Desert fish populations tolerate extreme salinity change to overcome hydrological constraints

Celia Schunter[1]*, Lucrezia C. Bonzi [2], Jessica Norstog[2], Jade Sourisse[1], Michael L. Berumen[2], Yoseline Angel[3], Stephen D. Parkes[3], Matthew F. McCabe[3] and Timothy Ravasi[4]*

[1]Swire Institute of Marine Science, School of Biological Sciences, The University of Hong Kong, Pokfulam, Hong Kong SAR

[2]Red Sea Research Center, Division of Biological and Environmental Sciences & Engineering, King Abdullah University of Science and Technology, Thuwal, Kingdom of Saudi Arabia

[3]Water Desalination and Reuse Center, Division of Biological and Environmental Sciences & Engineering, King Abdullah University of Science and Technology, Thuwal, Kingdom of Saudi Arabia

[4]Marine Climate Change Unit, Okinawa Institute of Science and Technology Graduate University, 1919–1 Tancha, Onna-son, Okinawa 904–0495, Japan

*corresponding authors: Celia Schunter (celiaschunter@gmail.com) and Timothy Ravasi (timothy.ravasi@oist.jp)

Abstract

The unstable nature of freshwater ponds in arid landscapes represent a sizable challenge for strictly aquatic organisms, such as fishes. Yet the Arabian Desert, bordering the coastline of the Red Sea, plays host to a species very well adapted to such extreme environments: the Arabian pupfish, *Aphanius dispar*. In this study, we estimated patterns of hydrological connectivity; population structure and stable isotope for samples of *A. dispar* living in small, isolated ponds of nearly-freshwater in the Arabian desert and highly saline coastal lagoons along the Red Sea. The genomic and hydrological analyses indicate that populations are largely separated by drainage origin, as fish from desert ponds appear to be transported to coastal lagoons of the Red Sea along ephemeral river systems arising from flash flood events. Further, our study indicates there is an ecological change when being washed from pond environments to coastal waters, due to a significant shift in muscle stable isotopes ratios between both groups. Considering that the genetic breaks are mostly observed between drainage origin, this study suggests that *A. dispar* can survive large changes in salinity and ecological regimes over small time-scales.

Introduction

Population dynamics are often defined and restricted by differences in environmental conditions. A growing body of research shows that environmental gradients are often a driver for selection, but also population structure, as adaptation to local conditions can lead to selection of particular genotypes on either side of an ecological break (Wang & Bradburd, 2014). Differences in fitness and conditioning across such ecological gradients can lead to restricted dispersal or even dispersal barriers, which in turn can significantly shape the population structure at relatively small geographic scales (Sexton, Hangartner, & Hoffmann, 2014). This concept is tremendously important in evolutionary biology, as ecological gradients have been suggested as one of the initial mechanisms of divergence that can promote speciation of lineages (Doebeli & Dieckmann, 2003).

A small number of species have the ability to traverse ecological barriers, even significant dispersal barriers and eventually conquer some of the world's extreme environments, via physiological and behavioural adaptations (Kelley et al., 2014; Moore, Cooper, Biewener, & Vasudevan, 2017). Among the most intriguing of these examples are fishes that live in desert environments, as they inhabit highly isolated and often ephemeral ponds in very arid landscapes. It is assumed that most of these species have high phenotypic plasticity, as these fishes encounter ephemeral streams (Furness, 2016), large temperature fluctuations (Bennett & Beitinger, 1997), extreme changes in water chemistry (Kavembe, Franchini, Irisarri, Machado-Schiaffino, & Meyer, 2015) and high spatiotemporal variability in water supply (Fisher, Gray, Grimm, & Busch, 1982). Among the main questions regarding these particular groups of fishes, the most intriguing

62 is how such isolated populations continue to survive after an initial colonization event. It
63 is well known that isolation of populations can result in a slew of detrimental conditions,
64 such as loss of genetic variability, inbreeding depression or the accumulation of
65 deleterious mutations (Gaggiotti, 2003). Considering that population persistence and
66 long-term survival is largely influenced by genetic diversity (Bouzat, 2010), living in
67 isolated and highly restricted water bodies might threaten the persistence of desert fishes
68 over long time-scales.
69
70 Among the few fish groups that inhabit arid environments, Cyprinodontiformes exhibit
71 remarkable adaptations to extreme conditions. For example, the African turquoise
72 killifish (*Nothobranchius furzeri*) evolved an embryonic diapause, allowing fertilized
73 eggs to survive a dry period (Furness, 2016). Furthermore, killifish can also be found in
74 habitats with widely varying salinities and they are therefore categorized as euryhaline
75 (Wood & Marshall, 1994). The genus *Fundulus*, for instance, displays a wide range of
76 osmotolerant physiologies (Whitehead, 2010), with *F. heteroclitus* being able to rapidly
77 acclimate to an osmotic shock by changing its transcriptional program and later
78 remodeling its tissues (Whitehead, Galvez, Zhang, Williams, & Oleksiak, 2011). Despite
79 these remarkable adaptations, some Cyprinodontiformes occupy very restricted habitats.
80 For example, the Devils Hole pupfish (*Cyprinodon diabolis*) lives only in Devils Hole
81 (in the Amargosa Desert of Nevada) and it is described as occupying the smallest known
82 natural range (a single pool < 80 m^2) for any vertebrate species (Martin, Crawford,
83 Turner, & Simons, 2016). Indeed, many species of fish inhabiting such arid landscapes
84 are currently listed as endangered (Hopken, Douglas, & Douglas, 2013; Van Haverbeke,

Stone, Coggins, & Pillow, 2013) and have been of conservation concern for decades (Meffe & Vrijenhoek, 1988).

The Arabian pupfish, *Aphanius dispar* (Cyprinodontidae), is present in the Middle East with landlocked populations in countries such as Oman, Iran and Saudi Arabia (Al-Kahem-Al-Balawi et al., 2008; Freyhof, 2014; Haas, 1982; Hrbek & Meyer, 2003). This species represents an interesting conundrum, as it has a large distribution range, but it inhabits highly restricted ephemeral ponds with no permanent rivers around them. Furthermore, from a pilot study we revealed that they are also present in coastal lagoons of the Red Sea. Phylogenetic analyses of the genus suggest that this group has saltwater ancestry, with the closing and drying of the Tethys Sea resulting in landlocked remnant populations which likely then diverged due to the resultant strong ecological changes (Hrbek & Meyer, 2003).

In this study we explore suitable habitats for *A.dispar* to live in the desert as well as the Red Sea coastline in Saudi Arabia and aim to understand the population connectivity of this species. We investigate the environmental conditions and population structure of *A. dispar* in the Arabian desert by employing a multi-disciplinary approach that combines hydrological predictive mapping, population genomics, stable isotope analysis, as well as chemical and physical analyses of water bodies.

Methods

Sample collection

In order to acquire information on the presence of water ponds and streams in the Saudi Arabian desert, extensive searches were performed using Google Earth, high-resolution satellite data and through contacting local landowners and regional police. Subsequently, a range of potential water ponds were identified and visited (see Figure 1). From a survey of 28 locations, *Aphanius dispar* was detected at 11 sites and was collected using a 3 m wide seine net. Fish length was measured, and a piece of the dorsal fin was cut and placed in 96% ethanol. Fish were then returned to the pond. Individuals with fins too small for finclip collection were euthanized with a blow to the head, with the entire tail preserved in ethanol. All procedures were performed in accordance with relevant guidelines and regulations and were approved and completed with the ethics permit 15IBEC35_Ravasi from the Institutional Biosafety and BioEthics Committee (IBEC) of the King Abdullah University of Science and Technology.

Stable isotopes

We employed two types of stable isotope measurements: fish muscle tissue and pond water isotopes. Fish muscle tissue can give indications on the diet of the fish in the ponds and the sea (Zanden & Rasmussen, 2001). Water isotope analysis can indicate the evaporative and flow processes of the water body. While these distinct isotopic measurements have different applications, they both help to understand the biological and hydrological connectivity of the ponds in the desert. Isotope ratios are reported relative to their respective standards:

129 1) $$\delta = \left(\frac{R_{sample}}{R_{standard}} - 1\right)‰$$

130 where R_{sample} and $R_{standard}$ are the isotope ratios (heavier isotopologue to the more
131 abundant isotopologue) measured in the sample.

132

133 *a) Water stable isotope analysis*

134 Water samples from the six desert ponds were analyzed for isotope analysis at the King
135 Abdullah University of Science and Technology on a Picarro Wavelength Scanning
136 Cavity Ringdown Spectrometer (WSCRDS L115-I, Picarro Inc., Sunnyvale, CA, USA)
137 interfaced to a liquid autosampler (CTC HTC Pal liquid autosampler; LEAP
138 Technologies, Carrboro, NC, USA). δ^2H and $\delta^{18}O$ were measured and are isotope ratios
139 for $^2H^1H^{16}O$ and $^1H_2{}^{18}O$ isotopologues, respectively. Samples were referenced to the
140 VSMOW scale using laboratory working standards previously referenced to the
141 VSMOW2/ SLAP2 (Standard Light Antarctic Precipitation) standards. The δ^2H and $\delta^{18}O$
142 values of samples was determined using the two-point linear normalization method
143 described in Paul *et al.* (2007) (Paul et al., 2007), which calibrate sample isotope ratios
144 using the linear relationship between the true and measured isotope ratios of two working
145 standards that bracket the sample. Each measurement of standards and samples consisted
146 of 10 injections of which the last 4 were used to determine isotope ratios. Uncertainty for
147 the reported isotope values is determined by propagating the standard deviations (1σ) of
148 the last 4 injections from measurement of standards and samples.

149

150 *b) Muscle tissue stable isotope preparation and analysis*

For each collection site, a subset of adult fish was sacrificed, and bodies were frozen with dry ice in the field. In the lab, fish were thawed, and white muscle tissue was de-skinned, descaled and dried for 24 h at 60 °C. Samples were ground using a MP Bio FastPrep-24 instrument (6.5 m/s for 60s until pulverized, repeated 1-4 times) and Lysing Matrix E 2 mL tubes with 1 x 4.0 mm ceramic sphere, 30±3 1.4 mm ceramic spheres. Tissues were rinsed in 1 mL 2:1 chloroform-methanol solution and shaken vigorously for 30 s to remove isotope differences due to a lipid content bias, as described by Ehrich *et al.* (2011) (Ehrich et al., 2011). The solution was transferred to 1.5 mL Eppendorf tubes, leaving the spheres behind. The samples were left to stand for 15 minutes at room temperature, and then centrifuged for 10 minutes at 3,400 rpm. The supernatant was discarded, and the chloroform-methanol rinse was repeated two additional times. Samples were then dried in 1.5 mL Eppendorf tubes for 24 h at 60 °C and broken apart using a scoopula. Approximately 1±0.2 mg of sample was placed in tin capsules for solid samples (5 x 9 mm) for the Isotope Ratio Mass Spectrometry analysis (IRMS). ^{13}C and ^{15}N isotopes were analyzed using a PDZ Europa ANCA-GSL elemental analyzer interfaced to a PDZ Europa 20-20 isotope ratio mass spectrometer (Sercon Ltd., Cheshire, UK) at the Stable Isotope Facility at the University of California, Davis. Briefly, samples were combusted at 1000°C with chromium oxide and silvered copper oxide, with a subsequent oxides removal in a reduction reactor. N_2 and CO_2 were separated on a Carbosieve GC column (65°C, 65 mL/min) before entering the IRMS. Samples were interspersed with replicates of two laboratory standards and isotope ratio delta values were measured relative to the international standards VPDB (Vienna PeeDee Belemnite) and Air for carbon and nitrogen, respectively (Sharp, 2007).

