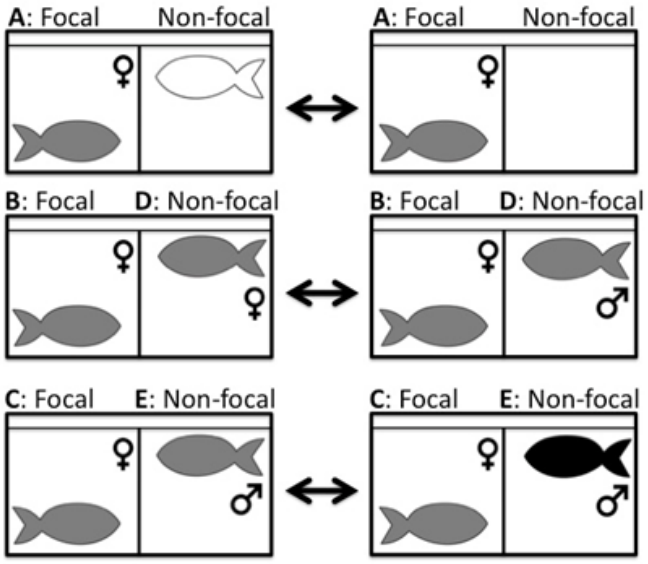

Fig. 11. Contrasts tested using PERMANOVA for the presentation experiments: With A) comparing the behaviour of focal females that were presented with another fish vs. with nothing (Co vs. CF, CM, HM), B) with a conspecific female vs. a conspecific male (CF vs. CM), C) and with a conspecific vs. a heterospecific male (CM vs. HM). We further compared D: the behaviour of non-focal conspecific males vs. females (NCF vs. NCM), and E: con- vs. heterospecific males (NCM vs. NHM). Black and grey fishes represent belonging to a different species

In order to reduce the number of comparisons, we restricted us to only biologically relevant contrasts. For focal specimens, we compared the behaviour between (a) females that were presented with another fish vs. with no fish, (b) focal females that were presented with a conspecific female vs. a conspecific male, and (c) females that were presented with a con- versus a heterospecific male. Two additional comparisons were tested for the non-focal individuals. We tested (d) whether conspecific females and males respond differently to a focal female and (e) whether con- and heterospecific males respond differently to a focal female. These tests were performed using PERMANOVA on the combined data of tracking parameters and point events collected 15 minutes before, and 15 minutes after the removal of the visual barrier (Fig. 11, Table IV).

For both the OV and the ON experiments, we found a significant difference in the behaviour of focal females that were not presented to another fish (controls) and focal females that were presented with another fish (comparison A). Additional Mann-Whitney U tests revealed that controls differed from other focal females by their higher values for the variable Dist.wall and their lower values for ‘ram’. Unexpectedly, in both sets of experiments, we did not observe a difference in behaviour between focal females that were presented with a conspecific female or a conspecific male

(comparison B). However, when comparing the behaviour of focal females presented with con- and heterospecific males (comparison C), a difference became evident between the ON and the OV experiments. In support of our hypothesis, females of *O. nasuta* responded differently towards con- and heterospecific males, whereas females of *O. ventralis* did not (Table V). Mann-Whitney U tests revealed that this was due to the lower number of observed 'ram' in *O. nasuta* females that were presented to *O. nasuta* males compared to those presented to *O. ventralis* males. Additionally, *O. nasuta* females never performed a 'flee' behaviour when being presented with an *O. ventralis* male, whereas this was observed in all but one of the *O. nasuta* females presented to a *O. nasuta* male. No difference was found in all tree comparisons performed on the data collected before the separation wall was removed, indicating that the differences found above were due to the presentation to another specimen. We also compared the behavioural responses of conspecific females and males (D), and con- and heterospecific males (E) presented to a focal female. Unexpectedly, permanova only revealed a difference in the behaviour of *O. nasuta* and *O. ventralis* males in the OV experiment, prior to the removal of the barrier. This was due to the higher average vertical swimming speed of *O. ventralis* males compared to *O. nasuta* males (SpY 1.7+-0.5 vs. 0.7+-0.4, p= 0.014, g=2.05).

Table V. PERMANOVA performed on the behavioral parameters of the ON (*O. nasuta*) and OV (*O. ventralis*) experiments. Tests were performed on the data collected during 15 minutes before (B), 15 minutes after (A1) the removal of the opaque wall. Comparisons as in Fig.12. Values in bold are significant at the 0.05 level.

	C	A	B	C	D	E
ON	B	F:0.31; p: 0.789	F:0.95; p: 0.423	F:0.91; p: 0.474	F:1.17; p: 0.339	F:1.15; p: 0.24
	A1	**F:7.28; p: 0.012**	F:0.42; p: 0.768	**F:11.33; p: 0.016**	F:0.45; p: 0.651	F:0.55; p: 0.73
OV	B	F:0.23; p: 0.762	F:1.62; p: 0.203	F:0.85; p: 0.387	F:3.19; p: 0.061	**F:5.13; p: 0.013**
	A1	**F:8.00; p: 0.005**	F:0.91; p: 0.386	F:0.40; p: 0.546	F:0.28; p: 0.633	F:0.48; p: 0.524

We also explored the datasets visually using CVA (Fig. 12). In the CVA's conducted on the behavioural data collected before the removal of the barrier, values of all females overlapped, suggesting a highly similar behaviour. However, values for males were separated from those of the females and from each other. This was mostly due to the more active swimming behaviour (Sp, SpX, SpY) for *O. ventralis* males (ON experiments) and a higher number of point events (Ram, Sand, Bite) performed at the floor of the aquarium (Height) for the *O. nasuta* males (OV experiments).

The CVA's conducted on the data collected 15 minutes after the removal of the barrier foremost showed the distinction of the focal females that were used as controls. These stood out by their high values for CV1, which could be explained by their higher values for Dist.Wall. In the ON experiments, (heterospecific) *O. ventralis* males stood out by their low values for CV1, which would be attributed to their more active swimming behaviour (Sp, SpX, SpY). Conspecific *O. nasuta* males stood out by their low values for CV2, which could be due to the higher occurrence of 'Sand' and 'Bite' behaviour. Values for females that were presented with another fish had more intermediate values for CV1 and CV2. However, values of *O. nasuta* females that were presented to a conspecific male clustered between values of those males and of those of the other females. Similarly, females that were presented to a heterospecific male had values that were intermediate between those of these males and those of the other females (Fig. 12). This suggest that, although females behaved relatively similar across treatments, their behaviour somewhat resembled that of the non-focal fish presented to them. In the CVA of the OV experiments, (heterospecific) *O. nasuta* males stood out by their low

values for CV1 and high values for CV2. This was most influenced by the higher occurrence of a 'Ram' behaviour. Values of *O. ventralis* males overlapped with those of female specimens that were presented with another fish.

Based on the results of Nevado et al. 2011, we hypothesised that *O. nasuta* females would discriminate stronger against heterospecific males, than females of *O. ventralis*. This hypothesis is supported by the outcome of our experiments as we found evidence for species recognition in females of *O. nasuta*, but not in those of *O. ventralis*. This suggests that a female's initial response towards a potential mate already predicts to a certain degree the outcome of the mating process that may follow. Unexpectedly, females of both species behaved similarly towards conspecific males and females. This could imply that we observed the routine behaviour of a (isolated) female that encounters a conspecific individual, rather than reproductive behaviour. In the wild, non-breeding females of both species aggregate in large feeding groups. Hence, being isolated for 12 hours, as was the case prior to the start of the experiment, represents an unnatural situation and it would not be unlikely if *Ophthalmotilapia* females are behaviourally hardwired to reunite immediately with conspecifics, regardless of whether these are female or male.

4.C Differential gene expression

4.B.1 Gene expression in the female brain during the pairing experiments

RNA extracts obtained from all six brain parts were sequenced for the eleven specimens for which we recorded mating. These include four *O. nasuta* females that mated with an *O. nasuta* male, four *O. ventralis* females that mated with an *O. ventralis* male and three *O. nasuta* females that mated with an *O. nasuta* male. Differential gene expression was compared between these 66 samples (one sample was removed because of low quality). For these samples, reads for 11,233 genes were retrieved, 10,138 of which were present in all samples and 0 were unique to one sample (Fig. 13). We visually explored the data, together with that of the control experiment with a MDS plot (Fig. 14). This, again, revealed that the bulk of the variation is explained by the differences between brain parts. Additionally, transcriptomes for the olfactory bulb were much less clustered than those from the other brain parts, again revealing the uncertainties we had with RNA extracts from this part of the brain.

