Charles University

Faculty of Science

Study programme: Biology

Branch of study: Zoology

Bc. Tomáš Dvořák

Phylogeography and population structure of two loach species (Teleostei: Nemacheilidae) in Southeast Asia

Fylogeografie a populační struktura dvou druhů mřenkovitých ryb (Teleostei: Nemacheilidae) z jihovýchodní Asie

Master thesis

Supervisor: Dr. Jörg Bohlen, PhD.

Consultant: Mgr. Zuzana Musilová, PhD.

Prague, 2018

Abstract

The freshwater fish species *Schistura robertsi* and *Paracanthocobitis zonalternans* belong to the family Nemacheilidae, which is distributed across whole Eurasia and with one species in northeast Africa.

P. zonalternans occurs in lowland habitats through western Southeast Asia from Central Myanmar until northern Malaysia. The distribution area is of biogeographic interest, because it crosses several known biogeographic barriers, namely the border between Indian and Indochinese freshwater fauna along the Salween River, the Isthmus of Kra, the Krabi – Surat Thani line and the Kangar - Pattani line. In the present study, around 250 specimens of *P. zonalternans* from 62 localities across the whole distribution area were investigated using genetic (nuclear and mitochondrial sequences), morphologic and geologic data.

The genetic data reveal the existence of seven major clades within the analysed material, each of them with a distinct geographic distribution area and only few cases of overlap, but with occurrence of some cases of secondary contact. Divergence time estimations suggested that *P. zonalternans* is about 18 my old, and a biogeographic analysis located the region of origin in the Tenasserim region (nowadays southern Myanmar). The global sea level fluctuations seem to have had a strong impact on the formation of the observed genetic lineages. Morphologic analyses supported the status of undescribed species for one of the lineages, which was also in phylogenetic analyses very distinct to all the others. The other clades showed only very little morphologic differentiation and are considered as conspecific. The results also show that some characters formerly used for species descriptions have no diagnostic value.

S. robertsi group includes five described species *S. robertsi, S. aurantiaca, S. balteata, S. crocotula, S. cincticauda* and also an undescribed species, *S.* sp „Sumo“. In the present study, around 200 specimens from 47 localities were used. Phylogenetic results show ten major lineages, seven of them are corresponding to known described or undescribed species. However, basing on the present analyses the *S. robertsi* is polyphyletic. Most of the lineages are genetically deeply isolated and part of them is fitting to biogeographical pattern, but with frequent co-occurrence in secondary contact areas.

Key words: *Paracanthocobitis*, *Schistura*, Nemacheilidae, Biogeography, Southeast Asia, Geology, Global sea water fluctuation, Morphometry, Colonisation, Bariers

Abstrakt

Sladkovodní druhy ryb ze skupin *Schistura robertsi* a *Paracanthocobitis zonalternans* patří do čeledi Nemacheilidae. Nemacheilidae jsou rozšířeni po celé Eurasii a jeden druh obývá severovýchodní Afriku.

Paracanthocobitis zonalternans žije v nížinném prostředí západní části jihovýchodní Asie od centrální Barmy po severní část Malajsie. Areál rozšíření je z biogeografického pohledu velmi zajímavý, protože prochází řadou biogeografických bariér, jmenovitě hranicí mezi Indickou a Indočínskou faunou podél řeky Salween, Krajskou šíjí, linií Surat thani – Krabi a linií Kangar – Patani. K této studii bylo za použití genetických (mitochondriální a jaderné sekvence), morfologických a geologických dat, využito přibližně 250 jedinců z 62 lokalit pokrývajících celý areál rozšíření

Genetická data z analyzovaného materiálu ukázala sedm hlavních linií odrážejících biogeografii areálu, s minimálním výskytem překryvu, ale s několika výskyty druhotného kontaktu. Datování událostí ukázalo, že *P. zonalternans* vznikl před 18 ml a biogeografické analýzy ukázaly, že se tak stalo v oblasti Tenasserim (jižní Barma). Změny hladiny světového oceánu měly velký vliv na formování zmiňovaných genetických linií. U jedné z linií, která již ve fylogenetických analýzách byla výrazně vzdálená ostatním, podpořily morfologické analýzy status nepopsaného druhu. Ostatní linie ukazují jen nevýrazné morfologické rozdíly a jsou považovány za stejný druh. Výsledky také ukazují, že některé znaky dříve použité pro popis druhů, nemají diagnostickou hodnotu.

Skupina kolem druhu Schistura robertsi obsahuje pět popsaných druhů: *S. robertsi, S. aurantiaca, S. balteata, S. crocotula, S. cincticauda* a také jeden nepopsaný druh *S.* sp. Sumo. V této studii bylo použito přibližně 200 jedinců ze 47 lokalit. Výsledky fylogenetických analýz ukázaly deset hlavních linií, z nichž sedm odpovídá popsaným, nebo známým nepopsaným druhům. Nicméně, na základě analýz je druh *S. robertsi* polyfyletický. Většina linií je hluboce izolovaná a některé z nich odrážejí biogeografii areálu, ale s velmi častým výskytem druhotného kontaktu.

Klíčová slova: *Paracanthocobitis*, *Schistura*, Nemacheilidae, Biogeografie, Jihovýchodní Asie, Geologie, Izolované linie, Morfometrie, Kolonizace, Bariéry

Acknowledgments

First big thank must go to my supervisor Jörg Bohlen and Vendula Bohlen Šlechtová for the selection of this great theme for my master thesis. I must also thank them for the huge amount of time invested to teaching me most of the methods and analyses and mainly for their tolerance with my postponing of everything to the last minute.

Big thank is also for my consultant Zuzana Musilová for giving me the best recommendations about the structure of the thesis.

I must thank also to the friends of my supervisor and fish collectors from Thailand Jens Kuehne, Kampul Udomritthiruji and Andreas Hartl and also to one of the best ichthyologist Maurice Kottelat, for collecting and sending to the collections of our lab huge amount of fantastic samples from the best localities.

Big thank must go also to my whole family for great support of my studies and for the tolerance of my nervous mood in last days, while I was finishing the master thesis.

On the end I must thank also for the financing of the study from the grants. Big thank must go to Gesellschaft für Ichthyologie for supporting the part about *Paracanthocobitis zonalternans* group and to GAČR for financing the part about *Schistura robertsi* group, because without them the study will never exist.

Declaration

I hereby declare that I have elaborated the present thesis on my own, under the supervision of my supervisor and support of the consultant, and that I have listed the used resources of scientific information in the list of literature and that neither this work nor any part of it was presented with aim to gain other or the same academic title.

Prohlášení

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně pod vedením školitele a odborné podpory konzultantky, řádně citoval všechny použité informační zdroje a literaturu, a tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze dne: 13.8.2018 ……

Tomáš Dvořák

Content

1.Introduction

1.1. Southeast Asia

Southeast Asia is formed by a continental part and an island part. The continental part is consisting from Myanmar on western borders, through Thailand, Laos Cambodia, part of Malaysia and Vietnam on eastern borders. The island part contains Indonesia, part of Malaysia, Brunei and Papua New Guinea. In the study will be more or less mentioned only the western part of Southeast Asia containing Myanmar, South-western Thailand and continental part of Malaysia.

The Southeast Asian region is characterized by seasonal weather. The seasonal weather is caused by monsoons. The wet season stretch from the end of May up to the end of September and is connected with summer monsoon blowing in north-west direction, which is bringing wet air from the ocean to the continent. The dry season starts in November and continues to April. In the dry season the winter monsoon is blowing in south-east direction and bringing dry air from the mountains through the continent to the ocean (Loo et al., 2014).

As in Europe the different seasons of the year have big impact on all natural processes including the freshwater fauna, also the seasonal weather changes in Southeast Asia trigger the life cycle of freshwater fishes. With the wet season starts for most species the spawning, the high amount of food and the fastest growth.

1.2. Biogeography and Geological history of Southeast Asia

Whole Southeast Asia is an area with very high biodiversity and is very often established like one big biodiversity hotspot for plants, invertebrates, terrestrial fauna and mainly freshwater fauna. The origin of this high diversity in terrestrial and Freshwater fauna was contributed by diversification as well as migration events (de Bruyn et al, 2013), but there is up to now only little known about the colonisation and other events, which made this high diversity (de Bruyn et al, 2014). The biodiversity in the Southeast Asia is now going through heavy crisis and quite high number of species is now extinct or critically endangered (Allen et al., 2010).

Western Southeast Asia could be divided into three subregions on the base of distribution area of its freshwater fauna. First one is the Burmese subregion, which contains Irrawaddy, Sittaung and Salween river basins, which share the majority of their freshwater fauna with the Indian region. Second is the Indochinese subregion, which contains Mae Klong, Chao Phraya and Mekong river systems. Third one is the subregion of the Malay

peninsula and Sundaland. The Burmese subregion is separated from the Indochinese subregion by the Tenasserim Mountain Ridge that separates the rivers that drain into the Sea of Andaman from those that drain into the Gulf of Thailand (Zakaria-Ismail 1994). For better orientation see the map in Fig.1. Despite their relatively scant altitude these mountains form an effective barrier for freshwater animals, making them the border of distribution for freshwater fishes, not only on species level, but even for several genera (the eastern border for e.g. *Botia* and *Psilorhynchus* from the Indian fauna and the western border for e.g. *Nemacheilus* and *Yasuhikotakia* from the Indochinese fauna) (e.g. Šlechtová et al., 2006, Šlechtová et al., 2007).

The fauna of the Malay Peninsula changes from north to south; north of the Isthmus of Kra, which is the narrowest point of the Malay Peninsula at about 10˚30'N, it contains Indian fauna in the west and Indochinese fauna in the east. Further south, the fauna changes considerably, so that about 20% of the roughly 300 freshwater fish species in western Malaysia are neither shared with the Indian nor the Indochinese fauna (Lim & Tan 2002). Despite the fact that the northern border of the Malay Peninsula as zoogeographic zone has been considered to be of similar biogeographic scale as the Wallace-line between the Asian and the Australian faunas (Parnell 2013), the exact point or latitude where the fauna changes to the Malay fauna has been controversially discussed as well as the cause of this faunistic shift. The first to mention that Indochina and the Malay Peninsula (and with it whole Sundaland) represent different biogeographic units was Wallace (1869). He assumed the transition between these units to be located in the Tenasserim region, meaning rather far north on the Malay Peninsula. Wells (1976) and Hughes et al. (2003) found a significant transition in the forest bird fauna just north of the Isthmus of Kra between 11˚ and 13˚N. They assigned this transition to a shift in vegetation from semi-evergreen rain forest to monsoon forest that has been identified to exist between 11˚ and 14˚N (Richards 1996). Although not studied in such detail as birds, a transition of species at the Isthmus of Kra has also been observed in mammals, reptiles, amphibians and butterflies (Hughes et al. 2003; Chen et al. 2018).

In contrast to these findings in terrestrial animals, a brief study on freshwater shrimps by de Bruyn et al. (2005) suggested that the faunistic transition occurs a few hundred km further south of the Isthmus of Kra at the Surat Thani – Krabi line, where a broad stripe of lowland crosses the whole Malay Peninsula in SW – NE direction. This stripe is not the point of vegetation change, but has been submerged under sea water during the periods of elevated sea levels. Two times during the Cenocoic, once during early-middle Miocene (24-13 mya) and once during the early Pliocene (5.5-4.5 mya) the global sea level has been about 100 m

higher than today and created marine transgressions across the Malay Peninsula, most pronounced at the Surat Thani – Krabi line (Woodruff 2003). This marine transgression has separated the populations of freshwater shrimps for such long period that this event still is reflected in the genetic structure of the animal.

The changes in global sea level of course did not only affect the Malay Peninsula, but all coastlines changed with fluctuations of the sea water. During the highest waters southern Thailand and northern Malaysia are best imagined as a shallow ocean with numerous islands (the present-day mountains) (Woodruff 2003). Marine water submerged a broad stripe of coastal land all around the Gulf of Thailand as well as a big part of Central Thailand (Woodruff 2003). The river system at these times was restricted to the (present-day) upper stretches and freshwater fish populations were isolated from each other. On the other hand, there have been periods with significantly lowered sea level. During these periods the shallow parts of the ocean emerged and as the land enlarged its area, also the rivers prolonged. In several cases, neighbouring rivers met and provided an opportunity for formerly isolated freshwater animal populations to expand or contact (Haq et al. 1987).

Altogether, western Southeast Asia has had a very active and complex geological and climatic past. Its fauna has been influenced by faunistic barriers like mountain ridges and marine transgressions as well as by repeated serious shrinking and expansion of their habitat due to sea water fluctuations. The fauna of the Indochinese region a perfectly suited model to study the impact these factors have on biogeography and evolution. As mentioned above, strict freshwater species are best suited as model for such investigations, and widespread and common taxa are better suited as model than local endemics or rare taxa because the later usually are the result of a single local isolation event, while widespread species may have been the object to various past events.

1.3. Nemacheilidae

The family Nemacheilidae (as subfamily Nemacheilinae) was first time established by Regan (1911). Since then, it usually was placed as subfamily Balitorinae into the family Cobitidae Swainson 1839. Later Sawada (1982) found the closer relation of Nemacheilinae with Balitorinae and they together were separated from Cobitidae under the family Balitoridae. As a separate family Nemachelidae were established by Šlechtová et al. (2007).

Nemachelidae are with approximately 700 species one of the biggest families of freshwater fishes and they are spread across whole Eurasia and with two known species also

from northeast Africa (Kottelat, 2012, Prokofijev and Golubtsov, 2013). The hotspot of the distribution of the family Nemacheilidae is in the Indian and Indochinese region. In Europe are known only two genera, *Oxynoemacheilus* (Banarescu & Nalbant, 1966), represented by three species in the Balkan Peninsula and *Barbatula* (Linck, 1790), widespread across Europe with several species, with one species *Barbatula barbatula* living also in Czech Republic.

Fishes from Nemacheilidae family are living in all types of freshwater habitat: from small fast flowing streams, through lakes, swamps, caves, periodic waters, up to the biggest rivers like Mekong. The quite big diversity of the habitats has also very high manifestation on their morphology. They vary in adult size between 3 (*Petruichthys* sp. 'rosy') and 50 cm (*Triplophysa siluroides* (Herzenstein, 1888), have highly elongated, stout and blind species. Colouration in many species is camouflage, but also brilliantly colourful or completely white species exist. The ecological plasticity of Nemacheilidae can be visualised by the fact that the highest freshwater fish species on earth (*Hedinichthys yarkandensis*, up to 5876 m a.s.l. (Day, 1877)) belongs to Nemacheilidae as well as the known deepest cave fish (*Triplophysa gejiuensis*, more than 400 m below surface (Chu and Chen,1979)).

The family Nemacheilidae is formed by several major lineages (clades) with only very small overlap in their distribution areas (Šlechtová and Bohlen, 2010b). The one that is relevant for the present study is the Indian clade that collects nearly all Nemacheilidae on the Indian subcontinent and the Near East south of the Balkan, Caucasus, Hindukush and Himalayan Mountains and west of the Tenasserim Mountains plus several species in Indochina and Sundaland. It is the clade with the highest diversity, containing 22 out of 48 valid genera. Both model groups studied here, the *Paracanthocobitis zonalternans* group and the *Schistura robertsi* group, belong to the Indian clade. This clade was isolated from the remaining clades of Nemacheilidae by the Himalayan orogenesis, which led to an efficient separation of the freshwater systems between the Indian subcontinent and Indochinese by the uplift of the above named mountains and the Tibetian Plateau. The isolation of the Indian clade can be dated back to 28-26 mya on the base of geologic data (Clark et al., 2004), and will be used in the frame of the present study as calibration point for the back calculation of the age of genealogic events.

The huge diversity inside the family Nemacheilidae is up to now still unexplored and it will take still quite long time to understand the whole diversity in this family. The family contains a high amount of undescribed species and genera. Most of the genera and species were described by morphological characters, but the molecular phylogeny shows that the

morphological characters are usually not corresponding with the evolutionary history and several genera and species are polyphyletic (Bohlen and Šlechtová, 2010a).

1.4. *Paracanthocobitis zonalternans* group

The genus *Paracanthocobitis* (Grant, 2007) was recently separated from the genus *Acanthocobitis* Peters, 1861 by having more caudal fin rays, emarginated or truncate (vs pointed) caudal fin, more rounded (vs more triangular) shape of the head, anus closer to anal fin than to pelvic fin and several other. The genus *Paracanthocobitis* contains presently 14 species, which live across most of the Indian faunistic region in India, Sri Lanka, Bangladesh and Myanmar. Interestingly, they occur additionally in a few places in the Indochinese region in Thailand (Mae Klong River basin, southern Thailand, lower Mekong) (Singer and Page, 2015). The only species that occurs in both faunistic regions, and therefore was chosen as model in the present study, is *P. zonalternans* (Kottelat, 1990).

The area inhabited by the *P. zonalternans* group starts in western part of India and continues through Bangladesh, whole Myanmar and into western Thailand (Singer and Page, 2015). They inhabit several main isolated river basins in Southeast Asia containing Irrawaddy, Sittaung, Salween, Tenasserim, Mae Klong and most of the smaller rivers on Malay peninsula, flowing to Andaman see and also to Gulf of Thailand. *P. zonalternans* is typically living on sandy or gravel ground of small or medium rivers. In comparison to many other Nemacheilidae they prefer slower flowing waters and consequently are found more downstream than other Nemacheilidae.

P. zonalternans is a small fish from five to seven centimetres, extraordinarily up to ten centimetres. Typical characters of *P. zonalternans* are dark bars or saddles and a dark midlateral stripe on a light body, the presence of an ocellus on the upper caudal-fin base, a short lateral line, presence of an axillary pelvic lobe and a suborbital flap in males (Kottelat, 1990). Inside the huge area of *P. zonalternans* were observed several small differences in morphological characters like colour pattern, axillary pelvic lobe, depth of caudal peduncle, anus position, number of caudal and pelvic fin rays, but most of the small differences did not show any consistent pattern useful for recognising the variable populations as formal taxonomic units (Kottelat, 1990). Very recently, Singer and Page (2017) suggested splitting *P. zonalternans* into several species, therefore the taxon is here referred to as the *Paracanthocobitis zonalternans* group.

1.5. *Schistura robertsi* group

With more than 200 valid species, *Schistura* is the largest genus within Nemacheilidae, but it is consensus among taxonomists working and the scare genetic data that it does not form a monophyletic lineage, but it is only polyphyletic collection of similar morphological forms or ecotypes (Kottelat 1990, Sember et al., 2016). The possible monophyletic position of the species inside the *S. robertsi* group was up to now never tested on the base of molecular analyses. The species inside the group were connected only by several morphological studies (Kottelat, 1990, Singer and Page, 2011 and Kottelat et al., 2013) and their monophyletic position will be tested in this study.

The *Schistura robertsi* group contains 5 described species *Schistura robertsi* (Kottelat, 1990), *Schistura aurantiaca* (Plongsesthee, Page & Beamish, 2011), *Schistura balteata* (Rendahl, 1948), *Schistura cincticauda* (Blyth, 1860), *Schistura crocotula* (Plongsesthee, Kottelat & Beamish, 2013) and also one undescribed species *Schistura* sp. Sumo known in aquarium trade, which is also very often placed under the species *S. balteata*.

The distribution area of *Schistura robertsi* group is quite smaller than that of the *P. zonalternans* group and ranges from southeast Myanmar, through most of western Thailand until northern Malaysia. The most northern species are *S. cincticauda* and *Schistura* sp. 'Sumo' in the Salween river basin and *S. balteata* known from the Dawei River in north Tenasserim. *S. aurantiaca* was known from the southern Mae Klong River system, but the present study shows it to occur in the whole Mae Klong River system. *S. crocotula* inhabits the Khanan river and usually several small streams in Prachuap Khiri Khan province. *S. robertsi* is known from the whole Thai part of the Malay Peninsula including the provinces Phang Nga, Phuket, Surat Thani, Ranong, Songkhla and Langkawi Island in Malaysia.

Most species of the *S. robertsi* group occupy small to medium rivers with moderate flow rate and gravel and rock structure. Some species (*S. balteata, S. aurantiaca, S. robertsi*) can reach the very small tip of mountain streams. The smallest species, *S. robertsi*, is nearly exclusively found in very small forest streams (often < 1 m broad and < 10 cm deep) with sand bottom and leaf litter. Since the species is nearly never found in medium sized streams, the populations in the small streams are strongly isolated from each other. Usually very few individuals are found in these small forest streams, therefore an exceptionally small effective population size can be assumed for these populations. The small effective population size

together with a strong isolation of populations opens the possibility of a strong genetic drift in these populations, which would result in a comparatively quick differentiation of populations from each other. In order to test if the populations of *S. robertsi* undergo such fast diversification this species was selected as second model in the present study.

Many species of *Schistura* have several dark bars with light interspaces on the body. In the *Schistura robertsi* group a tendency exists to reduce the number of bars and bands and to restrict them to only a part of the body. The most extreme result is visible in *S. balteata*, which has only 2 thin dark bands and 3-4 dorsal saddles on an otherwise light body (Kottelat, 1990). Within the group species differ by presence or absence of dark dots on lower lip, presence or absence of axillary pelvic lobe, position of anus, length of lateral line and different fin ray counts in pelvic, pectoral and caudal fin (Kottelat et al., 2013, Page et al., 2011).

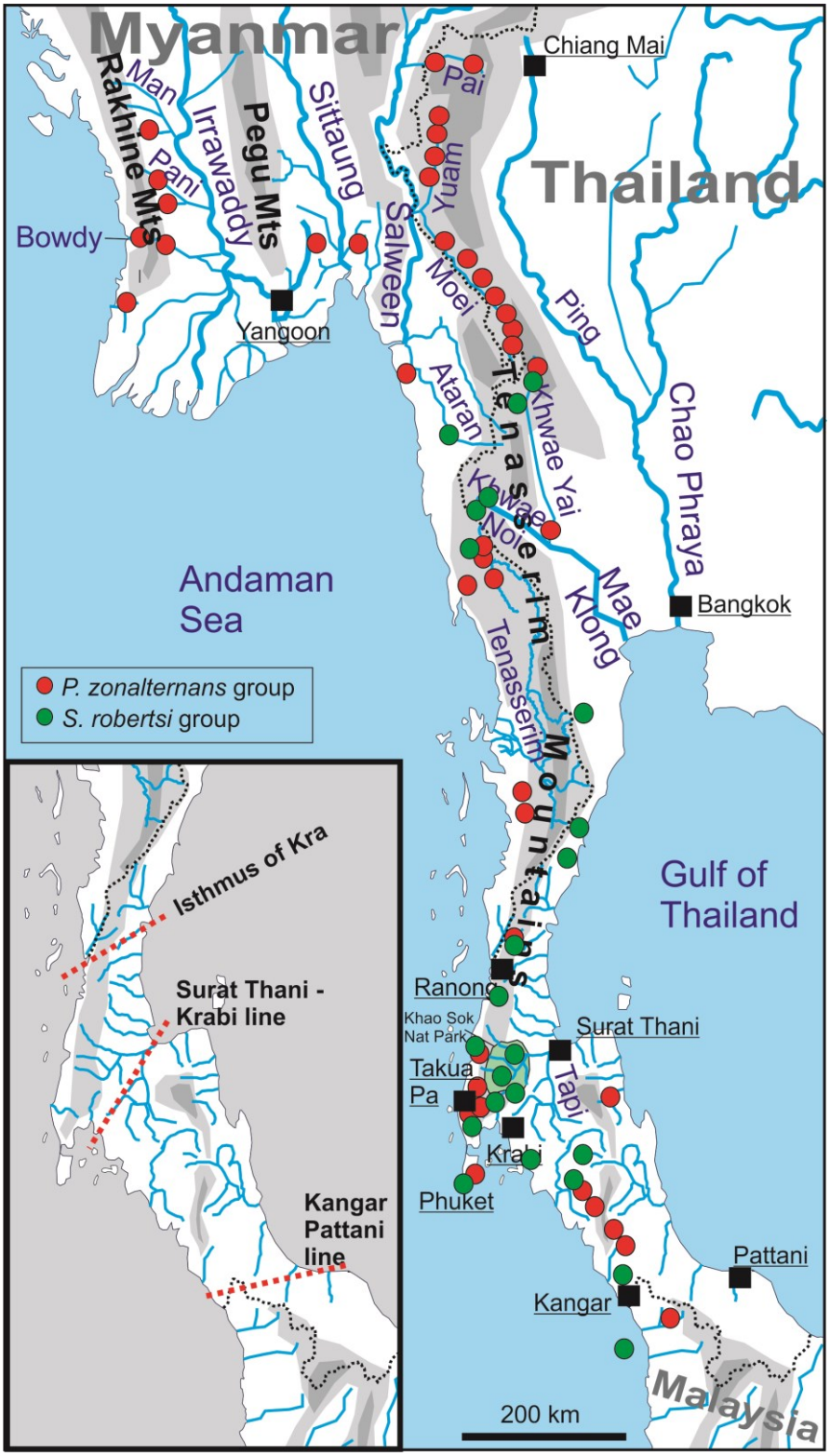


Fig. 1. Geographic overview about the studied area and sampling sites. Red circles indicate sampling sites for the *P. zonalternans* group, green circles for the *S. robertsi* group, black squares indicate cities. Country names in grey, river names in blue, mountain names in bold, city names underlined. the inlet map shows the position of three major biogeographic barriers, the Isthmus of Kra, the Surat Thani – Krabi line and the Kangar – Pattani line.

2. Aims of the study

First part of the study was focused on phylogeny and biogeography of *Paracanthocobitis zonalternans* group and second part was about concerned on phylogeny and biogeography of *Schistura robertsi* group. Fishes from *P. zonalternans* and *S. robertsi* group occur in many isolated river basins, across several high mountain ridges and few known biogeographical barriers like Isthmus of Kra.

By using one mitochondrial and one or two nuclear genes, I tried to show, how the phylogenetic structure corresponds with the geography of the area.

Next biogeographical event, which was tested, were the sea water level changes and their manifestation on phylogeny of the selected species groups.

After finishing the phylogenetic part of the study, I focused on analyses of morphological or morphometric characters in both species groups. Later, when all the results were ready, I compared them with actual taxonomy of both groups.

Part of the study is also focused on the comparison of the two species groups and how different is the level of isolation among the lineages or species in *P. zonalternans* group and *S. robertsi* group. I also compared the similarities in biogeography and dating of the divergence times of the both groups.