Statistical analyses for tissue stable isotopes were carried out using R (R version 3.4.0, 2017-04-21). Data were checked for normality with the Shapiro-Wilk test and for homogenous variance with the Fligner-Killeen test and considered significant at $p < 0.05$. The correlation between $\delta^{13}C$ and $\delta^{15}N$ was tested with a Spearman's correlation, as well as the allometric effect of size on isotope contents. Since data did not satisfy conditions for the use of parametric statistics, a two-way non-parametric ANCOVA in the *sm* package (Bowman & Azzalini, 2014) was used to test for differences in $\delta^{13}C$ and $\delta^{15}N$ among sites, with fish size as the covariate. Sampling sites of *A. dispar* were split into two water type groups (Desert Pond and Sea Water), as well as in four different clusters (determined through genetic and hydrological analysis, stated below), depending on the drainage basin they belong to (from north to south: Cluster 1 = DP1, DP2, SW1, SW2; Cluster 2 = DP3, SW3; Cluster 3 = DP4, SW4; Cluster 4 = DP5, DP6, SW5). Differences in fish muscle isotopic signatures between different desert pond were also evaluated. Muscle mean content in $\delta^{13}C$ and $\delta^{15}N$ and standard deviation for the different sites (n: DP3 =11, DP4 = 11, all others = 12) were calculated and visualized in R.

Water parameters and samples

To quantify the water chemistry of the desert ponds and better understand the environmental conditions of the fish habitat, we took *in situ* measurements as well as collected water samples for later analyses. A portable Ocean Seven 305 Plus CTD (Idronaut) was used to measure temperature, pressure, conductivity, salinity, percent oxygen saturation, parts per million of O_2 and pH on site. The CTD sonde was submersed at a depth of approximately 30 cm and left to acclimate for one minute. Five minutes of

197 measurements were taken at a sampling rate of 5 seconds, with the total of 60 data points
198 averaged to provide a single value for each parameter per site (referred to as "average
199 temperature"). Water samples at each site were collected in 500 ml HDPC containers,
200 either filled only with water for isotope and water analysis, or with 2ml of 5% nitric acid
201 for preservation for chemical water analysis. Water samples were analyzed using Ion
202 Chromotography. Seawater samples (sites SW1 – SW5) were diluted x10 and filtered
203 using Dionex OnGuard II Ba/Ag/H 2.5 cc cartridges to remove SO_4^{2-} and Cl^-. Standards
204 and water samples were transferred to Autoselect Polyvial 10 mL vials and covered with
205 septa and caps. Samples were run on Dionex ICS-3000 with the Chromeleon
206 Chromatography Management System (version 6.7) program and using an autosampler.
207 Photometric analysis was performed on water samples to measure for chlorides, sulphates
208 and silica. Standards were run accordingly (chlorides: 0 ppm, 20 ppm, 200 ppm;
209 sulphates: 0, 30 and 100 ppm; silica: 0, 100 and 500 ppb). Water samples were
210 transferred to cuvettes and loaded into an Aquakem 250 (Thermo Scientific) for analysis.
211 For samples that were outside of standard ranges, samples were diluted appropriately (see
212 Supplementary materials) and re-run.
213
214 *Streamflow mapping*
215 To establish the hydrological connectivity from the desert highlands to the Red Sea
216 coastline (and the resultant possible genetic links between the locations where pupfish
217 were present), we needed to determine the predominant flow direction at each cell of an
218 underlying topographic model. The ArcGIS-based hydrology toolset was used for the
219 extraction and analysis of watersheds and streamflow, using the Advanced Spaceborne

Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (ASTGTM, JPL 2009) to provide a topographic description of the region. Raw ASTER data, distributed by NASA as GeoTIFF files, have a 30 m spatial resolution and are referenced to the WGS84 coordinate system. Supplementary Figure 1 shows the flow process of delineating watershed boundaries and stream networks from a digital elevation model (DEM). To ensure that any water "flow" can move from one cell to any adjacent cell, a depressionless DEM was obtained by filling any localized "sinks" that might have formed as an artefact of the interpolation process. From this, the natural flow direction (as dictated by the direction of steepest descent) and the flow accumulation per cell (i.e. the number of upstream cells "draining" to that particular cell) can then be calculated. Only those cells with a high accumulation threshold (>1000 contributing cells) are considered to represent a dominant flow path. A streamline can be produced by connecting these cells. Watershed boundaries are delineated automatically based on the natural water divides, which follow the highest elevations in the DEM. Using Google Earth imagery (Google Earth 7.1.2.2041; December 31, 2016) as an underlying base map, we can determine the potential connectivity routes for fish by following delineated streamlines from the highest sampled locations on the mountains, to the lowest points near the shoreline. A visual interrogation of satellite images allowed elements such as dams, bridges, culverts and agricultural regions to be identified and to manually edit segments along the streamlines.

DNA extractions and Restriction Site Associated DNA Sequencing (RAD-Seq)

242 To understand the genetic population structure and genetic connectivity of *A. dispar* we
243 used 5 mm pieces of fin clip from each individual sample for 28-29 individuals per site
244 (at 11 sites) to extract DNA with 96-well DNA extraction kits from Qiagen (DNeasy 96
245 Blood and Tissue Kit) or Macherey-Nagel (NucleoSpin 96 Tissue). The manufacturer
246 protocols were followed with a deviation of 8-10 hours of lysis and elution in 50 µl H_20.
247 Concentrations were measured with a Qubit 2.0 fluorometer with a dsDNA High
248 Sensitivity reagent kit. We used a modified double digest (ddRad) protocol (Peterson,
249 Weber, Kay, Fisher, & Hoekstra, 2012). Briefly, DNA was digested using SphI High
250 Fidelity and MluCI High Fidelity enzymes with the appropriate CutSmart Buffer (New
251 England Biolabs) and cleaned with AMPURE XP beads. Adapters were ligated to 100 ng
252 of digested DNA with a combination of sixteen adapters and eleven indices to uniquely
253 identify individual samples out of 80 multiplexed samples in one sequencing lane. Pools
254 were created from equally concentrated and cleaned samples and size selected in a
255 BluePippin (Sage Science) with 2% Gel Cassettes to a size of 300bp. We used a KAPA
256 Hifi Ready Mix for the PCR amplification with ten PCR cycles and temperatures
257 according to the manufacturer. Samples were run on a bioanalyzer (Agilent) and a qPCR
258 (7900 HT Fast Real Time PCR system, ABI) was used to quantify and combine all 80
259 samples within one library at equimolar concentration. Four libraries were then
260 sequenced paired-end on an Illumina Hiseq2000 to a length of 100bp at the KAUST
261 Bioscience Core Lab facility.
262
263 *RAD-Seq data processing*

Raw sequence fastq files were de-multiplexed for each lane of Illumina using *process_radtags* in the software STACKS 1.40 (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013). All sequences were quality trimmed with Trimmomatic 0.33 (Bolger, Lohse, & Usadel, 2014) with a Phred score quality cutoff of 30. Individuals with less than 300,000 remaining first read sequences (from paired end reads) were removed from the analysis. Due to the lack of a reference genome for this species, we ran the quality trimmed first read files through the *denovo_map* perl script in STACKS. The final optimized parameters were set to a conservative number of mismatches allowed (-n 2 and -M 2), a minimum number of identical reads to form a stack was three and SNP calling was performed with an upper bound error rate of 0.05 (--bound_high 0.05). After the putative SNP detection, the function *populations* was run to select putative SNPs meeting several criteria. The variant had to have a minimum read number of 10 (-m 10), be present in at least seven of the eleven locations (-p 7) and in more than 20 individuals per locations (-r 0.72). A minimum allele frequency filter of 0.05 was applied and we chose to use one randomly putative SNP from each stack. The resulting vcf file was converted into different input file formats for further analysis using PGDSPIDER (Lischer & Excoffier, 2012). To avoid a bias due to duplicated sequences across the genome, putative SNPs were discarded if heterozygosity was higher than 0.5 (calculated with *vcftools*, following (Danecek et al., 2011) and if the mean read depth was a median absolute deviation away from the median depth (Seeb et al., 2014). Hardy-Weinberg exact tests were performed in Genepop version 4.6 (Rousset, 2008) and putative SNPs were removed if there was a significant deviation in more than four of the eleven

286 locations for which tests could be performed. All SNPs that did not pass the criteria were
287 blacklisted in *populations* and removed from further analysis.
288
289 *Population genetics and clustering analyses*
290 Population genetics metrics such as average allele number, private alleles, inbreeding
291 coefficient and locus and pair-wise F_{ST} were obtained for the set of filtered final SNPs
292 with *populations*. VCFtools v0.1.13 was used to look at minor allele frequencies, Hardy-
293 Weinberg equilibrium and locus and population heterozygosities (Danecek et al., 2011).
294 Pairwise AMOVA F_{ST} values for population differentiation measures and significance
295 ($p<0.05$) were obtained in GenoDive V2.0b27 using 10,000 permutations as well as the
296 pairwise kinship coefficient r (Meirmans & Van Tienderen, 2004). Effective population
297 sizes were calculated in NeEstimator using the linkage disequilibrium method including
298 all final selected SNP loci (Do et al., 2014). Here we present the results of the largest
299 harmonic mean sample size per population.
300 To evaluate genetic structure among all sampled individuals, we performed a principal
301 component analysis (PCA) as well as Bayesian clustering analysis. The PCA was
302 computed in the *adegenet* package v. 2.0.1 in R (Jombart, 2008). We represent the
303 eigenvalues of the analysis revealing the variance of each principal component and a
304 scatterplot summarizing the genetic diversity including the center ellipses per sampling
305 location. Clustering analysis was performed in Structure v.2.3.4 (Pritchard, Stephens, &
306 Donnelly, 2000) under the admixture model with a 10% burn-in period and 500,000
307 iterations of Markov Chain Monte Carlo (MCMC), which creates a probability
308 distribution and allows for the evaluation of the likeliness of different numbers of clusters

(K) within the dataset. Ten replicates were run for each putative number of clusters (K) with K ranging from 1 to 11. The results were then passed through STRUCTURE HARVESTER v0.6.94 (Earl & vonHoldt, 2012) to apply the *ad hoc* statistic delta K proposed by Evanno and coauthors (2005). Resulting individual and population files were used in CLUMPP v1.1 (Jakobsson & Rosenberg, 2007) and DISTRUCT v1.1 (Rosenberg, 2003) to combine all STRUCTURE runs and visualize the results.

Loci under selection

As *A. dispar* individuals were found in very different habitat types (from desert ponds to the highly saline waters of the Red Sea) we tested for possible selection in any of the analyzed loci possibly indicating adaptive processes to the environmental conditions. For this we re-ran the *populations* program with the same whitelist of 5,955 loci selecting specific locations. First, all locations were used by defining all desert pond locations as one population and all seawater location as another one, in order to evaluate global selection between desert and seawater habitats. However, adaptive processes might differ for the different desert ponds and we therefore sub-selected locations. We evaluate adaptive loci for DP1 & DP2 against SW1 & SW2, DP3 vs. SW3, DP4 vs. SW4, and DP5 & DP6 against SW5. Each resulting vcf files was format converted with PGDSpider 2.0.0.2 (Lischer & Excoffier, 2012). For outlier loci detection we used a Bayesian approach incorporated in Bayescan v2.1 (Foll & Gaggiotti, 2008). Briefly, posterior odds of a locus being under selection are obtained with MCMC with the help of the proportion of loci exhibiting large F_{ST} in comparison to other loci. To further minimize the number of false positives, the prior odds was increased to 100 (-pr_odds 100). Outlier loci were

then visualized and selected in R after applying a False Discovery Rate (FDR) of 0.05. Sequence reads for resulting putative loci under selection were then blasted against the NCBI nr database (blastn) and successful blast hits are presented if the evalue was below $1e^{-5}$.