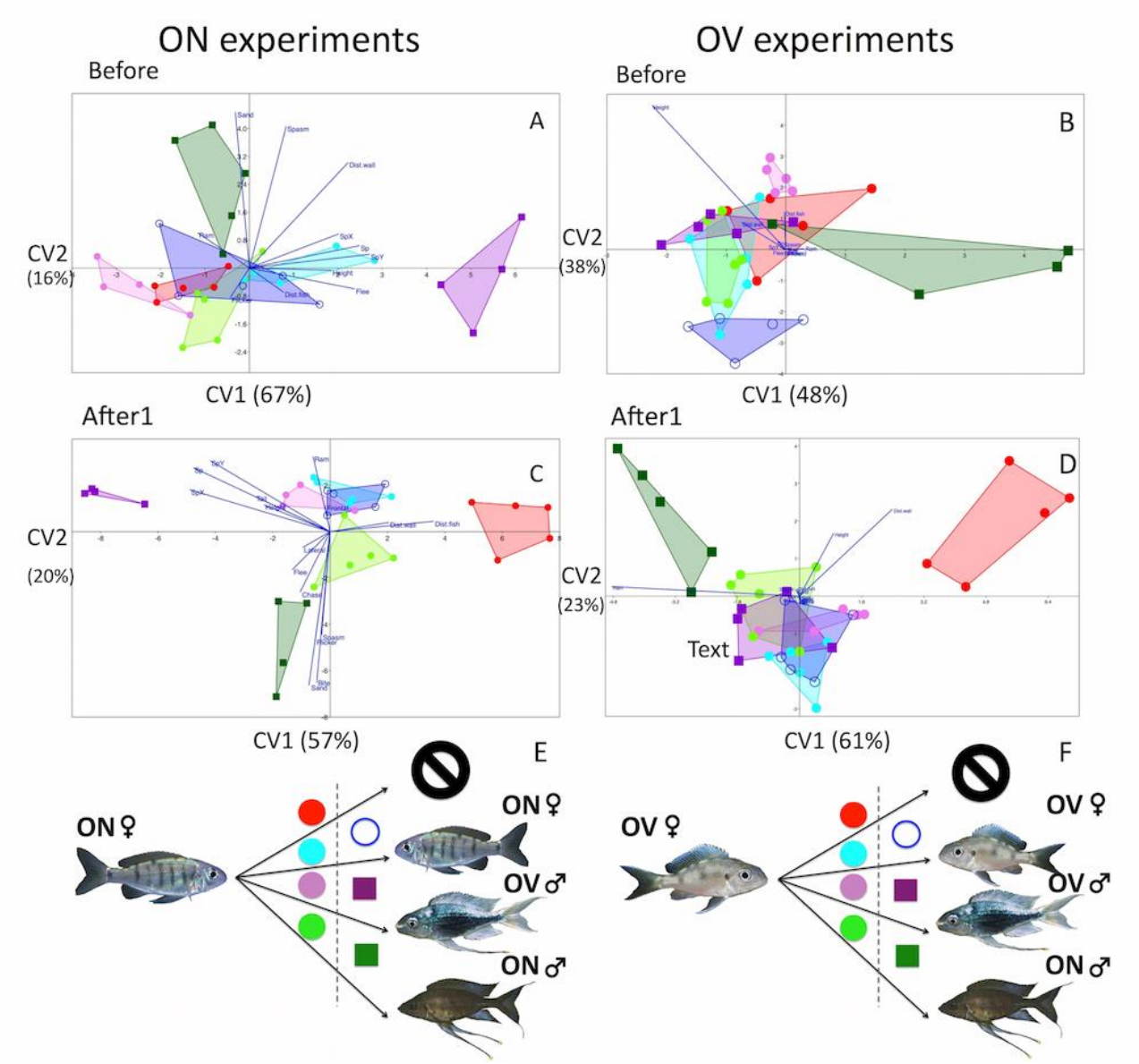

Fig. 12. Canonical variate analyses on the behavioural data collected 15 min before and 15 min after the visual barrier was removed in the ON (left) and the OV (right) experiments. Symbols on the scatter plots for the ON and OV experiments as in E and F, respectively, with full circles denoting focal females presented with no fish (red), a conspecific female (blue) an *O. ventralis* male (purple), and an *O. nasuta* male (green), empty circles denote non-focal conspecific females and full squares *O. ventralis (*purple) and *O. nasuta* (green) males. Explained variances are added to the axes. With *O. nasuta*: ON and *O. ventralis*: OV.

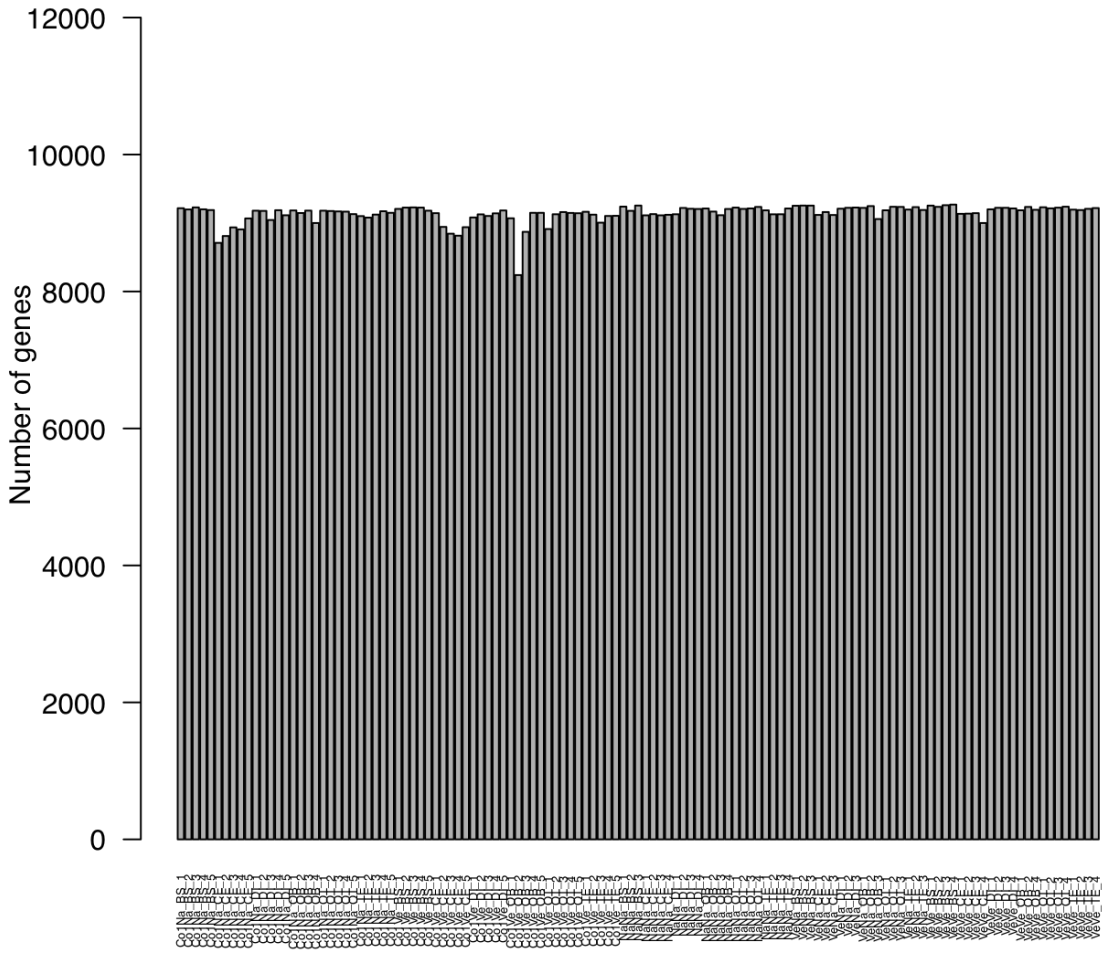

Fig. 13. Number of reads obtained from the retained 65 samples of the mating experiments, combined with samples from the control setting. Sample codes as treatment_brainpart_specimennumber, with Co1Na: *O. nasuta* female from the control experiments, Co1Ve: *O. ventralis* female from the control experiments; NaNa: *O. nasuta* female that mated with a *O. nasuta* male, VeNa: *O. ventralis* female that mated with a *O. nasuta* male; VeVe: *O. ventralis* female that mated with a *O. ventralis* male; BS: Brain Stem; CE: Cerebellum; OT: Optic Tectum; DI: Diencephalon; TE: Telencephalon; OB: Olfactory Bulbs

For the mating experiments, we examined differential gene expression in two different comparisons: A) between *O. nasuta* and *O. ventralis* females after mating with a conspecific male; B) between *O. ventralis* females that mated with an *O. nasuta vs.* an *O. ventralis* male (Fig. 15). In all comparisons, a sample from one individual for which the four brain parts were available was used as baseline to define up- and down regulation. We found that, out of the 10,138 shared genes, 2,525 were differentially expressed in at least one of the brain parts for comparison A.

