doi:10.20944/preprints201911.0322.v1

Review of the organismal biology of hill stream loaches.

Jay Willis (corresponding author), Oxford University , Department of Zoology

Theresa Burt De Perera, Oxford University , Department of Zoology

Adrian L. R. Thomas, Oxford University , Department of Zoology

Correspondence to be sent to:

Dr Jay Willis (jkwillis@gmail.com)

Abstract

Hill stream loaches are a group of fish that inhabit fast flowing shallow freshwater. The family has radiated over Asia. For some species their range is limited to single catchments; they provide an excellent example of biogeographical speciation on multiple scales. Hill stream loaches have a range of adaptations which help them exploit environments where competitors and predators would be washed away. They have streamlined bodies and keeled scales reminiscent of Mako sharks and potentially many other as yet undiscovered drag reducing features. They adhere to rocks, crawl over shallow films of water, glide over hard surfaces using ground effects and launch into currents to attack prey or evade predation. They offer a test of modern approaches to organismal biology and a broad range of biomimetic potential. In this paper we analyse what behaviour is associated with their physical adaptations and how this might relate to their evolution and radiation. Our intent here is to review information that is presently available but also to derive new insight through combination and statistical analysis of existing material. We also use our own observations and critically argue several contradictory, or poorly supported assertions in the existing scientific literature. Our conclusion is that, although extensively studied, there remain major gaps in knowledge about their organismal biology. We summarise the research opportunities in the form of hypotheses that are worth testing.

Keywords: loaches; suckers; sucker fish; rheophilic

Introduction

This review is organised as follows: After an explanation of the value of the study, we present an overview of the species and their geographical origins. Then we review phylogeny by combining existing studies with newly available genetic information. There follows information with respect to hydrodynamics of these fish, their scales, other external features and their various modes of locomotion. Colour changes are briefly reviewed. Finally background on the aquarium trade, notes regarding husbandry, and attempts to replicate natural habitat are outlined. The review concludes with a brief discussion and a selection of testable hypotheses.

Figure 1. Hill stream loaches (Sewellia sp.) kept in laboratory conditions. The fish generally remain attached to the substrate but will on occasion launch into the water to swim short distances but always remain close to the rocks.

Novelty of study

There are many rheophilic species across practically all rivers and across all families of freshwater fish (Lujan and Conway 2015). *Balitoridae* and *Gastromyzontidae* (hill stream loaches), however, are remarkable among teleosts as they use their whole body as a hydrodynamic adhesion device (De

Meyer & Geerinckx 2014) (Fig. 1). They do not have spines and odontodes or any other prominent adaptations to provide alternative adhesion; this simplifies the conceptual evolutionary development within hill stream loaches between species which are fully free swimming and those which predominantly adhere to the substrate. This is particularly important with respect to this family of fish because, as we shall review here, the present theories of radiation and speciation suggest that the evolutionary development from free swimming to bed adherent species has happened separately on multiple occasions. In their summary of rheophily in fish Lujan and Conway (2015) say, “Few species or morphological traits associated with rheophilic habitats have been empirically investigated from a functional, performance, or correlated eco-evolutionary perspective” and they go on to explain that their review is thus based predominantly on theory. This is the inspiration of this present study. Rheophilic fish are common in most rivers that are likely to be adversely impacted by human developments and yet there is a dearth of empirical evidence regarding their functional performance, habitat requirements and the behaviour associated with their elaborate physical adaptations.

Species and family radiations and invasions

All the loaches (superfamily *Cobitoidea*) are described by Kottelat (2012). The sucker loaches are now in the families *Balitoridae* and *Gastromyzontidae*. Those genera endemic to Borneo, (*Gastromyzon*, *Pseudogastromyzon* and *Hypergastromyzon*) are described by Tan (2006). For the purposes of the explanation of broad hypotheses about radiation of species it is common to examine two general groups: *Gastromyzonidae* and *Homalopteridae*. They are distributed across Asia in multiple overlapping and isolated groups (Fig. 2). *Gastromyzonidae* (example genera *Gastromyzon*, *Sinogastromyzon*, *Sewellia*, and others) are the most apparently elaborately adapted to fast rivers, with wide pectoral fins in a fixed horizontal position which overlap their equally wide pelvic fins and appear to form a single sucking disk along the majority of their flat ventral surface, they are moderately dorso-ventrally flattened (Fig. 3).

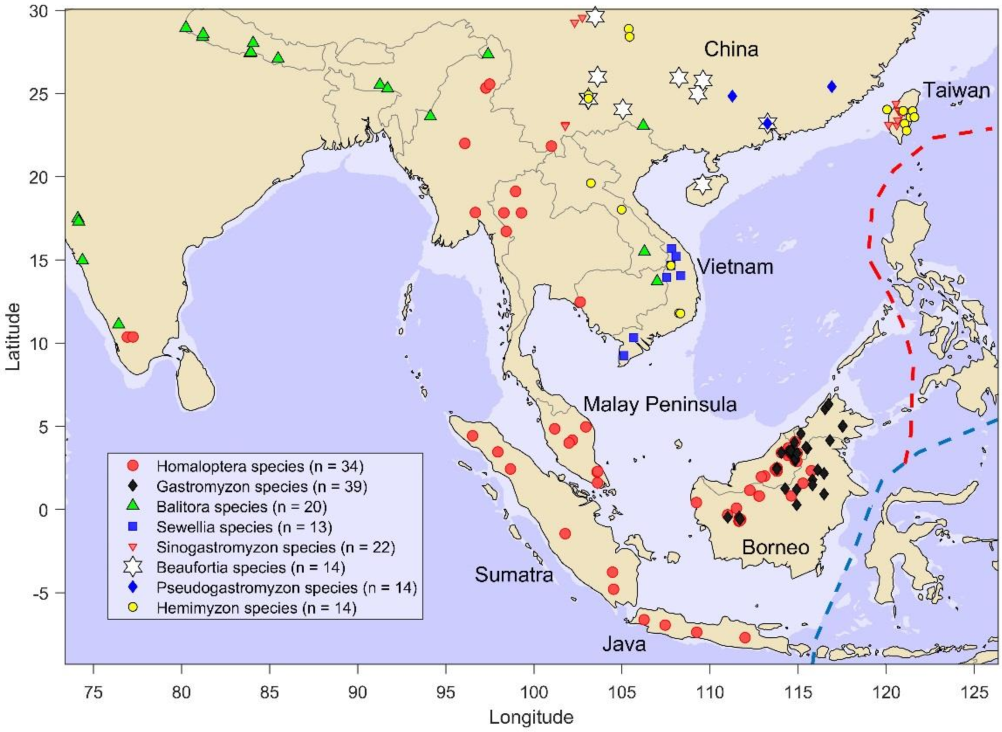


Figure 2. Map showing all locations listed for hill stream loach from www.fishbase.org by genus. Some are the type locations and others record presence, some species have no points associated with them in Fishbase (Froese and Pauly 2000). This map therefore only illustrates the general distributions by main genus (not exhaustive). Points also added from Hora (1952) and Tan (2006) to illustrate the argument that while the Gastromyzonidae (Gastromyzon, Sewellia, Beaufortia etc.) are isolated to Borneo, Taiwan, and Vietnam for instance, the Homalopteridae (Homaloptera and Balitora) are widely distributed, in particular on the Malay Peninsula, Java and Sumatra where no Gastromyzonidae are found. The lighter area in the sea shows the 200m depth contour which roughly coincides with the area that was land (called Sundaland) during the last ice age about 25 ka ago. The blue dashed line is called the Wallace Line and the red dashed line the Huxley Line, which separated two distinct faunal regions (For instance there are no hill stream loaches to the east of either of these lines.) Hora's (1952) theory is that the Homalopteridae were able to radiate through the lowlands in Sundaland, whereas the Gastromyzonidae were restricted to the highlands (fast flows) and thus remained isolated. The South China Sea opened between what is now Vietnam and Borneo about 30 Ma ago and so the theory is that the groups of Gastromyzonidae are the product of parallel evolution on Borneo and mainland Asia, and presumably their body shapes are an example of convergent evolution under the constraints of a life in fast hydrodynamic situations.

Homalopteridae (genera *Homalopteroides*, *Homaloptera*, and several others,) are more cylindrical in shape with more sinuous bodies and their fins are smaller and more flexible and may overlap slightly in some species but not in others (Fig. 3).

Hora (1952) outlined his hypothesis that two groups of Gastromyzonid fishes now separated by the South China Sea descended from a common standard loach type fish. Hora (1952) suggested a shared ancestor similar to *Nemacheilus* genus (stone loach) in the family *Cobitidae* (which is now placed in a family *Nemacheilidae*).

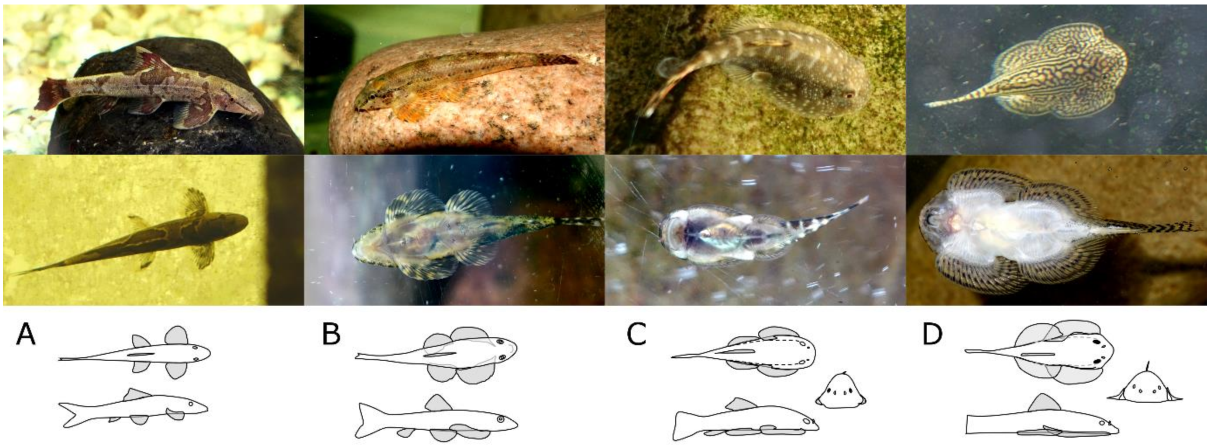


Figure 3. The main groups discussed in this study. Each of four columns shows a species of fish in photographic and cartoon form. Column A: Homaloptera orthogoniata B: Homalopteroides smithi C: Gastromyzon zebrinus D: Sewellia lineolata. The fish are representative of the main groups of hill stream loaches with general level of gross physical adaptation to a sucker loach increasing from left to right. Column A is typical of a lowland specialist Homalopterid and Balitoridae with wide distribution across Asia, Column B is a higher flow highland specialist Homalopterid which is often found in the same water body (albeit in the slower deeper sections) as the more elaborately developed suckers. Column C is a Gastromyzonid *endemic to Borneo, a key unique feature across the Bornean suckers is the very wide mouth. An important hydrodynamic feature is the suprapelvic flaps which are visible in the middle picture of column C as two white triangular shapes between the pelvic and pectoral fins, The structures can extend to the dorsal side just behind the margin of the pectoral fin as a small 'ear shaped' pelvic auxillary flap visible in the top photograph of column C. Column D is the most elaborate genus of sucker loaches, a Sewellia endemic to Vietnam. The frontal views in columns C and D show the limpet like shape of the sucker's cross section at the widest part of their bodies.*

Hora's (1952) idea was that the *Gastromyzonidae* formed two geographically separate groups, one group in Borneo and one in China and Vietnam. These two mountainous areas are similar in that they have abundant small streams; Hora (1952) suggested they may have been joined in the past. The evidence that there was a land bridge, suitable for the radiation of freshwater fish which inhabit lowland streams but not those that only inhabit highland streams, is provided by the present distribution of *Homalopteridae* and other fish. *Homalopteridae* are generalists that often co-habit the same stream sections as *Gastromyzonidae* (Tan 2006), they are present in South China, Vietnam, Taiwan and Borneo but crucially they are also present in India, Sumatra, Java, and the Malay Peninsula where *Gastromyzonidae* are not. Hora's (1952) concept is that *Homalopteridae* were an older group, spread more widely across low lying land bridges that no longer exist. This is broadly supported by detailed physiologically informed phylogeny (Sawada, 1982). Sulaiman and Mayden (2012) restate this thesis in light of recent discoveries and across a broad range of fish adding a suggestion that northern and eastern Borneo is distinct from south and western Borneo which was more closely connected with Sumatra and Java (respectively) due to the drainage of rivers creating lowland freshwater connection on Sundaland (Vorris 2000). This concept of separate parts of Borneo connecting through the lowland rivers at different times in geological history was first outlined by Inger and Chin (2002).

Molecular phylogenic analysis

There has been no specific study to analyse the phylogenic similarities and differences between the *Gastromyzonidae* of Borneo and Vietnam, and their congeners the *Homalopteridae*. We made a phylogenic tree from existing publically available gene sequences of the *Balitoridae* (Fig 4). We were restricted by the material available and its genetic overlap. The concept of several clades of hill stream loaches is confirmed. The three main groups being 1) the *Gastromyzonidae* of Borneo, 2) the *Sewellia/Beaufortia* group of China and Vietnam, and 3) the *SinoGastromyzonidae* of China and Taiwan. *Gastromyzon lepidogaster* is consistently positioned as developing at an earlier stage than other *Gastromyzonidae* of Borneo; this is interesting with respect to its geographic distribution (discussed following). The relative position of the *Homalopteridae, Barbatula* and other lowland inhabiting species and genera to the other groups remains unreliable (see Fig. 4 likelihood values).

This new molecular phylogenic treatment does perhaps provide a solution to a long running question in the phylogeny of the *Sinogstromyzon* and *Homaloperoides smithi*. Sawada was unable to determine through physiological comparisons of bone structure the branching patterns of *Homaloptera* (now *Homaloperoides*) *smithi*, *Hemimyzon formosanum*, and *Sinogastromyzon puliensi*s, and suggested three variant trees in his figure 99 (Sawada 1982). Of Sawada's (1982) three variants our molecular analysis here supports option H (Fig. 4).