Results

Pupfish locations, collection and pond properties

Due to the lack of previous knowledge on the locations of the Arabian pupfish populations in western Saudi Arabia, we sampled 28 sites, of which only eleven had pupfish present (Figure 1). 12 Red Sea coastline sites and 16 inland enclosed ponds or streams were visited (Table 1 & Supplementary Table 1); 6 inland sites and 5 seawater sites had *Aphanius dispar* specimen. For two inland ponds, no sonde measurements could be obtained due to complications transporting the CTD (conductivity, temperature and depth) sonde into the steep canyons. Five ponds contained freshwater with a salinity upper maximum threshold of 0.5 ppt, whereas salinity in the other nine measured ponds ranged from 0.54 to 1.45 ppt. Most of the ponds therefore consisted of brackish water. *Aphanius dispar* was only found in ponds with salinities higher than 0.74 ppt. pH ranged between 7.71 to 9.41 across all sites, with the majority exhibiting a pH between 8 and 9. The average temperature in the inland ponds was 28.9 °C (±4.9 SE), while in the Red Sea sites it was 21.9 °C (±4.17 SE). Sampling of fish was performed during the late boreal autumn and early winter months (i.e., November-January), hence during the period of lowest temperature of the year. The highest average water temperature, 38.4 °C, was observed at the Al Lith hot spring. At this site, pupfish live in higher temperatures in comparison to all other sites, as well as with larger water content of silica (80,470 ppb). These results illustrate the wide range of environments which pupfish populations are able to exploit along the Red Sea coast.

Water stable isotopes

To further understand environmental conditions of the six inland ponds inhabited by *A. dispar*, water samples were collected, and their stable oxygen and hydrogen isotopes measured (Supplementary Table 1). This can aid in understanding the source of the water in the desert pond and if evaporation occurs comparatively, which can indicate water flow or standing waters. A strong linear relationship between the $\delta^{2}H$ and $\delta^{18}O$ of these samples was observed, with a slope of 4.83. Supplementary Figure 2 shows the linear regression of measurements, with data exhibiting a low slope compared to the Global Meteoric Water Line (GMWL). The relatively low slope indicates samples were subject to significant evaporative enrichment (Gat, 1996; Gibson, Birks, & Edwards, 2008; Gibson & Reid, 2014). The linear relationship between $\delta^{2}H$ and $\delta^{18}O$ across all samples indicates a common water source derived from a single recharge event, with variability between samples caused by the degree of evaporative enrichment for each water pool. This is consistent with observations made at each site, with DP1 and DP2 being the two most stagnant ponds with the least oxygen and greatest $\delta^{2}H$ and $\delta^{18}O$ isotopic enrichment (Table 1).

Streamflow mapping

The overall topography of the study region is a mountainous inland region that is bounded to the east by the Arabian shield, and which drains westward to a flatter desert terrain towards the Red Sea coastline. In order to understand the hydrology of the area and the exact hydrological constraints for the Arabian pupfish we created streamflow maps for each sample location derived from a satellite-based digital elevation model. Using the derived stream networks, a total of 6 different migration pathways were

383 identified within the study area (Figure 2 & Supplementary Table 4). Figure 2 illustrates
384 the digital elevation model (DEM) based hydrological connectivity as streamlines
385 flowing from each of the desert site locations where pupfish were present. We find that
386 during periods of occasionally intense or sustained rainfall, the desert ponds may become
387 hydrologically connected to their downstream saltwater location through tributary and
388 main stem water flow. Four distinct areas of potential hydrological connectivity can be
389 determined from Figure 2. In the northern portion of our study region, water from DP1
390 and DP2 are hydrologically linked via defined streamlines to SW1. SW2 serves as a
391 regional seawater sample pair for SW1, as it is separated by a distance of approximately
392 40 km. Further south, there is delineated connectivity between DP3 and SW3, as well as
393 between DP4 and SW4. In the southernmost part of the sampling area, DP6 has a
394 tributary stream that connects with SW5, while DP5 has a separate watercourse to the
395 sea. It is important to note that these streamlines do not represent active flow paths, as
396 Saudi Arabia has no permanent rivers. Instead, the streamlines describe either
397 permanently or intermittently dry riverbeds, defined as a function of the topography, with
398 ephemeral flows only occurring in cases of sufficient rainfall.
399
400 *RAD-sequencing*
401 In order to verify the potential hydrological connectivity for the Arabian pupfish, we use
402 Single Nucleotide Polymorphism (SNPs) genetic markers to evaluate the genetic
403 population connectivity of *Aphanius dispar*. For this we genotyped between 28 and 30
404 individuals per site for the eleven *A. dispar* locations by means of RAD-sequencing
405 (Table 2 & Supplementary Table 2). Over 2 million sequence reads were obtained on

406 average for each individual. Six samples were removed from further analysis as they had
407 fewer than 300,000 reads after demultiplexing and quality trimming, resulting in 27 to 30
408 individual samples per location. The total final sample number was 314 individuals for 11
409 sites (five saltwater and six desert pond localities) (Figure 1). A total of 690,084 putative
410 single nucleotide polymorphisms (SNPs) were obtained with STACKS (Catchen et al.,
411 2013), which were stringently filtered to a final 5,955 SNP loci. The average depth of
412 coverage for these loci was 48x, with a minimum of 24x and an average minor allele
413 frequency of 0.21 (Supplementary Figures 4 & 5).
414
415 *Population genomics*
416 The number of alleles across the 5,955 SNP loci ranged from 1.3 to 1.8 for the different
417 locations (Table 2). Heterozygosity revealed lower levels for desert ponds and
418 particularly for desert pond DP4 the lowest heterozygosity among all locations, while
419 also having a negative inbreeding coefficient (F_{IS}). DP4 also exhibited the most private
420 alleles (27) followed by DP1 with 15 and SW2 with 13. Three seawater locations SW3-
421 SW5 had no private alleles. Pairwise genetic distance between populations (F_{ST}) was
422 highest for DP4, with almost all values above 0.2 and the closest genetic distance with
423 SW4 (Figure 3). The lowest value (0) was observed for DP1 and DP2. Interestingly, most
424 DP locations showed higher F_{ST} values when compared to other DP localities than with
425 SW locations, indicating a clear disconnection between desert pond sites. Exceptions to
426 this include DP5 and DP6, with an F_{ST} value of 0.048. Pairwise F_{ST} comparisons between
427 seawater localities ranged between 0.005 to a maximum of 0.105 (SW2 vs SW5).

Effective population size ranged between a very low 51 in DP4 to nearly 6,000 in SW2 (Table 2). DP3 was the only other site with an effective size below 1,000.

Bayesian clustering and *ad hoc* testing allowed for the approximation of the most likely number of clusters within the analyzed samples. The most probable cluster numbers were 2 or 8, but delta K *ad hoc* testing also showed a peak at K=4 (Supplementary Figure 6). K=2 divides the northern localities (DP1, DP1& SW1, SW2) from the southern sites, with DP3 and SW3 being an admixed group. When considering four different genetic units in our data set, a division from north to south is found, combining desert ponds and seawater locations together into clusters (Figure 4). In this scenario, DP4 stands out as the only locality with its own cluster. SW4 in turn seems to have the most genetically admixed individuals of all locations. By increasing the cluster number to 8, a further subdivision in the northern part can be found, separating the two desert ponds from the seawater locations (DP1 & DP2 and SW1 and SW2). Here SW4 and SW5, the two furthest southern seawater sites, appear to share some unique traits (represented by the yellow cluster in Figure 4b), which are not shared with the southern desert ponds.

Hydrological and genetic connectivity

The population clustering in the principal component analysis revealed a break between the four northern locations (SW1, SW2, DP1 and DP2) and the seven southern ones, regardless of them being desert ponds or seawater sites (Figure 5a). The two desert ponds DP1 and DP2 closely clustered together as well as the two seawater locations (SW1 and SW2), with some distance between the two different environments. This

disjunction, however, is not seen through hydrological mapping, as DP1 & DP2 have hydrological connectivity to SW1.

Since annual rainfall volume tends to increase from the north towards the south (i.e. heavier and more frequent rainfalls towards the south (El Kenawy & McCabe, 2016)), the likelihood of genetic connectivity might also be expected to increase amongst the southern sample sites. In fact, with the exception of DP4, a decrease in genetic distance between locations from the northern area to the south one was found. Moreover, in the southern parts of the study area, less genetic distance was found between desert pond and the nearest seawater location (DP3 with DP4; DP4 with SW4 as well as DP5, DP6 and SW5). This is in accordance with the hydrological map, where water flow is possible from the desert pond sites to the respective wadi (ephemeral river systems appearing during intense rain periods) outlets in the Red Sea. It is not unusual for significant flow events or even flash-floods to occur on an annual, or multi-annual basis (Deng et al., 2015), potentially providing the mechanism behind the observed genetic connectivity within the population clusters. When evaluating the third principal component, large genetic distance between DP4 and most other locations becomes apparent (Figure 5b).

Putative environmental selection

Possible selective processes to different environmental condition were evaluated using an F_{ST} outlier approach. All desert pond locations were grouped together and compared with all seawater locations, resulting in five outlier loci (Table 3). Based on the population genetic analyses, subsets of desert ponds and genetically close seawater locations were

473 evaluated separately as well. In the northern part, the analysis included two desert ponds
474 (Desert Pond (DP)1 & DP2) and two seawater locations together (Sea Water (SW) 1 &
475 SW2), and resulted in one outlier loci which was also detected during an analysis that
476 included all sites. The comparison of DP4 and SW4 resulted also in one outlier.
477 Comparing DP3 and SW3 recovered one locus only, whereas the three southernmost
478 locations exhibited ten outliers. There is no common overlapping outlier loci detected
479 among the independent salt versus desert pond comparisons. Out of the 21 unique outliers
480 detected in the different subsets of data, 10 could be successfully identified by homology
481 using BLAST, and include several genes involved in immune response and ion channels.
482 The list of NCBI description and accession number can be found in Table 3.
483
484 *Muscle tissue stable isotopes*
485 To investigate further into differences between locations and environmental influence on
486 *A. dispar* we analyzed nitrogen and carbon stable isotopes in fish white muscle tissues.
487 These analyses revealed a clear separation between pupfish from inland ponds and
488 saltwater Red Sea habitats (Figure 6). The mean values for $\delta^{13}C$ and $\delta^{15}N$ for desert pond
489 fish (DP) ranged between -21.49‰ and -26.28‰ (SD: 2.39‰ to 3.82‰) and 12.63‰ and
490 21.03‰ (SD: 1.62‰ to 4.31‰), respectively. Saltwater fish (SW) isotope mean values
491 ranged between -6.13‰ and -11.62‰ for $\delta^{13}C$ (SD: 1.21‰ to 1.58‰) and between
492 5.92‰ and 10.90‰ for $\delta^{15}N$ (SD: 0.61‰ to 1.25‰). Overall, DP sites were more
493 depleted in $\delta^{13}C$ than SW, but had the highest values of $\delta^{15}N$. The isotopic signatures of
494 $\delta^{13}C$ and $\delta^{15}N$ correlated significantly ($p < 0.0001$, $\rho = -0.85$). While gender did not result
495 in a significant covariate ($p > 0.05$), fish length was significantly correlated with both