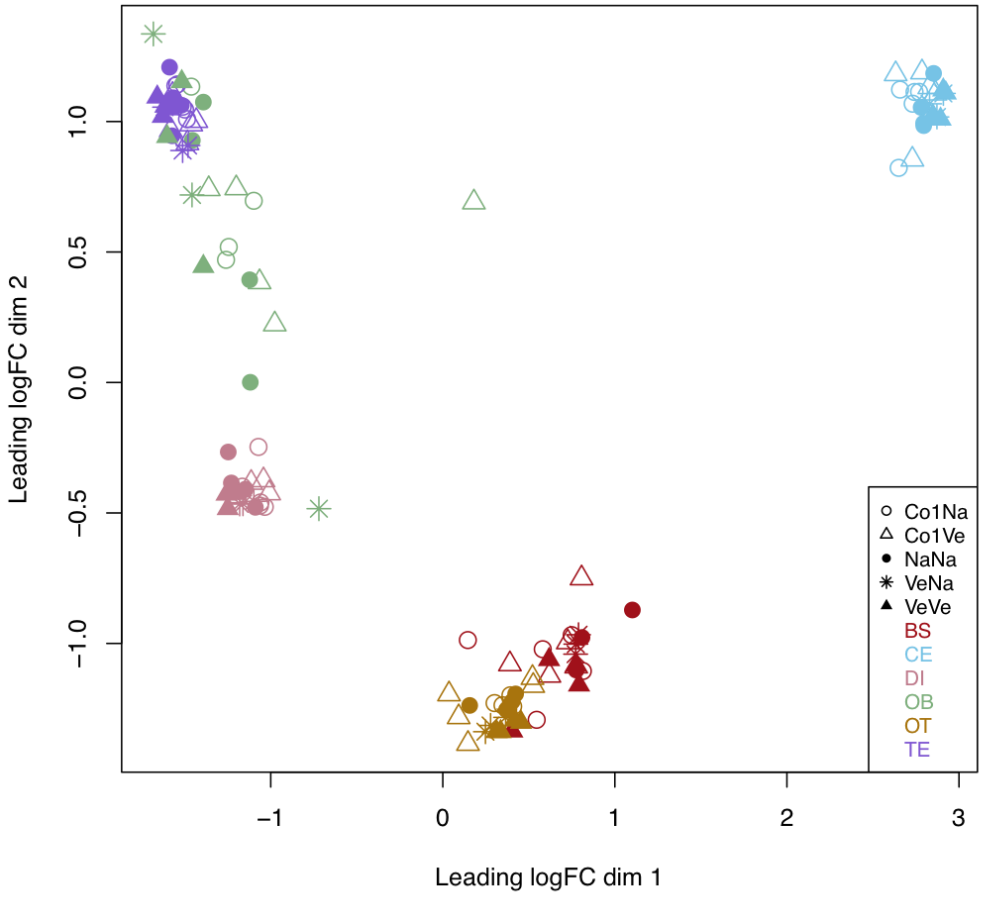

Fig. 14. MDS plot of 65 samples obtained from the mating experiments, . Colours indicate the different brain parts (BS: Brain Stem; CE: Cerebellum; OT: Optic Tectum; DI: Diencephalon; TE: Telencephalon; OB: Olfactory Bulbs), symbols denote the different treatments.

Although these comparisons stem from different species, this is surprising because in the absence of mating, transcriptomes of *O. ventralis* and *O. nasuta* females were almost identical (see 3.B.2). This implies that a large part of this difference should be explained by the mating behaviour. However, the mating repertoires of the two species were shown to be very similar. Differentially expressed genes were found in all brain parts although the largest numbers were found in the cerebellum and the diencephalon. These genes could be attributed to a large array of biological functions.

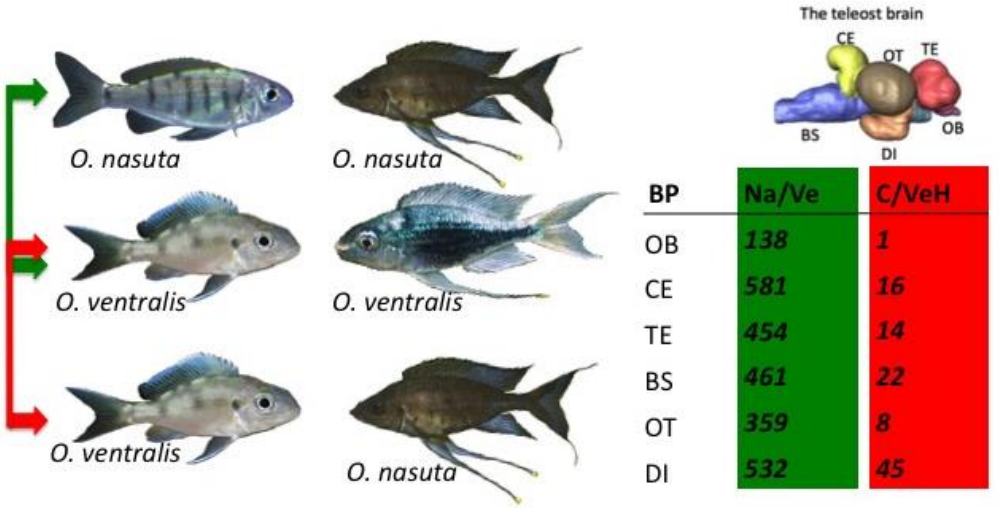

Fig. 15. Analysis of differential gene expression in the mating experiments. Right: Two different comparisons were made: between female specimens of *O. nasuta* that mated with an *O. nasuta* (N=4) male *vs.* female specimens of *O. ventralis* that mated with an *O. ventralis* male (N=4) (A: Na/Ve) and between female specimens of *O. ventralis* that mated with an *O. ventralis* male (N=4) and female specimens of *O. ventralis* that mated with an *O. nasuta* male (N=3). Left: The number of genes that were significantly differentially expressed in at least one of the brain parts for each of these contrasts with BS: Brain Stem; CE: Cerebellum; OT: Optic Tectum; DI: Diencephalon; TE: Telencephalon; OB: Olfactory Bulbs.

When comparing *ventralis* females that mated with an *O. nasuta vs.* an *O. ventralis* male, 106 genes were found to be differentially expressed. Here, the diencephalon proved to be the region were most differential gene expression was found. The two most differentially expressed genes were also found in this brain part and the gene with the highest observed upregulation has a known function in hormone production (pomc, Table VI).

4.B.2. Gene expression in the female brain during the presentation experiments

For the presentation experiments, RNA sequencing data of all of the OV experiments and some of the ON experiments has not yet been analysed (see 3.A.2). Hence, we will here only present the data for the 15 *O. ventralis* females used in the experiments performed prior to 2018 (Table III). After filtering, four samples were removed: two that yielded a low read count, and two olfactory bulb samples that clustered wrongly in the analyses (most likely because of errors in the dissections see 3.B.4). After removing these low quality samples, 86 samples (brain parts per individual) were retained for further analysis. For these samples, reads for 12,207 genes were retrieved, of which 9,745 were present in all samples and none were unique to one sample (Fig. 16).

Using the same statistical analysis as in 3.B.2, we tested differences in gene expression in three different contrasts (Fig. 17). These are the same contrast those tested on the behavioural parameters for the focal females (Fig. 11). In total, 96 genes were found to be differentially expressed in contrast A, 3 in contrast B and 14 in contrast C. This strongly agrees with the behaviour results since a strong difference was found in the behaviour of the focal females for comparison A (/-fish). No significant difference in behaviour was found for comparison B (F-M). A significant difference in behaviour was, however, encountered between female *O. nasuta* depending on whether they were presented with a con- and a heterospecific (*O. ventralis*) male (See 4.A.2). When

examining the function of the genes that were found to be differentially expressed, several had known functions in hormonal signalling or in behavioural processes. (Table VII)

Table VI. Results for the differential gene expression analysis between females of *O. ventralis* after mating with an *O. ventralis* and an *O. nasuta* male. The 20 genes (out of 106) with the highest degree of differential gene expression (measured ad logFC) are shown, logFC: log fold chage with a positive value representing higher expression in heterospecific compared to conspecific matings, FDR_AdjP: P-value adjusted for false discoveries, BP: brain part, GeneID, and GO_id: gene ontogeny (*O. niloticus* genome).

Gene	logFC	FDR_AdjP	BP	go_id	function
cga	7.69	0.0352	DI	NA	NA
pomc	7.34	0.0293	DI	GO:0010469	regulation of receptor activity
pomc	7.34	0.0293	DI	GO:0005179	hormone activity
pomc	7.34	0.0293	DI	GO:0005576	extracellular region
LOC102078435	5.95	0.0004	OT	NA	NA
LOC100704415	5.60	0.0245	DI		
LOC100695189	5.16	0.0032	DI	NA	NA
LOC100693802	4.69	0.0049	CE	GO:0005525	GTP binding
LOC102078435	4.65	0.0007	BS	NA	NA
LOC100689935	4.58	0.0055	DI	NA	NA
LOC102080970	4.57	0.0141	DI	GO:0016567	protein ubiquitination
LOC102080970	4.57	0.0141	DI	GO:0004842	ubiquitin-protein transferase activity
LOC102080970	4.57	0.0141	DI	GO:0016740	transferase activity
LOC102077096	4.54	0.0179	CE	NA	NA
LOC102078435	4.44	0.0124	CE	NA	NA
LOC102079967	4.37	0.0462	DI	NA	NA
LOC102078435	4.34	0.0000	TE	NA	NA
LOC102078435	4.32	0.0001	OB	NA	NA
LOC100697587	4.20	0.0055	DI	GO:0016567	protein ubiquitination
LOC100697587	4.20	0.0055	DI	GO:0016740	transferase activity
LOC100697587	4.20	0.0055	DI	GO:0004842	ubiquitin-protein transferase activity
LOC100693802	4.20	0.0269	DI	GO:0005525	GTP binding
LOC102078066	4.08	0.0062	DI		
LOC102079967	4.00	0.0000	CE	NA	NA
cmpk2	3.99	0.0245	DI		
LOC100689935	3.76	0.0073	TE	NA	NA
rsad2	3.58	0.0059	DI	GO:0009615	response to virus
rsad2	3.58	0.0059	DI	GO:0003824	catalytic activity
rsad2	3.58	0.0059	DI	GO:0051536	iron-sulfur cluster binding
rsad2	3.58	0.0059	DI	GO:0005789	endoplasmic reticulum membrane
rsad2	3.58	0.0059	DI	GO:0005811	lipid droplet
rsad2	3.58	0.0059	DI	GO:0051607	defense response to virus
pax7	3.53	0.0141	DI	GO:0042060	wound healing
pax7	3.53	0.0141	DI	GO:0048066	developmental pigmentation
pax7	3.53	0.0141	DI	GO:0005634	nucleus
pax7	3.53	0.0141	DI	GO:0006355	regulation of transcription. DNA-templated
pax7	3.53	0.0141	DI	GO:0003677	DNA binding
pax7	3.53	0.0141	DI	GO:0043565	sequence-specific DNA binding
pax7	3.53	0.0141	DI	GO:0050938	regulation of xanthophore differentiation
pax7	3.53	0.0141	DI	GO:0006351	transcription. DNA-templated
pax7	3.53	0.0141	DI	GO:0007275	multicellular organism development

