



universität
wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„Shedding light on the evolution of an enigmatic fish.
Comparative analyses of cavernicolous and
surface-dwelling populations of *Garra barreimiae*“

verfasst von / submitted by

Mag.rer.nat Sandra Kirchner

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

Wien, 2019 / Vienna, 2019

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on the student
record sheet:

A 794 685 437

Dissertationsgebiet lt. Studienblatt /
field of study as it appears on the student record sheet:

Biologie / Biology

Betreut von / Supervisor:

ao. Univ.-Prof. Mag. Dr. Harald Krenn
Priv.-Doz. Dr. Mag. Elisabeth Haring

Acknowledgements

Here and now is the place and time to say thank you to so many people that guided, helped and supported me on this long and bumpy PhD road.

First I would like to thank the University of Vienna, who funded this thesis in the course of a uni:docs fellowship programme and for defraying publication costs. My gratitude goes also to the Federal Ministry for Transport, Innovation and Technology Austria who provided financial support in the course of a FEMtech practical. And, of course, the Natural History Museum Vienna, not only for providing financial support in the form of lab consumables and collecting trips but especially for allowing me to come and work at such a wonderful place together with an incredible working group one can only wish for. Here, I found the perfect working environment and inspiring people I love to work with.

But above all, my deepest gratitude and appreciation goes to my team of supervisors. It was a privilege to have this “triangle” of people who made this journey even possible. Harald Krenn, who is one of the kindest and most understanding persons I know, was always there offering assistances with whatever obstacles crossed my way. Thank you, Harald, for patiently guiding me through this maze.

Luise Kruckenhauser and Elisabeth Haring, you both accompanied my academic history right from the beginning, you were the ones who introduced me to the academic world and who taught me to find my place. In the course of this journey, I truly learned to deeply value and appreciate the knowledge and experiences you shared with me, and I realised that these are amongst the most precious gifts one can receive. You gave me a push (or kick when necessary) into the right direction when I got lost in the jungle of task and challenges that come along with pursuing an academic career. You shaped my understanding of good scientific practice and you taught me how to establish and maintain scientific integrity and to always keep a critical mind. You have been (and still are) my remodels, not only scientifically but also personally. We shared not only stimulating discussions and heart-warming laughs over coffee or a glass of wine, but we also shared adventurous experiences and unforgettable moments during our numerous field trips. These were times, which I deeply enjoyed and will always remember fondly. In your own unique ways, you both are strong and independent women to whom I look up. Thank you for “adopting” me and allowing me to be part of this wonderful working group. Thank you for the way you let people be themselves, for rather allowing them to develop and blossom on their own terms than forcing them into the tight corset of expectations and requirements one has to face in this world.

At this point, I also want to express my deep admiration for Helmut Sattmann, whose open-mindedness and inspiring and refreshing life philosophy made a huge impact on me and my work. You always encouraged and challenged me to think outside the box. Besides, without your initiative and commitment, I would have never gotten the chance to work on this peculiar fish. Thank you for all of that.

A huge shout-out and thank you goes to my NHM-gang: Alex, Astrid, Arthur, Bärbi, Carina, Christoph, Florian, Hannah, Iris, Josef, Julia, Kathi, Marcia, Michi, Michl, Oliver, Paul, Susi (and all the others I forgot to mention here). Thank you all for making the daily work routine in the lab such a joy, for sharing laughs and cries, for partying together but also for standing by my side helping me through difficult times. You all make the NHM to such a fruitful and caring environment. You have become my second family over the last decade, and for that, I am incredibly grateful.

Last but not least, I want to thank my parents, although they do not always understand what I am doing or why I am doing it, they never get tired of supporting me and enabling me to follow my dreams.

Table of Content

Summary	9
Zusammenfassung	12
General Introduction	15
Aims	17
Thesis questions and corresponding publications	19
Chapter I: Adaptation to cave environments, gene flow between troglomorphic cavefish and conspecific surface populations	25
Paper 1 - Characterization of 19 new microsatellite loci for the Omani barb <i>Garra barreimiae</i> from 454 sequences	27
Paper 2 - Investigating gene flow between the blind cavefish <i>Garra barreimiae</i> and its conspecific surface populations	35
Chapter II: Phylogeography and cryptic diversity of <i>G. barreimiae</i>	51
Paper 3 - The morphological diversity of <i>Garra barreimiae</i> [Teleostei: Cyprinidae]	53
Paper 4 - Hidden diversity – Delimitation of cryptic species and phylogeography of the cyprinid <i>Garra barreimiae</i> complex in Northern Oman.....	69
Concluding Remarks	97
Bibliography	104
Appendix – Conference contributions	107

Summary

The climate in the southeastern Arabian Peninsula is hot and dry, and in the course of global warming the large-scale desertification in this region and the associated water scarcity continues to progress. Despite these unfriendly conditions, the fauna of this geographical region comprises also freshwater organisms like the cyprinid *Garra barreimiae*. This fish can be found in all kinds of water bodies, in the few permanent lakes and streams as well as in countless temporary waters throughout its distribution area in the Hajar mountains in northern Oman and the UAE. The harsh conditions under which this fish has managed to sustain present it as a true survivor who outlasts long periods of drought, probably by retreating into the water reservoirs in the numerous karst caves of the Hajar mountains temporising/holding out for the next rainfall.

In 1980 Dunsire & Gallagher discovered an unpigmented and blind cave population in the underground karst lakes of the Al Hoota cave in northern Oman. With a total length of approx. 5 km the Al Hoota cave is the largest cave in the east of the Arabian Peninsula and comprises four underground lakes. The cave can be reached through one of the two entrances. The fact that there is a height difference between these two entrances transforms the cave into a flow-through system with up to 100 m³/sec water flushing through the cave in case of heavy rainfall. Due to the great morphological similarity to the surface form, the cave form was identified as *G. barreimiae* (besides the obvious differences such as the absence of eyes and pigmentation). Interestingly, the larvae of the cave form possess eyes after hatching, which are overgrown by surrounding dermal tissue in the course of ontogenetic development until no externally visible structures are present in the adult individual. Given that these troglomorphic characters are still variable, this leads to the assumption that the Al Hoota population is of rather young age.

The present PhD thesis deals with the species (or species complex) of *G. barreimie* and the genetic diversity throughout its distribution area, presenting extant population structures between the countless populations, some of which are separated by geological barriers. Of special interest is how genetically distinct the two troglomorphic cave populations are from the nearby surface populations. Furthermore, the present PhD thesis investigated the prevailing distribution patterns and morphological variations in the context of harsh and unsteady environmental conditions.

The project started with developing and establishing adequate primers for the analysis of nuclear microsatellite markers. Next-generation sequencing (NGS) data was generated via 454 sequencing, suitable primer regions were extracted and tested extensively in the wetlab. After successful establishment of 19 variable microsatellite primers (Paper 1), these were used for population genetic analyses of selected populations and the results suggest that the cave-dwelling form is genetically differentiated from the nearby surface populations, but occasional gene flow from the cave to the surface populations occurs.

These findings show that although the cave population is genetically isolated, hybridisation between the two morphotypes is still possible and occurs under certain circumstances (cave habitat).

Further comparative studies on the development of the visual system (eye, optical nerve, lobus opticus, retina, lens...) between both morphotypes and the effect of absence/presence of light during development are already underway. Modern imaging methods as μ CT and histological methods are used to quantify the degree and variation of reduction of the visual apparatus in different developmental stages.

The next step was to gather data on the genetic variation within the distribution range of *G. barreimiae*. This was accomplished by analysing three mitochondrial genes as well as applying the previously established nc microsatellite markers on a broad scale. The data allowed us to identify a total of five clearly separated groups whose genetic distances are comparable to other species of the genus *Garra*. The occurrence of these groups is geographically separated, mostly by geological barriers such as deep valleys (wadis) or mountain ranges. One of these groups was previously valid as a subspecies of *G. barreimiae* (*G. b. gallagheri*, Krupp (1988)) and was recently elevated to species (*G. gallagheri*). The remaining four groups were named according to their geographical distribution: North, East, West and Central. These findings raised the question whether the genetic groups can also be differentiated morphologically. In the course of an MSc thesis, extensive morphometric measurements were carried out and evaluated. Although different trends of individual characters could be identified, a clear differentiation and assignment solely on the basis of measurements and meristics is not possible. The fact that we are dealing with several, genetically clearly separable species that cannot be distinguished based on morphology and have so far been treated as one species justifies the term "cryptic species" (according to Bickford et al. 2007). The review and description of the newly discovered species are already in preparation and will include two already known names (West - *G. barreimiae*, Central - *G. longipinnis*) along with two new species names (East - *G. sharq*, North - *G. shamal*).

The fact that this system comprises countless surface populations in addition to the troglomorphic Al Hoota cave morphotype enables a multitude of comparative studies on evolutionary adaptation to life under extreme conditions and the associated modifications. So far, only two isolated troglomorphic cave populations of this species are known, both occurring in neighbouring underground karst lakes. With such a limited occurrence, a population is always exposed to the threat of disappearing completely as consequence of single deleterious events (chemical pollution, desiccation, etc.). In light of the fact that the putative species *G. barreimiae* is classified as "Least Concern" in the IUCN red list yet comprises several cryptic species increases the pressure to re-evaluate the conservation status of these species individually. Progressive desertification, habitat destruction and water scarcity are factors that pose additional threats to a species living under harsh environmental conditions, especially if their occurrence is geographically restricted. The existence of both morphotypes enables comparative studies which focus on exploring the

effects of alternative evolutionary pathways. Hence, this system represents a valuable contribution to the scientific community.

Zusammenfassung

Das Klima im Südosten der Arabischen Halbinsel ist heiß und trocken, und im Zuge der globalen Erwärmung schreitet die Versteppung großer Landteile dieser Region und die damit verbundene Wasserknappheit immer weiter voran. Trotz dieser unwirklichen Umweltbedingungen gehören auch Süßwasserfische wie *Garra barreimiae* zur Fauna dieser geographischen Region. *G. barreimiae* ist in den wenigen permanenten sowie auch in unzähligen temporären Gewässern im Hadschar-Gebirge im Norden des Oman und den VAE anzutreffen. Die harschen Bedingungen unter denen sich dieser Fisch bisher behaupten konnte machen ihn zu einem wahren Überlebenskünstler, der trotz seiner Wasserabhängigkeit auch längere Dürreperioden ohne Niederschlag übersteht, höchstwahrscheinlich indem er sich in die Wasserreservoirs in den zahlreichen Karsthöhlen des Hadschar-Gebirges zurückzieht und dort auf den nächsten Regenfall wartet, bei dem er wieder zum Vorschein kommt.

Im Jahre 1980 haben Dunsire & Gallagher eine unpigmentierte und blinde Höhlenpopulation entdeckt, die in den unterirdischen Karst-Seen der Al Hoota-Höhle im Norden des Oman vorkommt. Die Al Hoota-Höhle ist mit ca. 5 km Länge die größte Höhle im Osten der Arabischen Halbinsel und umfasst vier unterirdische Seen. In die Höhle gelangt man durch einen der zwei unterschiedlich hoch gelegenen Eingänge, die die Höhle bei starken Regenfällen zu einem Durchflusssystem mit bis zu 100 m³/sek Wasserdurchfluss werden lassen und dabei die ganze Höhle durchspült.

Aufgrund der großen morphologischen Ähnlichkeit zur Oberflächenform (abgesehen von den offensichtlichen Unterschieden wie das Fehlen von Augen und Pigment) wurde die Höhlenform als *G. barreimiae* identifiziert. Faszinierend ist, dass die Larven dieser Höhlenform nach dem Schlupf Augen aufweisen, die erst im Laufe der ontogenetischen Entwicklung vom umliegenden Gewebe überwachsen und beim adulten Individuum äußerlich nicht mehr zu sehen sind. Dies lässt darauf zurückschließen, dass es sich bei der Al Hoota Population um eine relativ junge Höhlenform handelt, deren troglomorphe Merkmale noch relativ variabel sind.

Die Vorliegende Doktorarbeit beschäftigt sich mit der Art (bzw. dem Art-Komplex) *G. barreimie* und ihrer genetischen Vielfalt im Verbreitungsgebiet, mit vorherrschenden Populationsstrukturen zwischen den unzähligen, zum Teil durch geologische Barrieren getrennten Populationen, vor allem aber wie sich die bisher einzige troglomorphe Höhlenpopulation zu den umliegenden Oberflächenpopulationen verhält. Des Weiteren wurde in der vorliegenden Doktorarbeit auch untersucht wie sich Verbreitungsmuster und morphologische Variationen im Kontext von herausfordernden und unsteten Umweltbedingungen auswirken.

Am Beginn des Projekts stand die Entwicklung und Etablierung geeigneter Mikrosatelliten-Primer für die Analyse von nukleären Markern. Hierzu wurden „Next-Generation-Sequencing“ (NGS) Daten mittels 454-

Sequencing generiert, daraus entsprechende Primerregionen extrahiert und ausgiebig im Wetlab getestet. Nach erfolgreicher Etablierung von 19 variablen Mikrosatelliten-Primern (Paper 1) wurden diese für populationsgenetische Untersuchungen an ausgewählten Populationen herangezogen und zeigten, dass die höhlenbewohnende Form auch genetisch differenziert ist von den umliegenden Oberflächenpopulationen, es aber gelegentlich noch zu Genfluss von der Höhlen- an die Oberflächenpopulationen kommt. Diese Ergebnisse zeigen, dass die Höhlenpopulation zwar genetisch isoliert ist, Hybridisierung zwischen den Morphotypen aber noch möglich ist und unter bestimmten Umständen (Höhlen-Habitat) durchaus noch vorkommt.

Weiterführende vergleichende Untersuchungen zur Entwicklung des visuellen Systems in beiden Morphotypen (Auge, optischer Nerv, Lobus opticus, Retina, Linse...) in Abhängigkeit von Lichtexposition bzw. -ausschluss sind bereits im Gange. Hierzu werden moderne bildgebende Methoden via μ CT sowie histologische Verfahren angewendet um zu den Grad bzw. die Variation der Reduktion des visuellen Apparates zu quantifizieren.

Der nächste Schritt war, die genetische Variation innerhalb des Verbreitungsgebiets von *G. barreimiae* zu erheben. Dies wurde mittels Analyse drei mitochondrialer und der zuvor etablierten nc Mikrosatelliten Marker bewerkstelligt. Hier konnten wir insgesamt fünf klar getrennte Gruppen aufzeigen, deren genetische Distanz durchaus vergleichbar ist mit der von anderen Arten der Gattung *Garra*. Auch das Vorkommen dieser Gruppen ist geografisch getrennt, meist durch geologische Barrieren wie tiefe Täler (Wadis) oder Bergketten. Eine dieser Gruppen war zuvor als Unterart valide (*G. b. gallagheri*, Krupp (1988)), und wurde kurz darauf zur Art erhoben (*G. gallagheri*). Die übrigen vier Gruppen wurden entsprechend ihrer geografischen Verbreitung Nord-, Ost-, West- und Central benannt. Mit diesen Erkenntnissen drängte sich die Frage auf, ob die gefundenen genetischen Gruppen auch morphologisch differenzierbar sind. Im Zuge einer MSc-Arbeit wurden umfangreiche morphometrische Messungen vorgenommen und ausgewertet. Dabei konnten zwar unterschiedliche Trends einzelner Merkmale aufgezeigt werden, allerdings ist die klare Unterscheidung und Zuordnung anhand von den Mess- und Zählwerten alleine nicht möglich. Es handelt sich also um mehrere, genetisch klar trennbare Arten die morphologisch nicht unterschieden werden können und bis dato als eine Art behandelt wurden - womit die Bezeichnung „kryptische Arten“ (nach Bickford) gerechtfertigt ist. Die Beschreibung der neu entdeckten Arten ist bereits im Gange und wird in Kürze neben zwei bereits bekannten Namen (West - *G. barreimiae*, Central – *G. longipinnis*) auch zwei neue Artnamen vorstellen (East - *G. sharq*, North – *G. shamal*).

Die Tatsache, dass dieses System sowohl eine troglomorphe Höhlenform als auch unzählige Oberflächenpopulationen umfasst, ermöglicht eine Vielzahl an vergleichenden Untersuchungen die sich mit der evolutionären Anpassung an ein Leben in kompletter Dunkelheit und den damit verbundenen Modifikationen beschäftigt. Bisher sind nur zwei genetisch isolierte, troglomorphe Höhlenpopulationen dieser Art bekannt, beide kommen in unterirdischen Karstseen vor. Mit einem so stark begrenzten

Vorkommen ist eine Population immer der Gefahr ausgesetzt, durch einmalige Events (Einwaschung usw.) komplett zu verschwinden. In Anbetracht der Tatsache, dass die vermeintliche Art *G. barreimiae* mehrere kryptische Arten umfasst und in der roten Liste mit „Least Concern“ eingestuft ist, steigt auch der Druck, den Schutzstatus dieser Arten neu zu evaluieren. Voranschreitende Versteppung sowie Lebensraumzerstörung und Wasserknappheit sind Faktoren, die eine Art mit limitierter Verbreitung noch zusätzlicher Bedrohung aussetzen. Dieses System stellt in vielerlei Hinsicht eine Bereicherung im wissenschaftlichen Feld der Evolutionsforschung dar, die Existenz beider Morphotypen ermöglicht vergleichende Studien und Experimente, die prinzipiell zu wertvollen Erkenntnissen zur Anpassung an extreme Lebensräume führen und weiters Schlüsse über die evolutionäre Geschichte einer Art ermöglichen.

General Introduction

Studying the basic principles and mechanisms of evolution comprises many diverse subject areas. The present study deals with morphological reduction caused by colonization of extreme habitats and the thereby involved phenotypic alterations as well as with the phylogeographic history and cryptic diversity of the cyprinid freshwater fish *Garra barreimiae* FOWLER & STEINITZ, 1956.

This project introduces *G. barreimiae* as an ideal model organism on cave adaptation. What makes this organism especially interesting and valuable for evolutionary biology is the fact that both morphotypes, a troglomorphic cave form and a congeneric sighted surface form, coexist in close vicinity representing an ideal system for comparative analyses of evolutionary pathways. This is also the case for the well studied species of the genus *Astyanax*, but in many other known troglobiotic fish species the surface form has become extinct what limits the opportunity to comparatively investigate the two morphotypes of one species. On the other hand, this project is a perfect example to highlight the fact that the organismic diversity in the Middle East is understudied and is thereby heavily underestimated. Furthermore, the general knowledge gathered in the course of this study can be applied to many other study fields and organism groups, thus providing an adequate basis for a multitude of research topics.

Garra barreimiae and the Al Hoota Cave

G. barreimiae is endemic to Northern Oman and the United Arab Emirates, a geographic region where the present climate is hot and arid (Keith Edward Banister 1984; Feulner 1998). Precipitation is rare and irregular but can result in torrential rainfalls followed by massive floodings posing a potential threat to habitats as well as its inhabitants. Despite these putative hostile conditions, this species is quite abundant in fresh water habitats throughout its distribution area and inhabits various kinds of water bodies.

In addition to numerous surface populations, one blind and unpigmented cave population inhabiting the Al Hoota cave in the midst of the Hajar mountains was discovered by Dunsire and Gallagher in 1980. This spacious cave has developed in cretaceous limestone and is approximately 5 km long. It has two differently elevated entrances, thereby forming an underground flood system during occasional but heavy rainfalls, with flow velocities reaching up to 100 m³/s. Between such flood events, there is almost no running water in the cave, but four extensive subterranean lakes separated by several vertical drops (Kollmann, Sattmann, and Seemann 2007).

Living in Caves

Organisms colonizing such extreme habitats represent a unique opportunity to gain insight into mechanisms of adaptation and loss of function. Living under harsh environmental conditions often results

in rapid morphological changes. Animals that spend their whole life in the darkness of a cave environment are called troglobites. In general, they derived from surface dwelling ancestors, which were already preadapted to light-reduced conditions (troglophiles) (Wilkens and Strecker 2017). Animals colonize hypogean environments due to various reasons like protection from predators, protection against desiccation, to exploit different food sources or to use the cave environment as reproductive niches or hibernation places (Romero 2012). Evolution affecting troglobites under complete absence of light drives characteristic phenotypical modifications, the most striking are regression of pigmentation and the visual system (rudimentation) (Wilkens and Strecker 2017). The degree of regression is species-specific, and most cave dwelling organisms develop more sensitive constructive traits to compensate those regressive characters. They are often adapted to cave life by highly developed non-visual sensory organs (e.g. orientation, olfaction, taste) (Aden 2005; Wilkens and Strecker 2017). Both eye regression and loss of pigmentation are considered to be usually genetically based, but pigmentation can be also influenced by the environment (Wilkens 2005). These environmental-responsive morphological changes are caused by phenotypic plasticity, that is the ability of a genotype to express different phenotypes and specific traits in response to different environments. Thus, individuals with a higher plasticity often have a reproductive advantage over individuals with genetically fixed phenotype living under changing (and often extreme) environments (Romero 2009). The characteristic morphological changes of troglobites developed independently among various organism groups (salamanders, isopods, crayfish, beetles, spiders) highlighting the fact that they are the product of convergence. They have been studied most intensively in various cavefish species (e.g. *Poecilia mexicana*, *Phreatichthys andruzzii*, *Rhamdia zongolicensis*, *Typhlichthys subterraneus*). The expression of these troglomorphic traits can vary strongly and studying them can provide important insights into which evolutionary forces drive regressive adaptation.

Presently there are approximately 200 cavefish species known with numbers increasing each year due to newly described species, and the majority of them exhibit some degree of troglomorphy (Niemiller et al. 2019). The most prominent examples and best studied model organisms concerning cave adaptation in teleosts are the species of the characid cavefish genus *Astyanax*. The body of scientific studies on *Astyanax* is extensive and various fields of research, such as behaviour (Plath, Rohde, et al. 2006; Hüppop 2012; Hinaux, Rétaux, and Elipot 2016), morphology (Teyke 1990), metabolism (Hüppop 2000) or molecular phylogeny and evolution (Wilkens 1988; Strecker, Faúndez, and Wilkens 2004; Strecker, Hausdorf, and Wilkens 2012; Coghill et al. 2014) are addressed.

The Troglomorphic Cave Population of *G. barreimiae*

Despite the more obvious troglomorphic traits (lack of eyes and pigmentation), the Al Hoota cave population is considered conspecific with *G. barreimiae* by Banister (1984) due to general morphological similarities between the hypogean and epigean populations. Interestingly, larvae of the cave population exhibit externally visible eyes until they reach a certain standard length (~ 24 mm), at which point the eyes

stop to develop and are fully overgrown by surrounding dermal tissue until there is no external trace of the eyes anymore (Keith Edward Banister 1984). Banister (1984) also observed particular re-growth of the optic lobes as well as increased melanin production in cave specimens that were kept under light exposure for four months. Remarkably, we detected individuals with a phenotype in-between surface and cave morphotype that we collected from surface water bodies close to the Al Hoota cave entrance. Considering the rather recent origin of the Al Hoota cave population as suggested by Kruckenhauser et al. (2011) coupled with the occurrence of these individuals exhibiting an intermediate phenotype, the question arose whether these specimens are the product of hybridisation between the morphotypes or if they reflect the effects of phenotypic plasticity in the course of exposing cavefish to sunlight.

It is likely that the cave population of *G. barreimiae* originated at a time of climatic changes in the Late Pleistocene. In the course of rising temperature and decreasing precipitation, the connection between the cave and surface water bodies became disrupted due to a drop in water levels leaving a subpopulation with only a fraction of the species' genetic diversity trapped inside the cave and thus, genetically isolated from the nearby surface populations.

Phylogeography of *G. barreimiae*

Yet, climate and geology/geography did not only affect the cave population, it was also significantly involved in the present population structure of the many surface populations. The Northern Oman is not homogeneously populated by *G. barreimiae*. The landscape is characterized by many mountain ranges and wide plains pervaded by countless ephemeral riverbeds (wadis) that seem to act as either natural barriers or predetermined migrating routes. This may result in divergence of populations when given enough time of separation. Kruckenhauser et al. (2011) already detected a higher genetic variability within *G. barreimiae* and considering the lack of conspicuous morphological differentiations between populations this finding indicates cryptic diversity within this taxon.

Aims

In the present PhD thesis, I investigated the population structure in selected cave and surface populations of *G. barreimiae* via nuclear (nc) microsatellite analyses and tested if recent gene flow can be detected. Further, I analysed individuals exhibiting an intermediate phenotype to check if they are the products of hybridization between individuals of both morphotypes, or if they display the effects of phenotypic plasticity in the course of light exposure during early development of cave individuals. For this purpose, it was necessary to design and establish novel primers for microsatellite loci, which were isolated from genomic sequence data generated via 454 next generation sequencing. Further on, the morphological diversity between detected genetic groups was investigated via traditional morphometry, meristics and

μCT imaging analysis. The next step was the evaluation of the genetic diversity and phylogeography by means of mitochondrial (mt) sequence analyses of three different gene fragments as well as nuclear microsatellite analyses throughout the whole distribution area. Since the results indicated cryptic species within the so far known taxon "*G. barreimiae*", proper species description including a revision of this taxon are already in preparation. Thus, the taxa so far subsumed under the species *G. barreimiae* will henceforth be addressed as *G. barreimiae* species complex. In addition, the comparative study on development of the visual system between the troglomorphic cavefish and the sighted surface fish and the role of phenotypic plasticity concerning this matter is still ongoing. The topics and questions pursued in the course of this thesis and the corresponding manuscripts are listed below:

Thesis questions and corresponding publications

CHAPTER 1: ADAPTATION TO CAVE ENVIRONMENTS, GENE FLOW BETWEEN TROGLOMORPHIC CAVEFISH AND CONSPECIFIC SURFACE POPULATIONS

Adult individuals of the troglomorphic cave population of *G. barreimiae* lack pigmentation and externally visible eyes. They developed in complete darkness of the Al Hoota cave system in the midst of the Hajar mountains in Northern Oman. Of particular interest is that the juveniles of this cave population hatch with external visible eyes, which become overgrown by dermal tissue in the course of ontogenesis. In close vicinity to the Al Hoota cave are several surface water bodies inhabited by sighted individuals. In chapter 1, I followed the questions whether the two populations are still in contact with each other and if recent gene flow can be detected. It was already known that the cave population, even though slightly differentiated on a mitochondrial level, is closely related to the nearby surface populations. Since no suitable nuclear microsatellite markers were available, it was necessary to design and establish primers for microsatellite loci beforehand (see paper 1). Subsequently they were tested successfully in the lab. Population genetic analyses were performed on an extensive data set comprising individuals from the Al Hoota cave system (troglomorphic) and of selected surface populations (see paper 2).

Research Questions:

- a. Can gene flow between the troglomorphic cave population and its conspecific sighted surface populations be detected?
- b. If there is recent gene flow, does it occur on equal terms from the cave to the surface population and vice versa?
- c. What is the genetic structure and diversity of the investigated cave and surface populations?
- d. Do the nc markers confirm the presumed bottleneck events as suggested by previously published mt data?
- e. Are the individuals with intermediate phenotypes the result of recent hybridization between surface and cave individuals or do they represent cave specimens exhibiting the effects of phenotypic plasticity in the course of light exposure during early development?

1. Kirchner, S., Weinmaier, T., Rattei, T., Sattmann, H., and Kruckenhauser L. (2014): **Characterization of 19 new microsatellite loci for the Omani barb *Garra barreimiae* from 454 sequences.** BMC Research Notes 2014, 7:522.

- Focus on fulfilment of the technical/methodological requirements for the subsequent investigation on gene flow
 - Design and establishment of variable microsatellite markers specific to *G. barreimiae* from next generation 454 sequence data
 - Development of a useful tool to analyze the structure and genetic differences of populations of *G. barreimiae* and closely related species of the genus
2. Kirchner, S., Sattmann, H., Haring, E., Plan, L., Victor, R., and Kruckenhauser L. (2017): **Investigating gene flow between the blind cavefish *Garra barreimiae* and its conspecific surface populations.** Scientific Reports, 7:5130.
- Focus on questions [1] a – e
 - Utilization of the novel microsatellite primers on an extensive data set
 - First genetic analyses of population structure in *G. barreimiae*

To summarize the key findings of Chapter 1, we found that the cave populations are differentiated from nearby surface populations and show less genetic variation. Further on, we can confirm that gene flow between the troglomorphic cave and sighted surface populations is possible but occurs only rarely and unidirectional from the cave to the surface populations. The presence of individuals with cave genotype and intermediate phenotype suggests that exposure to light affects the phenotype of cave individuals during early development, hence phenotypic plasticity still plays a role in the rather young cave population of *G. barreimiae*.

As it was known from previous studies that the genetic diversity in *G. barreimiae* is higher than expected, I looked into that matter in detail applying an integrative approach on a broad data set including genetic and morphometric data, μ CT imaging techniques and geographic distribution. In the first paper, we evaluated the variability of morphological and meristic characters by means of traditional morphometry and μ CT imaging analyses. In the second paper, population genetic analyses were conducted via microsatellite loci and sequence analyses of three mt genes on an extensive sample set covering the whole distribution area.

Research Questions:

- a. How variable are morphological characters within *G. barreimiae*?
- b. Can significant morphological differences between the genetic groups of *G. barreimiae* be detected?
- c. What is the extent of genetic variation within the taxon *G. barreimiae* and its described subspecies across the distribution area?
- d. How much genetic variation can be detected between different sampling sites?
- e. Can genetic groups be distinguished and if so, are such groups concordant with their geographic distribution?
- f. Is there gene flow between the genetic groups?
- g. What are the phylogenetic relationships between *G. barreimiae* and closely related species of the genus?
- h. Is *G. barreimiae* monophyletic in the DNA-based phylogenetic tree?

3. Pichler, A., Ahnelt, H., Kirchner, S., Sattmann, H., Haring, E., Handschuh, S., Freyhof, J., Victor, R., and Kruckenhauser L. (2018): **The morphological diversity of *Garra barreimiae* [Teleostei: Cyprinidae]**. Environmental Biology of Fishes, 101:6.
 - Focus on questions [2] a – b
 - Evaluation of the morphological diversity in *G. barreimiae* by employing traditional morphometrics and μ CT imaging analyses
 - Identification of discriminating morphological characters
 - Morphological basis for an integrative approach on the revision of this taxon

4. Kirchner, S., Sattmann, H., Haring, Victor, R., and Kruckenhauser L. (submitted): **Hidden diversity – Delimitation of cryptic species and phylogeography of the cyprinid *Garra barreimiae* complex in Northern Oman**. Submitted to: Journal of Zoological Systematics and Evolutionary Research.
- Focus on questions [2] c – h
 - Evaluating the genetic diversity throughout the whole distribution area
 - Defining the geographic distribution of the genetic groups e.g. species within *G. barreimiae*
 - Checking for recent hybridization between genetic groups / species comprising the *G. barreimiae* complex
 - Using an integrative approach as basis for a Revision of the species complex

The central outcome of Chapter 2 is that the species so far valid as *G. barreimiae* comprises five distinct genetic lineages whose genetic distances are comparable to the distances between species of the genus. Both nc microsatellite data as well as mt sequence data support these findings. The distribution of the lineages is clearly separated, and their population structure explicitly arranged along geographic large-scale structures as mountain ridges or wadis that pervade the countries surface on a large scale. Besides, the monophyly of “*G. barreimiae*” in the phylogenetic tree reconstruction is disrupted as closely related species are placed amidst these lineages. However, they can not be distinguished based solely on single morphological traits, hence we address them as cryptic species.

CHAPTER 3: ONGOING INVESTIGATIONS

List of further planned studies and investigations already under progress.

Research Questions

- a. Which structures of the visual system are affected by regression in troglomorphic individuals from the cave in comparison to surface individuals with an inconspicuous phenotype?
- b. At which developmental stage does the regression of different structures of the visual system (lens, optic nerve, retina layer, cornea layer, optical nerve and lobus opticus) take place?
- c. How much variation can be detected in the degree of regression within the cave population?
- d. Has the absence/presence of light exposure during the early development of *G. barreimiae* an effect on the phenotype of both morphotypes?
- e. Can a combination of morphological characters help to discriminate the cryptic species comprised in the *G. barreimiae* species complex?

5. Kirchner, S., Weissenbacher, A., Handschuh, S., Sattmann, H., Haring, E., Krenn, H., Kruckenhauser, L. (in prep): **Comparative analyses of eye development and corresponding visual structures between troglomorphic cavefish and its congener sighted surface populations of *G. barreimiae* in dependence with light exposure.** In prep.

- Focus on questions [3] a – d
- First record of successful breeding experiments under lab conditions in *G. barreimiae* via hormone induced injection
- Collection of samples series from all developmental stages of both morphotypes
- Examining structures of the visual system via μ CT imaging and histological methods
- Comparative analyses of degree of eye degeneration (and accompanied structures) between troglomorphic cave and sighted surface individuals reared under different light regimes

6. Kirchner, S., Pichler, A., Kruckenhauser L., Borkenhagen, K., and Freyhof J. (submitted): **Review of the *Garra* species in the Hajar mountains in Oman and the United Arab Emirates with the description of two new species (Teleostei: Cyprinidae).** Submitted to: Zootaxa.

- Focus on question [3] e
- Species description of the two newly detected cryptic species based on the data from the manuscript “Hidden diversity – Delimitation of cryptic species and phylogeography of the

cyprinid *Garra barreimiae* complex in Northern Oman” and “The morphological diversity of *Garra barreimiae* [Teleostei: Cyprinidae]”

- Development of a key for species discrimination based on morphological characters

The data collected so far from the first few scanned individuals indicates diverging tendencies in eye development between the two morphotypes that were raised under natural conditions. The most striking differences between them are the missing lens and optical nerve in the cave form already at a very early developmental stage. On the other hand, we could observe a much thicker retina layer in the surface form. These are only preliminary results and much more individuals from various developmental stages need to be included and further structures investigated in order to make a well supported statement.

After inspecting the morphological/anatomical as well as genetical and distributional differences between populations of *G. barreimiae*, we consider the conjoined results sufficient to conclude that *G. barreimiae*, so far valid as a single species, actually comprises several cryptic species, two of which are newly described in the corresponding revision (submitted).

Chapter I: Adaptation to cave environments, gene flow between troglomorphic cavefish and conspecific surface populations

Paper 1: Characterization of 19 new microsatellite loci for the Omani barb
Garra barreimiae from 454 sequences

SHORT REPORT

Open Access

Characterization of 19 new microsatellite loci for the Omani barb *Garra barreimiae* from 454 sequences

Sandra Kirchner^{1*}, Thomas Weinmaier², Thomas Rattei², Helmut Sattmann³ and Luise Kruckenhauser⁴

Abstract

Background: *Garra barreimiae* is a cyprinid fish from the southeastern Arabian Peninsula, which inhabits regularly desiccating wadis and survives in isolated ponds or underground. In 1984 a cave-dwelling population was found in the Al Hoota cave system and previous genetic analyses revealed some differentiation with limited gene flow between the surface populations and the cave population. Since no suitable markers are available for evaluation of gene flow between the cave population and the adjacent surface populations, we focused on designing and establishing novel microsatellite markers from next generation sequencing data.

Findings: 19 microsatellite markers containing di- and tetranucleotide simple sequence repeats were developed from 454 sequences. Forty-four individuals from two surface populations (Wadi Al Falahi and Misfat Al Abriyeen) of *G. barreimiae* (sampling permission number 13/2012, export permission number 29/2012) were used for analyses and characterization of the loci. On average, the number of alleles per locus is 7.6 (range: 2–20). Two markers displayed indication of linkage disequilibrium in both populations (DL6X, 9XNC). Significant deviation from Hardy-Weinberg equilibrium was observed at four loci in the Misfat Al Abriyeen population (2PUM, 88CM, 1EHE, 3Z7M) and at two loci in the Wadi Al Falahi population (QLIM, 3 N43). Three of the microsatellite loci were significant for null alleles in one of the two populations (Misfat Al Abriyeen: CJHG; Wadi Al Falahi: PH8A, 3ROZ). Expected and observed heterozygosities ranged from 0 to 95.0% respectively from 0 to 95.8% (Wadi Al Falahi) and from 0 to 89.1% respectively from 0 to 95.0% (Misfat Al Abriyeen). Fourteen of these markers were successfully cross-amplified in *G. rufa*.

Conclusion: This 19 microsatellite loci provide a useful tool to understand the structure and genetic differences of populations. Moreover, these markers will help to evaluate species delimitation in *G. barreimiae* and potentially even in related species.

Keywords: *Garra barreimiae*, *Garra rufa*, Oman, microsatellites, 454 sequences

Findings

The cyprinid *Garra barreimiae* Fowler & Steinitz (1956) is endemic to the southeastern Arabian Peninsula, where it is common in ponds and rivulets of the Hajar mountains in the northern Oman and the United Arab Emirates [1] [2]. Most of these wadis dry out during summer and are only occasionally flooded, so the fish have to survive in isolated ponds or underground. In 1980 a hypogean troglomorphic population of this species was found in the Al

Hoota cave system near Al Hamra [3]. All adult individuals in this population are missing externally visible eyes, and the optic lobes are reduced. In addition, the squamation is weak and they lack pigmentation [3]. Until now the Al Hoota cave population is the only known troglomorphic form of *G. barreimiae*. Previous genetic analyses on the basis of the mitochondrial control region revealed a differentiation between cave and surface populations and possible gene flow as indicated by shared haplotypes [4]. The separation is not complete, and there seems to be limited gene flow from the cave to the adjacent surface populations. The small genetic differentiation between the two forms indicates that the troglomorphic cave

* Correspondence: sandra.kirchner@nhm-wien.ac.at

¹Natural History Museum Vienna, Central Research Laboratories, Burgring 7, 1010 Vienna, Austria

Full list of author information is available at the end of the article

population is of quite recent origin [4]. According to current taxonomy, the species comprises three subspecies: *Garra barreimiae barreimiae*, which can be found throughout the northern Oman, *G. b. gallagheri* Krupp [5] found only at one location in northern Oman and *G. b. shawkahensis* Banister & Clarke [1], found north of the United Arab Emirates [5]. Investigations of Kruckenhauser et al. [4] revealed that the first two subspecies are genetically distinct based on mtDNA analyses. Moreover, they found an additional clearly separated clade on the northeastern slope of the Hajar Mountains, suggesting that *Garra barreimiae* may be comprised of more than one species. In the present study, we analysed individuals of *G. barreimiae* from two surface populations: Misfat Al Abriyeen is a locality near the mountain village Misfat, north of Bahla. The fish were collected in the Falaj drainage system, which channels the water through the villages' gardens. About 15 km southwards from Misfat the small water body Wadi Al Falahi, a natural drain from the surrounding mountains, is located. Both localities hold numerous individuals of *G. barreimiae*. The individuals of *G. rufa*, which were used to test whether cross-species amplification works in another species of the genus, were commercially acquired.

To estimate gene flow between the different morphs and populations for this – in terms of evolutionary biology – highly interesting fish species specific genetic markers are needed. The method of choice for this purpose are microsatellite markers, which however, were up to now not available for *G. barreimiae* [6]. Next generation sequencing allows the isolation of a high number of microsatellite loci at moderate costs [7] and multiplexing of loci enables high through-put genotyping. Here, we describe the development of 19 microsatellite markers for *G. barreimiae* using the new technological advantages.

For DNA extraction, we used tissue from a finclip taken from one specimen of the cave population. We followed the DNA-extraction protocol (GEN-IAL First-DNA All-tissue DNA-Kit, GEN-IAL GmbH, Germany) and isolated genomic DNA. Approximate 500 ng DNA was used for 1/20 of a plate for 454 sequencing on a Roche/454 GSFLX sequencer at the Genomics Core Facility University of Leicester, UK. Sequencing generated a total of 36126 reads. The sequence reads were assembled into 193 contigs with size ranging from 500 – 3547 bp by using Newbler (454 Life Sciences, Roche, Switzerland). Of the remaining 35503 non-assembled reads, we also considered 29418 high quality reads (3' trimmed until quality 20; average quality 20; minimal length 100 bp) with sizes between 100 and 731 bp.

We designed primers with the program QDD version 2.1 [8], which incorporates the computer language Perl and the programs ClustalW2 [9], Primer3 [10] and BLAST + [11]. Criteria used to select microsatellite loci

included: appropriate product size (between 75 and 320 bp) and a minimum repeat motif of five subunits consisting of two to six nucleotides. For the primer design, conditions were set for a 20 bp optimum length (18–27 bp) and optimum temperature of 60.6°C, allowing a 12°C temperature difference between the two primers of a primer pair (52–63°C). Optimum GC-content was set on 50% (20–80). After these selections 919 contigs and single reads with at least one microsatellite and featuring suitable primers fulfilled the requirements. After selection of only di- and tetranucleotide repeats and sequences containing just one microsatellite, 767 sequences remained for further analysis. During primer selection, QDD grouped the primers into different primer classes (A to G). We chose only the two most restrictive classes, which fulfill the following conditions: (1) allowing assembled microsatellites, (2) not allowing nanosatellites in the primer region and (3) allowing nanosatellites in the flanking region, i. e., the region between the primer binding site and the actual microsatellite sequence (468 sequences). The final step was to select only sequences, whose microsatellites comprise a motif repeated eight to 22 times. After all this exclusions, 94 sequences retained for primer designing.

We selected primers with a high GC-content displaying no primer dimers and similar annealing temperature (comparing forward- and reverse primers of each primer pair). We tested the potential primer sequences against the original sequence with BioEdit version 7.2.5 [12]. In case we found no primers with matching annealing temperatures, we used NCBI Primer-BLAST [13] to find alternate primer sequences. We checked the primers for dimers with Amplify [14] and manually redesigned the primer sequences if necessary.

Finally 36 different microsatellite loci were left, which were first screened for consistent amplification via PCR. Those primers amplifying a product and working in different samples were chosen for microsatellite analyses of individuals from two surface populations (Wadi Al Falahi N = 24, Misfat Al Abriyeen N = 20). These individuals were collected in accordance with the local legislation of the Sultanate of Oman (Ministry of Environment and Climate Affairs: sampling permission number 13/2012, export permission number 29/2012).

Since we wanted to perform multiplex PCRs, we made use of the program Multiplex Manager version 1.0 [15], a tool for managing and optimizing a multiplex set by minimising the total number of reactions, minimising the annealing temperature differences among primer pairs within each reaction and maximising the size differences between the products amplified from the various loci within one reaction. The remaining 25 markers were put into three different multiplex reactions which were used for electrophoresis. For amplification

these multiplex reactions were further divided as some loci showed a better readable electrophoresis pattern when amplified with a multiplex kit (QIAGEN Multiplex PCR Kit, QIAGEN, Netherlands) and others using the TopTaq kit (TopTaq DNA Polymerase, QIAGEN, Netherlands).