496 $\delta^{13}C$ ($p < 0.0001$, $\rho = -0.39$) and $\delta^{15}N$ ($p = 0.0002$, $\rho = 0.32$) values. Our data did not
497 pass the normality test (Shapiro-Wilk test, $p < 0.0001$) nor the homoscedasticity test
498 (Fligner-Killeen test, $\delta^{13}C$ $p = 0.0003$; $\delta^{15}N$ $p < 0.0001$). For these reasons, to test the
499 differences between DP and SW isotope contents correcting for fish size, a nonparametric
500 ANCOVA was used. There was a strong water type effect on both $\delta^{13}C$ and $\delta^{15}N$ fish
501 content ($p < 0.0001$). When testing the differences within clusters (cluster 1: $\delta^{13}C$ $p =$
502 0.0037, $\delta^{15}N$ $p = 0.0039$; cluster 2: $\delta^{13}C$ $p = 0.0385$, $\delta^{15}N$ $p = 0.0352$; cluster 3: $\delta^{13}C$ p
503 $= 0.0370$; $\delta^{15}N$ $p = 0.0350$; cluster 4: $\delta^{13}C$ $p = 0.0121$; $\delta^{15}N$ $p = 0.0141$), the same result
504 was obtained, with a strong segregation between saltwater and desert pond samples,
505 irrespective of geographic proximity. When comparing between desert ponds, DP3 is
506 significantly different from all other ponds for $\delta^{13}C$ ($p = 0.0064$) and DP4 has a
507 significantly larger $\delta^{15}N$ than the other ponds ($p=0.0239$). Muscle tissue isotopes
508 analysis therefore suggests that fish in the desert pond have different diets and therefore
509 different ecological niches relative to the populations from the Red Sea coast and some
510 desert ponds have distinct isotope signatures in comparison to other desert ponds.
511

Discussion

Our sampling of *Aphanius dispar* along the Red Sea coastline of the Arabian Desert provides insight to previously uncharacterized sites hosting populations in both near-freshwater water ponds as well as in sheltered coastal marine habitats. In contrast to the often ephemeral habitats occupied by the African turquoise killifish, *Nothobranchius furzeri* (Furness, 2016), the survival and persistence of the Arabian pupfish seems to depend on constant groundwater discharge and irregular rainfall events. Hydrological mapping and genetic analyses reveal that due to sporadic flash floods occurring in the region (Deng et al., 2015) fish may literally be washed from desert ponds out to the Red Sea. The genetic population units comprised an admixture of pupfish from the two different environments, desert pond and seawater individuals. We detected four main genetic clusters, which mostly grouped desert fish with Red Sea individuals based on a latitudinal gradient. Due to little hydrological connectivity among desert ponds, pupfish from the desert (or 'oases') are alternately rapidly carried out to very different environmental conditions in the Red Sea along wadis during flash flood events. The change in environmental condition is though not accompanied by population structure, contrary to many other study systems across different environmental gradients with populations differentiating with distance or environment (Sexton et al., 2014).

Ecological acclimation

Surviving a 'washout' event from a desert pond with less than 1 ppt in salinity to the highly saline Red Sea (with an average salinity of 43 ppt) requires a considerable acclimation capacity. Many killifish species are euryhaline and can tolerate large salinity

535 changes in the environment. *Fundulus heteroclitus* has been shown to undergo a large
536 range of physiological changes, from drinking rates to changes in acid base regulation
537 (Wood & Marshall, 1994). Nonetheless, *Fundulus* species, albeit tolerant to large salinity
538 ranges, have diverged into species adapted to specific environmental conditions
539 (Whitehead, 2010). Interestingly, we found the same genetic populations of *A. dispar*
540 living in both desert ponds and seawater. The hypothesized saltwater evolutionary origin
541 of the *Aphanius* genus[19] provides a potential explanation for the tolerance of *A. dispar*
542 when moved from near-freshwater ponds to a highly saline environment. However, it is
543 not just temperature and salinity that differ between the desert ponds and the Red Sea
544 coastal environments. We also see an ecological divergence in tissue stable isotopes
545 between desert and seawater sites, even from within the same genetic population unit.
546 The seawater fish muscle tissues were depleted in ^{15}N and more enriched in ^{13}C than their
547 freshwater counterparts. Similar signatures have previously been found when comparing
548 freshwater and saltwater fish, although the fish species in each of the environments
549 differed (Fuller et al., 2012; Robson et al., 2016). Similarly, for anadromous species such
550 as salmon that travel from fresh to saltwater, an increase in ^{13}C is exhibited when resident
551 in seawater habitats (Litz et al., 2017). Higher levels of ^{15}N, found here for the desert
552 ponds, is often correlated with higher trophic hierarchy levels (Zanden & Rasmussen,
553 1999). For *A. dispar*, however, it is difficult to isolate the potential influences of habitat-
554 related variability in the basal isotopic values of the food web from the potential trophic
555 shifts (McMahon, Hamady, & Thorrold, 2013; Post, 2002).
556
557 *Genetic divergence*

Environmental differences leading to ecological divergence can drive adaptation and eventually speciation (Arnegard et al., 2014). Despite large ecological difference in prey and habitat niche, there is little genetic difference between the desert pond fish and the geographically-corresponding seawater fish. This could give a hint that 'washout events' occur frequently enough to create enough migration from the desert ponds to the seawater sites, or that, as previously shown in other killifish species (Whitehead, Roach, Zhang, & Galvez, 2012), the Arabian killifish has a large capacity for rapid and long-lasting plastic responses to environmental change. We found only five putative outlier single nucleotide loci that could be under selection between all desert pond inhabitants and seawater fish, and only three that can be functionally annotated. Perforin 1 (*prf1*), a gene to which one of the outliers was successfully blasted (see Table 3), has key functions in the immune response and forms part of killer T-cells, indicating a potential adaptive change to the highly saline environment in the Red Sea. This particular function has been seen to be conserved in many fish species (Nakanishi, Toda, Shibasaki, & Somamoto, 2011; Toda, Araki, Moritomo, & Nakanishi, 2011). Interestingly, in wild salmon, perforin-mediated apoptotic processes were important in survival when migrating back to freshwater spawning grounds from seawater (Miller et al., 2011). Environmental changes are also associated with a SNP in the ATPase Family gene 3 (*afg3*) previously attributed to stress and biosynthesis and with an increased expression in trout after a starvation period (Rescan et al., 2007). A hint towards its importance in seawater acclimation could be related to an increase in mitochondrion-rich cells and ATPase activity due to a physiological demand of acid-base and ion regulation in saltwater (Lee, Hwang, Shieh, & Lin, 2000).

One of the desert ponds, the Al Lith hot spring (site DP3), had high temperatures (>38^{o}C) and a different chemical signature, such as a high amount of silica. Life in such hot water could potentially provoke adaptive signals in the genome. However, only one putative outlier was found for the inhabitants of the hot spring in comparison to other desert ponds. therefore hinting to phenotypic plasticity, as for the case of the Magadi tilapia (*Alcolapia grahami*). *Alcolapia grahami* lives in hot springs in Kenya and was recently described to have the highest upper critical temperature recorded for a fish (45.6 $^{\circ}$C (Wood et al., 2016)). Despite the extreme environmental conditions, genetic studies did not find population differences when compared to tilapia of less extreme environmental conditions at close proximity (Wilson et al., 2004; Zaccara et al., 2014). Low numbers of putative outlier SNPs were also detected, suggesting some ongoing gene flow and admixture (Ford et al., 2015), which could be the case for our *A. dispar* samples from the Al Lith hot spring. Despite the lack of loci under selection, connectivity between desert ponds is low and there is limited gene flow, as indicated by large genetic distances between the hot spring and other desert ponds, suggesting isolation and divergence between these habitats. For the pupfish in the hot spring, we could detect a differentiated ecological signature in tissue stable isotopes compared to the other desert ponds. Although contradicting the findings on the Magadi tilapia, where the hot spring site revealed carbon isotope tissue enrichment (Kavembe, Kautt, Machado-Schiaffino, & Meyer, 2016), *A. dispar* from the Al Lith hot spring have lower values of ^{13}C in their muscle tissue. One possible explanation for this result might be the turbidity of the hot

603 spring site, due to a large amount of suspended silica, as carbon stable isotope depletion
604 was previously associated with turbidity in the environment (Nahon et al., 2013).
605
606 *Anthropogenic impacts on desert fish populations*
607 Besides natural environmental conditions that are reflected in the isotopic signature of the
608 pupfish, anthropogenic impacts can also be detected. In the case of DP4, we found an
609 isotopic as well as genetic signature of human disturbance. Fish tissues here are more ^{15}N
610 enriched in comparison to other locations, which indicates an accumulation of the heavier
611 isotope element possibly due to isolation of this pond (Amundson et al., 2003; Szpak,
612 2014). In this site, there is evidence of agricultural activity, most likely utilizing
613 groundwater in the area, which in turn might lower the water table and disconnect this
614 location from others. Furthermore, there is a dam structure 15 km upstream of DP4 most
615 likely restricting the water availability downstream. Even if topographic mapping shows
616 a hydrological connection between DP4 and SW4, it seems that human interference here
617 inhibits any flow of water to the sea. This hypothesis would explain the genetic
618 differentiation for this particular site, which seems to be undergoing a population
619 bottleneck. The fish in this site have in fact an increased number of private alleles and
620 low genetic diversity, indicated by low heterozygosity. Even though the inbreeding
621 coefficient was is low, this result is most likely due to low genetic differentiation within
622 the population used in the calculation. Pairwise kinship reveals most individuals within
623 this site to be highly related, and the effective population size is very low. It therefore
624 seems that use of water in this area has isolated this population, restricting its gene flow
625 and its resilience.

626

627 In the northern part of the Saudi Arabian coast, *A. dispar* in desert ponds and saltwater
628 locations do not cluster together genetically, as it is seen in the southern parts, albeit
629 hydrological connectivity potential. Here the desert pond fish are genetically closer to
630 each other, as are the seawater fish, with a clear division between desert and Red Sea
631 sample sites. There are two plausible reasons for this disconnection. The northern regions
632 receive less rainfall (El Kenawy & McCabe, 2016) and hence any hydrological
633 connectivity between the desert and the sea will be much lower. The observed strong
634 genetic divergence though is most likely caused by the 'upstream' construction of the
635 Rabigh Dam (completed 2008), which was built for municipal water supply and flood
636 control. Hence, even with rain events the water, and therefore the pupfish, can no longer
637 reach the sea. For this reason, *A. dispar* populations are now diverging without the
638 presence of gene flow through new migrants. A similar anthropogenic impact was found
639 for desert fish of the Colorado River area, where natural flooding occurred regularly until
640 the construction of dams that drastically changed the water availability and had a major
641 impact on the distribution of desert fish (Hillyard, Podrabsky, & van Breukelen, 2015).