3. Material and methods

3.1. Collecting of samples

Most samples used in this study were collected during several field trips by my supervisor Jörg Bohlen and his wife Vendula Bohlen Šlechtová between years 2005 and 2013. Further samples came from a professional fish exporter from Thailand and a professional field guide in southern Thailand. In nearly all cases the localities of collecting were known. Another lot of samples was provided by a Swiss ichthyologist, Maurice Kottelat. These fishes are particularly valuable, because they came from politically problematic areas like Tenasserim, which was inaccessible for foreigners until very recently and where very few scientists have sampled up to now.

Nowadays there are in the collection of Laboratory of Fish Genetics 340 individuals from *P. zonalternans* group and other closely related species. All these 340 individuals were collected on 62 localities and two different lots were bought from official fish trade or aquarium trade. 178 individuals are fixed in formaldehyde (Table 19 in attachment), 162 whole bodies and 35 tissue samples (usually fins clips removed from bodies before formaldehyde fixation), are fixed in ethanol (Table 18 in attachment).

From *Schistura robertsi* group 237 individuals from 47 localities were collected and five individuals of *Schistura* sp. Sumo from official fish trade or aquarium trade were bought. From these 237 individuals 134 are fixed in formaldehyde (Table 21 in attachment), while 103 specimens and 9 tissue samples are fixed in ethanol (Table 20 in attachment). The overview of sampled localities for *P. zonalternans* and *S. robertsi* groups is provided in Fig.1.

3.2. Laboratory procedures

All laboratory procedures were conducted in the Laboratory of Fish Genetics of the Institute of Animal Physiology and Genetics (Liběchov, Czech Republic).

3.2.1. DNA extraction

For genetic analyzes I used all samples preserved in ethanol. The first step of laboratory work was isolation of genomic DNA. DNA extractions were done by two different techniques. Most samples were isolated by Dneasy Blood & Tissue kit (QIAGEN) and some few were isolated by phenol – chloroform method (Sambrook et al., 1989). The phenol – chloroform method was used usually for isolation of samples with lower DNA concentration, usually

caused by bad fixation of the fish. For DNA extraction a small part of fin (pectoral or pelvic) was used, or in case of very low DNA concentration, I rather used part of the muscle tissue to obtain a higher amount of DNA. DNA extractions using the isolation kit followed the manufacturer's instructions. In some cases, in order to obtain higher concentration of DNA, only half of the volume of the elution buffer recommended in the protocol was used.

After isolation the quality of DNA was checked by electrophoresis. Samples were loaded on 0,8 % agarose gel with GelRed (BIOTUM) in proportion 1µl of GelRed to 20ml of the gel. 2µl of every DNA sample with 2 µl of blue loading buffer (Fermentas) was loaded on the gel placed into an electrophoresis bath filled with TBE buffer. For detection of length and level of fragmentation, 2 µl of 1kb ladder (Gene Ruler 1kb DNA ladder, Fermentas) was loaded into one well.

3.2.2. Markers

For this study one mitochondrial (mtDNA) and two nuclear (nDNA) genes were selected. The mitochondrial gene was cytochrome b (Cyt b), the second most frequently used gene for phylogenetic analyses after the "barcoding" gene Cytochrome oxidase 1 (CO 1). Cytochrome b is a gene, that is coding a group of different types of transmembrane proteins working in photosynthesis, respiration and steroids metabolism (Esposti et al., 1993, Blankenship, 2009).

The selected nuclear genes were the Recombination-activating gene 1(RAG 1) and the interphotoreceptor retinoid-binding protein gene 2 (IRBP 2). RAG 1 is a gene coding a protein, which has a function in activation of immunoglobulin. RAG 1 is rather a conservative gene that was often used also for analyses on higher taxonomical levels – for reconstruction of phylogenetic relationships among genera, families, orders, or even higher. It has been used also for phylogeny reconstruction among loach fishes on various taxonomical levels (e.g. Bohlen et al., 2011, Šlechtová et al., 2008, Šlechtová et al., 2007, Sember et al., 2016). The RAG 1 gene was not used for the analyses of S. robertsi group, because we did not find a useful combination of primers due to a high number of mutations in the priming sites.

The third selected marker, IRBP 2, is a gene that is coding a retinol binding protein. IRBP 2 was one of the genes, which were used for establishing the taxonomy of mammals (Lecompte et all., 2008, Tosi et all., 2003), and was also often used in the phylogeny of loach fishes (e.g. Mayden et al., 2009, Sember et al., 2016).

3.2.3. Primers for Polymerase Chain Reaction (PCR)

For Cyt b, I selected primers designed for cypriniform fishes by Kai Erik Witte, which proved to be very useful for loaches (Šlechtová et al., 2006). RAG 1 was used only for P. zonalternans group with selected combination of forward primer 1F and reverse primer Rv1. For IRBP 1 I selected combination of forward primer 109F and reverse 1162R for S. robertsi group and forward 109F plus reverse 1001R for P. zonalternans group. For some individuals, where the former forward primers failed (probably due to a mutation in the priming site), another primer (101F) had to be used (other information in Tab.1).

Tab.1 Primers used in the study, their sequences and source. F – forward, R – reverse.

Gene	Primer	Direction	Sequence 5' – 3'	Source
Cytochrome b	Glu L.Ca14337-14359	F	GAAGAACCACCGTTGTTATTCAA	Šlechtová et al., 2006
	Thr H.Ca15568-15548	*R*	ACCTCCRATCTYCGGATTACA	Šlechtová et al., 2006
RAG 1	RAG 1F	F	AGCTGTAGTCAGTAYCACAARATG	Quenouille et al., 2004
	RAG Rv1	*R*	TCCTGRAAGATYTTGTAGAA	Šlechtová et al., 2007
IRBP	101F	F	TCMTGGACAAYTACTGCTCACC	Chen et al., 2008
	109F	F	AACTACTGCTCRCCAGAAAARC	Chen et al., 2008
	1001R	*R*	GGAAATGCATAGTTGTCTGCAA	Chen et al., 2008
	1162R	*R*	TGGTGGWCTTYAGGCACTTGT	Chen et al., 2008

Tab.2 Volumes of chemicals.

Chemical	Volume (µl)
PCR H_20	17,25
Blue buffer	2,5
$MgCl_2$	1,5
dNTPs	0,5
Enhancer	0,5
Taq polymerase	0,25
Forward primer	0,25
Reverse primer	0,25
DNA sample	2
Total volume	25

3.2.4. Polymerase chain reaction (PCR)

Every PCR master mix was prepared for required number of samples plus one blank as a control of potential contamination. The master mix was distributed into the vessels and the DNA template was added. PCRs were performed in 25 µl reaction volume containing 10 mM Tris–HCl, 50 mM (HN4) 2SO4, 0.1% Triton X-100, 1.2–1.8 mM MgCl2 (PCR Blue Buffer), 2 mM TMA oxalate (PCR enhancer), 5 nmol of each nucleotide (PCR dNTP mix), 1.25 U of Taq DNA polymerase Unis (all chemicals by Top–Bio) and 12.5 pmol of each primer.

For the volumes add to the PCR master mix see Tab.2. Preparing own master mix instead of using commercial ones is more time consuming, but it is more flexible in case of any need of optimization. The PCRs were carried out on DNA Engine Peltier Thermal Cycler

(BioRad). The PCR profiles are listed in Tab.3. For Cyt b and RAG 1 was used program with touch-down (TD) profile of 1 min 30 s at 60– 55 °C (1 °C/cycle) and 2 min at 72 °C followed by 30 cycles with annealing temperature held at 54 °C.

Tab.3 Amplification profiles used for PCR of selected genes, Temp. = Temperature, TD = touch down.

Gene	Primers	Initial denaturation		Cycle denat.		Annealing		Elongation		Cycles	Final elongation	
		Time	Temp.	Time	Temp.	Time	Temp.	Time	Temp.		Time	Temp
Cyt b	L1+H2	5’	95°C	1’	94°C	1,5’	60-55 /54°C	2’’	72°C	6(TD) 30	7’	72°C
RAG 1	1F+Rv1											
IRBP	1R+3F	5’	95°C	1’	94°C	40’’	57°C	2’’	72°C	35	7’	72°C
	3R+2F	5’	95°C	1’	94°C	40’’	60°C	2’’	72°C	35	7’	72°C
	3R+3F	5’	95°C	1’	94°C	40’’	59°C	2’’	72°C	35	7’	72°C

After PCR reaction, the result was checked on electrophoresis (same conditions as after DNA isolation). In case of required result (PCR product of desired length, no side products), the PCR products were purified. Otherwise the PCR was repeated e.g. with use of other primers or higher amount of DNA template keeping 25 µl total volume. In some cases, multiple PCRs per sample were performed and afterwards purified and concentrated into one vessel.

3.2.5. PCR Purification

After the PCR amplifications all PCR products were purified. During this process, the rests of primers and other chemicals are washed away to get a clean PCR product for next steps. All samples were purified using QIAquick PCR Purification Kit (Qiagen). Most purifications followed strictly the manufacturer’s protocol. In few cases a smaller amount of elution buffer was used, to get higher concentration of the product to ensure the required quality for sequencing. As mentioned above, the samples resisting the proper amplification were PCRed in several tubes in parallel and during the purification procedure concentrated into one vessel, to obtain higher concentration. Quality of purified products was tested on the 0.8% agarose gel on the same way as mentioned above. The concentration of the samples sent for sequencing to a professional sequencing service (Macrogen, Amsterdam), was measured on nanodrop or qubit.

3.2.6. Sequencing PCR

Tab.4 Volumes of chemicals used in sequencing PCR reaction

Chemical	Volume (µl)
PCR H_20	10,36
Seq. buffer	6
Seq. kit (BDT)	2
Primer (F or R)	0,64
DNA sample	1-2*
Total volume	20

For sequencing PCR a Big Dye Terminator (BDT) 1.1 or 3.1 (Applied Biosystems) master mix was used according to manufacturer´s instruction. In order to lower the costs, the sequencing kit was diluted 4x by using 2µl of kit and 6 µl of BDT 1.1 or 3.1 5X Sequencing Buffer. It was performed in 20 µl reaction volume containing and 1.25 pmol of each primer. The Tab.4 provides the volumes of each component per sample. For every sample and gene, the reactions were done in both directions – one with forward and one with reverse primer. The PCR cycle profile followed the instructions of the producer. For details of the settings see the Tab.5

Tab.5 Sequencing PCR profile.

Sequencing PCR	Initial denaturation		Main cycle						Cycles
			Denaturation		Primer incorp.		Elongation		
	Time	Temp	Time	Temp	Time	Temp	Time	Temp	
	1’	96°C	10’’	96°C	5’’	50°C	4’’	60°C	25

After sequencing PCR, the products were purified, with use of DyeEx Spin Kit (Qiagen) following the manufacturer´s protocol. Samples were sequenced either in the Laboratory of Fish Genetics or via sequencing service in Macrogen (Amsterdam).

The absolute majority of the samples were sequenced in the Laboratory of Fish Genetics on a 3130 GA sequencer with use of POP-7 sequencing polymer and 30 or later with 50 cm long capillaries. Before the run, the amount and quality of gel were checked, water and sequencing buffer (1x Capillary Electrophoresis Running Buffer from SIGMA) in vessels exchanged. Also the cleanness of the capillaries was checked before every run.

Just before sequencing, the samples were denaturated for 2 min at 95 °C in the thermal cycler and immediately chilled on ice. The 3130 GA is a four capillary sequencer therefore four reactions are processed in parallel. After the sequencing procedure, the results must be analysed in the software Sequencing analyses 6 to obtain a chromatogram because the output are only raw data

3.2.7. Assembling of sequences

Since the lengths of single reads did not cover the complete selected regions, the final sequences had to be reconstructed by merging, the forward sequences with the

complementary reverse ones. The single reads (chromatograms) were usually 600 to 900 bp depending on the quality of PCR product and length of capillaries The overlap was big enough to reliably assemble the complete required fragment. The sequence assembly was conducted in the SeqMan II module of the DNAStar software package (LASERGENE). Every chromatogram was checked by eye for potential mistakes nucleotide by nucleotide.

3.3. Phylogenetic reconstructions

3.3.1. Alignments

The alignments of homological sequences were done in BioEdit (Hall, 1999) with use of ClustalW (Larkin et al., 2007) multiple alignment algorithm. Since all the selected markers were coding genes, there were no ambiguities in the alignments. For every gene, most of the sequences were of nearly the same length, although some of them were shorter due to e.g. noisy priming sites. These regions were treated as missing data for a given sequence. The alignments were also checked by eye for potential mistakes and in case any insertion or deletion was found, it was again confronted with the corresponding chromatogram.

The final alignments of the mitochondrial gene Cyt b were 1120 bp and 1109 bp in *P. zonalternans* group and *S. robertsi* group, respectively. Alignment of RAG 1 for *P. zonalternans* group was 950 bp long and alignments of IRBP 2 were 810bp and 970 bp for *P. zonalternans* and *S. robertsi* group, respectively. Nuclear combined datasets for *P. zonalternans* group had 1760 bp and combined dataset of all three genes was 2880 bp. Combined dataset of Cyt b and IRBP 2 for *S. robertsi* group was 2079 bp long.

3.3.2. Phylogenetic trees

The final alignments were imported into the MEGA v. 7.0 (Kumar et al., 2007). First step was a reconstruction of pilot neighbour joining (NJ) trees with p-distances to check for potential mistakes in dataset like sample mix-up or duplications When I found any samples in unexpected positions in the phylogenetic trees, I checked them with the existing voucher. When was suspected that the source of mistake comes from the laboratory process or sample mix-up, all the steps from DNA extraction up to sequencing were repeated, to make sure, that the unexpected position of the sample is reality and not mistake. Further the codon starting position was checked and the alignment translated to amino acid sequences to check for the presence of stop codons.

The model test (Lanfear et al., 2012) implemented in MEGA was performed to select the best model of nucleotide substitution for every codon position of every gen of all datasets. The models with the best Bayesian information criterion (BIC) were selected. Overview of the best-fit models is provided in Tab 6.

Tab.6 Selected substitution models, length of alignment, starting codon positions and % of variable positions

Taxon (group)	**Gene (alignment)**	**Length (bp)**	**Codon start**	**Models for codon position**			**Variable positions (%)**
				First	**Second**	**Third**	
P. zonalternans	Cyt b	1120	Second	TN93+G+I	HKY+G	GTR+G	47
	RAG	950	First	HKY+G	HKY+G	K2+G	34
	IRBP	810	First	TN93+G	HKY	K2+G	36
S. robertsi	Cyt b	1109	Third	K2+I+G	HKY+G	GTR+G	49,5
	IRBP	970	First	HKY+G	TN92	K2+G	40

The phylogenetic analyses were performed using the Bayesian inference (BI) in MrBayes 3.2 (Ronquist and Huelsenbeck, 2003). For this purpose, the datasets were exported into nexus format files and MrBayes command block with specifications of partitions, models and analyses conditions were added. Each of the analysis was performed in two parallel runs of 10 million generations with six Metropolis Coupled Markov Chains Monte Carlo (MCMCMC) of default heating conditions. Sampling frequency was set to every 100 Generations. The results were checked in Tracer v 1.6 – a software for analyses of MCMC chains (Rambaut et al., 2014) to see the effective sampling size (ESS) of the parameters. Besides, the stationarity of the log-likelihood scores was checked by plotting them against the generations. The relative burn in of 25% was used and from the remaining trees a 50% majority rule consensus tree was constructed.

As outgroup were used several published sequences of Nemacheilidae from Sember at all (2016) and Cobitis taenia from the related family Cobitidae (see Tab. 22 in Attachments).

3.2.3. Genetic difference

There are several statistical approaches how to define a genetic distance, working with different types of evolution models. However, in the current study the proportional genetic difference converted into percentages was used to allow the easy comparison with other studies. The genetic divergences among the main identified lineages were estimated in MEGA7.

3.3. Dating of events

The ages of cladogenetic events were estimated in BEAST 2.4.8 (Bayesian evolutionary analysis by sampling trees, (Bouckaert et al., 2014). The *.xml input files describing the data, partitions, models and MCMC specifications were prepared in BEAUti (Bouckaert et al., 2014). For calibration, the age of the Indian branch of Nemacheilidae (25-28 mya) was used. Shortly before 25 mya the uplift of the Himalayan Mountains and the Tibetian Plateau shifted the Southeast Asian rivers from a westward draining pattern to an eastward draining pattern and thus disconnected the freshwater systems on the Indian subcontinent from mainland Asia (Clark et al. 2004). This event isolated the Indian freshwater fauna from the Indochinese. collision of India with the main Asian continent.

For the analyses, the partitions were unlinked and assigned the estimated evolutionary models. As priors the Yule process of speciation and relaxed lognormal molecular clock were selected. The calibration point of "Indian lineage" was set to normal distribution with mean of 26.5 (+/- 0.5) MY. The MCMC analyses was set to 5x106 generations with sampling of every 1000 generations. The outputs were checked in Tracer 1.6 (Rambaut et al., 2014) to see if the effective sampling sizes (ESS) for all parameters were sufficient (> 200). A maximum clade credibility tree was built in TreeAnnotator v.1.8.0 (Rambaut and Drummond, 2010) after discarding the first 10% of trees. The final trees were visualised in FigTree 1.4.3 (Rambaut, 2009).

3.4. Biogeographical analyses

The biogeographical analyses were conducted using the Bayesian binary MCMC (BBM) analyses implemented in RASP 4.0 (Reconstruct Ancestral Stage in Phylogenies, Yu et al., 2015). Besides the distribution areas definition, the input files included the maximum clade credibility tree and 200 sample trees from the post burn-in trees inferred by BEAST. The distribution areas were defined based on the disjunct distribution areas of the major clades that reflected different river basins (see Fig.2).

The maximum number of unit areas at each node was set to two for *P. zonalternans* and to three for *S. robertsi* group. Ten MCMC chains were ran for 50000 generations under F81 + G model and sampling frequency of every 100 generations.

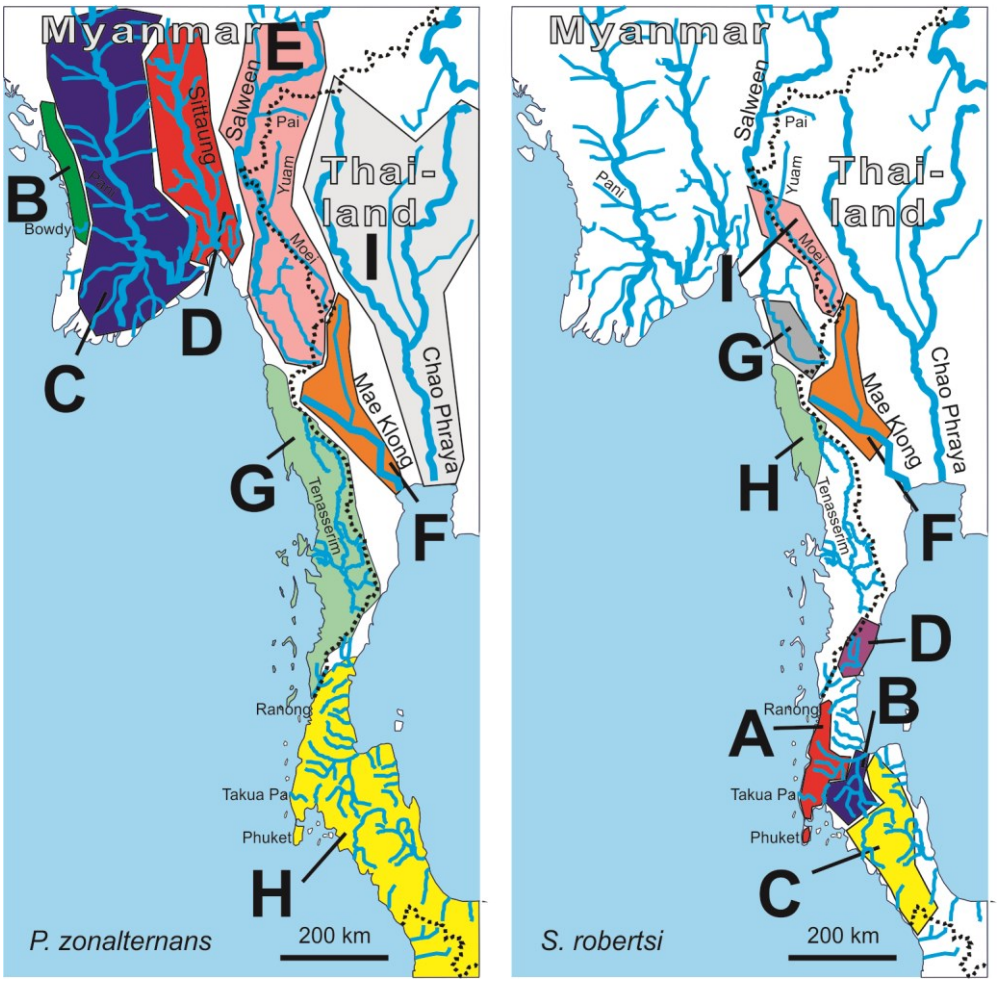


Fig. 2 Maps with defined distribution areas for the main lineages in both groups.

3.5. Morphologic analyses

For morphological analyzes usually fishes fixed in formaldehyde were used. Fishes preserved in ethanol were used only in cases where for a given genetic lineages material fixed in formaldehyde was not available, I must use them. Work with fishes fixed in ethanol must be really fast and careful, because of the risk of drying the material, making the bodies very fragile and the fins are often braking away, what is really unwanted process.

From *P. zonalternans* group 128 individuals and several individuals from closely related species were selected. First I made direct inspection of all selected individuals. That included checking of the colour pattern, body shape, anus position, length of lateral line and several other morphological characters. For measurements I used same characters used by Kottelat (1984). The overview of all the examined characters is provided in Tab.7 and Fig. 3.

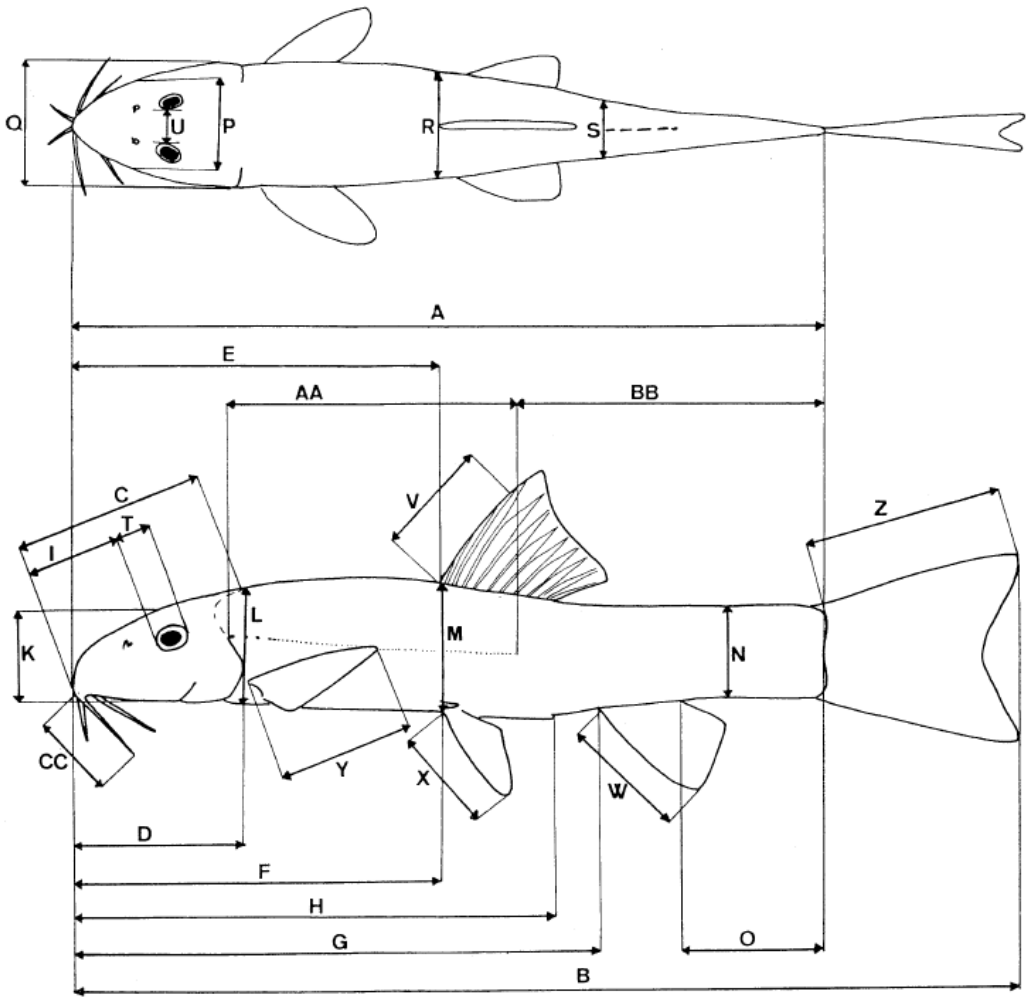


Fig. 3 Schematized Nemachelidae fish with selected characters from Kottelat (1984)

Tab.7 Characters selected for morphometric measurements.

In the picture	Character	In the picture	Character
A	Standard length	Q	Maximum head width
B	Total length	R	Body width at dorsal origin
C	Dorsal head length	S	Body width at anal origin
D	Lateral head length	T	Eye diameter
E	Predorsal length	U	Interorbital width
F	Prepelvic length	V	Height of dorsal fin
G	Preanal length	W	Depth of anal fin
H	Preanus length	X	Length of pelvic fin
I	Snout length	Y	Length of pectoral fin
K	Head depth at eye	Z(1)	Length of upper caudal lobe
L	Head depth at nape	Z(2)	Length of lower caudal lobe
M	Body depth	Z(3)	Length of median caudal rays
N	Depth of caudal peduncle	CC	Barbel length (not measured)
O	Length of caudal peduncle	AA	Lateral line (not measured)
P	Head width at nares	BB	No lateral line (not measured)

In *S. robertsi* group, I made a direct inspection of 193 individuals, where no difference in body shape or in some other measurable character was visible. However, there were differences in anus position, numbers of fin rays and big differences in colour pattern, which was also mentioned in the descriptions of the species, belonging to the group (Kottelat, 1990). I went also through all description papers, and checking every character mentioned there, and later I compare them with the results.

4. Results

4.1. *Paracanthocobitis zonalternans* group

4.1.1. Phylogenetic reconstruction

For construction of phylogenetic trees alignments of three genes were used. The final alignment of mitochondrial gene Cyt b was 1120 bp long, the final alignment of nuclear gene RAG 1 included 950 bp and the final alignment of second nuclear gene IRBP 2 had 810 bp. The length of the final alignment of combined dataset of all three genes was 2880 bp and of combined dataset of nuclear genes it was1760 bp.