Multiplex PCR was performed either using the multiplex kit in 10 µl reaction volumes (1 µl DNA, 2.5 µl QIAGEN multiplex PCR mastermix, 1 µl primermix, 1 µl Q-solution and 4.5 µl AD), or using the TopTaq kit also with a reaction volume of 10 µl (0.5 µl DNA, 2 µl

Table 1 Primer sequences, repeat motifs, amplification parameters and accession numbers of microsatellite loci in *Garra barreimiae*

Locus name	Primer sequences 5' -> 3'	Primer labelling	Primer C (pMol)	Repeat motif	Reaction	T _a (°C)	Accession Nr.
QLIM	F-ACGGTTGATAGCCACTGAC		2	AG	MPR1	65/60	KM099076
	R-ACTGTCGAGCAATGCAACAC	PET	2				
7GHL	F-GTATTACCGTGGTGTGTGGC	PET	4	AGAT	MPR1	65/60	KM099077
	R-GCGGCGTAGAAATCTGACTG		4				
PH8A	F-ACACTTCCAAAACGGTCGC	VIC	2	AGAT	MPR1	65/60	KM099078
	R-CGGAGCAAACCCAGTGACTA		2				
UHPE	F-ACCACAGCCAGTCAGATATGC	6-FAM	2	AC	MPR1	65/60	KM099079
	R-GCTCAACAAGCTGAACTCCG		2				
JQSO	F-TTGGCTGGAGCGATGGCTG	6-FAM	2	AC	MPR1	65/60	KM099080
	R-AGGGCTACATCACAGACTCAC		2				
DL6X	F-GTGACACATTGGATAATGAATCAGC	NED	2	AC	TTR1	65/60	KM099081
	R-AGCTGGCCTAGATACCTGAC		2				
FYP9	F-CGTCGACCAACGGACTAACG	NED	2	AC	TTR1	65/60	KM099082
	R-GACCGACATCAATCACGCGCT		2				
HOLN	F-ACTGCGCTCGTACCCTATG	6-FAM	2	AC	TTR1	65/60	KM099083
	R-CAGCAGCCGGTAAATAGCTG		2				
JMLR	F-CGATCAGATATGCAATGGCAG	NED	2	AC	MPR2	64/59	KM099084
	R-GTGATCAGGCCTGTGTAAACG		2				
CJHG	F-CTGACCTGATCTACGAAACCG	6-FAM	1.6	AC	MPR2	64/59	KM099085
	R-TGAGTTTGTGGAGTCCAGCG		1.6				
3Z71	F-CGAGGTTTGGTCTCTTCTC	VIC	1.6	AG	MPR2	64/59	KM099086
	R-CAGCCTGGCTAACTCTGAGC		1.6				
9XNC	F-GCCAGTCAGTAGCTGCAAC	PET	1.6	AC	TTR2	64/59	KM099087
	R-GTTGTGCTATTCTGTGTAAGACC		1.6				
3 N43	F-GTGCTGCTTCTCTTCTCG	PET	2	AC	TTR2	64/59	KM099088
	R-ACCAGGAGCCGATCCATAAAC		2				
1EHE	F-GCAGATGTAGTCAATTGACATGC	6-FAM	1	AC	TTR2	64/59	KM099089
	R-GCGAGGGCTTGAACGTATAC		1				
88CM	F-CTGTCGTCCATGGCTGTACC	VIC	1	AG	TTR2	64/59	KM099090
	R-AGGACTGGCATGTTTGCAAT		1				
I6G2	F-AGCATCTGAGCATTCTCTG	VIC	3	AT	MPR3	60/54	KM099091
	R-CCAAACGGTACGTCTTATGG		3				
3ROZ	F-TGGTGGTTTATATGCAGTCTGA	VIC	4	AC	MPR3	60/54	KM099092
	R-ACCAAGGATGTTGTGTCGG		4				
CB75	F-CCATGCATCTACCTACCCAC	PET	2	AGGT	MPR3	60/54	KM099093
	R-TGATCAAATAGATGGTGGCTG		2				
2PUM	F-CATGCCAATCACTCATTC	NED	1	AG	TTR3	60/54	KM099094
	R-ACGTCACGTGTTGTCTATCTTG		1				

Primer C (pMol) primer concentration in pMol, T_a (°C) annealing temperature, Primer labelling: DS-33 Dye set; accession numbers.

Table 2 Population genetic parameters of microsatellite loci of *Garra barreimiae* and cross-species amplification results

Locus name	N MF	N WF	Size (bp)	N _A	H _e MF	H _o MF	Oosterhout MF	HW p-value MF	H _e WF	H _o WF	Oosterhout WF	HW p-value WF	F _{ST}	CA
QLIM ^C	20	24	88 - 92	3	0.00	0.00	0	NV	19.42	20.83	-0.108	0	0.059	22
7GHL	20	24	166 - 250	19	89.10	90.00	-0.022	0.571	93.09	95.83	-0.027	0.777	0.025*	11
PH8A ^C	20	24	126 - 188	15	86.03	95.00	-0.076	0.634	91.58	75.00	0.082	0.995	0.020*	21
UHPE ^C	20	24	89 - 101	7	35.26	30.00	0.107	0.872	54.43	62.50	-0.087	0.437	0.071*	20
JQSO ^C	20	24	147 - 169	7	74.49	75.00	-0.021	0.092	75.89	62.50	0.077	0.966	0.160**	20
DL6X	20	24	112 - 126	8	60.00	65.00	-0.069	0.250	59.75	70.83	-0.120	0.273	0.156**	19
FYP9 ^C	20	24	169 - 235	20	78.46	90.00	-0.096	0.477	95.04	87.50	0.031	0.793	0.068**	22
HOLN ^C	20	24	191 - 233	7	59.49	65.00	-0.067	0.128	39.89	41.67	-0.033	0.419	0.195**	22
JMLR ^C	20	24	93 - 99	4	57.44	65.00	-0.064	0.360	71.37	70.83	-0.007	0.818	0.120**	20
CJHG ^C	20	24	198 - 215	7	54.23	35.00	0.172	0.968	77.66	75.00	-0.002	0.588	0.194**	21
3Z71 ^C	20	24	123 - 127	2	29.62	35.00	-0.194	0.000	0.00	0.00	0	NV	0.170*	22
9XNC ^C	20	24	101 - 112	6	71.67	70.00	-0.012	0.594	74.20	70.83	0.007	0.589	0.061*	22
3 N43	20	24	182 - 184	2	0.00	0.00	0	NV	28.37	33.33	-0.184	0	0.135*	22
1EHE ^C	20	24	106 - 112	4	68.08	80.00	-0.106	0.000	63.03	62.50	0.004	0.068	0.176**	22
88CM ^C	20	24	190 - 198	6	40.90	45.00	-0.069	0.000	73.40	54.17	0.120	0.995	0.216**	21
I6G2 ^C	20	24	115 - 137	8	58.21	60.00	-0.012	0.684	71.45	62.50	0.048	0.781	0.282**	22
3ROZ	20	24	262 - 276	5	14.23	5.00	0.172	0.923	56.34	4.55	0.386	1.000	0.172*	0
CB75	20	24	171 - 199	7	51.41	45.00	0.067	0.767	68.17	75.00	-0.069	0.074	0.005	2
2PUM ^C	20	24	152 - 176	8	38.33	45.00	-0.240	0.000	77.66	87.50	-0.082	0.212	0.121**	22

N number of individuals analysed, Ta (°C) annealing temperature, NA number of alleles, MF Misfat Al Abriyeen, WF Wadi Al Falahi, H_e expected heterozygosity in%, H_o observed heterozygosity in%, null allele estimates (Oosterhout values), HW p-value probability for deviation from Hardy-Weinberg equilibrium, F_{ST} values (*indicates significant values at p < 0.05 and **at p < 0.001); loci marked with ^Cwere positive and polymorphic in *Garra rufa*; CA number of individuals positive in cross-species amplification (tested in 22 individuals of *G. rufa*).

Q-solution, 1 µl buffer, 200 µM dNTPs, 1500 µM Mg⁺⁺, 0.5 µl primermix, 0.5 units TopTaq-polymerase and 5.1 µl AD). According to the three different multiplex reactions, those PCRs were named MPRI/2/3 for multiplex kit reaction 1/2/3 or TTRI/2/3 for TopTaq kit reaction 1/2/3. We followed a two-step cycling protocol for each reaction as follows: initial denaturation at 95°C for 15 min, 2 cycles denaturation at 94°C for 30 s, annealing at 65°C (R1), 64°C (R2), 60°C (R3) for 90 s, extension at 72°C for 60 s. This was followed by 35 cycles denaturation at 94°C for 30 s, annealing at 60°C (R1), 59°C (R2), 54°C (R3) for 90 s, extension at 72°C for 60 s and a final extension at 60°C for 30 min. PCR products were diluted 1:20 with AD, multiplexed according to the reactions and added to a mixture of Hi-Di formamide and size standard (Gene Scan™ 500 LIZ, Life Technologies, USA). After a denaturation at 94°C for 4 min, the products were run on a 3130xl Sequence Analyzer (Applied Biosystems, USA).

After fragment analyses and testing for polymorphic loci, 19 different microsatellite markers proved to be suitable for genotyping. Primer sequences, PCR conditions, and their respective GenBank accession numbers are listed in Table 1. The alleles of all 19 loci (44 individuals) obtained from the resulting electropherograms were identified and binned with GENEMAPPER v5.0 (Applied Biosystems, USA) and manually reviewed. For each population, we estimated the observed (H_o) and expected (H_e) heterozygosities and tested for linkage disequilibria and deviations from Hardy-Weinberg equilibria using Genepop v 4.2.2 [16]. Subsequently we used Microchecker v 2.2 [17] to test for null alleles. We calculated the fixation Index with the program ARLEQUIN v3.5 [18]. All loci proved to be polymorphic in both populations. Results of population genetic analyses are shown in Table 2. For the 19 examined loci, the number of alleles per locus ranged from two to 20. Two loci (DL6X and 9XNC) showed a significant deviation on the 5% level from linkage equilibrium in both populations (Wadi Al Falahi and Misfat Al Abriyeen). Microchecker results revealed evidence for null alleles in one locus (CJHG; frequency after Oosterhout: 0.172) in the population Misfat Al Abriyeen and two loci (PH8A, 3ROZ; frequency after Oosterhout: 0.0815, 0.3868) in the population Wadi Al Falahi. However, since none of these loci showed a constant significant deviation in both populations, we did not exclude the loci from our analyses. Considering each population separately, one locus (CJHG) in the Misfat Al Abriyeen population and 4 loci (PH8A, JQSO, 88CM, 3ROZ) in the Wadi Al Falahi population were not in Hardy-Weinberg equilibrium displaying a deficit of heterozygotes, which is likely due to the occurrence of potential null alleles in the markers CJHG, PH8A and 3ROZ. None of the loci showed a significant deviation in the population Misfat Al Abriyeen as

well as in the population Wadi Al Falahi. Heterozygosities at some loci were very low (UHPE) or even zero (3Z71, QLIM, 3 N43) for one of the populations. For the loci 3Z71, QLIM, 3 N43 this is due to the fact that the numbers of alleles are small (2–3) and in each case one of the two populations possessed only one of the alleles. In the locus UHPE the number of alleles is higher (7), but in each of the two populations one allele occurs with high frequency and most individuals are thus homozygous. The fixation index of 12.83% shows high degree of variation between the two populations (see Table 2 for single loci F_{ST} values).

All markers were tested for cross-species amplification in 22 individuals of *G. rufa* (fin clips were taken from bred individuals kept in the Zoo Vienna). Out of the 19 loci, 14 were positive in *G. rufa* and proved to be polymorphic in at least 75% of the individuals (Table 2). Thus, these 14 loci could be used for population genetic analyses in the commercially exploited species *G. rufa*.

The set of microsatellite loci isolated and tested in the present study will be useful DNA markers for further population genetic studies of *G. barreimiae* and might also be used in other species of this genus. The results from using these microsatellite markers will provide deeper insights into the species' population structure and its genetic diversity. Furthermore, it will be of utmost interest to use these loci in the cave population of *G. barreimiae* and estimate gene flow between the cave and the surface population. This will help to evaluate species delimitation of *G. barreimiae* and to assess whether any cryptic species exist besides *G. barreimiae*.

Competing interests

All authors declare that they have no competing interests.

Authors' contribution

LK and SK participated in the study conception. TW and TR supervised the bioinformatic analyses, SK carried out most laboratory work, searched and characterized microsatellite loci, LK and SK analysed the data and wrote the manuscript. HS initiated the study and provided the samples. All authors read and approved the final manuscript.

Acknowledgements

We thank Wolfgang Arthofer (University of Innsbruck) for coordinating the genome sequencing, Johannes Rath and Christian Baranyi from the University of Vienna for providing help with the Sequence Analyzer, Elisabeth Haring (NHM Vienna) for help with sample collection and critical discussions, Frank Zachos (NHM Vienna) for help and support with population genetic analyses and Toni Weissenbacher (Zoo Vienna) for providing the *G. rufa* samples. Special thanks to Lukas Raneburger for helpful comments and support. We want to express our gratitude to the Ministry of Environment and Climate Affairs of Oman for their cooperation, support and the sampling permits and to Michael Robinson and Reginald Victor, zoologists at the Sultan Qaboos University in Muscat, for their intellectual input. The Federal Ministry for Transport, Innovation and Technology Austria provided financial support in the course of a FEMtech practical.

Author details

¹Natural History Museum Vienna, Central Research Laboratories, Burggring 7, 1010 Vienna, Austria. ²Division of Computational Systems Biology, Department of Microbiology and Ecosystem Science, University of Vienna,

Althanstraße 14, 1090 Vienna, Austria. ³Third Zoological Department, Natural History Museum Vienna Austria, Burgring 7, 1010 Vienna, Austria. ⁴Natural History Museum Vienna, Central Research Laboratories, Burgring 7, 1010 Vienna, Austria.

Received: 17 March 2014 Accepted: 5 August 2014

Published: 12 August 2014

References

- Banister KE, Clarke MA: **The freshwater fishes of the Arabian peninsula.** In *Sci Results Oman Flora Fauna Surv 1975 J Oman Stud Spec Report No 1*. Edited by Harrison DL. Muscat, Sultanate of Oman: Ministry of Information and Culture; 1977:111–154 [The Journal of Oman Studies Special Report. No. 1].
- Krupp F: **Fishes of Saudi Arabia: freshwater fishes of Saudi Arabia and adjacent regions of the Arabian Peninsula.** *Fauna Saudi Arab* 1983, **5**:568–636.
- Banister KE: **A subterranean population of *Garra barreimiae* (Teleostei: Cyprinidae) from Oman, with comments on the concept of regressive evolution.** *J Nat Hist* 1984, **18**:927–938.
- Kruckenhauser L, Haring E, Seemann R, Sattmann H: **Genetic differentiation between cave and surface-dwelling populations of *Garra barreimiae* (Cyprinidae) in Oman.** *BMC Evol Biol* 2011, **11**:172.
- Krupp F: **Freshwater Fishes of the Wadi Batha drainage.** *J Oman Studies Spec Rept* 1988, **3**:401–404.
- Guichoux E, Lagache L, Wagner S, Chaumeil P, Léger P, Lepais O, Lepoittevin C, Malausa T, Revardel E, Salin F, Petit RJ: **Current trends in microsatellite genotyping.** *Mol Ecol Resour* 2011, **11**:591–611.
- Gardner MG, Fitch AJ, Bertozzi T, Lowe AJ: **Rise of the machines—recommendations for ecologists when using next generation sequencing for microsatellite development.** *Mol Ecol Resour* 2011, **11**:1093–1101.
- Megléczy E, Costedoat C, Dubut V, Gilles A, Malausa T, Pech N, Martin J-F: **QDD: a user-friendly program to select microsatellite markers and design primers from large sequencing projects.** *Bioinformatics* 2010, **26**:403–404.
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McLarkin MA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins D: **ClustalW and ClustalX version 2.** *Bioinformatics* 2007, **23**:2947–2948.
- Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen S: **Primer3—new capabilities and interfaces.** *Nucleic Acids Res* 2012, **40**:e115.
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL: **BLAST+: architecture and applications.** *BMC Bioinformatics* 2009, **10**:421.
- Hall TA: **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.** *Nucleic Acids Symp Ser* 1999, **41**:95–98.
- Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden T: **Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction.** *BMC Bioinformatics* 2012, **13**:134.
- Engels B: **Amplify - a freeware Macintosh program for simulating and testing polymerase chain reactions.** [http://engels.genetics.wisc.edu/amplify/]
- Holleley CE, Geerts PG: **Multiplex Manager 1.0: a cross-platform computer program that plans and optimizes multiplex PCR.** *Biotechniques* 2009, **46**:511–517.
- Rousset F: **Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux.** *Mol Ecol Resour* 2008, **8**:103–106.
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P: **Micro-Checker: Software for Identifying and Correcting Genotyping Errors in Microsatellite Data.** *Mol Ecol Notes* 2004, **4**:535–538.
- Excoffier L, Laval G, Schneider S: **Arlequin ver. 3.5: An integrated software package for population genetics data analysis.** *Evol Bioinformatics Online* 2005, **1**:47–50.

doi:10.1186/1756-0500-7-522

Cite this article as: Kirchner et al.: Characterization of 19 new microsatellite loci for the Omani barb *Garra barreimiae* from 454 sequences. *BMC Research Notes* 2014 **7**:522.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



Chapter I: Adaptation to cave environments, gene flow between troglomorphic cavefish and conspecific surface populations

Paper 2: Investigating gene flow between the blind cavefish *Garra barreimiae* and its conspecific surface populations

SCIENTIFIC REPORTS

OPEN

Investigating gene flow between the blind cavefish *Garra barreimiae* and its conspecific surface populations

Sandra Kirchner^{1,2}, Helmut Sattmann³, Elisabeth Haring^{1,2}, Lukas Plan⁴, Reginald Victor⁵ & Luise Kruckenhauser²

Cave-dwelling taxa often share the same phenotypic modifications like absence of eyes and pigmentation. These “troglomorphic characters” are expressed in the populations of *Garra barreimiae* from the Al Hoota Cave and nearby Hoti Pit in Northern Oman. Surface morphotypes of this cyprinid species are common throughout the distribution area. Very rarely individuals with intermediate phenotypes can be found. In the present study, potential gene flow between cave and surface populations was tested and population structure within five sampling sites was assessed. Overall, 213 individuals were genotyped at 18 microsatellite loci. We found that the cave populations have lower genetic diversity and are clearly isolated from the surface populations, which seem to be sporadically in contact with each other. The results indicate a recent genetic bottleneck in the cave populations. Thus, it can be assumed that during climatic changes the connection between cave and surface water bodies was disjoined, leaving a subpopulation trapped inside. Nevertheless, occasional gene flow between the morphotypes is detectable, but hybridisation seems only possible in cave habitat with permanent connection to surface water. Individuals from surface sites bearing intermediate phenotypes but cave genotypes imply that phenotypic plasticity might play a role in the development of the phenotype.

Troglomorphic cave fish always share a similar phenotype exhibiting reduction or absence of eyes and pigmentation. These “typical” cave characters developed independently among taxa and can vary strongly in their expression, often even within species or populations¹. The best studied organism concerning this matter is the characid *Astyanax mexicanus* (Filippi, 1853)^{2–10}. Besides *Astyanax*, the Afro-Asian cyprinid genus *Garra*, which comprises over 100 freshwater fish species can be considered as an exemplary system. It includes several cave-dwelling species or populations within species exhibiting varying degrees of troglomorphy. A mental disc (modified lower lip used to adhere to the substrate) is characteristic to most species of the strictly continental genus which is supposed to have originated in the early Miocene around 23 Myr in Southwestern China^{11–13}. Typical troglomorphic characters as complete lack of pigmentation and complete anophthalmy are displayed in the Iranian cave barb *Garra typhlops* (Bruun & Kaiser, 1944)¹⁴ and in the Iraqi blind barb *Garra widdowsoni* (Trewavas, 1955)¹⁵, both species being cave-dwelling without any surface occurrences known. This is also the case for the Somalian cave-dwelling species *Phreatichthys andruzzii* Vinciguerra, 1924, for which Colli and colleagues¹⁶ suggested a taxonomic update due to the position of *P. andruzzii* within the genus *Garra* in their phylogenetic tree. This species, which according to Calderoni and colleagues¹⁷ evolved in complete darkness for at least 2.8 million years, exhibits also pronounced troglomorphic traits¹⁸.

In the centre of the present study is *Garra barreimiae* FOWLER & STEINITZ, 1956, which, like *A. mexicanus* comprises both, surface and cave dwelling populations. *Garra barreimiae* is endemic to the southeastern Arabian Peninsula. It is quite common in the United Arab Emirates and northern Oman along the Hajar Mountains where

¹University of Vienna, Department of Integrative Zoology, Althanstraße 14, 1090, Vienna, Austria. ²Natural History Museum Vienna, Central Research Laboratories, Burgring 7, 1010, Vienna, Austria. ³Natural History Museum Vienna, Third Zoological Department, Burgring 7, 1010, Vienna, Austria. ⁴Natural History Museum, Department for Geology & Paleontology, Burgring 7, 1010, Vienna, Austria. ⁵Sultan Qaboos University, Department of Biology, Al Khoudh, Muscat, 123, Oman. Correspondence and requests for materials should be addressed to S.K. (email: sandra.kirchner@nhm-wien.ac.at)

it inhabits all kinds of water puddles, streams and even artificial water drainage systems (aflaj/falaj) used by the locals to water their gardens.

In 1980 Dunsire and Gallagher discovered a subterranean population of *G. barreimiae* in the Al Hoota Cave near the Jebel Akhdar in northern Oman. Due to the morphological similarities, it was assigned to the species *G. barreimiae* by Banister¹⁹. Yet, the individuals of this population exhibit a typical troglomorphic phenotype such as no pigmentation and lack of externally visible eyes, while scales are present. Besides the more obvious differences, Banister also described the troglomorphic specimens to exhibit smaller optic lobes and reduced squamation. In the closely located cave Hoti Pit²⁰, we discovered another population of *G. barreimiae* which also displayed the cave morphotype.

Kruckenhauser and colleagues²¹ were the first to conduct molecular genetic analyses in this species. They sequenced a part of the mitochondrial (mt) control region (CR) in individuals from both the Al Hoota Cave population as well as from several surface sampling sites. Their results showed that all individuals collected in the cave belong to the same haplogroup. However, this haplogroup also contained individuals from nearby surface localities. Most of the latter individuals exhibited an intermediate phenotype suggesting that the cave and surface populations might be still in contact or have been until recently. The genetic distances between individuals from Al Hoota Cave and individuals from nearby surface localities are quite small implying a recent origin of the cave population. Yet, the question remained whether a nuclear marker would confirm the close phylogenetic relationships between surface and cave populations as suggested by the mt CR. Within the genus *Garra*, *G. barreimiae* represents the rare case where both morphotypes (troglomorphic/surface) are still existing within one species. The aim of the present study was to better understand this system in which both forms co-exist. More specifically, we asked, if there is recent gene flow between those populations and whether it does occur equally in both directions? Are the individuals exhibiting an intermediate phenotype the result of recent hybridization between cave and surface individuals or a result of phenotypic plasticity? What is the genetic structure and diversity of cave and surface populations of *G. barreimiae*? Do the nuclear markers confirm the presumed bottleneck events as suggested by mt data?

To address those questions we utilized 18 of the 19 microsatellite loci designed for *G. barreimiae* by a 454 sequencing approach²² and analysed them in Al Hoota Cave as well as several surface populations in Northern Oman. Moreover, morphological measurements of the eye size were taken to assess the amount of reduction. The results were compared with those derived from mt CR sequences.

Material and Methods

Sampling. Tissue samples (finclips) were taken from a total of 213 individuals collected at five sampling sites located at the south side of Central Hajar Mountain chains in northern Oman, some 25 km northwest of the town Nizwa. These specimens were captured via fishing net and immediately anesthetized with clove oil and stored in 80% ethanol. Two of the sampling sites are mere cave locations, two are typical surface sites and one location consists of subterranean water bodies connected to surface water bodies (for more detail see Supplementary Figure S1):

- (1) The Al Hoota Cave (AH, also Khaf Hoti) is a 5 km long spacious cave southeast of the Jebel Shams mountain peak (2980 m above sea level (asl)) that developed in cretaceous limestone. The cave forms an underground flood system: during flood events water from the above lying Wadi Hoti sinks into the upper entrance of the cave (990 m asl) and re-emerges 262 m deeper at the lower entrance into the Wadi Falahi²⁰. These floods occur occasionally every few years or sometimes more than once a year. Between such events there is almost no running water in the cave, but there are four extensive lakes and some smaller ones. These water bodies hold only troglomorphic fish²⁰. Between the lakes there are several vertical drops that seem to be an insurmountable barrier for fish (from lower to higher lakes). From the passage morphology and transported cave sediment it can be estimated that during floods flow velocities in some parts reach more than 10 m/s and discharge rises up to few hundred litres per second. The average temperature in the cave is 25 °C and humidity between 40–70% up to 100% after floods when the water in the lakes is high and sumps stop air ventilation²³. Except for floods, the wadis up- and downstream of the cave are dry. The lower entrance of the cave was developed as a show cave. Geographic coordinates: longitude 57.3572; latitude 23.0844.
- (2) Hoti Pit (HP) is a pitfall in Wadi Hota 400 m upstream and some 20 m higher than the upper entrance to Al Hoota Cave. During floods Hoti Pit takes some water from the Wadi Hoti but according to sediments and morphology only few m³/s. The entrance is a 12 m deep pit and in the lake at its bottom no fish were found. The known end of the 380 m long cave is formed by another lake harbouring only troglomorphic fish. The part of the cave where the passage ceiling reaches into the water and forms a sump, has not been explored yet and the reappearance of the water is unknown. As there are no signs of backflooding at the terminal sump it can be concluded that the cave behind the sump is spacious. Geographic coordinates: longitude 57.3726; latitude 23.1015.
- (3) Ghubrat Tanuf (GT) is a cave located in the Wadi Falahi at 670 m asl and approx. 1 km downstream of the lower Al Hoota Cave entrance. A small stream emerges from the cave entrance that feeds a pool in the Wadi (called Ain Al Msalla). The rather narrow cave is an almost horizontal water table cave with a length of 690 m. Estimated discharge values reach from less than 1 l/s to few l/s. During floods, water from the Wadi enters the cave for some 100 m as indicated by wood and garbage transported into the cave. In the upstream end the cave splits up into two separate streams which both emerge from sumps. The upper stream is interrupted by a 9 m high cascade where no fish were found in the upstream part. The lower one is only marginally higher than the entrance, and surface fish can be found throughout this part. In this cave, mostly individuals displaying the surface morphotype occur, but also few individuals exhibiting

typical cave or intermediate phenotypes were found. Geographic coordinates: longitude 57.3681; latitude 23.0718.

- (4) The sampling site Wadi Falahi (WF) is a surface water body about 200 m underneath the lower Al Hoota Cave entry which represents the beginning of the Wadi Falahi, a riverbed which reaches ~4.5 km downstream comprising also the sampling site Ghubrat Tanuf/Ain al Msalla. Most of the time, the wadi lies dry, but single puddles may remain for several weeks or even months after rains. Fish were collected at a rather shallow water body located about 680 m asl which appears to be drying out regularly during summer. At this sampling site, usually the surface form occurs but also few individuals exhibiting intermediate phenotype were found. Geographic coordinates: longitude 57.3539; latitude 23.0804.
- (5) Misfat al Abriyeen (abbr. Misfat, MF) is a small mountain village located at an impressive wadi at 1000 m asl about 7 km northwest of Al Hoota Cave. At the end of (0.7 km behind) the village a spring emerges from big boulders in the wadi and is tapped into an artificial water drainage system (aflaj). All *G. barreimiae* from Misfat al Abriyeen had the surface phenotype and all specimens analysed were collected from this aflaj. Geographic coordinates: longitude 57.3159; latitude 23.1481.

In Summary, within the two enclosed caves (Al Hoota Cave and Hoti Pit), only the troglomorphic form was found. The water body of Ghubrat Tanuf cave is constantly connected to the surface water body of Ain Al Msalla and harbours mostly fish bearing a typical surface phenotype, but also few individuals with intermediate or troglomorphic phenotype. A similar phenotypic composition also applies for the surface water body Wadi Falahi, whereas in Misfat only fish with the surface phenotype were found.

DNA extraction, PCR amplification and genotyping. We extracted genomic DNA from tissue (finclips preserved in 80% ethanol) following the QIAGEN DNeasy Blood & Tissue extraction kit protocol. All samples were profiled with 19 microsatellite loci by using primers that were specifically designed for *G. barreimiae*²². One microsatellite marker published in ref. 22 turned out to be working improperly (~35% missing values over the whole dataset) and was therefore excluded from the data analyses. Thus, the comprehensive analysis was performed with 18 loci. In the current study, we applied the multiplex amplification protocol established in our previous study using also the same set of labelled primers. The microsatellite amplification products were visualized on a 3130xl Sequence Analyzer (Applied Biosystems, USA). The FASTA files delivered by the sequencer were imported into the program GeneMapper version 5.0 (Applied Biosystems, USA) and analysed using adapted settings for automatic allele recognition (bins). All allele calls were also checked manually to avoid misinterpretation of noise data or artefacts and to review ambiguous alleles. Amplifications that were too weak to resolve the peaks or which had any extra peaks were re-amplified and rerun up to four times. Finally, any remaining unresolved alleles were treated as missing data. A list containing the reviewed genotypes of all individuals analysed was used for data analysis (Supplementary Table S1). In this manuscript, the term “genotype” is used to describe the microsatellite allele composition characteristic for a population or morphotype.

In 96 selected samples, we amplified a section of the mt CR gene via PCR reaction (25 µl reaction volume: 1 µl DNA, 5 µl Q-solution, 2.5 µl buffer, 500 µM dNTPs, 1.5 µl (1.5 mM) Mg2+, 0.25 µl of each primer (50 pmol/µl of both primers, 0.1 µl TopTaq-polymerase (0.5 units), 13.9 µl a.d.). The following primers were used: Thr1 + 5'-GCATCGGTCTTGTAATCCGA-3', and CR4- 5'-TTGGGCGTCGGCGGTGAGAG-3'; designed by ref. 21). The cycling protocol was: initial denaturation at 94 °C for 4 min, 35 cycles denaturation at 94 °C for 30 s, annealing at 58 °C for 30 s, extension at 72 °C for 40 s, and final extension at 72 °C for 7 min. The resulting PCR products were purified and sequenced in both directions at LGC genomics (Berlin, Germany) using the PCR primers. The obtained sequences were edited in BioEdit version 7.2.5²⁴, and aligned with the program MEGA version 6.06²⁵ together with the 122 sequences retrieved from Kruckenhauser *et al.*²¹. The final alignment was 399 bp long and comprised a total of 218 sequences from nine sampling sites. The haplotype network was constructed using the software PopART version 1.7 (available at <http://popart.otago.ac.nz>) and the Median-joining (MJ) algorithm proposed by Bandelt, Forster and Röhl²⁶.

Data analysis. Genetic diversity. We explored the data mostly by employing the statistical computing software environment R version 3.1.3²⁷. After checking our data set for completeness and missing values (0.57% mean missing data within the whole data set), we calculated general information like the number of alleles, the diversity index and evenness per locus with the *R-package poppr* version 2.1.0²⁸. Furthermore, we calculated the expected (H_e) and observed (H_o) heterozygosity per locus in each population with the *R-package adegenet* version 2.0.0²⁹ and the mean H_e and H_o over all loci per population with the *R-package diveRsity* version 1.9.89³⁰. In addition, we tested for genotypic linkage disequilibrium with the software *Genepop* version 4.3³¹. Deviations from the Hardy-Weinberg equilibrium (HWE) were estimated using both the global test and exact test in *Genepop* executing 10000 Markov chain Monte Carlo (MCMC) runs for 100 batches, each with 5000 iterations. The software *Micro-Checker* v. 2.2.3³² was used to check for the existence of microsatellite null alleles in the data set. Although there was only one population, i.e., Hoti Pit, with no signal of microsatellite null alleles at any locus, the tests showed that microsatellite null alleles at a low frequency might be present in populations Ghubrat Tanuf (6 loci showing signs of microsatellite null alleles) and Wadi Falahi (7 loci showing signs of microsatellite null alleles), whereas in the populations Al Hoota and Misfat potential microsatellite null alleles are only present at one locus. However, since none of these loci showed a constant significant deviation in all populations, and further analysis indicated population admixture between Ghubrat Tanuf and Wadi Falahi (see below) we did not exclude these loci from our analyses. We tested the data set also for linkage disequilibrium and the results showed that some locus pairs exhibit evidence of linkage disequilibrium in single populations, but never in all populations.

In order to statistically describe the level of inbreeding or differentiation among and within populations, we calculated the fixation indices as implemented in the *R-package pegas* version 0.8.2³³. This function computes F_{it} ,

Population	N	Na	A	Ar	pA	pA%	H _o	H _e	F _{is}	p	Ld	Lp%	HWE
GT	73	119	6.61	4.91	10	8.40	0.48	0.56	0.14	0.24	PH8A*, UHPE*, FYP9*, I6G2*, 3ROZ*	94.44	0.000
AH	34	68	3.78	3.08	8	11.76	0.37	0.37	0.03	0.64	UHPE*	83.33	0.374
HP	17	41	2.28	2.15	1	2.44	0.26	0.25	−0.01	0.23		66.67	0.522
MF	39	77	4.28	3.68	9	11.69	0.49	0.50	0.04	0.43	CJHG*, 9XNC	88.89	0.027
WF	50	123	6.83	5.29	12	9.76	0.50	0.62	0.21	0.19	JQSO*, CJHG*, 88CM*, I6G2*, 3ROZ*, 2PUM*	94.44	0.000

Table 1. Sample information and population genetic parameters of the five populations investigated. N Number of analysed individuals, Na number of alleles per population, A mean number of alleles per locus, Ar allelic richness, pA number of private alleles, pA% percentage of private alleles per population based on the actual amount of alleles per population, H_o mean observed heterozygosity over all loci, H_e mean expected heterozygosity over all loci, F_{is} inbreeding coefficient, p (mean) p-value for Hardy-Weinberg equilibrium (global test, mean over all loci), Ld Loci, that deviated from HW-equilibrium (loci with potential microsatellite null alleles are marked with an asterisk), Lp% percentage of polymorphic loci, HWE p-value for Hardy-Weinberg equilibria (global test).

F_{st} and F_{is} for each locus in the data set according to the formulae of Weir and Cockerham³⁴. Subsequently, we also calculated the pairwise F_{st}-values between all populations with the *Microsoft Excel Add-In GenAlEx version 6.5*^{35,36}. The inbreeding coefficient F_{is} after Weir and Cockerham over all loci per population was calculated with the program *FSTAT*³⁷.

Allelic richness per locus per population was computed to take the differing sampling sizes of populations into account. This was done by using 1000 re-samples of the smallest sample in the data set (Hoti Pit with N = 17) via the *R-package diveRsity*. Private alleles were estimated with the *R-package poppr* and the percentage of private alleles per population was calculated based on the actual amount of alleles per population.

Population Structure Analysis and Differentiation. The software *structure*³⁸ was used to perform population structure analyses. We employed the admixture model and tested for clusters from K = 1 to K = 7 for the whole data set and K = 1 to K = 3 for the cave subset, and for each assumed K 20 iterations per run were performed to find the correct number of clusters (K). Length of burnin was set to 100000 and the number of MCMC repeats after burnin was 500000. The results were then revised with the online program *structure Harvester* web version 0.6.94³⁹ to explore which K has the highest probability and ΔK. The next step was to summarize the iterations and to minimize the variance across all twenty iterations of the selected ΔK value with the software *CLUMPP* version 1.1.2⁴⁰. For high quality visualization of the combined results obtained via *CLUMPP*, we used the software *Distruct* version 1.1⁴¹.

A principle component analysis using the *R-package ade4* version 1.7–2⁴² was performed. By using the *scale-Gen* function in the *R-package adegenet*, which replaces the missing values with generated mean allele values, it was possible to include all individuals.

To infer whether gene flow is recent we used the program *GeneClass2*⁴³ for real-time estimation of first-generation migrants. We used the L_{home} likelihood computation and the Bayesian method criterion⁴⁴ and applied the simulation algorithm suggested by Paetkau *et al.*⁴⁵. The calculation was conducted with a number of 10000 simulated individuals and Threshold was set on 0.05.

Morphological Eye Measurements. We measured eye diameter of intermediate individuals which were collected at the sampling sites Ghubrat Tanuf (N = 11) and Wadi Falahi (N = 14). As not all specimens were fully grown adults, we scaled the eye diameter as percentage proportional to the head length (from snout tip to end of operculum) to allow comparison between individuals of different sizes. Subsequently, we also measured eye diameter of surface individuals with external fully developed eyes and pigmentation (N = 9, three specimens of each surface site - Ghubrat Tanuf, Wadi Falahi and Misfat Al Abriyeen). The according measurements and photographs are provided in Supplementary Table S3 and Supplementary Figure S2.

Results

Genetic diversity within sampling locations. At first, in order to assess the genetic variation among and between specimens from different sampling sites in general, we calculated the descriptive statistics (Table 1) for the 18 unlinked microsatellite loci. All loci were polymorphic over the whole data set (between 2 and 23 alleles per locus). Concerning single sampling locations, there were a few monomorphic markers in each. In Ghubrat Tanuf the range of allele number per locus ranged from 1 (3Z7I) to 21, in Al Hoota from 1 (3Z7I, 3N43, 2PUM) to 14, in Hoti Pit from 1 (QLIM, JQSO, 3Z7I, 3N43, CB75, 2PUM) to 8, in Misfat from 1 (QLIM, 3N43) to 10 and in Wadi Falahi from 1 (3Z7I) to 18. None of these markers were monomorphic in all populations hence we did not exclude them from the data set. We repeated all descriptive statistic calculations with the data set divided into groups of morphotypes (surface/cave) according to the results of the structure analysis (see Fig. 1) disregarding the geographic origin (population) of the individuals (data shown in Supplementary Table S2).

The number of alleles (surface: 119 in Ghubrat Tanuf; 77 in Misfat, 123 in Wadi Falahi; cave: 68 in Al Hoota; 41 in Hoti Pit) and the proportion of polymorphic loci (surface: 17 in Ghubrat Tanuf, Misfat and Wadi Falahi; cave: 14 in Al Hoota, 12 in Hoti Pit) were generally higher in surface than in cave populations. Regarding the genetic

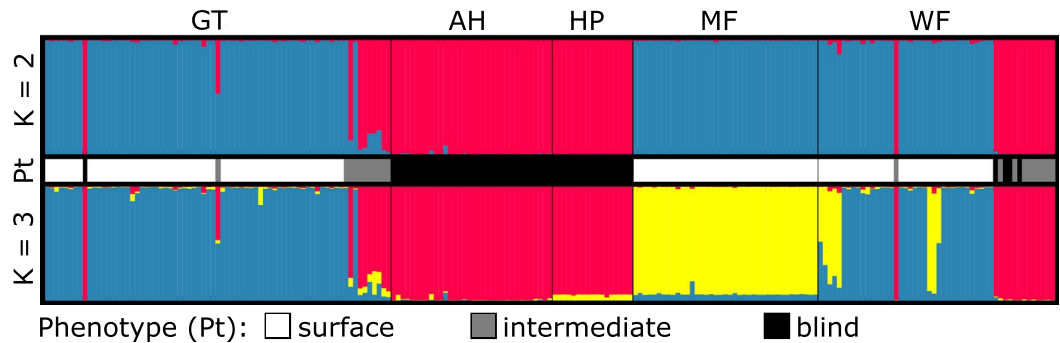


Figure 1. Population structure analysis based on the complete data set. The two colours in the top section ($K = 2$) represent the two most probable clusters (blue = surface cluster, magenta = cave cluster), whereas the bottom section shows the population structure analysis with the second highest probability ($K = 3$). The three colours represent two surface clusters (blue and yellow) and a cave cluster (magenta). The bar between the two sections resembles the phenotype (Pt) of the individuals analysed: white = surface, grey = intermediate, black = blind.

variability, we corrected these results by taking the variable sampling sizes into account and calculated the allelic richness according to the sample with the fewest individuals (Hoti Pit with 17 individuals). The average allelic number turned out to be slightly lower in the cave populations (3.08 in Al Hoota, 2.15 in Hoti Pit) than in the surface populations (4.91 in Ghubrat Tanuf, 3.68 in Misfat, 5.29 in Wadi Falahi). Furthermore, the calculated mean H_e over all loci was higher in surface populations (55.77% in Ghubrat Tanuf, 50.09% in Misfat, 62.29% in Wadi Falahi) compared to the cave populations (37.45% in Al Hoota, 25.31% in Hoti Pit). In the populations Ghubrat Tanuf and Wadi Falahi H_e was higher than H_o , whereas both parameters were similar in the other populations. The most private alleles were found in the Al Hoota and Misfat populations (proportion of private alleles: 11.76% and 11.69%, respectively). In the populations Ghubrat Tanuf and Wadi Falahi, less private alleles were detected (8.40% and 9.76%), whereas the fewest private alleles were found in the Hoti Pit population (2.44%).

Deviations from Hardy-Weinberg equilibrium (HWE) for each population and locus showed that the surface populations exhibit more loci with significant deviation (5% level) than the cave populations (surface: 6 Wadi Falahi, 4 Ghubrat Tanuf, 2 Misfat; cave: 1 Al Hoota, 0 Hoti Pit). The results from the global HW test indicate that the Ghubrat Tanuf, Misfat and Wadi Falahi populations are not in HW equilibrium, whereas the two cave populations Al Hoota and Hoti Pit showed no significant deviation for HW equilibrium. These results remained after repeating the global HW test excluding individuals with an intermediate phenotype (data not shown). The reason for deviations from HWE in the surface sites is most likely due to the presence of microsatellite null alleles. The few loci showing deviations from HWE are the same for which potential microsatellite null alleles were detected. Another reason could be occasional migration from one surface population to another (indicated in Fig. 1), which is a violation of the assumptions for testing HWE⁴⁶.

The inbreeding coefficient F_{is} is positive in all but the Hoti Pit populations indicating at least some degree of heterozygote deficiency. However, these F_{is} -values are close to zero (which would indicate complete panmixia) in Hoti Pit, Al Hoota and Misfat, while the highest values were found again in Ghubrat Tanuf (0.14) and Wadi Falahi (0.21).

Genetic differentiation between sampling localities and potential gene flow between cave and surface populations.