642

643 Fish living in desert regions have long been a conservation concern, with a large number
644 of such species being under threat or endangered, often due to the expansion of desert
645 agriculture and increasing global temperatures (Van Haverbeke et al., 2013). Although
646 the Dead Sea subspecies *(A. dispar richardsoni)* is considered endangered (Goren, 2014),
647 *A. dispar* itself is not considered to be endangered because it has stable populations
648 widely distributed throughout the Arabian region (Freyhof, 2014). The species' capacity

649 to acclimate and survive challenging environmental fluctuations likely plays a major role
650 in its success in this region. However, our results show that despite the large capacity of
651 *A. dispar* to acclimate and adapt to different environments and defy the constraints of
652 living in restricted desert environments, anthropogenic water use can dramatically alter
653 the population dynamics of the Arabian pupfish.
654

References

Al-Kahem-Al-Balawi, H. F., Al-Ghanim, K. A., Ahmad, Z., Temraz, T. A., Al-Akel, A. S., Al-Misned, F., & Annazri, H. (2008). A threatened fish species (Aphanius dispar) in Saudi Arabia, a case study. *Pakistan Journal of Biological Sciences : PJBS*, *11* (19), 2300–2307. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19137861

Amundson, R., Austin, A. T., Schuur, E. A. G., Yoo, K., Matzek, V., Kendall, C., … Baisden, W. T. (2003). Global patterns of the isotopic composition of soil and plant nitrogen. *Global Biogeochemical Cycles*, *17* (1). https://doi.org/10.1029/2002GB001903

Arnegard, M. E., McGee, M. D., Matthews, B., Marchinko, K. B., Conte, G. L., Kabir, S., … Schluter, D. (2014). Genetics of ecological divergence during speciation. *Nature*, *511* (7509), 307–311. https://doi.org/10.1038/nature13301

Bennett, W. A., & Beitinger, T. L. (1997). Temperature Tolerance of the Sheepshead Minnow, Cyprinodon variegatus. *Copeia*, *1*, 77. https://doi.org/10.2307/1447842

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)*, *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bouzat, J. L. (2010). Conservation genetics of population bottlenecks: the role of chance, selection, and history. *Conservation Genetics*, *11* (2), 463–478. https://doi.org/10.1007/s10592-010-0049-0

Bowman, A., & Azzalini, A. W. (2014). R package “sm”: nonparametric smoothing methods (version 2.2-5.4). Retrieved from http://www.stats.gla.ac.uk/~adrian/sm, http://azzalini.stat.unipd.it/Book_sm

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology*, *22* (11), 3124–3140. https://doi.org/10.1111/mec.12354

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., … Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27* (15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Deng, L., McCabe, M. F., Stenchikov, G., Evans, J. P., Kucera, P. A., Deng, L., … Kucera, P. A. (2015). Simulation of Flash-Flood-Producing Storm Events in Saudi Arabia Using the Weather Research and Forecasting Model. *Journal of Hydrometeorology*, *16* (2), 615–630. https://doi.org/10.1175/JHM-D-14-0126.1

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014). NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Molecular Ecology Resources*, *14* (1), 209–214. https://doi.org/10.1111/1755-0998.12157

Doebeli, M., & Dieckmann, U. (2003). Speciation along environmental gradients. *Nature*, *421* (6920), 259–264. https://doi.org/10.1038/nature01274

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4* (2), 359–361. https://doi.org/10.1007/s12686-011-9548-7

Ehrich, D., Tarroux, A., Stien, J., Lecomte, N., Killengreen, S., Berteaux, D., & Yoccoz,

N. G. (2011). Stable isotope analysis: modelling lipid normalization for muscle and eggs from arctic mammals and birds. *Methods in Ecology and Evolution*, *2* (1), 66–76. https://doi.org/10.1111/j.2041-210X.2010.00047.x

El Kenawy, A. M., & McCabe, M. F. (2016). A multi-decadal assessment of the performance of gauge- and model-based rainfall products over Saudi Arabia: climatology, anomalies and trends. *International Journal of Climatology*, *36* (2), 656–674. https://doi.org/10.1002/joc.4374

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, *14* (8), 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Fisher, S. G., Gray, L. J., Grimm, N. B., & Busch, D. E. (1982). Temporal Succession in a Desert Stream Ecosystem Following Flash Flooding. *Ecological Monographs*, *52* (1), 93–110. https://doi.org/10.2307/2937346

Foll, M., & Gaggiotti, O. (2008). A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective. *Genetics*, *180* (2). Retrieved from http://www.genetics.org/content/180/2/977

Ford, A. G. P., Dasmahapatra, K. K., Rüber, L., Gharbi, K., Cezard, T., & Day, J. J. (2015). High levels of interspecific gene flow in an endemic cichlid fish adaptive radiation from an extreme lake environment. *Molecular Ecology*, *24* (13), 3421–3440. https://doi.org/10.1111/MEC.13247

Freyhof, J. (2014). Aphanius dispar.

Fuller, B. T., Müldner, G., Van Neer, W., Ervynck, A., Richards, M. P., Cooremans, B., … Richards, M. P. (2012). Carbon and nitrogen stable isotope ratio analysis of freshwater, brackish and marine fish from Belgian archaeological sites (1st and 2nd millennium AD). *Journal of Analytical Atomic Spectrometry*, *27* (5), 807. https://doi.org/10.1039/c2ja10366d

Furness, A. I. (2016). The evolution of an annual life cycle in killifish: adaptation to ephemeral aquatic environments through embryonic diapause. *Biological Reviews*, *91* (3), 796–812. https://doi.org/10.1111/brv.12194

Gaggiotti, O. E. (2003). Genetic threats to population persistence. *Ann. Zool. Fennici*, *40* (40), 155–168. Retrieved from http://www.annzool.net/PDF/anz40-free/anz40-155.pdf

Gat, J. R. (1996). Oxygen and hydropgen iostopes in the hydrologic cycle. *Annual Review of Earth and Planetary Sciences*, *24* (1), 225–262. https://doi.org/10.1146/annurev.earth.24.1.225

Gibson, J. J., Birks, S. J., & Edwards, T. W. D. (2008). Global prediction of δ_A and δ^2 H- δ^{18} O evaporation slopes for lakes and soil water accounting for seasonality. *Global Biogeochemical Cycles*, *22* (2), n/a-n/a. https://doi.org/10.1029/2007GB002997

Gibson, J. J., & Reid, R. (2014). Water balance along a chain of tundra lakes: A 20-year isotopic perspective. *Journal of Hydrology*, *519*, 2148–2164. https://doi.org/10.1016/j.jhydrol.2014.10.011

Goren, M. (2014). Aphanius dispar ssp. richardsoni.

Haas, R. (1982). Notes on the ecology of Aphanius dispar (Pisces, Cyprinodontidae) in the Sultanate of Oman. *Freshwater Biology*, *12* (1), 89–95. https://doi.org/10.1111/j.1365-2427.1982.tb00605.x

Hillyard, S. D., Podrabsky, J. E., & van Breukelen, F. (2015). Desert Environments. In *Extremophile Fishes* (pp. 59–83). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-13362-1_4

Hopken, M. W., Douglas, M. R., & Douglas, M. E. (2013). Stream hierarchy defines riverscape genetics of a North American desert fish. *Molecular Ecology*, *22* (4), 956–971. https://doi.org/10.1111/mec.12156

Hrbek, T., & Meyer, A. (2003). Closing of the Tethys Sea and the phylogeny of Eurasian killifishes (Cyprinodontiformes: Cyprinodontidae). *Journal of Evolutionary Biology*, *16* (1), 17–36. https://doi.org/10.1046/j.1420-9101.2003.00475.x

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23* (14), 1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24* (11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Kavembe, G. D., Franchini, P., Irisarri, I., Machado-Schiaffino, G., & Meyer, A. (2015). Genomics of Adaptation to Multiple Concurrent Stresses: Insights from Comparative Transcriptomics of a Cichlid Fish from One of Earth's Most Extreme Environments, the Hypersaline Soda Lake Magadi in Kenya, East Africa. *Journal of Molecular Evolution*, *81* (3–4), 90–109. https://doi.org/10.1007/s00239-015-9696-6

Kavembe, G. D., Kautt, A. F., Machado-Schiaffino, G., & Meyer, A. (2016). Eco-morphological differentiation in Lake Magadi tilapia, an extremophile cichlid fish living in hot, alkaline and hypersaline lakes in East Africa. *Molecular Ecology*, *25* (7), 1610–1625. https://doi.org/10.1111/mec.13461

Kelley, J. L., Peyton, J. T., Fiston-Lavier, A.-S., Teets, N. M., Yee, M.-C., Johnston, J. S., … Denlinger, D. L. (2014). Compact genome of the Antarctic midge is likely an adaptation to an extreme environment. *Nature Communications*, *5*, ncomms5611. https://doi.org/10.1038/ncomms5611

Lee, T. H., Hwang, P. P., Shieh, Y. E., & Lin, C. H. (2000). The relationship between `deep-hole' mitochondria-rich cells and salinity adaptation in the euryhaline teleost, Oreochromis mossambicus. *Fish Physiology and Biochemistry*, *23* (2), 133–140. https://doi.org/10.1023/A:1007818631917

Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28* (2), 298–299. https://doi.org/10.1093/bioinformatics/btr642

Litz, M. N. C., Miller, J. A., Copeman, L. A., Teel, D. J., Weitkamp, L. A., Daly, E. A., & Claiborne, A. M. (2017). Ontogenetic shifts in the diets of juvenile Chinook Salmon: new insight from stable isotopes and fatty acids. *Environmental Biology of Fishes*, *100* (4), 337–360. https://doi.org/10.1007/s10641-016-0542-5

Martin, C. H., Crawford, J. E., Turner, B. J., & Simons, L. H. (2016). Diabolical survival in Death Valley: recent pupfish colonization, gene flow and genetic assimilation in the smallest species range on earth. *Proceedings of the Royal Society B: Biological Sciences*, *283* (1823), 20152334. https://doi.org/10.1098/rspb.2015.2334

McMahon, K. W., Hamady, L. L., & Thorrold, S. R. (2013). A review of ecogeochemistry approaches to estimating movements of marine animals.

Limnology and Oceanography, *58* (2), 697–714. https://doi.org/10.4319/lo.2013.58.2.0697

Meffe, G. K., & Vrijenhoek, R. C. (1988). Conservation Genetics in the Management of Desert Fishes. *Conservation Biology*, *2* (2), 157–169. https://doi.org/10.1111/j.1523-1739.1988.tb00167.x

Meirmans, P. G., & Van Tienderen, P. H. (2004). Genotype and Genodive: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, *4* (4), 792–794. https://doi.org/10.1111/j.1471-8286.2004.00770.x

Miller, K. M., Li, S., Kaukinen, K. H., Ginther, N., Hammill, E., Curtis, J. M. R., … Farrell, A. P. (2011). Genomic Signatures Predict Migration and Spawning Failure in Wild Canadian Salmon. *Science*, *331* (6014), 214–217. https://doi.org/10.1126/science.1196901

Moore, T. Y., Cooper, K. L., Biewener, A. A., & Vasudevan, R. (2017). Unpredictability of escape trajectory explains predator evasion ability and microhabitat preference of desert rodents. *Nature Communications*, *8* (1), 440. https://doi.org/10.1038/s41467-017-00373-2

Nahon, S., Richoux, N. B., Kolasinski, J., Desmalades, M., Ferrier Pages, C., Lecellier, G., … Berteaux Lecellier, V. (2013). Spatial and Temporal Variations in Stable Carbon (δ13C) and Nitrogen (δ15N) Isotopic Composition of Symbiotic Scleractinian Corals. *PLoS ONE*, *8* (12), e81247. https://doi.org/10.1371/journal.pone.0081247

Nakanishi, T., Toda, H., Shibasaki, Y., & Somamoto, T. (2011). Cytotoxic T cells in teleost fish. *Developmental & Comparative Immunology*, *35* (12), 1317–1323. https://doi.org/10.1016/j.dci.2011.03.033

NASA Jet Propulsion Laboratory. (2009). ASTER Global Digital Elevation Model. https://doi.org/10.5067/ASTER/ASTGTM.002

Paul, D., Skrzypek, G., & Fórizs, I. (2007). Normalization of measured stable isotopic compositions to isotope reference scales – a review. *Rapid Communications in Mass Spectrometry*, *21* (18), 3006–3014. https://doi.org/10.1002/rcm.3185

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. *PLoS ONE*, *7* (5), e37135. https://doi.org/10.1371/journal.pone.0037135

Post, D. M. (2002). Using Stable Isotopes to Estimate Trophic Position: Models,Methods and Assumptions. *Ecology*, *83* (3), 703–718. https://doi.org/10.1890/0012-9658(2002)083[0703:USITET]2.0.CO;2

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155* (2), 945–959. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1461096&tool=pmcentrez&rendertype=abstract

Rescan, P.-Y., Montfort, J., Rallière, C., Le Cam, A., Esquerré, D., & Hugot, K. (2007). Dynamic gene expression in fish muscle during recovery growth induced by a fasting-refeeding schedule. *BMC Genomics*, *8* (1), 438. https://doi.org/10.1186/1471-2164-8-438

Robson, H. K., Andersen, S. H., Clarke, L., Craig, O. E., Gron, K. J., Jones, A. K. G., … Heron, C. (2016). Carbon and nitrogen stable isotope values in freshwater, brackish