Alignment of complete Cyt b dataset contained 47% of variable position, inside the genus *Paracanthocobitis* there were 34% of variable positions and the ingroup of *P. zonalternans* contains 31% of variable positions. In complete IRBP 2 dataset I found 36% of variable positions, inside the genus *Paracanthocobitis* 9,8% of variable positions and inside the ingroup of *P. zonalternans* were 6,2% of variable positions. Complete RAG 1dataset contains 34% of variable positions, in the genus Paracanthocobitis 15% and the ingroup of *P. zonalternans* shows 11% of variable positions. The phylogenies of *Paracanthocobitis zonalternans* were reconstructed on the base of 367 sequences of three different genes. The phylogeny of mitochondrial gene Cyt b includes 88 sequences of *Paracanthocobitis zonalternans*, 13 sequences of other species of *Paracanthocobitis*, 2 sequences of *Acanthocobitis pavonacea* and additionally 18 sequences of various species of the family Nemacheilidae from Sember et al. (2016). For the phylogenetic reconstruction based on the nuclear gene IRBP 2 sequences of the same individuals were used. RAG 1included 92 sequences of *P. zonalternans*, other numbers were same. For rooting of all the trees sequences of *Cobitis taenia* from the family Cobitidae were used.

Bayesian phylogenetic trees were reconstructed from the following datasets: Cyt b, RAG 1, IRBP 2, nuclear combined dataset and datasets consisting of all three genes. The topologies of the final trees based on IRBP 2 and RAG 1 were identical. The final Bayesian trees of Cyt b, nuclear combined and combined datasets are in attachments (Fig. 9, Fig 10, Fig. 11 in attachments). For better overview schematic trees showing only main lineages are presented (see Fig. 4).

4.1.1.1. Phylogeny

Phylogenetic trees of Cyt b and nuclear dataset show seven quite isolated lineages. Those lineages, except several small differences, few conflicts (different position in mtDNA vs. nDNA based phylogeny) or secondary contact areas, show clear geographical distribution. Because of this clear biogeographical pattern, the lineages by the geographical names of their distribution areas were called (Fig. 4).

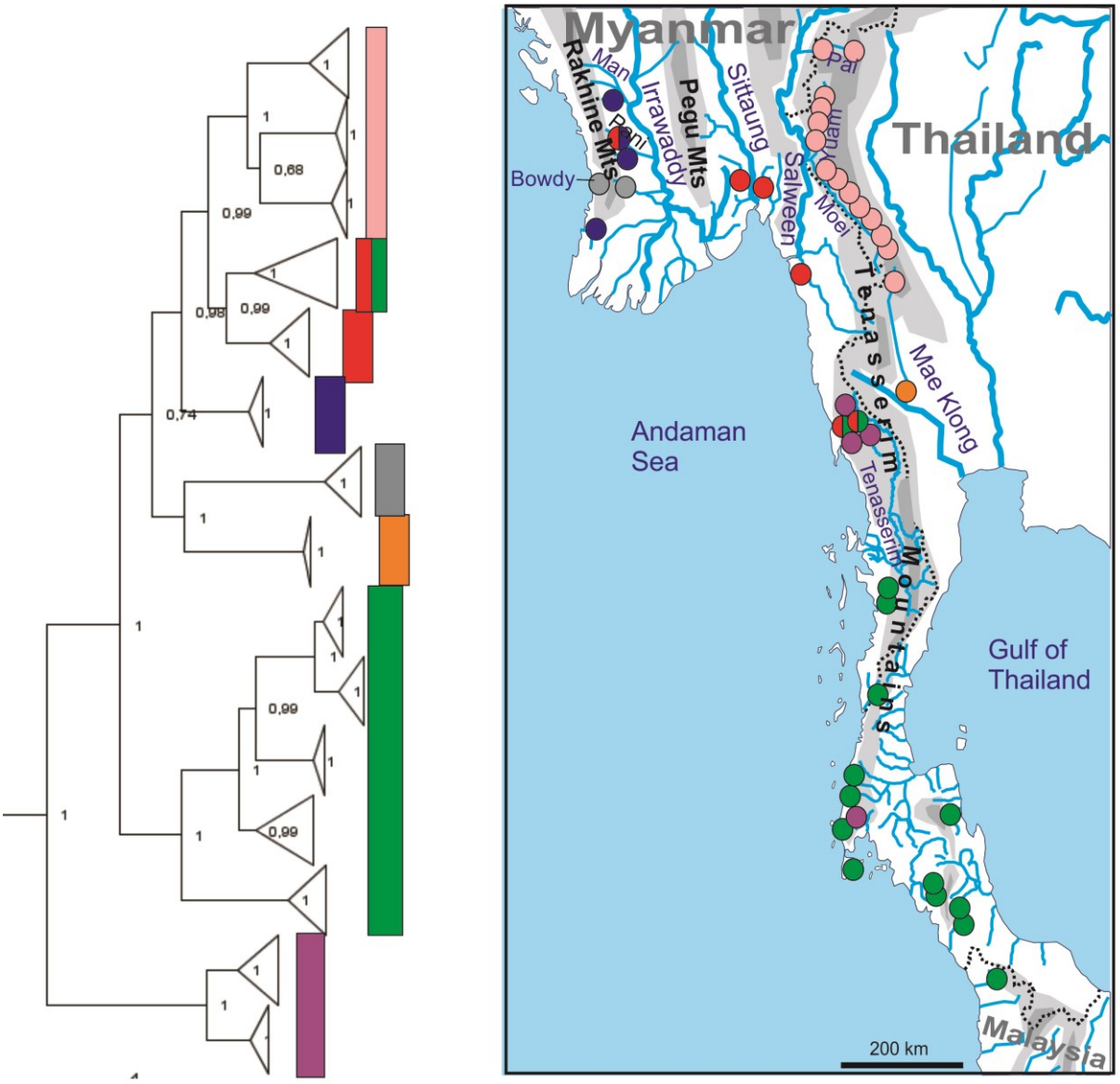


Fig.4 Bayesian tree basing on the combined dataset (Cyt b plus nuclear) of the P. zonalternans group and geographic distribution of the identified lineages. Colours in the tree match colours in the map. The values indicate the posterior probabilities.

Takua Pa lineage (called after the origin of samples) represents a sister clade to all other lineages in all trees. This lineage is formed by two sublineages, one comes from Takua Pa river in Malay peninsula and the other comes from several places in Tenasserim area in south Myanmar.

The **South Thailand lineage** contains samples from many small streams or big rivers from Thai part of Malay peninsula, one sample from north Malaysia, one sample from Phuket island and two samples from south Tenasserim. This lineage is formed by several sublineages. One of them, which is in sister relation to the others, contains the two samples from Tenasserim. It is also quite distant to all other samples within the lineage. The areas of the other four sublineages are exclusive and situated more south. Two out of those four sublineages occur north of Isthmus of Kra and the other two souths of it. Moreover, the north clade as well as the south clade consists of west and east sublineage separated by the Tenasserim mountains in the middle. In nuclear trees, this lineage includes also the samples from north and middle Tenasserim.

Another main clade was named the **Rakhine lineage**, because of the name for the historical Rakhine kingdom, which was occupying the area where the samples were collected. This lineage contains all samples from the river Bowdy Chaung flowing to the see in Rakhine area and also samples from small stream in Rakhine mountains, but flowing into Irrawaddy river.

Fourth lineage is called the **Mae Klong** lineage because except one it contains all samples from Mae Klong river basin. The **Irrawaddy lineage** contains, except few individuals, all samples from Irrawaddy river basin. The **Sittaung lineage** collects all samples from Sittaung river basin. In the Cyt b phylogenetic tree forms this lineage the main lineage also with north and middle Tenasserim sublineages. In the tree of nuclear dataset, the Sittaung and Irrawaddy lineages represent sister clades.

The **Salween lineage** contains all samples from the Salween river basin plus one sample from upper Mae Klong. The Salween consists of three distant sublineages corresponding to the three tributaries of Salween river: Moei, Yuam and Pai.

4.1.1.2. Genetic difference

Genetic differences among the main lineages of *Paracanthocobitis zonalternans* are surprisingly high, especially in the mitochondrial gene Cyt b (Tab.8). In nuclear genes the differences are much smaller, but the numbers are still very high (Tab.8 and also Tab.9). Genetic differences among *P. zonalternans* lineages and other species of *Paracanthocobitis* in Cyt b reach from 13,1% to 13,8%, in nuclear genes the differences range from 2,85% up to 3,5%. Within *P. zonalternans* main lineages divergences range from the difference between 5,7 up to 9,6% in Cyt b, but the Takua Pa lineage shows much higher (11,3-12,2%).

Tab.8 Genetic differences (in percent of sequence divergence) among the lineages in *Paracanthocobitis zonalternans* group in nuclear combined dataset (above diagonal, red) and the mitochondrial Cyt b gene (bellow diagonal, black)

	Salween	Mae Klong	Rakhine	Sittaung	Irrawaddy	South Thailand	Takua Pa	Paracantho-cobitis
Salween		1,33	1,54	0,85	1,12	1,79	1,23	3,26
Mae Klong	8,7		1,60	1,26	1,34	1,64	1,61	3,49
Rakhine	9,2	8,7		1,18	1,51	1,93	1,60	3,32
Sittaung	6,1	8,0	9,1		0,45	1,31	0,89	2,85
Irrawaddy	6,7	7,7	8,6	5,7		1,46	1,19	3,17
South Thailand	9,1	9,2	9,6	8,8	8,4		1,55	3,50
Takua Pa	12,1	12,2	12,0	11,5	11,3	11,6		3,01
Outgroup Paracant.	13,8	13,3	13,6	13,3	13,1	13,2	13,3	

The following table (Tab.9) provides a comparison of genetic differences of the two nuclear genes. From the results, the RAG 1 gene shows slightly higher genetic differences, than IRBP 2 gene.

Tab.9 Percent genetic difference between lineages in *Paracanthocobitis zonalternans* group in nuclear gene IRBP (red) and nuclear gene RAG (black)

	Salween	Mae Klong	Rakhine	Sittaung + Irrawaddy	South Thailand	Takua Pa	Paracantho-cobitis
Salween		1,07	1,89	0,65	1,25	1,07	3,08
Mae Klong	1,54		1,56	0,89	0,77	1,24	3,19
Rakhine	1,25	1,63		1,59	1,72	2,05	3,56
Sittaung + Irrawaddy	1,31	1,65	1,19		0,96	0,94	2,88
South Thailand	2,23	2,37	2,10	1,76		1,16	3,06
Takua Pa	1,46	1,95	1,33	1,24	1,78		3,02
Outgroup Paracant.	4,02	4,32	3,65	3,77	4,41	3,69	

4.1.1.3. Discrepancies between mtDNA and nDNA data

The phylogenetic trees that base on the mitochondrial dataset and the nuclear dataset differ from each other in a number of points:

Sample A4453 from upper Mae Klong

The sample A4453 originates from the uppermost Mae Klong river system at the watershed with the Salween River basin, but in all phylogenetic trees this individual belongs not to the Mae Klong lineage, but to the Salween lineage. However, it also shows in all analyses a certain distance from the other samples from Salween and the position inside the Salween lineage differs between the mtDNA dataset and the nDNA dataset. In the Cyt b tree it belongs to the Moei sublineage and in nuclear genes the sample belongs to the Yuam sublineage.

Sample A6622 from Irrawaddy river basin

Two specimens (A6621, A6622) were collected together at the same locality in the Pani River, a right tributary of the Irrawaddy River. In accordance with their geographic origin, both specimens join the Irrawaddy lineage in the nDNA dataset. However, in the mtDNA dataset, only specimen A6621 joins the Irrawaddy lineage, while specimen A6622 clusters with the Sittaung lineage.

Tenasserim area (visualised in Fig.5)

Most cases of discrepancy between results from nDNA and mtDNA datasets are found in the Tenasserim area. Inside both phylogenetic trees are the samples from Tenasserim area forming three sublineages, which I call North, Middle and South. The South sublineage belongs in all phylogenetic trees to the South Thailand lineage. The position of the North and Middle sublineages varies. In phylogenetic trees of nuclear dataset, they belong to the South Thailand lineage, while in the tree of Cyt b gene they belong to the Sittaung lineage.

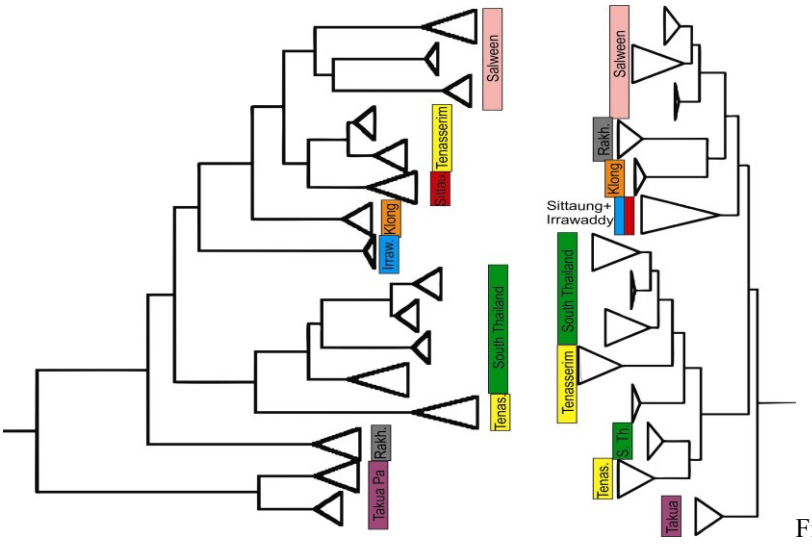


Fig.5 Different position of the three Tenasserim subpopulations (yellow colour) in the mitochondrial (Cyt b) tree and nuclear combined tree.

4.1.2. Dating of events

The results of divergence time analysis are presented in Fig. 15 in Attachments. The times of divergences are also visible in Fig. 6. It shows that the estimated age of genus *Paracanthocobitis* is around 19 Mya. The age of *Paracanthocobitis zonalternans* including the lineage of Takua Pa is around 13 Mya, the age of *P. zonalternans* without the Takua Pa lineage is around 10 Mya.

The results of the dating in most main nodes correlate with the global see water level fluctuations (see the eustatic curve Fig. 5).

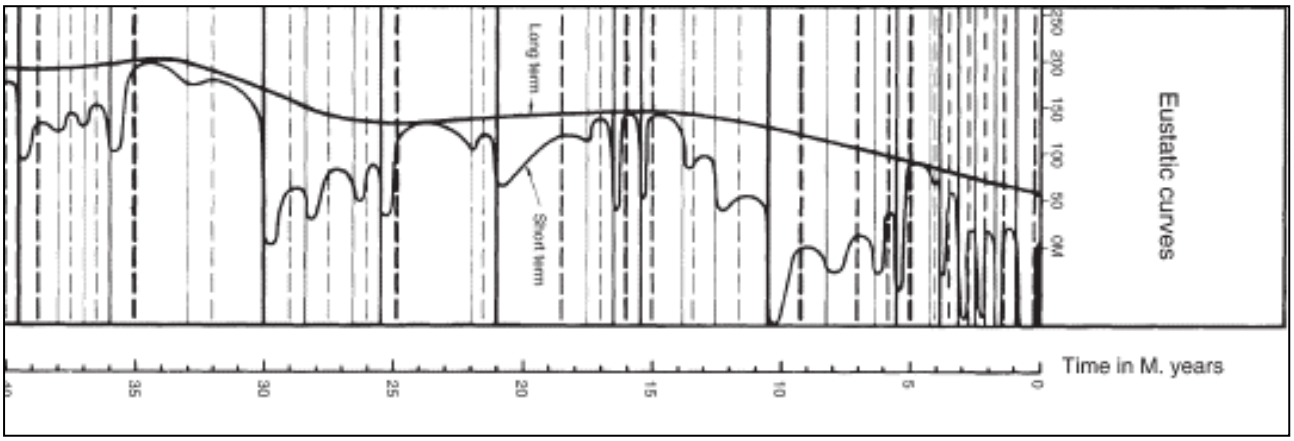


Fig. 5 Eustatic curve from Haq et al., (1987)

4.1.3. Biogeographical reconstruction

Biogeographical reconstruction in RASP was based the mtDNA tree generated in BEAST, which was used for the dating. The figure (Fig.6) shows probabilities of presence of each single area for a given node (represented by pie charts), dispersal and vicariant events The estimated ancestral area of the whole genus *Paracanthocobitis* is India or Bangladesh. The Ancestral area of *P. zonalternans* is most likely in the Tenasserim region (44% probability) The colonization was then from the Tenasserim area northwards (Salween, Irrawaddy, Sittaung and Rakhine) and southwards to South Thailand. The results did not show any extinction events inside the *Paracanthocobitis zonalternans* group. The only extinction event is expected to have happened within the lineage formed by other species of *Paracanthocobitis*. High amount of dispersal as well as vicariant events are estimated to have happened during the history of the *P. zonalternans*.

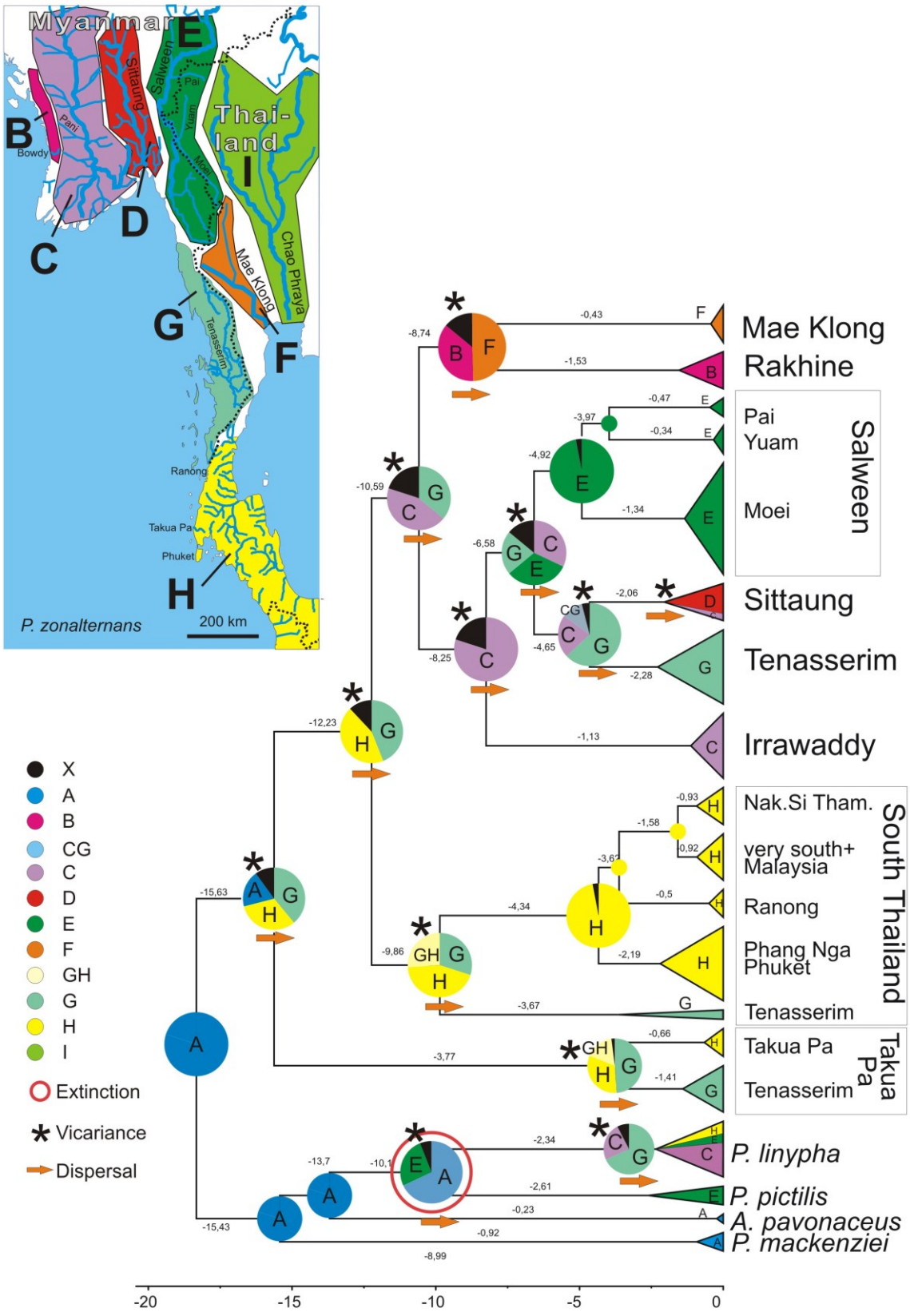


Fig. 6. Ancestral area reconstruction for *P. zonalternans* group performed with Bayesian binary MCMC analyses. Pie charts at the nodes represent the probabilities of each unit area in the ancestral range. X (black) - areas with probability lower than 10%. A (blue, not included in the map) - India/Bangladesh.

4.1.4. Morphology and morphometry

Colour pattern differences (Tab. 10)

At the first sight, there are many differences in colour pattern among the populations of *P. zonalternans*. One of the most visible characters is the dark stripe. This stripe starts close to the gill openings and in some individuals continues through whole body usually up to the end of the caudal peduncle. In other individuals it ends under dorsal fin origin and in some individuals it is completely missing. All these character states are found in all lineages and in all localities with a higher number of available specimens, meaning it shows a strong individual variability.

Next interesting character is associated with the previous one. There are several dark dots along midlateral line, which are overlaid by the dark stripe if present, but still visible in most cases. In some individuals these dots are missing, while others can have from 1 to 10 of them. In the Takua Pa lineage, individuals possess from 9 up to 16 of the dots. The ranges in number of the dots have a small overlap, but in combination with other characters, it could be diagnostic. When comparing all the other populations, this character shows only individual differences.

Another character in coloration is the shape of the bars on the dorsal part of the body. However, this character showed also only individual differences and no diagnostic differences between populations. Shape of the black dot on the base of the caudal fin is also very similar in all populations and only individual differences are present. Amount of dark stripes on dorsal and caudal fin shows, same as the previous characters, only individual differences and nothing useful for distinguishing the populations.

Some fishes have slightly different shade of the background colour on the whole body - while in some populations it is more yellowish, in others it is rather grey. However, this character is also not very significant and moreover may be influenced by preservation method and period of preservation.

Last checked character - the shape of black dot on upper caudal lobe also does not show any differences among lineages and populations, but only on the individual level. This character is in most populations connected with the size of the fishes. Juveniles have it usually small, without any strict shape and sometimes also invisible. In adult fishes the dot is usually more pronounced and already with strict shape. The shape of the dot is very variable among individuals.

Tab 10: Several colour pattern and morphological characters in P. zonalternans group

Colour pattern differences								
	Irrawaddy	Sittaung	Tenasserim	Mae Klong	Salween	Rakhine	South Thailand	Takua Pa
Individuals	18	10	10	8	16	6	19	10
Standard length	24 – 43,6mm	24,4 – 35,2mm	23,5 – 34,5mm	32,8 – 39,8mm	25,7 – 45,7mm	22,4 – 35,7mm	24,9 – 37,8mm	29,3 – 48,5mm
CoM*	Variable	Light Brown	Variable	Dark Brown	Light Brown	Dark Brown	Variable	Variable
CoB*	Dark Yellow	Dark Yellow	Light Yellow	Dark Yellow	Dark Yellow	Variable	Dark Yellow	Light Yellow
LoS*	1/3-1	1	0	2/3-1	1/3-2/3	1	1/3-1	0-1/3
ToS*	Thick	Thick	0	Thick	Thick	Thick	Variable	0/thin
BiS*	12/inv.	Inv.	10-11	0-10	6-12	0-2	7-13	12-16
ADB I*	4-6	5	5-7	7-8	3-7	0-3	4-8/inv.	3-5
ADB II*	4-5	4	5-6	5-6	4-5	0-3	4-6/inv.	2-4
ADB III*	4-6	4	5-7	6-7	4-7	0-4	3-7/inv.	4-9
BiDF*	3-4	3	3-4	4-5	3-4	3-4	2-3/inv.	3-4
BiCF*	5-6	6	4-5	6	4-5	4	3-5/inv.	3-6
SoPF*(ind)	5/18	10/10	2/10	4/8	11/16	2/6	8/19	7/10
BDoCF*	Variable	Variable	Variable	Variable	Variable	Variable	Variable	Variable

Legend for previous tab*: CoM - Colour of marks, CoB - Colour of background, LoS - Length of „lateral line“ stripe, ToS - Thickness of „lateral line“ stripe, BiS - Blotches in midlateral line stripe, ABD I - Amount of blotches before dorsal fin, ABD II - Amount of blotches in the area of dorsal fin, ABD III - Amount of blotches after dorsal fin, BiDF - Bars in dorsal fin, BiCF - Stripes in caudal fin, SoPF - Stripe on pectoral fin (individuals), BDoCF – black dot on caudal fin, L – light, D – dark, Br – brown, Bc – black, G – grey, Y – yellow, inv. - invisible

Morphological characters (Tab. 11)

Some morphological characters were selected after fast inspection of several individuals. Other characters were used from description of new species from Singer and Page (2017).

First checked character was the presence of axillary pelvic lobe. This character shows in most lineages only individual differences, but in Mae Klong lineage was found in all individuals. Inside several main lineages were found some differences between some small subpopulations. Another tested character was length of lateral line and this character show very nice result. In all populations the lateral line is incomplete and it is ending under the origin of dorsal fin, but in Takua Pa lineage the lateral line is complete and it is reaching up to the end of caudal peduncle. Next checked character was position of anus, but this character did not show any single difference. Last two characters are associated with sexual dimorphism, thus for them were comparing only males. Shape and size of suborbital flap did not show many differences, only in Takua Pa lineage is it little bit smaller. Surprisingly we found new type of sexual dimorphism for *P. zonalternans*, which was up to now never mentioned in the literature. This new character is presence of suborbital slid, which is starting nearly on the tip of the snout and continuing under the base of suborbital flap, known from several other species of *Paracanthocobitis* (Singer and Page, 2015). Another character we checked was the presence of tubercles on pectoral fins of males. This character was not found

in Sittaung and Tenasserim lineages, because from these lineages we had only very small males and we assume that they were still too young to have developed tubercles. In Takua Pa lineage there were the tubercles only in males with standard length bigger than 35mm and the tubercles were very small and nearly invisible. In other five lineages were the tubercles very big and usually visible by first look without using binocular.