The next step was to learn more about the genetic population structure and to test, if groups can be detected and whether these groupings are in accordance with the sampling sites and/or phenotypes. In the *structure* analysis $K = 2$ yielded the highest probability, and also the ΔK was highest for $K = 2$ (Fig. 1). In fact, the clear discrimination here lies between the populations from Ghubrat Tanuf, Misfat and Wadi Falahi (surface or constant contact to surface) on the one hand and the enclosed cave populations of Al Hoota and Hoti Pit. The latter seem to be strongly isolated from the surface populations. Nonetheless, the situation in the surface cluster is more complex: some individuals exhibit an intermediate phenotype. They belong genetically either to the cave cluster or they represent an intermediate genotype. These intermediate phenotypes are only found in Ghubrat Tanuf and Wadi Falahi. Of these two localities, only Ghubrat Tanuf has contact to underground water bodies.

When using the second highest probability at $K = 3$, further structuring can be found within the surface cluster. The Misfat population is differentiated from Ghubrat Tanuf and Wadi Falahi populations, although some individuals belonging genetically to Misfat can also be found in the Wadi Falahi cluster (Fig. 1).

In respect to the result of the *structure* analyses ($K = 2$), we split the data set accordingly into two groups, a cave cluster including the Al Hoota and Hoti Pit individuals and a surface cluster comprising individuals from Ghubrat Tanuf, Misfat and Wadi Falahi.

Besides, we also performed a *structure* analyses with a subset of individuals exclusively collected at the two cave sites (Al Hoota, Hoti Pit) to explore whether they are genetically homogenous or if they are differentiated from each other. The analysis resulted in the highest probability of $K = 2$, with the two clusters correlating with the two sampling sites (Fig. 2). Interestingly, the Al Hoota population contains also two individuals bearing at least part of the Hoti Pit genotype whereas no Al Hoota genotype was found in the Hoti Pit population. A summary

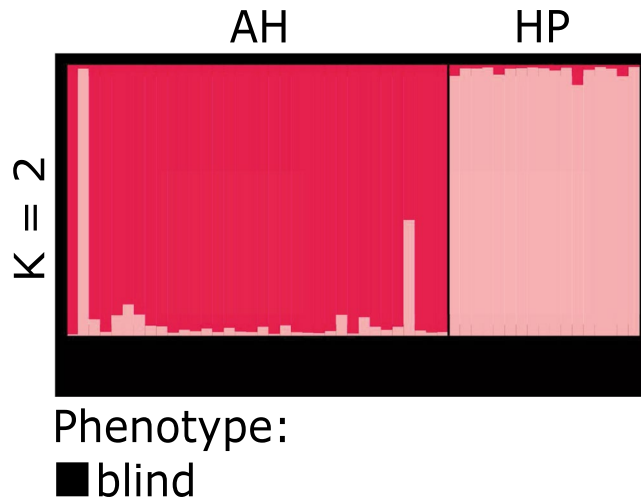


Figure 2. Population structure analysis based on the cave data set. The figure ($K=2$) represents the two most probable clusters (magenta = Al Hoota, light pink = Hoti Pit) resulting from the structure analysis. All individuals belonging to this data set bear a cave phenotype (indicated by the bar below).

of the geographic position, possible connections and the according phenotype and genotype affiliation of the individuals analysed at each sampling site is provided in Fig. 3.

A MJ network based on a 399 bp long fragment of the CR gene is shown in Fig. 4. The GenBank accession numbers for the 96 CR sequences (MF183116–MF183211) are listed in the Supplementary Table S4. The network consists of 41 haplotypes, of which five are rather frequently occurring in ten or more individuals (HT 1 to 5), twelve haplotypes occur in two to eight individuals (HT 6 to 17) and 24 haplotypes are represented by only one individual (HT 18 to 41). All samples of troglomorphic cave fish are grouped in a “cave” haplogroup (HT 2, 8, 11 to 13, 20 to 25) which also contains haplotypes of all intermediate/blind individuals from Wadi Falahi (HT 2, 8, 18, 19, 26, 27) as well as three intermediate and one blind individual from Ghubrat Tanuf (HT 2). Not a single surface phenotype individual is present in this haplogroup. The remaining intermediate individuals from Ghubrat Tanuf appear together with surface individuals from either Ghubrat Tanuf or Wadi Falahi (HT 1, 6). This outcome generally supports the findings of Kruckenhauser *et al.*²¹, but contains also samples from Hoti Pit, more samples of intermediate individuals and additional samples from the other sampling sites (see also Supplementary Table S4).

The fixation indices (F_{st}) and pairwise F_{st} -values (Nei’s unbiased genetic distance) between all pairs of populations were calculated (heterozygosities are weighted by group sizes). The values range from 0.296 (Misfat and Hoti Pit) to 0.020 (Ghubrat Tanuf and Wadi Falahi) (Table 2).

The results of the principal component analyses (PCA) display a tripartition roughly dividing the data set in a group comprising most of the Ghubrat Tanuf and Wadi Falahi samples, one cave group containing both Al Hoota and Hoti Pit samples and the third group including all samples from Misfat (Fig. 5A). In Fig. 5B, one can see a clear separation between samples of the two cave sites (as also depicted in Fig. 2). However, these separations between the surface samples are not so clear as some Wadi Falahi and Ghubrat Tanuf samples appear to be closer to the cave group than to the other surface clusters. This is due to individuals with an intermediate or blind phenotype from these localities (Fig. 5C). Regarding the samples of each locality, the Wadi Falahi population is by far the most variable one, with samples appearing in the diagram close to all other populations. This is also depicted by the corresponding inertia ellipses. The function reveals that ~24% of the information (variance) are explained by the first two components in the whole data set. Regarding the cave and surface subsets (Fig. 4B and C), the first two principal components explain ~19% of the variance in the cave subset and ~14% in the surface subset.

Origin of individuals with intermediate phenotype. The presence of individuals with an intermediate phenotype raises questions whether this phenotype is the result of hybridisation or different environmental conditions (e.g. exposure to sunlight) on the species’ phenotypic variability. These intermediate phenotypes show the whole palette from normally developed eyes without pigmentation to pigmentation without visible eyes at all (Supplementary Figure S2). In most cases, these specimens have microphthalmic eyes and reduced pigmentation. After we found such specimens at Wadi Falahi and Ghubrat Tanuf, we intensified the search for intermediate individuals, which are only rarely found. We added all these individuals to our investigation. Consequently, the proportion between surface and intermediate specimens presented in this study (see below) does not reflect their actual occurrence in the population. However, single juvenile individuals bear eyes only one-sided or appear to be extraordinary pale or dark, respectively, considering the habitat in which they were caught (see also Fig. 5 in ref. 21). Since larvae of the cave morphotype also exhibit eyes when they hatch (which stop to develop and are overgrown by surrounding dermal tissue later in development), it is hard to differentiate whether these individuals would have developed into a troglomorphic or intermediate phenotype¹⁹. Hence, we chose a conservative approach and treated them as intermediate phenotypes (see phenotype assignment in Fig. 1).

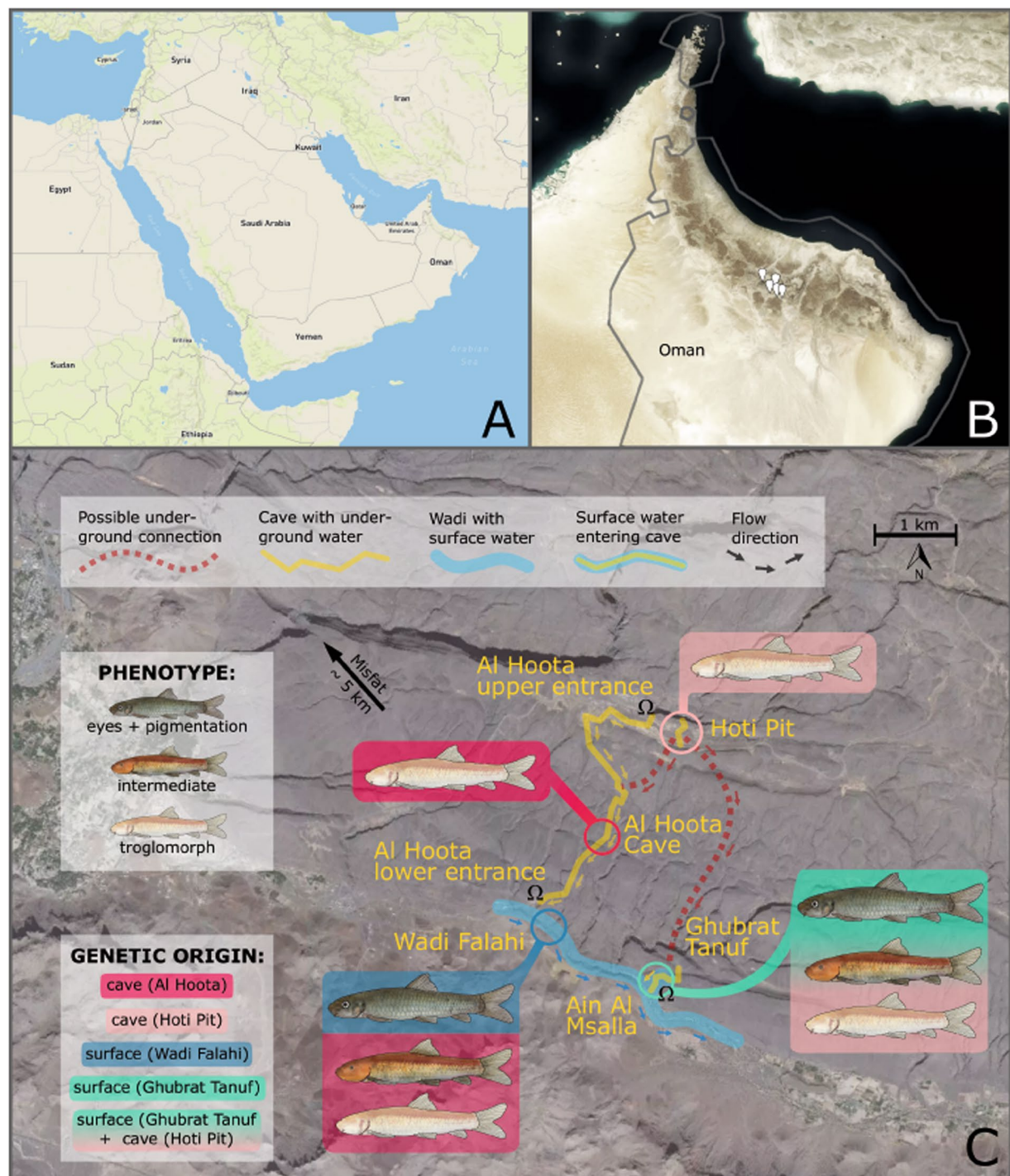


Figure 3. Geographic map of sampling area in detail. The figure shows a detailed map of the sampling area in northern Oman in the midst of the Hajar Mountains pointing out the sampling sites and showing the system in general (excluding Misfat). Information on which phenotypes were found at the sampling sites and their genetic assignment is provided. Caves with underground water bodies are displayed with a yellow line, the wadi (i.e., surface water bodies) is marked with a blue line and possible subterranean connections are indicated with the dotted red line. The contact zone between underground and surface water (passage from Ain al Msalla to Ghubrat Tanuf) is marked with a yellow line. Arrows indicate flowing direction of the water stream. The map was created with the program TileMill version 0.10.1 (<https://tilemill-project.github.io/tilemill/>).

Of the eleven intermediate, juvenile or blind individuals collected at Ghubrat Tanuf (among 62 normal surface individuals) one individual is truly troglomorphic (eyes only rudimentary, no pigmentation) with a 100% cave genotype, five individuals exhibit an intermediate phenotype (microphthalmic eyes and weak to no pigmentation)

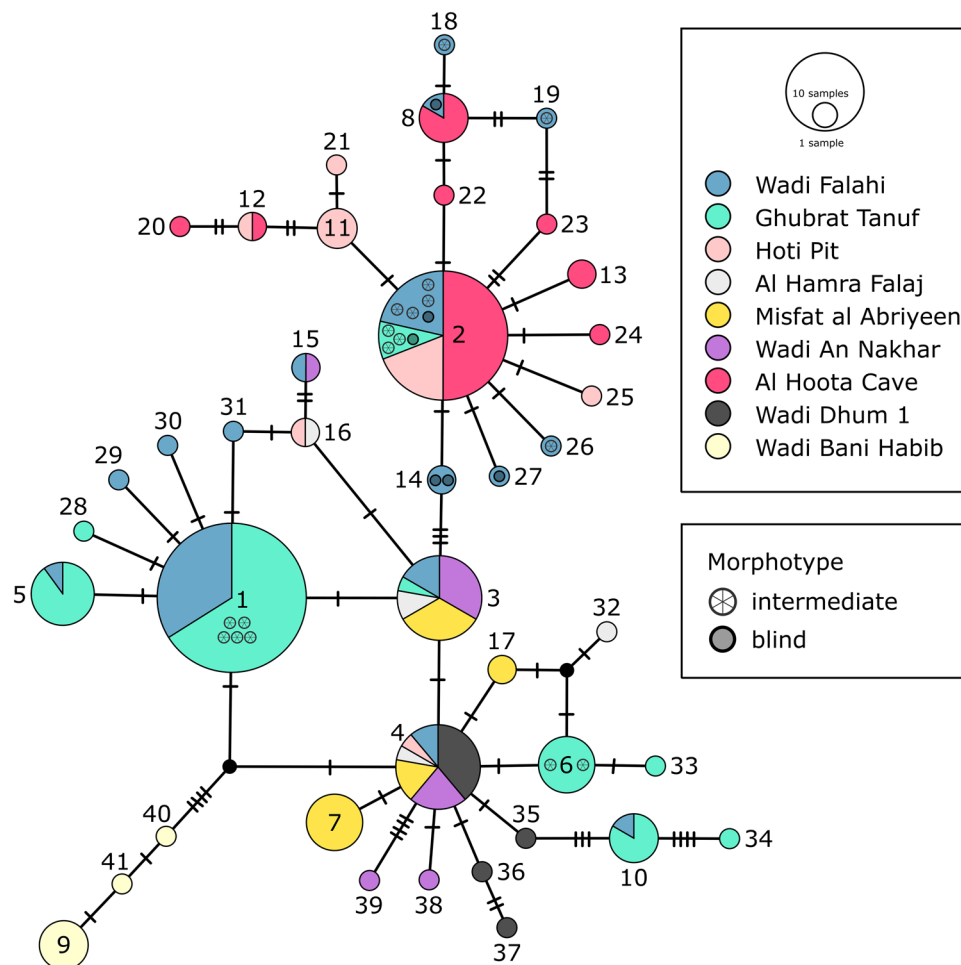


Figure 4. Median-Joining network based on CR sequences. Each circle represents a haplotype and the sizes of circles correspond to the number of individuals sharing this haplotype. Colours indicate sampling site. Every crossbeam on the connecting lines between haplotypes represents a substitution. Small black dots between haplotypes symbolize hypothetical haplotypes not present in the data set. Blind individuals from either Ghubrat Tanuf or Wadi Falahi are indicated by shaded circles, intermediate individuals are indicated by circles filled with an asterisk.

Fst	GT	AH	HP	MF
AH	0.172***			
HP	0.221***	0.103***		
MF	0.087***	0.223***	0.296***	
WF	0.020***	0.109***	0.159***	0.077***

Table 2. Fst-values between all pairs of populations calculated with GenAlEx. Fst fixation indices of the five populations investigated: GT Ghubrat Tanuf, AH Al Hoota Cave, HP Hoti Pit, MF Misfat, WF Wadi Falahi. P-values indicated by asterisk: *** ≤ 0.01 .

with 50% to 90% share of cave genotypes, and the remaining five individuals were juveniles with unusually small eyes and weak pigmentation and their genotypes exhibit 80% to 90% cave affiliation (Table 3, Supplementary Figure S2).

In Wadi Falahi 14 conspicuous individuals (among 36 normal surface individuals) were collected, five of which were completely troglomorphic without visible eyes and weak or no pigmentation, three were categorized as intermediate due to the presence of microphthalmic eyes and weak pigmentation and six individuals were still juvenile, but exhibited either microphthalmic eyes or no eyes at all and were only very weakly pigmented (Supplementary Figure S2). However, all of these 14 individuals possessed a 100% cave genotype (Table 4).

In order to explore the origin of these intermediate individuals, we ran a *structure* analysis comprising only cave individuals from Al Hoota and Hoti Pit and all individuals with a blind or intermediate phenotype (including also juveniles with conspicuous phenotype) from Ghubrat Tanuf and Wadi Falahi (Fig. 6). Again, the highest

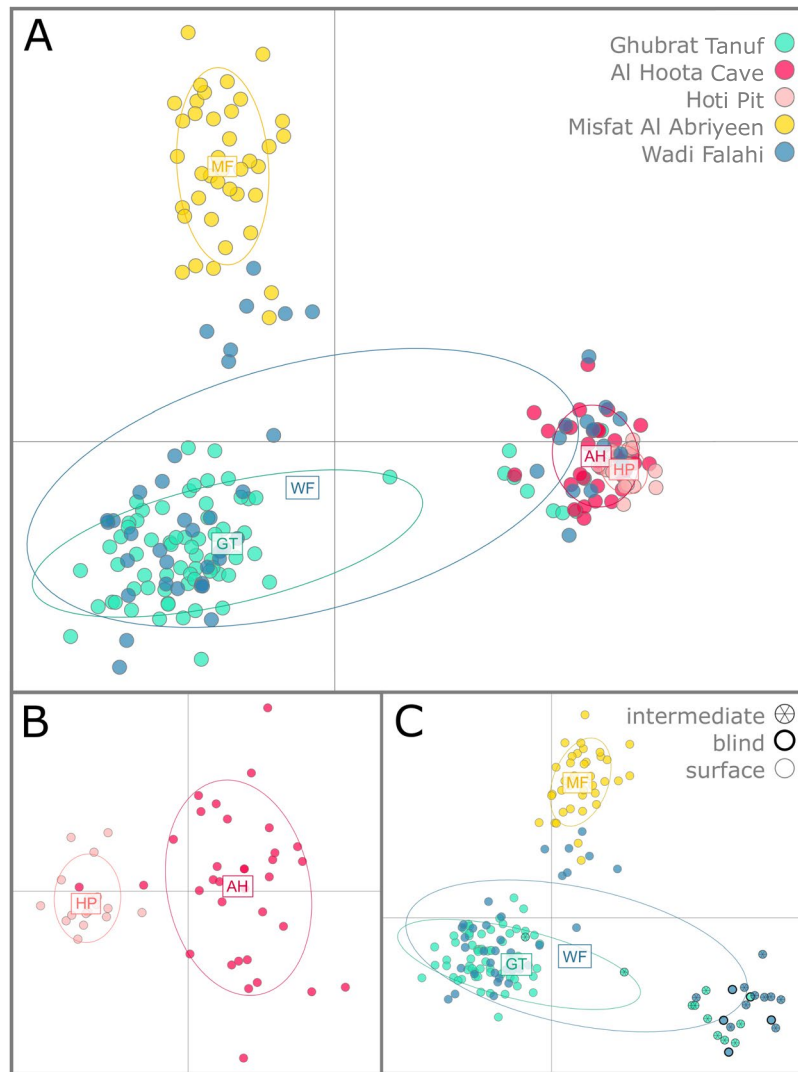


Figure 5. Results of PCA displaying the first two components of the analysis. A) PCA calculated for the whole data set, B) PCA calculated for the cave subset (Al Hoota and Hoti Pit), C) PCA calculated for the surface subset (Ghubrat Tanuf, Misfat Al Abriyeen and Wadi Falahi). The ellipses represent 95% of the inertia (variance) of this group. In C, specimens from surface sites exhibiting an intermediate phenotype are marked with an asterisk, blind individuals are marked with a black outline.

probability lies with $K = 2$ (ΔK) displaying that all specimen from Ghubrat Tanuf belong genetically to Hoti Pit (between 100% and 75% affiliation) except for one individual. This particular individual (Gbar221) is the only one from Ghubrat Tanuf exhibiting a true blind phenotype and belongs genetically 100% to the Al Hoota cluster. On the other hand, all individuals from Wadi Falahi belong genetically to the Al Hoota cluster (between 100% and 85% affiliation), whereas one specimen has a rather low (only a 60%) affiliation to the Al Hoota cluster (Gbar538). These findings are also supported by the results obtained from GeneClass2 analysis (Supplementary Table S5), where all intermediate individuals from Wadi Falahi were assigned to the Al Hoota population (one with high significance), and most intermediate individuals from Ghubrat Tanuf were assigned to the Hoti Pit population with the exception of Gbar221, which was assigned also to Al Hoota (highly significant).

Concerning variation in eye sizes, the results of the comparative eye diameter measurements revealed a clearly smaller eye size in intermediate specimens compared to surface individuals with fully developed eyes (see Table 5). This becomes obvious in Fig. 7, where the smallest eyes were observed in intermediate individuals from Wadi Falahi (mean size 9% of head length) followed by slightly bigger eyes in intermediate individuals from Ghubrat Tanuf (mean size 14% of head length) and the biggest eyes in surface form (mean size 22% of head length).

Discussion

Concerning the genetic diversity within populations, the fact that several parameters like the number of alleles, frequency of polymorphic loci, genetic variability as well as the mean heterozygosity are all lower in the investigated cave populations than in surface populations (which is similar to findings in cave and surface populations of

Nr.	Individual ID	Phenotype	Eye development	Eye diameter (%)	Pigmentation	Genotype affiliation (%) cave/surface	Cave population affiliation (%) Al Hoota/Hoti Pit
1	Gbar221	blind	no eyes		none	100/0	100/0
2	Gbar390	intermed	microphthalmic	17.939	very pale	~ 50/50	~ 25/75
3	Gbar346	juv	microphthalmic	7.506	very pale	~ 80/20	0/100
4	Gbar355	intermed	microphthalmic	19.320	none	0/100	~ 20/80
5	Gbar356	intermed	microphthalmic	16.595	none	~ 90/10	0/100
6	Gbar357	intermed	microphthalmic	12.230	none	~ 90/10	0/100
7	Gbar358	intermed	microphthalmic	6.667	none	~ 80/20	0/100
8	Gbar359	juv	microphthalmic	12.879	very pale	~ 80/20	0/100
9	Gbar360	juv	microphthalmic	15.686	very pale	~ 80/20	0/100
10	Gbar361	juv	microphthalmic	14.779	very pale	~ 90/10	0/100
11	Gbar362	juv	microphthalmic	11.187	very pale	~ 90/10	~ 10/90

Table 3. List of individuals with intermediate phenotype/genotype collected at Ghubrat Tanuf. Phenotype: blind = cave phenotype, intermed = intermediate phenotype, juv = juvenile individual. Eye development: no eyes = no eyes or eye structures externally visible, microphthalmic = small (underdeveloped) eyes visible. Eye diameter (in % of head length): if eyes were present, the eye diameter was measured relatively to head length (from snout tip to end of operculum of the specimen). Pigmentation: none = no sign of pigmentation visible, very pale = less pigmentation than normal surface phenotype. Genotype affiliation: assignment to the cave/surface cluster of each individual (%) according to the structure analysis (see Fig. 1). Cave population affiliation: assignment (%) to one of the cave populations (Al Hoota Cave or Hoti Pit) of each individual according to the structure analysis (see Fig. 6).

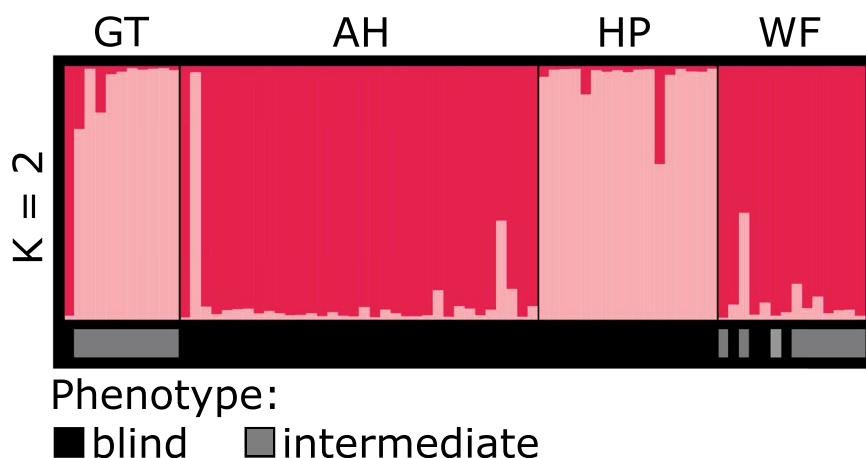


Figure 6. Population structure analysis based on the concatenated data set of cave individuals (AH and HP) and all blind or intermediate individuals collected from Ghubrat Tanuf or Wadi Falahi. The figure ($K=2$) represents the two most probable clusters (magenta = Al Hoota and blind/intermediate individuals mostly from Wadi Falahi, light pink = Hoti Pit and blind/intermediate individuals mostly from Ghubrat Tanuf) resulting from the structure analysis. The corresponding phenotypes of the individuals are indicated by the bar below: black = blind, grey = intermediate.

*Astyanax*⁹) supports the hypothesis that the cave populations probably went through repeated genetic bottlenecks as suggested by Kruckenhauser *et al.*²¹. Moreover, it implies that the cave populations originated from a group of surface ancestors holding only a fraction of the genetic diversity of the whole species. Given the low genetic distance between cave and surface populations as reported by Kruckenhauser *et al.*²¹ who further suggested that the cave population of *G. barreimiae* is phylogenetically young, it seems plausible that this population originated at a time, when the climate started to change from wet to arid resulting in dropping water levels^{23, 47, 48}. In southern Arabia the climate changed repeatedly in the Late Pleistocene, with two major wet periods, one of them coinciding with the last interglacial period (125 to 120 ka) and another during the early Holocene (9.7 to 6.2 ka)⁴⁸. One of these changes probably disrupted the connection between surface and cave water bodies, leaving a subpopulation trapped inside the cave and isolated from surface water bodies, which also became more fragmented. However, it cannot be excluded that the cave population emerged from parallel speciation as found in *Astyanax*⁹. Concerning the surface populations, H_e is higher than H_o in Ghubrat Tanuf and Wadi Falahi, whereas they are nearly similar in the other localities. Since the surface individuals from Ghubrat Tanuf, Wadi Falahi and Misfat all belong to the same cluster in the *structure* analyses (Fig. 1), it seems likely that gene flow between these sites occurs when waterways are easily traversable in season of high precipitations. This is also supported by the rather low fixation

Nr.	Individual ID	Phenotype	Eye development	Eye diameter (%)	Pigmentation	Genotype affiliation (%) cave/surface	Cave population affiliation (%) Al Hoota/Hoti Pit
1	Gbar18	intermed	microphthalmic	9.401	very pale	100/0	100/0
2	Gbar491	blind	no eyes		very pale	100/0	100/0
3	Gbar538	intermed	microphthalmic	6.053	none	100/0	~ 60/40
4	Gbar539	blind	no eyes		none	100/0	100/0
5	Gbar540	blind	no eyes		very pale	100/0	100/0
6	Gbar541	intermed	microphthalmic	4.800	very pale	100/0	100/0
7	Gbar542	blind	no eyes		none	100/0	100/0
8	Gbar543	intermed	microphthalmic	8.491	none	100/0	~ 85/15
9	Gbar551	juv	microphthalmic	6.040	very pale	100/0	100/0
10	Gbar552	juv	microphthalmic	11.876	very pale	100/0	~ 90/10
11	Gbar553	juv	no eyes		very pale	100/0	100/0
12	Gbar554	juv	no eyes		very pale	100/0	100/0
13	Gbar555	juv	microphthalmic	10.824	very pale	100/0	100/0
14	Gbar556	juv	no eyes		very pale	100/0	100/0

Table 4. List of individuals with intermediate phenotype/genotype collected at Wadi Falahi. Phenotype: blind = cave phenotype, intermed = intermediate phenotype, juv = juvenile individual. Eye development: no eyes = no eyes or eye structures externally visible, microphthalmic = small (underdeveloped) eyes visible. Eye diameter (in % of head length): if eyes were present, the eye diameter was measured relatively to head length (from snout tip to end of operculum of the specimen). Pigmentation: none = no sign of pigmentation visible, very pale = less pigmentation than normal surface phenotype. Genotype affiliation: assignment to the cave/surface cluster of each individual (%) according to the structure analysis (see Fig. 1). Cave population affiliation: percentage of genotype belonging to one of the cave populations (Al Hoota Cave or Hoti Pit) of each individual according to the structure analysis (see Fig. 6).

Nr.	Individual ID	Locality	Phenotype	Eye development	Eye diameter (%)	Pigmentation	Genotype affiliation (%) cave/surface
1	Gbar31	GT	surface	fully developed	20.46	present	0/100
2	Gbar67	GT	surface	fully developed	25.501	present	0/100
3	Gbar88	GT	surface	fully developed	24.386	present	0/100
4	Gbar280	MF	surface	fully developed	22.172	present	0/100
5	Gbar282	MF	surface	fully developed	27.199	present	0/100
6	Gbar286	MF	surface	fully developed	25.810	present	0/100
7	Gbar211	WF	surface	fully developed	20.494	present	0/100
8	Gbar212	WF	surface	fully developed	21.438	present	0/100
9	Gbar213	WF	surface	fully developed	20.582	present	0/100

Table 5. Eye diameter of selected surface individuals collected at Ghubrat Tanuf, Misfat Al Abriyeen and Wadi Falahi. All specimens bear a typical surface phenotype with fully developed eyes and pigmentation. Locality: GT = Ghubrat Tanuf, MF = Misfat Al Abriyeen, WF = Wadi Falahi. Eye diameter (in % of head length): eye diameter was measured relatively to head length (from snout tip to end of operculum of the specimen). Genotype affiliation: assignment to the cave/surface cluster of each individual (%) according to the structure analysis (see Fig. 1).

indices between the surface sites (0.021 to 0.083) suggesting increased gene flow compared to the fixation indices between both cave sites (0.093) and between surface and cave sites (0.101 to 0.244).

In the course of investigating potential gene flow and/or genetic differentiation between cave and surface populations, the *structure* analysis clearly shows that the troglomorphic individuals from Al Hoota Cave and Hoti Pit are differentiated from surface populations (except individuals with intermediate phenotype). Since not a single individual from one of the cave sites bears any degree of surface genotype, we conclude that there is no gene flow from the surface populations into the caves. The *structure* analysis of only the troglomorphic cave individuals revealed a differentiation between the individuals of Al Hoota Cave and Hoti Pit (Fig. 2), which became also apparent in the PCA (Fig. 5B). A possible explanation for the two individuals from Al Hoota cave bearing partial Hoti Pit genotype might be a temporal subterranean connection allowing unidirectional migration from Hoti Pit (located at higher altitude) to Al Hoota Cave during flood events. Apart from these two individuals, no sign for ongoing gene flow between both cave populations was detected, indicating that they are isolated from each other. As the individuals from Hoti Pit have a lower number of alleles which are mostly a subset of the Al Hoota alleles and have very few private alleles one can conclude that the Hoti Pit population originated from individuals of the Al Hoota Cave.



Figure 7. Boxplot based on eye sizes of intermediate individuals collected from Ghubrat Tanuf and Wadi Falahi compared to surface individuals with external fully developed eyes. The figure displays the difference in eye diameter between individuals with intermediate and surface phenotype. Y-axis represents the percentage of eye diameter measured relatively to head length. X-axis shows the three groups that were compared: intermediate individuals from Ghubrat Tanuf (N = 11, turquoise), intermediate individuals from Wadi Falahi (N = 14, blue), and surface individuals (N = 9, three specimens from each surface site - Ghubrat Tanuf, Wadi Falahi and Misfat Al Abriyeen, yellow). Each box is confined by its first and third quartiles (contains 50% of data), the thick notch within the box represents the median (95% confidence interval) and the whiskers extend to the largest/smallest values of each population.

The differentiation of the Misfat population in the *structure* analysis with the second highest probability ($K = 3$, Fig. 1) from the other two surface populations is in accordance with the results of Kruckenhauser *et al.*²¹ who investigated genetic distances (mt data) of individuals from several surface populations and found no geographic pattern within this group. Yet, this differentiation is only weak suggesting occasional gene flow. Although there is no direct connection between these sites during dry periods, there might be a temporal water junction (probably underground) enabling single individuals from Misfat to migrate into the Wadi Falahi when the water level rises during flood events.

Regarding individuals with intermediate phenotypes, the origin of these specimens was of particular interest. Those collected from Ghubrat Tanuf mostly have a mixed cave/surface genotype indicating that there must be at least some degree of gene flow from cave populations. These findings are also corroborated by the results of the CR network which clearly shows that some of these individuals share the same haplotype with typical surface individuals (HT 1 and 6, see Fig. 4). The results of the *structure* analyses (Fig. 6) show that the intermediate individuals found in Ghubrat Tanuf have genetic similarity to the Hoti Pit population. Since the water body in the Hoti Pit cave is secluded and no visible surface connection leading to Ghubrat Tanuf is known, we assume that either single individuals of Hoti Pit somehow managed to migrate downstream through the Al Hoota Cave, the wadi and Ain Al Msalla to Ghubrat Tanuf during flood events or there exists at least a temporary underground connection between these sites. This is supported by the finding of a single blind individual with typical cave phenotype and genotype in Ghubrat Tanuf. It seems reasonable that offspring of mixed mating couples can survive under cave conditions in Ghubrat Tanuf while in surface waters chances of survival and reproduction are low.

However, the situation with specimens bearing an intermediate phenotype from Wadi Falahi is different, because they all have a “pure” cave genotype (see *structure* analyses in Fig. 1) of Al Hoota Cave. In fact, all of these individuals also belong to the “cave” haplogroup in the CR network (Fig. 4). Most probably, these specimens are not the product of hybridisation, but cave individuals that managed to survive after they were flushed out of the lower entry of Al Hoota Cave. The average amount of rainfall in Oman is low and normally, rainwater drain away into the ground fast. However, sporadic rainfalls in this region can be heavy and lead to flooding of wadis. Flooding of the Wadi Hoota channels the water directly into the upper entry of the Al Hoota Cave and causes a flood wave that flushes through the cave with a rate of 100 m³ per second²⁰ connecting the cave with the surface water bodies. Anyhow, since not a single individual exhibiting any degree of surface genotype was found within the cave, we conclude that the strong flow rate during these flooding events prevent surface individuals from migrating into the cave. The intermediate phenotype of these “pure” cave individuals could be explained by phenotypic plasticity. One could assume that these individuals left the cave at an early developmental stage, when eyes were still visible. Subsequently, the increased exposure to sunlight leads to darker pigmentation and less reduced eyes than in cave specimens which developed in complete dark habitat. Banister¹⁹ observed similar effects of light

exposure in *G. barreimiae*. He described development of melanin in the skin and particular regrowth of optic lobes in cave individuals kept under light exposure within short periods of time (after four months). Phenotypic plasticity seems a reasonable explanation considering that the troglomorphic phenotype is of rather recent origin²¹. Similar observations were conducted with individuals of the Micos cave population of *Astyanax*, which is considered to be phylogenetically young. Therefore, individuals of this population bear rudimentary eyes that are less reduced compared to phylogenetically old troglodyte populations (e.g. El Abra region) whereas pigmentation is almost as pronounced as in the surface form but increases even more after exposure to light^{49,50}. Effects of light exposure can also be observed in the Atlantic molly *Poecilia mexicana*, a species occurring in the Mexican Cueva del Azufre – a system comprising both surface as well as cave habitats. Interestingly, individuals of this species show a gradual decrease of eye-size from the entrance and the front chambers, which receive dim light through numerous skylights, to the innermost chamber of the cave. Interestingly, the differences in eye sizes seem to be related to different amounts of light received in the chambers⁵¹. Subsequently, to clarify and quantify the phenotypic variability as a consequence of light exposure during ontogeny in individuals of *G. barreimiae*, comparative breeding and rearing experiments under controlled conditions are inevitable.

The fact that we found not a single hybrid individual at Wadi Falahi leads to the conclusion that hybridisation between cave individuals and surface individuals from Wadi Falahi is either not occurring at all or none of the F1 generation survives and reaches reproductive stage in the surface environment. Probably, hybrid larvae, which are pale and blind or with debility of sight, are at a disadvantage and more prone to predation compared to surface larvae. In addition, the cannibalistic nature of *G. barreimiae* as mentioned in Timmermann and Plath⁵² might also decrease the chance of survival. In contrast, in Ghubrat Tanuf the two morphotypes hybridize, which indicates that the individuals with a surface phenotype are capable of breeding in the cave environment. This is in accordance with our observation in the Wadi Bani Khalid cave (Oman), in which typical surface forms of *G. barreimiae* occur frequently and seem to survive well in the cave habitat.

Conclusion

Although our results showed significant genetic differentiation between the surface and cave form of *G. barreimiae*, we found that gene flow between the two morphotypes is still possible in cave environment but occurs only unidirectional from the cave to surface populations. Hence the integrity of the Al Hoota and Hoti Pit population with their peculiar phenotypes is currently not threatened by admixture with surface morphotypes. However, the restricted local occurrence in only two caves makes them vulnerable to extinction due to environmental impacts. Furthermore, there are indications for phenotypic plasticity, probably due to light exposure, in individuals collected from surface populations with an allele composition typical for cave specimens but exhibiting an intermediate phenotype. The finding of hybrid specimens and cave individuals outside the caves indicate that at least a temporary connection, probably underground, between Hoti Pit cave and surface water bodies must exist. These inferences are supported by results deduced from a combination of mitochondrial, nuclear and morphological data. Furthermore, we found a substructure between the investigated surface populations leading to the conclusion that gene flow between surface water bodies occurs sporadically.

References

- Romero, A. & Green, S. M. The end of regressive evolution: examining and interpreting the evidence from cave fishes. *J. Fish Biol.* **67**, 3–32 (2005).
- Jeffery, W. R. Regressive evolution in *Astyanax* cavefish. *Annu. Rev. Genet.* **43**, 25–47 (2009).
- Protas, M., Conrad, M., Gross, J. B., Tabin, C. & Borowsky, R. Regressive evolution in the Mexican cave tetra, *Astyanax mexicanus*. *Curr. Biol.* **17**, 452–4 (2007).
- Gross, J. B., Meyer, B. & Perkins, M. The rise of *Astyanax* cavefish. *Dev. Dyn.* **244**, 1031–1038 (2015).
- Jeffery, W. R. Cavefish as a Model System in Evolutionary Developmental Biology. *Dev. Biol.* **231**, 1–12 (2001).
- Strecker, U., Bernatchez, L. & Wilkens, H. Genetic divergence between cave and surface populations of *Astyanax* in Mexico (Characidae, Teleostei). *Mol. Ecol.* **12**, 699–710 (2003).
- Wilkens, H. Genetics and hybridization in surface and cave *Astyanax* (Teleostei): A comparison of regressive and constructive traits. *Biol. J. Linn. Soc.* 911–928, doi:10.1111/bj.12773 (2016).
- Bradic, M., Beerli, P., García-de León, F. J., Esquivel-Bobadilla, S. & Borowsky, R. L. Gene flow and population structure in the Mexican blind cavefish complex (*Astyanax mexicanus*). *BMC Evol. Biol.* **12**, 9 (2012).
- Strecker, U., Hausdorf, B. & Wilkens, H. Parallel speciation in *Astyanax* cave fish (Teleostei) in Northern Mexico. *Mol. Phylogenet. Evol.* **62**, 62–70 (2012).
- Moran, D., Softley, R. & Warrant, E. J. The energetic cost of vision and the evolution of eyeless Mexican cavefish. *Sci. Adv.* **1**, 1–9 (2015).
- Tang, Q., Getahun, A. & Liu, H. Multiple in-to-Africa dispersals of labeonin fishes (Teleostei: Cyprinidae) revealed by molecular phylogenetic analysis. *Hydrobiologia* **632**, 261–271 (2009).
- Menon, A. G. K. In *Memoirs of the Indian Museum* **14**, 173–260 (1964).
- Stiassny, M. L. J. & Getahun, A. An overview of labeonin relationships and the phylogenetic placement of the Afro-Asian genus *Garra* Hamilton, 1922 (Teleostei: Cyprinidae), with the description of five new species of *Garra* from Ethiopia, and a key to all African species. 41–83 (2007).
- Farashi, A. et al. Reassessment of the taxonomic position of *Iranocypris typhlops* Bruun & Kaiser, 1944 (Actinopterygii, Cyprinidae). *Zookeys* **77**, 69–77 (2014).
- Sayyadzadeh, G., Esmaili, H. R. & Freyhof, J. *Garra mondica*, a new species from the Mond River drainage with remarks on the genus *Garra* from the Persian Gulf basin in Iran (Teleostei: Cyprinidae). *Zootaxa* **4048**, 75–89 (2015).
- Colli, L., Paglianti, A., Berti, R., Gandolfi, G. & Tagliavini, J. Molecular phylogeny of the blind cavefish *Phreatichthys andruzzii* and *Garra barreimiae* within the family Cyprinidae. *Environ. Biol. Fishes* **84**, 95–107 (2008).
- Calderoni, L. et al. Relaxed selective constraints drove functional modifications in peripheral photoreception of the cavefish *P. andruzzii* and provide insight into the time of cave colonization. *Heredity (Edinb)* In press (2016).
- Ercolini, A. & Berti, R. Light sensitivity experiments and morphology studies of the blind phreatic fish *Phreatichthys andruzzii* Vinciguerra from Somalia. *Monit. Zool. Ital* **2**(Suppl.), 29–43 (1975).
- Banister, K. E. A subterranean population of *Garra barreimiae* (Teleostei: Cyprinidae) from Oman, with comments on the concept of regressive evolution. *J. Nat. Hist* **18**, 927–938 (1984).