838 and marine fish bone collagen from Mesolithic and Neolithic sites in central and
839 northern Europe. *Environmental Archaeology*, *21* (2), 105–118.
840 https://doi.org/10.1179/1749631415Y.0000000014
841 Rosenberg, N. A. (2003). distruct: a program for the graphical display of population
842 structure. *Molecular Ecology Notes*, *4* (1), 137–138. https://doi.org/10.1046/j.1471-
843 8286.2003.00566.x
844 Rousset, F. (2008). genepop'007: a complete re-implementation of the genepop software
845 for Windows and Linux. *Molecular Ecology Resources*, *8* (1), 103–106.
846 https://doi.org/10.1111/j.1471-8286.2007.01931.x
847 Seeb, L. W., Waples, R. K., Limborg, M. T., Warheit, K. I., Pascal, C. E., & Seeb, J. E.
848 (2014). Parallel signatures of selection in temporally isolated lineages of pink
849 salmon. *Molecular Ecology*, *23* (10), 2473–2485. https://doi.org/10.1111/mec.12769
850 Sexton, J. P., Hangartner, S. B., & Hoffmann, A. A. (2014). Genetic isolation by
851 environment or distance: which pattern of gene flow is most common? *Evolution*, *68*
852 (1), 1–15. https://doi.org/10.1111/evo.12258
853 Sharp, Z. (2007). *Principles of stable isotope geochemistry*. Pearson/Prentice Hall.
854 Szpak, P. (2014). Complexities of nitrogen isotope biogeochemistry in plant-soil
855 systems: implications for the study of ancient agricultural and animal management
856 practices. *Frontiers in Plant Science*, *5*, 288.
857 https://doi.org/10.3389/fpls.2014.00288
858 Toda, H., Araki, K., Moritomo, T., & Nakanishi, T. (2011). Perforin-dependent cytotoxic
859 mechanism in killing by CD8 positive T cells in ginbuna crucian carp, Carassius
860 auratus langsdorfii. *Developmental & Comparative Immunology*, *35* (1), 88–93.
861 https://doi.org/10.1016/j.dci.2010.08.010
862 Van Haverbeke, D. R., Stone, D. M., Coggins, L. G., & Pillow, M. J. (2013). Long-Term
863 Monitoring of an Endangered Desert Fish and Factors Influencing Population
864 Dynamics. *Journal of Fish and Wildlife Management*, *4* (1), 163–177.
865 https://doi.org/10.3996/082012-JFWM-071
866 Wang, I. J., & Bradburd, G. S. (2014). Isolation by environment. *Molecular Ecology*, *23*
867 (23), 5649–5662. https://doi.org/10.1111/mec.12938
868 Whitehead, A. (2010). The evolutionary radiation of diverse osmotolerant physiologies
869 in killifish (Fundulus sp.). *Evolution*. https://doi.org/10.1111/j.1558-
870 5646.2010.00957.x
871 Whitehead, A., Galvez, F., Zhang, S., Williams, L. M., & Oleksiak, M. F. (2011).
872 Functional genomics of physiological plasticity and local adaptation in killifish. *The*
873 *Journal of Heredity*, *102* (5), 499–511. https://doi.org/10.1093/jhered/esq077
874 Whitehead, A., Roach, J. L., Zhang, S., & Galvez, F. (2012). Salinity- and population-
875 dependent genome regulatory response during osmotic acclimation in the killifish (
876 Fundulus heteroclitus) gill, 1293–1305. https://doi.org/10.1242/jeb.062075
877 Wilson, P., Wood, C., Walsh, P., Bergman, A., Bergman, H., Laurent, P., & White, B.
878 (2004). Discordance between Genetic Structure and Morphological, Ecological, and
879 Physiological Adaptation in Lake Magadi Tilapia. *Physiological and Biochemical*
880 *Zoology*, *77* (4), 537–555. https://doi.org/10.1086/422054
881 Wood, C. M., Brix, K. V, De Boeck, G., Bergman, H. L., Bianchini, A., Bianchini, L. F.,
882 … Ojoo, R. O. (2016). Mammalian metabolic rates in the hottest fish on earth.
883 *Scientific Reports*, *srep26990*. https://doi.org/10.1038/srep26990

884 Wood, C. M., & Marshall, W. S. (1994). Ion Balance, Acid-Base Regulation, and
885 Chloride Cell Function in the Common Killifish, Fundulus heteroclitus: A
886 Euryhaline Estuarine Teleost. *Estuaries*, *17* (1), 34. https://doi.org/10.2307/1352333
887 Zaccara, S., Crosa, G., Vanetti, I., Binelli, G., Harper, D. M., Mavuti, K. M., … Britton,
888 J. R. (2014). Genetic and morphological analyses indicate high population mixing
889 in the endangered cichlid Alcolapia flock of East Africa. *Conservation Genetics*, *15*
890 (2), 429–440. https://doi.org/10.1007/s10592-013-0551-2
891 Zanden, M. J. Vander, & Rasmussen, J. B. (1999). Primary consumer□? 13 C and? 15 N
892 and the trophic position of aquatic consumers. *Ecology*, *80* (4), 1395–1404.
893 https://doi.org/10.1890/0012-9658 (1999)080[1395:PCCANA]2.0.CO;2
894 Zanden, M. J. Vander, & Rasmussen, J. B. (2001). Variation in δ^{15} N and δ^{13} C trophic
895 fractionation: Implications for aquatic food web studies. *Limnology and*
896 *Oceanography*, *46* (8), 2061–2066. https://doi.org/10.4319/lo.2001.46.8.2061
897
898

Acknowledgments

This study was supported by the King Abdullah University of Science and Technology (KAUST). We are very grateful to the KAUST Coastal and Marine Resources Core Lab and KAUST Government Affairs for their support in finding desert ponds, obtaining permits and aiding in the field. We thank the KAUST Integrative Systems Biology Laboratory and the KAUST Biosciences Core Laboratory for support and assistance.

Author contributions: C.S. designed and managed the field collection. C.S, L.C.B & J.N. performed the sample collection. L.C.B. and J.N. with the supervision of C.S. prepared the samples for DNA sequencing and tissue isotope measurements. M.L.B. provided reagents. J.N. performed chemical water analyses and L.C.B. analyzed the tissue isotope data. S.D.P., Y.A.L. and M.F.M. created hydrological mapping and analyzed water isotopes. C.S. analyzed the sequencing data, performed population genetic analyses with help from J.S. and integrated all datasets. C.S. & T.R. wrote the paper; all authors provided input to and approved the final version of the manuscript.

Additional information:

Raw sequencing data have been deposited on NCBI under BioProject ID PRJNA311159. Final SNP vcf file is available as Supplementary Material. Correspondence and requests for materials should be addressed to T.R. or C.S.

Competing interests

The authors declare no competing interests.

Figure legends

Figure 1: Locations of all sampling sites over a 1,100 km stretch of the Saudi Arabian coastline. Grey circles represent desert ponds without the presence of *Aphanius dispar*. Sites with pupfish are colored (n=11) and classified as desert ponds (DP) or seawater lagoons (SW). The top two pictures show a male and female *A. dispar*, respectively. The center picture is an example of a desert pond (DP2), below which is the Al Lith hotspring (DP3). The bottom picture shows an example of a seawater site (SW4). The map was created using ArcGIS 10.5 (www.arcgis.com).

Figure 2: Hydrological modeling. Latitudinal 2D overview of Saudi Arabian coastline with *Aphanius dispar* collection sites indicated with fish icons (left) and 3D close-ups on three regions with predicted hydrological flow. Maps were produced with Google Earth imagery (Google Earth 7.1.2.2041; December 31, 2016) and ArcGIS 10.5 software (www.arcgis.com).

Figure 3: Heatmap of F_{ST} values of pairwise population comparisons based on 5,955 SNPs. Location abbreviations as in Table 1 (DP=desert pond, SW=seawater).

Figure 4: *Aphanius dispar* population structure with a) K=4 and b) K=8. Each individual in the different sites (y-Axis) are plotted with colours representing the probability of each individual to be assigned to a certain cluster.

944 Figure 5: Scatterplot of principal component analysis evaluating the genetic structure
945 between all analyzed *Aphanius dispar* individuals based on 5,955 loci. Eigenvalues
946 represent the amount of genetic diversity shown by each principal component. a) PCA of
947 principal components 1 and 2. b) PCA of principal components 1 and 3.
948
949 Figure 6: Ratio of stable isotopes of carbon and nitrogen in muscle tissue of *Aphanius*
950 *dispar*, including fish from brackish water ponds and saltwater in the Red Sea. Error bars
951 represent the standard error per location
952

953 Table 1: Sampling locations, characteristics and selected water parameters (for full water
954 measurements refer to Supplementary Table 1) measured by CTD or Aquakem. A dash
955 indicates that no reliable measurement was obtained.

Site	Type	Temperature (°C)	Conductivity (µS/cm)	Salinity (ppt)	O2 (ppm)	pH	Chlorides (ppm)	Silica (ppb)	pupfish present
SW5	Sea	21.77	58,340.47	41.93	2.87	8.71	27.10	306.02	yes
SW4	Sea	27.90	66,952.99	42.75	3.39	9.01	29.59	314.61	yes
SW3	Sea	22.82	58,685.99	41.19	4.40	9.10	2.15	104.57	yes
SW2	Sea	24.80	64,966.91	44.20	3.43	8.37	36,677.27	517.96	yes
SW1	Sea	20.90	59,185.03	43.49	3.93	8.37	36.06	31.97	yes
DP6	Pond	31.42	1,663.76	0.74	7.75	9.41	282.41	34,511.49	yes
DP5	Pond	30.93	1,760.39	0.79	8.32	9.39	294.11	59,944.24	yes
DP4	Pond	26.44	2,004.49	0.99	6.14	8.31	278.57	63,358.87	yes
DP3	Hot Spring	38.42	1,916.49	0.75	5.16	8.20	361.21	80,469.67	yes
DP2	Pond	31.58	3,190.00	1.45	4.35	8.31	734.37	38,017.18	yes
DP1	Pond	29.93	1,720.00	0.78	4.35	8.50	469.77	33,586.98	yes
28	Sea	27.65	64,108.41	40.93	3.32	8.46	-	-	no
27	Sea	21.77	65,110.01	47.50	3.81	8.21	-	-	no
26	Sea	18.18	51,228.77	39.44	4.56	8.40	-	-	no
25	Sea	22.60	57,621.28	40.54	3.61	8.48	-	-	no
24	Sea	12.08	46,285.24	41.12	4.11	8.29	-	-	no
23	Sea	21.74	63,528.52	46.35	3.86	8.24	-	-	no
15	Pond	21.28	5,825.19	0.31	6.00	8.60	76.24	20,016.72	no
14	Pond	21.15	8,023.78	0.43	3.10	7.71	115.34	38,414.86	no
13	Pond	-	-	-	-	-	103.07	5,016.26	no
11	Pond	28.91	898.74	0.41	6.33	8.04	86.54	38,664.80	no
9	Pond	33.54	1,025.76	0.43	6.26	9.15	127.94	52,096.62	no
7	Pond	30.21	1,257.23	0.56	4.49	8.05	196.26	76,975.60	no
6	Pond	28.31	1,164.95	0.54	4.69	8.34	128.02	46,885.01	no
5	Pond	30.25	879.72	0.39	4.26	8.11	135.60	25,711.03	no
4	Pond	-	-	-	-	-	29.70	10,464.11	no
3	Pond	21.60	1,019.11	0.54	5.32	8.25	166.66	45,331.00	no