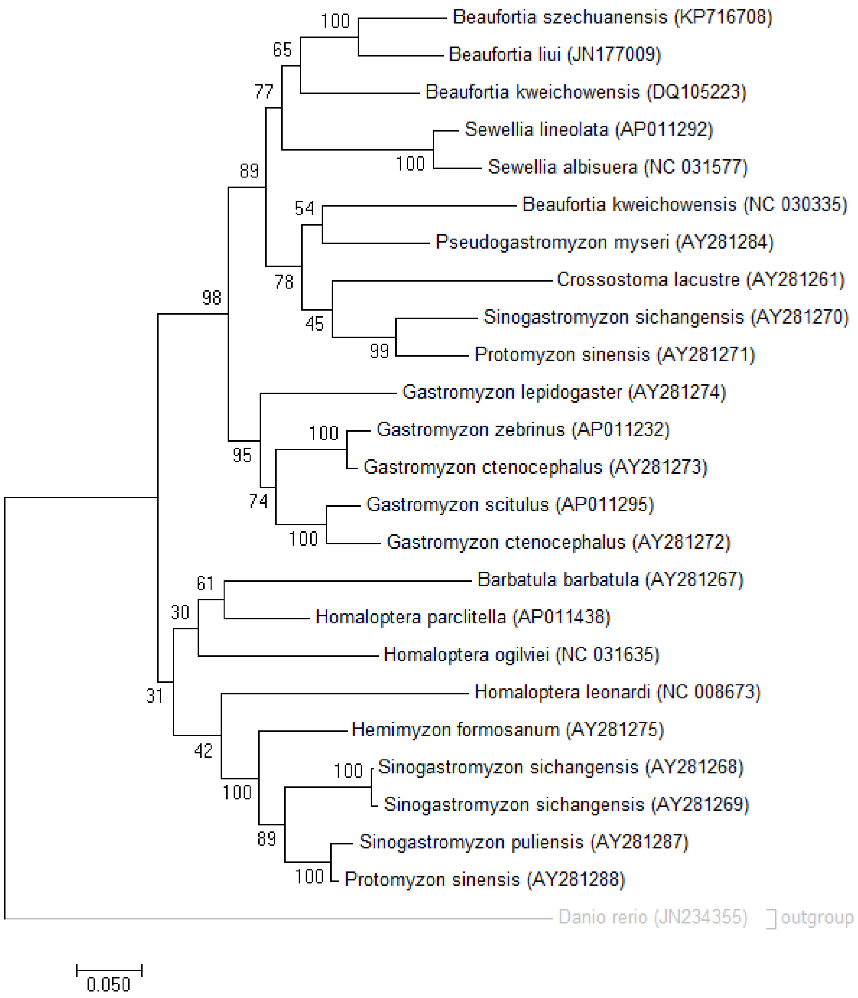


Figure 4. Molecular Phylogenetic analysis by Maximum Likelihood method from cytochrome b mitochondrial gene. Derived from Genbank, originally based on the popset 34733250 (Wang and Tzeng 2016) with additions and deletions dependent on material available in Genbank (accession numbers shown beside taxa) – duplicate species are not shown unless separated in the tree. Models (Time Reversible model) and methods selected by Bayesian Information Criteria from all available in MEGA7, aligned with MUSCLE. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. Further explanation of methods in supplementary section. A key finding is that Gastromyzon lepidogaster is reliably branched earlier from the other Gastromyzonidae of Borneo which may correlate to its position as an outlier in terms of the comparatively high number of catchments it inhabits (see main text). Also Gastromyzonidae (endemic to Borneo) are further distanced from SinoGastromyzonidae than some Homaloptera. Homaloptera are shaped quite differently from either of the two reasonably similar looking clades of sucker loaches.

This general concept of connectivity for lowland fish remains supported through the distribution and genetic consistency of *Homalopteridae*; *Barbucca diabolica* and *Balitoropsis zollingeri*. *Barbucca diabolica* is a loach physically similar and closely genetically related to *Homaloptera* genus which has existing populations in Borneo, Sumatra, The Malay Peninsula and Thailand (Šlechtová et al 2007). Balitoropsis zollingeri (very common and well known synonym: *Homaloptera zollingeri*) (Randall and

Riggs 2015) also exists in Borneo, Sumatra, Java, Malay Peninsula and Thailand with only minor differentiation between the species (now potentially a monophyletic genus) (Randall and Riggs 2015). This implies the *Gastromyzonidae* originated at different times in different places, well after the principle origins were geographically separate, from a common and mutable *Homalopterid* ancestor that inhabits low and high land streams.

Catchment scale radiation

Since the 1970’s these fish have entered the aquarium trade and we cannot overlook recent human movement of fish over long distances for trade, and previously for curiosity, food or other reasons. Tan (2006) mentions how one fisher returns small (juvenile) specimens in Borneo but not to the same stretch of river from where they were caught, presumably to intentionally increase their range. The financial incentives are evident (Tan 2006). Local people in Borneo have also eaten the *Gastromyzonidae* (Tan 2006) presumably since they (the people) have been resident (~40 ka ago) (Phillipps and Phillipps 2016) and there is evidence of other hominids in Borneo ~235 ka years ago from a stone tool factory (Phillipps and Phillipps 2016). There are also non-hominid candidates for egg distribution between catchments such as macaques (especially the southern pig-tailed macaque (*Macaca nemestrina*) and the crab-eating macaque (*Macaca fascicularis*)) which both are attracted to water and are omnivorous (Payne et al 1998). Macaques are co-habitant in all south eastern Asian areas with hill stream loaches (Abegg and Thierry 2002). But they might just share similar biogeological constraints and not come into contact, and there are many other potential animal vectors. Finally Tan (2006) mentioned the potential of a purely geological mechanism of river capture where the headwaters from one catchment may be diverted by erosion or other earth or rock movements. Small scale distributional differences between species are crucial in terms of hydrodynamic specialisation because; the higher up a catchment a fish climbs, the closer it is to an adjacent catchment.

Methods of analysis of catchment scale radiation

Tan (2006) provides a series of maps of distribution sites for the *Gastromyzonidae* of Borneo and Freyhof (2003) provides a similar map for the species of the *Sewellia* genus of Vietnam. We digitised these maps and scaled the images so that various numerical information could be derived for each species (Supplementary information Table S1). Each map depicted a number of sites at which each species had been collected. The maximum distance between two sites of the same species we defined as the range, the area of the minimum sized polygon that contained all the sites for a single species we defined as the area. We also counted the number of different catchments on which the sites were located. We used the riverine information on the supplied maps and counted a catchment as a separate coastal entry point. We also counted the catchments between inhabited catchments for each species using the depiction of catchments provided on the maps.

Results of catchment scale radiation

Gastromyzonidae information derived from Tans (2006) maps suggests that the distribution of ranges (with the addition of a constant term to model species with a single site (50 km)) between species is a reasonable approximation to a lognormal distribution (Fig. 5) (N = 42, mean = 172, standard deviation = 176, log likelihood = -250). The area data were similarly distributed (log likelihood = -269, in this case a constant of 2500 km^2 was added to account for single sites). The number of catchments inhabited by each species was reasonably well correlated to range ($r^2 = 0.74$, n = 42, $p < 0.001$). Range was correlated to area ($r^2 = 0.83$, n = 42,$p < 0.001$) and the square of range was slightly better correlated to area ($r^2 = 0.87$, n = 42, $p < 0.001$) as would be expected in each case. The distribution of catchments poorly fitted a Poisson distribution (N = 42, mean = 1.98, log likelihood = -72, Chi-squared test rejected fit at 95% significance) and there were two notable exceptions from a good fit in terms of a high number of singletons and a single large outlier. *Sewellia* species data from the map in Freyhof (2003) is similar for both analyses; range was again poorly fitted by a lognormal distribution (N = 8, mean = 102, standard deviation = 67, log likelihood = -42), and catchments fitted poorly by a Poisson distribution (N = 8, mean = 2.5, log likelihood -17, Chi-squared test rejected fit at

95% significance). The *Sewellia* species and Bornean Gastromyzonid groups were remarkably similar in their fitted distributions and the areas of poor fit in the case of the catchment data were similar. The two groups are statistically indistinguishable using either range or catchment.

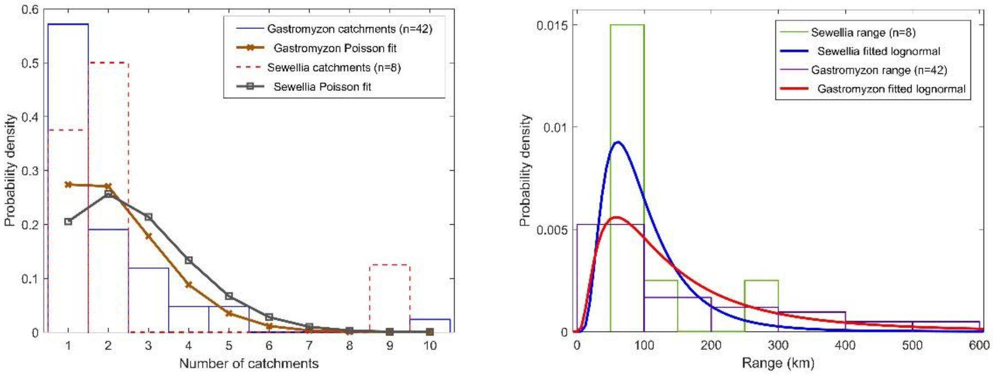


Figure 5. Probability density distributions for ranges and number of inhabited catchments for Gastromyzonidae of Borneo derived from the maps in Tan (2006) and for the Sewellia of Vietnam (Freyhof 2003). The left panel shows that the number of catchments inhabited by each species is reasonably well modelled by a Poisson distribution but that the number of singletons is higher than would be expected from a pure Poisson distribution. A pure Poisson distribution would suggest that there is a roughly equal probability for each species to attain a new catchment within a specified time period. The right panel shows that the ranges are reasonably well modelled by a lognormal distribution which suggests that the species and places of inhabitation are reasonably similar in terms of their impact on ranges. These range data suggest that the ranges attained by the species of fish are each dependent on a similar series of random factors (multiplicative normal probabilities) and that the species and their habitats are potentially well modelled as independent and similar entities in terms of their attainment of range. The similarity between the fitted Poisson distributions for the groups in Vietnam and Borneo is striking but there is little data for both, both fits are poor, and therefore it would not be wise to draw firm conclusions from these data alone.

Discussion of catchment scale radiation

A pure Poisson distribution, if followed, would suggest that a model would be justified in which there is a roughly equal probability for each species to attain a new catchment in a given time period. Poor fit at both extremes of the distribution means it is unlikely that any mono-modal probability distribution would fit better that the Poisson. Thus we can firmly reject the hypothesis that the distribution of catchments is randomly distributed in a way which suggests that there is a roughly constant chance of jumping catchments for all fish and that they have all been there for roughly the same time. It is possible to envisage an independent mechanism as say the transport of eggs on the fur of animals which exist in roughly constant abundance throughout the area. Whereas an alternative hypothesis, which is favoured by this analysis, is that the species are different in their distribution abilities – either they are better adapted to get higher in any given catchment, or their eggs are better designed to be carried by certain animals. But there are other possible hypotheses such as differing geography (i.e. some whole groups of catchment gaps easier to jump than others) and the possibility that all the species have not been resident for a similar length of time. Species may have been resident for less time, for instance, if hybrids are regularly formed after a catchment jump. Finally the other groups of causations that might account for the non-random distribution are those involving a lack of independence between species and inconsistent data collection; for instance if species compete and make each other extinct in a catchment, if they hybridise (as mentioned above) and if they genetically diverge to form new species through adaptation or genetic drift. Observations may be biased by many factors such as physical access, or identification familiarity, but since we have used two sets of data that were collated and established independently and consistently, but without any prior intention of the analysis here, biases related to selection of data are independent to our analysis. Nevertheless, caution is warranted due to the paucity of data, especially in the *Sewellia* case.

The lack of a good fit to a random model is also interesting due to its possible divergence and the counter implications for sub-groups or geography. For instance it would be interesting to focus on the *Sewellia lineolata* and the *Gastromyzon lepidogaster*, together and in comparison with their sister genera, as they are both outliers and are both abnormally prolific in catchment attainment. They are both the largest of their families and share anatomical adaptation to flow (see following sections). It may be significant that in the phylogenic rapid assessment (Fig. 4) the *Gastromyzon lepidogaster* is consistently branched earlier in the *Gastromyzonidae* than any others. The conformance of range to a lognormal distribution is a possible expression of these same effects and root causes in a continuous variable. (The catchment data are discrete and thus a Poisson distribution is appropriate).

Other sources of distribution data are publically available. The Global Biodiversity Information Facility (GBIF) has 171 records of location of collection of 17 species of Gastromyzon on Borneo for instance (GBIF 2017). These data were derived from 7 different Museum collections and other databases such as FishBase and the European Molecular Biology Laboratory. One major advantage of these data is that the locations are provided in precise digital format and thus further computational analysis is convenient. The data were superficially similar to the records in Tan (2006) albeit with fewer species represented and many instances of multiple individuals collected from a single site. There certainly were not any data that contradicts our tentative hypotheses derived from Tan (2006). We decided against incorporating these data in this analysis because we were unsure if the taxonomy had been as thoroughly contemporarily reviewed as Tan (2006). The taxonomy of these species has been changing and expanding rapidly over the past 20-30 years and many of the data on GBIF had apparently been input from records originating in the 1950's. We felt that for simplicity and reproducibility it would be more helpful if we concentrated on the two most thorough and consistent contemporary studies and the information contained explicitly in these. However this not to say that the GBIF is not an excellent and convenient resource and we feel that it is likely to be the way in which this information is provided in the future. We would advocate that Tan's (2006) data are uploaded to GBIF.