Tab. 11 Comparison of morphological characters

Morphological characters								
	Irrawaddy	Sittaung	Tenasserim	Mae Klong	Salween	Rakhine	South Thailand	Takua Pa
Individuals	18	10	10	8	16	6	19	10
Standard length	24 – 43,6mm	24,4 – 35,2mm	23,5 – 34,5mm	32,8 – 39,8mm	25,7 – 45,7mm	22,4 – 35,7mm	24,9 – 37,8mm	29,3 – 48,5mm
Pelvic lobe	Variable	Variable	Variable	Yes Big	Variable	Variable	Variable	Variable
Lateral line	Incomplete	Incomplete	Incomplete	Incomplete	Incomplete	Incomplete	Incomplete	Complete
Anus position	Closer to anal fin	Closer to anal fin	Closer to anal fin	Closer to anal fin	Closer to anal fin	Closer to anal fin	Closer to anal fin	Closer to anal fin
Suborbital flap	Big	Big	Big	Big	Big	Big	Big	Small
Tubercles	4-5rows	No *	No *	1-5 rows	3-6 rows	4-6 rows	4-6 rows	3 rows

4.1.5. Morphometric differences

Comparison of age classes don't show us to much useful information, only when we compare real juveniles of standard length maximally 30mm with adults of standard length bigger than 40 mm, then we could find some few trends of allometric growth.

In following table (Tab. 12.) is visible, that pectoral and pelvic fins could grow slowly than whole body. Eye diameter shows nearly no, or only very slow growth through time. Other differences are visible in body depth, head depth (at eyes) and body width, what is usually fixed with growing of gonads inside the body. Tendency of allometric growth is also visible in predorsal and prepelvic length.

Tab 12. Comparing tab of age classes with marked characters with tendency of allometric growth, SL = standard length, yellow colour = different character without overlap, green colour = tendency with small overlap

Differences in age classes		> 40 mm SL		35 - 39 mm SL		30 - 34,9 mm SL		< 30 mm SL	
	Number of individuals	10		34		37		42	
		MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX
mm	Standard length	40,7	42,7	35	39,5	30,9	34,9	25,7	29,4
%SL	Total length	126,8	127,7	126,1	130,4	125,0	132,1	126,1	129,3
%SL	dorsal head length	19,2	20,8	19,5	22,1	19,6	23,6	20,5	22,6
%SL	lateral head length	23,3	23,4	22,6	24,0	22,0	24,6	22,8	25,7
%SL	predorsal length	46,7	47,8	44,5	48,1	44,4	49,5	47,6	50,4
%SL	prepelvic length	52,3	55,0	51,1	53,7	50,2	52,8	50,2	53,3
%SL	preanus length	73,0	73,1	69,7	76,5	71,4	73,5	70,9	74,3
%SL	preanal length	77,9	79,1	76,5	81,3	75,5	79,0	77,2	79,7
%SL	head depth (eye)	11,1	11,9	11,3	13,9	11,2	12,6	11,9	12,8
%SL	head depth (nape)	14,3	15,2	13,7	15,6	13,4	15,6	13,8	15,6
%SL	body depth	20,0	23,7	17,4	23,8	17,4	21,3	16,8	19,4
%SL	depth of caudal peduncle	12,2	12,9	12,0	14,5	11,9	13,5	12,3	13,8
%SL	length of caudal peduncle	15,2	15,6	13,4	15,9	13,2	17,1	14,5	16,0
%SL	snout length	8,4	8,7	7,5	9,6	6,9	9,2	7,8	9,0
%SL	head width (at nares)	8,7	9,1	9,1	9,9	8,9	10,5	8,9	10,3
%SL	maximum head width	14,6	15,0	14,3	15,4	14,0	15,6	14,5	16,0
%SL	body width (dorsal)	13,4	15,2	12,3	17,0	11,6	15,3	10,1	13,1
%SL	body width (anal)	9,1	9,5	7,9	10,0	6,3	9,5	6,0	9,2
%SL	eye diameter	5,2	5,8	5,2	6,3	5,1	6,9	6,1	7,8
%SL	interorbital width	7,9	8,0	7,1	9,1	6,6	9,2	7,8	8,8
%SL	height of dorsal fin	14,3	17,2	13,7	19,1	15,4	19,2	15,3	18,8
%SL	length of upper caudal lobe	24,6	26,0	24,4	29,6	23,6	28,4	24,6	28,3
%SL	length of lower caudal lobe	24,1	26,0	23,8	28,8	23,3	27,9	25,0	27,9
%SL	length of median c. rays	23,0	24,6	22,8	26,9	22,4	26,0	22,8	26,1
%SL	depth of anal fin	19,9	21,4	19,5	23,7	21,1	23,8	20,4	22,3
%SL	length of pelvic fin	18,5	19,0	18,2	20,9	19,4	21,6	19,0	20,7
%SL	length of pectoral fin	19,7	23,1	21,3	24,6	21,0	26,1	22,4	26,1

There were not big differences in the whole dataset of measurements between the Takua Pa lineage and the remaining lineages, but when the dataset was separated for size classes a good amount of differences became visible. In the size class of real adults, there is big difference in prepelvic length, in length of caudal peduncle, in length of most fins and in most characters of width and depth. All of these differences show us that the Takua Pa lineage has whole body more elongated (Tab. 13).

In other size classes, the differences are such smaller, but there is still visible in all of them the big difference in length of caudal peduncle, some small differences in length of fins and several tendencies of small differences in most of the characters connected with depth or width (Tab. 23 and 24 in attachments).

Tab. 13 Comparison of individuals from the biggest size class (> 40mm SL) of *P. zonalternans* and Takua Pa lineage. SL = standard length, yellow colour = different character without overlap, green colour = tendency with small overlap

40 – 48,5 mm SL		P. zonalternans		Takua Pa lineage	
Number of individuals		7		3	
		MIN	MAX	MIN	MAX
mm	Standard length	40,7	43,6	40,4	48,5
%SL	Total length	124,8	127,7	123,7	125,2
%SL	Dorsal head length	19,2	22,8	19,2	21,0
%SL	Lateral head length	22,7	24,1	21,2	23,3
%SL	Predorsal length	46,7	47,9	43,9	46,6
%SL	Prepelvic length	52,3	55,0	48,9	50,5
%SL	Preanus length	72,2	73,2	72,2	75,0
%SL	Preanal length	77,9	79,2	77,9	79,5
%SL	head depth (eye)	11,1	11,9	10,3	11,9
%SL	head depth (nape)	13,9	15,1	12,1	12,9
%SL	body depth	16,0	23,1	14,2	16,1
%SL	depth of caudal peduncle	12,2	12,8	10,5	11,2
%SL	length of caudal peduncle	14,0	15,6	16,7	17,1
%SL	snout length	8,4	9,1	7,8	9,4
%SL	head width (at nares)	8,5	10,2	7,2	8,3
%SL	maximum head width	14,6	15,3	12,6	12,9
%SL	body width (dorsal)	11,4	13,5	10,2	10,9
%SL	body width (anal)	6,5	9,5	5,2	6,8
%SL	eye diameter	5,3	6,6	6,3	6,4
%SL	interorbital width	7,9	9,8	6,6	6,9
%SL	height of dorsal fin	15,4	18,8	17,1	17,5
%SL	length of upper caudal lobe	24,6	26,0	24,0	24,5
%SL	length of lower caudal lobe	24,1	26,0	22,9	24,5
%SL	length of median caudal rays	22,2	24,6	19,4	22,0
%SL	depth of anal fin	20,4	21,7	17,3	19,1
%SL	length of pelvic fin	18,5	22,5	16,9	19,1
%SL	length of pectoral fin	22,1	24,4	20,2	23,0

A sexual dimorphism is present in nearly every single population, but across the whole dataset was not observed.

There is visible difference or usually trend in every population in length of pectoral fin, which is in males bigger. Another visible difference in several populations is in body depth and body width, where females have bigger values, than males, what is most often caused by full bellies of eggs in adult females. Table (Tab. 14) contains example of the two populations with the strongest sexual dimorphism.

Tab. 14 Example of sexual dimorphism in the Takua Pa lineage and in two population of *P. zonalternans* from Takua Pa and Sittaung, characters same as in Tab. 13, yellow colour = different character without overlap, green colour = tendency with small overlap

	Takua Pa lineage				Takua Pa (South Thailand lineage)				Sittaung lineage			
	MALES		FEMALES		MALES		FEMALES		MALES		FEMALES	
	7 individuals		6 individuals		7 individuals		8 individuals		3 individuals		4 individuals	
	MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX
1	38,2	41,2	33,1	48,5	26,9	37,8	28,5	32,1	32,8	35,2	30,3	43,6
2	125,0	126,4	123,0	126,8	129,6	132,0	129,8	131,8	124,1	126,1	124,1	124,8
3	19,9	22,0	19,2	20,5	20,1	22,3	20,3	22,8	19,3	20,1	20,5	21,1
4	22,1	23,5	21,2	22,5	22,2	24,5	22,2	25,6	22,2	22,5	21,9	23,1
5	46,5	49,4	43,9	45,6	45,6	47,3	44,5	48,6	48,0	52,4	47,9	49,8
6	50,0	52,7	48,9	50,8	51,7	57,2	50,0	56,5	53,5	56,7	52,5	55,3
7	73,1	75,0	71,9	73,2	70,1	73,1	69,0	73,7	72,0	79,0	72,2	74,9
8	78,2	79,5	75,7	77,9	77,7	80,8	76,1	80,8	78,7	83,5	77,5	79,5
9	10,7	11,9	10,0	10,7	12,2	13,8	11,5	13,3	10,5	11,0	10,2	11,9
10	12,1	13,6	11,5	12,8	14,6	16,0	13,9	15,1	12,5	13,1	12,5	14,2
11	15,5	16,5	13,6	14,7	19,3	21,1	16,4	19,0	18,0	18,6	15,8	18,3
12	10,3	11,8	10,0	10,8	12,2	13,6	11,6	13,5	12,5	12,6	11,6	12,2
13	16,7	16,8	17,1	18,2	14,1	15,3	12,6	16,4	14,4	15,9	14,2	15,0
14	8,5	9,4	7,5	8,2	7,4	9,1	8,0	10,3	7,6	7,8	6,9	8,7
15	8,2	9,2	7,2	8,5	9,3	10,1	8,4	9,7	8,4	9,8	8,1	9,2
16	12,6	13,9	12,7	13,2	13,9	15,1	14,0	15,8	13,5	13,7	12,5	14,7
17	10,1	11,0	9,4	10,5	9,7	13,8	11,3	15,3	12,5	13,8	11,7	12,5
18	6,4	7,1	5,2	6,0	5,8	10,4	8,2	12,9	8,7	9,7	8,5	8,9
19	6,3	6,7	6,4	7,3	6,1	7,4	6,2	7,4	6,5	6,7	5,3	7,3
20	6,8	7,3	6,4	7,2	6,3	8,2	7,4	10,0	6,9	7,9	7,5	9,2
21	17,1	17,5	16,0	17,7	16,4	19,4	17,8	19,0	-	-	-	-
22	24,0	25,9	23,0	26,0	28,3	32,3	27,1	30,2	25,0	26,7	25,2	26,1
23	23,5	25,7	22,9	25,7	28,0	31,2	28,1	30,2	24,4	27,0	24,8	25,7
24	21,6	22,5	19,3	21,2	23,6	28,6	25,2	27,7	21,6	23,7	19,9	22,2
25	17,8	19,1	16,6	18,2	20,8	23,8	19,3	22,4	20,5	20,7	18,5	20,7
26	17,0	19,1	16,0	17,9	20,4	21,9	19,3	21,1	19,9	20,4	19,3	19,8
27	21,4	23,0	20,1	22,3	25,1	29,0	22,1	25,1	23,6	25,0	21,5	23,9

4.2. *Schistura robertsi* group

4.2.1. Phylogenetic reconstruction

Phylogenetic trees of *S. robertsi* group were reconstructed on the base of three final alignments: 1109 bp of the mitochondrial Cyt b, 970 bp of the nuclear gene IRBP 2 and 2079 bp of the concatenated dataset.

The final Cyt b alignment of the complete dataset contains 49,5% variable positions and the in-group dataset contains 42,5% variable positions. The complete IRBP 2 dataset includes 40% of variable positions, while in the in-group only 13,8% of variable positions were found.

The phylogeny of *S. robertsi* group was reconstructed using 182 sequences of two genes. The phylogenetic analyses of Cyt b, IRBP 2 and combined dataset included 95, 87 and 86 sequences of ingroup, respectively. Same as for *P. zonalternans* group analyses, as outgroup I used 18 sequences of other fishes from the family Nemacheilidae published by Sember et al. (2016) and for rooting of all the trees I used sequences of *Cobitis taenia* from the family Cobitidae. The final Bayesian trees of all datasets are provided in Attachments (Fig. 12, Fig. 13, Fig. 14 in attachments). The Fig.7 provides schematized trees showing only main lineages.

4.2.1.1. Phylogeny

The phylogenetic trees basing on Cyt b and IRBP 2 show ten well isolated lineages with very high statistical supports. The topologies of the Cyt b and IRBP 2 trees, except one small differences, were congruent. Five of the ten main lineages correspond to the known species *S. aurantiaca*, *S. balteata*, *S. cincticauda*, *S. crocotula* and *S.* sp 'Sumo'. The samples of *S. robertsi* form three main lineages, which are in both trees in polyphyletic position. Despite several cases of co-occurrence, the main lineages have disjunctive geographic distribution. For better understanding and to avoid confusion with the terminology in the *P. zonalternans* group, we named the lineages by their species names for lineages corresponding with described species, by specific colour and by numbers in case of *S. robertsi* (Fig.7)

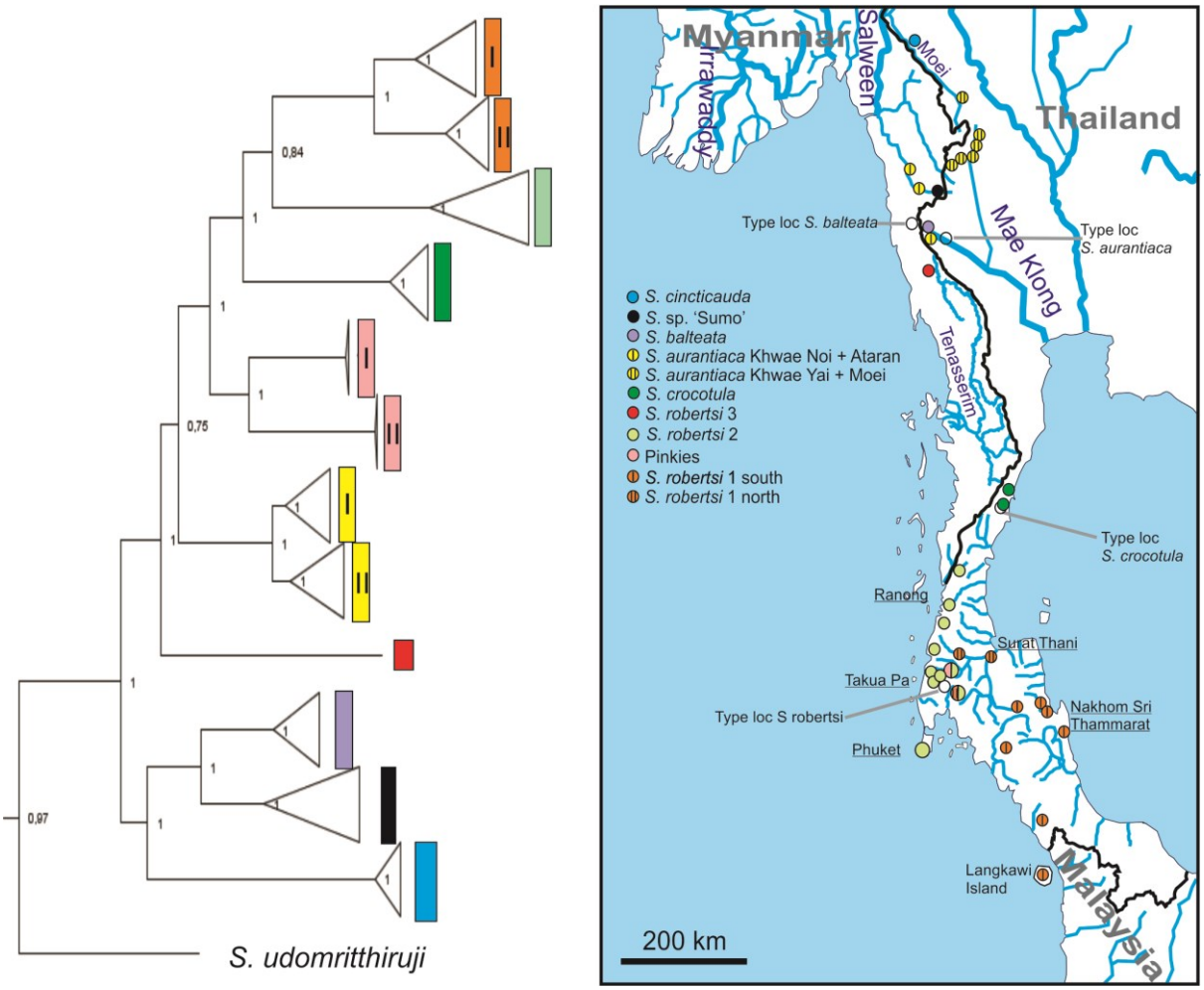


Fig. 7. Bayesian tree based on the combined dataset (Cyt b plus IRBP 2) of the *S. robertsi* group and geographic distribution of the identified lineages. Colours in the tree match colours in the map. The values at the nodes indicate the posterior probabilities.

The first of the main lineage gets here the name **Cincticauda**, because these two individuals are in all characters corresponding with the former description of *Schistura cincticauda* (Kottelat, 1990), they come from the upper Moei River.

The second main lineage gets the name **Balteata** because this clade collects all individuals of *Schistura balteata* (Rendahl, 1948). The Balteata lineage is in Cyt b tree formed by two quite distant sublineages, which are in IRBP 2 tree invisible. The first sublineage collects samples from the Khwae Noi River (upper Mae Klong river system) plus one sample from the ornamental fish trade. The other lineage contains samples from an unknown locality in Myanmar, probably from the Salween River basin.

The third lineage was named **Sumo** because those fishes are known in the aquarium trade as *Schistura* sp. 'Sumo'. The Sumo lineage also forms two distant sublineages in the Cyt b tree, but since all specimens came from the ornamental fish trade the precise sampling

locality for most specimens is unknown. Up to now *S.* sp. 'Sumo' is known only from the Ataran River.

The **Aurantiaca** lineage contains all samples of *Schistura aurantiaca* (Plongsesthee et al., 2011). In the Cyt b tree, five sublineages are visible inside the Aurantiaca clade. The first contains all but one samples from the upper Ataran River, the second contains samples from the Khwae Noi River plus one from the upper Ataran, in the third, there are samples from the Khwae Yai River (upper Mae Klong river system), the fourth and the fifth are formed by samples from different places in the river Mae Klong Noi basin.

The **Crocotula** lineage consist of the samples of *Schistura crocotula* (Plongsesthee et al., 2013). Moreover, some individuals (A10513, A9591) were collected from the Khanan river in Prachuap Khiri Khan province, the type locality of *S. crocotula*. All specimens were collected in several small streams flowing to the Gulf of Thailand on the east side of the Tenasserim Mountains.

The next two lineages **Pinkies 1** and **Pinkies 2** are in both trees sister lineages, but in Cyt b with a very low statistical support, therefore we consider them rather two different lineages. Both lineages come from the Khao Sok National Park in south Thailand, but while the sampling locality of the specimens in Pinkies 1 lineage is known in detail, the sampling locality of the two specimens in lineage Pinkies 2 is known only vaguely.

The last three main lineages were originally considered to represent *S. robertsi* and therefore are named **Robertsi 1**, **Robertsi 2** and **Robertsi 3**. The lineage Robertsi 3 contains only one sample collected from the Tenasserim area. The specimens that make the main lineages Robertsi 1 and Robertsi 2 are morphologically very similar to each other and are all matching the original description of *Schistura robertsi* (Kottelat, 1990). Genetically however, these two lineages are rather distantly related and only in the tree of Cyt b they form monophyletic sister lineages. In IRBP Robertsi 1 forms sister lineage to Crocotula and the topology is (Robertsi 2 (Crocotula + Robertsi 1)). The Robertsi 1 clade is formed by two quite distant sublineages, one containing samples from Ranong province, Phang Nga province and Phuket island, the other samples from Surat Thani and Krabi provinces. Robertsi 2 lineage contains samples from north Malaysia, from Nakhon Si Thammarat and all other more south provinces in Thailand.

In both phylogenetic trees the main lineages cluster into two big sister clades. Into the first clade belong three main lineages: Cincticauda, Balteata and Sumo. The topology inside the first branch is (Cincticauda (Balteata + Sumo)). The second big clade in both trees

contains the other seven main lineages Aurantiaca, Crocotula, Pinkies 1-2, Robertsi 1-3. The topology of the second branch in Cyt b tree is (Robertsi 3, Aurantiaca ((Pinkies 1 + Pinkies 2) (Crocotula (Robertsi 1 + Robertsi 2)))). In IRBP tree the topology is (Robertsi 3 (Aurantiaca ((Pinkies 1 + Pinkies 2) + (Robertsi 2 (Robertsi 1 + Crocotula))))).

4.2.1.2. Genetic difference

For calculation of the genetic differences the same alignments as for BI. For comparison the most closely related species from the outgroup, *Schistura udomritthiruji* (Bohlen and Šlechtová, 2009), was included.

The genetic differences in the *Schistura robertsi* group are comparably high (Tab. 15). In the mitochondrial gene Cyt b the difference reaches from 9,1 % between lineages Pinkies 1 and Pinkies 2 up to 16,7% between lineages Robertsi 3 and Cincticauda. In comparison, the difference between the main lineages of the *S. robertsi* group and *S. udomritthiruji* ranges from 13,5% to 17,1%.

As expected, in the nuclear IRBP 2 the values are much lower than in the mtDNA. The genetic differences among the main lineages range from 1,5% between lineages Pinkies 1 and Pinkies 2 up to 3,4% or 3,6% between Cincticauda and Robertsi 1a or 1b. The differences between the main lineages of the *S. robertsi* group and *S. udomritthiruji* are 4,5% - 5,4%.

Tab 15. Genetic difference (in percent of sequence divergence) among the lineages in *Schistura robertsi* group in the nuclear IRBP 2 (above diagonal, red) and mitochondrial Cyt b (bellow diagonal, black).

	Auran-tiaca 1	Balteata	Cincti-cauda	Croco-tula	Pinkies 1	Pinkies 2	Robertsi 1a	Robertsi 1b	Robertsi 2	Robertsi 3	Sumo	Udomri-tthiruji
Aurantiaca 1		1,82	2,95	2,37	2,32	2,00	2,29	2,54	1,83	2,36	2,65	4,82
Balteata	12,9		2,13	2,32	2,34	1,88	2,57	2,77	1,87	2,37	1,64	4,51
Cincticauda	15,6	14,9		3,14	3,01	2,99	3,44	3,62	2,82	3,14	2,73	4,93
Crocotula	13,4	15,1	15,6		2,67	1,89	2,08	2,26	1,78	2,94	2,88	4,72
Pinkies 1	11,7	13,1	15,0	12,4		1,47	2,87	3,04	2,14	2,67	2,92	5,19
Pinkies 2	11,1	11,6	14,3	12,3	9,1		2,13	2,31	1,83	2,57	2,64	4,51
Robertsi 1a	13,2	15,4	15,6	13,9	13,0	12,1		1,21	1,86	2,89	3,10	5,07
Robertsi 1b	13,6	15,5	15,0	14,9	13,8	12,6	8,5		2,02	3,10	3,36	5,34
Robertsi 2	14,3	14,9	16,3	15,2	13,2	12,7	14,4	14,7		2,41	2,36	4,85
Robertsi 3	14,0	14,4	16,7	15,2	13,6	13,8	15,1	16,5	15,1		2,64	5,35
Sumo	13,4	9,9	13,4	14,9	12,5	11,8	15,0	15,1	14,8	14,8		4,80
Udomritthiruji	15,0	13,5	17,1	15,4	14,7	15,0	16,4	15,8	16,8	16,2	14,3	

4.2.1.3. Cases of co-occurrence

In both phylogenetic trees several places were found, where the main lineages co-occur without any case of gene flow.

Robertsi 1 with Robertsi 2

In the southern tip of the Phang Nga province in south Thailand 12 individuals were collected in a small forest stream and seven of them were sequenced. One sample (A2448) turned out to belong to lineage Robertsi 2, while the remaining six samples belong to the lineage Robertsi 1 (A2445-A2446 and A2449-A2451).

Pinkies 1 with Robertsi 2

In a small stream in the Khao Sok National Park were on two occasions collected individuals from the main lineages Pinkies 1 and Robertsi 2 at the same spot. From the first collection, the samples A 4672 and 4674-4676 represent Pinkies 1 lineage in the phylogenetic trees visible while the samples A4673 and 4680-4681 belong to the lineage Robertsi 2. Next time a group of unidentifiable juveniles was collected, which were reared up in aquarium by Jens Kuehne. From this lot stem the sample A11261 representing lineage Pinkies 1 and sample A11262 representing Robertsi 2 lineage (figure x).

Balteata with Aurantiaca

From the Khwae Noi River in upper Mae Klong river system from one locality come 4 individuals of the Balteata lineage (A4413-A4416) and five individuals from the Aurantiaca lineage (A4417-A4421).

Aurantiaca with Sumo in upper Ataran

Samples A0954 – 0957 of Aurantiaca lineage were collected in the locality of upper Ataran river and on a place situated very close was collected the sample A11005 from Sumo lineage.

4.2.2. Dating of events

The divergence time estimations of S. robertsi group was also based on the Cyt b dataset, to have the results of both species groups comparable.

The dating estimates very high ages of the events and suggests that the main lineages are much older than the expectation was (Fig. 16 in attachments). The time tree shows the age of the whole S. robertsi group around 24 my. The diversification into the two main clades happened 21 Mya. The main lineages inside the two main branches were diversified between 18,5 to 11,5 Mya. The ages of present subpopulation are between 7,5 up to 0,5 my.

4.2.3. Biogeographical reconstruction

Similarly, like in P. zonalternans group, for biogeographical reconstruction I used the Cyt b tree from BEAST, which was used for time estimation.

The pie charts at the nodes (Fig.8) show the possible ancestral areas and their probabilities and the dispersal, vicariant and extinction events are indicated The analyses suggest the ancestral area of whole S. robertsi group to be in the Ataran river (percent of likelihood), which is also a distribution range of the Sumo and Balteata. Also several samples from Aurantiaca main lineages come from this area. The origin of the ancestor for whole group was in India or Bangladesh.

The colonization went from Ataran or Salween river southwards by some river flow change to Mae Klong river basin. From the Mae Klong river basin were colonized all the streams on the west part of Malay peninsula and in similar time also the Tenasserim area.