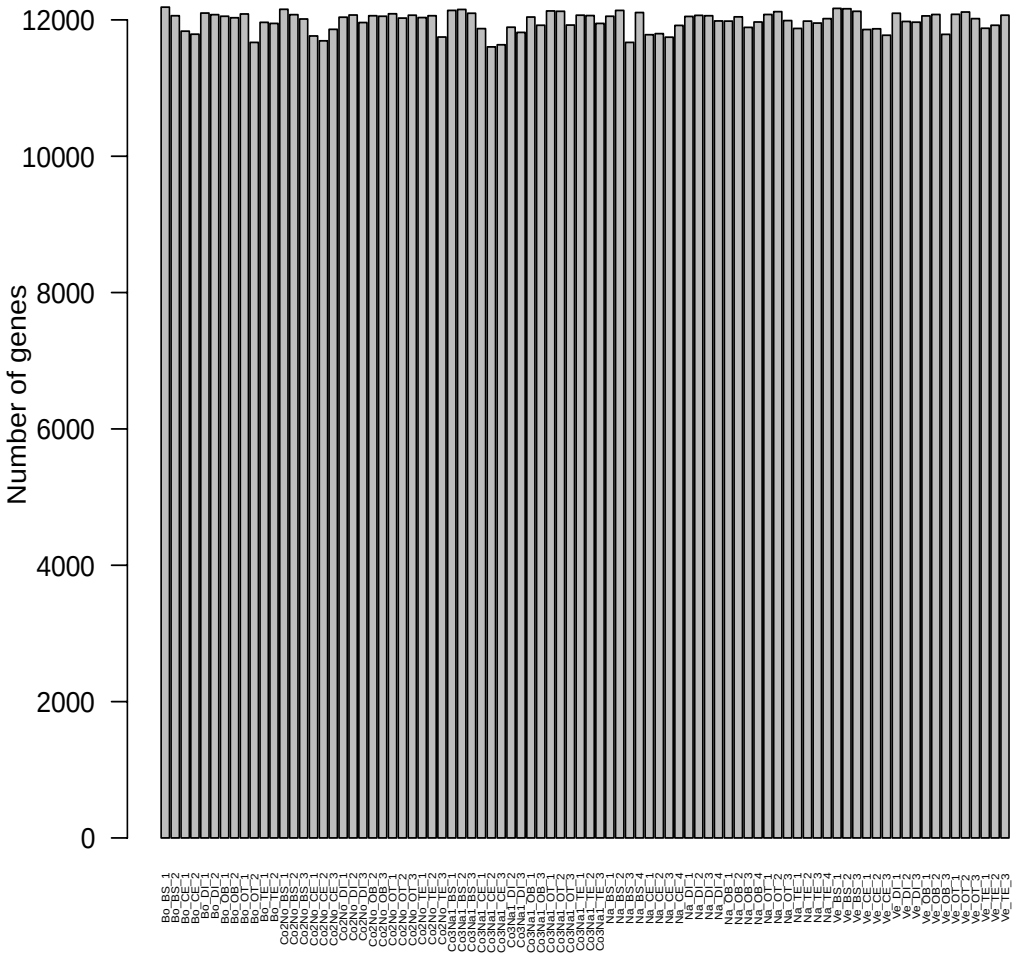

Fig. 16. Number of reads obtained from the retained 86 samples of the presentation experiments performed with 15 focal *O. ventralis* specimens. Sample codes as: treatment_brainpart_specimennumber, with Bo: presented to a male *O. boops*, Na: presented to a male *O. nasuta,* Ve: presented to a male *O. ventralis,* Co3No: presented to a *O. ventralis* female, Co2No: presented to no fish; BS: Brain Stem; CE: Cerebellum; OT: Optic Tectum; DI: Diencephalon; TE: Telencephalon; OB: Olfactory Bulbs

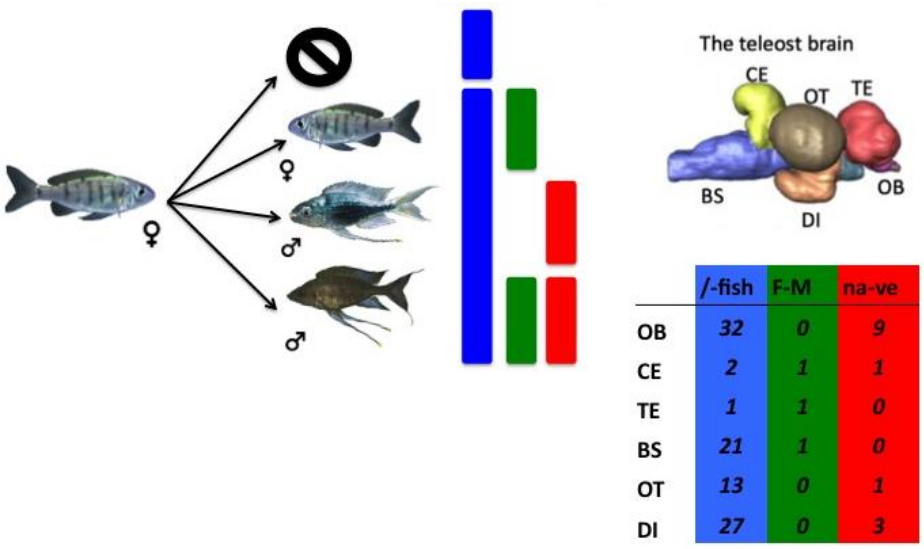

Fig. 17. Analysis of differential gene expression in the presentation experiments. Right: focal female specimens of *O. nasuta* were presented to nothing, a conspecic female, a heterospecific male or a conspecific male. Differential gene expression was tested between three contrasts, as shown in Fig. 11 A: fish *vs.* no fish (/-fish; blue), B: conspecific female *vs.* male (F-M, green) and C: con- *vs.* heterospecific male (na-ve; red). Left: The number of genes that were significantly differentially expressed in at least one of the brain parts for each of these contrasts with BS: Brain Stem; CE: Cerebellum; OT: Optic Tectum; DI: Diencephalon; TE: Telencephalon; OB: Olfactory Bulbs.

Table VII. Results for the differential gene expression analysis between females of *O. nasuta* being presented with an *O. ventralis* and an *O. nasuta* male. With logFC: log fold chage with a positive value representing higher expression in heterospecific compared to conspecific encounter, FDR_AdjP: P-value adjusted for false discoveries, BP: brain part, GeneID, and GO_id: gene ontogeny (*O. niloticus* genome).

Gene	logFC	FDR_AdjP	BP	go_id	function
ansn	5.13	0.0038	OT	GO:0016805	dipeptidase activity
ansn	5.13	0.0038	OT	GO:0008237	metallopeptidase activity
ansn	5.13	0.0038	OT	GO:0006508	proteolysis
ansn	5.13	0.0038	OT	GO:0016787	hydrolase activity
ansn	5.13	0.0038	OT	GO:0008152	metabolic process
LOC100690881	3.63	0.0393	CE	NA	NA
ajuba	-3.56	0.0043	OB	GO:0007507	heart development
ajuba	-3.56	0.0043	OB	GO:0046872	metal ion binding
crhbp	3.38	0.0058	OB	GO:0005576	extracellular region
crhbp	3.38	0.0058	OB	GO:0051424	corticotropin-releasing hormone binding
gnrh1	5.29	0.0035	OB	GO:0007275	multicellular organism development
gnrh1	5.29	0.0035	OB	GO:0005179	hormone activity
gnrh1	5.29	0.0035	OB	GO:0005183	gonadotropin hormone-releasing hormone activity
gnrh1	5.29	0.0035	OB	GO:0005576	extracellular region
lhx8	3.66	0.0121	OB	GO:0046872	metal ion binding
lhx8	3.66	0.0121	OB	GO:0005634	nucleus
lhx8	3.66	0.0121	OB	GO:0006355	regulation of transcription, DNA-templated
lhx8	3.66	0.0121	OB	GO:0043565	sequence-specific DNA binding
lhx8	3.66	0.0121	OB	GO:0003677	DNA binding
LOC100701631	2.05	0.0443	OB	GO:0005244	voltage-gated ion channel activity
LOC100701631	2.05	0.0443	OB	GO:0006810	transport
LOC100701631	2.05	0.0443	OB	GO:0008076	voltage-gated potassium channel complex
LOC100701631	2.05	0.0443	OB	GO:0016021	integral component of membrane
LOC100701631	2.05	0.0443	OB	GO:0071805	potassium ion transmembrane transport
LOC100701631	2.05	0.0443	OB	GO:0006813	potassium ion transport
LOC100701631	2.05	0.0443	OB	GO:0005267	potassium channel activity
LOC100701631	2.05	0.0443	OB	GO:0005249	voltage-gated potassium channel activity
LOC100701631	2.05	0.0443	OB	GO:0051260	protein homooligomerization
LOC100701631	2.05	0.0443	OB	GO:0005216	ion channel activity
LOC100701631	2.05	0.0443	OB	GO:0006811	ion transport
LOC100701631	2.05	0.0443	OB	GO:0055085	transmembrane transport
LOC100701631	2.05	0.0443	OB	GO:0016020	membrane