Natural habitat of hill stream loaches

General descriptions of habitat indicate fast flowing small rivers, about 10 m to 20 m wide, high in catchments, with average water velocities of between 0.5 ms^{-1} and 2 ms^{-1}. Depths are shallow around 0.2 m to 0.6 m, very clear water (Inger and Chin 2002), cool temperatures, high oxygen, abundant algae, little or no macro fauna in water, but much riparian vegetation on banks, and preferred substrate always rocky and often variable with boulders, rocks and pockets of gravel and sand (Tan 2006, Yu and Lee 2002). In mixed flow rivers (riffle, run, pool) hill stream loaches are riffle specialists (Martin-Smith 1998, Tan 2006). Some species exclusively graze algae and eat algae and the animals that live in it (predominantly *crustacea* and diatoms) (Tan 2006), others supplement an algal diet with striking at animals drifting, swimming or themselves attached to rocks. Their predators are unknown. Inger and Chin (2002) suggest otters, snakes and several species of bird (auhingas, egrets and kingfishers) as likely predators of all fish in Borneo, although otters are rare in Borneo (Payne et al 1998). There is one report of a *Gastromyzonid* (most likely a *Sinogastromyzon* sp. From Taiwan) in the stomach of a perch known for eating smaller fish (*Siniperca*) (Hora 1950). *Siniperca* are known to inhabit fast moving highland streams but stay mainly in the deeper pools (Hora 1950). Hill stream loaches may have adapted to life in shallow fast moving water as a defensive strategy; a micro habitat selection considering other fish such as the *Siniperca sp.* can occupy the same wider habitat in small rivers with deeper sections of fast water. *Gastromyzonidae* in Borneo are consistently only associated with riffles in Sabah (Borneo) whereas a higher diversity of other species are associated with pools, runs or ubiquitously with multiple zones (Martin-Smith 1998). Martin-Smith (1998) highlights the importance of micro-habitat zones across 14 different river sites with a wide range of water velocities and substrates. However our field observations suggest that larger adult *Gastromyzon lepidogaster* only inhabit clear, slow moving pools, under waterfalls, and thus contradict the observations of Marytin-Smith (1998) while juveniles were found in riffles in more mixed sections of rivers (Fig. 6). *Homalopterid* congeners in that study were associated with multiple zones, including riffles (Martin-Smith 1998). Another possibility is that some hill stream loaches are uniquely adapted to scale barriers in smaller rivers; once in a section that is not accessible for predators or competitors they then inhabit all areas of that section – so they only occasionally use their extreme adaptations.

Figure 6. A large adult Gastromyzon lepidogaster (~0.08 m Total Length (TL)) in a slow moving clear pool at a depth of about 0.2 m (left). The pool was 20 m downstream of a waterfall and over 2 m deep in places. Inset is a picture of juvenile Gastromyzon lepidogaster (~0.05 m TL) which was caught in a riffle about 2 km downstream of the waterfall. The other fish is of an unknown species, but is an open water fish unlikely to have the same station holding and climbing abilities of hill stream loaches. Photos taken in upper Danum Valley Sabah, Borneo by author (JW). This contradicts earlier literature which suggests hill stream loaches are only found in riffles.

In Taiwan, the work of Yu and Lee (2002) provides thorough habitat preferences for endemic hill stream loaches (*Sinogastromyzon puliensis*) (sampling at over 60 sites), and Wu and Wang (2002) examine substrate preferences and associated water velocities. Yu and Lee (2002) describe small rocky streams high in catchments typified by extreme variation in currents and depths (they found maximum preference for around 0.6 m depth with preferred velocities around 1.6 m s^{-1}) whereas Wu and Wang (2002) report a preference of water velocity of 0.4 ms^{-1} (max 1.5 ms^{-1}) and depth of 0.35 m (max 0.7 m) from a three year survey in a single catchment performed by the Taiwan Endemic Species Research Centre specifically to identify habitat preferences of this species. Yu and Lee (2002) found a broad range of preference for both velocity (roughly between 0.8 ms^{-1} and 2 ms^{-1}) and depth 0.3 m to 0.6m, with no collections made > 0.9 m depth, and observations declining sharply at velocities in excess of 1.8 ms^{-1}. They observed highest preference of substrate for small rocks, and similar but lower acceptance for pebbles and big rocks. Yu and Lee (2002) conclude that habitat preference is co-dependent on both velocity and depth. Substrate sizes are also correlated to velocity and depth so it is worth considering if they observed a derived preference for a particular substrate type or vice versa. It is interesting to note that the maximum water speed where small cobbles are likely to remain in position is about 2 ms^{-1}(Giller and Malmqvist 1998 (Table 3.1 p. 39)) so if fish were targeting small cobbles, their preference for water velocity would also be limited accordingly. The association between substrate size, density, shape, water speed and depth was characterised by Shields (1936) in the form of Shield's curves. Wu and Wang (2002) suggested that *Sinogastromyzon puliensis* had an increasing preference for larger substrate increasing monotonically all the way to solid rock substrate.

The association between *Gastromyzonidae* and *Homalopteridae* and substrate, water velocity and depth is also confirmed in Borneo (Martin-Smith 1998). Yu and Lee (2002) also noted a correlation to water conductivity but since they were using low voltage electrofishing this might be an artefact of the method rather than any preference of the fish. A plausible working hypothesis is that these preferences correlate with habitat preferences of stream invertebrates. A study, albeit in a completely unrelated river system, found a positive correlation between invertebrate species diversity and current velocity where velocities of 0.6 ms^{-1} to 0.8 ms^{-1} contained the greatest density of faunal preference whereas highest depth preferences ranged between 0.05 m and 0.6 m (Degani et al. 1993). In another study, of 225 sites again in a completely unrelated river system, the conditions of highest faunal diversity were 0.75–1.25 ms^{-1} current velocity at 0.2 m to 0.4 m depth and the optimum condition appeared to be 0.76 ms^{-1} at a depth of 0.28 m over medium cobble substrates (Gore 1978), which is remarkably similar to the positioning of fish described above. It is reasonable to assume *Sinogastromyzon puliensis* preferences were primarily related to prey preferences which were themselves related to floral abundances and ultimately all these are correlated to current and depth.

Hydrodynamics of body shape

The literature is typically not detailed or robust in terms of reports of the shape of the fish with respect to hydrodynamic performance or the potential reasons for the shape. Here it is necessary to augment a literature review with some new measurements to identify the range of body shape within the groups of hill stream loaches.

Firstly with respect to the naïve assumption that the shape of hill stream loaches is determined by their requirement to avoid being swept away by strong currents. The standard comparative example for an animal that adheres to rocks in fast moving sheets of water is the marine limpet (Hora (1930) describes the progressive development of torrential fauna toward the limpet shape). The marine limpet (*Lottia gigantea*) superficially resembles hill stream loaches from certain angles (anterior cross section of *Sewellia* sp. for instance). Limpets adhere to rocks in water flows of up to 20 ms^{-1} (perhaps 10 times faster than the normal hill stream loaches endures). However, the limpet is not shaped in the way it is to avoid dislodgement by hydrodynamic forces such as lift and drag. In fact, a limpet is capable of such strong adhesion to the rock that it effectively escapes from the danger of dislodgement due to hydrodynamic forces, regardless of the shape of its shell (Denny and Blanchette 2000). It is shaped in a sub-optimal way from a hydrodynamic perspective to facilitate its aggressive territorial behaviour (Denny and Blanchette 2000). Extreme capabilities are only relevant in terms of the behaviour to which they are adapted. We should therefore assume that dislodgement due to extreme flow situations to which hill stream loaches have voluntarily moved themselves is an unlikely adaptive constraint.

With respect for the static (non-swimming) body shape there appears to be contradiction in the scientific literature with respect to the theoretical hydrodynamics of these fish; De Meyer and Geerinckx (2014) explain “The abruptly rising head profile, … causes a strong downward hydrodynamic force: the body is pressed against the substrate when facing the water flow.” Whereas Vogel (1994) criticised Steinmann (1907 cited in Vogel 1994) for a similar explanation of the shape of insects that inhabit fast water; “Part of the problem was that Steinmann viewed the flow as pushing the flattened insect down against the substratum by striking its inclined, upstream facing surface.” Vogel (1994) goes on to explain how positive rather than negative lift will ordinarily be developed in this type of case, and that furthermore, flow separation and its position along the body will also determine lift and downforce balance (this indicates different functional effects of posterior fins and head not picked up by De Meyer and Geerinckx (2014)). Vogel's view is supported for animals of a similar size to our fish by Denny and Blanchette (2000) in a wide range of currents. Vogel's view of lift being more of an issue than downforce seems more likely to be true (based on reference and hydrodynamic theory) than De Meyer and Geerrinckx's (2014) opinion, which appears to be based on intuition. One may also question the need for the major anatomical structure of the body sucker, if downforce could be generated by head shape alone. Downforce may incidentally be a problem for a fish wishing to lift off to strike prey, or to avoid predation. However, we notice that dead fish of these species appear to remain attached to a smooth flat substrate even in reasonably quick flows (~0.5 ms^{-1}), and return to attachment if lifted off and dropped in mid-water. So the situation is not as clear cut as either opinion above. We are left unsure if the body shape of these fish creates lift, downforce, both or neither, in water flows of any speed, in mid water, or on or near the bed. Clearly neither theory nor intuition is adequate for this most simple static case, let alone for the situation when the fish moves or changes its fin positions or body shape. This other important aspect, long overlooked in all these assessments, is the ability of these fish to change their hydrodynamic profile, for instance a *Homalopteroides smithi* takes up different body poses, in exactly similar water conditions, dependent on its intention to strike (Fig. 7)

Figure 7. Two images of a Homalopteroides smithi taken several seconds apart in our aquarium. The pose on the left was maintained for several minutes before food was introduced into the tank. The fish adopted the pose on the right when the food entered the tank and struck at the food several seconds later. The water current was about 0.2-0.3 ms^{-1} moving from right to left in these images. The fish habitually face upward into currents. Prior to striking the fish stiffens its body and expands its caudal fin. The change in pose would change the hydrodynamic profile of the fish and cause water to flow in different patterns around its body; presumably the striking pose reduces suction onto the rock to allow a quick release, whereas previously suction would be a benefit for stability of attachment. This is an example of many different situations where hill stream loaches change their static hydrodynamic profile in the water currents.

Another point very often (almost universally) stated throughout the literature without question or explanation is the fact that hill stream loaches are dorso-ventrally flattened to reduce drag in fast water (Hora 1930, Hora 1950 and many others later), and a related fact that they have an enlarged ventral area to aid adhesion by enhanced friction (De Meyer and Geerinckx 2014). In fact they are not as flattened as other fish which are not assumed to need extreme hydrodynamic capabilities (marine flatfish such as plaice (*Pleuronectes platessa)* maximum height to standard length from photographs in Brainerd et al (1997) = 0.11 whereas for *Gastromyzon praestans* identification photograph (Tan 2006) = 0.20, indicating *Gastromyzonidae* are roughly half as ventro-dorsally flattened as plaice). However marine flatfish often inhabit shelf and estuarine areas where currents can be strong, and as juveniles (around the same dimensions as the fish in this study (0.04 m SL)) some species (European flounder (*Platichthys flesus*) for instance) climb rivers by means of selective tidal stream transport and face mid-water currents of around ~1 ms^{-1} (Colclough et al. 2002). Plaice (*Pleuronectes platessa)* can sustain currents of around 0.5-0.7 ms^{-1} in a flat based flume, of 0.3 m depth, generally correlated to fish lengths from 0.02 m to 0.3 m (Arnold 1969). Like hill stream loaches flat fish prefer to face upstream. Plaice eject water from under their body to remain attached by using their dorsal peripheral fins (Arnold 1969). Hill stream loaches can apparently sustain current speeds up to twice as high (Tan 2006) and are not so flattened as plaice which clearly questions the simple relationship of dorso-ventral flattening for life in fast currents.

Another directly related and interesting parallel between flat fish and hill stream loaches is the need for fast starts. The main problem for flat fish is overcoming Stefan adhesion (Brainerd et al. 1997). – which is the adhesion caused when a flat plate is moved off a surface and a pressure drop is developed between the objects being separated as fluid has to enter the space between the two objects. Stefan adhesion force is highly sensitive to the size of the objects being separated (fourth power) and to the separation (third order inverse power) and effective only during rapid separation (Brainerd et al. 1997). Plaice generally breathe using only their upper operculum, but during a fast start they pump water strongly through their lower (blind side) operculum and peel off their body head first in a wave like motion to almost entirely overcome Stefan adhesion (Brainerd et al. 1997). Thus since hill stream loaches need to detach very quickly for feeding and protection, it is again not entirely clear if a wide flat base is advantageous and required only (or at all) for adhesion and if it could be an impediment to attack or flight. It is further not clear if they have special adaptations for fast starts to overcome Stefan adhesion which would be an interesting case of convergent evolution considering their very different skeletal arrangement to marine flat fish. The only orifice which is on the ventral surface is the mouth in hill stream loaches and that is usually very close to the anterior edge so jetting like the plaice described above seems unlikely although a distinct possibility. Perhaps the moderate dorso-ventral flattening of hill stream loaches is primarily to facilitate conformance to a curved substrate (of smooth spherical boulders and rocks) and to allow for arching head rises to overcome Stefan adhesion. Pelagic osseous fish are usually not flexible in the plane perpendicular to their normal swimming action (lateral flexibility) (See Aleyev 2012 for detailed observations of fish flexibility). Again there is a clear scope for experimental observations about basic body shape.