20. Waltham, A. C., Brown, R. D. & Middleton, T. C. Karst and Caves in the Jabal Akhdar, Oman. *Cave Sci* **12**, 69–79 (1985).
21. Kruckenhauser, L., Haring, E., Seemann, R. & Sattmann, H. Genetic differentiation between cave and surface-dwelling populations of Garra barreimiae (Cyprinidae) in Oman. *BMC Evol. Biol.* **11**, 172 (2011).
22. Kirchner, S., Weinmaier, T., Rattei, T., Sattmann, H. & Kruckenhauser, L. Characterization of 19 new microsatellite loci for the Omani barb Garra barreimiae from 454 sequences. *BMC Res. Notes* **7**, 522 (2014).
23. Kollmann, H. A., Sattmann, H. & Seemann, R. „Al Hoota“, ein karst- und höhlenkundliches Forschungs- und Tourismusprojekt im Sultanat Oman. *Abhandlungen der Geol. Bundesanstalt* **60**, 107–113 (2007).
24. Hall, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**, 95–98 (1999).
25. Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729 (2013).
26. Bandelt, H. J., Forster, P. & Röhl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **16**, 37–48 (1999).
27. R Development Core Team. R: A Language and Environment for Statistical Computing. at <http://www.r-project.org> (2015).
28. Kamvar, Z. N., Tabima, J. F. & Grünwald, N. J. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* **2**, e281 (2014).
29. Jombart, T. Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403–1405 (2008).
30. Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W. & Prodöhl, P. A. diveRsity: An R package for the estimation of population genetics parameters and their associated errors. *Methods Ecol. Evol.* **4**, 782–788 (2013).
31. Roussel, F. Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Mol. Ecol. Resour.* **8**, 103–106 (2008).
32. Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M. & Shipley, P. Micro-Checker: Software for Identifying and Correcting Genotyping Errors in Microsatellite Data. *Mol. Ecol. Notes* **4**, 535–538 (2004).
33. Paradis, E. pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* **26**, 419–420 (2010).
34. Weir, B. S. & Cockerham, C. C. Estimating F-Statistics for the Analysis of Population Structure. *Evolution (NY)* **38**, 1358–1370 (1984).
35. Peakall, R. & Smouse, P. E. GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Notes* **6**, 288–295 (2006).
36. Peakall, R. & Smouse, P. E. GenALEX 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* **28**, 2537–9 (2012).
37. Goudet, J. Fstat Version 1.2.: a computer program to calculate F-statistics. *J. Hered.* **86**, 485–486 (1995).
38. Pritchard, J. K. Documentation for structure software: Version 2. 3. 1–38 (2010).
39. Earl, D. A. & VonHoldt, B. M. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**, 359–361 (2012).
40. Jakobsson, M. & Rosenberg, N. A. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801–1806 (2007).
41. Rosenberg, N. A. DISTRUCT: A program for the graphical display of population structure. *Mol. Ecol. Notes* **4**, 137–138 (2004).
42. Dray, S. & Dufour, A. B. The ade4 package: implementing the duality diagram for ecologists. *J. Stat. Softw.* **22**, 1–20 (2007).
43. Piry, S. *et al.* GENECLASS2: A software for genetic assignment and first-generation migrant detection. *J. Hered.* **95**, 536–539 (2004).
44. Rannala, B. & Mountain, J. L. Detecting immigration by using multilocus genotypes. *Proc. Natl. Acad. Sci. USA* **94**, 9197–9201 (1997).
45. Paetkau, D., Slade, R., Burden, M. & Estoup, A. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Mol. Ecol.* **13**, 55–65 (2004).
46. Frankham, R., Briscoe, D. A. & Ballou, J. D. *Introduction to conservation genetics* (Cambridge University Press, 2010).
47. Clark, I. D. & Fontes, J.-C. Paleoclimatic reconstruction in northern Oman based on carbonates from hyperalkaline groundwaters. *Quat. Res.* **33**, 320–336 (1990).
48. Burns, S. J., Matter, A., Frank, N. & Mangini, A. Speleothem-based palaeoclimate record from northern Oman. *Geology* **26**, 499–502 (1998).
49. Wilkens, H. & Burns, R. J. A new Anoptichthys cave populations (Caracidae, Pisces). *Ann. Spéleol.* **27**, 263–270 (1972).
50. Peters, N. & Scholl, A. & Wilkens, H. Der Micos-Fisch, Höhlenfisch in statu nascendi oder Bastard? Ein Beitrag zur Evolution der Höhlentiere. *J. Zool. Syst. Evol. Res.* **13**, 110–124 (1975).
51. Fontanier, M. E. & Tobler, M. A morphological gradient revisited: Cave mollies vary not only in eye size. *Environ. Biol. Fishes* **86**, 285–292 (2009).
52. Timmermann, M. & Plath, M. Phototactic response and light sensitivity in an epigeal and a hypogean population of a barb (Garra barreimiae, Cyprinidae). *Aquat. Ecol.* **43**, 539–547 (2008).

Acknowledgements

The authors want to thank Johannes Rath and Christian Baranyi from the University of Vienna for providing help with the Sequence Analyzer, Harald Krenn (University of Vienna) for his general help and support, Barbara Däubel and Julia Schindelar (NHM Vienna) for technical assistance in the laboratory, Robert Illek (NHM Vienna) for support in practical field work, Frank Zachos (NHM Vienna) and Stephan Koblmüller (University of Graz) for help and support with population genetic analyses and Toni Weissenbacher (Zoo Vienna) for providing additional samples. Special thanks to Lukas Raneburger for helpful comments and support. We want to express our gratitude to the Ministry of Environment and Climate Affairs of Oman for their cooperation, support and the sampling permits and to Michael Robinson, a zoologist at the Sultan Qaboos University in Muscat, for his intellectual input. The study was funded by the University of Vienna (uni:docs fellowship), the Natural History Museum of Vienna (Central Research laboratories), by the Ministry of Commerce and Industry of the Sultanate of Oman and by the Ministry of Tourism of the Sultanate of Oman. Open access funding provided by University of Vienna. The Federal Ministry for Transport, Innovation and Technology Austria provided financial support in the course of a FEMtech practical.

Author Contributions

S.K. and L.K. planned the research, S.K. carried out the molecular genetic and statistical data analyses and drafted the manuscript. L.K. participated in the molecular genetic analyses and helped to draft the manuscript. L.P. contributed to the manuscript with his speleological and geographical expertise and helped with collecting samples from the cave sites. H.S. initiated the study, coordinated the sampling and helped to draft the manuscript. E.H. helped with the data collection and to draft the manuscript. R.V. helped with the sampling and provided detailed information about geographic conditions. All authors read and approved the final manuscript.

Additional Information

Supplementary information accompanies this paper at doi:[10.1038/s41598-017-05194-3](https://doi.org/10.1038/s41598-017-05194-3)

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017

Chapter II: Phylogeography and cryptic diversity of *G. barreimiae*

Paper 3: The morphological diversity of *Garra barreimiae* [Teleostei: Cyprinidae]

The morphological diversity of *Garra barreimiae* [Teleostei: Cyprinidae]

Arthur Pichler  · Harald Ahnelt · Sandra Kirchner · Helmut Sattmann · Elisabeth Haring · Stephan Handschuh · Jörg Freyhof · Reginald Victor · Luise Kruckenhauser

Received: 15 September 2017 / Accepted: 2 April 2018
© The Author(s) 2018

Abstract *Garra barreimiae* (Fowler and Steinitz, Bull Res Counc Isr 5B:262–289, 1956) is a freshwater fish that lives throughout the Hajar Mountains of Oman and the United Arab Emirates. Previously, four different genetic clades (West clade, Central Clade, North clade and East clade) have been identified within this species. This study observes morphological differences between these clades using morphometrics and meristics as well as micro CT imaging, further strengthening the assumption that the taxon of *G. barreimiae* should be restricted to one of the genetic clades detected. Although many morphometric and meristic characteristics are highly variable within the clades, the West clade fits the original description of *G. barreimiae* and is distinguishable by its higher number of fifteen to seventeen gill rakers on the lower limb of the first gill arch (compared to the other clades with ten to fourteen gill rakers), its seven

branched pelvic fin rays (in contrary to the other clades with eight branched pelvic fin rays) and its number of vertebrae, with typically nineteen precaudal and one intermediate vertebrae, compared to eighteen precaudal and two intermediate vertebrae in the Central and North clade and seventeen precaudal and two intermediate vertebrae in the East clade.

Keywords *Garra barreimiae* · Labeonini · Micro CT · Morphometrics · Oman · Hajar Mountains

Introduction

Labeonine cyprinids of the genus *Garra* (Hamilton, 1822) are widespread in tropical and subtropical Asia, Africa and the Middle East (Menon 1964; Yang and

A. Pichler (✉) · S. Kirchner · E. Haring · L. Kruckenhauser
Central Research Laboratories, Laboratory of Molecular Systematics, Museum of Natural History Vienna, Burgring 7, 1010 Vienna, Austria
e-mail: arthur.pichler@univie.ac.at

A. Pichler · S. Kirchner · E. Haring
Department of Integrative Zoology, University of Vienna, Althanstraße 14, 1090 Vienna, Austria

H. Ahnelt
Department of Theoretical Biology, University of Vienna, Althanstraße 14, 1090 Vienna, Austria

H. Ahnelt
1st Zoological Department, Museum of Natural History Vienna, Burgring 7, 1010 Vienna, Austria

H. Sattmann
3rd Zoological Department, Museum of Natural History Vienna, Burgring 7, 1010 Vienna, Austria

S. Handschuh
VetImaging, VetCore Facility for Research, University of Veterinary Medicine Vienna, Veterinärplatz 1, 1210 Vienna, Austria

J. Freyhof
Leibniz-Institute of Freshwater Ecology and Inland Fisheries (IGB), Müggelseedamm 310, 12587 Berlin, Germany

R. Victor
Department of Biology, Sultan Qaboos University, Al Khoudh, 123 Muscat, Oman

Mayden 2010). Following species of *Garra* occur on the Arabian Peninsula: *Garra buettikeri* (Krupp 1983), *Garra dunsirei* (Banister 1987), *Garra lautior* (Banister 1987), *Garra longipinnis* (Banister and Clarke 1977), *Garra mamshuqa* (Krupp 1983), *Garra sahilia* (Krupp 1983), *Garra smarti* (Krupp and Budd 2009), *Garra tibania* (Trewavas, 1941) and *Garra barreimiae* (Fowler and Steinitz 1956) (Freyhof et al. 2015). Only recently *Garra barreimiae gallagheri* was raised to species level and *Garra sindhi* has been described from the Dhofar region in Oman (Lyon et al. 2016). New *Garra* species have also been described from areas adjacent to the Arabian Peninsula, including *Garra jordanica* (Hamidan et al. 2014), *Garra mondica* (Sayyadzadeh et al. 2015), *Garra lorestanensis* (Mousavi-Sabet and Eagderi 2016), *Garra tashanensis* (Mousavi-Sabet et al. 2016) and *Garra amirhosseini* (Esmaeili et al. 2016), demonstrating, that more undescribed species might be found in the area. This is especially true for the Arabian Peninsula, where most freshwater fish now live in isolated refuges. These fish live in very dry and often unstable habitats only inhabited by a restricted number of fish species which have adapted to massive changes in water level, temperature and therefore oxygen content as well as availability of food (Krupp 1988).

Garra barreimiae is one of these species having a very fragmented distribution area with many small and isolated populations. The species is widespread in Northern Oman and the United Arab Emirates (Banister and Clarke 1977; Krupp 1983; Freyhof et al. 2015). Fowler and Steinitz (1956) described *G. barreimiae* from Buraimi (also Al Buraimi, Baraimi, Barreimi) in Oman. Later, two additional subspecies were described: *Garra barreimiae shawkahensis* (Banister and Clarke 1977) from the United Arab Emirates and *Garra barreimiae gallagheri* (Krupp 1988) from Wadi Bani Khalid in Oman. In 1980 Dunsire and Gallagher reported a subterranean population of *G. barreimiae* from the Al Hoota Cave close to Nizwa (Banister 1984). This troglomorphic population seems to be of recent origin less than one million years ago (Kruckenhauser et al. 2011), however occasional gene flow between the cave population and the surface population seems to take place (Kirchner et al. 2017). Although many populations of *G. barreimiae* are very isolated from each other and inhabit different drainage systems around the Hajar Mountains (Kruckenhauser et al. 2011; Freyhof et al. 2015), in former times during

intermittent pluvial intervals, the river network was much denser, so populations might have been more connected. Kruckenhauser et al. (2011) were the first to use molecular genetic methods to assess genetic differentiation among populations of *G. barreimiae*. On the basis of DNA sequences of the mitochondrial control region (mt CR) two well differentiated clades were found, one in the surroundings of the Al Hoota cave, with a separated subclade which consists of individuals from Wadi Bani Khalid (= *G. gallagheri*), and one in the northeast of the Hajar Mountains. Subsequent genetic analysis (Kirchner et al. 2017) distinguished at least four different clades with p-distances between 6% and 10%, as measured in a section of the mt *cytochrome c oxidase subunit 1* gene (*COI*). These clades are geographically separated but their distribution does not reflect the distribution of the subspecies previously described. Furthermore, including *G. rufa* into the tree revealed that *G. barreimiae* in the sense of Kruckenhauser et al. (2011) might not be monophyletic. With the newly discovered and strikingly different mt clades, which may indicate cryptic species, the morphological characteristics of these clades ought to be examined.

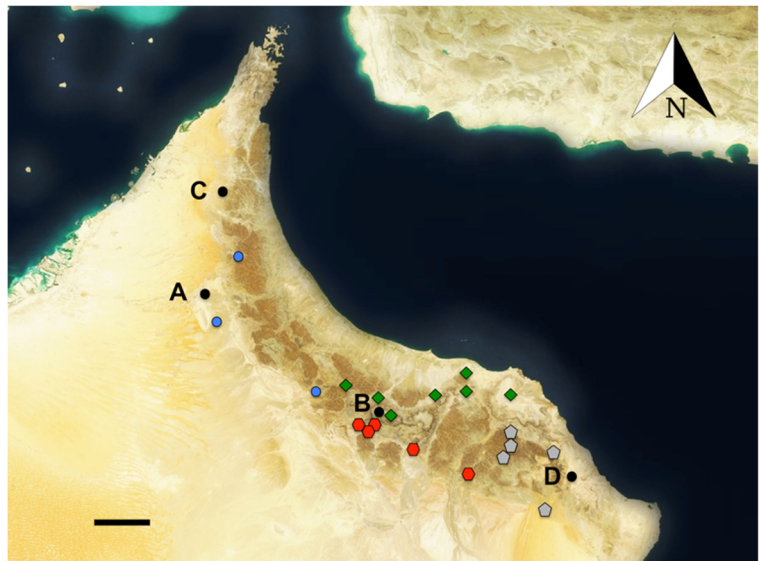
Water extraction in Oman has increased drastically in the last three decades (Al-Rawahi et al. 2014) which poses a threat to isolated fish populations. *G. barreimiae* is widely distributed throughout the Northern Oman, but since many populations are isolated and potentially vulnerable to water extraction (Freyhof et al. 2015) it appears even more urgent to study potentially cryptic species within *G. barreimiae*.

This study is part of a broader phylogenetic analysis of the *G. barreimiae* complex and aims to reveal whether morphological differences can be found between previously determined genetic clades using morphometrics and meristics as well as micro CT imaging, which is used for the first time to compare the osteology within *G. barreimiae*.

Material and methods

Eighty-seven specimens of *G. barreimiae*, collected at 20 localities in Oman (Fig. 1; Table 1), were included in the morphological analysis. Specimens were captured with fishing nets, immediately anesthetized with clove oil and fixed in 80% ethanol. The individuals, which are part of the collection of the Natural History Museum Vienna (NHMV), represent four mitochondrial (mt)

Fig. 1 Sampling sites of specimens used in morphological analysis and sites of original descriptions across the Al Hajar Mountains, Oman. West clade: blue; Central clade: red; North clade: green; East clade: grey; A: Buraimi, type locality of *Garra barreimiae* (Fowler and Steinitz 1956); B: Wadi Sahtan; assigned to *Garra barreimiae barreimiae* by Banister and Clarke (1977); C: Wadi Shawkah, type locality of *Garra barreimiae shawkahensis* (Banister and Clarke 1977); D: Wadi Bani Khalid; type locality of *Garra gallagheri* (Krupp 1988), not included in this study. Scale bar: 50 km



clades (Kruckenhauser et al. 2011; Kirchner et al. 2017): West clade ($n = 19$), Central clade ($n = 25$), North clade ($n = 19$) and East clade ($n = 24$).

Table 1 Sampling sites for each clade with corresponding coordinates as decimal degrees [longitude, latitude] and number of morphologically examined specimens (n)

Clade	Sampling Site	Coordinates	n
West clade	Al Hayyal	[23.43, 56.74]	7
	Al Juwayf	[24.54, 56.10]	7
	Hafeet	[23.99, 55.84]	5
Central clade	Al Hamra	[23.11, 57.28]	1
	Birkat al Mawz	[22.92, 57.66]	2
	Lizq	[22.71, 58.17]	4
	Misfat al Abriyeen	[23.14, 57.31]	10
	Wadi an Nakhar	[23.17, 57.20]	8
North clade	Wadi al Mayh	[23.40, 58.52]	4
	Wadi Hammam	[23.37, 57.82]	4
	Wadi Hat	[23.18, 57.41]	1
	Wadi Khawd	[23.57, 58.11]	5
	Wadi Mansah	[23.40, 58.11]	1
	Wadi Sahtan	[23.38, 57.30]	2
	Wadi Shkar	[23.45, 57.06]	2
	Wadi Dayn	[22.89, 58.42]	2
East clade	Al Hajir	[22.98, 58.50]	7
	Wadi Khabbah 1	[22.91, 58.87]	7
	Wadi Khabbah 2	[23.11, 58.50]	6
	Mintirib	[22.43, 58.79]	2

Twenty-five measurements (Table 2) were taken at least two times with a digital calliper from point to point to the nearest 0.1 mm. Methods for counts and measurements largely followed Armbruster (2012). Due to the condition of preserved specimens (crooked fixation, mutilated fins, cuts in the ventral region) not all measurements and counts could be taken from all specimens. Measurements were described in percent of standard length unless stated otherwise. Measurements were tested for normality (Shapiro Wilk's test) and equality of variance (Levene's Test) before conducting One-way ANOVA. Kruskal-Wallis test was used to test for significant differences in non-parametric data. Tukey's honest significant difference test was used as post-hoc test after ANOVA. Principal component analysis (PCA) and linear discriminant analysis (LDA) were conducted using the following characters, which were chosen because of their parametric nature and significant differences between the mt clades: anal fin - caudal fin, head length, head width (in % of head length), internostril distance (in % of head length), orbit diameter (in % of head length), snout length (in % of head length), pectoral fin - pelvic fin, pelvic fin - anal fin and prepelvic length. All tests and analyses were conducted using Past3 (Hammer et al. 2001). Violin plots were constructed using Rstudio version 1.0.136 and the R package ggplot2 (R Development Core Team 2015; Wickham 2009).

Meristic counts of gill rakers of the lower limb of the first gill arch (ceratobranchial), simple and branched dorsal fin rays, anal fin rays, pectoral fin rays, pelvic fin rays, caudal fin rays as well as lateral line scales and

Table 2 Morphometric data of the four mitochondrial (mt) clades

	West clade				Central clade				North clade				East clade			
	mean	min	max	SD	mean	min	max	SD	mean	min	max	SD	mean	min	max	SD
SL (mm)	40.1	28.3	54.7	8.8	35.5	26.0	50.9	6.9	41.5	27.5	61.8	9.0	35.4	25.4	50.9	6.8
% SL																
Anal fin - Anus	4.8	3.4	6.4	0.9	4.2	3.4	5.0	0.5	5.2	3.8	6.5	0.7	5.6	4.4	6.6	0.6
Anal fin - Caudal fin	21.6	19.0	25.0	1.7	20.2	17.6	22.3	1.1	20.9	18.8	24.4	1.3	20.9	19.0	23.6	1.3
Anal fin height	15.9	14.4	17.9	0.9	16.2	14.3	17.7	0.8	16.7	14.5	19.4	1.2	16.0	13.9	18.7	1.3
Body depth	23.0	20.9	25.7	1.7	24.8	21.7	28.7	1.6	25.2	21.4	29.2	2.1	22.6	19.9	24.9	1.4
Caudal peduncle depth	12.1	11.1	13.1	0.7	12.4	11.4	13.2	0.5	13.1	11.8	13.8	0.6	12.4	11.7	13.0	0.3
Caudal peduncle length	15.1	12.9	17.5	1.4	13.8	11.9	15.6	0.8	14.6	12.3	16.9	1.1	14.1	12.6	16.3	1.0
Dorsal fin height	21.3	19.9	23.7	1.2	20.6	19.0	22.9	1.1	22.2	20.3	28.0	1.6	21.9	20.2	25.6	1.3
Head depth	16.6	14.8	19.0	1.0	17.8	16.0	19.5	0.9	17.4	16.1	18.8	0.9	16.9	15.7	18.4	0.8
Head length	23.9	21.2	27.2	1.4	25.3	22.0	27.1	1.2	24.2	21.9	27.5	1.5	24.8	21.7	27.3	1.3
Head width	18.8	16.6	21.4	1.2	19.5	16.2	21.9	1.4	18.7	16.7	21.0	1.3	18.0	16.4	19.4	0.9
Internostral distance	7.2	5.8	8.4	0.6	7.8	6.9	9.0	0.5	7.7	6.4	8.6	0.5	7.8	6.9	8.5	0.4
Interorbital width	11.0	9.9	11.9	0.7	12.7	11.9	14.3	0.6	11.5	10.6	12.4	0.6	12.1	11.2	13.0	0.4
Pectoral fin height	20.2	18.7	23.7	1.2	20.8	17.9	24.8	1.8	20.2	17.2	22.8	1.4	19.7	17.4	21.9	1.0
Pectoral fin - Pelvic fin	33.0	30.2	35.1	1.3	30.9	27.5	33.8	1.6	31.7	28.6	33.8	1.4	31.3	29.4	34.1	1.4
Postorbital length	9.23	7.8	10.6	0.8	10.2	8.6	12.0	0.8	9.1	8.0	10.5	0.6	9.1	7.8	11.1	0.8
Postdorsal length	35.8	32.0	38.4	1.8	34.8	31.6	37.8	1.4	35.9	33.4	38.5	1.7	35.1	32.2	39.0	1.6
Preal length	78.6	76.0	81.1	1.5	79.3	77.4	81.3	1.0	78.6	75.3	81.2	1.4	79.0	76.0	81.0	1.2
Predorsal length	51.6	49.5	54.5	1.4	51.9	50.5	54.2	1.1	50.9	49.0	53.6	1.2	50.8	47.7	53.3	1.4
Prepelvic length	56.0	52.2	59.2	2.0	55.6	52.9	58.2	1.4	54.6	52.2	58.5	1.4	54.7	51.4	57.5	1.4
Pelvic fin - Anal fin	22.5	20.3	25.1	1.6	23.3	20.5	25.1	1.2	23.6	22.1	26.4	1.2	24.0	19.7	26.8	1.6
Pelvic fin - Caudal fin	44.2	39.9	48.5	1.9	43.9	40.7	46.4	1.3	45.0	39.9	47.5	1.8	45.0	41.3	49.0	1.6
Pelvic fin height	16.6	15.2	18.8	1.0	16.8	15.2	19.2	1.0	16.9	15.0	19.8	1.2	16.8	14.5	18.4	1.0
Snout length	10.5	8.2	12.2	1.0	10.3	9.0	12.1	1.0	10.8	9.7	12.5	0.8	10.6	9.2	12.7	1.0
% HL																
Head depth	69.4	62.4	74.8	3.1	70.3	62.4	74.7	2.9	72.4	63.2	78.2	4.0	68.1	63.6	78.6	3.4
Head width	78.7	70.0	90.2	5.4	77.2	64.8	89.9	5.2	77.6	70.3	84.2	4.0	72.8	65.4	79.6	3.2
Internostral distance	30.2	27.3	33.7	1.6	30.8	27.0	34.8	1.7	31.7	27.2	34.7	1.9	31.4	26.9	34.1	1.7
Interorbital width	46.4	41.7	50.2	2.9	50.1	46.3	54.9	2.3	48.0	38.5	55.7	4.4	49.0	44.0	54.0	2.6
Orbit diameter	20.8	19.5	23.5	1.1	21.4	18.9	25.0	1.6	22.1	19.4	25.0	1.6	22.7	20.4	24.7	1.3
Postorbital length	38.7	33.6	43.9	2.8	40.4	36.3	46.6	2.6	38.0	32.7	43.2	3.4	36.6	34.1	41.5	1.9
Snout length	44.1	38.1	49.0	3.2	40.6	35.1	46.3	3.4	44.7	36.9	48.5	2.2	42.9	36.9	48.5	3.0

numbers of scale rows were taken. Vertebrae were counted as precaudal, intermediate and caudal vertebrae, without including the urostyle. Precaudal vertebrae have no closed haemal arches and caudal vertebrae have a joined parapophyses forming an arch. The parapophyses of the ultimate precaudal vertebrae may be joined by a bony bridge, forming a narrow arch, but their ends are not joined (Ahnelt and Duchkowsch 2004). In this case

these vertebrae were counted as intermediate vertebrae. Since the posterior two rays of the dorsal fin and the anal fin are branches of the same ray, they were counted as 1½. Unbranched dorsal fin rays and unbranched anal fin rays could only be counted via micro CT images.

For micro CT scanning, 30 individuals were used (West clade: $n = 7$; Central clade: $n = 7$; North clade: $n = 8$; East clade: $n = 8$). The fish were mounted

vertically in plastic tubes containing 70% ethanol. The diameter of sample holder tubes was 12 mm. For each tube two fish were mounted in opposite direction. Micro CT scans were acquired using a Scanco μ CT35 (SCANCO Medical AG, Brüttisellen, Switzerland) at 70 kVp source voltage and 114 μ A intensity. Projections were recorded with 420 ms integration time (camera binning = 2) and an angular increment of 0.36°. Reconstructed slices measured 1024 $\times$ 1024 pixel, and isotropic voxel resolution of reconstructed volumes was 12 μ m. Image processing and analysis was done using Fiji (Schindelin et al. 2012; Fiji ImageJ version 2.0.0) and Amira 6 (FEI Visualization Sciences Group, Mérignac Cédex, France).

Morphometrics

Morphometric measurements of the preserved specimens of *G. barreimiae* are given for each clade in Table 2. All measurements show overlaps between all of the mt clades of *G. barreimiae*, although many measurements show significant differences (Table 3). A probability value of less than 0.05 of the statistical tests was considered significant. The following measurements are not normally distributed and are thus not suitable for ANOVA, Tukey's HSD test and PCA: body depth (West clade), dorsal fin height (North clade), head depth in % of head length (East clade and Central clade), predorsal length (Central clade), pectoral fin length (West clade). The following measurements have no equality of variance and are thus not suitable for ANOVA, Tukey's HSD and PCA: pectoral fin length,

caudal peduncle length, interorbital width in % of head length, anal fin - anus, postorbital length in % of head length, caudal peduncle depth. Parametric measurements which show significant differences between the clades in ANOVA are plotted in Fig. 2.

Results for Levene's test, ANOVA and Kruskal-Wallis test are given in the appendix. Results for Tukeys HSD test are shown in Table 3 based on the results of ANOVA.

In the PCA (Fig. 3) PC1 explains 51.65% of the variance and PC2 explains 19.61% of the variance. Ellipses with 75% tolerance are used in the plot. The four different clades are widely overlapping. In the LDA Axis 1 (Fig. 4) explains 53.19% of the variance. Axis 2 explains 41.16% of the variance. The four clades are overlapping but also differentiated to some extent, especially the West clade and the Central clade. 75.86% (63.22% Jackknife tested) of all specimens are correctly classified in the confusion matrix.

Meristics

Vertebral numbers, positions of dorsal fin and anal fin pterygophores, pharyngeal teeth and unbranched dorsal and anal fin rays were counted using micro CT. Other structures were counted using light microscopy. Meristics were summarized in Table 4. Meristics of gill rakers, vertebrae and branched pelvic fin rays can be used to distinguish the West clade from the other clades. Intermediate vertebral numbers: one in West clade (Fig. 5) versus mostly two in the other clades (Fig. 6), mostly i,7 pelvic fin rays in the West clade and i,8 in the

Table 3 Tukey's honest significant difference test of parametric morphometric measurements of the four clades depicts which clades differ significantly from one another in specific measurements

West Central	West North	West East	Central East	Central North	North East
AFCDF 0.0069	INTN%HL 0.0152	HW%HL 0.0005	BD 0.0006	HL 0.0266	BD 0.0002
HL 0.0036	OD%HL 0.0159	OD%HL 0.0003	HW%HL 0.0125	SNL%HL 0.0003	HW%HL 0.0049
PCTPV 0.0002	PCTPV 0.0217	PCTPV 0.002	OD%HL 0.0158		
SNL%HL 0.0016	PRPV 0.019	PRPV 0.038			
	PVFAF 0.0464	PVFAF 0.0045			

AFCDF: anal fin - caudal fin; BD: body depth; HL: head length; HW%HL: head width in % of head length; INTN%HL: internostril distance in % of head length; OD%HL: orbit diameter in % of head length; PCTPV: pectoral fin - pelvic fin; PRPV: prepelvic length; PVFAF: pelvic fin - anal fin; SNL%HL: snout length in % of head length

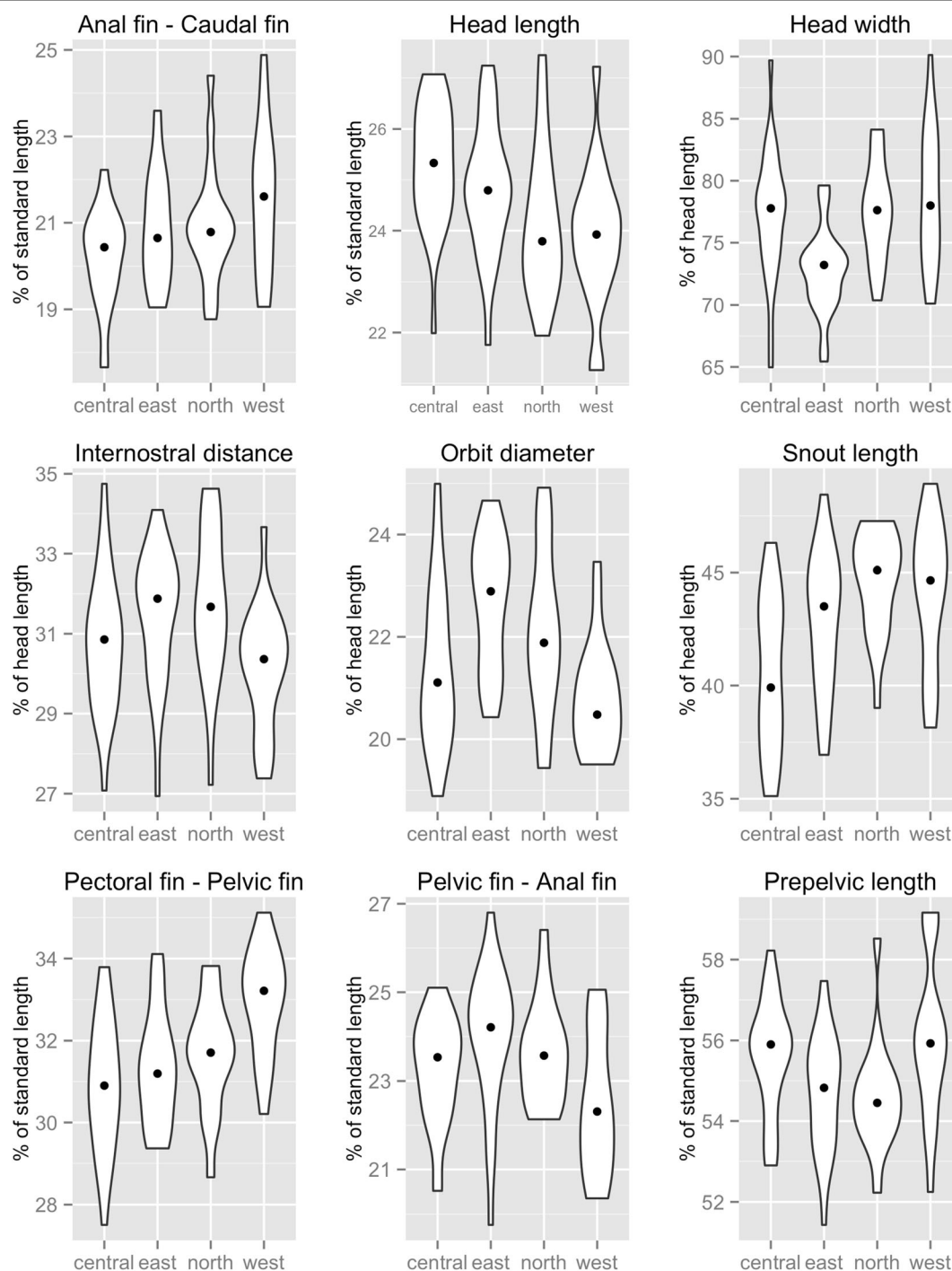


Fig. 2 Violin plots based on nine parametric measurements which show significant differences between the four mt clades (x-axis) in One-way ANOVA. The median is shown in each plot as a black dot

other clades (Fig. 7). Typically, there are 17 precaudal vertebrae in the East clade, 18 in the Central clade and North clade and 19 in the West clade, although there are overlaps. There are 15 to 17 gill rakers on the lower limb

of the first gill arch in the West clade, clearly separating it from the other clades with 12 to 14 in the Central clade and 10 to 14 in the North clade and East clade. Across all clades the following traits are quite stable: Dorsal fin

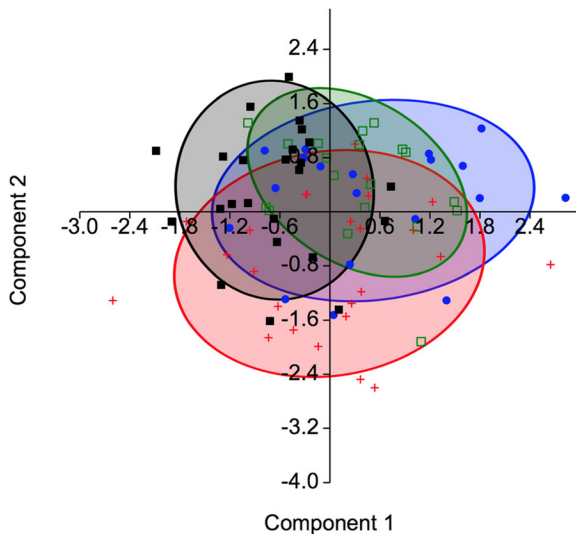


Fig. 3 PCA of significant parametric measurements with ellipses indicating 75% tolerance. PC1 explains 51.65% of the variance and PC2 explains 19.61% of the variance. West clade: blue, Central clade: red, North clade: green, East clade: black

rays of all clades are $\text{iii}, 7\frac{1}{2}$ and anal fin rays are $\text{iii}, 5\frac{1}{2}$. Individual cases of two unbranched anal fin rays as well as $4\frac{1}{2}$ branched anal fin rays occur. The first two pectoral fin rays of *Garra* are usually unbranched (Lundberg and Marsh 1976; Stiassny and Getahun 2007). In *G. barreimiae*, however, one or two unbranched fin rays are present, but one unbranched pectoral fin ray is the common case. The number of branched rays is very

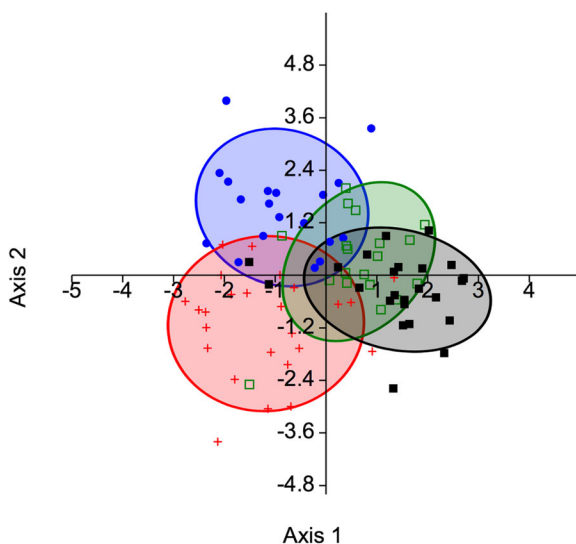


Fig. 4 LDA of significant parametric measurements with ellipses indicating 75% tolerance. Axis 1 explains 53.19% of the variance. Axis 2 explains 41.16% of the variance. West clade: blue, Central clade: red, North clade: green, East clade: black

variable. The first unbranched dorsal fin ray and anal fin ray as well as the final branched pectoral fin rays often are extremely small and thus hard to detect. The caudal fin usually has 10 upper and nine lower principal fin rays (equivalent to $9 + 8$ branched caudal fin rays). Gill rakers are small, sometimes only knob-like, very fragile and present on the lower limb (ceratobranchial), but sometimes missing on the upper limb (epibranchial). There are 12 or 13 predorsal scales with individual cases of 11 or 14 predorsal scales, however the pattern very often is irregular. There are four scales above the lateral line and three scales below the lateral line. Individual cases of five scales above the lateral line and four scales below the lateral line occur. All fish have four barbels. Pharyngeal teeth are in three rows: typically 2.4.5–5.4.2. Teeth numbers of 1.2.5, 1.3.5 or 2.3.5 occur.

Discussion

With the introduction of molecular genetic methods into phylogenetic questions, distinct lineages in *G. barreimiae* have been detected recently (Kirchner et al. 2017). Our results show that at least some of these lineages are differentiated by morphological characters. All specimens of this study have originally been caught for genetic analyses only. Therefore, the samples sometimes did not represent the full size spectrum of a population and contained many small individuals. However, this study is based on the largest number of morphologically examined specimens of the *G. barreimiae* species complex to this point. The morphometric work that had been previously performed on *G. barreimiae* yielded contradicting results, which could be interpreted as even higher morphological variability than observed in the present study. Yet they could also be explained by different techniques of measurement and counting. As an example, the number of dorsal fin rays was stated as $\text{iii}, 7$ by Fowler and Steinitz (1956), $\text{iii}, 6$ by Menon (1964), $\text{iv}, 6$ or 8 by Banister and Clarke (1977) and $\text{iv}, 6–8$ by Krupp (1983). As demonstrated in the results, all specimens in this study had $\text{iii}, 7\frac{1}{2}$ dorsal fin rays. Especially the first unbranched dorsal fin ray often is rather small and could be observed in greater detail using micro CT.

Banister & Clarke divided *G. barreimiae* into *G. b. shawkahensis* (from one locality in Wadi Shawkah) and *G. b. barreimiae* (from all other mentioned localities and collections) based on a different number of gill rakers on the lower limb of the first gill arch (*G. b. barreimiae*:

Table 4 Meristic counts and associated number of specimens per clade. Dorsal fin pterygophore insertion refers to the precaudal vertebrae number anterior to the first major dorsal fin pterygophore. Anal fin pterygophore insertion refers to the caudal vertebrae number anterior to the first major anal fin pterygophore

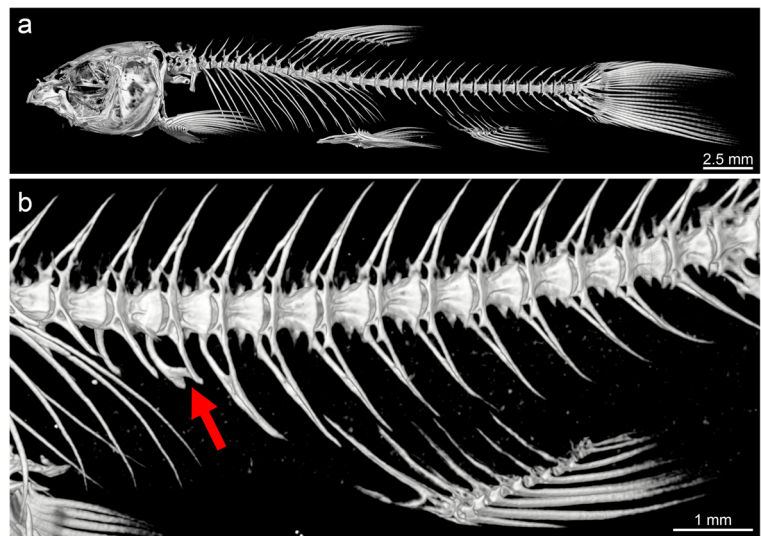
	Count	West	Central	North	East
Precaudal vertebrae	17	–	–	2	7
	18	1	6	5	–
	19	6	1	1	–
Intermediate vertebrae	1	7	1	1	2
	2	–	6	6	5
	3	–	–	1	–
Caudal vertebrae	12	1	1	1	–
	13	5	4	6	4
	14	1	1	1	3
Total vertebrae	31	–	–	–	1
	32	1	1	1	4
	33	6	4	7	2
	34	–	1	–	–
Dorsal fin pterygophore insertion	10	–	3	4	6
	11	6	4	4	1
	12	1	–	–	–
Anal fin pterygophore insertion	1	2	1	5	1
	2	5	6	2	5
	3	–	–	1	1
Lateral line scales	30	–	6	6	2
	31	11	16	11	13
	32	6	3	2	8
	33	2	–	–	–
Lateral scales on caudal fin base	1	–	1	–	–
	2	11	23	8	17
	3	8	1	11	6
Total lateral scale count	32	–	6	–	2
	33	6	16	13	7
	34	8	3	5	14
	35	5	–	1	–
Scales between anus and anal fin	1	–	2	1	3
	2	17	20	10	11
	3	2	3	8	7
	4	–	–	–	1
Scales between anus and pelvic fin	4	2	–	–	2
	5	4	1	5	3
	6	11	16	10	16
	7	1	8	3	2
Postdorsal scales	13	–	1	–	6
	14	4	3	2	4
	15	7	8	14	10

Table 4 (continued)

	Count	West	Central	North	East
Circumpeduncular scales	16	2	5	2	1
	12	7	1	12	11
	13	1	–	3	4
	14	9	8	3	7
	15	2	8	–	2
	16	–	8	1	–
Breast scales	yes	12	18	11	10
	no	7	7	8	14
Gill rakers on the lower limb of the first gill arch	10	–	–	3	2
	11	–	–	3	6
	12	–	12	3	5
	13	–	6	6	4
	14	–	2	2	5
	15	7	–	–	–
	16	3	–	–	–
	17	1	–	–	–
Branched pelvic fin rays	7	18	3	4	4
	8	1	22	15	19
Unbranched pectoral fin rays	1	16	24	15	18
	2	3	1	4	4
Branched pectoral fin rays	11	–	1	–	4
	12	3	3	1	4
	13	6	15	8	9
	14	5	3	6	5
	15	5	3	4	–
Tubercles	yes	13	11	11	13
	no	6	14	8	11

11–14; *G. b. shawkahensis*: 15–18) and a different mean prepelvic distance (*G. b. barreimiae*: 53.9%; *G. b. shawkahensis*: 56.2%). They also described *G. longipinnis* from the Saiq region based on eight specimens with strikingly long paired fins (Banister and Clarke 1977) like the pectoral fins with 23.3–32.9%. Saiq is within the region of this study's Central clade and Hamidan et al. (2014) identified fish from that area as *G. cf. longipinnis*. The pectoral fin is variable in the Central clade with a range of 17.9–24.8%, thus some specimens would fit the lower range values of the original species description of *G. longipinnis*, however the mean height of the pectoral fins is not longer in the Central clade than in the others. Krupp (1983) published a species key for Arabian *Garra* in which *G. b. barreimiae* should have less than fifteen gill rakers on

Fig. 5 Micro CT image of selected specimens of the West clade: all examined West clade specimens have one intermediate vertebra (indicated with red arrow). The upper image (a) shows the whole skeleton, the lower image (b) represents a detailed view on the intermediate vertebra, caudal vertebra and the anal fin

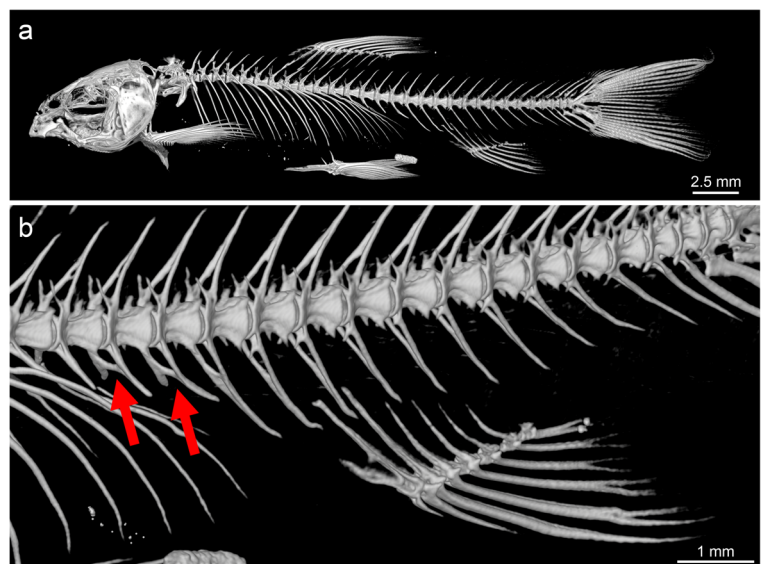


the lower limb of the first gill arch and *G. b. shawkahensis* should have more than fourteen gill rakers, which was the substantial morphological difference between the two subspecies.