LOC100701631	2.05	0.0443	OB	GO:0034765	regulation of ion transmembrane transport
LOC100704176	3.10	0.0087	OB	GO:0051424	corticotropin-releasing hormone binding
LOC100704176	3.10	0.0087	OB	GO:0005576	extracellular region
LOC100705548	-1.83	0.0152	OB		
LOC100708195	-2.99	0.0043	OB	GO:0050795	regulation of behavior
LOC100708195	-2.99	0.0043	OB	GO:0040012	regulation of locomotion
LOC100708195	-2.99	0.0043	OB	GO:0006939	smooth muscle contraction
LOC100708195	-2.99	0.0043	OB	GO:0005887	integral component of plasma membrane
LOC100708195	-2.99	0.0043	OB	GO:0005886	plasma membrane
LOC100708195	-2.99	0.0043	OB	GO:0007186	G-protein coupled receptor signaling pathway
LOC100708195	-2.99	0.0043	OB	GO:0098664	G-protein coupled serotonin receptor signaling pathway
LOC100708195	-2.99	0.0043	OB	GO:0007165	signal transduction
LOC100708195	-2.99	0.0043	OB	GO:0004930	G-protein coupled receptor activity
LOC100708195	-2.99	0.0043	OB	GO:0004871	signal transducer activity
LOC100708195	-2.99	0.0043	OB	GO:0016021	integral component of membrane
LOC100708195	-2.99	0.0043	OB	GO:0016020	membrane
LOC100708195	-2.99	0.0043	OB	GO:0007268	chemical synaptic transmission
LOC100708195	-2.99	0.0043	OB	GO:0004993	G-protein coupled serotonin receptor activity
LOC100708195	-2.99	0.0043	OB	GO:0042310	vasoconstriction
rorb	2.94	0.0085	OB	GO:0005634	nucleus
rorb	2.94	0.0085	OB	GO:0006355	regulation of transcription, DNA-templated
rorb	2.94	0.0085	OB	GO:0006351	transcription, DNA-templated
rorb	2.94	0.0085	OB	GO:0003677	DNA binding
rorb	2.94	0.0085	OB	GO:0046872	metal ion binding
rorb	2.94	0.0085	OB	GO:0043565	sequence-specific DNA binding
rorb	2.94	0.0085	OB	GO:0003700	transcription factor activity, sequence-specific DNA binding
rorb	2.94	0.0085	OB	GO:0004879	RNA polymerase II transcription factor activity, ligand-activated sequence-specific DNA binding
rorb	2.94	0.0085	OB	GO:0043401	steroid hormone mediated signaling pathway
rorb	2.94	0.0085	OB	GO:0003707	steroid hormone receptor activity
rorb	2.94	0.0085	OB	GO:0030522	intracellular receptor signaling pathway
rorb	2.94	0.0085	OB	GO:0008270	zinc ion binding
crhbp	3.42	0.0005	DI	GO:0051424	corticotropin-releasing hormone binding
crhbp	3.42	0.0005	DI	GO:0005576	extracellular region
gal	1.91	0.0003	DI	GO:0005576	extracellular region
gal	1.91	0.0003	DI	GO:0005179	hormone activity
sox1	2.51	0.0013	DI	GO:0006355	regulation of transcription, DNA-templated
sox1	2.51	0.0013	DI	GO:0003677	DNA binding
sox1	2.51	0.0013	DI	GO:0007399	nervous system development
sox1	2.51	0.0013	DI	GO:0005634	nucleus

4.C Population genomics

4.C.1 Population genetics

All specimens for which tissue samples were available were identified morphologically. Hence, we expanded the dataset of Nevado et al. (2011) by sequencing all *Ophthalmotilapia* specimens available in the collections. Novel specimens from the Western side of the Lake, less covered in previous studies, were collected in 2010 and in 2014. As preliminary analyses revealed a large amount of haplotype diversity in *O. nasuta*, this species was subdivided in seven groups, based on distribution ranges and on classifications presented in the aquarist literature. In order to investigate differences in the degrees of gene flow*, O. nasuta* originating from the zone where three species of *Ophthalmotilapia* occur in sympatry were also assigned to a different group. Hence, *O. nasuta* specimens from the western, the southern and the south-eastern shores that did or did not co-occur with *O. boops* were classified as *O. nasuta* 'All' (allopatric) and *O. nasuta* 'Sym' (sympatric), respectively.

Out of the 515 haplotypes in the dataset, 161 were unique. In the TCS haplotype network, seven clusters could be identified, which will be referred to, in what follows, as clusters A to G (Fig. 17). Cluster A contained most of the specimens belonging to the *O. ventralis*/*O. heterodonta* complex,

together with two *O. boops* sequences and eight sequences of *O. nasuta* from the south of the Lake. The two *O. nasuta* specimens from Nyanza Lac were also found within this cluster as they carried a haplotype that they shared with *O. heterodonta*. Cluster A could also be divided into several subclusters. Haplotypes belonging to these subclusters often were often restricted to a well-defined part of the Lake's shoreline (Fig. 18, top). Cluster B contained all *O. nasuta* 'Malagarasi' specimens, together with two sequences of *O. nasuta* 'paranasuta'. Cluster C consisted of haplotypes assigned to *O. nasuta* from the western, the southern and the south-eastern shores of Lake Tanganyika. One *O. heterodonta* haplotype was also found in this cluster. Cluster D contained all but two of the *O. boops* sequences together with 30 sequences of *O. nasuta*. All of these belonged to specimens from locations where *O. nasuta* occurs in sympatry with *O. boops*. Two *O. ventralis* sequences were also found in this cluster. Cluster E is made up entirely of all the *O. nasuta* 'Mahale' sequences, whereas cluster F contains all but two of the *O. nasuta* 'paranasuta' haplotypes. Cluster G, finally, contains all sequences of *Cyathopharynx*.

Our results showed a clear difference in the mitochondrial diversity present in the different species of *Ophthalmotilapia*. Especially *O. nasuta* was highly diverse as its haplotypes could be found in all *Ophthalmotilapia* clusters. Hence, at least in terms of MtDNA, its diversity equalled that of the entire genus. In some cases, this could be attributed to unidirectional hybridisation (Nevado et al., 2011). At other parts of the shoreline, *O. nasuta* populations carried isolated haplotypes. This pointed towards geographical differentiation or even undescribed taxonomical diversity.

The *O. ventralis/O. heterodonta* complex also contained geographical structure, but on a smaller scale as most specimens carried haplotypes that grouped within cluster A. Whether *O. ventralis* and *O. heterodonta* are separate species remains debated. However, even though no one-to-one correspondence between the valid species and the MtDNA haplotypes was found, our data did supported the distinction of the southern and the northern populations. This somewhat agrees with the distribution patterns of the nominal species *O. ventralis* and *O. heterodonta*, supporting their validity.

The genetic structuring of the two wide spread lineages in *Ophthalmotilapia*: *O. nasuta* and the *O. ventralis*/*O. heterodonta* sister species was also calculated using a genetic landscape shape analysis, implemented in alleles in space (Miller 2005) (Fig 19). For *O. nasuta*, this revealed an large degree of structuring in the lake, although some homogenous regions were found along the western and, especially, the south-western part of the lake (note that the western edge is visible on Fig. 18D). For the species belonging to the *O. ventralis*/*O. heterodonta* complex homogeneity was overall large although clear boundaries were found in the south-east and central-west. These agree with the proposed boundaries between *O. ventralis* (south) and *O. heterodonta* (north), giving further support for the distinction of the two species.