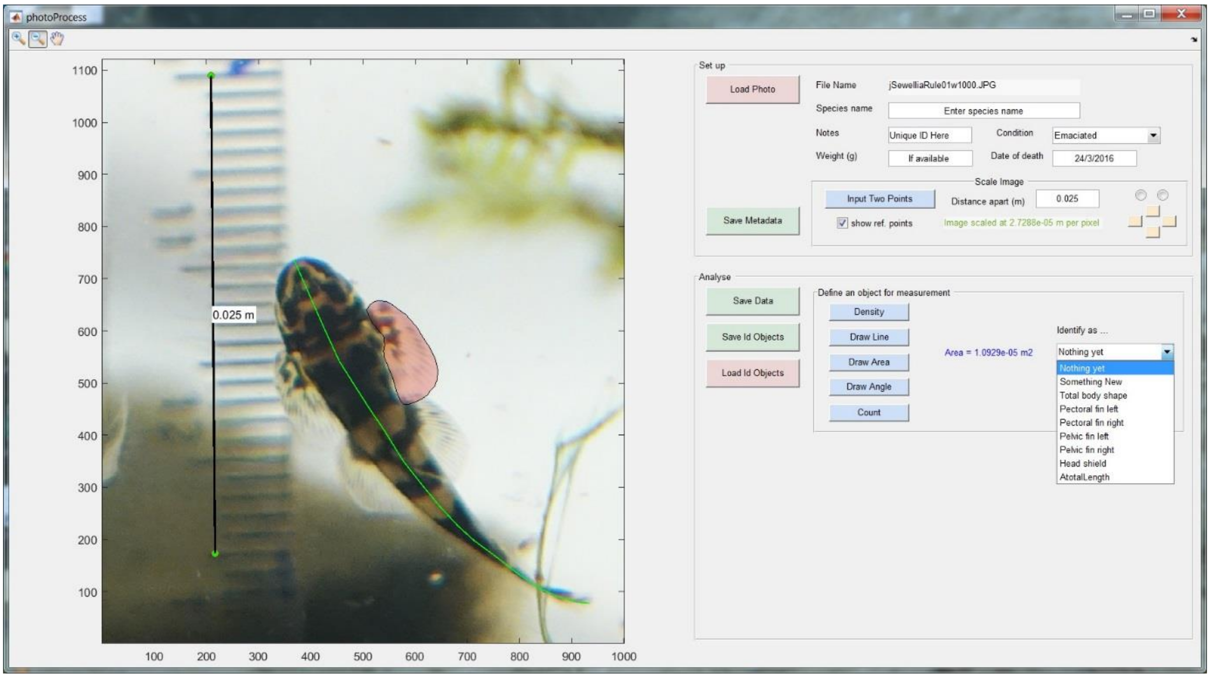
Introduction to hill stream loach body shape

The primary assumptions are that hill stream loaches are adapted to 1) minimise drag (skin and form), 2) maximise surface contact with rocks for physical suction and for downforce generation. The aim is to differentiate 4 species of hill stream loach in respect of physical hydrodynamic based metrics using the above assumptions and to use this as a basis for a wider classification of hill stream loaches, contrasting this with their behaviour and habitat selection.

Form drag in a fish is dependent on the body lateral cross section perpendicular to standard orientation in flow. A suitable dimensionless statistic to compare body shapes for adaptation to form drag is the ratio of cross-sectional width to overall length. Skin drag will be dependent on surface area exposed to flow, and potentially would be differentiated by a similar statistic to form drag (area to length etc.). We examined the ratio of body area (dorsal cross sectional area as proxy) to length, with the limits representing either a straight line with close to zero area-length ratio, or a maximum area-to-length ratio with a cross sectional shape of a circle (assumed half sphere in 3D). The statistic was body area divided by half-length squared in order to constrain the value between 0 and 1. Thirdly we looked at fin area alone. In these fish this area may be related to suction potential, downforce creation from flow, and frictional resistance to drag using a hard surface. Following the same process to derive a dimensionless statistic bounded by 0 and 1, as above, we used fin area divided by half-length squared.

Methods of hill stream loach body shape

A Matlab based Graphical User Interface (GUI) was made to extract data from photographs of specimens. The GUI allows for lines, areas, and angles to be drawn onto an image. The image can be scaled against a reference distance in the photo (fig. 8). The subsequent named measures of distance, area and angle are output in a standard format (CSV, Excel etc.).



File Name jSewelliaRule01w1000.JPG
Species name Enter species name
Notes Unique ID Here
Condition Emaciated
Weight (g) If available
Date of death 24/3/2016
Distance apart (m) 0.025
☑ show ref. points
Image scaled at 2.7288e-05 m per pixel
Area = 1.0929e-05 m2
Identify as ... Nothing yet

Figure 8. Example of GUI in use with a photograph of a live juvenile Sewellia sp. The thick black line on the image is a scaling line between the two green circles. The green line is input to measure total length, and the red area is input to measure right pectoral fin area. The code for this GUI is available in supplementary material.

Fish from 4 species of hill stream loach were measured. *Sewellia lineolata* (n = 3 + 1 juvenile), *Homaloptera smithi* (n = 9), *Gastromyzon sp.* (n = 5) and *Balitoropsis zollingeri* (n = 5). Twenty two sets of measurements were taken from photos of dead specimens including areas of body and fins and linear dimensions of bodies, and one set of measurements was taken from a live juvenile specimen.

The following measurements were taken (all from dorsal aspect):

1. Body cross sectional area

2. Total body length (including tail)

3. Body width at anterior union of pectoral fin

4. Body Width at posterior union of pectoral fin

6. Body Width at posterior union of pelvic fin

7. Areas of 4 individual (non-overlapping) pectoral and pelvic fins.

The derived statistics were:

1) Bluntness of body (body maximum width / total length) assumed to vary between 0 and 1. Where 0; a line and 1; a circular shape. Body width at anterior pectoral fin union was a proxy used as maximum body width.

2) Body shape (body cross-sectional area / half total length squared) assumed to vary between 0; a line, and 1; a circular shape.

3) Fin contact area (fin area / half total length squared) constrained by 0; a line with no ground contact, and 1; fins arranged in a circular structure with diameter equal to body length.

The statistics were plotted as histograms and fitted with beta distributions (bounded by 0 and 1). The statistics were also plotted as 2 and 3 dimensional scatter plots. The derived statistics (referred to hereafter as dimensions) distinguished between the 4 groups of fish in alternative ways (See S3 Supplimentary information). Once the least number of dimensions required to make reliable classification of species were identified the sample group could be extended to any available material – such as photographs published in the internet, published in books or other scientific papers or in our existing library. We used 23 additional data to test the classifier.

Results of hill stream loach body shape

The three statistics used in the initial classifer performed well as a classifier with the material derived from other sources (Fig. 9). Since a mono-modal bi-variate normal distribution (any ellipsoid shape) on the graph in Figure 9 could exclusively contain each whole group without overlap, no further probabilistic treatment is necessary, and the groups are statistically distinct. If more data are collected to the point where this is not the case, then distributions would need to be fitted to calculate relative probabilities of membership of one group or another.

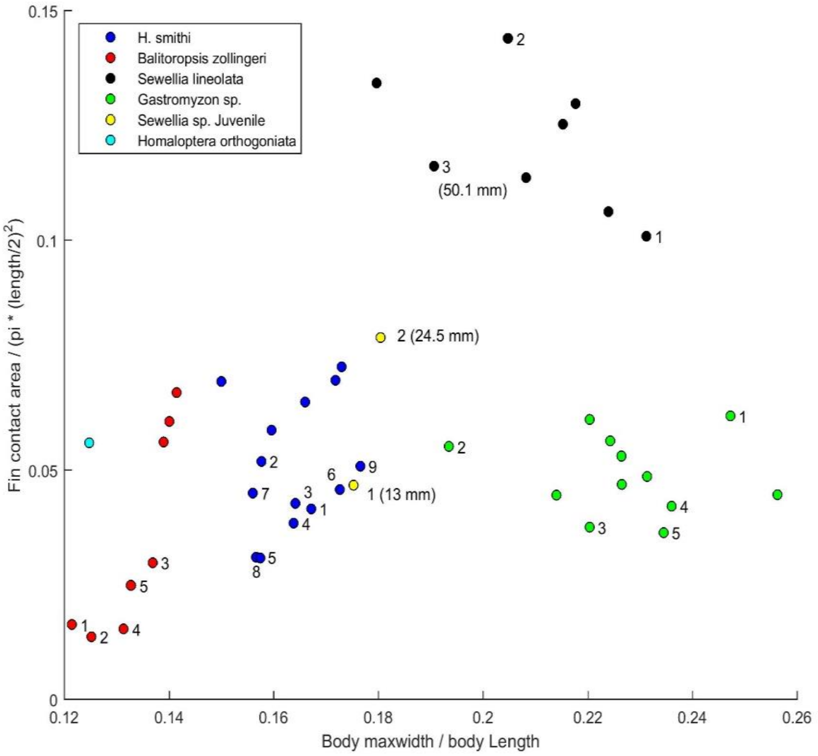


Figure 9. Scatter plot in two dimensions of body fineness and fin area ratio to length, both are ratios without units. The original 22 data from existing specimens are plotted with the additional of 23 additional data from other sources (points without numbers by them). The small numbers beside some of the points indicate the relative absolute size (standard length) with 1 being the shortest. There was no relationship between length and derived statistics. The additional points fit the existing patterns reasonably well with the exception of the blue and red groups where the new points are all translated in the positive Y direction. The new points are from living fish and their fins were more extended for these species. The additional juvenile Sewellia shows a pattern of progression in fin size with maturity.

Discussion of hill stream loach body shape

The results show that the *Gastromyzonidae* of Borneo are similar in body shape to the *Sewellia* of Vietnam but that they differ significantly in terms of relative size of paired pectoral and pelvic fins (sucker size). The bodies of both the *Gastromyzonidae* and *Sewellia* are more blunt that the *Homalopteridae* and *Balitoropsis sp*. The data also show that a developing juvenile *Sewellia* moves from a body shape that resembles a *Homalopteroides* through to the full adult shape predominantly through an increase in relative fin size (in combination with a slight blunting of body shape). This mirrors the habitat use of juvenile *Sewellia* which is more similar to a *Homalopterid* when it is small (less suction type attachment and more open water habit – although remains close to substrate). So in terms of refining Hora's (1932) original 4 classifications of hill stream adaptation; bluntness of body is sufficient to differentiate lowland and highland species and thus one could propose a 2 stage classification of body bluntness and a 2 stage classification of relative fin size to differentiate the 4 main groups in this section of the study. This raises the question of the behaviour associated with the difference in fin and sucker adaptation between the *Sewellia* of Vietnam and *Gastromyzonidae* of Borneo.

Hydrodynamics of Scales

The hill stream loaches have cycloid scales which can be keeled, flat or wrinkled, and in the *Homalopteridae* have lateral line scales with a perforated duct (a hollow tubular structure leading from a hole in the centre of the scale to the posterior margin of the scale where the tube is open at the end). This type of lateral line structural scale is common in osseous fish (Günther 1880). Since the *Homalopteridae* scales are definitive in species identification they have been the source of interest for some time (Cockerell 1909, Law 1950). They have been comprehensively described for a range of species of *Gastromyzonidae* and *Homalopteridae* (Law 1950). The original opinion is that they are not extraordinary: '[*Homalopterid* scales are] all essentially the same type…more or less degenerate, but hardly specialised…earlier type of Cyprinoid scale in a weak form…' (Cockerell 1909). The scales of large (0.08 m) *Gastromyzon lepidogaster* are prominently keeled and their arrangement defining the probable line of hydrodynamic separation of their bodies is strong *prima facie* evidence of their role in drag reduction – but caution should be exercised as a rain forest cohabitant not known for its aquatic abilities also has similarly keeled scales (Fig. 10)

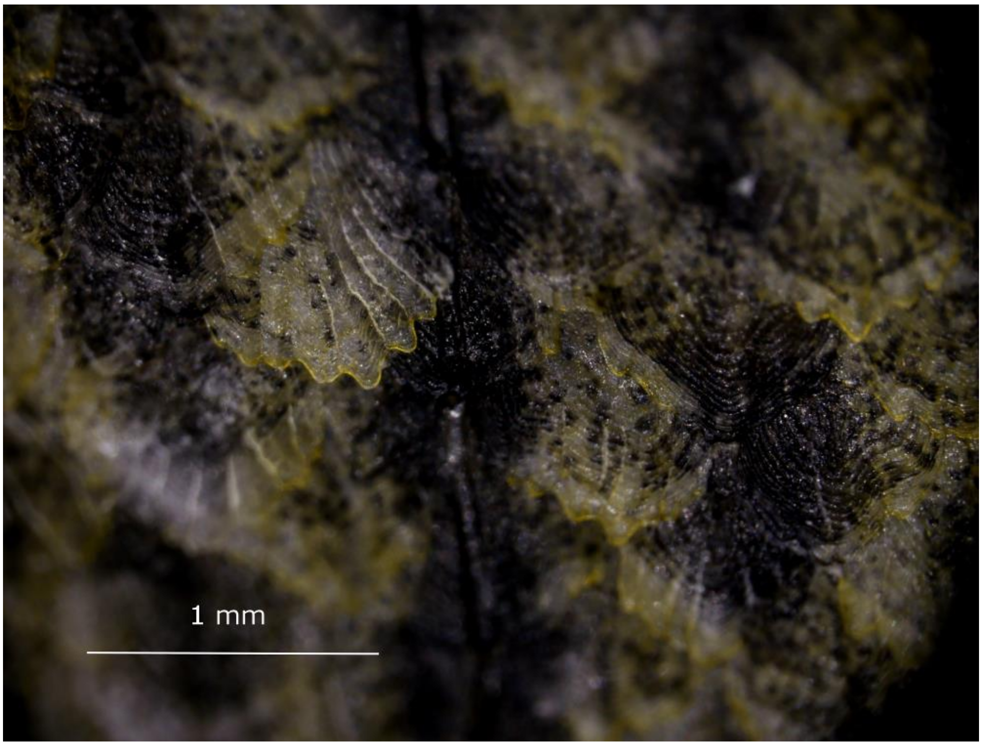
Figure 10. Scales of aquatic and terrestrial animals that are superficially similar. Left panel is a Bornean Anglehead Lizard (Gonocephalus bornensis) roughly 0.4 m in total length, with an inset panel of its leg magnified. The scales on its leg are about 3 mm at their widest. On the right is a Gastromyzon lepidogaster about 0.08 m in total length with the inset panel a magnified anterior view. In this case the scales are about 1 mm at their widest. In both cases the keels of the scales are prominent and coloured light orange. The linear separation between the head and the body of the fish in the right panel delineated by the beginning of the prominent scales is strong evidence of their hydrodynamic role as they define the hydro-dynamic separation region (Vogel 1994). But the lizard rarely if ever enters water and a hydrodynamic role for its scales is unlikely. Thus perhaps the scales of the fish are shaped for some other purpose (shared with the lizard) but this does not preclude their distribution based on hydrodynamic efficiency. In any case the situation is a great deal more complex than at first it would seem.