One extinction event was suggested on the node formed by Aurantiaca and Tenasserim lineage. Vikariance and dispersal events suggested for most of the nodes and on several they must have happened shortly one after the other.

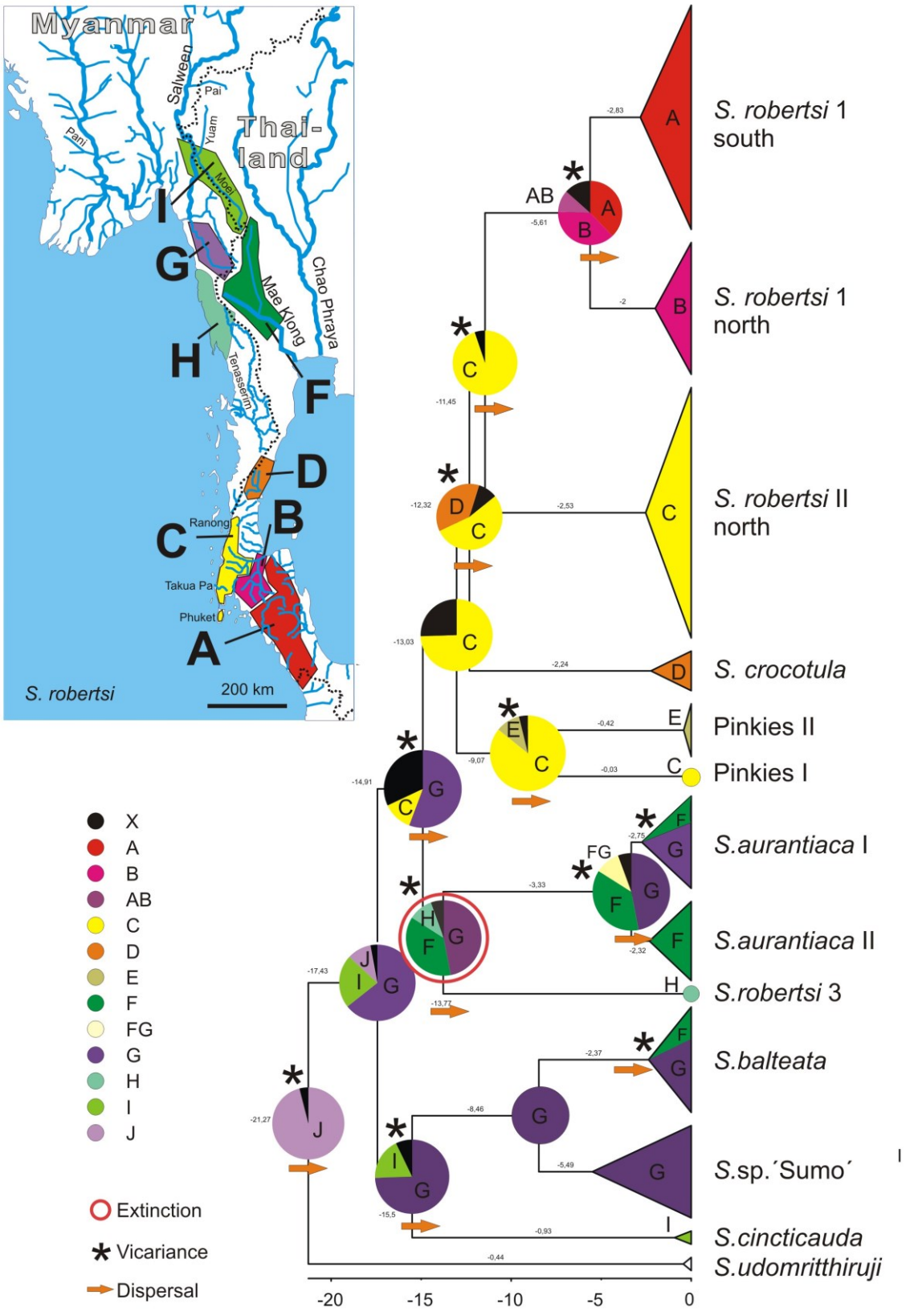


Fig.8. Ancestral area reconstruction for S. robertsi group performed with Bayesian binary MCMC analyses. Pie charts at the nodes represent the probabilities of each unit area in the ancestral range. X (black) - areas with probability lower than 10%. J (light violete, not included in the map) - India/Bangladesh.

Morphological characters

Colour pattern differences

Several informative characters were extracted in the *S. robertsi* group from the dark bars and light interspaces on the body. On the one hand there is a tendency in the *S. robertsi* group to reduce the number of bars and interspaces, their position on the body and their width. The most extreme state is reached in Balteata, which has light brown body (interspace) and only from two up to four thin dark bars under the base of dorsal fin. It is similar in Sumo, which have also only two to four bars, but there they are interrupted by thick and light interspaces. For Sumo, also the colour of the interspaces is specific, which could be from yellow, through orange up to pink, sometimes with completely yellow at the caudal peduncle. It is similar also in Crocotula and Aurantiaca, which both have usually the interspaces orange, but Sumo has the interspaces only under base of dorsal fin. Crocotula has from three up to four orange interspaces on the predorsal part of the body and Aurantiaca has on predorsal usually only one. Pinkies 1 and Pinkies 2 have several pink interspaces, where the number is very variable, but nice character is their position. The interspaces are visible only on the predorsal part of the body and under dorsal fin origin, but they are not visible on caudal peduncle. In all three lineages of Robertsi is this character very variable. From Tenasserim lineage, we have only one individual, that there we could not speak about strict characters and in other two lineages, there could be from zero up to twelve bars. The bars could be light and thin up to dark brown and also such thick. Very similar colour pattern has also Cincticauda, but the bars are much thicker and the number is very stable. They have usually seven and only exceptionally eight.

Another useful character is the black dot on dorsal fin (Tab. 16.), which separate the lineages into two different groups. Aurantiaca, Cincticauda, Crocotula, both Pinkies and all three lineages of Robertsi have it visible, but Balteata with Sumo don't have it. Instead, Balteata and Sumo have two or three dark and thick stripes on the dorsal fin.

Third interesting character is the black bar on caudal fin (Tab. 16), which is in Balteata completely missing. In Sumo is this character variable, because in one lot it is missing, in next lot it is light and thin, and in another lot it is dark and thick. In both Aurantiaca lineages is the bar light and very thin. Thin, light and also usually incomplete is the bar in Pinkies 1. In other lineages is the bar usually dark and thick.

Last useful colour pattern character is the pair of black dots on the lower lip (Tab. 16), which separate the lineages into two groups. One group contain Balteata and Cincticauda,

where the dots are absent and into second group belong all other lineages and the dots are visible.

Morphological differences (all characters are visible in the following Tab 16.)

By anus position, the lineages could be split to three different groups. In one group are three lineages (Cincticauda, Sumo and Tenasserim), which have the anus closer to pelvic fin base, than to anal fin base. Second group contain five lineages (Aurantiaca 1, Aurantiaca 2, Balteata and both lineages of Pinkies), which have the anus more or less in the middle between pelvic and anal fin base. Third group contain three lineages (Robertsi 1, Robertsi 2 and Crocotula), which have the anus much closer to anal fin base, than to pelvic fin base.

By the lateral line length, the lineages from *S. robertsi* group could be split to three such different groups. First group, where belongs all three Robertsi lineages and Crocotula, have the lateral line very short and it is ending much closer to pectoral fin origin, than to dorsal fin origin. Second group, where belongs Aurantiaca 1, Cincticauda and both lineages of Pinkies, have the lateral line longer, but it is ending before the dorsal fin origin. Third group, which contains Aurantiaca 2, Balteata and Sumo, have the lateral line in the comparison with other lineages very long, because it is ending under the dorsal fin origin.

By the axillary lobe, the fishes are separated into two groups. One group have the axillary lobe present, and there belongs three lineages (Balteata, Sumo, Aurantiaca 2). The second group contains other seven lineages, where is the axillary lobe absent.

In Balteata, Sumo, Cincticauda and also Aurantiaca 2 we found some structure, like fold on the base of pectoral fin, which looks like a pectoral lobe.

Length of barbells also shows some differences, but this character could be really tricky, because barbells are very soft structures. There were found two lineages, which has the barbells very short and the length was maximally same like the potential size of eye. In other lineages the barbells are usually two or three times longer than the potential size of eye.

Tab 16. Morphological characters and colour pattern in *S. robertsi* group

Species	Anus position	Bars		Black dot - lower lip	Black line caudal fin	Black dot on dorsal	Lateral line	Pectoral lobe	Pelvic lobe
		Comp.	Incomp.						
Aurantiaca 1	„In the middle“	2-9	2-8	Yes	Light, thin	Yes	„Middle“	No	No
Aurantiaca 2	„In the middle“	2-3	1-4	Yes	Light, thin	Yes	„Long“	Yes	Yes
Balteata	„In the middle“	0-2	1-4	No	No	No	„Long“	Yes	Yes
Cincticauda	„Closer to pelvic fin“	7	0-1	No	Dark, thick	Yes	„Middle“	Yes *	No
Crocotula	„Closer to anal fin“	3-5	2-6	Yes	Dark, thick	Yes	„Short“	No	No
Pinkies 1	„In the middle“	5-7	1-3	Yes	Dark, thick	Yes	„Middle“	No	No
Pinkies 2	„In the middle“	2-12	2-4	Yes	Light, thin incomplete	Yes	„Middle“	No	No
Robertsi 1 “North-west”	„Closer to anal fin“	2-12	0-9	Yes	Dark, thick	Yes	„Short“	No	No
Robertsi 2 “South-east”	„Closer to anal fin“	0-11	0-10	Yes	Dark, thick	Yes	„Short“	No	No
Sumo	„Closer to pelvic fin“	2-4	0-5	Yes	Variable	No	„Long“	Yes **	Yes
Robertsi 3 “Tenasserim”	„Closer to pelvic fin“	7	1	Yes	Dark, thick	Yes	„Short“	No	No

Differences in fin ray counts

For fin ray counts three different methods could be used. In pectoral, pelvic, anal and also dorsal is every time the first fin ray unbranched and this one is usually counted. Another several fin rays are branched and the number could be variable. There is also one last unbranched ray, but this ray is on the base connected with the previous one and someone is counting it, someone not and someone count it like one half. For our comparison table (Tab. 17) only the numbers of branched fin rays were used, because this is the variable character. Counting of fin rays in caudal fin is simpler, because usually only the branched rays are counted.

Tab 17. The fin ray counts of lineages in *S. robertsi* group

Species	Individuals	Pectoral	Pelvic	Anal	Dorsal	Caudal
Aurantiaca 1	34	8	6	5	7	9+8
Aurantiaca 2	6	9-10	6	5	7	9+8
Balteata	9	9-10	6	5	7	8+8
Cincticauda	2	8	5	5	6	9+8
Crocotula	13	8	6	5	7	8+8
Pinkies 1	2	8	6	5	7	8+8
Pinkies 2	10	8	6	5	7	9+8
Robertsi 1	53	8	6	5	7	„9+8”*
Robertsi 2	46	7	6	5	7	„7+8“**
Sumo	11	8	6	5	7	8+8
Robertsi 3	1	7	6	5	7	8+7
Outgroup Udomritthiruji	6	9	6	5	8	9+8

The number of anal fin rays is not variable in this group, but in other fins are some differences visible. In pelvic fin all fishes, also with most closely related outgroup *Schistura udomritthiruji*, has 6 branched fin rays and only Cincticauda has 5. This character is nicely fitting with original description, where is mentioned, that S. cincticauda is different from all other Schistura species by having only 5 branched pelvic fin rays (Kottelat, 1990). Small differences are also in dorsal fin ray counts. Schistura udomritthiruji has 8 branched fin rays in dorsal fin, S. cincticauda has 6 and all other lineages have 7.

Pectoral fin shows little bit bigger variability. Balteata and Aurantiaca 2 have in pectoral fin 9 or 10 branched fin rays. *S. udomritthiruji* has 9 branched rays. Aurantiaca 1, Cincticauda, Crocotula, both pinkies, Robertsi 1 and Sumo has 8. Robertsi 2 and also the single individual from Tenasserim have 7 branched rays there.

Caudal fin ray counts show several interesting differences also. There are six lineages (Aurantiaca 1, Aurantiaca 2, Cincticauda, Pinkies 2, *S. udomritthiruji* and also 42* individuals of Robertsi 1), which has 9 branched rays in upper caudal lobe and 8 in lower caudal lobe. Balteata, Crocotula, Pinkies 2, Sumo and 10* individuals from Robertsi 1 has in both caudal lobes 8 branched fin rays. Single individual from Tenasserim lineage has in upper caudal lobe 8 and in lower 7 branched fin rays. Robertsi 2 has the number of branched caudal fin rays very variable, they could have 7+8 (18** individuals), 7+7 (14** individuals) and 7+6 (6** individuals).

5. Discussion

5.1. *Paracanthocobitis zonalternans*

5.1.1. Phylogenetic reconstruction

Evolutionary units within the *P. zonalternans* group

All analyses identified seven major lineages within the *P. zonalternans* group. This robustness of their appearance leads to the conclusion that these seven lineages should be considered as evolutionary units of *P. zonalternans*. In general, the major lineages have disjunct distribution areas and only few cases of secondary contact were detected.

Cases of secondary contact

Cases of secondary contact can be identified by presence of genetic markers of different main lineages in the same river or best at the same locality. Within the present study two cases of mitochondrial introgression were observed and these cases are best explained by secondary contact.

The phylogenetic tree of the mitochondrial gene Cyt b shows one sample (A6622) from Irrawaddy river basin inside the Sittaung lineage. In both nuclear markers this sample joins the Irrawaddy lineage. Sample A6622 was collected in one locality with sample A6621, which in mtDNA as well as in nDNA datasets belongs to the Irrawaddy lineage. Both samples share the same haplotype in both nuclear genes, without any single different nucleotide. This observation indicates that in the sampling locality exist mitochondrial genes that evolved in the Sittaung River basin. Most likely, at some time in the past female specimens of *P. zonalternans* from the Sittaung basin entered the Irrawaddy basin and bred with the Irrawaddy population. After numerous backcrossing with the Irrawaddy population the nuclear DNA of the Sittaung lineage was replaced completely, but the maternally inherited mitochondria were being still present (mitochondrial introgression). Some genetic exchange between Irrawaddy and Sittaung rivers are not unlikely, since both are neighbouring rivers. They might be well separated in the upper and in the middle part, but in the lower part they are together forming one huge delta, which is nearly flat and the river arms are there changing continually. These rivers were probably connected several times in history. From the time tree this event appears to be around 1,5-2 million years old, because the sample A6622 is quite distant from the remaining Sittaung samples. However, possibility exists that the event could be also considerably younger if similar mitochondrial haplotypes still exist in unsampled parts of the Sittaung River basin.

The second case of secondary contact indicated by mitochondrial introgression is found in the Tenasserim region. Both phylogenetic trees show three different sublineages formed by samples from whole Tenasserim area and these sublineages could be named North, Middle and South Tenasserim sublineages. In phylogenetic tree of nuclear dataset are all three lineages form one main lineage together with all sublineages from the South Thailand lineage, but in the tree of mitochondrial dataset is only the South sublineage with the South Thailand lineage and the other two form another main lineage with samples from Sittaung river basin. It looks like, that the North and Middle subpopulations share their mitochondrial genome with the Sittaung lineages and the nuclear genome with the South Thailand lineage. This indicates that the Tenasserim area is a secondary contact area, in which the South Thailand lineage and the Sittaung lineage met and interbred.

In both cases of secondary contact, the lineages interbred to the extent that the nuclear genetic markers of one lineage got lost from our dataset. It demonstrates that Irrawaddy and Sittaung lineage on the one hand and the Sittaung and North Thailand lineage on the other hand are not reproductively isolated when coming in contact. Therefore, I consider these evolutionary units as conspecific.

Takua Pa lineage as separate species

A different situation is found with the Takua Pa and South Thailand lineages. In all sampling localities, the Takua Pa lineage occur sympatrically with the South Thailand lineage to the degree that specimens of both lineages can be collected in the same movement of a hand net. In opposite to the before mentioned cases of secondary contact, no sign of hybridisation between the Takua Pa and South Thailand lineages has been observed. Two related populations that occur sympatrically but without hybridisation are reproductively isolated and must be considered two distinct species. This assumption will be discussed further when dealing with the morphological data.

5.1.2. Biogeographical reconstruction + Dating of events

The big bang: origin of the *P. zonalternans* group

The phylogenetic reconstructions demonstrate that *P. zonalternans* is a member of the *Paracanthocobitis* genus and that this genus has its roots in the Indian fauna. The age of the diversification of *P. zonalternans* from the other species of *Paracanthocobitis* is by the results of the time estimation analyses around 18,3 million years The Biogeographical analyses revealed, that the ancestral area of the species *P. zonalternans* should have been in the Tenasserim region. Such result matches the fact that one of the sister species of *P.*

zonalternans, *P. pictilis*, is an endemic of the Ataran River, which flows through the extreme north of the Tenasserim region. It is further supported by the fact that the oldest lineage of the *P. zonalternans* group, the Takua Pa lineage, lives in Tenasserim. Therefore, the beginning of the *P. zonalternans* group should have been about 18.3 mya in the Tenasserim region.

The first split: Takua Pa and *P. zonalternans*

The first split in the present phylogenetic reconstructions separates the Takua Pa lineage from *P. zonalternans* and appeared about 15.6 mya. This event could not have been linked with any geologic or climatic event, therefore the driving force for this split remain unknown. The fact that the two lineages life in sympatry could point on ecological reasons, but the possibility of a small-scaled geographic separation with subsequent secondary contact cannot be excluded.

12.2 mya: northern and southern group

The next genealogic events separate *P. zonalternans* into a southern group (Tenasserim and South Thailand) and a northern group (all remaining lineages). This event coincidence with a significant drop in the global sea level from 120 m above present level to 40 m above present (Haq et all., 1987). Such 80 m drop of sea level should have enlarged the available land in the Tenasserim region, which can have caused range expansions in the fauna of the region. Since Tenasserim is long in north-south direction, but very narrow in east-west direction, it seems plausible that the ancestral population of *P. zonalternans* northwards. For unknown reasons, this northwards expansion was not joined by the Takua Pa lineage (since both species co-occur it is unlikely that the Takua Pa lineage went extinct later on in the northern realm).

10,5 to 9,5 mya: going south and going north, seriously

Between about 10,5 – 9,5 mya the global sea level dropped for the first time in earth's history below the present day level (Haq et all., 1987), and with > -80 m it was one of the deepest sea level stands ever

since. This drop did not only connect Mainland Southeast Asia, the Malay Peninsula and the Great Sunda Islands Sumatra, Java and Borneo into one large land mass, it also retreated the coastline along the Andaman Sea for several kilometres. The short coastal rivers that run from the Tenasserim Mountains prolonged and potentially joined, allowing range expansions for aquatic fauna.

At this period, in the phylogeny of *P. zonalternans* in the southern group a split between the Tenasserim populations and the populations in southern Thailand could be seen. We conclude that this group was enabled by the low sea level to expand its range southwards to (nowadays) southern Thailand. By doing so, the fishes surrounded a major biogeographic barrier, the Isthmus of Kra.

At the same time, in the northern group a split into a Mae Klong plus Rakhine lineage and a lineage collecting the lineages in Central Myanmar could be seen. In much the same way like the southern group expanded southwards, the northern group utilised the low sea level and the retreated coastline to expand northwards, one branch followed the new coastline until it reached behind the Rakhine Mountains (the founder of the Rakhine lineage), one branch colonised Central Myanmar (ancestor of the Irrawaddy, Sittaung and Salween lineages) and one branch managed to cross the Tenasserim Mountains, most likely via stream capture at the Three Pagoda Pass, one of the very few passes in these mountains, and founded the Mae Klong lineage. It is likely, but not necessary, that *P. zonalternans* after these range expansions already occurred in its total present day distribution area.

7,5 – 6,5 and 5,3 – 4 mya: the big floods

The Cenocoic fluctuations in the global sea level were not restricted to low sea periods, but there have been also periods with increased sea level. The longest and most prominent happened in the period 5,3 – 4 mya, but also 7,5 – 6,5 mya and 3,5 – 3 mya such increases of the sea level occurred. During these periods, most of the Malay Peninsula is best imagined as a group of islands (the peaks of the present day mountains) and the lowland areas e.g. in Central Thailand and Central Myanmar should have been flooded at least 200 km landwards (Woodruff, 2003).

During these periods, the populations of *P. zonalternans* especially in Central Myanmar and southern Thailand must have been strongly fragmented and restricted to a number of inland refuge areas. These periods of fragmentation should have shaped out the present day lineages, which after the retreat of the sea water re-colonised the certain river basins. It was this period when in Central Myanmar the Irrawaddy, Sittaung and Salween lineages differentiated and when in the South Thailand lineage, the populations north and south of the Surat Thani – Krabi line and north and south of the Kangar – Pattani line developed to differences observed in our phylogenetic reconstructions.

4 mya until today: The story goes on - Secondary contacts

During the last 4 million years, also periods with lowered sea level existed, but it seems that faunistic exchanges during these periods did not lead to major changes in the picture. Most likely range expansions were hampered by the fact that the potential range was already inhabited by *P. zonalternans*. However, these periods should be responsible for the observed cases of secondary contact.

Faunistic exchange between Mae Klong and Moei rivers

The Mae Klong River is situated on the western side of the Tenasserim Mountains and the Salween River is flowing on the eastern side. In general, the Tenasserim Mountains are known as a geographical barrier between the Indian (including *Paracanthocobitis*) and the Indochinese fresh water fauna (Zakaria-Ismail, 1994), and there are only few freshwater fish species living in both of these rivers. One of these few is *P. zonalternans*. As mentioned above, *P. zonalternans* entered the Mae Klong basin from the Tenasserim region about 10 mya. This colonisation event founded the populations in the middle Mae Klong basin.

However, the present analyses include sample A4453 collected in the uppermost Mae Klong river, very close to the watershed with the Moei River, a tributary of the Salween River. Indeed, genetically sample A4453 belongs to the Salween lineage. I was trying to find, how this sample comes from Salween river basin to Mae Klong river basin and there are two hypotheses. By the first one, I expect, that this sample was transported from Moei river to Mae Klong river by some fisherman or fish collectors by mistake. This hypothesis will be useful only when I will count with some heavy genetic drift, because the time estimations show the age of this event minimally 0,5 my old. Next hypothesis is more biogeographical and by this second hypothesis, I expect, this individual or another group of fishes before several generations comes from Moei River into the Mae Klong River by flow exchange of some small mountain stream. By checking the maps, I found several valleys, where are the spring of streams, flowing to Moei river, nearly in contact with the springs of streams that flow to the Mae Klong River. This sample A4453 was found very close to one of those places, where the streams are only 2-3 kilometres far from each other and the horizontal difference is only around 20 meters.

5.1.3. Morphology and morphometry

Comparison of main lineages

The data of the analyses of the morphological and morphometrical characters did not show any useful difference between the main lineages. Most of the morphological and morphometrical characters are showing only individual differences or differences between age (size) classes and there is no possibility for using them like characters for distinguishing species, as it was made by Singer and Page (2017). The very small differences in the characters are most likely caused by quite big population sizes and not enough long isolation between the lineages. Also the areas of the main lineages are ecologically quite similar, so the

different lineages undergo a similar selection what leads to a stabilisation of the phenotype and balances against effects of genetic drift. Comparison of different age (or size) classes show a strong signal of allometric growth, then for next comparing of the main lineages of P. zonalternans, and for comparison between males and females, were used only fishes of similar size. The results of the comparison show clear differences between the Takua Pa lineage and the remaining main lineages in every selected age class. Most of the differences in the characters were caused by elongation of the whole body. Why are the fishes from Takua Pa lineage elongated like this, I am not able to say, but maybe it could be caused by some little bit different type of swimming performance.

Sexual dimorphism

In most of the populations males and females show different length or minimally trend of different length of pectoral fins, a character that is commonly observed in Nemacheilidae (Kottelat, 1990) as well as in the related family Cobitidae (Šlechtová et al, 2008)

In the biggest males through most populations I found also structure very similar to suborbital slid mentioned in other species of Paracanthocobitis (Singer and Page, 2015).

The two populations (Takua Pa and Tenasserim) of Takua Pa lineage have nicely visible sexual dimorphism, but there is the trouble, that all individuals from Takua Pa are females and all individuals from Tenasserim are males. This makes all the nice characters little bit useless, because I could not say which differences are real sexual dimorphism, which are population differences and which are mixed sexual dimorphism characters with population differences.

5.1.4. Taxonomic implications

Paracanthocobitis zonalternans was recently split by Singer and Page (2017) into three different species named Paracanthocobitis zonalternans, Paracanthocobitis nigrolineata and Paracanthocobitis phuketensis. Under name P. zonalternans they consider fishes from the Salween river, since the neotype locality is in the Moei River (Kottelat 1990). This conclusion matches the genetic results of the present study, because in our phylogenetic trees all samples from Salween river basin form one isolated lineage. The name P. phuketensis was used by Singer & Page for all South Thailand populations, what is also fitting with my results, because all samples, except few individuals from Takua Pa, from South Thailand are in all phylogenetic trees forming one isolated lineage also.

Conflicts starts when I compare the species P. nigrolineata, where belongs according to Singer and Page the populations from the Mae Klong, Irrawaddy and Sittaung rivers.

Samples from those three river basins are in the trees forming three different isolated lineages which are polyphyletic.

Singer and Page used morphological characters for the delimitation of P. phuketensis and P. nigrolineata, not genetic data, what is maybe the reason for the observed difference. However, in the present study the morphologic characters of X specimens from these lineages were analysed and it turned out that the diagnostic characters postulated by Singer & Page are not diagnostic. P. phuketensis was diagnosed by the absence of an axillary pelvic lobe (vs. present in all other lineages). Present data show that only about half of the specimens of the South Thailand miss the axillary pelvic lobe completely, while it is present in the other specimens. The absence of the lobe is also not diagnostic for any of the sublineages within the South Thailand lineage, but usually specimens with a normally developed, a very small and an absent axillary lobe occur in nearly every locality. Moreover, also in most localities of the Rakhine, Mae Klong and Salween lineages coexist specimens with and without the lobe, and in Irrawaddy and Salween lineage the lobe is very small in the majority of specimens. Altogether, the presence or absence of the axillary pelvic lobe is not a diagnostic character for any group of specimens within P. zonalternans. The same turned out for the characters that were used by Singer & Page to diagnose P. nigrolineata: presence of a dark stripe along midlateral line and of a short black stripe from behind the operculum in midlateral level ventrally to the base of the pectoral fin. These character states are not present in about half of all analysed specimens from Irrawaddy, Sittaung and Mae Klong basins but present in about one third of specimens from the remaining lineages. Therefore, these characters are not suited to diagnose groups of specimens within P. zonalternans.