In this study, the mean prepelvic distance is the largest in the West clade with 56.0% (Fig. 2, Table 2) and thus in fact very close to that reported by Banister and Clarke (1977) (56.2%) for *G. b. shawkahensis*. In addition, all West clade specimens have more than 14 gill rakers on the lower limb of the first gill arch, further fitting the description of *G. b. shawkahensis* of Banister and Clarke as well as that of Krupp (1983). One might be tempted to expand the known range of *G. b.*

shawkahensis by our three West clade sampling sites, but things may be more complicated. Fowler and Steinitz did not measure the prepelvic distance and did not count gill rakers in their original description of *G. barreimiae*. Banister and Clarke did not take measurements of specimens from Buraimi, but considered them to belong to the same taxon as specimens from Wadi Sahtan, a locality that clearly belongs to another genetic clade (North clade), as well as fishes from undisclosed regions of Oman. In an analysis of mt sequences (Kirchner et al. 2017) fishes from Wadi Shawkah did cluster together with specimens from the West clade (Hafeet, Hayyal and Al Juwayf).

Fig. 6 Micro CT image of an East clade specimen. The Central, North, or East clade usually have two intermediate vertebrae (indicated by red arrows). The upper image (a) shows the whole skeleton, the lower image (b) shows details of the intermediate vertebrae, caudal vertebrae and the anal fin



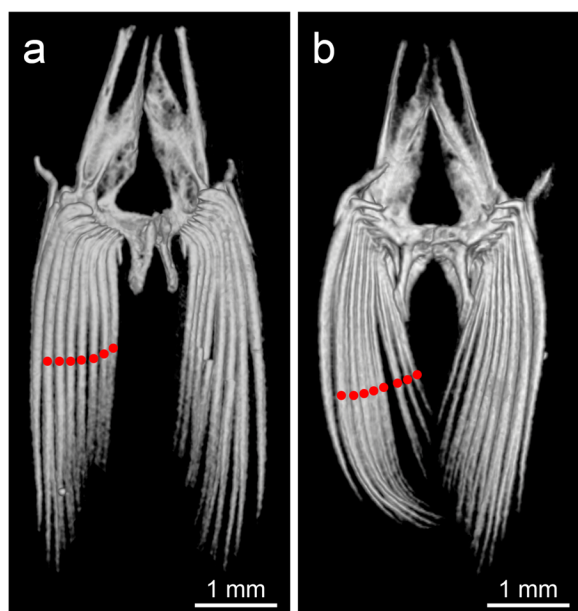


Fig. 7 Variations of pelvic fin ray numbers. West clade specimens (a) typically have i,7 pelvic fin rays, while the other clades (b), with a Central clade specimen depicted, usually have i,8 pelvic fin rays. Branched fin rays are marked by red dots

Furthermore, all sampling sites from the West clade and the original description sites of *G. barreimiae* and *G. b. shawkahensis* are found in the same region, the western slopes of the Al Hajar Mountains (Fig. 1). In the Catalog of Fishes *G. b. shawkahensis* is considered to be a synonym of *G. barreimiae* (Eschmeyer et al. 2017). During two collection trips of the Natural History Museum Vienna, the original description site of *G. barreimiae*, Buraimi, was visited, but no fish could be found. Type specimens have not been examined in this study, thus comparisons refer exclusively to literature. However, it is very likely that Fowler & Steinitz' *G. barreimiae*, Banister & Clark's *G. b. shawkahensis* and this study's specimens from the sampling sites Hafeet, Hayyal and Al Juwayf belong to the same taxon. The West clade fits the original description of *G. barreimiae* best, both concerning distribution and morphology. Whether the classification of *G. b. shawkahensis* as subspecies is justified, remains open at the current stage of knowledge.

The sampling locations of the mt clades Central and East have not been mentioned in any valid species description. Individuals belonging to the North clade have at least partially falsely been grouped together with individuals of *G. b. barreimiae* in the past. All morphometric measurements overlap between the clades and

are very variable within the clades (Fig. 2, Table 2). The clades are not separated in the PCA (Fig. 3), although in the LDA (Fig. 4) the West clade, Central clade and East clade are distinguishable and most specimens can be correctly classified based on the same measurements in the confusion matrix. A noteworthy difference is the pelvic fin ray count with seven branched rays in most of the West clade specimens and eight branched rays in most specimens of the other clades (Table 4; Fig. 7). The West clade also has generally nineteen precaudal and one intermediate vertebrae, both of which are very uncommon in the other clades. All East clade specimens have seventeen precaudal vertebrae, a feature apart from that only found in two of the North clade specimens.

Conclusion

Besides the number of gill rakers, which distinctly separates the West clade from the other clades, no clearly well-defined morphological character distinguishing the clades could be identified with the methods applied. Several morphometric measurements however, display significant differences between the clades. Measurements of head structures are particularly conspicuous in all clades and therefore should be considered as primary targets of more precise morphological methods (e.g. geometric morphometrics). Besides a few overlaps, the numbers of precaudal and intermediate vertebrae as well as those of branched pelvic fin rays indicate that the West clade is morphologically distinct from the other clades. The original descriptions of *G. barreimiae* and *G. b. shawkahensis* correspond to the West clade both geographically and morphologically. According to the concept of Integrative Taxonomy multiple disciplines should be used for species delimitation with morphology and methods revealing nuclear genetic information always being applied (Schlick-Steiner et al. 2010). Thus, concerning the Central clade, North clade and East clade, the question whether they represent separate taxa has to be addressed combining genetic and morphological data.

Acknowledgements Open access funding provided by University of Vienna. We thank Oliver Macek, Julia Schindelar, Barbara Tautscher and Lorena Toscano for their continued support in the lab, as well as Robert Illek and Lukas Plan for helping with the collection of samples. We want to express our gratitude to the Ministry of Environment and Climate Affairs of Oman and the Austrian Economy Chambers for their cooperation, support and sampling permits.

Appendix

Table 5 Results of Levene's test for homogeneity of variance, ANOVA and Kruskal-Wallis test for each morphometric measurement. In the case of non-parametric data, results of the Kruskal-Wallis test are preferable over ANOVA. In this case results of the Kruskal-Wallis test are written in bold letters. A probability value of less than 0.05 is considered significant

	Levene's test	ANOVA	Kruskal-Wallis test
Anal fin - Anus	0.0012	2.460E-10	3.652E-08
Anal fin - Caudal fin	0.2325	0.0125	0.0430
Anal fin height	0.0881	0.0808	0.1639
Body depth	0.6192	1.263E-06	1.445E-05
Caudal peduncle depth	0.0035	1.649E-06	9.943E-05
Caudal peduncle length	0.0170	0.0006	0.0033
Dorsal fin height	0.9175	0.0003	0.0002
Head depth (% of HL)	0.4949	0.0007	0.0005
Head length	0.8436	0.0023	0.0013
Head width (% of HL)	0.1002	0.0002	0.0002
Interostralar dist. (% of HL)	0.8597	0.0196	0.0114
Interorbital width (% of HL)	0.0367	0.0012	0.0032
Orbit diameter (% of HL)	0.2250	0.0001	0.0002
Pectoral fin height	0.0154	0.0625	0.1778
Pectoral fin - Pelvic fin	0.7127	0.0001	0.0030
Postorbital length (% of HL)	0.0040	4.836E-05	0.0001
Postdorsal length	0.3254	0.0539	0.0777
Preanal length	0.1987	0.1715	0.1955
Predorsal length	0.7126	0.0079	0.0335
Prepelvic length	0.3664	0.0105	0.0036
Pelvic fin - Anal fin	0.4907	0.0063	0.0180
Pelvic fin - Caudal fin	0.6056	0.0895	0.1220
Pelvic fin height	0.8987	0.8523	0.7766
Snout length (% of HL)	0.1126	9.016E-05	0.0003

Table 6 Confusion matrix of LDA. 75.86% of specimens were correctly classified based on significant parametric measurements. Rows show the given groups, columns show the predicted groups

	West	Central	North	East	Total
West	15	0	4	0	19
Central	2	19	1	3	25
North	2	1	12	4	19
East	1	1	2	20	24
Total	20	21	19	27	87

Table 7 Confusion matrix of LDA. After Jackknife resampling 63.22% of specimens were correctly classified based on significant parametric measurements. Rows show the given groups, columns show the predicted groups

	West	Central	North	East	Total
West	12	1	6	0	19
Central	2	18	2	3	25
North	2	1	9	7	19
East	1	1	6	16	24
Total	17	21	23	26	87

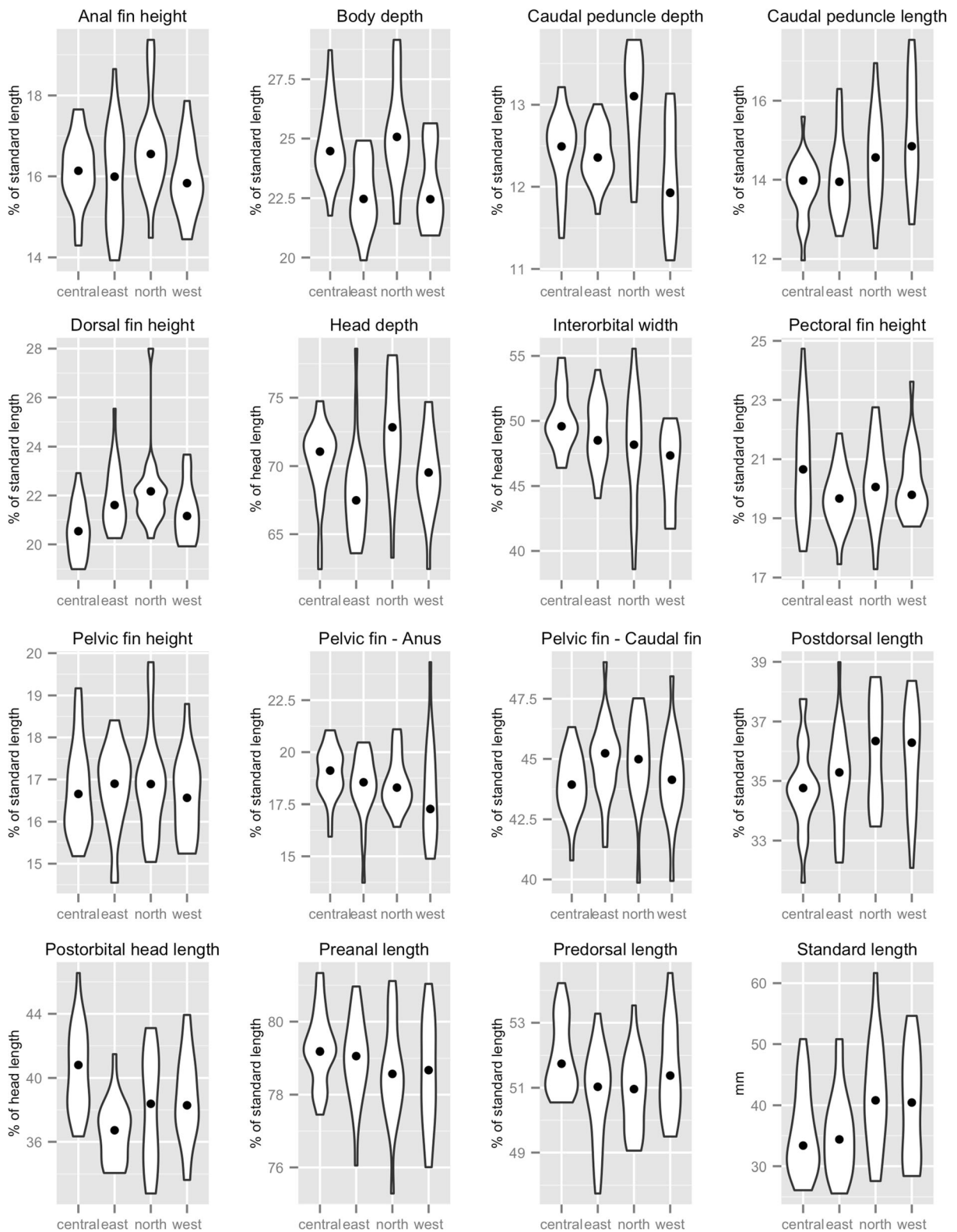


Fig. 8 Violin plots of non-parametric measurements of the four mt clades (in % of standard length and % of head length) and absolute measurements of the standard length. The median is shown in each plot as a black dot

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Ahnelt H, Duchkowsitch M (2004) The postcranial skeleton of *Proterorhinus marmoratus* with remarks on the relationships of the genus *Proterorhinus* (Teleostei: Gobiidae). *J Nat Hist* 38:913–924. <https://doi.org/10.1080/0022293021000047873>
- Al-Rawahi MN, Brinkmann K, Schlecht E, Buerkert A (2014) Effects of changing water availability on land use in irrigated mountain oases of Al Jabal Al Akhdar, northern Oman. *Erde* 145:197–211. <https://doi.org/10.12854/erde-145-18>
- Armbruster JW (2012) Standardized measurements, landmarks, and meristic counts for cypriniform fishes. *Zootaxa* 3586:8–16
- Banister KE (1984) A subterranean population of *Garra barreimiae* (Teleostei: Cyprinidae) from Oman, with comments on the concept of regressive evolution. *J Nat Hist* 18: 927–938. <https://doi.org/10.1080/00222938400770811>
- Banister KE (1987) Two new species of *Garra* (Teleostei: Cyprinidae) from the Arabian peninsula. *Bull Br Museum Natural Hist* 52:59–70
- Banister KE, Clarke MA (1977) The freshwater fishes of the Arabian peninsula. *J Oman Stud*:132–144
- Eschmeyer WN, Fricke R, van der Laan R (eds) (2017) CATALOG OF FISHES: GENERA, SPECIES, REFERENCES. (<http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp>). Electronic version accessed 11 12 2017.
- Esmaili HR, Sayyadzadeh G, Coad BW, Eagderi S (2016) Review of the genus *Garra* Hamilton, 1822 in Iran with description of a new species: a morpho-molecular approach (Teleostei: Cyprinidae). *Iran J Ichthyol* 3(2):82–121
- Fowler HW, Steinitz H (1956) Fishes from Cyprus, Iran, Iraq, Israel and Oman. *Bull Res Counc Isr* 5B:262–289
- Freyhof J, Hamidan NA, Feulner G, Harrison I (2015) The status and distribution of freshwater fishes of the Arabian peninsula. In: García N, Harrison I, Cox N, Tognelli MF (eds) The status and distribution of freshwater biodiversity in the Arabian peninsula. IUCN, Gland, pp 16–29
- Hamidan NA, Geiger MF, Freyhof J (2014) *Garra jordanica*, a new species from the Dead Sea basin with remarks on the relationship of *G. ghorensis*, *G. tibanica* and *G. rufa* (Teleostei: Cyprinidae). *Ichthyol Explor Freshwaters* 25:193–288
- Hammer Ø, DAT H, Ryan PD (2001) PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontol Electron* 4(1):9
- Kirchner S, Sattmann H, Haring E, Plan L, Victor R, Kruckenhauser L (2017) Investigating gene flow between the blind cavefish *Garra barreimiae* and its conspecific surface populations. *Sci Rep* 7(5130). <https://doi.org/10.1038/s41598-017-05194-3>
- Kruckenhauser L, Haring E, Seemann R, Sattmann H (2011) Genetic differentiation between cave and surface-dwelling populations of *Garra barreimiae* (Cyprinidae) in Oman. *BMC Evol Biol* 11(172). <https://doi.org/10.1186/1471-2148-11-172>
- Krupp F (1983) Freshwater fishes of Saudi Arabia and adjacent regions of the Arabian peninsula. In: Fauna of Saudi Arabia. Basel, Jeddah, pp 568–636
- Krupp F (1988) Freshwater fishes of the Wadi Batha drainage. *J Oman Stud* 401–404
- Krupp F, Budd K (2009) A new species of the genus *Garra* (Teleostei: Cyprinidae) from Oman. *Aqua* 15(2):117–120
- Lundberg JG, Marsh E (1976) Evolution and functional anatomy of the pectoral fin rays in cyprinoid fishes, with emphasis on the suckers (family Catostomidae). *Am Midl Nat* 96:332–349
- Lyon RG, Geiger MF, Freyhof J (2016) *Garra sindhi*, a new species from the Jebel Samhan nature Reserve in Oman (Teleostei: Cyprinidae). *Zootaxa* 4154:79–88. <https://doi.org/10.11646/zootaxa.4154.1.5>
- Menon A (1964) Monograph of the cyprinid fishes of the genus *Garra* Hamilton. *Mem Indian Museum* 14:173–260
- Mousavi-Sabet H, Eagderi S (2016) *Garra lorestanensis*, a new cave fish from the Tigris River drainage with remarks on the subterranean fishes in Iran (Teleostei: Cyprinidae). *FishTaxa* 1:45–54
- Mousavi-Sabet H, Vatandoust S, Fatemi Y, Eagderi S (2016) Tashan cave a new cave fish locality for Iran; and *Garra tashanensis*, a new blind species from the Tigris River drainage (Teleostei: Cyprinidae). *FishTaxa* 1(3):133–148
- R Development Core Team (2015) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.
- Sayyadzadeh G, Esmaili HR, Freyhof J (2015) *Garra mondica*, a new species from the Mond River drainage with remarks on the genus *Garra* from the Persian Gulf basin in Iran (Teleostei: Cyprinidae). *Zootaxa* 4048:75–89. <https://doi.org/10.11646/zootaxa.4048.1.4>
- Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez JY, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9(7):676–682
- Schlick-Steiner BC, Steiner FM, Seifert B, Stauffer C, Christian E, Crozier RH (2010) Integrative taxonomy: a multisource approach to exploring biodiversity. *Annu Rev Entomol* 55: 421–438. <https://doi.org/10.1146/annurev-ento-112408-085432>
- Stiassny MLJ, Getahun A (2007) An overview of labeonin relationships and the phylogenetic placement of the afro-Asian genus *Garra* Hamilton, 1922 (Teleostei: Cyprinidae), with the description of five new species of *Garra* from Ethiopia, and a key to all African species. *Zool J Linnean Soc* 150:41–83
- Wickham H (2009) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York
- Yang L, Mayden RL (2010) Phylogenetic relationships, subdivision, and biogeography of the cyprinid tribe Labeonini (*sensu* Rainboth, 1991) (Teleostei: Cypriniformes), with comments on the implications of lips and associated structures in the labeonin classification. *Mol Phylogenet Evol* 54:254–265

Chapter II: Phylogeography and cryptic diversity of *G. barreimiae*

Paper 4: Hidden diversity – Delimitation of cryptic species and phylogeography of the cyprinid *Garra barreimiae* complex in Northern Oman

Hidden diversity – Delimitation of cryptic species and phylogeography of the cyprinid *Garra barreimiae* complex in Northern Oman

*Sandra Kirchner^{1,2}, Helmut Sattmann³, Elisabeth Haring^{1,2}, Reginald Victor⁴, Luise Kruckenhauser²

¹ University of Vienna, Department of Integrative Zoology, Althanstraße 14, 1090 Vienna, Austria

² Natural History Museum Vienna, Central Research Laboratories, Burgring 7, 1010 Vienna, Austria

³ Natural History Museum Vienna, Third Zoological Department, Burgring 7, 1010 Vienna, Austria

⁴ Sultan Qaboos University, Department of Biology, Al Khoudh, Muscat 123, Oman

Email: * sandra.kirchner@nhm-wien.ac.at, helmut.sattmann@nhm-wien.ac.at, elisabeth.haring@nhm-wien.ac.at,
regvictor@gmail.com, luise.kruckenhauser@nhm-wien.ac.at

Keywords: *Garra barreimiae*, Phylogeography, Oman, Hajar Mountains, Microsatellite Markers, Cryptic Species, Cyprinid, Population genetics

INTRODUCTION

Phylogeographic studies are an important tool to help understanding a taxons' historic background and create knowledge about its genetic variation, distribution and its limits which is crucial to evaluate biodiversity (Moyle & Leidy, 1992; Zachos & Habel, 2011).

The generally known major causes for biodiversity loss include habitat destruction, pollution and climate change. Besides, also the lack of information on hidden species diversity as well as the non-description of newly found cryptic species has severe consequences as these unrecognized or uncatalogued taxa are not protected, leading further to a major loss of biodiversity (Zachos & Habel, 2011).

With the advance of DNA-based techniques for detection of genetic diversity it became evident that biodiversity has been grossly underestimated in the past (Adams *et al.*, 2014). Cryptic species have been overlooked by taxonomists for a long time, because conventional assessment was based mainly on morphological traits. Adams *et al.* (2014) demonstrated, that even a well-studied organism like the Australian freshwater fish *Galaxias olidus* exhibited species diversity many times higher than previously

assumed after including molecular genetic techniques in the assessment. With this in mind, one can only speculate how much undetected diversity is hidden in taxa that are not that well studied.

In the present study we exemplify this issue using the genus *Garra* (Cyprinidae) from the southeastern Arabian Peninsula. According to Durand et al. (2002) the subfamily Cyprininae in Middle East is much more diversified than in the Euro-Mediterranean region. Por (1989) stated that this geographic region was an important centre of radiation before the Pliocene orogenesis. Even though scientific positions are discordant whether the Middle East was also a centre of speciation, there is broad agreement that it was a biogeographic interchange area of great importance. Hence, the Middle East has potentially a much higher, yet undiscovered species richness than so far assumed.

The landscape of Northern Oman is shaped by high mountain ridges up to 3000 m asl and countless wadis (ephemeral riverbeds) that channel the water bodies resulting from the scarce precipitation events. The climate is hot and arid, precipitation is irregular but can result in heavy rainfalls followed by massive floodings. The amount of water pouring down during one of these torrential flash floods can have massive impacts and can exceed the average annual precipitation in this area (Al-Shaqsi, 2010; Gunawardhana & Al-Rawas, 2014). These conditions are challenging for freshwater dependant animals and only a few well adapted organisms like the cyprinid barb *Garra barreimiae* FOWLER & STEINITZ, 1956, managed to inhabit this region permanently. This species is abundant throughout the Hajar mountains in Northern Oman and the United Arab Emirates (UAE) where it inhabits all kinds of water bodies, and even artificial water drainage systems (aflaj/falaj). Kruckenhauser et al. (2011) were the first to uncover an unexpected genetic diversity of *G. barreimiae* and consequently questioned the monophyly of this taxon. This finding was more or less a by-product and not discussed in detail since the study aimed to look for differentiation between the troglomorphic cave population inhabiting the Al Hoota cave and surface populations of *G. barreimiae*. Subsequently Pichler et al. (2018) investigated the morphological variation of the species throughout its distribution range in Oman and found only scarce morphological differentiation. Studies like these illustrate that the level of genetic variation is not necessarily coherent with morphological differentiation, emphasizing once more the importance of thorough and careful taxonomic decisions based on data from multiple sources (e.g. not only single-gene barcoding, nor only morphological data) as it is pointed out under the concept of integrative taxonomy (Schlick-Steiner *et al.*, 2010).

The presented molecular genetic study pursues the following research questions: What is the extent of genetic variation within the taxon *G. barreimiae* and its corresponding subspecies? How genetically variable are the samples collected throughout the distribution area in Northern Oman? Can genetic groups be distinguished and if so, are such groups concordant with their geographic distribution? What are the phylogenetic relationships between *G. barreimiae* and closely related taxa? Is *G. barreimiae* non-monophyletic in the DNA based phylogenetic tree as suggested earlier? Are the genetic results concordant with the previously published morphological data?

In the course of finding answers to the posed questions, we investigated in detail how much genetic diversity can be found within *G. barreimiae* and its corresponding subspecies. This was done by analysing three mitochondrial (mt) marker sequences as well as 19 polymorphic nuclear (nc) microsatellite markers. Samples from numerous sites covering the whole distribution area in Northern Oman have been analysed to detect potential phylogeographic structures. Furthermore, it was examined if the subspecies of *Garra barreimiae* are genetically differentiated and compared the ranges of genetic diversification to other species of the genus *Garra*.

MATERIAL AND METHODS

Study area and sampling

Tissue samples (finclips) of *G. barreimiae* were taken from 297 individuals. Fish were collected from 56 sampling sites covering the whole distribution area in Oman (Fig. 1). Specimens were captured with fishing nets and anesthetized with clove oil, then fixed in 80% ethanol. Specimens and their DNA are stored in the Natural History Museum in Vienna (NHMW; Fish Collection and DNA and Tissue Collection).

Altogether 632 specimens were included in the present study comprising *G. barreimiae* and *Garra gallagheri* as well as representatives of other closely related taxa: *Garra rufa*, *Garra sahilia*, *Phreatichthys andruzzii*. *G. gallagheri* was listed as a subspecies of *G. barreimiae* until it was recently raised to species status by Lyon, Geiger, and Freyhof (2016) based on the data of Kruckenhauser et al. (2011). The latter study already provided the information that *G. gallagheri* is placed in the phylogenetic analysis within *G. barreimiae*, thus it was included in our analysis of the intraspecific genetic variation. While 302 individuals were analysed for the first time in the present study, the remaining individuals were already used in earlier studies. Table S1 provides details on specimens and the markers analysed for each specimen. For the microsatellite analyses, altogether 524 samples were analysed, including also 213 samples that had been already genotyped in a previous study (Kirchner *et al.*, 2017). For the analyses of the *CR* sequences 500 individuals were included, among them 117 sequences published by Kruckenhauser et al. (2011) (accession numbers: JF709013 - JF709129). From 47 individuals the *Cytb* and the *COI* genes were analysed in addition to the *CR*.



Figure 1 Study area and sampling site. Overview of the study area on the Arabian Peninsula and detailed map of Northern Oman with sampling sites indicated as white dots.

DNA extraction, PCR amplification, genotyping and sequencing

Genomic DNA was extracted from tissue of fin clips preserved in 80% ethanol following the QIAGEN DNeasy Blood & Tissue extraction kit protocol. Partial sequences of three mt markers were analysed (*CR*, *COI*, *Cytb*). For the *CR*, a ~940 bp long amplicon (*CR*-long) was generated which then was trimmed in order to include also shorter sequences (*CR*-short) generated in a previous study (Kruckenhauser et al. 2011). Primers and PCR conditions used for PCR amplification of the various mt sequences are shown in Table 1. DNA amplifications were conducted in 25 µl reaction volume containing 1 µl DNA solution, 5 µl 5x Q-Solution, 2.5 µl 10x PCR Buffer, 0.5 mM of each dNTP, 1.5 mM Mg²⁺, 0.5 µM of each primer, 0.5 units TopTaq DNA polymerase, 13.9 µl AD for both the *COI* and *Cytb* section, and 1 µl DNA, 2.5 µl 10x PCR buffer, 0.5 mM of each dNTP, 1.5 mM Mg²⁺, 0.5 µM of each primer, 0.5 units Platinum Taq DNA Polymerase, 18.9 µl AD for *CR*. The cycling protocol for the three different fragments (*COI/Cytb/CR*) was as follows: initial denaturation at 94 °C for 3/2/3 min, 35 cycles denaturation at 94 °C for 20/20/15 s, annealing at 58 °C for 30 s/52 °C for

10 s/65 °C for 30 s, extension at 72 °C for 60 s/72 °C for 90 s/68 °C for 60 s, and final extension at 72 °C for 7 min/10 min.

Table 1 Primer sets used in the present study.

Gene	Amplicon length (bp)	N	Forward primer (5'-3')	Reverse primer (5'-3')	Analysis	Reference
<i>COI</i>	964	47	Tyr_Garra_1_fw GGCTACAACCCACCGCCTA	COI_Garra_4_rv GTTGCGGATGTAAATATGC	BI tree	ps
<i>Cytb</i>	1185	47	Glu1mod+ GACTTGAAGAACCACCGTTGT	Thr1-rv GATCTTCGGATTACAAGACCG	BI tree	a
<i>CR-short</i>	~440	481	Gar-Thr1+ GCATCGGTCTTGTAAATCCGA	TDK_DGbar1- CCTGAAATAGGAACCAGATG	Network	b
<i>CR-long</i>	~940	314	Gar-Thr1+ GCATCGGTCTTGTAAATCCGA	CR4-_Gbar TTGGGCGTCGGCGGTGAGAG	BI tree, NJ tree	b

Primer sets used in this study to obtain sequences of the mt COI, Cytb and two overlapping amplicons of CR (CR-short, CR-long). The column "Analysis" provides information on which fragment was used for the different analyses. References: ps present study, a Calderoni et al. (2016), b Kruckenhauser et al. 2011.

The resulting PCR products were purified and sequenced in both directions at Microsynth AG (Balgach, Switzerland) using the PCR primers.

For the microsatellite analyses samples were profiled with 19 microsatellite loci by using primers that were specifically designed for *G. barreimiae* (Kirchner *et al.*, 2014). The same multiplex protocols have been applied as described in detail in Kirchner *et al.* (2017). In this manuscript, the term "genotype" is used to describe the microsatellite allele composition characteristic for a population.

Alignment, sequence diversity statistics and phylogenetic reconstructions

The obtained sequences were edited manually using BioEdit v. 7 (Hall, 1999), and aligned with the program MEGA v. 6 (Tamura *et al.*, 2013). The software TrimAl v. 1.3 (Capella-Gutiérrez, Silla-Martínez, & Gabaldón, 2009) implemented in the Phylemon 2.0 web tools was employed, using the "no gap" option for automated removal of gap regions for the final *CR-long*/*CR-short* alignment.

Pairwise p-distances as well as mean p-distances within and between clades were calculated with MEGA. The program DnaSP v. 5 (Librado & Rozas, 2009) was used to estimate nucleotide and haplotype diversity. The substitution saturation in the *CR-long* fragment have been examined with DAMBE v. 6 (Xia, 2017) (pairwise deletion) and employed the software PartitionFinder v. 2 (Lanfear *et al.*, 2017) to determine the substitution model for the data set of each fragment separately (AICc model selection, search greedy, linked branch lengths).

The nucleotide substitution pattern showed that the sequences have not reached substitution saturation and are therefore well applicable for phylogenetic analyses. The best-fit substitution model for each

amplicon and codon position was TVM+G for *CR*, GTR+I for *COI* (codon position 1/2/3 = TRN+G/TRN+I/F81+I) and TRN+G for *Cytb* (codon position 1/2/3 = GTR+I+G/K80+I+G/HKY+I).

Based on the *CR*-short alignment, haplotype networks were constructed using the software PopART v. 1.7 (<http://popart.otago.ac.nz>) and the Median-joining (MJ) algorithm proposed by Bandelt, Forster, and Röhl (1999) was applied. The NJ and ML phylograms were calculated with MEGA and the BI phylograms were calculated with MrBayes v. 3 (Huelsenbeck & Ronquist, 2001). The results were visualized with FigTree v. 1.4.2.

Genetic diversity of nc microsatellite data

The nc microsatellite data has been explored mostly by employing the statistical computing software environment R version 3.3.1 (R Development Core Team, 2015). After checking the data set for completeness (0.78% missing values in complete data set), basic information was calculated like number of alleles with *R base*, allelic richness with *R diveRsity* (Keenan *et al.*, 2013), private alleles with *R poppr* (Kamvar, Tabima, & Grünwald, 2014), expected (H_e) and observed heterozygosity (H_o) per locus as well as the mean H_e and H_o over all loci per population with *R diveRsity*. In addition, the software Genepop v. 4 (Rousset, 2008) was used to test for genotypic linkage disequilibrium and the results showed that some locus pairs exhibit evidence of linkage disequilibrium in single populations, but never in all populations. Deviations from HWE were estimated using both the global test and exact test in Genepop executing 10000 Markov chain Monte Carlo (MCMC) runs for 100 batches, each with 5000 iterations.

The software Micro-Checker v. 2.2 (Van Oosterhout *et al.*, 2005) was used to check for the existence of microsatellite null alleles. In order to statistically describe the level of inbreeding or differentiation among and within populations, the fixation indices as implemented in the R-package *pegas* version 0.8.233 (Paradis, 2010) was calculated. This function computes F_{IT} , F_{ST} and F_{IS} for each locus in the data set according to the formulae of Weir and Cockerham (1984). Subsequently, the pairwise F_{ST} values between all populations have been calculated with the Microsoft Excel Add-In GenAlEx version 6.5 (Peakall & Smouse, 2012). The inbreeding coefficient F_{IS} after Weir and Cockerham (1984) over all loci per population was calculated with the program FSTAT v. 2.9 (Goudet, 1995).

Allelic richness per locus and population was computed to take differing sampling sizes of populations into account. This was done by using 1000 re-samples of the smallest sample in the data set (*gallagheri* with $N = 29$) via the *R diveRsity*. Private alleles were estimated with *R poppr* and the percentage of private alleles per population was calculated based on the actual number of alleles per population.

Population structure analysis and differentiation

The software STRUCTURE v 2.3 (Pritchard, Stephens, & Donnelly, 2000) was used to perform population structure analyses on both the complete microsatellite data set as well as for separate subsets corresponding to the mt clades. The admixture model was employed and tested for clusters from $K = 1$ to $K = 7$ for the complete data set, and 20 iterations per run were performed for each assumed K to find the most reasonable number of clusters (K). Length of burnin was set to 100000 and the number of MCMC repeats after burnin was 500000. This analysis was repeated for each cluster (subsets) that resulted from the analysis of the complete data set and tested for as many K as sampling sites were included (Central: $K = 14$, $N = 255$; *gallagheri*: $K = 2$, $N = 29$; East: $K = 5$, $N = 67$; North: $K = 9$, $N = 62$; West: $K = 4$, $N = 35$).

All results were then evaluated with the online program STRUCTURE HARVESTER web version 0.6 (Earl & VonHoldt, 2012) to explore which K has the highest probability and ΔK . Concurrently, the Structure results were loaded onto the CLUMPAK online pipeline (Kopelman *et al.*, 2015) which automatically summarizes the results and offers a graphical representation.

In addition, a discriminant analysis of principal components (DAPC) implemented in *R adegenet* (Jombart, 2008) was performed based on the complete data set.

RESULTS

Variation in mitochondrial (mt) DNA sequences

After trimming the primer regions, the length of the final alignments for each of the three sequences was 859 bp for the mitochondrial *cytochrome c oxidase subunit 1* (*COI*), 1109 bp for the mitochondrial *cytochrome b* (*Cytb*) and ~818 bp for the mitochondrial *control region* (*CR*). The alignment of the *CR*-short sequences was ~401 bp. After trimming, the *CR*-long/*CR*-short alignment had 765/382 positions.

All three mitochondrial markers were analysed in 41 specimens of *G. barreimiae* from 23 sampling sites representing all clades. Combined with sequences from other taxa the concatenated alignment comprised 47 sequences and had a length of 2739 bp (*COI*: position 1 - 859, *Cytb*: position 860 – 1968, *CR*-long: position 1969 – 2739).

The phylogenetic Bayesian Inference (BI) / Maximum Likelihood (ML) tree reconstruction based on the concatenated mt alignment resulted in trees presenting the same topology with five distinct and geographically separated clades corresponding to geographic regions (Fig. 2). Accordingly, we will refer to them as Central, East, North and West clade as well as a *gallagheri* clade as this nomenclature was already established in the study of Pichler *et al.* (2018). The clade *gallagheri* comprised samples of *G. gallagheri*

from Wadi Bani Khalid (WBK). The topology between the clades Central, East and *gallagheri* was highly supported whereas the relationship to both West and North clade remained unclear.

In the *CR*-long alignment, 143 haplotypes were found with a max. genetic distance of ~ 13%. The phylogenetic Neighbor Joining (NJ) tree reconstruction resulted in the same clades as the BI/ML tree (Fig. S1). Only the split between Central and *gallagheri* clade is highly supported, the topology of the residual clades remained unclear.

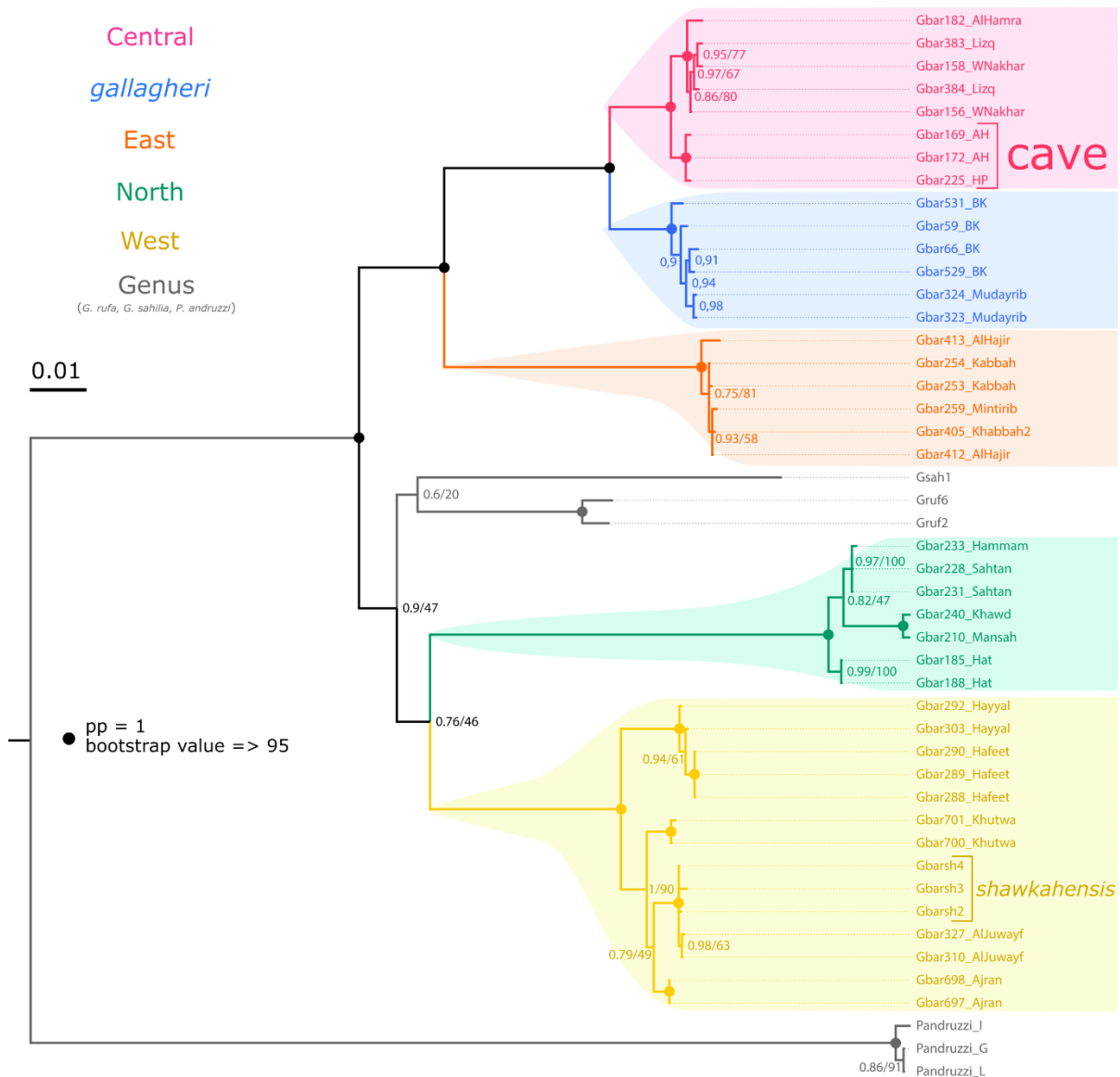


Figure 2 Phylogenetic BI tree based on the concatenated alignment (*COI* + *Cytb* + *CR*-long) of 44 individuals of the genus *Garra* and three individuals of *P. andruzzii* (outgroup). Posterior probabilities (pp) and bootstrap values from the ML analysis are indicated at the nodes. Nodes which are highly supported by both analyses (pp = 1, bootstrap values >= 95) are marked with a filled dot. Individuals with "cave genotype" as well as sequences originated from the subspecies *G. b. shawkahensis* are indicated on the right side.

The genetic variation was quite high in all clades, and within group distances were highest in the West clade (mean p-distances: *CR*: 2.3%, *Cytb*: 0.8%) and North clade (*COI*: 0.7%) whereas the lowest were detected

in the East and *gallagheri* clade (*CR*: 0.3%, *Cytb*: 0.2%, *COI*: 0.1%). Concerning haplotype diversity (*Hd*), the highest values were found within the West clade (*CR*: 0.97, *COI*: 0.87) and the *gallagheri* clade (*Cytb*: 1.0), nucleotide diversity (π) was highest in the West clade (*CR*: 0.023, *Cytb*: 0.008) and the North clade (*COI*: 0.007). Again, lowest values of *Hd* were detected in the East clade (*CR* 0.734, *Cytb*: 0.6) and the *gallagheri* clade (*COI*: 0.6), *gallagheri* was also the group that exhibits lowest π values for all three fragments (*CR*: 0.008, *COI*: 0.001, *Cytb*: 0.002) (Table 2, Additional File Table S2).