The scales of *Sewellia* sp are also of interest (Fig. 11). Some *Sewellia* scales have 4 or 5 prominent ridges on the posterior surface creating a series of rounded crests orientated along the major axis of the body of the fish. These are standard ctenoid fish scales with a scalloped trailing edge (Videler 1993). The scales with most prominent ridges are in the central lateral pelvic region and are strongly reminiscent of riblets on sharks (in particular Mako sharks *Isurus oxyrinchus*) which have been the subject of intense scrutiny for 30 years in the hope that they have some drag reducing potential. However when considering the hydrodynamic functionality of Mako shark scales Bechert et al.'s (2000) complete theoretical treatment is considered in combination with Aleyev's (2012) physical observations, Reid et al's (2012) models and Oeffner and Lauder's (2012) cyborgs, it appears more

doi:10.20944/preprints201911.0322.v1

likely that the scale architecture of Mako sharks provides a dynamic balance between thrust generation and drag reduction across a wide range of laminar and turbulent flow and pressure conditions on each swimming stroke. Nevertheless, it is interesting from a hydrodynamic perspective that despite the dimensional difference in overall body size, the scales of Mako sharks and hill stream loaches are similar sizes, and both species are specialists at acceleration into water already moving quickly relative to their own movement.

Alternatively the scales of hill stream loaches (e.g. Fig. 11) can be considered in relation to other teleost ctenoid fish scales. Videler (1993) outlined the relationship between cycloid and ctenoid scales and Re number using the ontogenetic development of Mullet (*Mugil* sp.). (The Reynolds number (Re) is a dimensionless number that is a useful way to compare flow patterns and turbulence thresholds of different sized objects at different flow speeds (Vogel 1993)). Observation of the mullet shows that at Re below 10^3 the juvenile has no scales; between 0.01 m and 0.03 m body length - cycloid scales develop as Re increase to 10^4. During growth up to 0.1 m and Re increase to 3×10^5 – there is gradual complete replacement of cycloid with ctenoid scales, at Re above 10^6 the ctenoid scales regress and by Re above 5×10^6 all the scales are cycloid again. This is confirmed physiologically (Jacot 1920). This is entirely consistent with hill stream loaches. Juvenile *Sewellia* (2 cm) have cycloid scales (see following figures), whereas the largest adults have ctenoid scales (Fig. 1) from about 5 cm along their body – so in water speeds of 2.5 ms^{-1} (e.g. background flow <2 ms^{-1} + burst swimming at 7-8 body lengths s^{-1}) ctenoid scales are beginning to occur at Re of 1.2×10^5. Since the ctenoid scales are largest and most elaborate on the flanks of the Sewellia around the points of separation during swimming manoeuvres (Aleyev 2012) it seems likely that these are more to do with dynamic minimisation, delay or spatial control of an ephemeral hard transition to turbulence. This is rather than a static drag management device, during coasting or when attached to the substrate, because in those cases they'd be expected to get more pronounced toward the rear (Re increases with length). In any case the occurrence of cycloid scales is known, while their function remains a mystery – although there is strong evidence of hydrodynamic function through the above analysis.



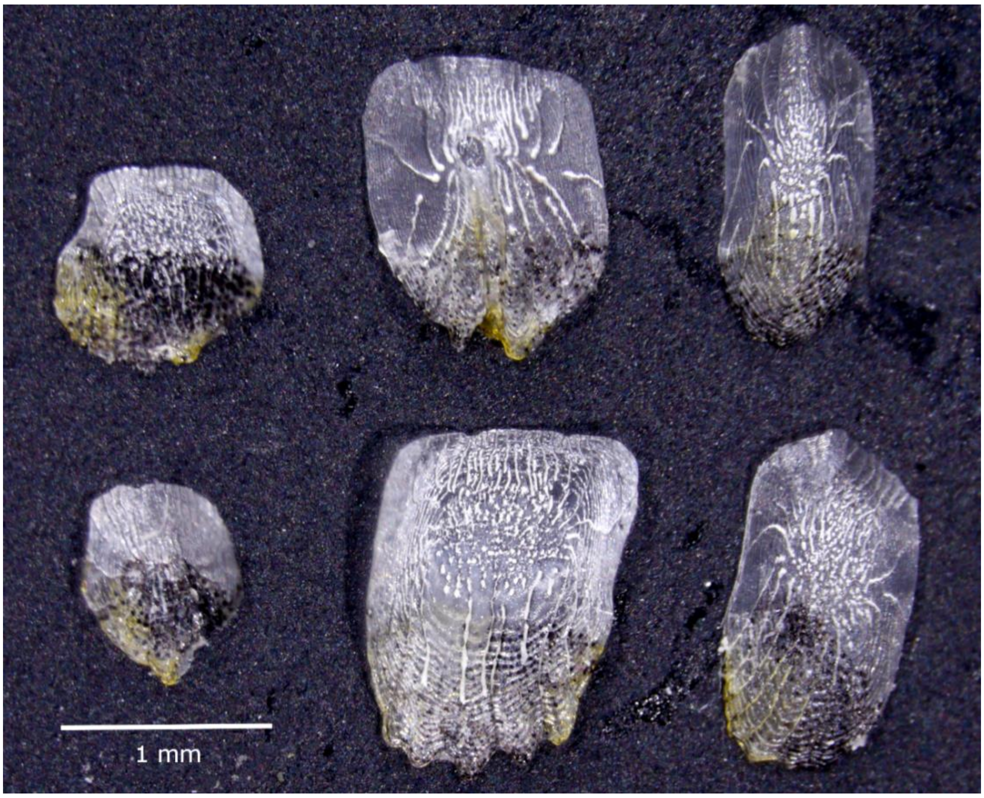


Figure 11. Scales of Sewellia 'spotted'. Upper panel scales in-situ on the lateral mid-section of intact specimen . Lower panel: left column: scales just posterior of head shield, middle column scales in line with anterior attachment of dorsal fin, right column close to tail. Upper row 1/5 body height below dorsal midline, lower row 1/5 body height above ventral margin. The scales are noticeably different in size with the largest and broadest in the mid section of the body. The scales on the tail are large but relatively long in proportion to width.

Crawling and walking

The pelvic skeletal structure of a blind cavefish, which moves around in a primitive crawling manner, may be a precursor to tetrapod structures which facilitate walking on land (Flammang et al 2016). The cavefish (*Cryptotora thamicola*) is a hill stream loach endemic to a single geological formation in northern Thailand. It is shaped like a *Homalopterid*. The *Gastromyzonidae* have one method of walking by slightly twisting the pelvic and pectoral sections of their adhesion disk independently (Described in detail in Roberts 1982 derived from Wicker 1971 cited therein). This could be a similar adaptation to the cavefish above, although physically they are not as flexible about the midsection and so the walking gait is not as pronounced as the cavefish. In a similar way to the cavefish *Gastromyzonidae* such as the *Sewellia* species can walk up wet rocks and across rocks with sheets of flowing water a few mm deep. However with their more elaborate suction devices *Sewellia* can also walk up vertical dry glass (for about 4 cm), up vertical glass in strong flowing water several cm deep, and upside down on the underside of smooth rocks removed from the water (personal observation). It is unknown if they have similar pelvic skeletal adaptations to the cavefish in order to support their fins independently. *Homalopteridae* walk in a similar way to the cave fish but underwater, and often walk backwards (Supplementary material video of *Homalopteriodes smithi*) and can also climb a vertical glass waterfall for more than 0.01 m as we observed when one entered the apparatus of an experiment to observe the climbing of Sewellia sp accidently and voluntarily began climbing (Fig. 12). The *Cryptotora* may be a useful subject to study early tetrapod walking but it is endangered and difficult to capture (Flammang et al 2016), and consequently a *Homalopteroides* which is readily available through the aquarium trade may be a convenient alternative.

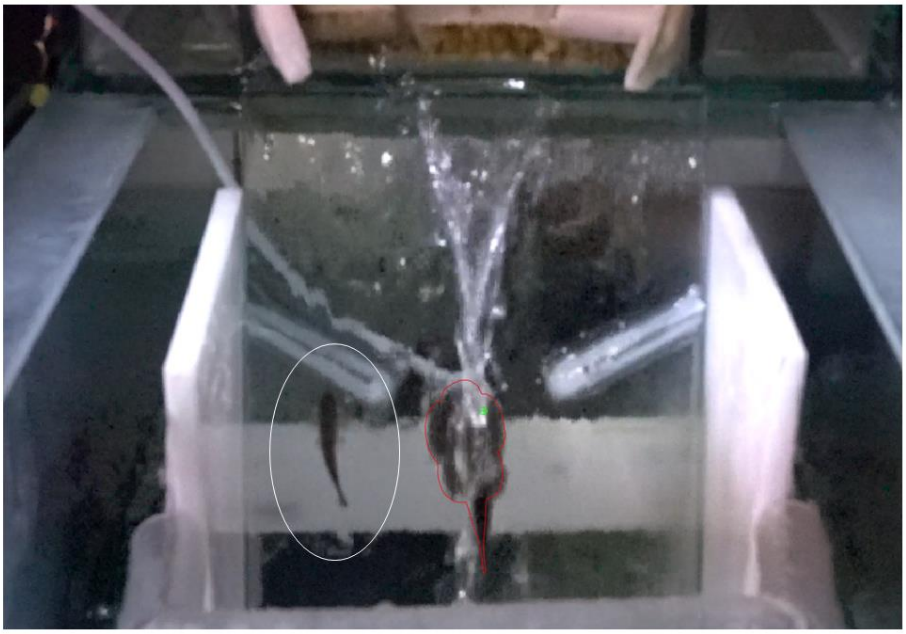


Figure 12. Homalopteroides smithi (circled in white) climbing a vertical glass surface under a waterfall. In this study (as yet unpublished) the Sewellia sp. (outlined in red) was climbing the wall when a Homalopteroids smithi accidently entered the area and voluntarily participated in the observations. It was able to climb the wall for about 0.01 m and sheltered under a baffle designed to direct the water flow to the centre – after several minutes it returned to the main tank under the waterfall.

Unculi and other protrusions

Gastromyzonid and Homalopterid species have hard conical single cell protrusions on their ventral surfaces called unculi (Roberts 1982). These are most prevalent on the ventral surfaces of the most anterior two fin rays of the pectoral and pelvic fins and they are densely packed in a regular pattern. They are about 15 µm high (Roberts 1982). The assumption is that they provide frictional resistance to hydrodynamic drag force, and considering their shape and position this seems likely. Unculi are a very common adaptation of bottom dwelling ostariophysan fish (Roberts 1982). Although considering the requirement for fast starts discussed above, it is conceivable that they prevent over strong adhesion of the leading edge of the fins so that they can be lifted quickly if necessary. Their function may warrant further analysis. Roberts (1982) also describes a series of flaps or ridges on the ventral surface of the fins which are reminiscent of the ridges on Gecko (infraorder: *Gekkota*) feet supposed to aid in their adhesion, and the regular pattern of toothed ridges on the adhesion disk of the Remora (*Echeneidae*). The hair like structures, on the ridges of the Gecko foot, adhere through Van De

Vaals forces (electrical) which are expected to be much less effective as an adhesive force in water (Sameoto and Menon 2010). The Remora is an interesting analogue as the ridges on its disk act as a hydrodynamic clamp, in that the stronger the force being applied to the anterior of the fish the stronger the adhesive disk adheres (Fulcher and Motta 2006, Nadler et al 2013). We speculate that the stiff plate like ridges of the Remora create trapezoidal channels and if the fish is forced backward the trapezoids are deformed into a squarer cross-section which increases the suction by increasing the volume of each channel. The crucial point being that the suction can be released by forward movement of the fish, so it can release itself smoothly and quickly without residual suction if it swims forward. It is perhaps possible that the flap or ridge structures on the hill stream loaches provide the same kind of one-way hydrodynamic clamp (although it is on a smaller scale and without teeth on the ridges). Both Gecko's feet and the Remora's headgear have proven rich sources of biomimetic inspiration (Sameoto and Menon 2010, Xie and Terry 2014). Again further empirical observations of the suction physics of hill stream loaches are warranted.

Colour changes and orientation with respect to patterns on substrate

Gastromyzon lepidogaster change colour from bright green to black and barred, with orange edges on fins in a few minutes (Tan 1998). Casual observation of *Homalopteroides smithi* suggests that they can change colour between orange-brown, off white, mid-grey, and black within 30 seconds to several minutes, closely matching the colour of rocks on which they sit, while the basic patterning of light and dark shaded areas is maintained to a varying degree dependent on the overall colour with it being least evident in the darkest and lightest shades (personal observation). Colour changes in fish are classified as 'physiological' or 'morphological' depending on the time taken for the change. Physiological changes can take place within seconds to minutes, while morphological usually take several weeks (Sumner 1940). Colour changes have been recorded for many species of fish, including loaches (Šećerov 1909). Šećerov (1909) demonstrated mediation through the eyes of colour changes in the Stone Loach (*Nemacheilidae* family (*Barbatula barbatula*) by recording colour changes and pattern enhancement in blinded specimens. Blinding in one eye causes darkening to the opposite side of the body. The colour palette of *Barbatula* ranges from dark green, light tan, brown, and black with patch patterning (Šećerov 1909). Colours and patterns similar to *Homalopteroides smithi* mentioned above (with the possible absence of green). Rapid colour changes in fish are widespread and are caused by concurrent movement of pigment organelles inside chromatophores in the skin, or sometimes within the body of translucent or transparent fish (Nilsson Sköld et al. 2013). Colour changes in fish that are rapid (minutes) are thought to be neurohumorally regulated (similarly to inflammation or trauma response involving the whole body system) and thus are controlled directly through the nervous system whereas morphological colour changes are thought to be controlled through the bloodstream and perhaps two hormonal pathways (Nilsson Sköld et al. 2013). Some chromatophores can change within seconds (iridophores (Nilsson Sköld et al. 2013)) whereas skin colour of Sole (*Solea solea*) for instance can take up to two weeks to change in response to a change in substrate (Ellis et al 1997). Physiological changes in colour exhibited by *Homalopteroides smithi* are likely camouflage considering the above initial observations, but other fish show rapid colour changes in response to status after combat (Beeching 1995), the presence of predators, fitness or stress (Nilsson Sköld et al. 2013). There is scope for further quantitative work on the colour of hill stream loaches. It is unknown if they use color changes for any other behavioural reason than camouflage. Their colour changing adaptations may be uniquely elaborate because they inhabit shallow, crystal clear streams and are obliged to spend a high proportion of their time stationary on sunlit exposed rocks in order to graze algae or keep a good look out for mobile prey. Other intraspecific signal methods may be less effective in fast flowing water (olfaction for instance) or not. Again casual observation suggests that these fish may orient their bodies (which are permanently banded as explained above) to coincide with bands of colouration of the rocks on which they reside, which is an opportunity for quantitative observation, and interesting if confirmed.