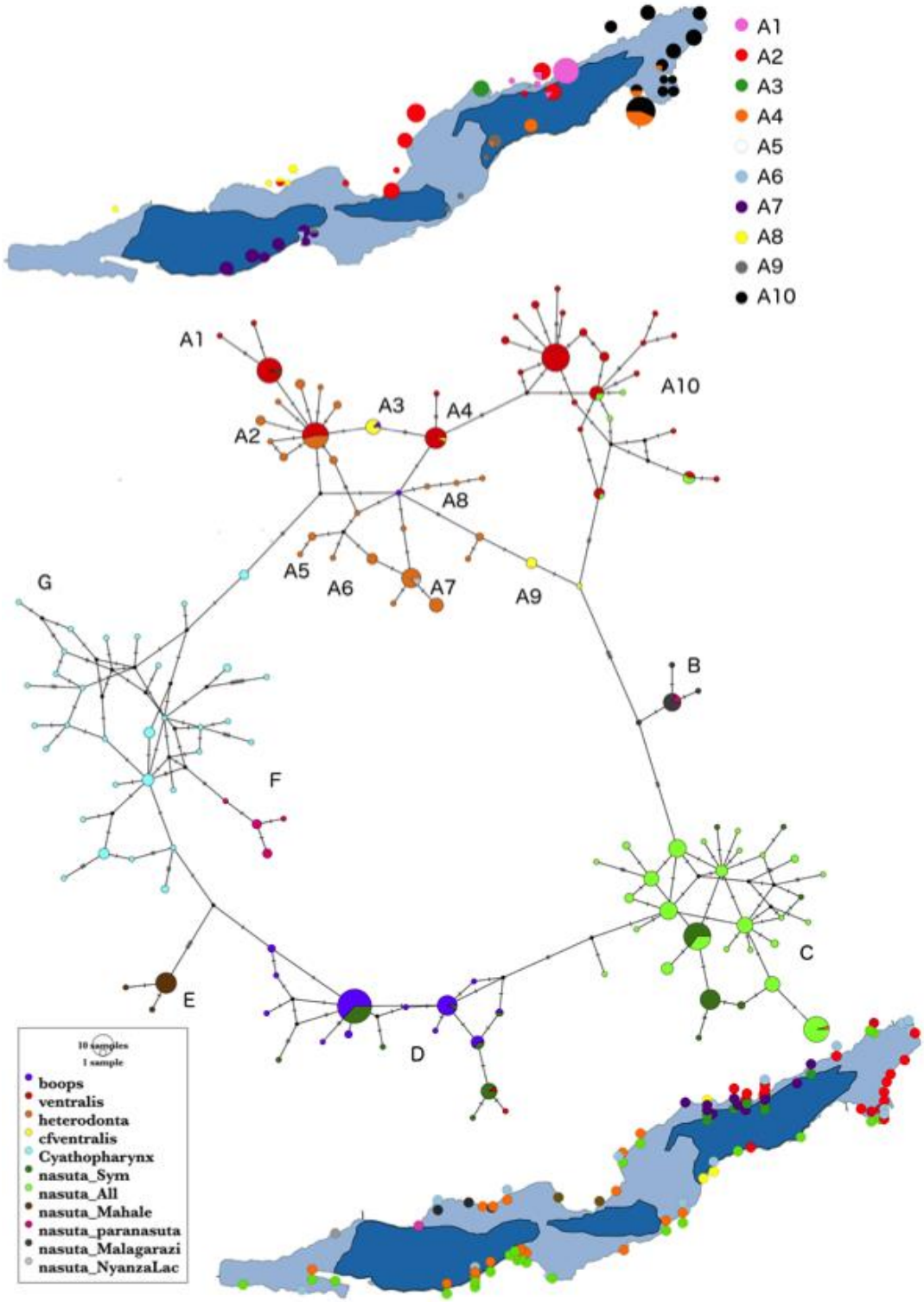

Fig. 18. Analysis of differential gene expression in the presentation experiments. With, top: map of sampling localities of the *O. ventralis*/*O. heterodonta* complex, colours indicate the different subclusters identified in the haplotype network. centre: TCS haplotype network of *Ophthalmotilapia* and *Cyathopharynx*, vertex size denotes the number of sequences per haplotype, colour the identity of specimen (species identify, or assignment to a certain *O. nasuta* variety), below: Map of the sampling localities of the different species of *Ophthalmotilapia* and *Cyathopharynx* and of the different varieties of *O. nasuta*.

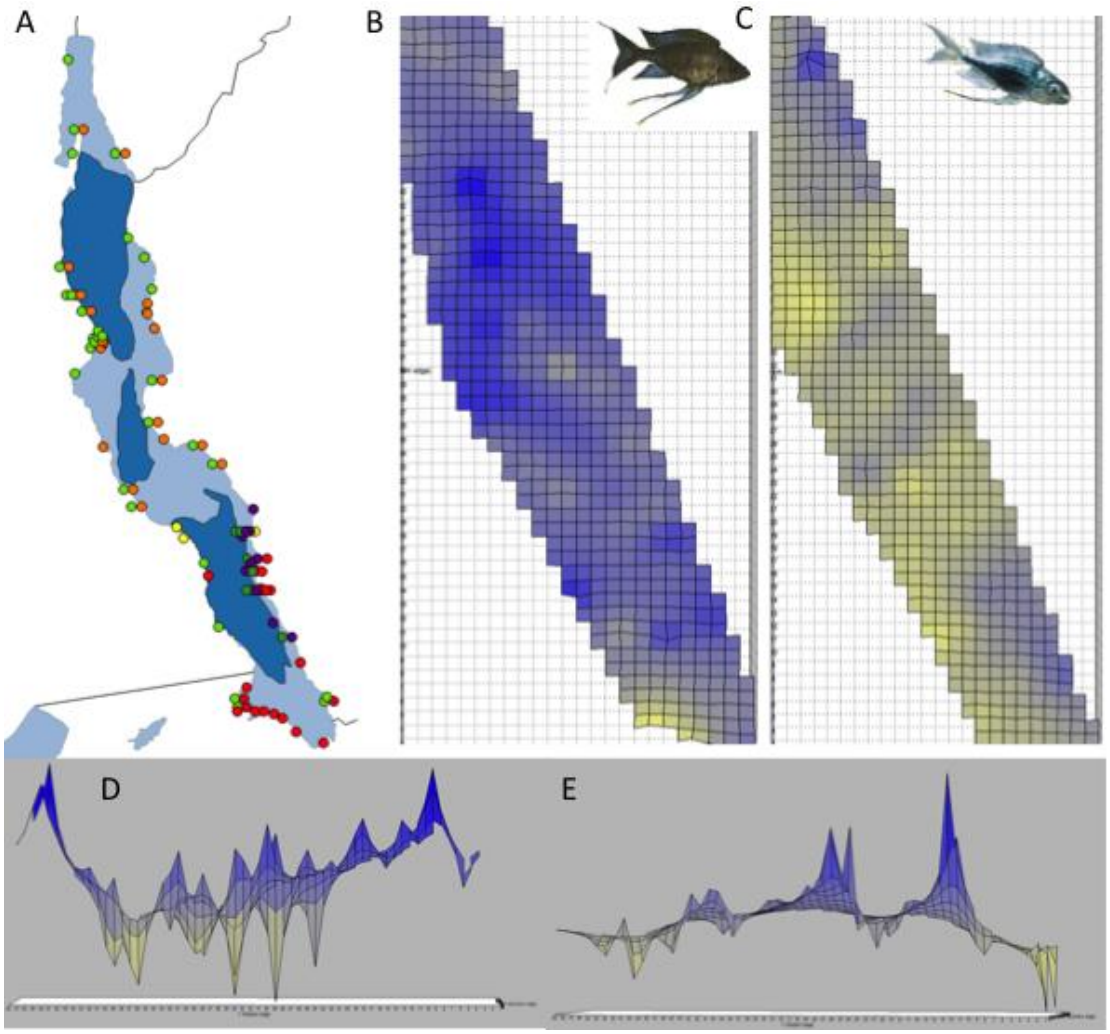

Fig. 19. Genetic landscape shape analysis using Dloops on *O. nasuta* and on the *O. ventralis/O. heterodonta* complex across their distribution. With **A**: Lake Tanganyika with sampling for *O. nasuta* (green), *O. ventralis* (red), *O. cf. ventralis* (yellow), *O. heterodonta* (orange), O. *boops* (purple). **B** and **C**: visualisation of the genetic landscape shape on a grid (10km x 10km) representing the Lake, blue and yellow squares denotes zones that form strong or weak barriers to gene flow for *O. nasuta* (**B**) and for the *O. ventralis*/*O. heterodonta* complex (**C**). **D** and **E**: cross section across the Lake, in which zones that form strong or weak barriers to gene flow are visualised as blue mountains and yellow valleys for *O. nasuta* (**D**) and for the *O. ventralis*/*O. heterodonta* complex (**E**).

4.C.2 Population genomics and Genotyping By Sequencing (GBS)

We investigated the structure of the genus using the GBS dataset with ADMIXTURE. As the cross validation plot did not reveal a clear value for K, several analyses for distinct K-values were explored. This lack of a clear K-value indicates that genomic structuring is layered in *Ophthalmotilapia* with no clear distinction in the degree of variation between the species, variety and population level (Fig. 20).

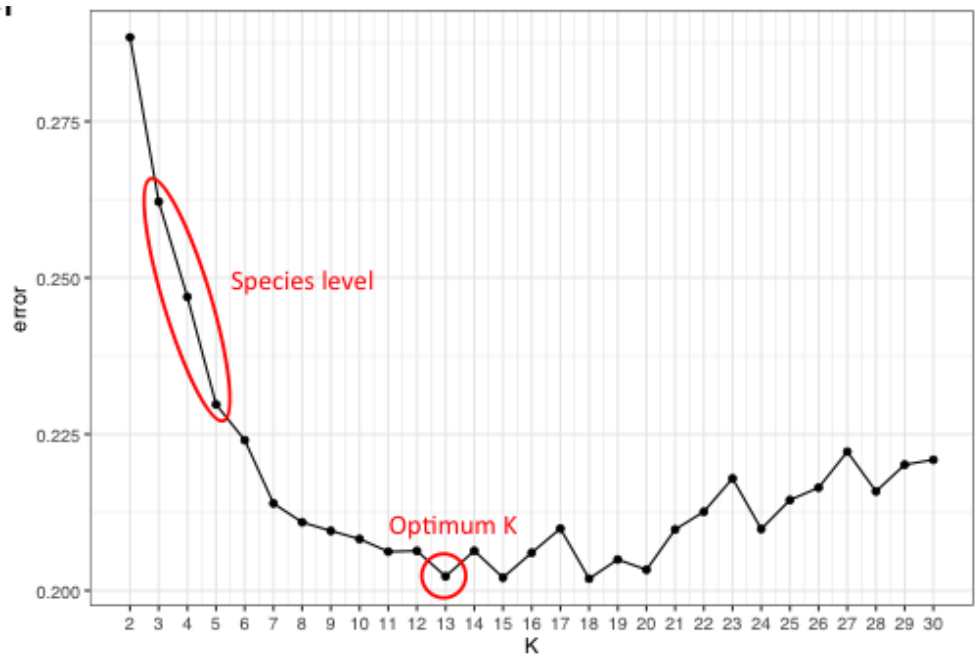

Fig. 20. Cross validation plot to discover the appropriate K value for ADMIXTURE. In the absence of a clear valley, K values that agree with sharp drops as well as those that have low values were examined.