Table 2 General sequence information, mean and maximum *p*-distances, nucleotide and haplotype diversity of *G. barreimiae*.

Group	All	Central	<i>gallagheri</i>	East	North	West
N sequence	309	177	12	35	21	54
N haplotype	132	55	7	14	15	34
<i>Hd</i>	0.968	0.916	0.879	0.734	0.962	0.973
π	0.070	0.012	0.008	0.003	0.012	0.023
mean distance	7.0	1.2	0.8	0.3	1.2	2.3
max. distance	12.7	3.3	1.6	1.4	2.4	4.6

Nucleotide (π) and haplotype diversity (*Hd*) as well as mean and max. distances within mt clades of *G. barreimiae* (*CR*-long).

When comparing the genetic groups to each other, highest between group differences were detected between North/Central (*CR*: 11.4%, *COI*: 8.5%) and North/*gallagheri* (*Cytb*: 8.5%) whereas lowest between group differences were found between Central/*gallagheri* in all three fragments (*CR*: 4%; *COI*: 2.0%; *Cytb*: 1.9%) (Table 3).

Table 3 Mean *p*-distances between clades based on three mt fragments (*CR*-long, *COI* and *Cytb*).

	<i>CR</i> (N = 309)				<i>COI</i> (N = 42)				<i>Cytb</i> (N = 42)			
	Central	<i>gallagheri</i>	North	East	Central	<i>gallagheri</i>	North	East	Central	<i>gallagheri</i>	North	East
<i>gallagheri</i>	0.040				0.020				0.019			
North	0.114	0.112			0.085	0.082			0.084	0.085		
East	0.102	0.095	0.108		0.056	0.059	0.080		0.055	0.055	0.080	
West	0.102	0.100	0.104	0.106	0.071	0.067	0.084	0.066	0.078	0.074	0.076	0.078

All Median Joining networks of the mt clades are depicted in Fig. 3 and Fig. 4. The network of the Central clade reflects no clear pattern between the geographic origin of the samples. There is a “cave haplogroup” containing all samples from Al Hoota cave, all but two (Gbar504 and Gbar505) from the Hoti Pit cave, as well as most specimens with intermediate phenotype from WF and from GT (some are in HT 1 and 9).

The *gallagheri* network reveals no geographic structured pattern at all. The samples that originated from the second site Mudayrib all share the same haplotype. This haplotype is also found in samples retrieved from Wadi Bani Khalid (HT 1).

Depicted in the East clade network is one central haplotype (HT 1) comprising 60% of the samples and containing samples from all sampling sites except Arbeieen, Mintirib and Mayh. The remaining 40% are divided into nine haplotypes which all but one (HT 3) comprise samples from only one sampling site.

The network of the North clade is more structured and comprises 19 haplotypes. It can roughly be divided into two geographically separated parts: The first comprising a central (HT 2) and nine associated starlike arranged haplotypes (HT 7, 8, and 10 to 16) containing exclusively all samples from the northeastern located sites Khawd, Mahal, Manal, Samail and Mansah. The site Mayh is even further east but is associated with the second part of the network, that comprises a major (HT 1) and several minor haplotypes containing samples from more southwestern located sites.

From the network of the West clade it can be seen this clade is significantly more differentiated compared to the other clades. It comprises 25 haplotypes and shows also some geographical structuring. All samples from the most northwards located sites (Wadi Shawkah, AlJuwayf and Zihaymi) are clustered exclusively in five haplotypes (HT 2, 11, 23 to 25). At least seven mutational steps are between this “northern cluster” and the next closest haplotypes (HT 7 or 19). Another 14 mutational steps separate a geographically rather broad haplogroup comprising all and exclusively samples from more southwards lying sites (HT 1, 3 to 6, 8, 12 to 17).

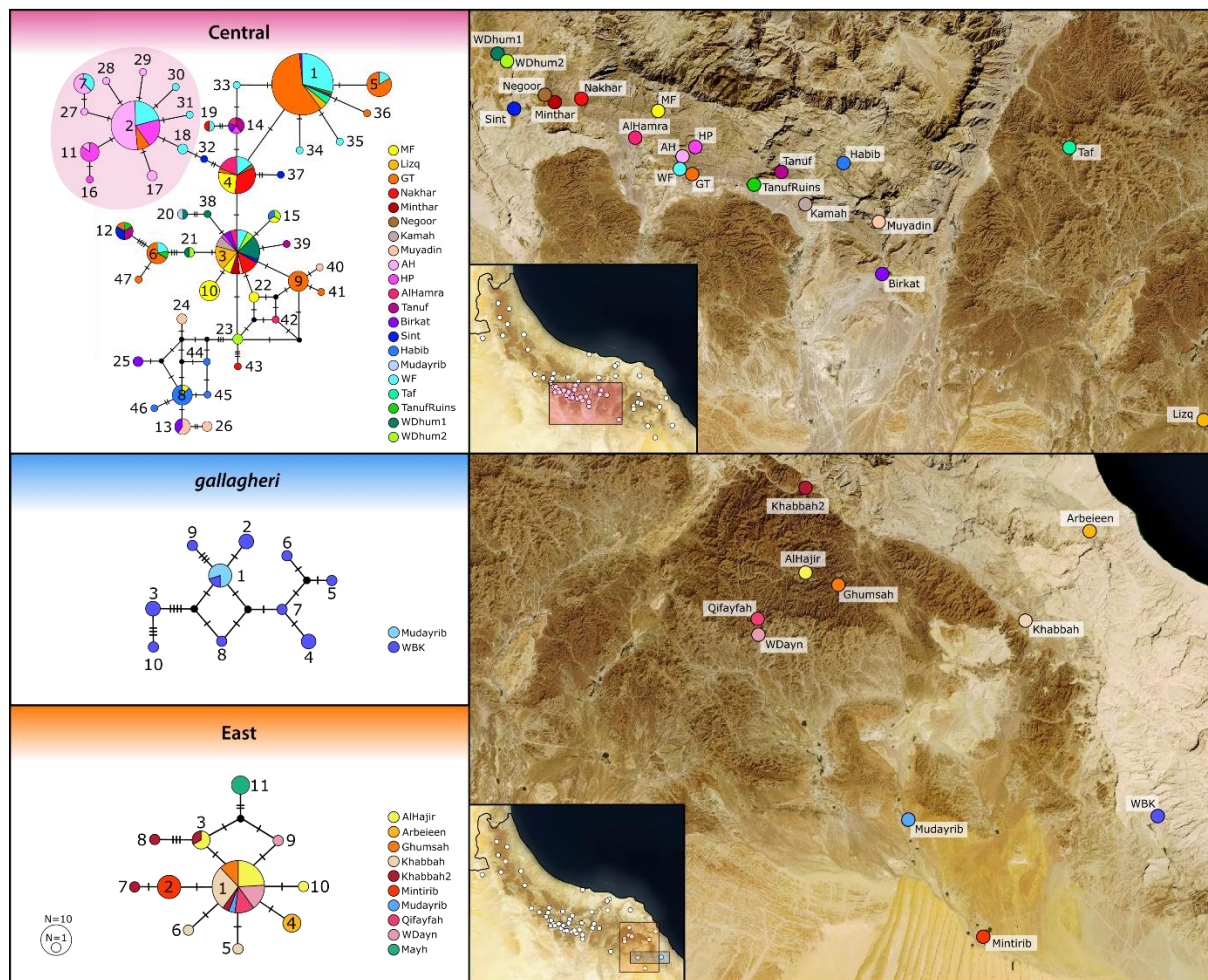


Figure 3 Median-Joining networks based on CR haplotypes of mt clades (Central = pink, gallagheri = blue, East = orange). Each circle represents a haplotype, size of circles corresponds to the number of individuals sharing the haplotype. Colours indicate sampling sites. Every crossbeam on the connecting lines between haplotypes represents a substitution. Small black dots between haplotypes symbolize hypothetical haplotypes not present in the data set. Distribution area of each mt clade (small map) as well as detailed geographic origin of the samples are depicted in the maps to the right.

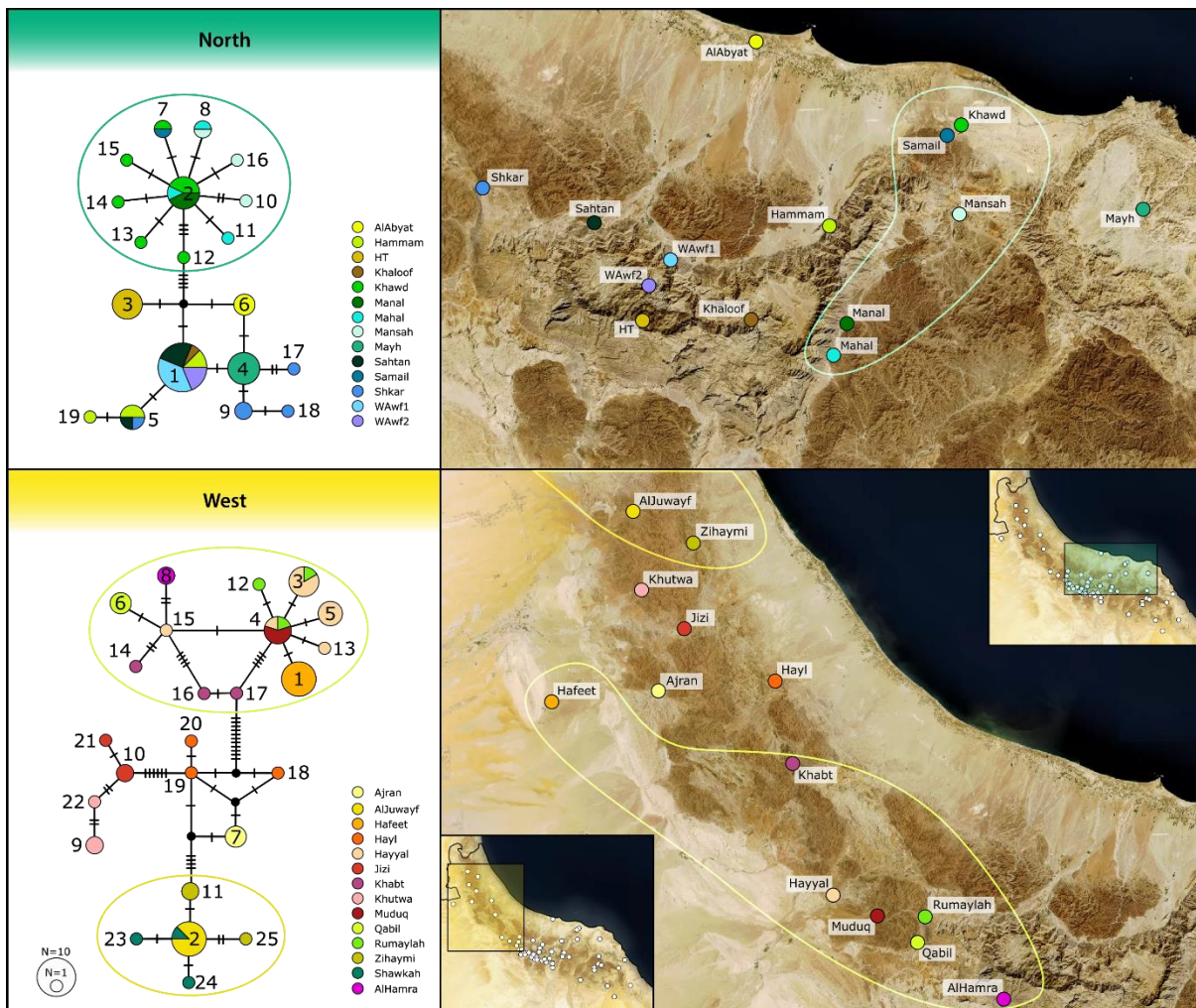


Figure 4 Median-Joining networks based on CR haplotypes of mt clades (North = green, West = yellow). Each circle represents a haplotype, size of circles corresponds to the number of individuals sharing the haplotype. Colours indicate sampling sites. Every crossbeam on the connecting lines between haplotypes represents a substitution. Small black dots between haplotypes symbolize hypothetical haplotypes not present in the data set. Distribution area of each mt clade (small map) as well as detailed geographic origin of the samples are depicted in the maps to the right.

Microsatellite variation (population structure analysis)

Genetic differentiation between groups and potential gene flow

The nuclear markers were used to test whether the same genetic groups would be detected via microsatellite data as with mitochondrial DNA sequence data. Given that the clusters of the initial STRUCTURE analysis resulted in similar groupings as the mt data analysis, the subsets of the microsatellite data were prepared accordingly. Since not all of the 19 microsatellite loci (Kirchner *et al.*, 2014) worked equally good in the different clades, subsets were used for the analysis of the genetic structure within the clades. Hence, the non-working loci were excluded leaving: 17 in subset Central, 18 in subset *gallagheri*, 18 in subset East, 13 in subset North and 12 in subset West. Only nine loci worked well across all subsets and were used for STRUCTURE analysis of the complete data set. The reviewed genotypes of all individuals analysed are listed in Supplementary Table S1.

After inspecting the results of the initial STRUCTURE analysis, it became evident that it was necessary to reduce the number of samples belonging to the Central subset (N = 331) since the number of samples comprised in the remaining subsets is far lower and unequal sampling can affect the outcome of the analysis (Sebastien, 2016). Hence, for analysis we chose sampling sites covering the whole distribution area where at least 10 specimens were available. Furthermore, also samples with full (Al Hoota Cave: 37, Hoti Pit: 17) or partial cave genotypes (Ghubrat Tanuf: 11, Wadi Fallahi: 14) from the Central subset were excluded, because the differentiation between cave and surface populations produces a stronger signal than any other population structure within the subset Central. Besides, these samples have already been examined in detail by Kirchner et al. (2017). This left us with 64 samples from the subset Central that were included in the STRUCTURE analysis of the complete data set and with 255 samples for the analysis of the Central subset only. The other subsets comprised 29 (*gallagheri*), 67 (East), 62 (North) and 35 (West) samples.

The reduced data set yielded the highest probability for K = 5. These five clusters correspond to the mt clades. The analyses was repeated with these five subsets corresponding to these clusters, which resulted in K = 2 for subset Central (N = 255, without cave/intermediate individuals), K = 1 for *gallagheri* (N = 29), K = 2 for East (N = 67), K = 3 for North (N = 62) and K = 3 for West (N = 35) as depicted in Fig. 5. A map illustrating the distribution of the clusters from each clade is shown in Fig. 6. Only few sites showed ambiguity of belonging (< 90% affiliation to one cluster): two sites from Central clade (Sint, WF) and one from East clade (Dayn).

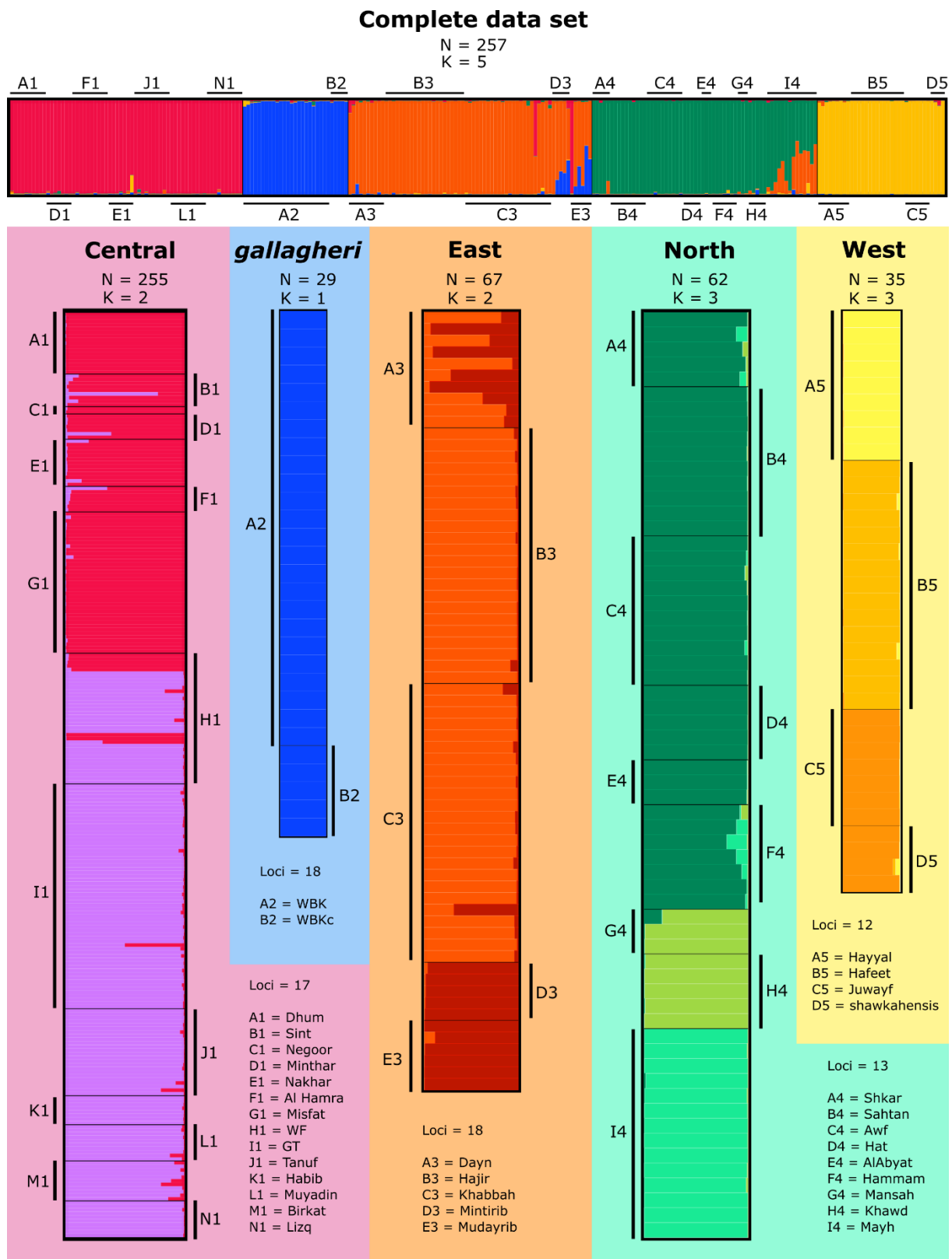


Figure 5 Population structure analyses based on the complete data set (top) and the subsets (bottom). For the complete data set, the most probable number of clusters was five. Concerning the subsets, STRUCTURE analyses resulted in K = 2 in Central (magenta), K = 1 in gallagheri (blue), K = 3 in North (green) and K = 3 in West clade (yellow).

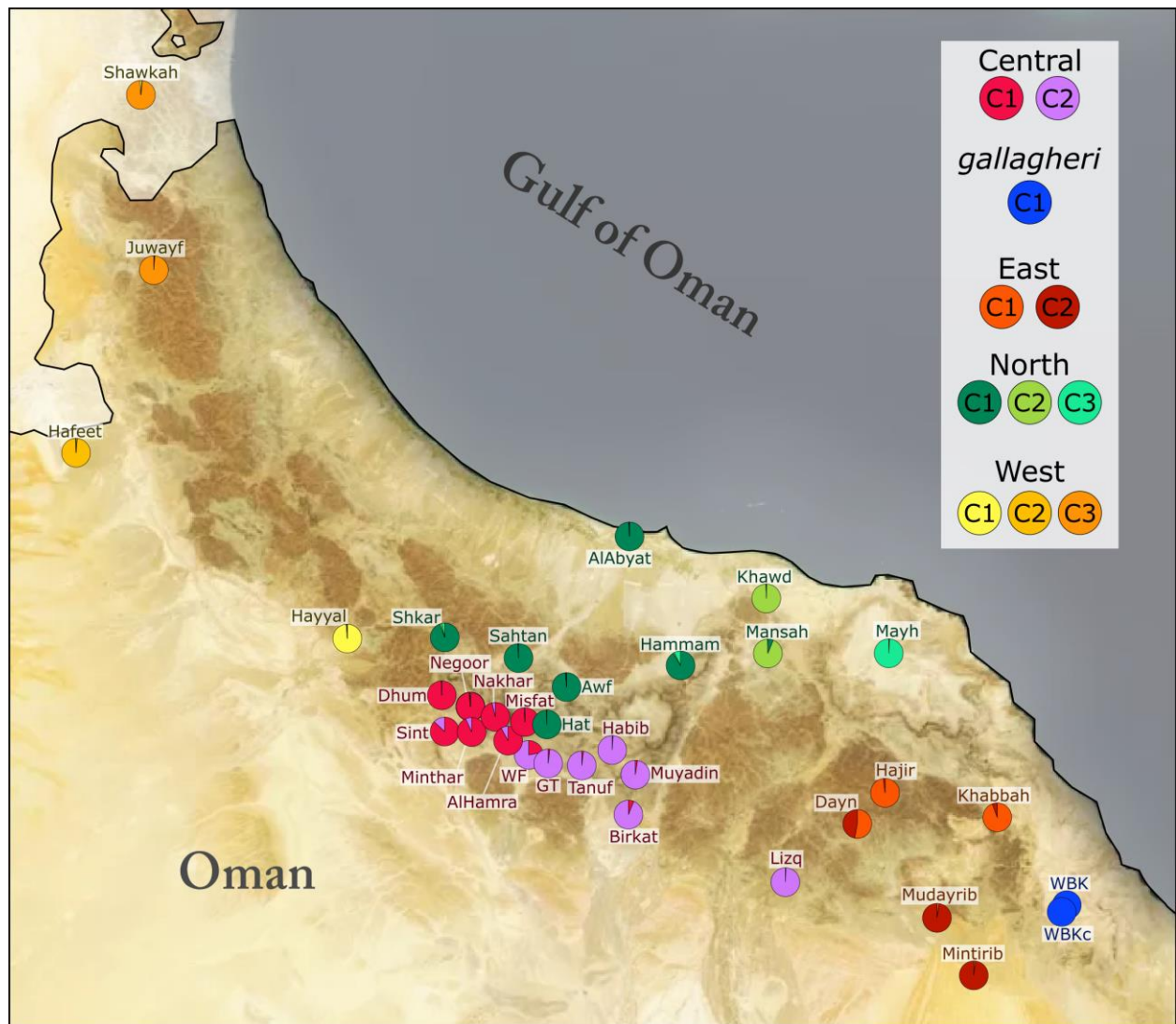


Figure 6 Geographic map of study area in Northern Oman. Sampling sites used for microsatellite analyses are marked via pie chart. The different colours of the pie chart fractions resemble the portion of cluster affiliation from this sampling site in each clade.

There are a few sites where STRUCTURE results indicated some degree of gene flow between the clusters. The samples from Mintirib and Mudayrib, which belong to the mt East clade show mixed genotypes between the clusters East and *gallagheri*. In addition, one individual from Mudayrib belongs 100% to the Central cluster according to nc and mt data. Individuals which originated from Mayh of the mt North clade, show high degree of mixed signals from both North as well as East group (Fig. 5, K = 5).

Five clusters were identified with the 'find.clusters' function of the discriminant analysis of principal components (DAPC) based on the complete data set. Only a few individuals were grouped ambiguously when comparing the DAPC clusters with the mt clades (Fig. S2).

Genetic variability within groups/clusters

In order to assess genetic variation among and between populations, we calculated the descriptive statistics (Table 4) for the nine polymorphic unlinked microsatellite loci that worked well across the complete data set. In general, variation was high with 32 to 142 alleles per cluster across all loci. These results were corrected by taking the sample size into account and calculating the allelic richness according to the group with the fewest individuals. The highest allelic richness was found within the cluster Central_C2 (8.35) whereas the lowest was detected in West_C2 (2.25). Furthermore, the calculated mean H_e over all loci was highest in North_C3 (0.69) and lowest in West_C3 (0.24). Overall, the expected heterozygosity (H_e) was higher than the observed heterozygosity (H_o) in all clusters but West_C1 and West_C2. Most private alleles were detected in West_C3 (16.67%) and fewest in West_C1 (4.76%).

Null alleles were detected at least at one locus in most clusters. However, since none of these loci showed a constant significant deviation in all clusters, we did not exclude these loci from our analyses. Deviations from Hardy-Weinberg equilibrium (5% level) were detected in at least one locus in almost all clusters but West_C1 and West_C2, whereas Central_C1 had most loci with significant deviations. However, the reason for deviations from Hardy-Weinberg equilibrium (HWE) in the majority of these loci is most likely due to the presence of null alleles. The results from the exact test indicate that all clusters are in HW equilibrium.

The inbreeding coefficient F_{IS} is positive in all populations except the West_C1 and West_C2, indicating at least some degree of heterozygote deficiency especially in the cluster East_C2 (0.27). The pairwise F_{ST} values between pairs of clusters of each subset (Table 5) revealed highest values between West_C2 and West_C3 (0.36), whereas lowest F_{ST} values were found between the two clusters of the Central clade (0.053).

Table 4 Population genetic parameters of the five subsets and their corresponding clusters.

Central - 17 Loci													
Pop	N	Na	A	Ar	pA	pA %	H_o	H_e	F_{IS}	p	Ld	Lp%	HWE
All	255	173	10.18	-	24	13.87	0.51	0.63	0.18	0.08	JMLR	100.00	0.00
C1	101	130	7.65	7.07	9	6.92	0.49	0.59	0.17	0.09	PH8A, DL6X, JMLR	100.00	0.00
C2	154	142	8.76	8.35	15	10.56	0.53	0.60	0.11	0.19	-	100.00	0.00
Gallagheri - 18 Loci													
Pop	N	Na	A	Ar	pA	pA %	H_o	H_e	F_{IS}	p	Ld	Lp%	HWE
C1	29	130	7.22	-	15	11.54	0.52	0.60	0.13	0.25	JQSO, 3Z7I, 3N43, I6G2	100.00	0.00
East - 18 Loci													
Pop	N	Na	A	Ar	pA	pA %	H_o	H_e	F_{IS}	p	Ld	Lp%	HWE
All	67	171	9.50	-	11	6.43	0.51	0.61	0.16	0.02	UHPE, HOLN, 9XNC, 3N43, I6G2, 3ROZ	100.00	0.00
C1	51	138	7.67	5.18	11	7.97	0.52	0.55	0.07	0.34	JMLR, 9XNC, I6G2	94.44	0.00
C2	16	116	6.44	5.47	0	-	0.50	0.65	0.27	0.13	3Z7I, 3ROZ	100.00	0.00
North - 13 Loci													
Pop	N	Na	A	Ar	pA	pA %	H_o	H_e	F_{IS}	p	Ld	Lp%	HWE
All	62	147	11.31	-	28	19.05	0.49	0.68	0.28	0.01	UHPE, 1EHE, CJHG, JMLR, 3N43, 2PUM, CB75	100.00	0.00
C1	40	94	7.23	4.07	13	13.83	0.45	0.54	0.18	0.20	UHPE, CB75	92.31	0.00
C2	8	60	4.62	3.78	9	15.00	0.48	0.53	0.15	0.44	-	84.62	0.00
C3	14	72	5.54	4.41	6	8.33	0.62	0.69	0.14	0.35	JQSO, JMLR	100.00	0.00

West - 12 Loci													
Pop	N	Na	A	Ar	pA	pA %	H _o	H _e	F _{IS}	p		Ld	
All	35	69	5.75	-	13	18.84	0.32	0.49	0.35	0.19	JQSO		
C1	9	42	3.50	3.07	2	4.76	0.48	0.43	-	0.69	-		
C2	15	32	2.67	2.25	5	15.63	0.32	0.31	-	0.64	-		
C3	11	36	3.00	2.38	6	16.67	0.20	0.24	0.22	0.60	HOLN		
												Lp%	HWE
												91.67	0.00
												75.00	0.62
												66.67	0.61
												75.00	0.00

Pop Population, N Number of analysed individuals, Na number of alleles per population, A mean number of alleles per locus, Ar allelic richness, pA number of private alleles, pA% percentage of private alleles per population based on the actual amount of alleles per population, H_o mean observed heterozygosity over all loci, H_e mean expected heterozygosity over all loci, F_{IS} inbreeding coefficient, p (mean) p-value for Hardy-Weinberg equilibrium (exact test, mean over all loci), Ld Loci, that deviated from HW-equilibrium and do not exhibit potential microsatellite null alleles, Lp% percentage of polymorphic loci, HWE p-value for Hardy Weinberg equilibria (global test).

Table 5 Pairwise F_{ST} values between the clusters of *G. barreimiae*.

Cluster 1	F_{ST}	Cluster 2
Central_C2	0.053	Central_C1
East_C2	0.069	East_C1
North_C2	0.131	North_C1
North_C3	0.176	North_C1
North_C3	0.202	North_C2
West_C2	0.167	West_C1
West_C3	0.274	West_C1
West_C3	0.366	West_C2

Pairwise F_{ST} values between the clusters of *G. barreimiae* based on 17 (Central), 18 (East), 13 (North), 12 (West) microsatellite loci. Calculated with GenAlEx, all values are highly significant (***).

Geographic distribution of the clades

With more than 50 sampling sites spread throughout Northern Oman, one can delimit distribution areas of each clade of *G. barreimiae* quite precisely (Fig. S3). The gathered data gives detailed information on the distribution margins and thereby enables to identify potential barriers or paths to gene flow.

The Central clade with its distribution area amidst the Hajar Mountains on the Saiq Plateau was the most extensively sampled group and shows moderate degree of variation. The microsatellite data reveals a clear differentiation between a western (Central_C1) and eastern (Central_C2) cluster, a pattern that could be explained by the change of direction of water drainage, one arm of the wadi close to the city Al Hamra leading southwards, another extending eastwards comprising all sites belonging to the eastern cluster (Central_C2). The distribution range of the Central clade is clearly restricted by geographic boundaries. All (but two) sampling sites lie on the southern slope of the Al Jebel al-Akhdar mountain range, which reaches altitudes up to 3000 m asl (highest peak in Arabia: Jebel Shams) (Kusky, Robinson, & El-Baz, 2005). This mountain ridge limits the distribution of the Central clade to the North. To the east lies the Samail Gap, a valley stretching in north-south direction from the city Bidbid to Izki, that separates the eastern part of the Hajar mountains (Al Hajar Al Sharqi). South to the Central clade distribution the sand desert Rub al Khali

("empty quarter") expands, which acts as a natural boundary for freshwater fish migration. The western foothill of Jebel Shams represents the westernmost occurrence of the Central clade.

The sister group of the Central clade is the *gallagheri* clade which consists of all samples from the easternmost site Wadi Bani Khalid and most samples from nearby located Mudayrib. Since the geographic distribution of this clade is the most restricted, the genetic variation within this group is also rather low.

The East clade is genetically the most homogenous clade with lowest mt distances and very little substructure. This can be explained by the fact that the distribution covers only the eastern part of the Hajar mountains Al Hajar Al Sharqi and appendant geological structures which is not disrupted by natural barriers like deep valleys or high mountain ridges. Besides the *gallagheri* clade, the distribution of the East clade is geographically the most isolated one seemingly not connected to any other clade. It is likely that aquatic passageways between sampling sites of the East clade are more permanent and thereby, creating contact zones between populations and hence, preventing diversification within the clade. Concerning nc data, we found two clusters that roughly follow a north (East_C1) to south (East_C2) pattern but the presence of mixed genotypes at the site Dayn in the middle of the distribution area as well as very low F_{ST} values (Table 5) between these clusters indicate little genetic differentiation (Wright, 1978).

The distribution of the North clade stretches across the northern part of the mountains and reaches the coastal area at the Gulf of Oman. This clade exhibits high genetic variation and can be divided in at least two subgroups. Again, the Samail Gap seems to form a barrier dividing a western cluster (North_C1) from an eastern cluster (North_C2). In addition, the easternmost site Mayh represents a third cluster based on nc data (North_C3) but shows no such differentiation based on mt data. Since the distribution of the North clade extends from high mountain peaks to coastal lowland areas of the Al Batinah plain, it comprises high habitat diversity including many extensive wadis that channel water directly towards the Gulf of Oman. These geological structures could be accountable for the high genetic structure, preventing contact between populations of different wadis even during high water level periods.

The West clade is distributed along the western ridge of the Hajar mountains and reaches northwards at least until Wadi Shawkah in the UAE. This clade shows highest within group differences, which are well reflected in the phylogenetic tree and network reconstructions (Fig. 2, 3, S1). Two main subclades were found, one northern exhibiting high degree of substructure and including samples from the subspecies *G. b. shawkahensis* (also represented by nc cluster West_C3), and a second comprising all southern samples (reflected by two nc cluster West_C1 and West_C2). In this case, the distribution limits of the subclades run along the Wadi Hibi, which leads from the mountains in northeastern direction towards the coast. The highest F_{ST} values were detected within the West clade indicating more genetic differentiation than in any other clade (Table S3). The Al Batinah plain also overlaps with the northeastern distribution of the West

clade. Therefore, we assume that these geological barriers caused the high degree of genetic differentiation within the West clade.

DISCUSSION

Phylogeography

The most striking result of this study was the high genetic variation within *Garra barreimiae* found throughout the rather small geographic range around the Hajar mountains. Using a comprehensive dataset of 632 individuals and employing both mt sequences as well as microsatellite data, five genetically distinct groups were detected, which were named due to their distinct geographic distribution North, East, West, Central and, according to the taxon name, *gallagheri* clade. We found strong correlations between genetic and geomorphological structures (see chapter “Geographic Distribution of the Clades” in Results).

Within the Central clade, the mt tree presents a “cave subclade” comprising only troglomorphic fish originated from Al Hoota cave and Hoti Pit as well as hybrid individuals between cave and nearby surface sites (Kirchner *et al.*, 2017). This supports the findings of the previous study by Kruckenhauser *et al.* (2011). Originally, the species *G. gallagheri* (former subspecies of *G. barreimiae*) was described only from the site Wadi Bani Khalid and thereby represents the geographically most restricted clade. However, members of this taxon have been discovered the first time also at another site (Mudayrib), together with individuals from the East clade. The East clade is genetically the most homogenous clade. In contrast, both the North clade and the West clade exhibit a much higher genetic variation in respect of mt as well as nc data. Pichler *et al.* (2018) investigated the morphological variation among the clades of *G. barreimiae* and found that the West clade is also the most diverging clade. It would be highly interesting to study whether the genetic substructure of this clade is reflected also in morphological differences (this clade contains individuals sampled closest to the *locus typicus* of the nominate form as well as from the *locus typicus* of *G. b. shawkahensis*). For this purpose, it would be necessary to measure also individuals assigned to *G. b. shawkahensis* from the Wadi Shawkah in the UAE. Unfortunately, only tissue samples from these specimens were available for this study.

Although most sampling sites of each cluster are not connected by any perennial waterbody, the microsatellite analysis shows no genetic differentiation between single sampling sites. Therefore, we assume that there must be some recent gene flow, either due to underground water connections in the karst area, as it has been reported for other cyprinids (Palandačić *et al.*, 2012) or because of sporadic heavy rainfalls, which quickly fill wadis with water and thereby generate temporal surface water connections.

Hybridization

Although the nuclear and mitochondrial data give concordant results, following the geographic distribution, there are some exceptions: a few individuals possess mixed genotypes from two clades indicating recent hybridization. In the present study, indication for hybridization between the clades was only detected at three out of 37 sites that were analysed with nc microsatellite data. This affects the closely related clades East and *gallagheri* but also the more distant clades East and North. Most samples from the site Mudayrib cluster with the *gallagheri* clade in the mt tree but exhibit a mixed nc genotype between the clades East and *gallagheri*. A similar pattern is found in the close-by site Mintirib where all samples belong to the East clade (mt) but show mixed genotype between the clades East and *gallagheri* (nc). Both sites are located on the southeastern margin of the distribution area, southbound their distribution is limited in general by the sand desert Ramlat al-Wahiba. In this case, when the water-level is rising due to rainfall, it seems plausible that also distribution expands along temporary water routes and thereby enables sporadic contact zones between geographically separated taxa. The third site where hybrids were detected is Mayh. Based on mt data, half of the individuals originating from Mayh belong to the North clade whereas the other half belongs to the East clade. Most individuals of this site exhibit mixed nc genotypes irrespective of their mt clade affiliation.

Interspecific hybridisation is common in cyprinids. Genetic evidence is not that frequent, but there are some striking examples, like hybrids between taxa of the European *Phoxinus* species complex (Palandačić *et al.*, 2017), sometimes also over multiple generations like between bighead and silver carp (Lamer *et al.*, 2015) or even between different genera like roach and bream in Ireland (Hayden *et al.*, 2010) or goldfish and introduced common carp in Australia (Haynes *et al.*, 2012).

Overall, it can be stated that hybridization between the clades of *G. barreimiae* and between *G. barreimiae* and *G. gallagheri* is not a common event and does not lead to blurring of the genetic patterns. It seems plausible that the sites in which hybridization occurs are zones of secondary contact between the clades. This might be due to anthropogenic influence. At least the samples of Mudayrib and Mintirib were collected from aflaj, which are manmade irrigation canal systems transporting water over kilometres from mountain springs to the cultivated gardens (personal communication with a local from Mintirib). Introductions either deliberately or as a consequence of channelling water via falaj might play a role.

Another indication that hybridization occurs only very rarely is that even though three out of the five clades occur in close vicinity around the mountain peak of Jebel Shams (West, North, Central), not a single individual with mixed genotype is detected via microsatellite analyses from this area. There are two individuals from the site Al Hamra that belong to the West clade based on mt data but are grouped perfectly with cluster C1 of the Central clade, to which all other individuals from Al Hamra belong.

Relationship between clades and possible origin and age of clades

Although we found clearly distinct clades within our data set, the phylogenetic relationships between these clades as well as to other closely related species of the genus are still not completely resolved and therefore, we cannot evaluate whether the five investigated clades are monophyletic with respect to *G. rufa*.

Because of the unresolved phylogeny of the clades, it cannot be concluded if the Arabian Peninsula was colonized in one or several independent migration events. It is difficult to deduce when and where the first migration of *Garra* into the Arabian Peninsula occurred, but it is quite likely that migration happened along the Persian Gulf during a glacial maximum when the sea level was lower than today. The flooding of Persian Gulf Basin at the end of the last glacial maximum (LGM; 20 kya) marks the latest possible time of migration along this route. With a maximum depth of 110 m the Persian Gulf was exposed for ensuing 70 kya until sea level rose after the end of the last Glaciation (Clark *et al.*, 2009). During that time the climate was subtropical, and the Persian Gulf formed the fertile Ur-Schatt River Valley that was recharged by the rivers Euphrat, Tigris, Karun and Wadi Batin meaning the majority of freshwater of Southwest Asia drained into the Ur-Schatt River Valley (Rose, 2010). Considering the distances found between clades it is unlikely that this amount of differentiation arose after a migration onto the Arabian Peninsula during the LGM. The question whether several invasions occurred over a longer period of time remains open. To further address this and other questions, comparative genome-wide analyses of a broader data set including more closely related species from this region, specifically from Iran as well as southern Oman would be necessary.

In order to get insight into the phylogenetic history of *Garra* in Northern Oman it would be of high interest to get an estimate when the splits happened. Yet, a molecular clock needs to be calibrated with an independent source of information such as fossils. Unfortunately, despite the wide distribution of extant *Garra* spp. in Africa and Asia, no fossil records were found so far from this region. The only known fossil evidence of *Garra* was retrieved from Armenia and dated back to the Pliocene (Vasilyan & Carnevale, 2013). However, as a first approximation the substitution rates from the literature for the *Cytb* gene were used to estimate the approximate age of the clades. Two different substitution rates have been proposed concerning cyprinids: 0.53% per million years (Boore, 1999) and 0.76% per million years (Dowton *et al.*, 2009). According to our estimates based on these rates, three time intervals of diversification were identified: the first interval reflects the split between the Central and *gallagheri* clade, which corresponds to an estimate of 1.3 (respectively 1.8) Myr and represents the youngest split. The second interval was dated to 3.6 (5.2) Myr and comprises the split between East clade and Central/*gallagheri* clades. The splits between both the West and North clade and all other clades represent the third and oldest interval, estimates reach from 4.9 up to 5.6 (7.2 to 8.0) Myr. Accordingly, the age of the cave population in comparison to the genetically closest surface population (Nakhar) must have occurred not earlier than the Middle Pleistocene (0.8 to 0.1 Myr). In those days, the climate in Oman was characterized by gradually

increasing aridity interrupted by several pluvial periods caused by spikes in Indian Ocean monsoon intensity (Burns *et al.*, 1998; Blechschmidt *et al.*, 2009). It is plausible that these climate changes triggered the colonialization of the Al Hoota cave, either to retreat to the few remaining water reservoirs when temperature was rising, and surface water level dropped, or to occupy new freshwater habitats established in the course of heavy rain periods.

Species delimitation

In the past few years, several new species of the genus *Garra* from the Middle East have been described: *G. jordanica* (Hamidan, Geiger, & Freyhof, 2014) from the Dead Sea Basin, *G. mondica* (Sayyadzadeh, Esmaeili, & Freyhof, 2015) and *G. amirhosseini* (Esmaeili *et al.*, 2016) from Iran, *G. sindhi* (Lyon *et al.*, 2016) and *G. smarti* (Krupp & Budd, 2009) from southern Oman. Besides, also new cave species have been described, as *G. lorestanensis* (Mousavi-Sabet & Eagderi, 2016) and *G. tashanensis* (Mousavi-Sabet *et al.*, 2016) from Iran. Obviously, the diversity within this genus had been underestimated in the past and needs thorough revision. Since this genus comprises over 130 described species (Froese & Pauly, 2018) and has a vast Afro-Asian distribution, a proper revision based on an integrative approach is unfeasible to conduct at once. Thus, this study aims to take one of many steps necessary to elucidate the unfathomed diversity this taxon holds.

A study by Pichler *et al.* (2018) stated, that the clades of *G. barreimiae* are morphologically quite variable and not to distinguish on the basis of single characters. The West clade can be discriminated by a few characters like number of gill rakers, precaudal and intermediate vertebrae, and branched pelvic fin rays. But in general, the variation of characters of all clades overlap considerably. Hence, the genetically well separated clades of *G. barreimiae* cannot be distinguished morphologically based solely on single traits that are commonly accepted and prevalently used to classify Cypriniformes.