Husbandry and the aquarium trade

Loaches Online (LOL) is a comprehensive source of amateur scientific information on the keeping and behaviours of loaches (about 90,000 posts on 10,000 topics) (http://www.loaches.com/). After the forum had been operating for around 10 years participants summarised it in a book (Macdonald & Thorne 2007). There is a large subsection of the book and the forum devoted to hill stream loaches. Thus there is an abundance of anecdotal evidence about suitable aquarium and water parameters.

Figure 13 Hill stream loaches (Sewellia sp.) of various ages between about 3 weeks and 4 months feeding on a slab of gel based food in our laboratory aquarium. The fish co-habit without aggression and are not territoirial. They are tactile and regularly touch each other and generally remain in contact or close proximity to rocks. They are visually attentive to objects inside and outside the tank.

Tan (2006) reports 'field notes' for the *Gastromyzon* which are informal lists of information about the places that specimens were found. Field notes for 13 species of Bornean *Gastromyzonidae* indicate pH, mean 7.5, range 6 to 8.7 (tan 2006). Yu and Lee (2002) report pH (and other water parameters) for over 60 sites where *Sinogastromyzon puliensis* were sampled over several years. They found fish in pH between 6.2 and 10.9 with an average of 8.2 and no correlation of abundance to pH. We have kept *Sewellia sp*., *Gastromyzon sp*., *Homaloperoides sp*., and *Pseudogastromyzon sp*. In our tanks over three years and bred over 20 cohorts (Fig. 13). We designed an aquarium which allows water to be pumped to a shallow overhead section and flow back down a ramp into the main tank.

The tank system also provides dark spaces, aeriation and white noise through falling water, and deeper sections of slower moving water. The growth of algae was promoted through strong (500 W) incandescent overhead lighting, nutrient fertilization, high pH (~8.5) buffered by medium to high carbonate hardness (kH > 150 ppm), standard fish food and variable flow conditions. We supplement algae feeding with gel based foods (Repashy Soilent Green (repashy.com)). The substrate is only large smooth rocks which are removed and replaced, in different positions, on a weekly basis. We have kept snails, shrimp and other fish (*Danio rerio*) co-habiting in these tanks in an attempt to replicate the natural habitat.

Summary Discussion and Hypotheses

The hill stream loaches hydrodynamic adaptations should provide a system of physical constraints to evolution which can be compared and contrasted with members of the same and different families. It is clear that it is possible that the family has converged on a similar set of solutions to the same set of physical constraints on multiple occasions with minor but important pattern changes of adaptation. This hints at a genetic predisposition for a set of physical characteristics to develop on multiple occasions which are not necessarily evident in the physiology of the common closest ancestor. These may be viewed as genetic constraints to evolution – or perhaps genetic evolutionary paths of least resistance. The geographical and geological linkages in space and time are well understood.

Five Hypotheses (to falsify) with the associated questions.

1. Fin shape and size is related to rock sucking performance. Why are they capable of such strong suction?

2. Scale rib architecture is for drag reduction, propulsive performance or some other function such as resilience to abrasion. Why are the scales shaped like they are?

3. Geology is not a determinant of body shape. Why are fish that live in similar areas so dissimilar in fin shape?

4. Camouflage, sucking habit and rapid colour changes are an adaptation to avian predators, sucking piscine predators or humanoids.

5. Body shape enhances surface attachment in fast flow. When is downforce or lift created and how?

References

Abegg, C. and Thierry, B., 2002. Macaque evolution and dispersal in insular south-east Asia. *Biological Journal of the Linnean Society*, *75*(4), pp.555-576.

Aleyev, Y.G., 2012. *Nekton*. Springer Science & Business Media.

Angilletta, Michael J., and Michael W. Sears. "Coordinating theoretical and empirical efforts to understand the linkages between organisms and environments." *Integrative and Comparative Biology* (2011): icr091.

Arnold (1969) The reactions of the plaice (*Pleuronectes platessa* L.) to water currents. *J. Exp. Biol.* 51, 681-697

Barger, V. and Olsson, M., Classical Mechanics, A Modern Perspective, McGraw-Hill, 1973

Beeching, S.C. (1995). Colour pattern and inhibition of aggression in the cichlid fish Astronotus ocellatus. *J. Fish Biol.* 47, 50–58.

Bechert, D.W., Bruse, M., Hage, W. and Meyer, R., 2000. Fluid mechanics of biological surfaces and their technological application. *Naturwissenschaften*, *87*(4), pp.157-171.

Block, B.A. and Stevens, E.D. eds., 2001. *Tuna: physiology, ecology, and evolution* (Vol. 19). Gulf Professional Publishing.

Bohlen, J. and Šlechtová, V., 2009. Phylogenetic position of the fish genus Ellopostoma (Teleostei: Cypriniformes) using molecular genetic data. *Ichthyological Exploration of Freshwaters*, *20*(2), p.157.

Brainerd, E., Page, B. and Fish, F., 1997. Opercular jetting during fast-starts by flatfishes. *Journal of experimental biology*, *200*(8), pp.1179-1188.

Brown, R.M. and Guttman, S.I., 2002. Phylogenetic systematics of the Rana signata complex of Philippine and Bornean stream frogs: reconsideration of Huxley's modification of Wallace's Line at the Oriental–Australian faunal zone interface. *Biological Journal of the Linnean Society*, *76*(3), pp.393-461.

Carlston, C.W., 1969. Downstream variations in the hydraulic geometry of streams; special emphasis on mean velocity. *American Journal of Science*, *267*(4), pp.499-509.

Clark, P.U., Dyke, A.S., Shakun, J.D., Carlson, A.E., Clark, J., Wohlfarth, B., Mitrovica, J.X., Hostetler, S.W. and McCabe, A.M., 2009. The last glacial maximum. *science*, *325*(5941), pp.710-714.

Choi, H., Moin, P. and Kim, J., 1993. Direct numerical simulation of turbulent flow over riblets. *Journal of fluid mechanics*, *255*, pp.503-539.

Cockerell T. D. A., 1909. The scales of the Cobitid and Homaloptera Fishes. Proc. Biol. Soc. Washington XXII. Pp. 205-207.

Colclough, S.R., Gray, G., Bark, A. and Knights, B., 2002. Fish and fisheries of the tidal Thames: management of the modern resource, research aims and future pressures. *Journal of Fish Biology*, *61*(sA), pp.64-73.

Cox, J.P., 2008. Hydrodynamic aspects of fish olfaction. Journal of The Royal Society Interface, 5(23), pp.575-593.

Degani, G., Herbst, G.N., Ortal, R., Bromley, H.J., Levanon, D., Netzer, Y., Harari, N. and Glazman, H., 1993. Relationship between current velocity, depth and the invertebrate community in a stable river system. *Hydrobiologia*, *263*(3), pp.163-172.

De Meyer, J. and Geerinckx, T., 2014. Using the whole body as a sucker: combining respiration and feeding with an attached lifestyle in hill stream loaches (*Balitoridae*, Cypriniformes). *Journal of morphology*, *275*(9), pp.1066-1079.

Denny, M.W. and Blanchette, C.A., 2000. Hydrodynamics, shell shape, behavior and survivorship in the owl limpet Lottia gigantea. *Journal of Experimental Biology*, *203*(17), pp.2623-2639.

Du, Y. and Karniadakis, G.E., 2000. Suppressing wall turbulence by means of a transverse traveling wave. *Science*, *288*(5469), pp.1230-1234.

Elhimer, M., Alaoui, A.E.M., Croci, K., Gabillet, C. and Jacques, N., 2015, May. Numerical study of hydrodynamic impact on bubbly water. In *ASME 2015 34th International Conference on Ocean, Offshore and Arctic Engineering* (pp. V007T06A042-V007T06A042). American Society of Mechanical Engineers.

Eliason, E.J., Clark, T.D., Hague, M.J., Hanson, L.M., Gallagher, Z.S., Jeffries, K.M., Gale, M.K., Patterson, D.A., Hinch, S.G. and Farrell, A.P., 2011. Differences in thermal tolerance among sockeye salmon populations. *Science*, *332*(6025), pp.109-112.

Favret, E.A., 2009. *Functional properties of bio-inspired surfaces: characterization and technological applications*. World Scientific.

Flammang, B.E., Suvarnaraksha, A., Markiewicz, J. and Soares, D., 2016. Tetrapod-like pelvic girdle in a walking cavefish. *Scientific reports*, *6*.

Froese, R. and D. Pauly, Editors. 2000. FishBase 2000: concepts, design and data sources. ICLARM, Los Baños, Laguna, Philippines. 344 p

Freyhof, J., 2003. *Sewellia* albisuera, a new balitorid loach from Central Vietnam (Cypriniformes: *Balitoridae*). *Ichthyological Exploration of Freshwaters*, *14*(3), pp.225-230.

Fulcher, B.A. and Motta, P.J., 2006. Suction disk performance of echeneid fishes. *Canadian journal of zoology*, *84*(1), pp.42-50.

GBIF.org (10th January 2017) GBIF Occurrence Download http://doi.org/10.15468/dl.fqm7oz

Gerstner CL. 2007. Effect of oral suction and other friction-enhancing behaviors on the station-holding performance of suckermouth catfish *(Hypostomus spp.).* Can J Zool 85:133–140.

Giller, P.S. and Malmqvist, B., 1998. *The biology of streams and rivers*. Oxford University Press.

Gore, J.A., 1978. A technique for predicting in-stream flow requirements of benthic macroinvertebrates. *Freshwater Biology*, *8*(2), pp.141-151.

Gould, S.J. 2007. *Punctuated equilibrium*. Harvard University Press.

Günther, A.C.L.G., 1880. *An introduction to the study of fishes*. A. and C. Black.

Hall, B.G., 2013. Building phylogenetic trees from molecular data with MEGA. *Molecular biology and evolution*, *30*(5), pp.1229-1235.

Hora, S.L., 1930. Ecology, bionomics and evolution of the torrential fauna, with special reference to the organs of attachment. *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character*, *218*, pp.171-282.

Hora, S.L., 1950. Notes on homalopterid fishes in the collection of certain American museums. *Records of the Indian Museum*, *48*, pp.45-57.

Hora, S.L., 1952. Parallel evolution of the Gastromyzonid fishes on the mainland of Asia and in the island of Borneo. In *Proc. Nat. Inst. Sci. India* (Vol. 18, pp. 407-16).

Inger, R.F. and Chin, P.K., 2002. F*resh-water fishes of North Borneo*. Natural History Publications (Borneo).

Jacot, A.P., 1920. Age, growth and scale characters of the mullets, *Mugil cephalus* and *Mugil curema*. Transactions of the American Microscopical Society, 39(3), pp.199-229.

Kottelat, M., 2012. Conspectus cobitidum: an inventory of the loaches of the world (Teleostei: Cypriniformes: Cobitoidei). *The Raffles Bulletin of Zoology*, *26*(S1), pp.1-199.

Lappan, S. and Whittaker, D. eds., 2009. *The gibbons: new perspectives on small ape socioecology and population biology*. Springer Science & Business Media.

Lauder, G.V., 2015. Fish locomotion: recent advances and new directions. *Annual review of marine science*, *7*, pp.521-545.

Law, N.C., 1950. The scales of the Homalopterid Fishes. *Rec. Ind. Mus*, *48*, pp.69-84.

Lighthill, M.J., 1971. Large-amplitude elongated-body theory of fish locomotion. *Proceedings of the Royal Society of London B: Biological Sciences*, *179*(1055), pp.125-138.

Liu, S.Q., Mayden, R.L., Zhang, J.B., Yu, D., Tang, Q.Y., Deng, X. and Liu, H.Z., 2012. Phylogenetic relationships of the Cobitoidea (Teleostei: Cypriniformes) inferred from mitochondrial and nuclear genes with analyses of gene evolution. *Gene*, *508*(1), pp.60-72.

Martin-Smith, K.M., 1998. Relationships between fishes and habitat in rainforest streams in Sabah, Malaysia. *Journal of Fish Biology*, *52*(3), pp.458-482.

Macdonald M., Thorne M., 2007. *Loaches: Natural History and Aquarium Care.* TFH Publications Limited Neptune City NJ USA.

McConkey, K.R., Ario, A., Aldy, F. and Chivers, D.J., 2003. Influence of forest seasonality on gibbon food choice in the rain forests of Barito Ulu, Central Kalimantan. *International Journal of Primatology*, *24*(1), pp.19-32.

Motta, Philip, et al. 2012 Scale morphology and flexibility in the shortfin Mako Isurus oxyrinchus and the blacktip shark Carcharhinus limbatus. *Journal of morphology* 273.10 (2012): 1096-1110.

Nadler, J.H., Mercer, A.J., Culler, M., Ledford, K.A., Bloomquist, R. and Lin, A., 2013. Structures and function of remora adhesion. In *MRS Proceedings* (Vol. 1498, pp. 159-168). Cambridge University Press.

Nilsson Sköld, H., Aspengren, S. and Wallin, M., 2013. Rapid color change in fish and amphibians–function, regulation, and emerging applications. *Pigment cell & melanoma research*, *26*(1), pp.29-38.

Oeffner, J. and Lauder, G.V., 2012. The hydrodynamic function of shark skin and two biomimetic applications. *Journal of Experimental Biology*, *215*(5), pp.785-795.

Ott, G., 2009. Redescription of Homaloptera ripleyi (Fowler, 1940) from Sumatra, Indonesia (Teleostei: *Balitoridae*). *Bulletin of Fish Biology Volume*, *11*(1/2), pp.73-86.

Payne, J., Francis, C.M. and Phillipps, K., 1998. *Field guide to the mammals of Borneo*. Sabah Society; World Wildlife Fund Malaysia.

Phillipps, Q. and Phillipps, K., 2016. *Phillipps' Field Guide to the Mammals of Borneo and Their Ecology: Sabah, Sarawak, Brunei, and Kalimantan*. Princeton University Press.

Quinn, T.P., 2011. *The behavior and ecology of Pacific salmon and trout*. UBC press.

Randall, Z.S. and Riggs, P.A., 2015. Revision of the hill stream lizard loaches, genus Balitoropsis (Cypriniformes: *Balitoridae*). *Zootaxa*, *3962*(1), pp.206-225.