956

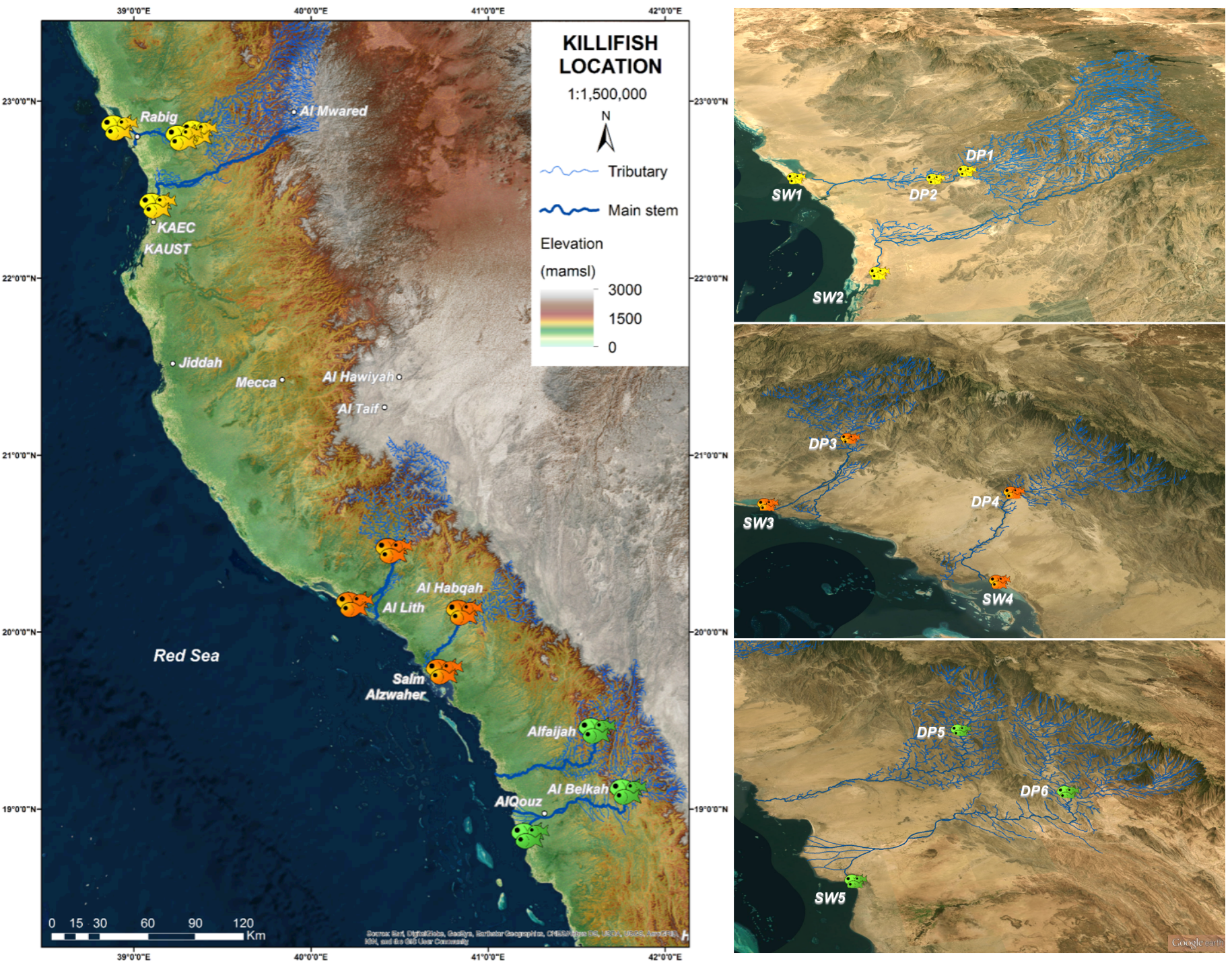
957

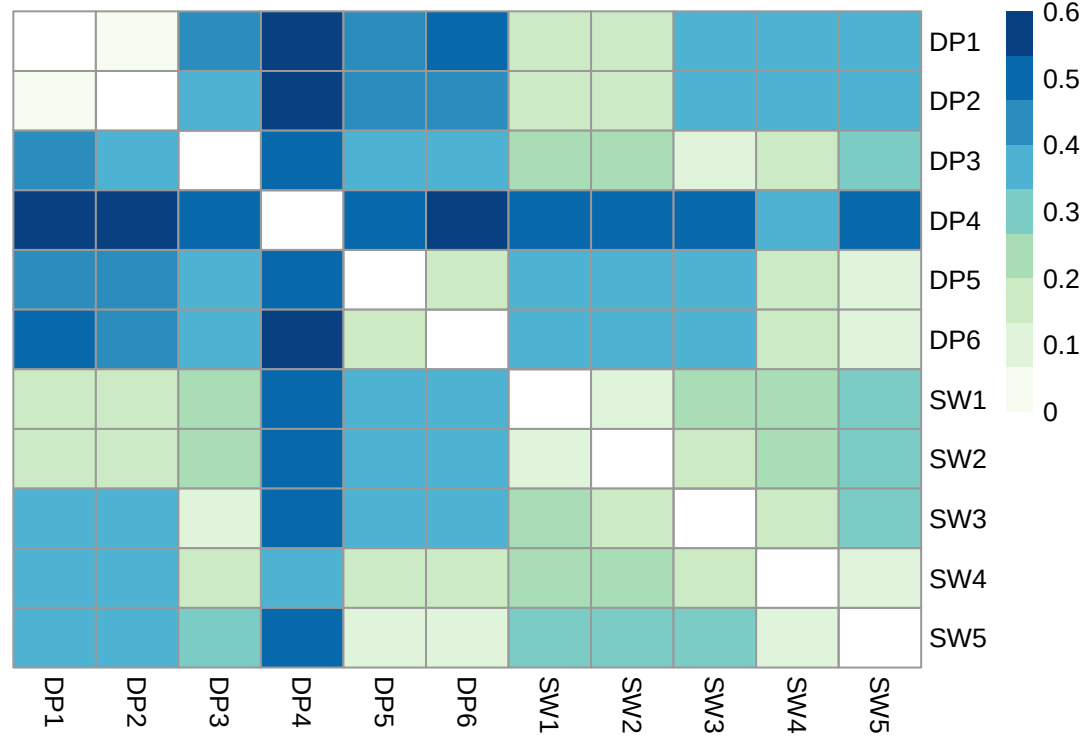
Table 2: Genetic metrics for all 11 pupfish locations for 5,955 SNP loci per sampling site. N = number of individuals, Private a = number of private alleles, Na = average number of alleles, Ho = observed heterozygosity, He = expected heterozygosity, F_{IS} = inbreeding coefficient, Ne = effective population size, 95% CI = 95% confidence intervals for Ne, mean r = average of pairwise kinship coefficient.

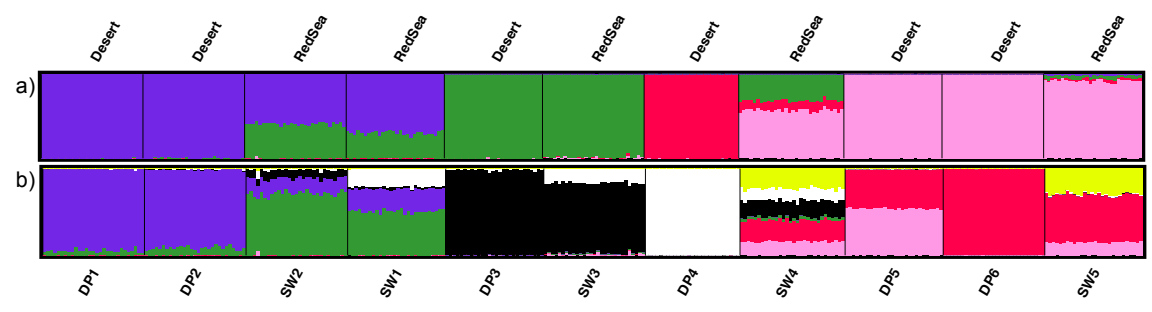
Location	N	Private a	Na	Ho	He	F_{IS}	Ne	95% CI	Mean r
DP1	29	15	1.74	0.21	0.23	0.07	1222	983-1614	0.22
DP2	29	8	1.78	0.24	0.25	0.05	1851	1256-3511	0.12
DP3	28	6	1.57	0.17	0.19	0.07	464	421-515	0.12
DP4	27	27	1.31	0.08	0.08	-0.02	51	49-52	0.20
DP5	28	4	1.53	0.16	0.17	0.06	1497	1084-2416	0.25
DP6	29	9	1.50	0.15	0.16	0.09	1713	1236-2782	0.13
SW1	28	7	1.80	0.24	0.27	0.08	1149	882-1646	0.27
SW2	29	13	1.80	0.23	0.26	0.09	5910	2229-inf	0.06
SW3	29	0	1.66	0.19	0.20	0.06	1702	1199-2925	0.15
SW4	30	0	1.77	0.20	0.22	0.10	1645	1284-2286	0.19
SW5	28	0	1.70	0.18	0.20	0.12	3917	2175-19493	0.42

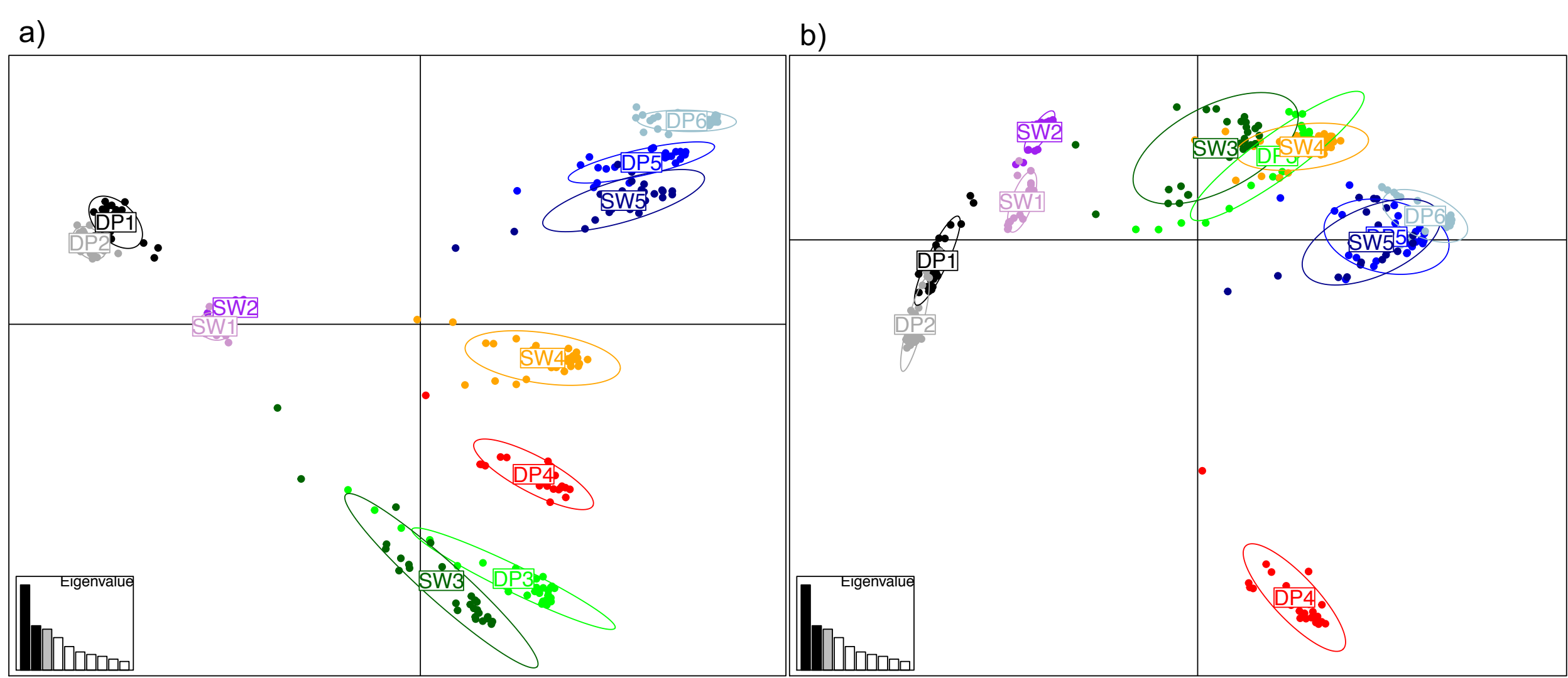
Table 3: Putative outlier loci between desert pond and red sea locations. Several pairwise comparisons were performed, and the compared locations are indicated under "Locations included". Sequences were blasted again NCBI and only successful blast hits are listed.