The broad overlap and lack of distinct characters to discriminate the clades of *G. barreimiae* morphologically impede a straightforward species delimitation. Eventually, a decision to classify the four newly detected clades as separate species had to be made. Subsequently, there is a need to state our applied concept of species. De Queiroz (2005, 2007) presented the unified species concept that proposes that the existence of separately evolving metapopulation lineages is the common element reconciling alternative concepts of species. Yet, this concept still embraces the various properties that have been considered defining in alternate species concepts and sees them as evidence for separate evolving lineages. In the case of *G. barreimiae*, taxa are morphologically similar and delimiting species boundaries based solely on morphological characters is futile. Molecular genetic methods on the other hand reveal well differentiated and geographically separated clades. In this case, the diverging mt and nc marker represent the lines of evidence relevant for the phylogenetic species concept.

Since the use of molecular genetic methods in biodiversity research, specifically since DNA barcoding intensified, the term cryptic or pseudocryptic species has been used increasingly in the past decade. Yet, there are no commonly agreed criteria to fulfil the status of a cryptic species. Hence, the term has been used inconsistently depending on which organism group was investigated and on previous knowledge about the taxon (Fišer, Robinson, & Malard, 2018). Fišer et al. (2018) identified three mechanisms causing cryptic diversity: recent divergence, phylogenetic niche conservatism (morphological stasis) and morphological convergence. Since our data confirm close relationship between the *G. barreimiae* clades, one can exclude convergence as an explanation. Considering the genetic distances and age estimations, it becomes evident that the clades are comparably old enough for morphological changes to take place. It seems more likely that extreme environmental conditions caused stabilizing selection on morphology, as suggested by Bickford et al. (2007). They stated that evolving under severe environmental extremes could limit changes in morphology. The environmental fluctuations in Northern Oman affected all clades of *G. barreimiae* similarly, thus it might have prevented morphological diversification. This can also be confirmed for water quality parameters (temperature, conductivity, pH) as well as habitat structures, which are strongly varying between different sampling sites according to our field observations. Following the suggestions of Bickford et al. (2007) who consider species to be cryptic if they have been classified as a single nominal species and cannot be distinguished by morphology, we regard the term cryptic species as justified in this case at least for the North, Central and East clades.

However, the rising number on studies revealing cryptic diversity highlight also the problem of underestimating the actual (so far overseen) biodiversity, which has a major effect on conservation efforts. A taxonomic revision including the description of the newly discovered cryptic species of *G. barreimiae* complex from Northern Oman is already in progress and will hopefully lead to a re-evaluation of the IUCN Red List conservation status, according to which *G. barreimiae* is currently assigned to “Least Concern”.

CONCLUSIONS

In the course of this study the genetic variability and phylogeographic pattern within the genus *Garra* from Northern Oman was investigated. We applied an integrative approach incorporating all available data gathered from several independent sources (mt sequences and nc microsatellite markers, geographic and morphological information). The results clearly show high genetic variation with genetic distances comparable to distances between other species of this genus. In conjunction with the fact that these genetic groups can not be discriminated morphologically yet, we address them as cryptic species in the sense of Bickford et al. (2007). The taxon *G. barreimiae* is currently listed as “Least Concern” according to the IUCN Red List. Hence, the newly discovered *Garra* species complex in Northern Oman is in urgent need for re-evaluation, as desertification, habitat destruction as well as water shortage and potential chemical

pollution pose serious threats to organisms occupying extreme habitats. Contemporaneously, the species descriptions of these newly discovered species are already in preparation and will provide the basis for a conservation reassessment.

LIST OF ABBREVIATIONS

BI - bayesian inference
COI - cytochrome C oxidase subunit I
CR - control region
Cytb - cytochrome b
DAPC - discriminant analysis of principal components
 F_{is} - inbreeding coefficient
Hd - haplotype diversity
 H_e - expected heterozygosity
 H_o - observed heterozygosity
HT - haplotype
HWE - Hardy Weinberg equilibrium
K - cluster
kya - thousand years
Ld - loci, that deviated from HW
LGM - last glacial maximum
Lp% - percentage of polymorphic loci
MJ - median joining
ML - maximum likelihood
mt - mitochondrial
Myr - million years
N - number of analysed individuals
Na - number of alleles per population
nc - nuclear
NJ – neighbour-joining
pA - number of private alleles
pa% - percentage of private alleles per population based on the actual number of alleles per population
PCR - polymerase chain reaction
Pop - population
pp - posterior probability
 π - nucleotide diversity

DECLARATIONS

Ethics approval and consent to participate

The study was conducted in accordance with local legislation and comply with national guidelines. Specimens used in this study were legally collected from wild stocks by the authors with permission from the Ministry of Environment and Climate Affairs of the Sultanate of Oman (Permit No.: 3/2017).

Availability of data and materials

All data generated or analysed during this study are included in this published article [and its supplementary information files (Table S1)]. DNA sequences are available under the Genbank accession numbers XXX-XXX (accession numbers will be added in proof).

Competing interests

The authors declare that they have no competing interest.

Funding

The study was funded by the University of Vienna (uni:docs fellowship for S. Kirchner), the Natural History Museum of Vienna (Central Research laboratories; materials and infrastructure), by the Ministry of Commerce and Industry of the Sultanate of Oman and by the Ministry of Tourism of the Sultanate of Oman (travel support). The Federal Ministry for Transport, Innovation and Technology Austria provided financial support in the course of a FEMtech practical (for S. Kirchner). Open access funding provided by University of Vienna. The funding body had neither an effect/impact on the design of the study nor on the collection, analysis and interpretation of the data.

Authors contributions

S.K. and L.K. planned the research, S.K. carried out the molecular genetic and statistical data analyses and drafted the manuscript. L.K. participated in the molecular genetic analyses and helped to draft the manuscript. H.S. initiated the study, coordinated the sampling and helped to draft the manuscript. E.H. helped with the data collection and to draft the manuscript. R.V. helped with the sampling and provided detailed information about geographic conditions. All authors read and approved the final manuscript.

Acknowledgement

The authors want to thank Johannes Rath and Christian Baranyi from the University of Vienna for providing help with the Sequence Analyzer, Harald Krenn (University of Vienna) for his general help and support, Arthur Pichler for his valuable input and overall help, Barbara Däubel (NHM Vienna) and Julia Schindelar (NHM Vienna) for technical assistance in the laboratory, Robert Illek (NHM Vienna) for support in practical field work, Lukas Plan (NHM Vienna) and Katharina Bürger (NHM Vienna) for their speleological and geographical expertise and for helping with collecting samples from the cave sites, Frank Zachos (NHM Vienna) and Stephan Koblmüller (University of Graz) for help and support with population genetic analyses, and Cristiano Bertolucci (University of Ferrara) and Silvia Fuselli (University of Ferrara) as well as Jörg

Freyhof (IGB Leibniz) and Matthias Geiger (ZFMK Bonn) for providing additional samples. We want to express our gratitude to the Ministry of Environment and Climate Affairs of Oman for their cooperation, support and the sampling permits and to Michael Robinson, a zoologist at the Sultan Qaboos University in Muscat, for his intellectual input.

REFERENCES

- Adams M, Raadik TA, BurrIDGE CP, Georges A. 2014. Global biodiversity assessment and hyper-cryptic species complexes: More than one species of elephant in the room? *Systematic Biology* 63: 518–533.
- Al-Shaqsi S. 2010. Care or Cry: Three years from Cyclone Gonu. What have we learnt? *Oman medical journal* 25: 162–167.
- Bandelt HJ, Forster P, Röhl a. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular biology and evolution* 16: 37–48.
- Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I. 2007. Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution* 22: 148–55.
- Blechs Schmidt I, Matter A, Preusser F, Rieke-Zapp D. 2009. Monsoon triggered formation of Quaternary alluvial megafans in the interior of Oman. *Geomorphology* 110: 128–139.
- Boore JL. 1999. Animal mitochondrial genomes. *Nucleic Acids Research* 27: 1767–1780.
- Burns SJ, Matter A, Frank N, Mangini A. 1998. Speleothem-based palaeoclimate record from northern Oman. *Geology* 26: 499–502.
- Calderoni L, Rota-Stabelli O, Frigato E, Panziera A, Kirchner S, Foulkes NS, Kruckenhauser L, Bertolucci C, Fuselli S. 2016. Relaxed selective constraints drove functional modifications in peripheral photoreception of the cavefish *P. andruzzii* and provide insight into the time of cave colonization. *Heredity* 117.
- Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl Tutorial. *Bioinformatics (Oxford, England)* 25: 1972–3.
- Clark PU, Dyke AS, Shakun JD, Carlson AE, Clark J, Wohlfarth B, Mitrovica JX, Hostetler SW, McCabe a M. 2009. The Last Glacial Maximum. *Science (New York, N.Y.)* 325: 710–4.
- Dowton M, Cameron SL, Dowavici JI, Austin AD, Whiting MF. 2009. Characterization of 67 mitochondrial tRNA gene rearrangements in the Hymenoptera suggests that mitochondrial tRNA gene position is selectively neutral. *Molecular Biology and Evolution* 26: 1607–17.
- Durand JD, Tsigenopoulos CS, Ünlü E, Berrebi P. 2002. Phylogeny and biogeography of the family cyprinidae in the Middle East inferred from cytochrome b DNA - Evolutionary significance of this region. *Molecular Phylogenetics and Evolution* 22: 91–100.
- Earl DA, VonHoldt BM. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359–361.
- Esmaili HR, Sayyadzadeh G, Coad BW, Eagderi S. 2016. Review of the genus *Garra* Hamilton , 1822 in Iran with description of a new species : a morpho- molecular approach (Teleostei : Cyprinidae). *Iranian Journal of Ichthyology* 3: 82–121.
- Fišer C, Robinson CT, Malard F. 2018. Cryptic species as a window into the paradigm shift of the species concept. *Molecular Ecology* 27: 613–635.
- Froese R, Pauly D. 2018. FishBase. www.fishbase.org.version (06/2018). World Wide Web Electronic Publication.
- Goudet J. 1995. Fstat Version 1.2.: a computer program to calculate F-statistics. *Journal of Heredity* 86: 485–486.
- Gunawardhana LN, Al-Rawas GA. 2014. Trends in extreme temperature and precipitation in Muscat, Oman. *Proceedings of the International Association of Hydrological Sciences* 364: 57–63.
- Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95–98.
- Hamidan NA, Geiger MF, Freyhof J. 2014. *Garra jordanica*, a new species from the Dead Sea basin with remarks on the relationship of *G. ghorensis*, *G. tibana* and *G. rufa* (Teleostei: Cyprinidae). *Ichthyological Exploration of Freshwaters* 25: 223–236.

- Hayden B, Pulcini D, Kelly-Quinn M, O'Grady M, Caffrey J, McGrath A, Mariani S. 2010. Hybridisation between two cyprinid fishes in a novel habitat: Genetics, morphology and life-history traits. *BMC Evolutionary Biology* 10.
- Haynes GD, Gongora J, Gilligan DM, Grewe P, Moran C, Nicholas FW. 2012. Cryptic hybridization and introgression between invasive Cyprinid species *Cyprinus carpio* and *Carassius auratus* in Australia: Implications for invasive species management. *Animal Conservation* 15: 83–94.
- Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17: 754–755.
- Jombart T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24: 1403–1405.
- Kamvar ZN, Tabima JF, Grünwald NJ. 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2: e281.
- Keenan K, McGinnity P, Cross TF, Crozier WW, Prodöhl P a. 2013. DiveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution* 4: 782–788.
- Kirchner S, Sattmann H, Haring E, Plan L, Victor R, Kruckenhauser L. 2017. Investigating gene flow between the blind cavefish *Garra barreimiae* and its conspecific surface populations. *Scientific Reports* 7: 5130.
- Kirchner S, Weinmaier T, Rattei T, Sattmann H, Kruckenhauser L. 2014. Characterization of 19 new microsatellite loci for the Omani barb *Garra barreimiae* from 454 sequences. *BMC Research Notes* 7: 522.
- Kopelman NM, Mayzel J, Jakobsson M, Rosenberg N a, Mayrose I. 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources* 15: 1179–1191.
- Kruckenhauser L, Haring E, Seemann R, Sattmann H. 2011. Genetic differentiation between cave and surface-dwelling populations of *Garra barreimiae* (Cyprinidae) in Oman. *BMC evolutionary biology* 11: 172.
- Krupp F, Budd K. 2009. A new species of the genus *Garra* (Teleostei : Cyprinidae) from Oman. *International Journal of Ichthyology* 15: 117–120.
- Kusky T, Robinson C, El-Baz F. 2005. Tertiary-Quaternary faulting and uplift in the northern Oman Hajar Mountains. *Journal of the Geological Society* 162: 871–888.
- Lamer JT, Ruebush BC, Arbieva ZH, McClelland MA, Epifanio JM, Sass GG. 2015. Diagnostic SNPs reveal widespread introgressive hybridization between introduced bighead and silver carp in the Mississippi River Basin. *Molecular Ecology* 24: 3931–3943.
- Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2017. Partitionfinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution* 34: 772–773.
- Librado P, Rozas J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452.
- Lyon RG, Geiger MF, Freyhof J. 2016. *Garra sindhi*, a new species from the Jebel Samhan Nature Reserve in Oman (Teleostei: Cyprinidae). *Zootaxa* 4154: 79–88.
- Mousavi-Sabet H, Eagderi S. 2016. *Garra lorestanensis*, a new cave fish from the Tigris River drainage with remarks on the subterranean fishes in Iran (Teleostei: Cyprinidae). *FishTaxa* 1: 45–54.
- Mousavi-Sabet H, Vatandoust S, Fatemi Y, Eagderi S. 2016. Tashan Cave a new cave fish locality for Iran; and *Garra tashanensis*, a new blind species from the Tigris River drainage (Teleostei: Cyprinidae). *FishTaxa* 1: 133–148.
- Moyle PB, Leidy RA. 1992. Loss of Biodiversity in Aquatic Ecosystems: Evidence from Fish Faunas
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley PF. 2005. MicroChecker User Guide.
- Palandačić A, Matschiner M, Zupančič P, Snoj A. 2012. Fish migrate underground: The example of *Delminichthys adspersus* (Cyprinidae). *Molecular Ecology* 21: 1658–1671.
- Palandačić A, Naseka A, Ramler D, Ahnelt H. 2017. Contrasting morphology with molecular data: An approach to revision of species complexes based on the example of European *Phoxinus* (Cyprinidae). *BMC Evolutionary Biology* 17.
- Paradis E. 2010. pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* 26: 419–420.
- Peakall R, Smouse PE. 2012. GenALEX 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics (Oxford, England)* 28: 2537–9.
- Pichler A, Ahnelt H, Kirchner S, Sattmann H, Haring E, Handschuh S, Freyhof J, Victor R, Kruckenhauser L. 2018. The morphological diversity of *Garra barreimiae* [Teleostei : Cyprinidae].
- Por FD. 1989. *The Legacy of Tethys. An Aquatic Biogeographic of the Levant*. Dordrecht: Kluwer Academic.

- Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics Society of America* 155: 945–59.
- de Queiroz K. 2005. A Unified Concept of Species and Its Consequences for the Future of Taxonomy. *Proceedings of the California Academy of Sciences* 56: 196–215.
- De Queiroz K. 2007. Species concepts and species delimitation. *Systematic biology* 56: 879–86.
- R Development Core Team. 2015. R: A Language and Environment for Statistical Computing.
- Rose JL. 2010. New Light on Human Prehistory in the Arabo-Persian Gulf Oasis. *Current Anthropology* 51: 849–883.
- Rousset F. 2008. Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources* 8: 103–106.
- Sayyadzadeh G, Esmaeili HR, Freyhof J. 2015. *Garra monica*, a new species from the Mond River drainage with remarks on the genus *Garra* from the Persian Gulf basin in Iran (Teleostei: Cyprinidae). *Zootaxa* 4048: 75–89.
- Schlick-Steiner BC, Steiner FM, Seifert B, Stauffer C, Christian E, Crozier RH. 2010. Integrative Taxonomy: A Multisource Approach to Exploring Biodiversity. *Annual Review of Entomology* 55: 421–438.
- Sebastien P. 2016. The program STRUCTURE does not reliably recover the correct population structure when sampling is uneven: sub-sampling and new estimators alleviate the problem. *Molecular Ecology Resources* 16: 608–627.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* 30: 2725–2729.
- Vasilyan D, Carnevale G. 2013. The Afro-Asian labeonine genus *Garra* Hamilton, 1822 (Teleostei, Cyprinidae) in the Pliocene of Central Armenia: Palaeoecological and palaeobiogeographical implications. *Journal of Asian Earth Sciences* 62: 788–796.
- Weir BS, Cockerham CC. 1984. Estimating F-Statistics for the Analysis of Population Structure. *Evolution* 38: 1358–1370.
- Wright S. 1978. *Evolution and the Genetics of Populations, Volume 4. Variability within and Among Natural Populations*. Chicago: University of Chicago Press.
- Xia X. 2017. DAMBE6: New tools for microbial genomics, phylogenetics, and molecular evolution. *Journal of Heredity* 108: 431–437.
- Zachos FE, Habel JC. 2011. *Biodiversity Hotspots. Distribution and Protection of Conservation Priority Areas*.

Concluding Remarks

To summarize the outcome of chapter 1 “Adaptation to cave environments, gene flow between troglomorphic cavefish and congener surface populations”, it was possible to answer all the stated questions based on our dataset. All newly designed microsatellite markers worked well across the individuals sampled and hence, for the first time enabled the analysis of population structure in *G. barreimiae* on a large scale. The microsatellite primers are applicable not only for the study organism, but also for other species of the genus, thus representing a precious contribution for the scientific community in population genetics.

From the start of this project, the cave population of *G. barreimiae* was of particular interest. This population is completely isolated, and its occurrence is restricted to only two adjacent caves, making it vulnerable and prone to extinction. The fact that many populations of a conspecific surface morphotype are still extant and geographically in close vicinity makes this system extremely valuable for evolutionary biologists. Gathering information about the population structure and understanding the evolutionary background of these diverging populations adds valuable information. The conclusions drawn from this study will complement the body of work on cave colonization and the thereby associated adaptations and thus, will ultimately aid in the protection and conservation of this fascinating organism.

In chapter 2 “Phylogeography and cryptic diversity of *G. barreimiae*”, I chose to follow an integrative approach to base the interpretations on various sources of information. Molecular genetic methods (mt and nc) revealed well differentiated and geographically separated groups. On the other hand, it was shown that the taxa of the *G. barreimiae* species complex are morphologically very similar and delimiting species boundaries based solely on morphological characters was not possible. In essence, we arrived at the conclusion that the momentarily valid taxon *G. barreimiae* comprises four cryptic species (in the sense of Bickford et al. 2007). The present study represents a well-supported foundation for species description which is already in preparation and will be the stepping stone for further investigations on the genus.

With these new tools and insights at our disposal, we accomplished to answer the questions that were raised prior to this project and thereby opened a door for further investigations and subsequent studies (some of them already in preparation, see Chapter 3).

CHAPTER 1: ADAPTATION TO CAVE ENVIRONMENTS, GENE FLOW BETWEEN TROGLOMORPHIC CAVEFISH AND CONGENER SURFACE POPULATIONS

Even though the results clearly separate the troglomorphic cave individuals from the nearby surface populations, some of the individuals with intermediate phenotype pose an exception by exhibiting a mixed genotype. These specimens were collected from two sampling sites: the first is an underground water body close to the Al Hoota cave entrance with constant connection to surface water and harbours individuals with intermediate phenotypes that exhibit also mixed genotypes. This strongly indicates an origin via hybridisation between the two morphotypes. Accordingly, gene flow between the troglomorphic cave and sighted surface populations is possible but occurs only unidirectional from the cave to the surface populations. Nevertheless, the population genetic parameters and the few detected hybrid individuals point out that these hybridisation events occur rarely. All collected individuals with intermediate phenotype from the second site have a “pure” nc cave genotype and belong to the cave-haplogroup in the mt CR network, hence they are not the product of hybridisation but rather cave individuals that survived in the surface water body. Even though the average amount of rainfall in Oman is low, sporadic rainfalls can be heavy resulting in massive floodings reaching a flow rate of 100 m³ per second when channelled through the Al Hoota cave system (Waltham, Brown, and Middleton 1985). It seems likely that young cave individuals are swept along the water flow exiting the cave through the lower entry and ending up in the surface water body. The intermediate phenotype of these “pure” cave individuals could be explained by phenotypic plasticity. One could assume that these individuals left the cave at an early developmental stage, when eyes were still visible. Subsequently, the increased exposure to sunlight may lead to darker pigmentation and less reduced eyes compared to cave specimens which developed in complete dark habitat.

Several population genetic parameters show that the overall genetic diversity is much lower in the cave individuals compared to the surface individuals. These results are similar to the findings in cave and surface populations of *Astyanax* (Bradic et al. 2012), and they support the hypothesis that the cave population originated from a small group of surface ancestors that got trapped in the cave after the water level dropped in the course of climate change from wet to arid in the Late Pleistocene (Clark and Fontes 1990; Burns et al. 1998; Kollmann, Sattmann, and Seemann 2007). Subsequently, the isolated cave population went through repeated genetic bottleneck events inducing such low genetic diversity (as already suggested by Kruckenhauser et al. 2011). On the other hand, the different surface populations seem to be in contact with each other more regularly as supported by rather low fixation indices. It seems likely that gene flow between these sites occurs when waterways are easily traversable in seasons of high precipitations.

For the first time, extensive molecular genetic analyses have been conducted on *Garra* from northern Oman. This study highlights the fact that the diversity within this genus has been profoundly underestimated in the past. Compared to the ratio of its rather small geographic distribution range in the Hajar mountains, an astounding high genetic variation was detected that strongly correlates with geomorphological structures. We detected five distinct genetic clades that were supported by both nc microsatellite data as well as mt sequence data with high genetic distances between groups. The geographic distribution of these clades is clearly separated, and their population structure explicitly arranged along geographic large-scale structures as mountain ridges or deep ephemeral riverbeds (wadis) that pervade the countries surface on a large scale. Thus, the clades were named accordingly: North clade, East clade, West clade, Central clade and *gallagheri* clade.

Originally, three subspecies have been described. Only recently, one of these (*gallagheri* clade) has been raised to species status and is now valid as *G. gallagheri* (Lyon, Geiger, and Freyhof 2016). For the first time, members of this taxon have been discovered also at a second site and thus adding significant information on its strongly restricted distribution range. The species *G. gallagheri* is confirmed by our results (also on the basis of nc data) and represents the group with the lowest genetic diversity. The nominate subspecies *G. b. barreimiae* FOWLER & STEINITZ (1956) was described to be distributed throughout the northern Oman with the type locality in Al Buraimi. The distribution of *G. b. shawkahensis* BANISTER & CLARKE (1977) described from Wadi Shawkah lies northwards along the Hajar mountains reaching into the northern part of the UAE. According to our findings, samples from both type localities belong to the West clade, which is the most diversified clade. Consequentially, we found out that behind the species formerly named *G. barreimiae* actually four species are hidden.

While these clades can be genetically clearly separated, they cannot be distinguished based solely on single morphological traits. However, some characters can be consulted to discriminate single clades like the number of gill rakers. Even though most of the measurements are partially overlapping between the clades, several traits display significant differences, in particular measurements of the head structures provide valuable information for discrimination. Besides, also the number of precaudal and intermediate vertebrae as well as the branched pelvic fin rays add information. However, after an extensive morphological assessment, it became clear that the revision of the taxon *G. barreimiae* demands an integrative approach before taxonomic decisions and species delimitations can be drawn.

Although the genetic clades are strongly supported, the phylogenetic relationships between them as well as to other closely related species of the genus are still not completely resolved. In order to infer about the phylogenetic history and place of origin of the *G. barreimiae* species complex, I conducted approximate age estimations for the clades based on substitution rates from the literature for the cytochrome b (*Cytb*) gene,

which yielded in three time intervals marking the separation of the single clades roughly at 1.5 Myr, 4.5 Myr and 6.5 Myr. The same method was utilized to get an estimation for the age of the cave population as well. Accordingly, their emergence was not earlier than in the Middle Pleistocene when the climate in Oman was characterized by gradually increasing aridity interrupted by several pluvial periods (Blechs Schmidt et al. 2009; Burns et al. 1998). These climatic changes might have triggered the colonialization of the Al Hoota cave, either to retreat to the few remaining water reservoirs, or to occupy new freshwater habitats established as consequence of rain periods.

Regarding the genetic diversity between the sampling sites, although most of them are not connected by any perennial waterbody and thus, obviate permanent contact between them, the microsatellite analysis detected no genetic differentiation between single sampling sites of each clade. Therefore, we assume that there must be some recent gene flow, probably due to underground water connections in the karst area, as it has been reported for other cyprinids (Palandačić et al. 2012). Another reasonable explanation for this pattern is heavy precipitation, when wadis are quickly filled up with water and thereby generating temporal surface water connections enabling genetic interchange between individuals from different sites. However, the situation is different concerning contact between individuals of different clades. Out of over 500 investigated individuals, only few exhibit mixed genotypes featuring traces from two clades, thus indicating recent gene flow. In fact, hybridisation between the clades was detected only at three out of 37 sites and affects the closely related East and *gallagheri* clades but also the more distant East and North clades. Individuals collected from these sites exhibit ambiguous nc genotypes while the mt affiliation does not always reflect the same pattern. There are many striking examples highlighting the fact that interspecific hybridisation is a rather common event especially in cyprinids (f.e. Gante, Collares-Pereira, and Coelho 2004; Broughton et al. 2011; Palandačić et al. 2017). This does not only occur in closely related species as one might suspect, hybridisation can also be observed over multiple generations (Lamer et al. 2015) and even between different genera (Hayden et al. 2010; Haynes et al. 2012). However, hybridization between the clades of *G. barreimiae* and *G. gallagheri* is not a common event and does not lead to blurring of the genetic patterns. It seems plausible that these secondary contact zones between the geographically separated clades are established when water-level rises on account of heavy rainfall, thus expanding the distribution area along temporary water routes. Additional factors might play a role like anthropogenic influence either directly by introduction into so far unoccupied water bodies or indirectly by providing potentially new migration routes via manmade irrigation canal systems or possible (temporary) underground connections through the karst area of the Hajar mountains.

With the currently valid taxonomic position, our results clearly rule out monophyly of the “superspecies” *G. barreimiae*. The placement of other species of this genus (*G. gallagheri*, *G. rufa* and *G. sahilia*) in the phylogenetic tree reconstruction disrupts the samples of *G. barreimiae* and thereby its monophyly. Instead, it is evident that this group is in need of a taxonomic revision, which is already in preparation and will result

in the splitting into four cryptic species. Since this genus comprises over 130 described species (Froese and Pauly 2018) and has a vast Afro-Asian distribution, a proper revision of the whole genus is unfeasible to conduct at once. With that in mind, we prefer to make an allegedly small but elaborated start with revising *Garra* from Northern Oman. However, the incorporation of several sources of data by applying an integrative approach and interpreting the combined results is labour-intensive and time-consuming but also essential to keep up with high standards and to maintain scientific integrity.

CHAPTER 3: ONGOING INVESTIGATIONS

The breeding experiments and thereby associated work steps required by far the longest time span for data collection and evaluation. First, we started with attempts to breed our fish stock via environmental induced breeding, where specific environmental cues ought to trigger ovulation and spawning (Lam 1983). We experimented with changing water current and temperature as well as amount and source of food. None of these attempts were successful, so after four months of trying, we decided to switch to hormone induced breeding (Roy 2016). This method is commonly used in commercially important fish, where specimens were anaesthetized with clove oil and then injected intramuscular with a mixture of synthetically produced hormone (LH RH) and Pimozide (dopamine antagonist). We followed the breeding protocol of Cristiano Bertolucci (University of Ferrara), a specialist in physiology who investigated the regulation of the circadian clock in *Phreatichthys andruzzii*, a closely related species originally from Somalia (Colli et al. 2009, Cavallari et al. 2011, Calderoni et al. 2016). The dosage was calculated individually per gram body weight. Hormone induced breeding has several advantages: it guarantees control over time of reproduction, it is a simple procedure that does not need special training and it does not require expensive equipment nor substances. Also, it provides a great opportunity to breed endangered or rare species for conservational purposes. All breeding experiments were conducted on the premises of the Zoo Schönbrunn (Vienna) under the supervision of the head of the department Anton Weissenbacher. After successful injection, fertilized eggs were harvested after 24 hours. We transferred the eggs into tanks under different light regimes, so half of the eggs from both morphotypes were hatched and larvae reared under complete darkness, the other half were exposed to light. Approximately 20 individuals of each group (cavefish/dark, cavefish/bright, surface fish/dark, surface fish/bright) were collected regularly to be able to follow all steps of eye development in the course of ontogeny. Banister (1984) already reported that troglomorphic specimens of *G. barreimiae* displayed elevated melanin production and redevelopment of optic lobes after they were kept under light exposure. Interestingly, juveniles of cavefish seem to have normally developed eyes until they reach a standard length of > 24 mm, at which point eyes stop to develop and are fully overgrown by surrounding tissue (Keith Edward Banister 1984). By breeding both morphotypes of *G. barreimiae* and raising them under different light regimes, we can evaluate the effects of phenotypic plasticity in the course of exposure to sunlight on the development of eyes and corresponding structures.

In particular, we will investigate which structures are missing or present from very early larval stages and how they develop. We will employ μ CT imaging techniques for qualitative and quantitative screening of the structures of interest (lens, optic nerve, retina layer, cornea layer, optical nerve and lobus opticus). Histological methods are planned to provide detailed characterisation of potential microanatomical differences concerning eye development in selected individuals of both populations and especially in specimens featuring an intermediate phenotype.

Just a few individuals were analysed so far, hence we have only preliminary data to discuss, but the final results will culminate in a manuscript in the near future. The striking differences between the morphotypes that were raised under “natural” conditions are the missing lens and optical nerve in the cave form already at a very early developmental stage. On the other hand, we could observe a much thicker retina layer in the surface form. Wilkens (2005) reported that the retina in *G. barreimiae* cavefish can contain either all characteristic layers of a functional eye or is no longer layered at all. However, for statistical support and exact delimitation of differences, much more individuals also from additional developmental stages need to be investigated. Further on, the control groups (cavefish reared under light exposure, surface fish reared in darkness) need to be included for quantifying the effects of different light regimes on the visual system in detail.

In the course of this PhD thesis, it became evident that *G. barreimiae* is genetically much more diverse than previously assumed and comprises four cryptic species. The taxon *G. barreimiae* is currently listed as “Least Concern” according to the IUCN Red List. Hence, the newly discovered *G. barreimiae* species complex in Northern Oman is in urgent need for re-evaluation, as desertification, habitat destruction as well as water shortage and potential chemical pollution pose serious threats to organisms occupying such extreme habitats. The findings gathered in the manuscripts of chapter 2 on morphological and genetic diversity within *G. barreimiae* form the basis for a subsequent species description. The manuscript is already in preparation and will assemble all gathered information and draw taxonomic conclusions by introducing two new species formerly comprised in the *G. barreimiae* species complex, thereby providing the basis for a conservation reassessment.

Bibliography

- Aden, Elke. 2005. "Adaption to Darkness." In *Encyclopedia of Caves*, edited by David C. Culver and William B. White, 1–3. London: Elsevier Academic Press.
- Banister, K.E. 1984. "A Subterranean Population of Garra Barreimiae (Teleostei: Cyprinidae) from Oman, with Comments on the Concept of Regressive Evolution." *Journal of Natural History* 18 (6): 927–38. <https://doi.org/10.1080/00222938400770811>.
- Banister, Keith Edward. 1984. "A Subterranean Population of Garra Barreimiae (Teleostei: Cyprinidae) from Oman , with Comments on the Concept of Regressive Evolution." *Journal of Natural History* 18 (6): 927–38.
- Bickford, David, David J. Lohman, Navjot S. Sodhi, Peter K. L. Ng, Rudolf Meier, Kevin Winker, Krista K. Ingram, and Indraneil Das. 2007. "Cryptic Species as a Window on Diversity and Conservation." *Trends in Ecology and Evolution* 22 (3): 148–55. <https://doi.org/10.1016/j.tree.2006.11.004>.
- Blechs Schmidt, Ingo, Albert Matter, Frank Preusser, and Dirk Rieke-Zapp. 2009. "Monsoon Triggered Formation of Quaternary Alluvial Megafans in the Interior of Oman." *Geomorphology* 110 (3–4): 128–39. <https://doi.org/10.1016/j.geomorph.2009.04.002>.
- Bradic, Martina, Peter Beerli, Francisco J García-de León, Sarai Esquivel-Bobadilla, and Richard L Borowsky. 2012. "Gene Flow and Population Structure in the Mexican Blind Cavefish Complex (*Astyanax Mexicanus*)." *BMC Evolutionary Biology* 12 (1): 9. <https://doi.org/10.1186/1471-2148-12-9>.
- Broughton, Richard E., Krishna C. Vedala, Tessa M. Crowl, and Lauren L. Ritterhouse. 2011. "Current and Historical Hybridization with Differential Introgression among Three Species of Cyprinid Fishes (Genus *Cyprinella*)." *Genetica* 139 (5): 699–707. <https://doi.org/10.1007/s10709-011-9578-9>.
- Burns, Stephen J., Albert Matter, Norbert Frank, and Augusto Mangini. 1998. "Speleothem-Based Palaeoclimate Record from Northern Oman." *Geology* 26 (6): 499–502. [https://doi.org/10.1130/0091-7613\(1998\)026<0499](https://doi.org/10.1130/0091-7613(1998)026<0499).
- Calderoni, Luca, Omar Rota-Stabelli, Elena Frigato, Alex Panziera, Sandra Kirchner, Nicholas S Foulkes, Luise Kruckenhauser, Cristiano Bertolucci, and Silvia Fuselli. 2016. "Relaxed Selective Constraints Drove Functional Modifications in Peripheral Photoreception of the Cavefish *P. Andruzzii* and Provide Insight into the Time of Cave Colonization." *Heredity* 117 (5). <https://doi.org/10.1038/hdy.2016.59>.
- Clark, Ian D., and Jean-Charles Fontes. 1990. "Paleoclimatic Reconstruction in Northern Oman Based on Carbonates from Hyperalkaline Groundwaters." *Quaternary Research* 33 (3): 320–36. [https://doi.org/10.1016/0033-5894\(90\)90059-T](https://doi.org/10.1016/0033-5894(90)90059-T).
- Coghill, Lyndon M, C Darrin Hulsey, Johel Chaves-Campos, Francisco J García de Leon, and Steven G Johnson. 2014. "Next Generation Phylogeography of Cave and Surface *Astyanax Mexicanus*." *Molecular Phylogenetics and Evolution* 79: 368–74. <https://doi.org/10.1016/j.ympev.2014.06.029>.
- Colli, Licia, Annalisa Paglianti, Roberto Berti, Gilberto Gandolfi, and James Tagliavini. 2009. "Molecular Phylogeny of the Blind Cavefish *Phreatichthys Andruzzii* and *Garra Barreimiae* within the Family Cyprinidae." *Environmental Biology of Fishes* 84 (1): 95–107. <https://doi.org/10.1007/s10641-008-9393-z>.
- Feulner, Gary R. 1998. "Wadi Fish of the UAE." In *Tribulus. Bulletin of the Emirates Natural History Group*, 16–22. Abu Dhabi: Tribulus.
- Froese, R, and D Pauly. 2018. "FishBase. Wwww.Fishbase.Org.Version (06/2018). World Wide Web Electronic Publication." 2018.
- Gante, Hugo F., Maria J. Collares-Pereira, and Maria M. Coelho. 2004. "Introgressive Hybridisation between Two Iberian *Chondrostoma* Species (Teleostei, Cyprinidae) Revisited: New Evidence from Morphology, Mitochondrial DNA, Allozymes and NOR-Phenotypes." *Folia Zoologica* 53 (4): 423–32.
- Hayden, Brian, Domitilla Pulcini, Mary Kelly-Quinn, Martin O'Grady, Joe Caffrey, Aisling McGrath, and Stefano Mariani. 2010. "Hybridisation between Two Cyprinid Fishes in a Novel Habitat: Genetics, Morphology and Life-History Traits." *BMC Evolutionary Biology* 10 (1). <https://doi.org/10.1186/1471-2148-10-169>.
- Haynes, G. D., J. Gongora, D. M. Gilligan, P. Grewe, C. Moran, and F. W. Nicholas. 2012. "Cryptic Hybridization and Introgression between Invasive Cyprinid Species *Cyprinus Carpio* and *Carassius Auratus* in Australia: Implications for Invasive Species Management." *Animal Conservation* 15 (1): 83–94. <https://doi.org/10.1111/j.1469-1795.2011.00490.x>.
- Hinaux, Hélène, Sylvie Rétaux, and Yannick Elipot. 2016. "Social Behavior and Aggressiveness in *Astyanax*." In *Biology and Evolution of the Mexican Cavefish*, 335–59. <https://doi.org/10.1016/B978-0-12-802148-4.00017-7>.
- Hüppop, Kathrin. 2000. "Hüppop_2000.Pdf." In *ECOSYSTEMS OF THE WORLD 30. Subterranean Ecosystems*, 159–88.
- . 2012. "Food-Finding Ability in Cave Fish (*Astyanax Fasciatus*)." *International Journal of Speleology* 16 (1/2): 59–66. <https://doi.org/10.5038/1827-806x.16.1.4>.

- Kollmann, Heinz A, Helmut Sattmann, and Robert Seemann. 2007. „Al Hoota“, Ein Karst- Und Höhlenkundliches Forschungs- Und Tourismusprojekt Im Sultanat Oman.“ *Abhandlungen Der Geologischen Bundesanstalt* 60: 107–13.
- Kruckenhauser, Luise, Elisabeth Haring, Robert Seemann, and Helmut Sattmann. 2011. “Genetic Differentiation between Cave and Surface-Dwelling Populations of Garra Barreimiae (Cyprinidae) in Oman.” *BMC Evolutionary Biology* 11 (1): 172. <https://doi.org/10.1186/1471-2148-11-172>.
- Krupp, Friedhelm. 1988. “Freshwater Fishes of the Wadi Batha Drainage.” *Journal of Oman Studies*, no. 3: 401–4.
- Lam, T. J. 1983. *Environmental Influences on Gonadal Activity in Fish. Fish Physiology*. Vol. 9. [https://doi.org/10.1016/S1546-5098\(08\)60302-7](https://doi.org/10.1016/S1546-5098(08)60302-7).
- Lamer, James T., Blake C. Ruebush, Zarema H. Arbieva, Michael A. McClelland, John M. Epifanio, and Greg G. Sass. 2015. “Diagnostic SNPs Reveal Widespread Introgressive Hybridization between Introduced Bighead and Silver Carp in the Mississippi River Basin.” *Molecular Ecology* 24 (15): 3931–43. <https://doi.org/10.1111/mec.13285>.
- Lyon, Robert Gary, Matthias F Geiger, and Jörg Freyhof. 2016. “Garra Sindhi, a New Species from the Jebel Samhan Nature Reserve in Oman (Teleostei: Cyprinidae).” *Zootaxa* 4154 (1): 79–88.
- Niemiller, Matthew L., Maria E. Bichuette, Prosanta Chakrabarty, Dante B. Fenolio, Andrew G. Gluesenkamp, Daphne Soares, and Yahui Zhao. 2019. *Cavefishes. Encyclopedia of Caves*. 3rd ed. Elsevier Inc. <https://doi.org/10.1016/B978-0-12-814124-3.00026-1>.
- Palandačić, Anja, Michael Matschiner, Primož Zupančič, and Aleš Snoj. 2012. “Fish Migrate Underground: The Example of Delminichthys Adspersus (Cyprinidae).” *Molecular Ecology* 21 (7): 1658–71. <https://doi.org/10.1111/j.1365-294X.2012.05507.x>.
- Palandačić, Anja, Alexander Naseka, David Ramler, and Harald Ahnelt. 2017. “Contrasting Morphology with Molecular Data: An Approach to Revision of Species Complexes Based on the Example of European Phoxinus (Cyprinidae).” *BMC Evolutionary Biology* 17 (1). <https://doi.org/10.1186/s12862-017-1032-x>.
- Plath, Martin, Matthias Rohde, Thekla Schröder, Angelika Taebel-Hellwig, and Ingo Schlupp. 2006. “Female Mating Preferences in Blind Cave Tetras Astyanax Fasciatus (Characidae, Teleostei).” *Behaviour* 143 (1): 15–32. <https://doi.org/10.1163/156853906775133560>.
- Plath, Martin, Uta Seggel, Heike Burmeister, Katja U. Heubel, and Ingo Schlupp. 2006. “Choosy Males from the Underground: Male Mating Preferences in Surface- and Cave-Dwelling Atlantic Mollies (Poecilia Mexicana).” *Naturwissenschaften* 93 (3): 103–9. <https://doi.org/10.1007/s00114-005-0072-z>.
- Romero, Aldemaro. 2009. *Cave Biology: Life in Darkness*. Cambridge: Cambridge University Press.
- . 2012. “Caves as Biological Spaces.” *Polymath: An Interdisciplinary Arts and Sciences Journal* 2 (3): 1–15.
- Roy, Koushik. 2016. “Rapid Review on the Use of New Age Induced Breeding Agent ‘LHRHa’ in Indian Finfish Seed Production Sector.” *Journal of Fisheries* 4 (2): 401. <https://doi.org/10.17017/jfish.v4i2.2016.145>.
- Strecker, Ulrike, Víctor H. Faúndez, and Horst Wilkens. 2004. “Phylogeography of Surface and Cave Astyanax (Teleostei) from Central and North America Based on Cytochrome b Sequence Data.” *Molecular Phylogenetics and Evolution* 33 (2): 469–81. <https://doi.org/10.1016/j.ympev.2004.07.001>.
- Strecker, Ulrike, Bernhard Hausdorf, and Horst Wilkens. 2012. “Parallel Speciation in Astyanax Cave Fish (Teleostei) in Northern Mexico.” *Molecular Phylogenetics and Evolution* 62 (1): 62–70. <https://doi.org/10.1016/j.ympev.2011.09.005>.
- Teyke, Thomas. 1990. “Morphological Differences in Neuromasts of the Blind Cave Fish Astyanax Hubbsi and the Sighted River Fish Astyanax Mexicanus.” *Brain, Behavior and Evolution* 35: 23–30.
- Waltham, A C, R D Brown, and T C Middleton. 1985. “Karst and Caves in the Jabal Akhdar, Oman.” *Cave Science* 12 (3): 69–79.
- Wilkens, Horst. 1988. “Evolution and Genetics of Epigeal and Cave Astyanax Fasciatus (Characidae, Pisces). Support for the Neutral Mutation Theory.” *Evolutionary Biology* 23: 271–367. https://doi.org/10.1007/978-1-4613-1043-3_8.
- . 2005. “Fish.” In *Encyclopedia of Caves*, edited by David C. Culver and William B. White, 241–51. London: Elsevier Academic Press.
- Wilkens, Horst, and Ulrike Strecker. 2017. *Evolution in the Dark. Darwin’s Loss Without Selection. Evolution in the Dark*. https://doi.org/10.1007/978-3-662-54512-6_1.