Reid, Daniel AP, et al. "Fluid dynamics of moving fish in a two-dimensional multiparticle collision dynamics model." *Physical Review E* 85.2 (2012): 021901.

Roberts, T.R., 1982. Unculi (horny projections arising from single cells), an adaptive feature of the epidermis of ostariophysan fishes. *Zoologica Scripta*, *11*(1), pp.55-76.

Rodman, P.S., 1991. Structural differentiation of microhabitats of sympatricmacaca fascicularis andM. nemestrina in East Kalimantan, Indonesia. *International Journal of Primatology*, *12*(4), pp.357-375.

Sawada, Y., 1982. Phylogeny and zoogeography of the superfamily Cobitoidea (Cyprinoidei, Cypriniformes). Memoirs of the Faculty of Fisheries Hokkaido University, 28(2), pp.65-223.

Sameoto, D. and Menon, C., 2010. Recent advances in the fabrication and adhesion testing of biomimetic dry adhesives. *Smart Materials and Structures*, *19*(10), p.103001.

Schwenk, K., Padilla, D.K., Bakken, G.S. and Full, R.J., 2009. Grand challenges in organismal biology. *Integrative and Comparative Biology*, *49*(1), pp.7-14.

Šećerov, S., 1909. Farbenwechselversuche an der Bartgrundel (Nemachilus barbatula L.). *Development Genes and Evolution*, *28*(4), pp.629-660.

Shields, A., 1936. *Application of similarity principles and turbulence research to bed-load movement*. Soil Conservation Service.

Šlechtová, V., Bohlen, J. and Tan, H.H., 2007. Families of Cobitoidea (Teleostei; Cypriniformes) as revealed from nuclear genetic data and the position of the mysterious genera Barbucca, Psilorhynchus, Serpenticobitis and Vaillantella. *Molecular Phylogenetics and Evolution*, *44*(3), pp.1358-1365.

Sumner, F.B., 1940. Quantitative changes in pigmentation, resulting from visual stimuli in fishes and amphibia. *Biological Reviews*, *15*(3), pp.351-374.

Sulaiman, Z.H. and Mayden, R.L., 2012. Cypriniformes of borneo (Actinopterygii, otophysi): an extraordinary fauna for integrated studies on diversity, systematics, evolution, ecology, and conservation. *Zootaxa*, *3586*, pp.359-376.

Sun, Z., Zhou, D., Zhong, Z., Xia, B., Qiu, X., Zeng, Z. and Jiang, J., 2006. Research on the dynamics of the South China Sea opening: Evidence from analogue modeling. *Science in China Series D: Earth Sciences*, *49*(10), pp.1053-1069.

Tan, H.H., 2006. *The Borneo Suckers: Revision of the Torrent Loaches of Borneo (Balitoridae, Gastromyzon, Neogastromyzon)*. Natural History Publications (Borneo).

Tan, H.H., 1998. Diversity of freshwater fishes from eastern Sabah: annotated checklist for Danum Valley and a consideration of inter-and intra-catchment variability. *The Raffles Bulletin of Zoology*, *46*(2), pp.573-604.

Tang, Q.Y., Liu, S.Q., Yu, D., Liu, H.Z. and Danley, P.D., 2012. Mitochondrial capture and incomplete lineage sorting in the diversification of balitorine loaches (Cypriniformes, *Balitoridae*) revealed by mitochondrial and nuclear genes. *Zoologica Scripta*, *41*(3), pp.233-247.

Turner E., 2016. Practical Fishkeeping. http://www.practicalfishkeeping.co.uk/features/articles/how-to-keep-hill stream-and-brook-loaches-in-the-aquarium Accessed September 30 2016.

Uijttewaal, W.S., 2014. Hydrodynamics of shallow flows: application to rivers. *Journal of Hydraulic Research*, *52*(2), pp.157-172.

Videler JJ. Fish swimming. Springer Science & Business Media; 2012 Dec 6. (First edition: 1993 Chapman and Hall, London UK)

Vogel, S., 1994. *Life in moving fluids: the physical biology of flow*. Princeton University Press.

Voris, H.K., 2000. Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography*, *27*(5), pp.1153-1167.

Wallace, A.R., 1863. On the physical geography of the Malay Archipelago. *The Journal of the Royal Geographical Society of London*, *33*, pp.217-234.

Wang,T.Y. and Tzeng,C.S., Molecular phylogeny of *Balitoridae* fish in China. Unpublished. Accessed 28/10/2016 (e.g. https://www.ncbi.nlm.nih.gov/nuccore/AY281285.1).

Webb, Paul W. "Form and function in fish swimming." *Scientific American* 251 (1984): 72-82.

Wu, F.C. and Wang, C.F., 2002. Effect of flow-related substrate alteration on physical habitat: a case study of the endemic river loach *Sinogastromyzon* puliensis (Cypriniformes, *Homalopteridae*) downstream of Chi-Chi Diversion Weir, Chou-Shui Creek, Taiwan. *River Research and Applications*, *18*(2), pp.155-169.

Wu, J., He, Y., Ren, H., Zhang, Y., Du, Z., Xie, M., Zhu, G., Wang, Q., Jiang, Y., He, T. and Wen, A., 2016. The complete mitochondrial genome sequence of Beaufortia szechuanensis (Cypriniformes, *Balitoridae*). *Mitochondrial DNA Part A*, *27*(4), pp.2535-2536.

Xie, W. and Terry, B.S., 2014. Biomimetic Attachment to the Gastrointestinal Tract. *Journal of Medical Devices*, *8*(3), p.030909.

Shyi-Liang Yu and Teh-Wang Lee, 2002. Habitat preference of the stream fish, *Sinogastromyzon puliensis* (*Homalopteridae*). *Zoological Studies* **41**(2): 183-187.

Supplementary information

S1 Catchment scale habitat analysis

Table S1. Information derived from Tan (2006) and Freyhof (2003) digitised maps. Species are genus G (Gastromyzon), H (Hypergastromyzon), N (Neogastromyzon) from Tan (2006) and S (*Sewellia*) from Freyhof (2003). Catchments were defined as isolated coastal connections. Interrupts were catchments in between inhabited catchments of a single species. Ranges were the furthest distances between two distribution sites and areas were smallest polygons containing all sites of a single species.

Species	**Number of sites**	**Area**	**Range**	**Catches**	**Interrupts**
G. borneensis	2	756	30	1	0

G. monticola	2	535	38	2	0
G. ornaticauda	6	8,166	210	5	3
G. cornusaccus	1			1	0
G. extrosus	2	446	35	1	0
G. introsus	6	4,969	164	2	8
G. cranbrooki	1			1	0
G. bario	6	6,685	124	3	1
G. punctulatus	8	49,170	521	5	14
G. katibasensis	1			1	0
G. aeroides	4	5,200	214	3	1
G. fasciatus	6	29,577	522	4	11
G. praestans	6	44,191	427	4	12
G. contractus	2	4,119	136	1	0
G. megalepis	6	11,846	240	2	3
G. umbras	2	1,252	37	2	1
G. ctenocephalus	2	948	42	2	1
G. scitulus	1			1	0
G. lepidogaster	18	56,474	401	10*	1
G. psiloetron	8	33,444	295	3	0
G. ridens	2	1,789	120	1	0
G. stellatus	3	3,053	215	3	1
G. zebrinus	1			1	0
G. crenastus	1			1	0
G. danumensis	3	1,452	121	1	0
G. aequabilis	1			1	0

G. ingeri	2	616	26	1	0
G. pariclavis	1			1	0
G. embalohensis	5	3,524	113	1	0
G. spectabilis	1			1	0
G. russulus	4	5,972	282	2	5
G. viriosus	1			1	0
G. venustus	1			1	0
G. auronigrus	1			1	0
N. nieuwenhuisii	2	2,480	288	2	5
N. pauciradiatus	1			1	0
N. crassiobex	2	2,509	111	2	4
N. chini	5	8,842	300	3	3
N. kottelati	1			1	0
N. brunei	1			1	0
H. humilis	2	2,124	107	1	0
H. eubranchus	1			1	0

	Number of sites	**Area**	**Range**	**Catches**	**Interrupts**
S. lineolata	18	11,187	244 874**	9	2
S. elongata	1	0	0	1	0
S. speciosa	1	0	0	1	0
S. marmorata	3	360	46	2	1
S. pterolineata	3	280	62	2	2

S. breviventralis	1	0	0	1	0
S. patella	3	309	42	2	20
S. albisuera	2	343	34	2	20

* Tan (2006) mentions 14 catchments in text, but to be consistent we identified only 10 from the maps. This does not impact the analysis.

** *Sewellia* lineolata is also resident at locations outside of Freyhof's (2003) map in Southern Vietnam (Hora 1952) and LOL(2016), when these are calculated into the scheme range is the only statistic that is available.

S2 Phylogenic tree details.

The evolutionary history was inferred by using the Maximum Likelihood method based on the General Time Reversible model [1]. The tree with the highest log likelihood (-9051.2727) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 1.4722)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 51.7629% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 25 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 1073 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 [2].

Phylogenic tree references

1. Nei M. and Kumar S. (2000). Molecular Evolution and Phylogenetics. Oxford University Press, New York.

2. Kumar S., Stecher G., and Tamura K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets.Molecular Biology and Evolution 33:1870-1874.

3. Felsenstein J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783-791.

4. Tamura K., Battistuzzi FU, Billing-Ross P, Murillo O, Filipski A, and Kumar S. (2012). Estimating Divergence Times in Large Molecular Phylogenies. Proceedings of the National Academy of Sciences 109:19333-19338.

S3 Body shape analysis

Calibration of classifier – initial results analysis

The three statistics distinguished between the 4 groups of fish in alternative ways (Fig. 3). In the frequency histograms it is evident that each statistic provides alternative ‘leverage’ on the classification. For instance the body maxwidth ratio to body length (right panel Fig. 3) shows ‘clear water’ between *B. zollingeri* and *H. smithi* whereas neither of the other statistics do that. There are also situations in which one statistic appears to provide reinforcement or replication of an effect already provided by another, for instance the right most and left most panels in Figure 3 are similar in pattern.

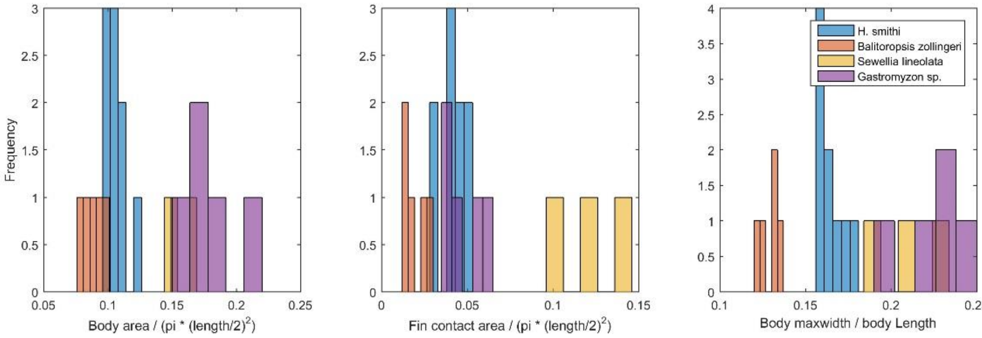


Figure S3.1. Frequency histograms of 3 statistics designed to distinguish between 4 species of fish. The left panel shows the body area ratio to length appears to differentiate well between 2 groups of 2 species. The centre panel shows a single species well differentiated by the fin contact area to length metric. The right panel differentiates between the B. zollingeri species and the others.

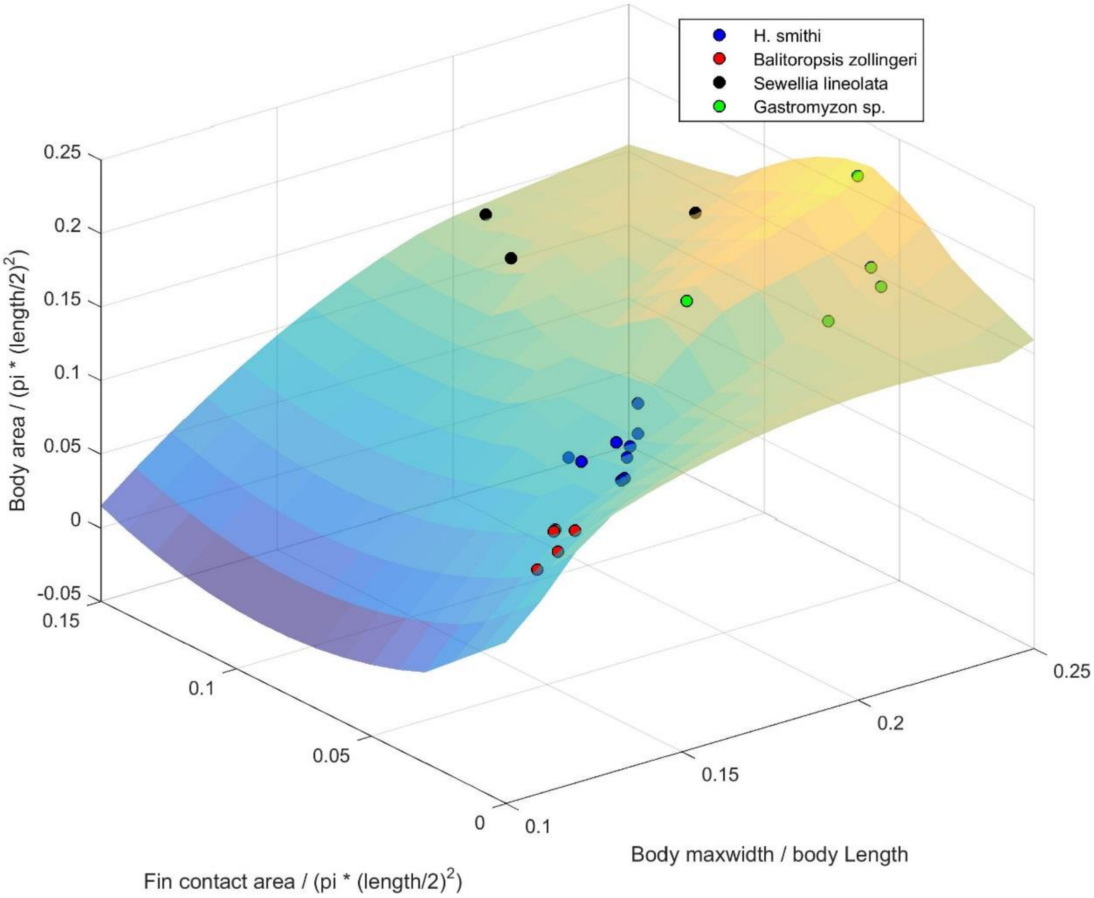


Figure S3.2. 3D scatter plot of the data. The points represent the statistics for each of 22 individual adult fish. The surface is an interpolated surface based on all the data treated similarly. This plot highlights the fact that the fin area statistic – the left-hand side horizontal axis (increasing from the centre to the left in this view) – provides a clear differentiation for one group of fish (Sewellia lineolata – black dots), while the other two axes show a potential pattern in the data for the other three species in an ascending line from centre to right.