For K=3, *O. nasuta*, *O. boops* and the combined *O. ventralis*/*O. heterodonta* complex were supported as distinct groups (not shown). However, for K=4, *O. nasuta* was divided into a western/southern and a central-eastern group (Fig. 21). The separation of O. nasuta somewhat agrees with the results of the population genetics analysis (4.C.1). The former group corresponded to the specimens that bore the wide spread haplotype C, whereas the latter group contained the specimens with haplotypes B, E and F (Fig. 18). These contain the specimens identified as 'paranasuta', which is sometimes considered a distinct species. Only for K= 5, a separation within the *O. ventralis*/*O. heterodonta* complex could be seen, although it did not agree with the current delineation of the species (not shown). For larger values of K (eg. K=13) separations were supped in which *O. ventralis* and *O. heterodonta* were supported as harbouring mutually distinct groups. With this higher K-value, specimens of potential hybrid origin could also be retrieved. These included a specimen that was morphologically identified as potential hybrid between *O. nasuta* and *O. boops* (Fig. 22).

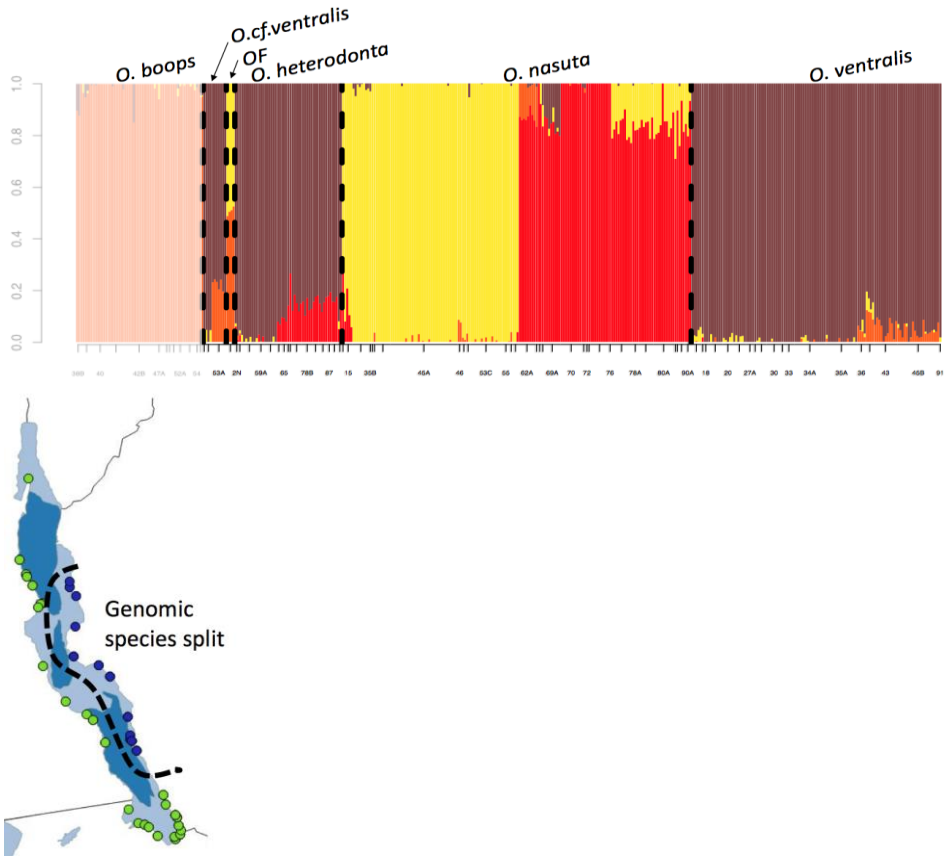

Fig. 21. ADMIXTURE analysis for the GBS dataset of all *Ophthalmotilapia* specimens with K=4. For this value of K, *O. boops* and the *O. ventralis/heterodonta* complex are supported, whereas *O. nasuta* is split in a central eastern and a western/southern group. OF: lab bred *O. ventralis* x *O. nasuta* hybrids.

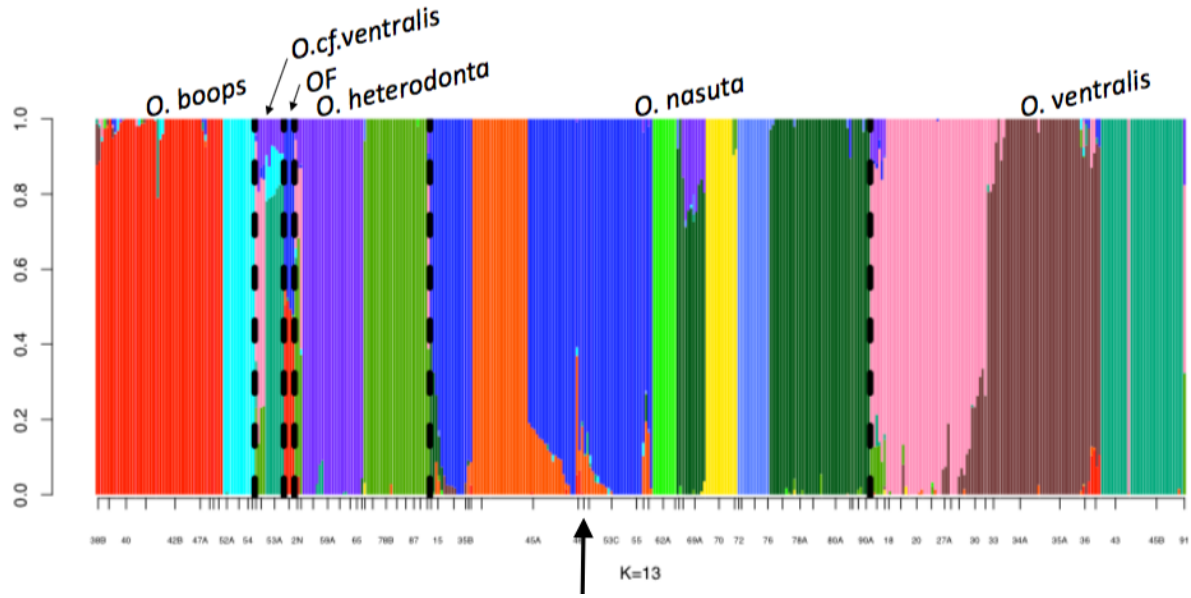

Fig. 22. ADMIXTURE analysis for the GBS dataset of all *Ophthalmotilapia* specimens with K=13. For this value of K, *O. boops* and the *O. ventralis/heterodonta* complex are supported, whereas *O. nasuta* is split in a central eastern and a western/southern group. OF: lab bred *O. ventralis* x *O. nasuta* hybrids. The arrow denotes a specimen that was morphologically identified as of potential *O. boops* x *O. nasuta* hybrid origin.

4.D References

- Allender, C. J., O. Seehausen, M. E. Knight, G. F. Turner, and N. Maclean. 2003. Divergent selection during speciation of Lake Malawi cichlid fishes inferred from parallel radiations in nuptial coloration. Proceedings of the National Academy of Sciences of the United States of America 100:14074–9
- Brawand D, Wagner CE, Li YI, Malinsky M, Keller I, Fan S, Simakov O, Ng AY, Lim ZW, Bezault E, et al. 2014. The genomic substrate for adaptive radiation in African cichlid fish. Nature 513(7518):375–381
- Elmer, K. R., S. Fan, H. M. Gunter, J. C. Jones, S. Boekhoff, S. Kuraku, and a Meyer. 2010. Rapid evolution and selection inferred from the transcriptomes of sympatric crater lake cichlid fishes. Molecular Ecology 19 Suppl 1:197–211
- Elshire, R. J., J. C. Glaubitz, Q. Sun, J. a Poland, K. Kawamoto, E. S. Buckler, and S. E. Mitchell. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PloS one 6:e19379.
- Friard, O., Gamba, M., 2016. BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. Methods Ecol. Evol. 7, 1325–1330.
- Hanssens, M., J. Snoeks, and E. Verheyen. 1999. A morphometric revision of the genus Ophthalmotilapia (Teleostei, Cichlidae) from Lake Tanganyika (East Africa). Zoological Journal of the Linnean Society 125:487–512
- Hulsey, C. D., M. C. Mims, N. F. Parnell, and J. T. Streelman. 2010. Comparative rates of lower jaw diversification in cichlid adaptive radiations. Journal of evolutionary biology
- Koblmüller, S., W. Salzburger, and C. Sturmbauer. 2004. Evolutionary relationships in the sand-dwelling cichlid lineage of lake Tanganyika suggest multiple colonization of rocky habitats and convergent origin of biparental mouthbrooding. Journal of molecular evolution 58:79–96.
- Kocher, T. D. 2004. Adaptive evolution and explosive speciation: the cichlid fish model. Nature reviews. Genetics 5:288–98.
- Loh, Y.-H. E. et al. 2013. Origins of Shared Genetic Variation in African Cichlids. Molecular biology and evolution 30:906–917.
- Meier JI, Marques DA, Mwaiko S, Wagner CE, Excoffier L, Seehausen O. 2017. Ancient hybridization fuels rapid cichlid fish adaptive radiations. Nat Commun. 8(1):14363.
- Mendelson TC, Shaw KL. 2012. The (mis)concept of species recognition. Trends Ecol Evol. 27(8):421-427.
- Nevado B, Koblmuller Sturmbauer C, Snoeks J, Verheyen E. 2009. Complete mitochondrial DNA replacement in a Lake Tanganyika cichlid fish. Molecular Ecology 1820 : 4240–55
- Nevado, B., V. Fazalova, T. Backeljau, M. Hanssens, and E. Verheyen. 2011. Repeated unidirectional introgression of nuclear and mitochondrial DNA between four congeneric Tanganyikan cichlids. Molecular Biology and Evolution 28:2253–2267.
- Panfili J, Thior D, Ecoutin J-M, Ndiaye P, Albaret J-J. 2006. Influence of salinity on the size at maturity for fish species reproducing in contrasting West African estuaries. J Fish Biol. 69(1):95-113.
- Poll, Max & H. Matthes. 1962. "Trois poissons remarquables du lac Tanganika". Annales du Musée royal de l'Afrique Centrale (Série Zoology). (Serie-8) Zoologique (111); pp. 1-26
- Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. Nat Met. 9(7):671-675.
- Turner GF, Seehausen O, Knight ME, Allender CJ, Robinson RL. 2001. How many species of cichlid fishes are there in African lakes. Mol Ecol. 10(3):793-806.