Appendix – Conference contributions

In 2014 at the beginning of my PhD, I presented a poster introducing the concept and research questions of the project on *G. barreimiae* at the annual meeting of the “**Network of Biological Systematics Austria**” (NOBIS) in Munich (Germany).

Chapter 1 examines the “Adaptation to cave environments and gene flow between troglomorphic cave fish deals with the question if the blind cave population inhabiting the Al Hoota cave is still in contact with the nearby surface populations and whether gene flow can be detected by means of microsatellite analyses. The first (preliminary) findings of these population genetic analyses were presented in 2015 at the meeting of the “**Society for Molecular Biology and Evolution**” (SMBE) in Vienna (Austria). Later this year, after finishing the analyses, the concluding results were presented at the “**XV European Congress of Ichthyology**” (ECIXV) in Porto (Portugal) in the course of an oral presentation, which was rewarded with the Student Award for “best oral presentation” by the European Ichthyological Society. These results were presented to the scientific community on an international level at the “**23rd International Conference on Subterranean Biology**” (ICSB) in Fayetteville (USA) in 2016.

The findings of Chapter 2 on “Phylogeography and cryptic diversity of *G. barreimiae*” were presented at the “**110th Annual Meeting of the German Zoological Society**” (DZG) in Bielefeld (Germany) in 2017. In the same year, these results were presented at the “**Young researchers’ day 2017**” organized by NOBIS Austria in Vienna (Austria) in the form of a poster, which was awarded best poster.

In 2018, all information gathered in the course of this project was put into context with the potential natural (climate change and desertification) and anthropogenic (pollution and ground water extraction) threats to *G. barreimiae* and presented at the “**3rd Annual Meeting in Conservation Genetics**” in Vienna (Austria). So far, only two cave populations harbouring the cave morphotype of *G. barreimiae* are known, making them vulnerable and prone to become extinct. By highlighting the potential threats and danger of losing such a fascinating study organism, I wanted to raise awareness and highlight the urgency for a re-evaluation of the conservation status of *G. barreimiae* to the scientific community.

3rd Annual Meeting in Conservation Genetics 2018

at the Natural History Museum Vienna, AUSTRIA (26. – 28. February 2018)

DANGER IN THE DARK: POTENTIAL THREATS TO ISOLATED CAVE POPULATIONS OF *GARRA BARREIMIAE*

Kirchner, S.^{1,2}, Sattmann, H.³, Haring, E.^{1,2}, Victor, R.⁴ & Kruckenhauser, L.¹

¹ Central Research Laboratories, Natural History Museum Vienna, Burgring 7, 1010 Vienna,

² Department of Integrative Zoology, University of Vienna, Althanstraße 14, 1090 Vienna, Austria,

³ Third Zoological Department, Natural History Museum Vienna, Burgring 7, 1010 Vienna,

⁴ Department of Biology, Sultan Qaboos University, Al Khoudh, Muscat 123, Oman.

Garra barreimiae is a cyprinid freshwater fish endemic to the northern Oman and United Arab Emirates, where it inhabits all kinds of water bodies. Due to the harsh climate and scarce rain, *G. barreimiae* has become a true survivalist which endures the dry season in the crevices of the Hajar mountains. Since the mountains are mostly cretaceous limestone rock, many of such lacunas which serve as important water reservoirs can be found. The largest known cave in Oman is the Al Hoota cave, a spacious 5 km long cave system that is roughly 2 million years old and holds several underground lakes. In 1980, a troglomorphic cave population of *G. barreimiae* has been discovered in these lakes. The adult individuals are characterized by the complete lack of pigmentation and external visible eyes. Our genetic analyses show that the population of the Al Hoota cave is completely isolated from the surface populations. The Al Hoota cave has two differently elevated entries, without any natural barrier, therefore harmful substances can easily be introduced from upstream and thereby potentially menacing the whole blind cave fish population.

Unlike in many other cave fish species in *G. barreimiae* the troglomorphic and the conspecific surface phenotype still exist and are accessible. That allows comparative analyses, hence making it a fascinating system to study evolutionary adaptation to life in complete darkness.

So far, the Al Hoota cave and the nearby Hoti Pit cave are the only known occurrences of the cave morphotype of *G. barreimiae*, increasing the threat of losing this valuable study system, especially considering that its conservation status was changed from “vulnerable” to “least concern” in 2015. By investigating the genetic diversity and population structure of this organism, we want to raise awareness and subsequently increase conservation status for the cave population of *G. barreimiae*.

DANGER IN THE DARK

POTENTIAL THREATS TO ISOLATED CAVE POPULATIONS OF *GARRA BARREIMIAE*

Sandra Kirchner^{1,2}, Helmut Sattmann³, Elisabeth Haring^{1,2}, Reginald Victor⁴, Luise Kruckenhauser²

¹University of Vienna, Department of Integrative Zoology, Althausstraße 14, 1090 Vienna, Austria
²Natural History Museum Vienna, Central Research Laboratories, Burggasse 7, 1180 Vienna, Austria
³Natural History Museum Vienna, Third Zoological Department, Burggasse 7, 1180 Vienna, Austria
⁴Sultan Qaboos University, Department of Biology, Al Khodh, Muscat 123, Oman

Cave morphotype

Garra barreimiae is a cyprinid freshwater fish endemic to northern Oman and UAE. The troglomorphic cave form of this taxon occurs only in the Al Hoota cave and the nearby Hoti Pit. The cave phenotype differs from the surface form by the lack of eyes and pigmentation, reduced optic lobes and squamation.

The cave population is genetically isolated from the surface populations. Although the population experienced at least one bottleneck event in the past, it is currently not threatened by inbreeding depression. Our *nc* microsatellite data suggests that gene flow from the cave to surface populations is still possible, but occurs only rarely.

Furthermore, our results suggest that the cave population is phylogenetically rather young.



The lakes in the Al Hoota cave are densely populated by the cave morphotype of *G. barreimiae*.

Waste management & pollution

In Oman, over 1.2 kg waste are produced per person every day, most of it winds up at one of the > 350 dumpsites, creating environment and health issues. In 2011, the first engineered sanitary landfill was built, several others are planned. However, this resembles only the first step leading to more sustainable waste management practices (e.g. waste reduction, reuse, recycling).

Besides the problems caused by waste, we observed mycosis infected and deceased fish in front of the lower Al Hoota cave entrance.

If waste/pathogens are introduced to the cave, this poses a serious threat to the unique cave population of *G. barreimiae*.



Heavily polluted water bodies are not uncommon, still they are sometimes inhabited by *G. barreimiae*.

Al Hoota cave

The Al Hoota cave is located in the midst of the Al Hajar mountains (northern Oman). There are four permanent lakes which are inhabited by the cave morphotype of *G. barreimiae*.

The ~ 5 km long cave system has two differently elevated entrances which form an underground flood system during rare but heavy rainfalls. Except for floods, the wadis up- and downstream of the cave are dry. Temperature ~ 25 °C, humidity up to 100 %.



The Al Hoota cave is > 2 myr old and developed in Cretaceous limestone. It is characterized by spacious halls and big boulders.



Groundwater overpumping results in seawater intrusion, there are currently > 160 000 plastic water pump engines.

Freshwater shortage

Oman appears on 9th place on the list of countries that will face extreme high water stress by 2040. Water consumption is still increasing, reasons are:

- rapid population growth,
- increasing temperature due to climate change,
- intensive water usage for petroleum production.

To address the water demand, construction of desalination plants has intensified. Also, over 30 major recharge dams have been built, thus threatening the habitat of *G. barreimiae* (desiccation of natural wadis and falaj system).

Water pumps are used to tap groundwater, especially for water-intensive agriculture, resulting in rapid drop of groundwater levels leading to seawater intrusion.



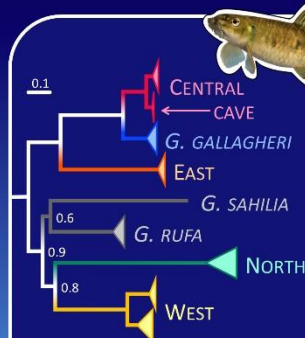
Oman's climate is arid and dry, rain is scarce. Water consumption has increased by 20 % from 2010 to 2015.

Surface populations & genetic diversity

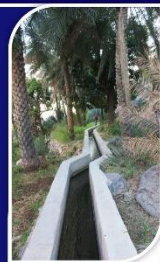
Garra barreimiae is a cyprinid freshwater fish endemic to the northern Oman and UAE. Due to the harsh climate and scarce rain, *G. barreimiae* has become a true survivalist which endures the dry season in the crevices of the Hajar Mountains, which serve as important water reservoirs.

Our analyses show high genetic diversity within what is currently known as '*Garra barreimiae*'. These results are in accordance with the geographic data, since genetic groups are detected along the watershed of the Hajar mountains. The maximum differences between the mt clades (13% in CR) are even higher than compared to well acknowledged species of the genus. This is also supported by *nc* (microsatellite) data.

A taxonomic revision will result in the splitting of *G. barreimiae* in its four genetic / geographic groups. Hence, the Red List Assessment should be revised and re-evaluated as well.



Bayesian inference tree of concatenated mt sequences (COI + Cytb + CR: 2739 bps) of *G. barreimiae* and *G. gallagheri* collected from 20 sampling sites as well as *G. rufo* / *G. sahila*.



The falaj system (water drainage) is used by the locals to transport freshwater from the spring to the villages and gardens. Quite often, one can also find high numbers of *G. barreimiae* in these drains.

Map of northern Oman with sampling sites coloured according to mt clades. The clades are congruent with the watershed along the Hajar mountains: central (magenta), *G. gallagheri* (blue), east (orange), north (green), west (yellow), outgroup (grey). *G. rufo* from Turkey, *G. sahila* from Saudi Arabia. Polygons indicate distribution range of mt clades. Location of the Al Hoota cave is indicated by a black arrow in the middle of the central clade distribution range.



Sultanate of Oman
 Al Hoota cave

IUCN RED LIST INFO

"This species is common and widespread throughout the mountains of Oman and the United Arab Emirates in the Arabian Peninsula. It is therefore assessed as **Least Concern**."

Previously published Red List assessments:

1996 – Vulnerable (VU)

1990 / 1994 – Rare (R)

TAKE HOME MESSAGE

Unlike in many other cave fish species both cave and surface phenotypes of *G. barreimiae* still exist and are accessible, what allows comparative analyses, hence making it a fascinating study system for evolutionary adaptation to life in complete darkness.

So far, the Al Hoota cave and the nearby Hoti Pit cave are the only known occurrences of the cave morphotype of *G. barreimiae*, increasing the threat of losing it. By investigating the genetic diversity and population structure, we want to raise awareness and subsequently increase conservation status for the cave population of *G. barreimiae*.

The study was funded by the funded research programme (University of Vienna, the Bank Austria and the Austrian Economic Chamber) and the Ministry of Science and Research (Austria). We thank the Ministry of Environment and Climate Affairs of Oman for cooperation, support and the sampling permits. We also want to thank Robert Bask, Lukas Han and Arthur Fichtler for collecting samples. The Federal Ministry for Transport, Innovation and Technology Austria provided financial support during the course of a PhD thesis.



phm
 naturhistorisches museum wien

Kirchner, S., H. Sattmann, E. Haring, L. Plan, R. Victor, L. Kruckenhauser (2017) Investigating gene flow between the blind cavefish *Garra barreimiae* and its conspecific surface populations. In: Scientific Reports 7:5130.
 Kruckenhauser, L., E. Haring, R. Seemans, and H. Sattmann (2011) Genetic differentiation between cave and surface-dwelling populations of *Garra barreimiae* (Cyprinidae) in Oman. In: BMC evolutionary biology 11:172.
 Kirchner, S., T. Weinmann, T. Rattai, H. Sattmann and L. Kruckenhauser (2014) Characterization of 19 new microsatellite loci for the Oman barb *Garra barreimiae* from 454 sequences. In: BMC research notes 7:522.
 Pichler, A., H. Ahnelt, S. Kirchner, H. Sattmann, E. Haring, S. Handschuh, J. Freyhof, L. Kruckenhauser (submitted) The morphological diversity of *Garra barreimiae* (Teleostei: Cyprinidae).

universität wien

110th Annual Meeting of the German Zoological Society (DZG) 2017

at the University Bielefeld, GERMANY (12. – 15. September 2017)

AND THEN THERE WERE FOUR – HIDDEN GENETIC DIVERSITY WITHIN THE CYPRINID FRESHWATER FISH *GARRA BARREIMIAE* FROM OMAN

S. Kirchner, H. Sattmann, E. Haring, R. Victor, L. Kruckenhauser

The North of the Sultanate of Oman is characterized by the Al Hajar mountains, which consist of anticlinal folds of limestone and dolomites. It is a harsh and arid environment with high mountain peaks and many wadis, which occasionally dry out. Only some of the bigger ones carry water permanently. These water bodies are inhabited by the cyprinid *Garra barreimiae*, which is endemic to the northern Oman and the UAE. Within this species, three subspecies have been described: *G. b. barreimiae* Fowler & Steinitz (1956), *G. b. shawkahensis* Banister & Clarke (1977) and *G. b. gallagheri* Krupp (1988). The latter has been lifted to species status recently (Lyon et al., 2016).

In the present phylogeographic study, we investigated the species' genetic diversity using mitochondrial (CR, COI) and nuclear (microsatellite) markers. The analyses based on mt data display a high genetic diversity, resulting in four well defined clades. This is also supported by the microsatellite data, which highlights substructuring within the clades. These clades have different geographic distributions which are in accordance with the drainage systems of the Al Hajar mountains. Between these clades the genetic distances are high and in the range of those between *G. barreimiae* and the closely related *Garra rufa*. A recent study showed that these groups are morphologically variable and only few meristic characters separate the West clade from the others (Pichler et al., submitted). While the West clade comprises samples of both subspecies *G. b. barreimiae* as well as *G. b. shawkahensis*, the remaining three clades correspond to not yet described taxa.

Our results clearly show cryptic variability within *G. barreimiae* and will help to decipher the taxons' historic background. Consequently, as *G. barreimiae* is not monophyletic according to our results, this study depicts the necessity for a revision of this group from Oman, which should cover the three newly discovered clades and establish them as species.

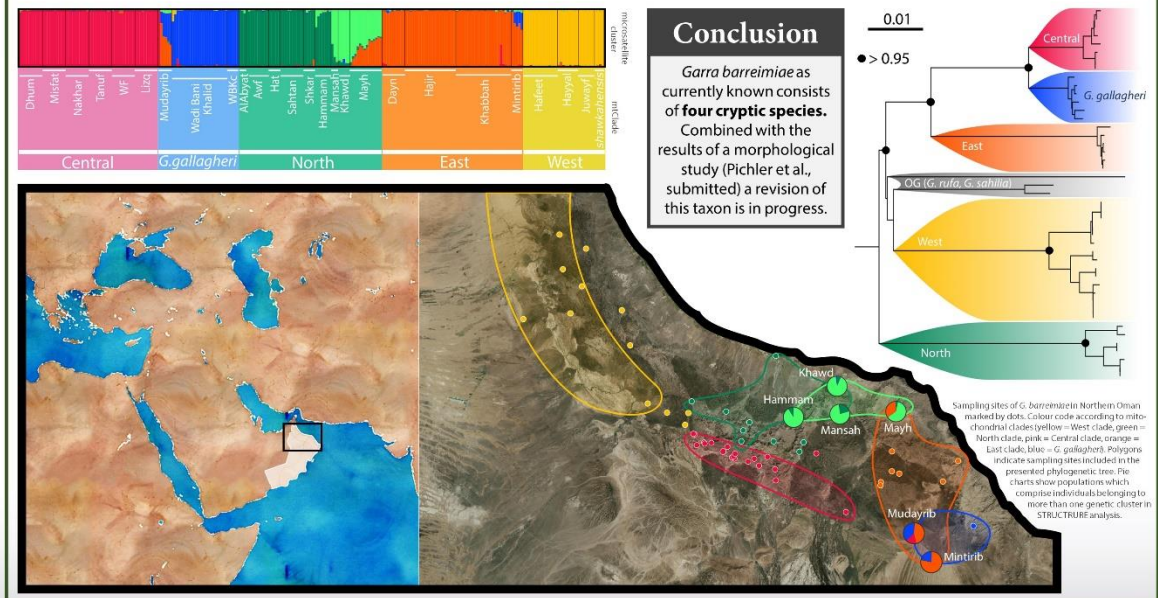
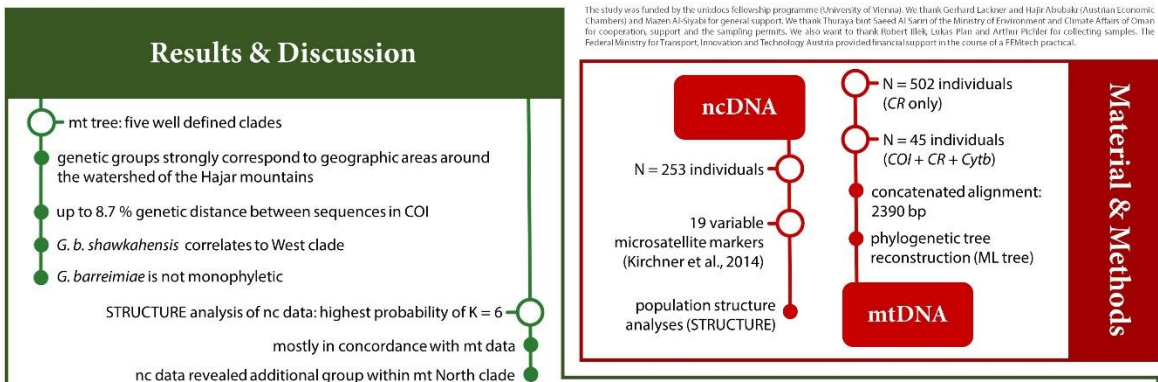
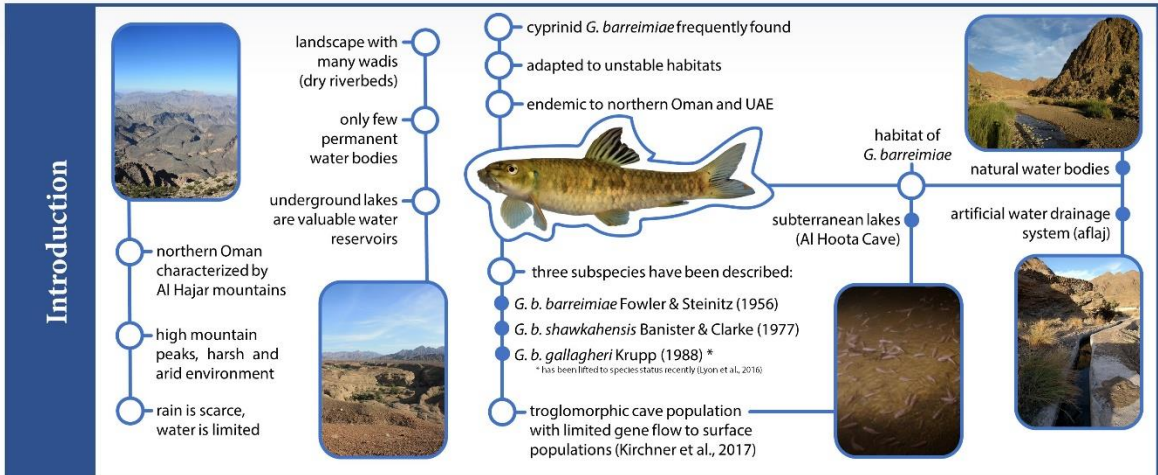


And then there were four - Hidden genetic diversity within the cyprinid freshwater fish *Garra barreimiae* from Oman



Sandra Kirchner^{1,2}, Helmut Sattmann¹, Elisabeth Haring^{1,2}, Reginald Victor⁴, Luise Kruckenhauser²

¹ University of Vienna, Department of Integrative Zoology, Althanstrasse 14, 1090 Vienna, Austria
² Natural History Museum Vienna, Central Research Laboratories, Burggasse 7, 1010 Vienna, Austria
³ Natural History Museum Vienna, Third Zoological Department, Burggasse 7, 1010 Vienna, Austria
⁴ Sultan Qaboos University, Department of Biology, Al Khoudh, Muscat 123, Oman



Kruckenhauser, L., E. Haring, R. Seemann, and H. Sattmann (2011). "Genetic differentiation between cave and surface-dwelling populations of *Garra barreimiae* (Cyprinidae) in Oman". In: BMC evolutionary biology 11:172.
 Kirchner, S., T. Weinmaier, T. Rattai, H. Sattmann and L. Kruckenhauser (2014). "Characterization of 19 new microsatellite loci for the Omani barb *Garra barreimiae* from 454 sequences". In: BMC research notes 7:522.
 Lyon, R. G., Geiger, M. F. and J. Freyhof (2016). "Garra sindhi, a new species from the Jebel Samhan Nature Reserve in Oman (Teleostei: Cyprinidae)". In: Zootaxa 4154:1.
 Kirchner, S., H. Sattmann, E. Haring, L. Pichler, R. Victor, L. Kruckenhauser (2017). "Investigating gene flow between the blind cavefish *Garra barreimiae* and its conspecific surface populations". In: Scientific Reports 7:5130.
 Pichler, A., H. Ahnelt, S. Kirchner, H. Sattmann, E. Haring, S. Handschuh, J. Freyhof, L. Kruckenhauser (submitted). "The morphological diversity of *Garra barreimiae* (Teleostei: Cyprinidae): Evidence for cryptic species?"



23rd International Conference on Subterranean Biology (ICSB) 2016

at the University Fayetteville, Arkansas, USA (13. – 17. June 2016)

AS ABOVE, SO BELOW? TESTING FOR GENE FLOW BETWEEN CAVE AND SURFACE-DWELLING POPULATIONS OF *GARRA BARREIMIAE*.

Kirchner, Sandra^{1,2}; Sattmann, Helmut³; Plan, Lukas⁴; Krenn, Harald¹; Victor, Reginald⁵; Haring, Elisabeth^{1,2}; Kruckenhauser, Luise²

¹ Department for Integrative Biology, University of Vienna, Althanstraße 14, 1090 Vienna, Austria

² Central Research Laboratories, Natural History Museum, Burgring 7, 1010 Vienna, Austria

³ Third Zoological Department, Natural History Museum, Burgring 7, 1010 Vienna, Austria

⁴ Department for Geology & Paleontology, Natural History Museum, Burgring 7, 1010 Vienna, Austria

⁵ Department of Biology, Sultan Qaboos University, Al Khoudh, Muscat 123, Oman

Garra barreimiae is a cyprinid freshwater fish endemic to the southeastern Arabian Peninsula, where it inhabits regularly desiccating water bodies. A hypogean population with troglomorphic traits found in the Al Hoota cave system in northern Oman is considered as conspecific with the normally developed surface morphotypes. Normally the fish live in quiet cave pools but after heavy rain the cave is flushed and limited contact with the surface population is likely. Therefore *G. barreimiae* is a perfect organism to study the effects of life in complete darkness and mechanisms of adaption.

The DNA-study of Kruckenhauser et al. (2011) based on a section of the mitochondrial control region showed that specimens from the cave population are genetically differentiated from the adjacent surface populations. Some surface individuals possess haplotypes characteristic for the cave population and most of these individuals display intermediate phenotypes. These results suggest a quite recent separation of the cave population. For a better understanding of the evolution of the troglomorphic phenotype it is crucial to know if and to what extent gene flow between these morphotypes occurs. Hence, we investigated individuals from two troglomorphic cave populations (including Al Hoota), one cave with constant water outflow giving connection to the surrounding surface water body and two surface sites. In order to estimate potential gene flow between cave and surface populations we utilized 19 variable microsatellite markers designed specifically for *G. barreimiae* (Kirchner et al. 2014). The results show that both caves inhabited by troglomorphic fish are clearly separated from the surface populations. Some of the specimens with intermediate phenotypes possess a hybrid genotype, while others are genetically clearly of cave origin. For the latter, it can be assumed that cave fish were washed out during floods and the intermediate phenotype is a result of phenotypic plasticity induced by light exposure.

XV European Congress of Ichthyology (ECIXV) 2015

in Porto, PORTUGAL (07. – 11. September 2015) (talk)

TO SEE OR NOT TO SEE – THAT IS THE QUESTION. POPULATION GENETIC ANALYSES IN CAVE-DWELLING POPULATIONS OF *GARRA BARREIMIAE*

S. Kirchner^{1,2}, H. Sattmann³, E. Haring^{1,2}, L. Kruckenhauser²

¹ Department for Integrative Biology, University of Vienna, Althanstraße 14, 1090 Vienna, Austria

² Central Research Laboratories, Museum of Natural History, Burgring 7, 1010 Vienna, Austria

³ Third Zoological Department, Museum of Natural History, Burgring 7, 1010 Vienna, Austria

E-mail: sandra.kirchner@nhm-wien.ac.at

The cyprinid *Garra barreimiae* is endemic to the southeastern Arabian Peninsula, where it inhabits regularly desiccating water bodies. Since a hypogean population with troglomorphic traits (loss of eyes/pigmentation) is conspecific with the “normally” developed surface population, *G. barreimiae* seems to be a perfect organism to study the effects of cave life and the origin of troglomorphic characters.

Previous genetic analysis (Kruckenhauser et al. 2011) showed that specimens from the cave population are genetically differentiated from the adjacent surface populations. Few surface individuals were found, which exhibit haplotypes characteristic for the cave population and most of these display an intermediate phenotype. This indicates a quite recent separation of the cave population from the surface populations. For a better understanding of the evolution and the maintenance of the troglomorphic phenotype it is crucial to know if and to what extent gene flow between the cave and surface populations occurs. For this purpose, we will investigate individuals from three different cave localities, including individuals with an intermediate phenotype, as well as several surrounding surface populations with 19 variable microsatellite markers, which were designed and established especially for *G. barreimiae* in a previous study (Kirchner et al. 2014). The results of the present study will deliver the basis for interpreting the recent population genetic history of the species, detect hybridisation and evaluate genetic variability within the species in general.

2015 Meeting of the Society for Molecular Biology and Evolution (SMBE)

at the Hofburg Vienna, AUSTRIA (12. – 16. July 2015) (poster)

TO SEE AND TO BE SEEN – A CLOSER LOOK INTO THE EVOLUTION OF A BLIND FISH

S. Kirchner, H. Sattmann, E. Haring, L. Kruckenhauser

The cyprinid *Garra barreimiae* is endemic to the southeastern Arabian Peninsula, where it inhabits regularly desiccating water bodies. Since a hypogean population with troglomorphic traits (loss of eyes/pigmentation) is conspecific with the “normally” developed surface population, *G. barreimiae* seems to be a perfect organism to study the effects of cave life and the origin of troglomorphic characters.

Previous genetic analysis (Kruckenhauser et al. 2011) showed that specimens from the cave population are genetically differentiated from the adjacent surface populations. Few surface individuals were found, which exhibit haplotypes characteristic for the cave population and most of these display an intermediate phenotype. This indicates a quite recent separation of the cave population from the surface populations. For a better understanding of the evolution and the maintenance of the troglomorphic phenotype it is crucial to know if and to what extent gene flow between the cave and surface populations occurs. For this purpose, we will investigate individuals from three different cave localities, including individuals with an intermediate phenotype, as well as several surrounding surface populations with 19 variable microsatellite markers, which were designed and established especially for *G. barreimiae* in a previous study (Kirchner et al., 2014). The results of the present study will deliver the basis for interpreting the recent population genetic history of the species, detect hybridisation and evaluate genetic variability within the species in general.

TO SEE AND TO BE SEEN

A closer look into the evolution of a blind fish

Sandra Kirchner^{1,2}, Helmut Sattmann³, Elisabeth Haring², Luise Kruckenhauser²

¹ University of Vienna, Department of Integrative Zoology, Althanstraße 14, 1090 Vienna, Austria
² Central Research Laboratories, Natural History Museum Vienna, Burggasse 7, 1010 Vienna, Austria
³ Third Zoological Department, Natural History Museum Vienna, Burggasse 7, 1010 Vienna, Austria



Garra barreimiae is a cyprinid barb endemic to the Hajar mountains in northern Oman and the UAE. In 1980 Dunsire discovered a population of this species inhabiting a subterranean lake in the Al Hoota cave in the midst of the Jabal Akhdar mountains (Oman). Specimens of this cave population feature troglomorphic traits like imperfectly developed eyes, demagnified optic lobes, reduction of squamation and loss of pigmentation. Besides these differences, the cave and surface individuals are morphologically similar.

Kruckenhauser et al. (2011) analysed part of the mitochondrial control region (CR) and found, that all individuals collected at the Al Hoota cave belong to one haplogroup. This cave haplogroup also contains individuals from nearby surface populations, of which most exhibit an intermediate phenotype (less pigmentation, reduced or missing eyes, see Fig. 1). This indicates a quite recent separation of the cave population from the surface population and the question arises whether there is still some gene flow between the populations.



Fig. 1. Morphotypes of *G. barreimiae*: typical surface (top), intermediate (mid) and cave morphotype (bottom).

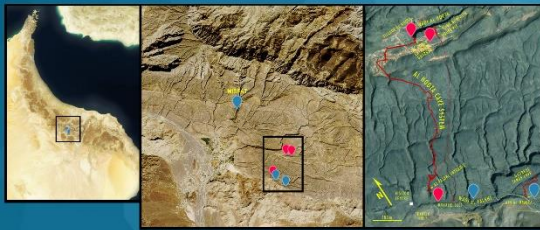


Fig. 2. Sampling locations of *G. barreimiae* in northern Oman, blue marks indicate surface sampling sites and red marks indicate cave sampling sites (troglomorphic individuals).

To infer the potential hybridization between the cave and surface populations, we analyzed 19 variable microsatellite markers which were designed specifically for *G. barreimiae* (Kirchner et al., 2014) on 208 individuals from different cave and surface sites. Troglomorphic individuals were collected at the Al Hoota cave (35) and at Hoti Pit (17). The surface individuals were collected at Wadi Falahi (51), the closest Wadi to the Al Hoota cave, at Ain Al Msalla (66) which is downstream of Wadi Falahi and includes surface as well as underground waters (Ain Ghubrat Tanuf) and at Misfat (39), a water channel which lies a few km northwest of the Al Hoota cave (Fig. 2).

We calculated population genetic parameters (Fig. 4, 5) and analysed the putative population subdivision with the program Structure. The expected heterozygosity and the number of alleles show that the cave populations are genetically less variable than the surface populations. ΔK pinpoints the highest probability at $K=2$, whereat individuals from the Al Hoota cave and Hoti Pit represent a „cave“ population (red) and the other individuals are united in a „surface“ population (blue, see Fig. 3). All individuals of the two „cave“ populations display a pure cave genotype and show no sign of hybridization between cave and surface populations. Also individuals from the „surface“ population Misfat show no sign of hybridization. However, some individuals from the population Ain Al Msalla are assigned partly to the „cave“ population and partly to the „surface“ population. Most of these individuals exhibit an intermediate or blind phenotype. It is most likely that this is the result of hybridization between cave and surface individuals.

There are also individuals from the „surface“ population Wadi Falahi featuring an intermediate phenotype, but all of them are completely assigned to the „cave“ population. The most reasonable explanation for this is, that single individuals originated in the cave and were flushed out into the lower lying epigeal water bodies of Wadi Falahi during rainfall periods and develop an intermediate phenotype induced by the exposure to light.

The second highest probability of population subdivision ($K=3$) shows that the „surface“ population is substructured into individuals from Misfat (yellow) and a rather mixed group of individuals from Wadi Falahi (Fig. 3).



Fig. 3. Results of the Structure analyses with $K=2$ (top) and $K=3$ (bottom): the whole data set is divided into a „cave“ group comprising the populations Al Hoota and Hoti Pit (red) and a „surface“ group including the populations Ain Al Msalla, Wadi Falahi (blue) and Misfat (K=2 blue, K=3 yellow). Columns marked with an * indicate intermediate or blind individuals collected at surface populations.



Conclusion: The results clearly show that hybridization between the cave and surface population is still possible and can occur in areas where epigeal and hypogean water bodies are connected. Besides, the findings also show that the Al Hoota population is genetically completely isolated.

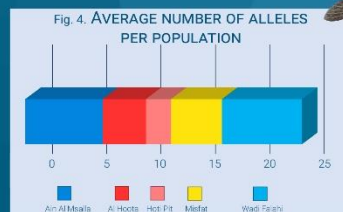


Fig. 4. AVERAGE NUMBER OF ALLELES PER POPULATION

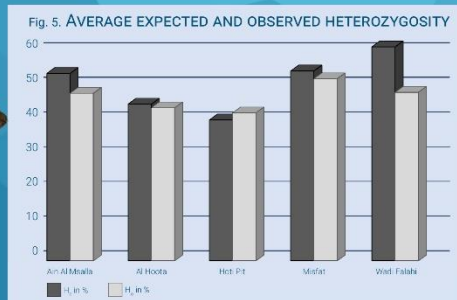


Fig. 5. AVERAGE EXPECTED AND OBSERVED HETEROZYGOSITY

References: Kruckenhauser, L., E. Haring, R. Seemann, and H. Sattmann (2011). „Genetic differentiation between cave and surface-dwelling populations of *Garra barreimiae* (Cyprinidae) in Oman“. In: BMC evolutionary biology 11:172.
 Kirchner, S., T. Weinmaier, T. Rattai, H. Sattmann and L. Kruckenhauser (2014). „Characterization of 19 new microsatellite loci for the Omani barb *Garra barreimiae* from 454 sequences“. In: BMC research notes 7:522.



universität
wien

This study was funded by the unidos fellowship of the University of Vienna and the Natural History Museum in Vienna. The Ministry of Commerce and Industry and the Ministry of Tourism of the Sultanate of Oman supported this project as well as Salim al Maskery (Oman Geological Consultant). We thank the Ministry of Environment and Climate Affairs of Oman for their cooperation, support and the sampling permits and Michael Robinson and Reginald Victor, zoologists at the Sultan Qaboos University in Muscat, for their intellectual input. The Federal Ministry for Transport, Innovation and Technology Austria provided financial support in the course of a FEMTech practical.



Network of Biological Systematics Austria (NOBIS) 8 – Crossing Borders

at the Zoologische Staatssammlung Munich, GERMANY (28. – 29. November 2014)

SHEDDING LIGHT ON THE EVOLUTION OF AN ENIGMATIC FISH. COMPARATIVE ANALYSES OF CAVERNICOLOUS AND SURFACE-DWELLING POPULATIONS OF *GARRA BARREIMIAE*

Sandra Kirchner^{1,2}, Anton Weissenbacher³, Helmut Sattmann⁴, Elisabeth Haring², Harald Krenn¹, Luise Kruckenhauser²

¹ Department for Integrative Zoology, University of Vienna, Althanstraße 14, 1090 Vienna, Austria

² Central Research Laboratories, Museum of Natural History, Burgring 7, 1010 Vienna, Austria

³ Aquaria & Terrarium division, Zoo Vienna, Maxingstraße 13b, 1130 Vienna, Austria

⁴ Third Zoological Department, Museum of Natural History, Burgring 7, 1010 Vienna, Austria

sandra.kirchner@nhm-wien.ac.at

Many cave-dwelling vertebrates display the same morphological modifications, such as eye regression and loss of pigmentation. It is still unclear how these regressive traits are genetically determined and whether phenotypic plasticity plays an important role.

The cyprinid barb *Garra barreimiae* is endemic to the southeastern Arabian Peninsula, where it inhabits regularly desiccating water bodies. Since a hypogean population with troglomorphic traits is conspecific with the “normally” developed surface population, *Garra barreimiae* seems to be a perfect organism to study the effects of cave life and the origin of troglomorphic characters. Previous genetic analysis of mtDNA (Kruckenhauser et al. 2011) revealed that individuals from the cave population are genetically differentiated from the adjacent surface populations. However, some individuals from surface populations near the cave entry possess haplotypes characteristic for the cave population. As those individuals mostly display an intermediate phenotype, it is reasonable to assume that hybridization still occurs and that the populations have separated quite recently.

In this project we want to answer the question whether and to what extent hybridization between cave and surface populations occurs by analysing microsatellite markers (Kirchner et al. 2014). To evaluate the consequences of light exposure during development and crossbreeding between individuals from cave and surface populations we will examine if offspring display different development of eyes and pigmentation in the presence or absence of light. For the first time extensive morphometric and morphological analyses will be conducted with *G. barreimiae*. This will include histology of eye development, various distance measurements and shape analyses. By comparing size and shape of the different species of the genus and subspecies of *G. barreimiae* we want to evaluate if status of these species and the subspecies is justified. In order to explain how the current distribution patterns of the species emerged, we will investigate the

phylogeography by analysing several mitochondrial and nuclear markers. The combined results shall help to draw a conclusive picture of the evolution of this interesting species and to obtain new insights into mechanisms leading to the transition to troglotrophic life style in general.



SHEDDING LIGHT ON THE EVOLUTION OF AN ENIGMATIC FISH

Comparative analyses of cave- and surface-dwelling populations of *Garra barreimiae*

Sandra Kirchner^{1,2}, Harald Krenn¹, Helmut Sattmann³, Elisabeth Haring², Luise Kruckenhauser²

¹ University of Vienna, Department of Integrative Zoology, Althanstraße 14, 1090 Vienna, Austria
² Central Research Laboratories, Natural History Museum Vienna, Burgring 7, 1010 Vienna, Austria
³ Third Zoological Department, Natural History Museum Vienna, Burgring 7, 1010 Vienna, Austria





Troglobites are animals that spend their whole life in complete darkness. Typical troglomorphic traits are: lack of eyes and pigmentation, often in combination with highly developed non-visual sensory organs.



Introduction

Garra barreimiae is a cyprinid barb endemic to the Hajar mountains in northern Oman and the UAE. In 1980 Dunsire discovered a subterranean population of this species inhabiting a lake in the Al Hoota cave at the southern slopes of the Jabal Akhdar mountains (Oman). Specimens of this cave population feature troglomorphic traits like imperfectly developed eyes (overgrown by dermal tissue), demagnified optic lobes, reduction of squamation and loss of pigmentation. They were described as a conspecific population of the cyprinid barb *Garra barreimiae* by Banister (1984) due to concordance of morphological characters between hypogean and epigeal populations.

Cavernicolous animals are eminent models to study general principles of evolution as well as the cause of regressive evolutionary processes during the adaptation to a new habitat. Thus, *G. barreimiae* is an excellent organism for comparative studies of these differing traits since both cave and surface populations are still extant. Kruckenhauser et al. (2011) analysed part of the mitochondrial control region (CR) and found, that all individuals collected at the Al Hoota cave belong to the same haplogroup. This cave haplogroup also contained individuals from nearby surface populations (most of them with intermediate phenotype) while other, more distant populations are genetically well separated. Thus, it can be assumed that there is some gene flow between the populations.

1. Hybridisation



The Al Hoota cave system (top) shelters the only known troglomorphic cave fish population of *G. barreimiae* (bottom)

- What is the extent of gene flow between the populations?
- Do adults of both morphotypes mate under lab conditions and do hybrid-offspring display an intermediate phenotype?
- Do different light conditions during the rearing of juveniles have any effect on the phenotype?

Gene flow: Following the study of Kruckenhauser et al. (2011), we will test whether and to what extent specimens of the cave population hybridise with specimens from surface populations (in nature). For this purpose, 454 sequence data and bioinformatic methods were used to design and establish suitable microsatellite markers (Kirchner et al., 2014). These markers will be utilised to perform population genetic analyses in order to infer about gene flow between the populations. If individuals exhibit population-specific alleles of both populations, it may be an indication that they originate from hybridisation.

Breeding experiments: Juveniles of both populations will be held in presence/absence of light during the rearing to see whether exposure to light/darkness has any effect on their phenotype. Furthermore we will try to mate adults of both populations under lab conditions to check if crossbreeding between the two morphotypes is possible and to test if hybrid-offspring display an intermediate phenotype.



Surface fish (mid) inhabit natural water bodies (top) and Wadis as well as falaj (water drainage system, right)

2. Phylogeography

- How genetically distinct are specimens of *G. barreimiae* and other species of the genus?
- Is the geographical distribution pattern in concordance with the genetic distances?

To infer about the extent of genetic variation within *G. barreimiae* including its morphotypes and subspecies and between other species of the genus, we will use several genetic markers (mt: CR, COI, cyt b and nc: RAG1, RH) to analyse their relationship genetically. Based on these phylogenetic relationships associated with the current distribution areas, we will try to find patterns which might help to explain the results of the genetic analyses and reveal past migration patterns. Combined with morphometric results, these findings will help to evaluate the status of the various sub-/species.



Sampling locations of *G. barreimiae* in northern Oman, black and blue marks indicate typical localities of the three subspecies of *G. barreimiae*, detailed map shows position of Al Hoota cave

3. Morphology

- What are the morphological differences in eye development between the two morphotypes?
- Are the three subspecies morphologically clearly differentiated?
- What differences in size and shape can be found between the (sub)species?

Histology: Juveniles of cave fish seem to have normally developed eyes which are later overgrown by surrounding tissue. Histological methods will provide detailed characterisation of potential microanatomical differences in different growth stages of juveniles of both morphotypes which have been reared under different light conditions and especially in specimens featuring an intermediate phenotype.

Morphometry: For the first time an extensive number of specimens of *G. barreimiae* will be analysed morphometrically (via landmark analyses) and morphologically (via distance based measurements). In doing so, we try to capture the morphological variation between and within (sub)species occurring on the Arabian Peninsula with a clear focus on species delimitation.



Morphotypes of *G. barreimiae*. Top: surface, mid: intermediate, bottom: cave



References:
 Kruckenhauser, L., E. Haring, R. Serrano, and H. Sattmann (2011): "Genetic differentiation between cave and surface-dwelling populations of *Garra barreimiae* (Cyprinidae) in Oman". In: BMC evolutionary biology 11:172.
 Kirchner, S., T. Widemayer, T. Ratsch, H. Sattmann and L. Kruckenhauser (2014): "Characterization of 19 new microsatellite loci for the Oman barb *Garra barreimiae* from 454 sequences". In: BMC research notes 7:522.