Additionally, the observations of hybridisation in cases of secondary contact suggest that the lineages within P. zonalternans are not reproductively isolated. I therefore consider all identified lineages except the Takua Pa lineage to represent evolutionary units within one widely distributed species, P. zonalternans.

The situation is different for the Takua Pa lineage. It forms the sister lineage to P. zonalternans in all genetic analyses and differs from it by more than 11% genetic difference. In all known populations it co-occurs with the South Thailand lineage without any trace of hybridisation, indicating reproductive isolation in sympatry.

This lineage is also distinguished from all other lineages by some morphological and morphometric characters. These fishes have longer lateral line, which is nearly complete. In the Takua Pa lineage continue the lateral line up to the end of caudal peduncle and in other P. zonalternans lineages is the lateral line ending approximately under the hind part of dorsal fin.

Other character visible already by first look is completely elongated body of these fishes from the Takua Pa lineage. There are also some morphometrically characters different in this lineage like longer and thinner head and longer caudal peduncle. This lineage is a different species than P. zonalternans and presently undescribed.

5.2. *Schistura robertsi* group

5.2.1. Phylogenetic reconstruction

Schistura robertsi group is in all phylogenetic trees forming clearly supported monophyletic lineage. From all phylogenetic trees is also visible the polyphyletic position of the genus named *Schistura*, what was already mentioned by Sember et al., (2016). The genus *Schistura* appears to be a variable collection of morphological types distributed through the whole family Nemacheilidae.

The *S. robertsi* group contains in all phylogenetic trees ten highly supported lineages. Because of the very high support in mitochondrial as well as nuclear dataset these lineages can be referred to as evolutionary units.

The topology is in all phylogenetic trees of *S. robertsi* group consistent, only the Robertsi 3 lineage from Tenasserim had a little bit unstable position. Most likely this was caused by the fact that only a single specimen was available for analysis. The position will come clearer when several more samples can be added to the analyses.

Only one conflict was found between mitochondrial and nuclear tree. In the mitochondrial tree the Crocotula lineage is in sister position to the Robertsi 1 and Robertsi 2, but in nuclear tree the Crocotula lineage is in sister position only to Robertsi 1 and together they are sister to Robertsi 2 lineage. Since these lineages also do not have the highest statistic support I will rather consider them three sister lineages. By my hypothesis, those lineages developed by some very fast colonization followed by very fast isolation.

All phylogenetic trees show quite long and deeply isolated branches, with only very small and young substructure. There is the explanation that those long branches were caused by heavy genetic drift. The fishes are ecologically specialized for living in very small forest streams. Because of the seasonal weather in this area, the streams are often dried and the fishes must stay alive only in several small hidden puddles inside some depressions usually covered by leaves. From the fish exporter Jens Kuehne I also know, that those fishes are

strongly territorial, what could usually decrease the population sizes also. By my opinion the expected genetic drift was caused by heavy bottleneck effect in the long dry periods.

5.2.2. Biogeographical reconstruction + dating of events

Origin of *S. robertsi* group

The *Schistura robertsi* group separated from other related species about 21,5-22,5 mya, what means, that this group was differentiated from other fishes of the Indian clade of Nemacheilidae quite early. Because of the absence of any single fish from *Schistura robertsi* group, or any other closely related species in Sittaung or Irrawaddy river basins, the ancestral area of this group was somewhere in northern Tenasserim region.

The biogeographical reconstruction placed the most possible ancestral area of *Schistura robertsi* group into Ataran river, but with quite high possibility of Moei river also. Those places are occupied by the oldest lineages like Balteata or Cincticauda.

Biogeographic events

Until about 11 mya the sea level was such high that the ancestral area of the *S. robertsi* group and the first genealogic events happened in a landscape that in best imagined as islands in the sea (Fig. X). In such scenario, a drop of sea level from +130 m to +70 m (as occurred about 16,5 mya and again 15,5 mya) could connect two islands, by that lead to a range extension and during the following period (of up to 3 million years) of high sea level separate the populations until they appear as main lineages in a recent phylogeny. In accordance with this assumption, the *S. robertsi* group separated into two main branches16,5-17,5 mya. Both branches of the *S. robertsi* group show another split around 15 mya that could have been caused by such drop of sea water level. However, reconstructions of such old events that took place in such small area remain suggestive, but the sheer coincidence between genealogic event and change in sea level is inviting.

About 13 mya an important genealogic event took place, again after a 50 m drop of global sea level: a massive range expansion southward by the common ancestor of the both Pinky lineages, Crocotula and Robertsi 1 + 2. It is difficult to reconstruct the pathway of colonisation, but due to the facts that 1) the related lineage Aurantiaca occurs in the Mae Klong basin (Gulf of Thailand drainage) and 2) the northernmost place of occurrence of the new southern group (Crocotula lineage) lays in the Gulf of Thailand drainage and 3) no members of the S. robertsi group are known from the Andaman Sea drainage in Tenasserim

south of Dawei a pathway on the eastern (Gulf of Thailand) slopes of the Tenasserim Mountains seems more likely than a pathway on the western (Andaman Sea) slopes.

The southern group radiated into the present day lineages, but the exact causes for genealogic events remain unknown. Moreover, it can be assumed that the strong impact of genetic drift on the small isolated populations in the southern group added to mislead the age calculations.

Faunistic exchange between upper Mae Klong and upper Moei river

One sample A9854 of Aurantiaca lineage was collected in the upper Moei river (Salween river basin) and the phylogenetic analyses show that this sample is forming one sublineage with all samples from Khwae Yai river (upper Mae Klong river). By my hypothesis this fish (or some ancestor) come to the upper Moei River, by some river flow exchange, usually in same time like the sample A4453 in *P. zonalternans*, only by opposite direction. Those places where were collected the Aurantiaca samples are corresponding with the places, where was found the possible faunistic exchange from Moei river to Mae Klong river in *P. zonalternans* group.

5.2.3. Morphology

Lineages in the *S. robertsi* group show several nice differences in colour pattern, morphology and meristic. There is not one single character, which could be used for distinguishing all lineages, but the combination of them is for identification of every lineage enough.

Variability in colour pattern

Main lineages of *S. robertsi* group are very interesting group of fishes mainly because of their colour pattern. They are typical by their really expressive colour of the lighter interspaces and also sometimes by the expressive colour of the darker bars visible in the lineage Sumo. The colour of interspaces could be from yellow up to pink or violet and could be connected with the territoriality and quite small population sizes of the fishes. For the lineages Robertsi 1-3 and Cincticauda is the colour of interspaces yellow, what is for the species *Schistura robertsi* and *Schistura cincticauda* specific (Kottelat, 1990). The lineages Crocotula and Aurantiaca have the interspaces orange, what was also mentioned in the original descriptions. For the lineages Pinkies 1 and 2 is specific pink colour of interspaces. Sumo has first two or three interspaces usually orange and the others yellow with nearly violet bars.

Balteata have the colour different to all other lineages, for them is specific the combination of light brown body (interspace) only with two or three dark thin bars. Into the

species *S. balteata* is quite often given the Sumo lineage, but already by the colour pattern, those fishes are completely different.

Very specific could be in the main lineages the shape and position of bars. For Crocotula and Aurantiaca is specific, that the first two orange interspaces are thicker, then others, what was already mentioned in the descriptions of *S. crocotula* (Plongsesthee et al., 2013) and *S. aurantiaca* (Plongsesthee et al., 2011). In Pinkies 1 and 2 are the interspaces only before and under dorsal fin origin, but they are missing in the caudal peduncle. For *S. balteata* is specific, that the dark bars are only under dorsal fin origin (Rendahl, 1948), what is also specific for Sumo, but there are under dorsal fin origin visible orange interspaces and not dark bars.

Another useful character is presence or absence of the black dots on lower lip. The black dots are absent in Balteata and Cincticauda and present in all other lineages. This character is very good for distinguishing Sumo from Balteata, because is stricter, that the differences in colour of bars or interspaces. Quite specific for Balteata and Sumo is the absence of the black dot on the front part of dorsal fin, which is in all other lineages present. They have on the dorsal fin instead of the black dot three quite thick stripes.

Variability in morphological characters

Lateral line length in *S. robertsi* group could be also very good character. Without strict measurements, could not be differentiate all lineages, but the lineages could be split to three such different groups without overlapping.

Anus position is one of the characters used in original description of most species in this group (Kottelat, 1990). The anus position is inside the group very variable, but same as in previous character it must be used with the combination of other ones, because it helps only for splitting the lineages into three groups. For Aurantiaca, Balteata and Pinkies is specific the position of anus more or less in the middle between pelvic and anal fin, what was for *S. balteata* and *S. aurantiaca* mentioned also in original descriptions. In Cincticauda. Sumo and Robertsi 3 is the anus closer to pelvic fin base, what is for *S. cincticauda* specific position (Kottelat, 1990). Robertsi 1,2 and Crocotula have the anus closer to anal fin base, what was also mentioned in original descriptions of *S. robertsi* (Kottelat, 1990) and *S. crocotula* (Plongsesthee et al., 2013).

Checking of presence or absence of axillary pelvic lobe show us also some interesting information. There are three lineages, which have it present. In Balteata it was already mentioned in former descriptions of this species S. balteata (Kottelat, 1990). Next lineage is Sumo, which is sister to Balteata. But very surprising was the presence of such big lobe in

Aurantiaca 2, what means, that this lineage is usually not S. aurantiaca, because in original description is mentioned absence of this character (Plongsesthee et all, 2011) and also in all our populations of Aurantiaca 1 is the lobe absent.

Variability in meristic characters

The meristic counts inside the group are much more variable than it is usual in other species of Nemacheilidae. Only the number of anal fin rays is stable, but the others are useful for identification of several lineages, but again, only in combination with other characters.

Cincticauda has only 5 (5+1/2) branched pelvic fin rays, what was already mentioned by Kottelat, (1990), but all other lineages have 6 (6+1/2) branched ones. For Cincticauda is also specific 8 branched dorsal fin rays (Kottelat, 1990) and all other lineages have 7.

Caudal fin ray counts are in the lineages which are corresponding to described species fitting with the original descriptions. The number of rays in the lower caudal-fin lobe is not informative, since nearly all specimens have 8 branched fin rays, except the single Robertsi 3 specimen with only 7. The number of branched rays in the upper lobe was informative, even in the case of the otherwise very similar lineages Robertsi 1 and Robertsi 2. Robertsi 1 has usually 9, or exceptionally 8 branched fin rays in upper caudal lobe and Robertsi 2 has usually 7 or 6 but never 8 or 9. By this character it is possible to connect the lineage Robertsi 1 with original description of S. robertsi from Kottelat (1990). Also for distinguishing of Pinkies 1 and 2 it is a suited character, because Pinkies 1 have 9 branched fin rays in upper caudal lobe and Pinkies 2 only 8.

Pectoral fin ray counts in the lineages, which are corresponding with the described species, are fitting with the original description. Interesting is this character in Aurantiaca 2 lineage, where is same number like in Balteata, what means two more branched pectoral fin rays then in Aurantiaca 1 lineage, which have same number as is mentioned in the original description of *S. aurantiaca* (Page et al., 2011). The number of pectoral fin ray is also the second suited character for identification of Robertsi 1 from Robertsi 2, because Robertsi 1 has 8 (8+1/2) branched rays and Robertsi 2 has only 7 (7+1/2) branched rays.

5.2.4. Cases of co-occurrence

Robertsi 1 with Robertsi 2

Our Phylogenetic trees show that one from 7 samples collected in the south tip of the Phang Nga province in south Thailand belongs to a different main lineage. The six samples A2445-A2446 and A2449-A2451 belong in all phylogenetic trees into the Robertsi 1 lineage

and the other one sample A2448 belongs to the Robertsi 2 lineage (see Fig. 12 – 14 in Attachments). Those samples were collected in a very small forest stream, which continue only several kilometres and then disappearing somewhere in the ground. The closest streams in the west are flowing to the area of Robertsi 1 lineage and the closest streams on east are flowing into the area of Robertsi 2 lineage. By our hypothesis the stream changes the flow from one side to the other and this could be the reason, why it contains fishes from both main lineages.

Pinkies 1 with Robertsi 2

In a small stream in Khao Sok National Park were two times on same place collected such big groups of fishes containing several individuals from both main lineages Pinkies 1 and Robertsi 2. From the first collection are in the phylogenetic trees visible samples A 4672 and 4674-4676 representing Pinkies 1 lineage and 4673 and 4680-4681 (Fig. 12, 13, 14 in Attachments). Next time there were collected unidentifiable juveniles and they were grown by Jens Kuehne in aquarium and from them comes samples A11261 representing lineage Pinkies 1 and A11262 representing Robertsi 2 lineage (Fig. 12 in Attachments). Another ten individuals fixed in Formaldehyde are by the morphological characters perfectly identifiable also. From the ten fishes are four representing Pinkies 1 lineage and six are representing Robertsi 2 lineage. Khao Sok National park was most likely the place of origin of pinkies and it was later recolonised by the fishes from Robertsi 2 lineage.

Balteata with Aurantiaca

In the Khwae Noi river in upper Mae Klong river system were on one locality collected 4 individuals of the Balteata lineage (A4413-A4416) and five individuals from the Aurantiaca lineage (A4417-A4421) (Fig. 12, 13, 14 in Attachments). Schistura balteata is a very rarely recorded species and was quite long time known only from the single type specimen from Dawei river in north Tenasserim (Kottelat, 1990). But in last few years they were collected on several other localities including Mae Klong river basin (Page et al., 2012, Beamish and Plongsesthee, 2015) and it looks like they are more widely spread.

Aurantiaca with Sumo in upper Ataran

In upper Ataran river were on two places not far to each other collected four specimens (A954-957) from Aurantiaca lineage and one specimen (A11005) from Sumo lineage (Fig. 12, 13, 14 in Attachments). By my hypothesis those localities are close enough for the fishes to come into contact, because this part of the Ataran river is up to now quite natural and without any known barrier for fishes.

5.2.5. Comparison of our results with present taxonomy

Inside S. robertsi group were found ten deeply isolated lineages and five of them are by morphological characters perfectly corresponding with described species. The lineage Cincticauda perfectly fit to the former description of *Schistura cincticauda* by Kottelat (1990), Balteata perfectly fit to original description of *Schistura balteata* (Rendahl, 1948), Aurantiaca 1 perfectly fit to original description of *Schistura aurantiaca* (Page et al., 2011), Crocotula perfectly fit to original description of *Schistura crocotula* (Kottelat et al., 2013), The lineage Robertsi 1 by meristic count and distribution area perfectly fit to original description of *Schistura robertsi* (Kottelat, 1990).

Sumo lineage is the sister lineage to *S. balteata* and theoretically both could represent local forms of one species, but Sumo lineage has several severe differences in morphology, and therefore we consider it a distinct but unnamed species. The other four are quite different and they have significant morphological characters. By the phylogenetic position those lineages could not be placed under any other species, because they will make it polyphyletic.

5.2.6. Taxonomic implication

Schistura robertsi group contains 5 described species and two known undescribed ones, but our new phylogenetic results show, that one of the species, *Schistura robertsi* is polyphyletic, because it is forming several isolated lineages thorough whole tree. By our opinion all the main lineages represent distinct species. In phylogenetic trees all the main lineages are deeply isolated and also the percent genetic difference between 14 to 16 % in mitochondrial and between 2 and 4 % in nuclear DNA is such high number.

There is also possibility for identification of every lineage by several morphological characters and only small knowledge of geography. *Schistura* sp sumo could be distinguished from all other species by presence of axillary pelvic lobe, presence of black marks on lower lip and by the position of anus. *Schistura aurantiaca* 2 could be distinguished from all other lineages by presence of axillary pelvic lobe, by presence of black marks on lower lip and by presence of black mark on dorsal fin. *Schistura* sp. pinkie 1 could be distinguished from all other lineages by position of anus, length of lateral line and different number of fin rays in caudal fin. *S. robertsi* 2 could be distinguished from all other lineages by position of anus, specific number of branched pectoral fin rays and by specific number of caudal fin rays. *Schistura robertsi* 3 could be from other lineages distinguished by length of lateral line,

position of anus and specific number of pectoral and also caudal fin rays, but from this lineage we have only one specimen and for making some taxonomic results, we will need to check more specimens.

Problematic identification could be only in two lineages. *Schistura* sp. Pinkie 2 has same fin ray counts and also other morphological characters like *Schistura aurantiaca* 1. For identification of these two lineages must be used colour pattern differences. *S.* sp. pinkie 2 has four or five (pink) interspaces on predorsal part of the body and no single interspace visible on caudal peduncle. *Schistura aurantiaca* 1 has on predorsal part of the body only two (orange) interspaces, one directly after head and one shortly before dorsal fin. *S. aurantiaca* 1 also has two or three (orange) interspaces on caudal peduncle. The difference in colour is visible only on live fishes and on preserved ones, we must count only with the position and number of the interspaces. Useful character to distinguish these two lineages is also the position of distribution areas, which are minimally 500 km far away from each other. *S.* sp pinkie 2 is living only in Khao Sok national park in middle part of Malay peninsula and *S. aurantiaca* 1 is living in upper Mae Klong river system in north part of Thailand.

Next problematic lineage is *Schistura robertsi* 1, where 40 from 53 individuals have 9+8 fin rays in caudal fin, and these individuals, we are able to identify from all other lineages, by this number of caudal fin rays in combination with anus position and number of pectoral fin rays, but there are 10 individuals, which have 8+8 fin rays in caudal fin, what is same number like in *Schistura crocotula*. Identification of these individuals from *S. crocotula* is then really hard, because the only character, which makes them different, is the number of caudal fin rays. In live fishes, there is the strict difference in colour of interspaces, which are in *S. crocotula* orange and in *S. robertsi* 1 from white up to light yellow. Next little bit tricky character is, that *S. crocotula* has usually two interspaces (first before and the second under dorsal fin base) broader, than the others in whole body and *S. robertsi* 1 have all the interspaces same broad.

6. Conclusion

Phylogenetic reconstructions show in both groups several isolated lineages with distinct geographic distribution (evolutionary units).

Lineages in *Schistura robertsi* group are more isolated and usually older than lineages in *Paracanthocobitis zonalternans* group. The difference in age and level of isolation could be caused by strong bottleneck effects and fast genetic drift connected with the specific ecology of *S. robertsi*.

The results of biogeographical reconstruction and time estimations show that the large mountain ridges are not necessarily a barrier for fishes, but that sea water flooding has permanent impact on their phylogeny. See water level was the most important factor in the evolutionary history of the analysed fishes.

Ways of colonisation and geologic period of genealogic events were in both group quite different, but they share few features: Both groups evolved in the northern Tenasserim region, then they spread in several waves and these events were in periods of lowered global sea level. On the other hand, during periods of higher sea level the separation between different lineages were manifested. Both groups managed to colonise the Mae Klong River basin, in case of *P. zonalternans* even two times independently and on different pathways, meaning they crossed the Tenasserim Mountains, an otherwise known biogeographic barrier for freshwater fishes. Both groups colonised southern Thailand and by that crossed the Isthmus of Kra, the Surat Thani - Krabi line and the Kangar – Pattani line. These lines obviously did not represent a barrier for colonisation, but at least in the case of the Surat Thani - Krabi line the prolonged presence of a seawater belt caused a separation between the populations north and south from it.

In both groups a detailed sampling allowed producing phylogenetic trees with fine resolution. This was the base for the reconstruction of the genealogic events, their ages and the biogeographic causes. The combination with morphologic data and the identification of hybridisation events added further information on separation processes. The combined dataset basing on a fine-scaled sampling and several analytic approaches allowed the reconstruction of the evolutionary history of two groups of model organisms from their origin to the present day.

7. References

Allen, D. J., Molur, S. & B. A. Daniel. 2010. The Status and Distribution of Freshwater Biodiversity in the Eastern Himalaya. Cambridge, UK and Gland, Switzerland: IUCN, and Coimbatore, India: Zoo Outreach Organisation. ISBN: 978-2-8317-1324-3

Bănărescu, P. & T. T. Nalbant. 1966. The 3rd Danish Expedition to Central Asia. Zoological results 34. Cobitidae (Pisces) from Afghanistan and Iran. *Videnskabelige Meddelelser fra Dansk Naturhistorisk Forening i København*, 129: 149–186, pls. 19–21.

Beamish, F. W. H & R. Plongsesthee. 2015. Species Diversity, Abundance and Environmental Associations of Loaches (Nemacheilidae and Balitoridae) in the Central region of Thailand. *Nat. Hist. Bull. Siam. Soc.* 60(2): 89–111.

Blankenship, R. E. (2009). Molecular Mechanisms of Photosynthesis. *Blackwell Publishing*. pp. 124–132

Bohlen, J., V. Šlechtová, H. H. Tan & R. Britz. 2011. Phylogeny of the Southeast Asian freshwater fish genus *Pangio* (Cypriniformes; Cobitidae). *Mol. Phylogenet. Evol*, 61(3): 854–865.

Bohlen, J. & V. Šlechtová. 2009. *Schistura udomritthiruji*, a new loach from southern Thailand (Cypriniformes: Nemacheilidae). *Ichthyol. Explor. Freshwaters*, 20: 319-324.

Bohlen, J. & V. Šlechtová. 2010. Genetic lineages and genera in river loaches (Nemacheilidae), a puzzle for patient players. *Book of abstracts*, International Loach Conference.

Bouckaert, R., Heled J., Kühnert, D., Vaughan, T., Xie W. C-H. D., Suchard, M. A., Rambaut A. & A. J. Drummond. 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Comput Biol* 10(4): e1003537

Chen, W.-J., Miya M., Saitoh K. & R. L. Mayden. 2008. Phylogenetic utility of two existing and four novel nuclear gene loci in reconstructing Tree of Life of ray-finned fishes: The order Cypriniformes (Ostariophysi) as a case study. *Gene*. 423:125-34.

Chen, J.-M. et al. 2018. Large-scale phylogenetic analyses provide insights into unrecognized diversity and historical biogeography of Asian leaf-litter frogs, genus *Leptolalax* (Anura: Megophryidae). *Mol. Phylogenet. Evol* 124: 162–171.

Chu, X.-L. & Y.-R. Chen. 1979. A new blind cobitid fish (Pisces, Cypriniformes) from subterranean waters in Yunnan, China. *Acta Zoologica Sinica*, 25(3): 285–287.

Clark, S.H.B., Zhongcheng, W. & F.G. Poole,2004. Comparison of some sediment-hosted, stratiform barite deposits in China, the United States, and India. *Ore Geology Reviews* 24: 85 – 101

de Bruyn, M., Nugroho, E., Hossain, M. M., Wilson, J.C. & P. B. Mather. 2005. Phylogeographic evidence for the existence of an ancient biogeographic barrier: the Isthmus of Kra Seaway. *Heredity* 94:370-8.

de Bruyn, M., Rüber, L., Nylinder, S., Stelbrink, B., Lovejoy, N. R. & S. Lavoué. 2013. Paleo-Drainage Basin Connectivity Predicts Evolutionary Relationships across Three Southeast Asian Biodiversity Hotspots. *Syst Biol*. 62: 398-410.

de Bruyn, M., Stelbrink, B., Morley, R. J., Hall, R., Carvalho, G. R. & C. H. Cannon. 2014. Borneo and Indochina are Major Evolutionary Hotspots for Southeast Asian Biodiversity. *Syst Biol*. 63: 879-901.

Day, F. 1877. The fishes of India; being a natural history of the fishes known to inhabit the seas and freshwaters of India, Burma, and Ceylon. *Quaritsch, London*. pls. 79–138;

Esposti, M.D., Crimi, M., Ghelli, A., Patarnello, T., Meyer, A. & S. De Vries. 1993. Mitochondrial cytochrome b: evolution and structure of the protein. *Biochim. Biophys.Acta*.**1143** (3): 243–271.

Grant, S. 2007. A new subgenus of Acanthocobitis Peters, 1861 (Teleostei: Nemacheilidae). *Ichthyofile* 2:1–9.

Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids. Symp. Ser.* 41, 95–98.

Haq, B.U., Hardenbol, J. & Vail, P.R. 1987. Chronology of fluctuating sea levels since the Triassic. *Science*, 235, 1156- 1167.

Herzenstein, S. 1888. Wissenschaftliche Resultate der von N. M. Przewalski nach Central-Asien unternommenen Reisen. Band 3, Abth. 2. Fische. Kaiserliche Akademie der Wissenschaften, St. Petersburg, *Fasc.* 1 (1888): i–vi, 1–90, pls. 1–8, 2

Hughes, J.B., Round, P.D., & D. S. Woodruff. 2003. The Indochinese–Sundaic faunal transition at the Isthmus of Kra: an analysis of resident forest bird species distributions. *Journal of Biogeography* 30: 569-580.

Kottelat, M. 1990. Indochinese nemacheilines. A revision of nemacheiline loaches (Pisces: Cypriniformes) of Thailand, Burma, Laos, Cambodia and southern Viet Nam. *Raffles Bull. Zool,* 38(1): 85–86.

Kottelat M. 2012. Conspectus Cobitidum: an inventory of the loaches of the world (Teleostei: Cypriniformes: Cobitoidei). *Raffles Bull. Zool. Suppl.,* 26: 1–199.

Kumar, S., Stecher, G. & Tamura, K. 2016.) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular biology and evolution*, 33 (7), 1870–1874.

Lanfear, R., Calcott, B., Ho, S.Y. & Guindon, S. 2012. Partition Finder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, 29 (6), 1695–1701.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. & D.G. Higgins. 2007. Clustal, W & Clustal, X. Version 2.0. *Bioinformatics*, 23, 2947–2948.

Lecompte, E., Aplin, K., Denys, Ch., Catzeflis, F., Chades, M. & P. Chevret. 2008 Phylogeny and biogeography of African Murinae based on mitochondrial and nuclear gene sequences, with a new tribal classification of the subfamily *BMC Evolutionary Biology*, 8:199

Lim, K. K. P. & H. H. Tan. 2002. Freshwater fish diversity in Penninsular Malaysia: A review of recent findings. pp. 461-481 in: A. Ali et al. (eds.) The Asian wetlands: bringing partnership into good wetland practices. Penerbit Universiti Sains Malaysia, Penang.

Linck, H. F. 1790. Versuch einer Eintheilung der Fische nach den Zähnen. *Magazin für das Neueste aus der Physik und Naturgeschichte, Gotha*, 6(3): 28–38. *[Fascicle 1 dated 1789, fascicle 2–4 1790]*.