When plotted on a 3D scatter plot two patterns become evident (Fig. 4). Firstly the fin area statistics differentiate the *Sewellia lineolata* and secondly there is a clear progressive pattern between the other three species based on body shape to body length. It is also evident that the two statistics based on body shape to length are reasonably similar and that the discrimination may be effective with only two dimensions (the plan view of the 3D scatter plot in Fig 4 appears to separate the species). The effectiveness of two dimensions to differentiate the species is clear when beta distributions are fitted to the density histograms (Figs. 5 and 6). Body fineness (width ratio to length) (Fig. 5) differentiates between all groups except for *Sewellia* and *Gastromyzon* which are similar in body shape. Therefore another statistics that differentiates between these two species will provide the 'leverage' to a classifier to be able to classify all the different species when used in combination. The fin area statistic provides differentiation between *Sewellia* and *Gastromyzon* (Fig. 6). Therefore the two statistics of body fineness and fin area ratio are likely to provide good discrimination between all the species (Fig. 7).

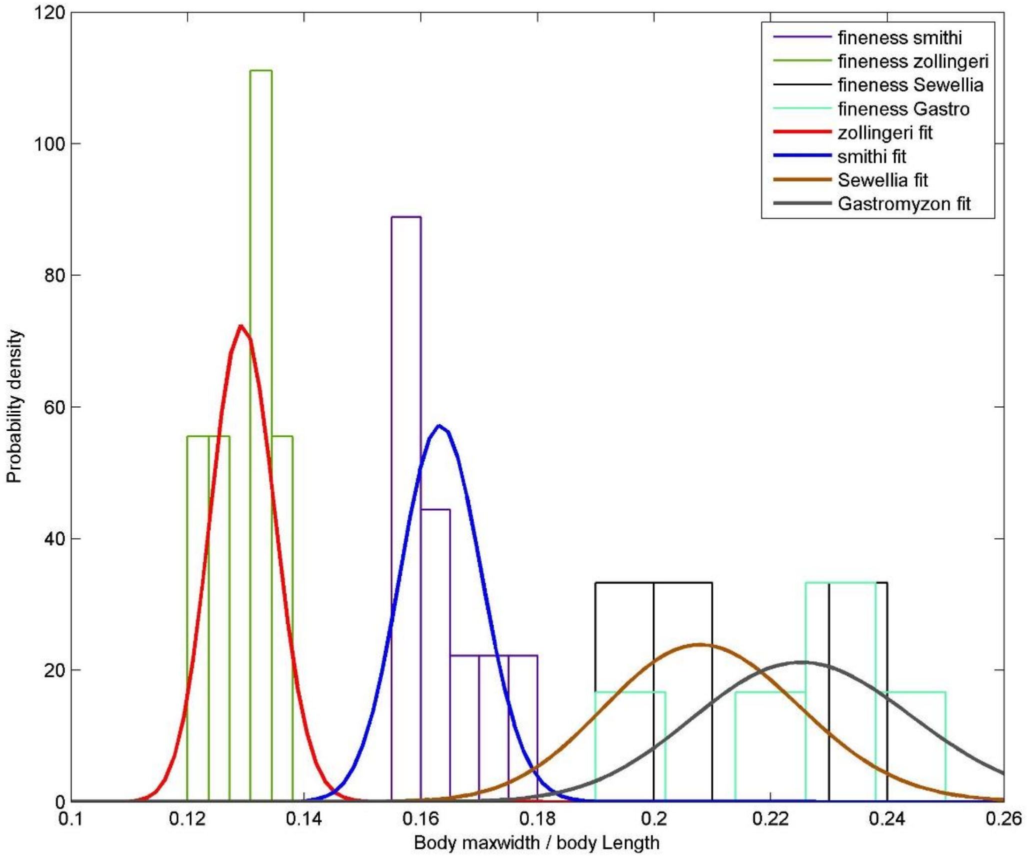


Figure S3.4. Probability density histogram of body fineness (max width ratio to length) for the fish in the study. Beta distributions have been fitted to the density data. The fitted distributions show clear statistically significant differentiation between H. smithi (blue) and B. zollingeri (red) with minimal overlap. The only two groups that are not differentiated by this statistic are Sewellia (brown) and Gastromyzon (grey) because they have similar body shapes and consequently their fitted beta distributions overlap.

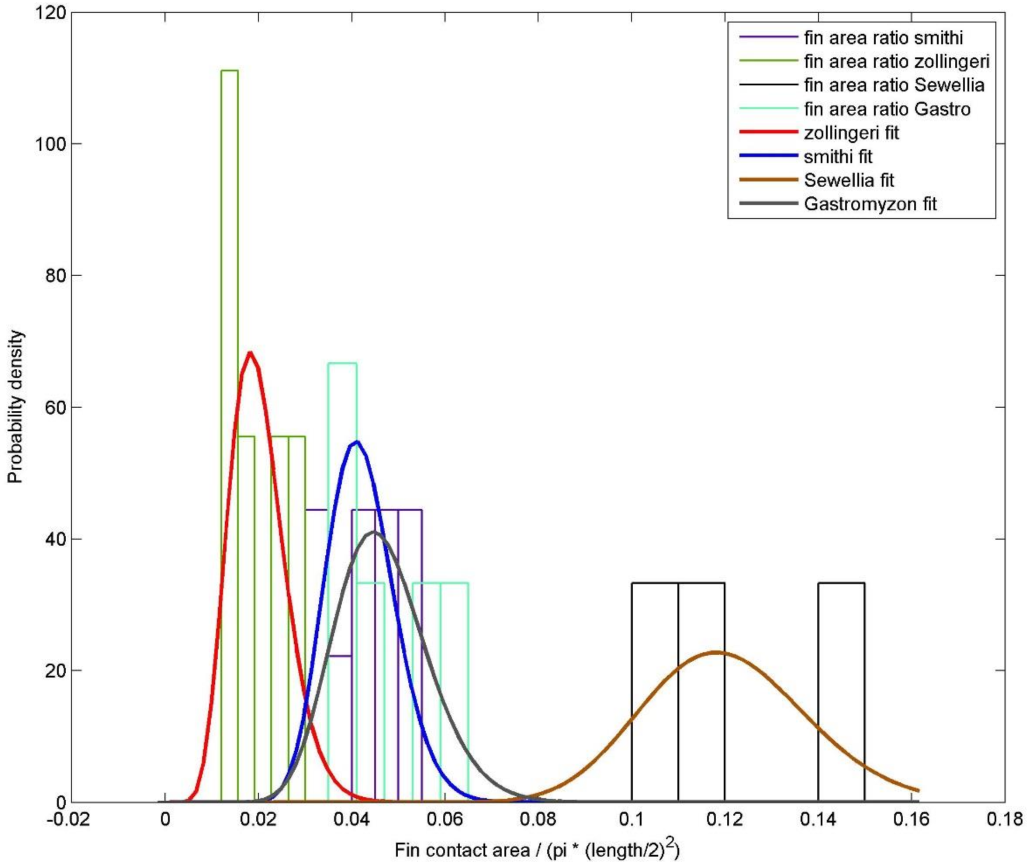


Figure S3.5. Probability density histogram of body fineness (max width ratio to length) for the fish in the study. Beta distributions have been fitted to the density data. The fitted distributions show clear statistically significant differentiation between Sewellia (brown) and all other species.

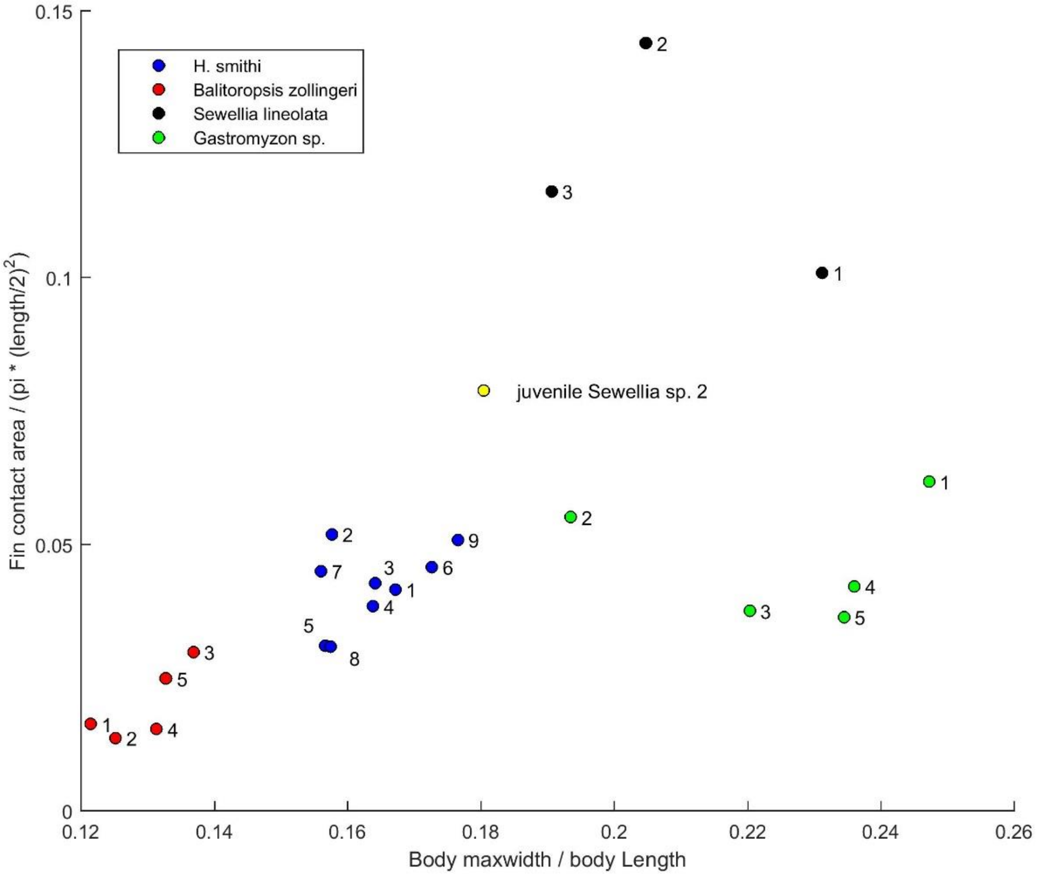


Figure 7. Scatter plot in two dimensions of body fineness and fin area ratio to length. This plot shows the species are all reasonably well separated. The Sewellia (black dots) are clearly differentiated by their larger relative fin area, or it could be viewed that the Gastromyzon (green dots) are differentiated by their relatively small fin area ratio if the patterns of the other species are extended in a linear relationship. The juvenile Sewellia falls outside all the other species (is further from the centre of mass of any other group than any other member of those groups). But it does appear to follow the pattern from left bottom to top right in this graph during its development. The numbers beside the points indicate their length ranking in their species group, so green dot with a numeral '1' beside it is the shortest of the Gastromyzon sp. In the Gastromyzon group and that with a 5 is the longest. No patterns of length are evident.

Effects of age and size

It is assumed that fish continue to grow throughout their lives, and that their rate of growth is related to food intake rather than age alone. I examined if absolute length (length measured in metres) for each species was related to the patterns formed in the putative classifier. There was no evidence that spread of individuals within the dimensions of the classifier were related to age/size, except for the juvenile *Sewellia lineolata*, where the progression seems to be from areas of lower fin area ratio to length toward the higher ratios of adult *Sewellia*.

S3.1 Other sources of data

Photographs were extracted from our existing library of photographs. We have a large number of photographs of our existing stock, for record keeping and explorative purposes. We have not kept formal id shots on arrival (but it is perhaps about time we did!). Photographs were also extracted from a general search for various species on the internet and were extracted from the definitive taxonomy guide to *Gastromyzon* sp. 'The Borneo Suckers' (Tan Heok Hui 2006).

Table S3. Other sources of loach data.

File ID	Species	Source	Credit
DSC00061.jpg	*H. smithi* *B. zollingeri* *G. punctulatus*	Oxford University	Author
DSC00179.jpg	*H. smithi*	Oxford University	Author
DSC00291.jpg	*S. lineolata*	Oxford University	Author
DSC00818.jpg	*S. lineolata*	Oxford University	Author
Sewellia-lineolata-baby_13mm.jpg	*S. lineolata Juvenile*	http://www.joerg-bohlen.de/?p=357	Jörg Bohlen
imagesGas2.jpg	*G. stellatus*	http://www.loaches.com/species-index/photos/g/gastromyzon_stellatus_09.JPG	Martin Thorne
imagesGastro.jpg	*G. bario?*	http://www.estalens.fr/articles.php?lng=fr&pg=1023	Estalens.fr
imagesInet2.jpg	*S. lineolata*	http://www.aquaportal.bg/Fish/details.php?Fish=346	Aquaportal.bg
imagesInet.jpg	*S. lineolata*	http://www.poissons-exo.com/*Sewellia*-lineolata-1285.html	poissons-exo
imagesInternet.jpg	*S. lineolata WF*	http://blog.sina.com.cn/s/blog_99496cc80101lgeo.html	Miroslav Farkak
photo1Tan.jpg	*G. fasciatus*	Tan Heok Hui 2006 (p 94)	Tan Heok Hui
photo2tan.jpg	*G. praestans*	Tan Heok Hui 2006 (p 97)	Tan Heok Hui
DSCF2033.jpg	*H. orthogoniata*	Oxford University	Author