5. DISSEMINATION AND VALORISATION

A significant part of the obtained results were presented at international scientific meetings and published in international journals (see below). Besides that, the project's output was presented on the project's website, that was taken dopwn after the project's funding ended.

6. PUBLICATIONS

International scientific journals

- Derycke S, Kéver L, Herten K, Van den Berge K, Van Steenberge M, Van Houdt J, Clement L, Poncin P, Parmentier E, Verheyen E. 2018. Neurogenomic Profiling Reveals Distinct Gene Expression Profiles between Brain Parts that are Consistent in Ophthalmotilapia Cichlids. Frontiers in Neuroscience 1, 13 - https://doi:10.3389/fnins.2018.00136
- Kéver L, Parmentier E, Derycke S, Verheyen E, Snoeks J, Van Steenberge M, Poncin P. 2017. Limited possibilities for prezygotic barriers in the reproductive behaviour of sympatric Ophthalmotilapia species (Teleostei Cichlidae). Zoology - https://doi.org/10.1016/j.zool.2017.12.001
- Van Steenberge M, Jublier N, Kéver L, Gresham S, Derycke S, Snoeks J, Parmentier E, Poncin P, Verheyen E (submitted)). The immediate response of females towards congeneric males matches their propensity to hybridize . Behavioural Ecology

Conference abstracts

- Gresham S, Van Steenberge M, Derycke S, Kéver L, Lescroart J, Herten K, Van Houdt J K J, Maes G E, Calboli F, Poncin P, Parmentier E, Snoeks J, Verheyen E. 2019. GBS data confirms past hybridisation among congeneric cichlid species from Lake Tanganyika. Cichlid Science 2019, National Museum of Natural Science (MNCN, CSIC) / Royal Botanical Garden (RJB, CSIC), September 9-12, 2019, Madrid, Spain. Abstract Book page na
- Van Steenberge M, Derycke S, Jublier N, Kéver L, Gresham S, Jublier N, Van Den Berge K, Clement L, Herten, Van Houdt J K J, Maes G E, Poncin P, Parmentier E, Snoeks J, Verheyen E. 2019. Behaviour and neural gene expression predict patterns of asymmetric hybridisation in a mouthbrooding cichlid genus (Ophthalmotilapia) from Lake Tanganyika. Cichlid Science 2019, National Museum of Natural Science (MNCN, CSIC) / Royal Botanical Garden (RJB, CSIC), September 9-12, 2019, Madrid, Spain. Abstract Book page na
- Van Steenberge M, Deryck S, Kéver L, Van den Berge K, Clement L, Gresham S, Herten K, Van Houdt J K J, Maes G E, Koblmüller S, Poncin P, Parmentier E, Snoeks J, Verheyen E. 2018. Is differential gene expression in the female brain linked to assortative mating in a mouthbrooding cichlid (Ophthalmotilapia) from Lake Tanganyika? Speciation in Ancient Lakes (SIAL 8), Celebrating 25 Years and Moving Towards the Future, 29 July–3 August 2018, Entebbe (Uganda), Abstract Book page 61
- Van Steenberge M, Derycke S, Kéver L, Herten K, Van Houdt J, Maes G, Koblmüller S, Poncin P, Parmentier E, Snoeks J, Verheyen E. 2017. The genetic and behavioural basis of female mate recognition in sympatric Ophthamotilapia species. 4–7 September 2017, Prague, Czech Republic. Unpublished
- Derycke S, Herten K, Kéver L, Van Houdt J, Van Steenberge M, Snoeks J Parmentier E, Poncin P, Verheyen E. 2016. Comparison of genome wide SNP signatures within and between four Ophthalmotilapia species (Pisces Cichlidae) : Speciation Unveiled? Annual Symposium of the Fisheries Society of the British Isles. 18-22 July 201, Bangor University UK : Pontio Arts and Innovation Centre, United Kingdom. Abstract Book page #
- Derycke S, Herten K, Van Steenberge M, Kéver L, Maes G, Snoeks J Poncin P, Parmentier E, Verheyen E. 2016. E. Speciation genomics of cichlids (Ophthalmotilapia) from Lake Tanganyika Zoology 201, 16 and 17 December 201, University of Antwerp. Abstract Book page 41
- Derycke S, Kéver L, Herten K, Van Houdt J, Maes G, Van Steenberge M, Snoeks J Parmentier E, Poncin P, Verheyen E. 2016. Pinpointing behavioral responses during mating using differential gene expression in the female brain of cichlid fish. Joint symposium on Eco-Evolutionary dynamics and Flanders Annual Meeting of Ecology-symposium 4 : Ecological and evolutionary genomics. Abstract Book page 24
- Derycke S, Herten K, Kéver L, Van Houdt J, Van Steenberge M, Snoeks J Parmentier E, Poncin P, Verheyen E. 2016. Comparison of genome wide SNP signatures within and between four

Ophthalmotilapia species (Pisces Cichlidae) : Speciation Unveiled? Annual Symposium of the Fisheries Society of the British Isles. 18-22 July 201, Bangor University UK : Pontio Arts and Innovation Centre, United Kingdom. Abstract Book page #

- Derycke S, Herten K, Van Steenberge M, Kéver L, Maes G, Snoeks J Poncin P, Parmentier E, Verheyen E. 2016. E. Speciation genomics of cichlids (Ophthalmotilapia) from Lake Tanganyika Zoology 201, 16 and 17 December 201, University of Antwerp. Abstract Book page 41
- Derycke S, Kéver L, Herten K, Van Houdt J, Maes G, Van Steenberge M, Snoeks J Parmentier E, Poncin P, Verheyen E. 2016. Pinpointing behavioral responses during mating using differential gene expression in the female brain of cichlid fish. Joint symposium on Eco-Evolutionary dynamics and Flanders Annual Meeting of Ecology-symposium 4 : Ecological and evolutionary genomics. Abstract Book page 24
- Derycke S, Herten K, Van Steenberge M, Snoeks J, Van Houdt J, Kéver L, Parmentier E, Poncin P, Verheyen E. 2015. Speciation genomics of Ophthalmotilapia species from Lake Tanganyika. Cichlid Science meeting 2015 in Gra, 6-9 September 2015. Abstract Book page 136
- Kéver L, Poncin P, Derycke S, Verheyen E, Van Steenberge M, Snoeks J Parmentier E. 2015. The reproductive behavior of two sympatric Ophthalmotilapia species (Ectodini) from Lake Tanganyika. Cichlid Science Meeting 201, 6-9 September 201, University of Gra, Austria. Abstract Book page 147
- Derycke S, Kéver L, Parmentier, Poncin P, Snoeks J, Van houdt J, Van Steenberge M, Verheyen E. 2014. GENBAS : Unravelling the genomic basis of speciation in African cichlids from Lake Tanganyika. Zoology 201, 21st Benelux Congress of Zoology LiegE, Belgium (12-13 December 2014). Abstract Book page 145

7. ACKNOWLEDGEMENTS

We greatfully acknowledge the experts who have guided us during the execution of the project.
In chronologiocal order we thank Chris Jiggings (University of Cambridge, UK) in view of his work on speciation in *Heliconius* butterflies, and Lieven Clement and Koen Van Den Berghe (Ughent, Belgium) who contributed with their expertise in statistical analyses of gene expression data and Walter Salzburger (University of Basel, Switzerland), Hannes Svardal (University of Antwerp, Belgium) for their advise on the phylogenomic analyses.

ANNEXES

1. Derycke S, Kéver L, Herten K, Van den Berge K, Van Steenberge M, Van Houdt J, Clement L, Poncin P, Parmentier E, Verheyen E. 2018. Neurogenomic Profiling Reveals Distinct Gene Expression Profiles between Brain Parts that are Consistent in Ophthalmotilapia Cichlids. Frontiers in Neuroscience 1, 13 - https://doi:10.3389/fnins.2018.00136
2. Kéver L, Parmentier E, Derycke S, Verheyen E, Snoeks J, Van Steenberge M, Poncin P. 2017. Limited possibilities for prezygotic barriers in the reproductive behaviour of sympatric Ophthalmotilapia species (Teleostei Cichlidae). Zoology - https://doi.org/10.1016/j.zool.2017.12.001
3. Van Steenberge M, Jublier N, Kéver L, Gresham S, Derycke S, Snoeks J, Parmentier E, Poncin P, Verheyen E (submitted). The immediate response of females towards congeneric males matches their propensity to hybridize . Behavioural Ecology