Loo, Y. Y., Billa, L. & A. Singh. 2014. Effect of climate change on seasonal monsoon in Asia and its impact on the variability of monsoon rainfall in Southeast Asia. *Geoscience Frontiers*: 1-7

Mayden, R. L., Chen, W. -J., Bart, H. L., Doosey, M. H., Simons, A. M., Tang, K. L., Wood, R. M., Agnew, M. K., Yang, L., Hirt, M. V., Clements, M. D., Saitoh, K., Sado, T., Miya, M. & M. Nishida. 2009. Reconstructing the phylogenetic relationships of the earth's most diverse clade of freshwater fishes – order Cypriniformes (Actinopterygii: Ostariophysi): a case study using multiple nuclear loci and the mitochondrial genome. *Mol. Phylogenet. Evol.,*51(3): 500–514.

Parnell, J. 2013. The biogeography of the Isthmus of Kra region: a review. *Nordic Journal of Botany* 31: 1-15.

Peters, W. 1861. Über zwei neue Gattungen von Fischen aus dem Ganges. Monatsberichte der Königlichen Preuss[ischen] *Akademie der Wissenschaften zu Berlin* 1861: 712–713.

Plongsesthee, R., L. M. Page. & F. W. H. Beamish. 2011. *Schistura aurantiaca*, a new species from the Mae Khlong basin, Thailand (Teleostei: Nemacheilidae). *Ichthyol. Explor. Freshwaters*, 22: 169-178.

Plongsesthee, R., M. Kottelat & F. W. H. Beamish. 2013. Schistura crocotula, a new loach from Peninsular Thailand (Teleostei: Nemacheilidae). *Ichthyol. Explor. Freshwaters* v. 24 (no. 2): 171-178.

Prokofiev A. M. & A. S. Golubtsov. 2013. Revision of the loach genus *Afronemacheilus* (Teleostei: Balitoridae: Nemacheilinae) with description of a new species from the Omo-Turkana basin, Ethiopia. *Ichthyol. Explor. Freshwaters,* 24 (1): 1-14.

Quenouille, B., Bermingham, E. & S. Planes. 2004. Molecular systematics of the damselfishes (Teleostei: Pomacentridae): Bayesian phylogenetic analyses of mitochondrial and nuclear DNA sequences. *Mol Phylogenet Evol*. 31:66-88.

Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & M. A. Suchard. 2018. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*.

Rambaut, A. & A.J. Drummond. 2010. TreeAnnotator. Version 1.8, Distributed as Part of the BEAST Package

Rambaut, A. 2009. FigTree v.1.4. Available at: http://tree.bio.ed.ac.uk/software/figtree/

Regan, C. T. 1911. The classification of teleostean fishes of the order Ostariophysi. I. Cobitoidea. *Ann. Mag. Nat. Hist*., 8: 13–32.

Richards, P.W. 1996. The tropical rain forest: an ecological study, 2nd edn. *Cambridge University Press*, Cambridge.

Ronquist, F. & J. P. Huelsenbeck. 2003 MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 19, 1572-1574.

Sambrook, J., Fritsch, E.F., & T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual, *second ed. Cold Spring Harbor Laboratory Press*, Cold Spring Harbor, NY, USA.

Sember, A., Bohlen, J., Šlechtová, V., Altmanová, M., Symonová, R. & P. Ráb. 2016. Karyotype differentiation in 19 species of river loach fishes (Nemacheilidae, Teleostei): extensive variability associated with rDNA and heterochromatin distribution and its phylogenetic and ecological interpretation. *BMC Evolutionary Biology,* 15:251.

Singer, R. A., Pfeiffer, J. M. & L. M. Page. 2017. A revision of the *Paracanthocobitis zonalternans* (Cypriniformes: Nemacheilidae) species complex with descriptions of three new species. *Zootaxa* 4324 (1): 085–107.

Swainson, W. 1838. The natural history of fishes, amphibians, & reptiles, or monocardian animals. Vol. 1. Longman, Brown, Green & Longmans, London. vi+368 pp.

Šlechtová, V., Bohlen, J., Freyhof, J. & P. Ráb. 2006. Molecular phylogeny of the Southeast Asian freshwater fish family Botiidae (Teleostei: Cobitoidea) and the origin of polyploidy in their evolution. *Mol Phylogenet Evol*., 39:529-41.

Šlechtová, V., Bohlen, J. & H. H. Tan. 2007. Families of Cobitoidea (Teleostei; Cypriniformes) as revealed from nuclear genetic data and the position of the mysterious genera Barbucca, Psilorhynchus, Serpenticobitis and Vaillantella. *Mol Phylogenet Evol*. 44:1358-65.

Šlechtová, V., J. Bohlen & A. Perdices. 2008. Molecular phylogeny of the freshwater fish family Cobitidae (Cypriniformes: Teleostei): delimitation of genera, mitochondrial introgression and evolution of sexual dimorphism. *Mol. Phylogenet. Evol*, 47: 812–831.

Šlechtová, V. & J. Bohlen. 2010. Main lineages of river loaches (Nemacheilidae) as inferred from molecular genetic markers. *Book of abstracts*, International Loach Conference.

Tosi, J. C. M., & D. J. Melnick. 2003. Paternal, maternal, and biparental molecular markers provide unique windows onto the evolutionary history of macaque monkeys. *Evolution*, 57(6), pp. 1419–1435

Wells, D.R. 1976. Resident birds. The birds of the Malay peninsula, Vol. V (by Lord Medway & D.R. Wells), pp. 1–33. Witherby, London and Penerbit Universiti Malaya, Kuala Lumpur.

Woodruff, D. S. 2003. Neogene Marine Transgressions, Paleogeography and Biogeographic Transitions on the ThaiMalay Peninsula. *Journal of Biogeography*, Vol. 30, No. 4, pp. 551-567

Yu, Y., Harris, A. J., Blair, Ch., & X. He. 2015. RASP (Reconstruct Ancestral State in Phylogenies): A tool for historical biogeography. *Mol. Phylogenet. Evol* 87: 46–49

Zakaria-Ismail, M. 1994. Zoogeography and biodiversity of the freshwater fishes of Southeast Asia. *Hydrobiologia* 285: 41-48.

8. Attachment

Tab. 18 Collection of samples from *Paracanthocobitis* group fixed in ethanol.

	A NUMBER		BODIES	FIN	COUNTRY	PROVINCE	RIVER
Total number			**162**	**35**			
A. pavonaceus	1863	1864		2			
P. botia	3437	3437	1		Nepal		
P. linypha	2562	2566	5		Myanmar	Kachin	Irrawaddy
P. linypha	2567	2568	2		Myanmar	Kachin	Irrawaddy
P. linypha	9209	9209	1		Myanmar	Kachin	Irrawaddy
P. linypha	9836	9836	1		Myanmar	Kachin	Irrawaddy
P. mackenziei	82	85	4			Ornamental fish trade	
P. mackenziei	494	496	2		Bangladesh		
P. pictilis	4512	4512	1		Myanmar	Mon state	Salween
P. pictilis	6940	6941		2	Thailand	Aquarium trade	
P. spec.	9840	9843	4		Myanmar	Tanintharyi	Tenasserim
P. spec.	9844	9845	2		Myanmar	Tanintharyi	Tenasserim
P. spec.	2460	2465	6				
P. spec.	9837	9839	3		Myanmar	Tanintharyi	Tenasserim
P. urophthalma	330	331	2				
P. zonalternans	772	776		5	Thailand	Mae Hong Son	Salween
P. zonalternans	793	798		6	Thailand	Mae Hong Son	Salween
P. zonalternans	811	816		6	Thailand	Mae Hong Son	Salween
P. zonalternans	900	909	10				
P. zonalternans	1354	1354	1		Thailand	Mae Hong Son	Salween
P. zonalternans	1355	1355	1		Thailand	Mae Hong Son	Salween
P. zonalternans	1364	1366	3		Thailand	Mae Hong Son	Salween
P. zonalternans	1367	1371	5		Thailand	N. Si T.	Pak Paying
P. zonalternans	1680	1681	2				
P. zonalternans	2355	2358	4		Thailand	Ranong	Kra Buri
P. zonalternans	2466	2468	3				
P. zonalternans	2691	2691	1		Thailand	Phang Nga	Bang Niang
P. zonalternans	3005	3005	1		Myanmar		
P. zonalternans	4025	4025	1		Myanmar	Bago	Sittaung
P. zonalternans	4102	4130	29		Myanmar	Mon state	Sittaung
P. zonalternans	4356	4356		1	Myanmar		
P. zonalternans	4427	4428	2		Malaysia	Kedah	Sungai Padang
P. zonalternans	4447	4450	4		Thailand	Satun province	
P. zonalternans	4453	4453	1		Thailand	Tak province	Mae Klong
P. zonalternans	4515	4515	1		Myanmar	Mon state	
P. zonalternans	4797	4797	1		Thailand	Phuket province	
P. zonalternans	4920	4921	2		Thailand		Salween
P. zonalternans	4934	4935	2		Thailand		Salween
P. zonalternans	4937	4939	3		Thailand	Tak	Salween
P. zonalternans	4973	4973	1		Thailand		Salween
P. zonalternans	4979	4986	4		Thailand		Salween
P. zonalternans	4992	5001	10		Thailand	Tak	Salween
P. zonalternans	5175	5175	1		Thailand	Trang	Palian
P. zonalternans	5178	5178	1		Thailand	N. Si T.	
P. zonalternans	5185	5188	4		Thailand	Phatthalung	Pak Phaniat
P. zonalternans	5190	5191	2		Thailand	Phatthalung	Pak Phaniat
P. zonalternans	5236	5246	6		Thailand	Ranong	
P. zonalternans	5327	5332		3	Myanmar		
P. zonalternans	5559	5559	1		Myanmar	Rakhine	Irrawaddy
P. zonalternans	5774	5778	5		Myanmar	Magway	Irrawaddy

P. zonalternans	6309	6310		1	Myanmar	Magway	Irrawaddy
P. zonalternans	6531	6562	32		Myanmar	Tanintharyi	
P. zonalternans	6573	6578	6		Myanmar	Ayeyerwady	Irrawaddy
P. zonalternans	6621	6622	2		Myanmar	Magway	Irrawaddy
P. zonalternans	6660	6660	1		Myanmar	Magway	Irrawaddy
P. zonalternans	6709	6710	2		Myanmar	Rakhine	Bowdi chaung
P. zonalternans	7057	7059	3		Thailand	Phang Nga	Takua Pa
P. zonalternans	7060	7060	1		Thailand	Kanchanaburi	
P. zonalternans	7554	7558		3	Myanmar	Aquarium trade	
P. zonalternans	8327	8327	1		Thailand	Tak	Salween
P. zonalternans	9713	9714		2	Thailand	Phang Nga	Tam Nang
P. zonalternans	9725	9730		6	Thailand	Mae Hong Son	Salween
P. zonalternans	9834	9834	1		Myanmar	Tanintharyi	Lenya
P. zonalternans	9835	9835	1		Myanmar	Tanintharyi	Lenya

Tab. 19 Collection of samples from *Paracanthocobitis* group fixed in formaldehyde

SPECIES	A NUMBER		FORMOL	COUNTRY	PROVINCE	RIVER
Total number			**178**			
A. pavonaceus	1863	1864	2			
A. pavonaceus	9606	9606	1			
P. botia	5560	5646	87	Myanmar	Tanintharyi	
P. botia	9200	9200	1	India	Uttarakhand	Ganges
P. mackenziei	494	496	1	Bangladesh		
P. pictilis	6940	6941	2	Thailand	Aquarium trade	
P. zonalternans	772	776	5	Thailand	Mae Hong Son	Salween
P. zonalternans	793	798	6	Thailand	Mae Hong Son	Salween
P. zonalternans	811	816	6	Thailand	Mae Hong Son	Salween
P. zonalternans	4356	4356	1	Myanmar		
P. zonalternans	4474	4474	1	Thailand		
P. zonalternans	4940	4943	4	Thailand	Tak	Salween
P. zonalternans	4979	4986	4	Thailand		Salween
P. zonalternans	5236	5246	5	Thailand	Ranong	
P. zonalternans	5248	5251	4	Thailand	Ranong	
P. zonalternans	5327	5332	3	Myanmar		
P. zonalternans	5678	5678	1	Myanmar	Magway division	Irrawaddy
P. zonalternans	5772	5772	1	Myanmar	Rakhine	Bowdi chaung
P. zonalternans	5790	5799	10	Myanmar	Magway division	Irrawaddy
P. zonalternans	5817	5820	4	Myanmar	Ayeyerwady	Irrawaddy
P. zonalternans	6293	6293	1	Myanmar	Magway division	Irrawaddy
P. zonalternans	6309	6310	2	Myanmar	Magway division	Irrawaddy
P. zonalternans	6816	6816	1	Myanmar	Ayeyerwady	Irrawaddy
P. zonalternans	7502	7503	2	Thailand	Phang Nga	
P. zonalternans	7509	7511	3	Thailand	Ranong	
P. zonalternans	7554	7558	5	Myanmar	Aquarium trade	
P. zonalternans	9713	9714	2	Thailand	Phang Nga	Tam Nang
P. zonalternans	9716	9724	9	Thailand	Mae Hong Son	Salween
P. zonalternans	9725	9730	6	Thailand	Mae Hong Son	Salween

Tab. 20 Colection of samples from *Schistura robertsi* group fixed in ethanol

SPECIES	A NUMBER		BODIES	FIN	COUNTRY	RIVER	STREAM
Total number			**103**	**9**			
S. aurantiaca	954	957	4		Myanmar	Ataran	
S. aurantiaca	4417	4421	5		Thailand	Mae Klong	Khwae Noi
S. aurantiaca	9580	9580	1		Thailand	Mae Klong	Khwae Yai
S. aurantiaca	9584	9584	1		Thailand	Mae Klong	Khwae Yai
S. aurantiaca	10979	10979	1		Thailand	Mae Klong	Mae Klong Noi
S. aurantiaca	11000	11000	1		Thailand	Mae Klong	Mae Klong Noi
S. aurantiaca	11067	11067	1		Thailand	Mae Klong	Mae Klong Noi
S. aurantiaca	11180	11180		1	Thailand	Mae Klong	Mae Klong Noi
S. balteata	324	324	1			OFT	
S. balteata	4413	4416	4		Thailand	Mae Klong	Khwae Noi
S. cf *aurantiaca*	10902	10902	1		Thailand	Mae Klong	Mae Klong Noi
S. cf *aurantiaca*	11010	11010	1		Thailand	Mae Klong	Mae Klong Noi
S. cf *robertsi*	10966	10966	1		Thailand	Takua Pa	
S. cf *robertsi*	11002	11002	1		Thailand	Takua Pa	
S. cf *robertsi*	11018	11019	2		Thailand	Tapi	
S. cf *robertsi*	11248	11248	1		Thailand	Tapi	Phum Duan
S. cf 'Pinky'	4672	4679	8		Thailand	Tapi	Khlong Sok
S. cf 'Pinky'	11261	11261	1		Thailand	Tapi	Phum Duan
S. cf. *robertsi*	3689	3690	2		Thailand	Tapi	
S. cf. *robertsi*	10888	10888	1		Thailand	Bang Patong	
S. cf. *robertsi*	10905	10905	1		Thailand	Pranburi	
S. cincticauda	8312	8313	2		Thailand	Salween	Mae Moei
S. crocotula	9589	9590	2		Thailand		
S. crocotula	9591	9591	1		Thailand		
S. crocotula	10513	10517		1	Thailand		
S. crocotula	10518	10518	1		Thailand		
S. robertsi	1667	1667	1		Thailand	Phuket	
S. robertsi	2351	2354	4		Thailand	Kra Buri	Bang Yai
S. robertsi	2424	2427	4		Thailand	Phang Nga	
S. robertsi	2445	2453	9		Thailand	Phang Nga	
S. robertsi	2553	2553	1		Thailand	Kapoe	Thonglang
S. robertsi	4210	4210		1	Malaysia		
S. robertsi	4396	4412	17		Thailand		
S. robertsi	4680	4681	2		Thailand	Tapi	Khlong Sok
S. robertsi	5084	5093	10		Malaysia		
S. robertsi	5179	5179	1		Thailand	Trang	Lam Phu Ra
S. robertsi	5247	5247	1		Thailand		
S. robertsi	6892	6893	2		Thailand	Rangu	
S. robertsi	6974	6975	2		Thailand	Krabi River	
S. robertsi	6988	6989	2		Thailand	Tapi	
S. robertsi	7049	7053	5		Thailand	Kapong	Takua Pa
S. robertsi	7146	7148	3		Thailand	Tapi	
S. robertsi	9916	9916	1		Myanmar	Tenasserim	Tenasserim
S. robertsi	10512	10512	1		Thailand	Klong Tadi	
S. robertsi	11030	11030	1		Thailand		
S. robertsi	11031	11032	2		Thailand	Klong Tadi	
S. robertsi	11262	11262	1		Thailand	Tapi	Phum Duan
S. robertsi	11263	11263	1		Thailand	Krabi Noi	
S. robertsi	11264	11264	1		Thailand	Klai	Khlong Krung Ching
S. sp 'Sumo'	2560	2561	2		Myanmar	Irrawaddy	Tanai
S. sp 'Sumo'	11264	11264	1		Thailand	Salween	Ataran
S. sp 'Sumo'	5062	5064	3				
S. sp 'Sumo'	11005	11005	1		Thailand	Salween	Ataran

Tab. 21 Collection of samples from *Schistura robertsi* group fixed in formaldehyde

SPECIES	A NUMBER		FORMOL	COUNTRY	RIVER	STREAM
Total number			**134**			
S. aurantiaca	9557	9559	3	Thailand	Mae Klong	Khwae Yai
S. aurantiaca	10980	10998	19	Thailand	Mae Klong	Mae Klong Noi
S. aurantiaca	10999	10999	1	Thailand	Mae Klong	Mae Klong Noi
S. aurantiaca	11001	11001	1	Thailand	Mae Klong	Mae Klong Noi
S. aurantiaca	11265	11271	7	Thailand	Mae Klong	Mae Klong Noi
S. aurantiaca	11272	11274	3	Thailand	Mae Klong	Mae Klong Noi
S. balteata	4510	4511	2	Thailand	Mae Klong	Khwae Noi
S. balteata	7512	7512	1			
S. cf *aurantiaca*	11011	11016	6	Thailand	Mae Klong	Mae Klong Noi
S. cf *aurantiaca*	11024	11029	6	Thailand	Mae Klong	
S. cf *balteata*	10943	10948	6	Thailand		
S. cf *robertsi*	10967	10970	4	Thailand	Takua Pa	
S. cf *robertsi*	11003	11004	2	Thailand	Takua Pa	
S. cf *robertsi*	11020	11023	4	Thailand	Tapi	
S. cf 'Pinky'	11249	11260	12	Thailand	Tapi	Phum Duan
S. cf. *robertsi*	10841	10844	4	Thailand	Bang Patong	
S. cf. *robertsi*	10845	10848	4	Thailand	Pranburi	
S. crocotula	10513	10517	5	Thailand		
S. robertsi	4210	4210	1	Malaysia		
S. robertsi	4671	4671	1	Thailand	Tapi	Khlong Sok
S. robertsi	6990	6995	6	Thailand	Tapi	
S. robertsi	7040	7048	9	Thailand	Kapong	Takua Pa
S. robertsi	7452	7454	3	Thailand	Phang Nga	Unknown name
S. robertsi	7493	7495	3	Thailand		
S. robertsi	7496	7498	3	Thailand		
S. robertsi	7504	7508	5	Thailand		
S. robertsi	9533	9533	1	Thailand	Tapi	Phum Duang
S. robertsi	10504	10504	1	Thailand	Klong Tadi	
S. robertsi	10505	10511	7	Thailand	Klong Tadi	
S. robertsi	11040	11040	1	Thailand		
S. sp 'Sumo'	11007	11009	3	Thailand	Salween	Ataran

Tab. 22 Sequences used like outgroup in final Bayesian trees from Sember et al. (2016)

Species	Number	Origin	GenBank number		
			Cyt b	RAG 1	IRBP 2
Schistura savona	A7530	OFT	KP738598	KP738558	KP738518
Petruichthys brevis	A4184	OFT	KP738571	KP738531	KP738491
Schistura pridii	A7548	OFT	KP738602	KP738562	KP738522
Physoschistura sp	A7545	OFT	KP738600	KP738560	KP738520
Pteronemacheilus luciodorsum	A8465	OFT	KP738606	KP738566	KP738526
Seminemacheilus lendlii	A4833	OFT	KP738577	KP738537	KP738497
Schistura corica	A6945	OFT	KP738592	KP738552	KP738512
Schistura hypsiura	A6922	OFT	KP738584	KP738544	KP738504
Mesonoemacheilus guentheri	A6935	OFT	KP738588	KP738548	KP738508
Schistura notostigma	A7519	OFT	KP738595	KP738555	KP738515
Physoschistura elongata	A7541	OFT	KP738608	KP738568	KP738528
Nemachilichthys rueppelli	A4341	OFT	KP738573	KP738533	KP738493
Nemacheilus binotatus	A6926	OFT	KP738586	KP738546	KP738506
Schistura bolavensis	A4618	OFT	KP738575	KP738535	KP738495
Schistura fasciolata	A5300	OFT	KP738579	KP738539	KP738499
Lefua costata	A6942	OFT	KP738591	KP738551	KP738511
Barbatula barbatula	A8393	OFT	KP738604	KP738564	KP738524
Cobitis taenia	A1860	Germany	EF508508	EF056334	Present study
Schistura udomritthiruji	A1129	Thailand	Present study	-	Present study
Schistura udomritthiruji	A1131	Thailand	-	-	Present study
Schistura udomritthiruji	A2547	Thailand	Present study	-	-

Fig. 9 Phylogenetic tree of *P. zonalternans* group from cytochrome b dataset

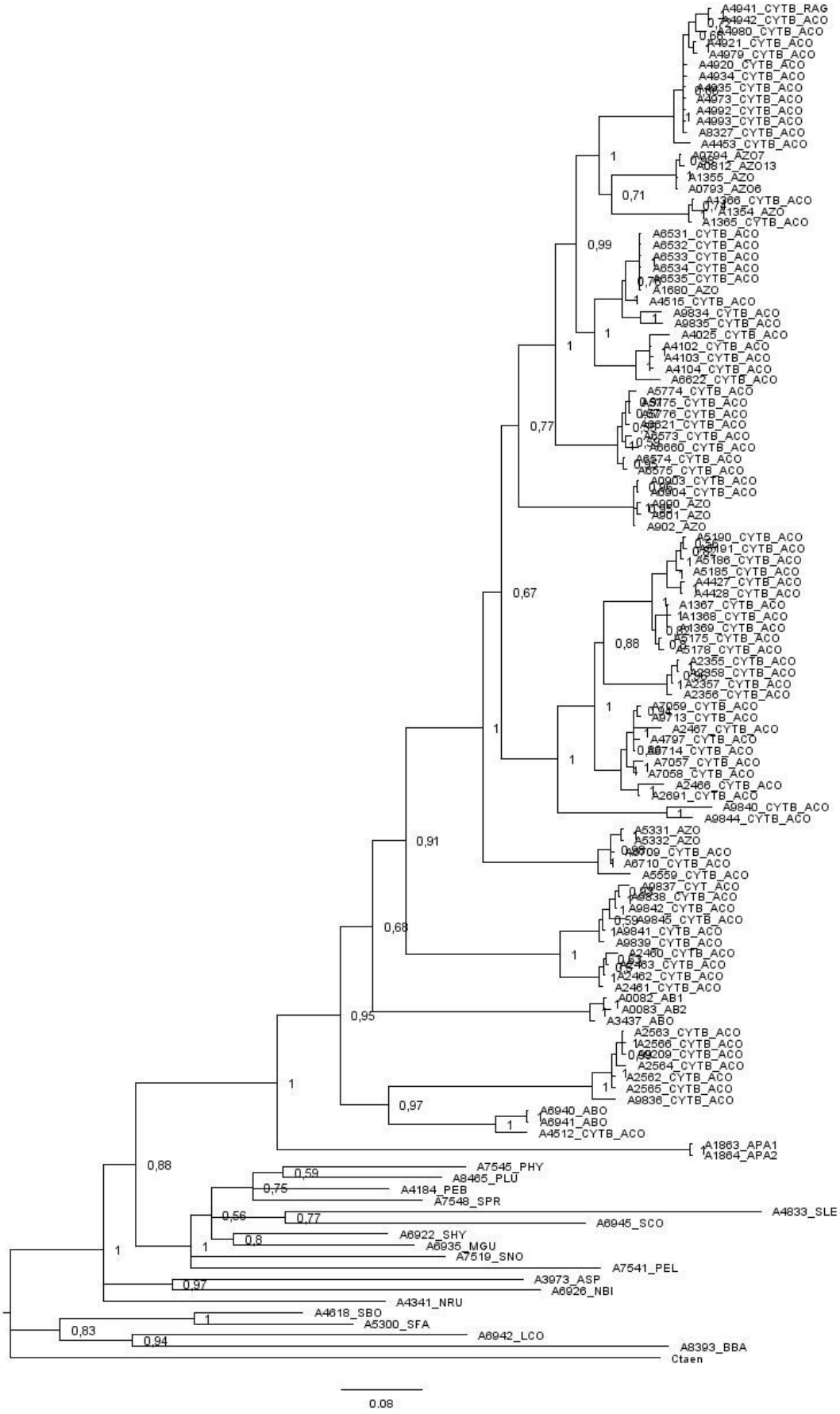


Fig. 10: Phylogenetic tree of *P. zonalternans* group from combined nuclear dataset

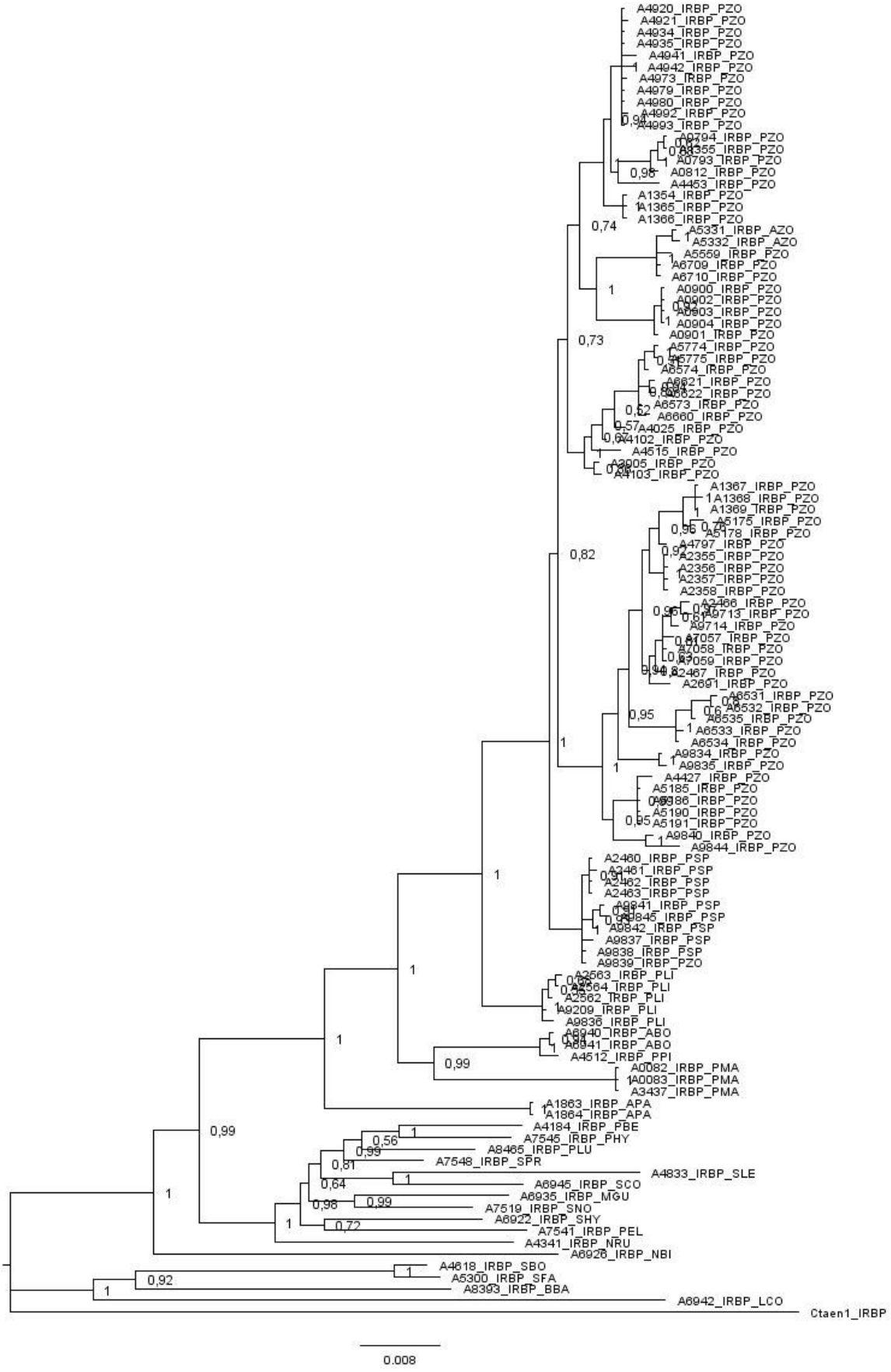


Fig. 11 Phylogenetic tree of *P. zonalternans* group from combined dataset

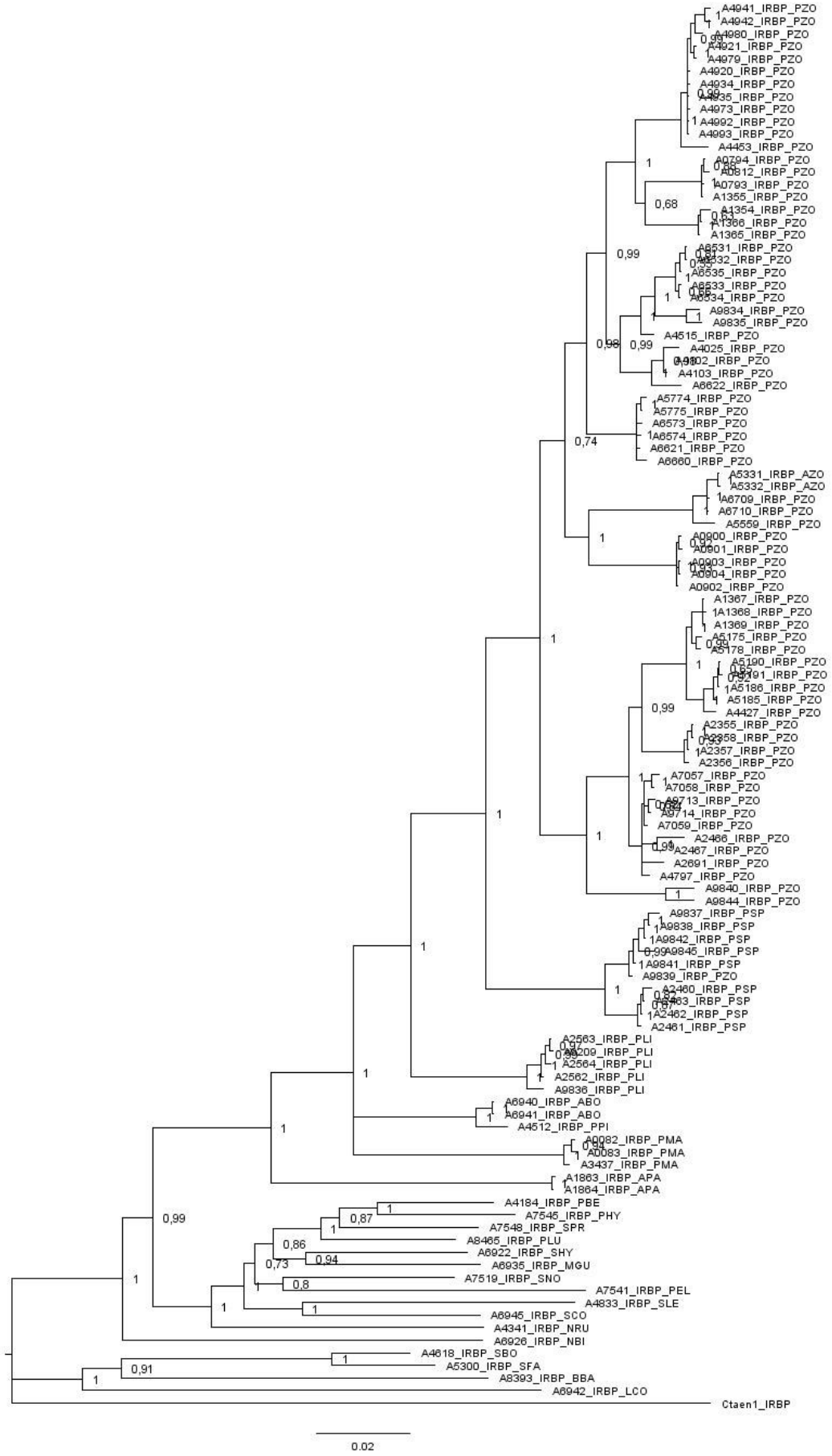


Fig. 12 Phylogenetic tree of *S. robertsi* group from cytochrome b dataset

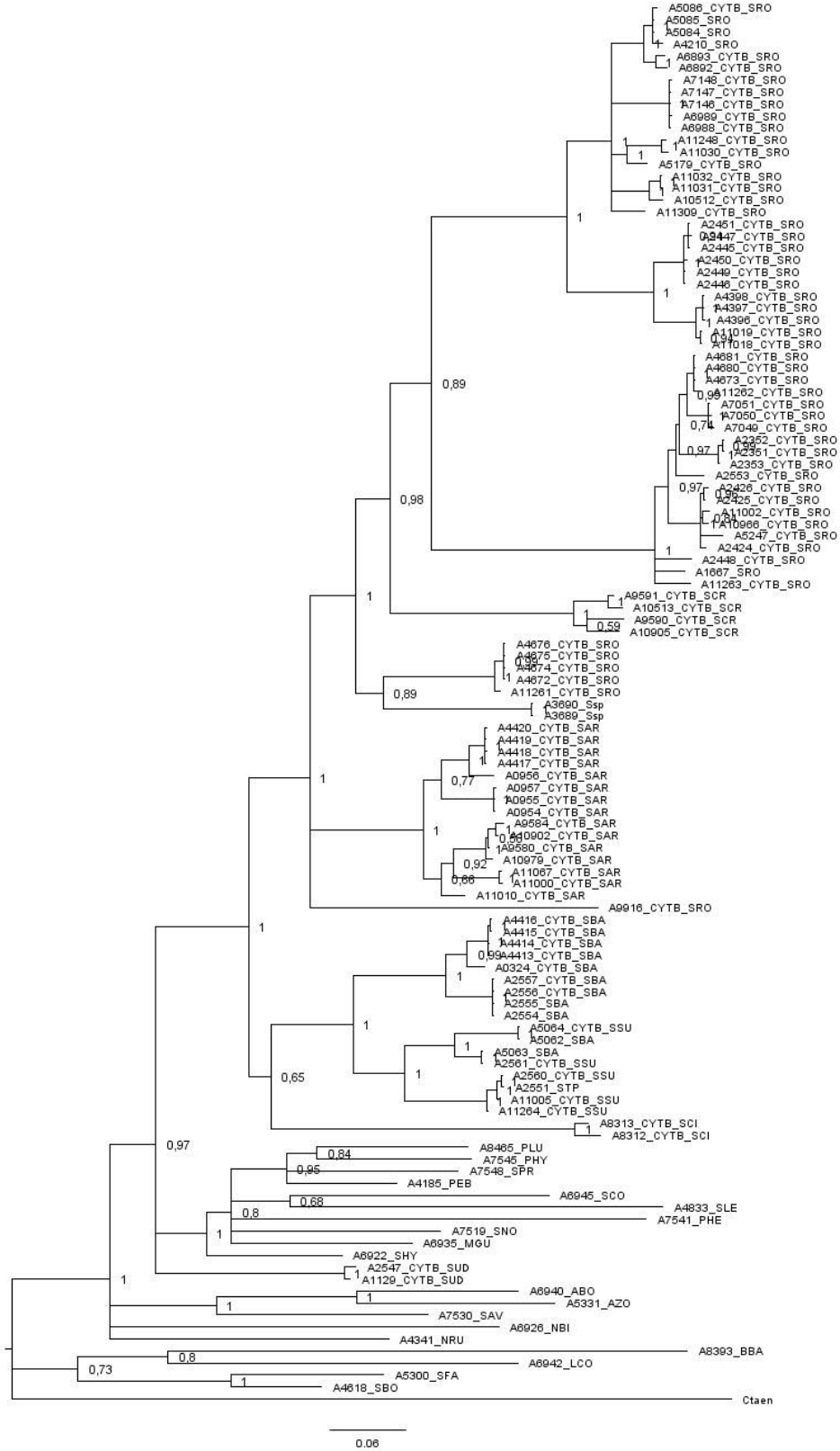


Fig. 13 Phylogenetic tree of *S. robertsi* group from nuclear (IRBP) dataset

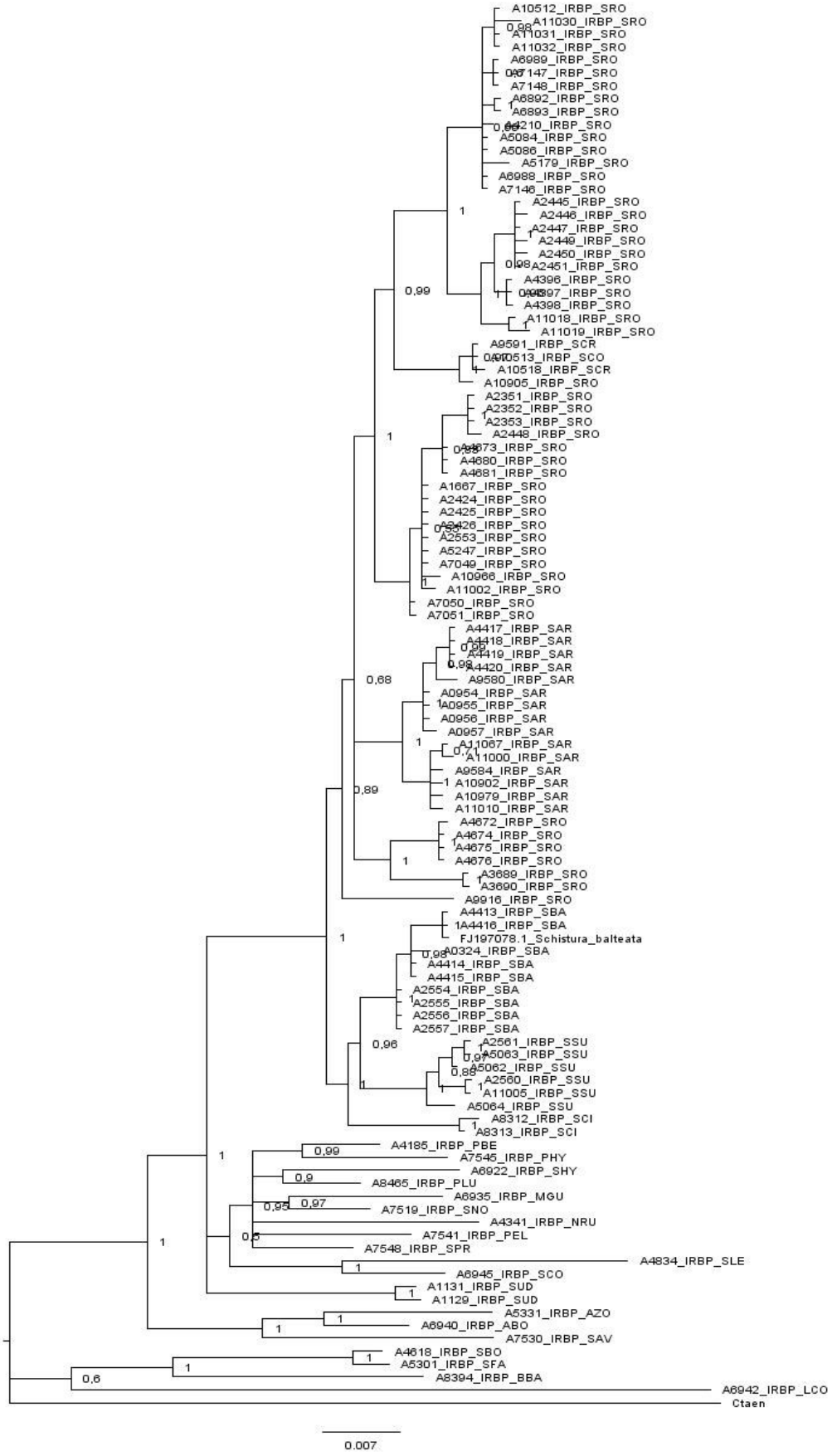
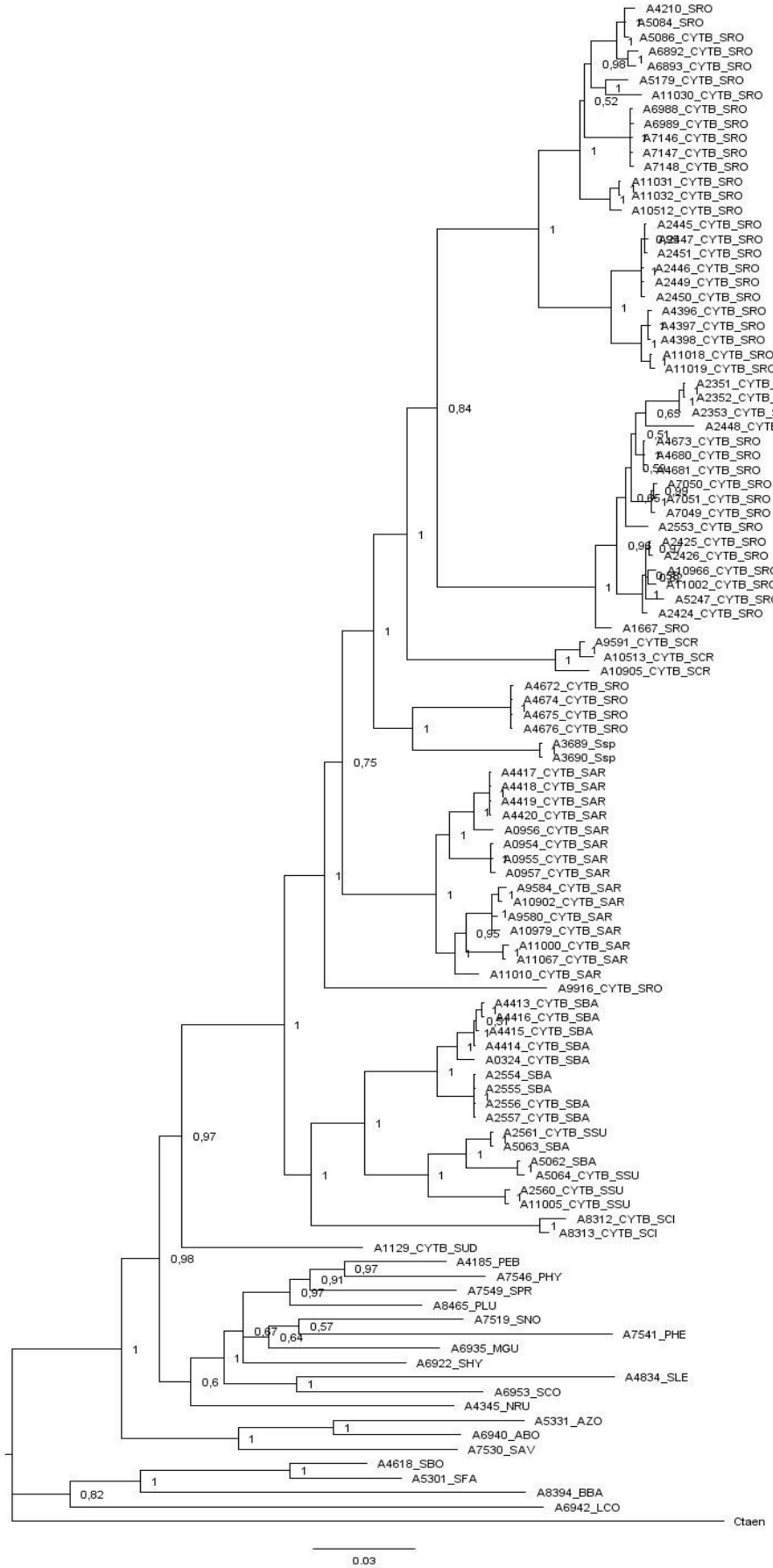


Fig. 14 Phylogenetic tree of *S. robertsi* group from combined dataset



Tab. 23 Comparison of individuals of the second size classes (35 – 39,9 mm SL) of *P. zonalternans* and Takua Pa lineage. SL = standard length, yellow colour = different character without overlap, green colour = tendency with small overlap

35 - 39,9mm SL		P. zonalternans		Takua Pa lineage	
Number of individuals		34		4	
		MIN	MAX	MIN	MAX
mm	standard length	35	39,8	35,8	38,7
%SL	total length	125,3	130,4	125,7	126,8
%SL	dorsal head length	19,3	22,1	20,4	22,0
%SL	lateral head length	21,6	25,1	22,1	23,5
%SL	predorsal length	44,5	48,7	44,4	49,4
%SL	prepelvic length	49,9	56,3	49,1	52,7
%SL	preanus length	71,4	79,0	72,9	74,1
%SL	preanal length	77,5	83,5	75,9	79,1
%SL	head depth (eye)	10,8	13,9	10,3	11,9
%SL	head depth (nape)	13,1	16,5	12,0	13,6
%SL	body depth	16,2	22,3	14,2	16,5
%SL	depth of caudal peduncle	11,2	15,1	10,3	11,8
%SL	length of caudal peduncle	13,4	16,8	16,8	18,2
%SL	snout length	7,3	10,2	7,5	9,3
%SL	head width (at nares)	8,8	10,7	7,8	9,2
%SL	maximum head width	13,4	16,7	12,8	13,9
%SL	body width (dorsal)	10,3	17,0	9,4	11,0
%SL	body width (anal)	6,1	10,2	5,4	7,1
%SL	eye diameter	5,2	7,0	6,5	7,0
%SL	interorbital width	7,1	9,1	6,4	7,3
%SL	height of dorsal fin	8,8	19,1	16,2	17,7
%SL	length of upper caudal lobe	24,4	29,6	24,0	26,0
%SL	length of lower caudal lobe	23,8	28,8	23,5	25,7
%SL	length of median caudal rays	22,2	26,9	20,4	22,5
%SL	depth of anal fin	19,5	23,7	17,2	19,1
%SL	length of pelvic fin	18,2	21,5	16,9	17,9
%SL	length of pectoral fin	21,3	27,6	21,4	23,0

Tab. 24 Comparison of individuals of the third size classes (30 – 34,9 mm SL) of *P. zonalternans* and Takua Pa lineage. SL = standard length, yellow colour = different character without overlap, green colour = tendency with small overlap

30 - 34,9mm SL		P. zonalternans		Takua Pa lineage	
	Number of individuals	37		5	
		MIN	MAX	MIN	MAX
mm	standard length	30	34,9	31,7	34,6
%SL	total length	124,1	132,1	122,3	128,6
%SL	dorsal head length	18,5	24,0	20,1	23,0
%SL	lateral head length	21,4	25,7	22,4	24,6
%SL	predorsal length	44,4	52,4	44,7	48,6
%SL	prepelvic length	46,3	57,2	50,2	52,3
%SL	preanus length	68,1	75,7	71,9	74,6
%SL	preanal length	74,0	79,6	75,7	79,2
%SL	head depth (eye)	10,2	13,8	10,0	11,7
%SL	head depth (nape)	12,5	17,5	11,5	13,9
%SL	body depth	14,0	23,2	13,5	17,3
%SL	depth of caudal peduncle	10,4	15,7	10,0	12,3
%SL	length of caudal peduncle	10,9	17,1	17,4	17,6
%SL	snout length	5,7	10,5	7,8	9,8
%SL	head width (at nares)	8,1	11,0	8,2	9,8
%SL	maximum head width	13,3	17,1	12,7	15,8
%SL	body width (dorsal)	10,4	16,3	9,4	11,7
%SL	body width (anal)	6,9	12,9	5,0	7,6
%SL	eye diameter	5,1	7,3	6,9	7,6
%SL	interorbital width	6,6	10,0	6,9	8,2
%SL	height of dorsal fin	9,9	19,5	16,0	17,6
%SL	length of upper caudal lobe	23,6	30,2	23,0	27,4
%SL	length of lower caudal lobe	23,3	30,2	21,6	27,1
%SL	length of median caudal rays	19,9	27,7	18,8	24,9
%SL	depth of anal fin	20,0	23,8	16,6	20,5
%SL	length of pelvic fin	19,1	22,9	16,0	19,4
%SL	length of pectoral fin	21,0	27,9	20,1	23,4

Fig. 15 Time estimation tree of *P. zonalternans* group, scale bars with the approximate age of nodes

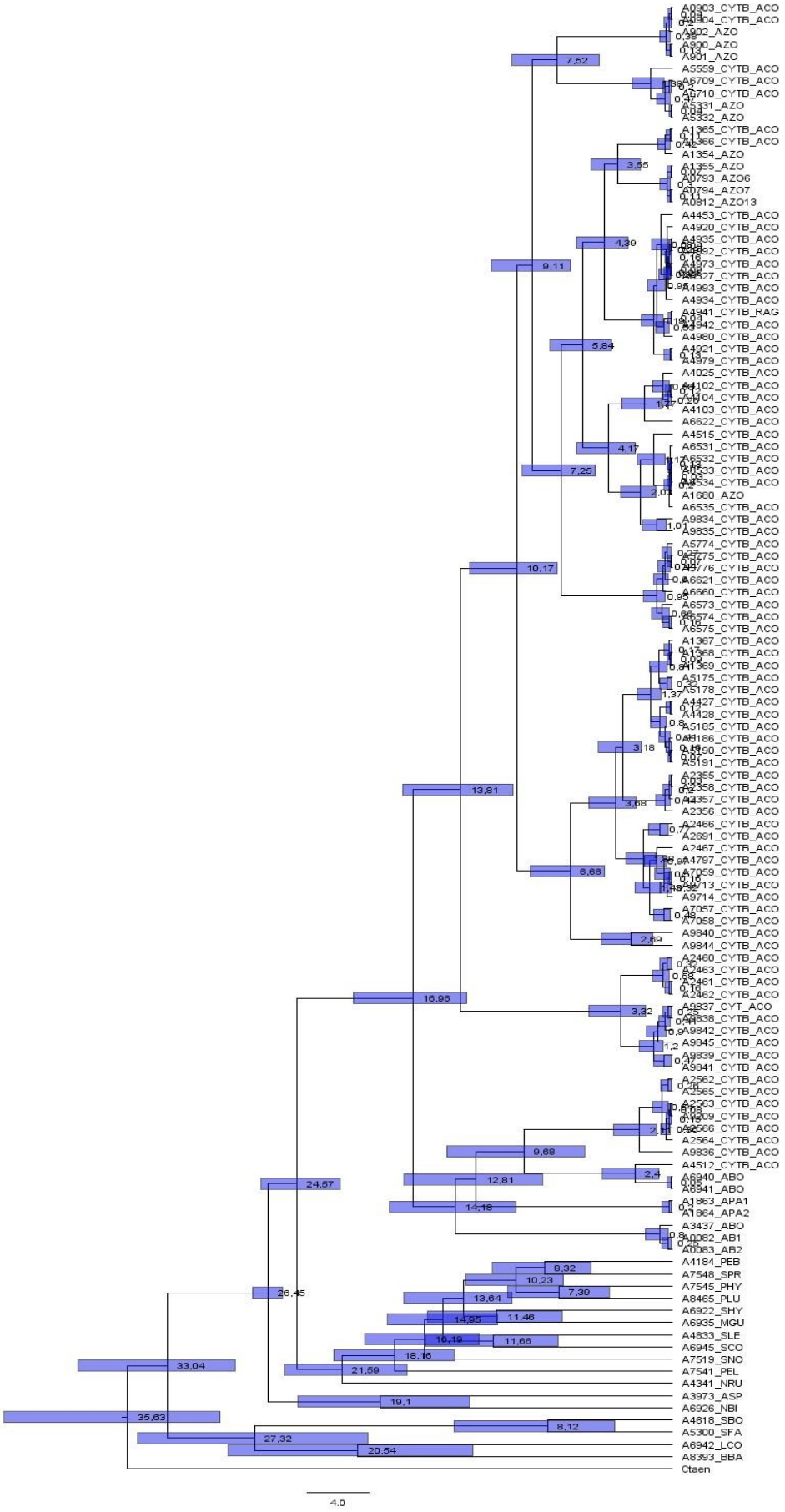


Fig. 16 Time estimation tree of *S. robertsi* group, scale bars with the approximate age of nodes