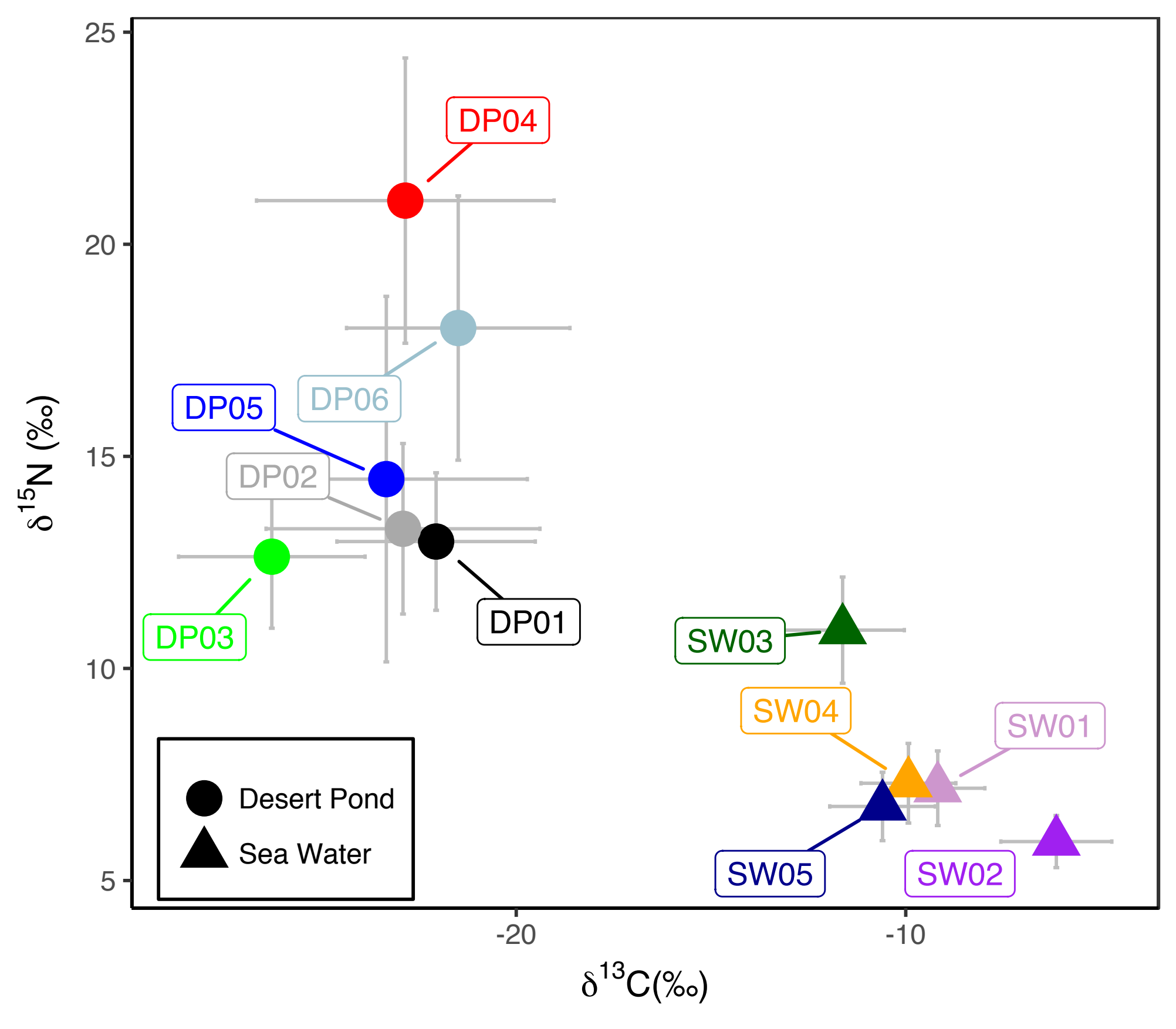
Locations included: DP	Locations included: SW	SNP loci included	Number of outliers	SNP name	NCBI description	Species of best blast hit	gene name	e-value	coverage	NCBI accession
DP all	**SW all**	4507	**5**	19937_75						
				35939_106	perforin-1-like	*Salmo salar*	perf	2.00E-05	86%	XM_014185448.1
				76967_58						
				90197_26	integrin alpha-6-like	*Poecilia formosa*	itga6	3.00E-10	79%	XM_016678011.1
				112486_90	AFG3-like protein 1	*Nothobranchius furzeri*	afg3	8.00E-30	95%	XM_015952175.1
DP1&DP2	**SW1&SW2**	2449	**1**	90197_30	integrin alpha-6-like	*Poecilia formosa*	itga6	3.00E-10	79%	XM_016678011.1
DP1&DP2	**SW1**	2586	**1**	126619_29	unchar. protein K02A2.6-like	*Oreochromis niloticus*		4.00E-08	75%	XM_019364105.1
DP1&DP2	**SW2**	2890	**6**	26511_91						
				33001_35						
				76967_58						
				90197_30	integrin alpha-6-like	*Poecilia formosa*	itga6	3.00E-10	79%	XM_016678011.1
				107780_39						
				120257_43	potassium voltage-gated channel subfamily H member 2-like	*Poecilia mexicana*	kcnh2	0.003	70%	XM_014994260.1
DP3	**SW3**	3704	**1**	116615_106						
DP4	**SW4**	2631	**1**	112486_90	AFG3-like protein 1	*Nothobranchius furzeri*		8.00E-30	95%	XM_015952175.1
DP5&DP6	**SW5**	3963	**10**	19073_13	immunoglobulin light chain genomic sequence	*Takifugu rubripes*		0.003	84%	KU365392.1
				20996_81						
				39387_14						
				62495_79						
				63109_75	v-rel avian reticuloendotheliosis viral oncogene homolog	*Nothobranchius furzeri*	rel	1.00E-21	82%	XM_015944097.1
				67796_71						
				68351_79	piezo type mechanosensitive ion channel component 1	*Nothobranchius furzeri*	piezo1	2.00E-05	73%	XM_015945536.1
				80013_20	zinc finger MIZ domain-containing protein 1-like	*Astyanax mexicanus*	zbtb17	3.00E-22	89%	XM_007235440.2
				83760_40	uncharacterized LOC103909955	*Danio rerio*		5.00E-07	81%	XM_017354508.1
				89871_81						
DP4	**SW all**	4761	**1**	19939_80						

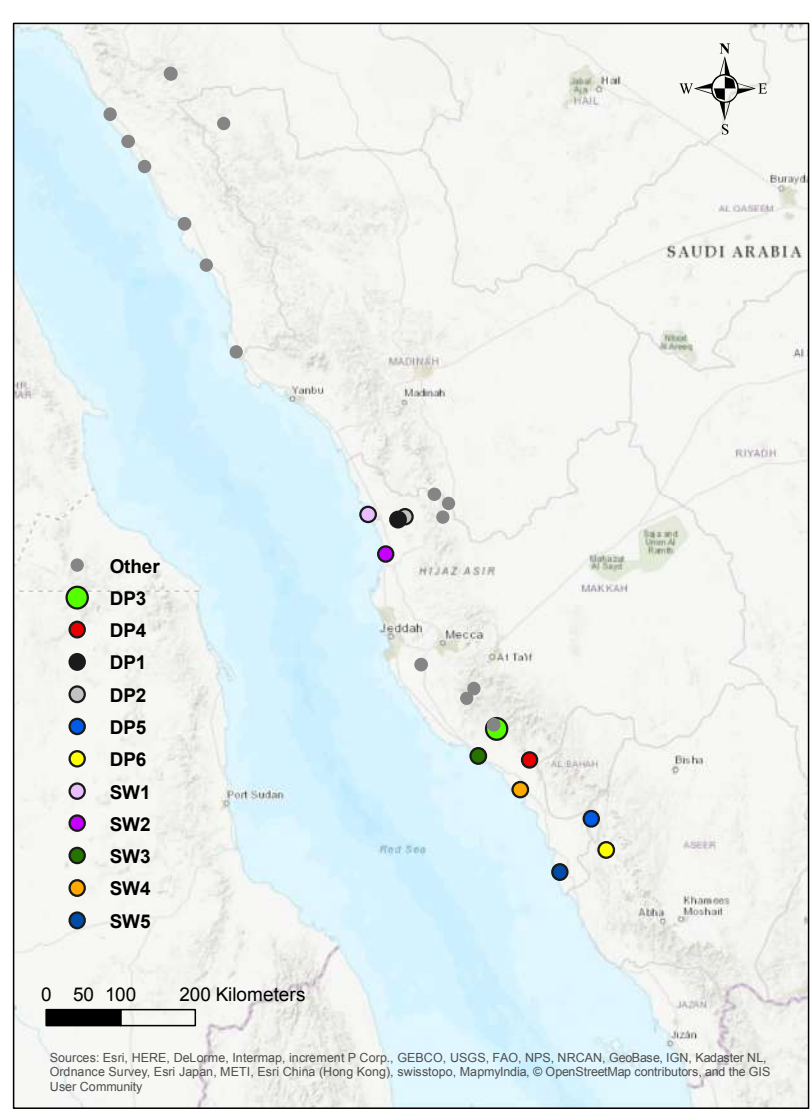

KILLIFISH LOCATION
1:1,500,000
N
Tributary
Main stem
Elevation (mamsl)
3000
1500
0
Rabig
Al Mwared
KAEC
KAUST
Jiddah
Mecca
Al Hawiyah
Al Taif
Al Habqah
Al Lith
Red Sea
Salm Alzwaher
Alfaijah
Al Belkah
AlQouz
0 15 30 60 90 120 Km
39°0'0"E
40°0'0"E
41°0'0"E
42°0'0"E
23°0'0"N
22°0'0"N
21°0'0"N
20°0'0"N
19°0'0"N
DP1
SW1
DP2
SW2
DP3
SW3
DP4
SW4
DP5
DP6
SW5


DP1
DP2
DP3
DP4
DP5
DP6
SW1
SW2
SW3
SW4
SW5
0.6
0.5
0.4
0.3
0.2
0.1
0


Desert
Desert
RedSea
RedSea
Desert
RedSea
Desert
RedSea
Desert
Desert
RedSea
a)
b)
DP1
DP2
SW2
SW1
DP3
SW3
DP4
SW4
DP5
DP6
SW5


a)
DP1
DP2
SW2
SW1
DP6
DP5
SW5
SW4
DP4
SW3
DP3
Eigenvalue
b)
SW2
SW1
SW3
DP3
SW4
DP1
DP2
DP6
SW5
DP4
Eigenvalue


25
20
15
10
5
δ15N (‰)
-20
-10
δ13C(‰)
DP04
DP06
DP05
DP02
DP03
DP01
SW03
SW04
SW01
SW05
SW02
Desert Pond
Sea Water


N
W
E
S
SAUDI ARABIA
Hail
HAIL
Buraydah
AL QASEEM
MADINAH
Yanbu
Madinah
RIYADH
HIJAZ ASIR
MAKKAH
Jeddah
Mecca
AT Taif
AL BAHAH
Bisha
Port Sudan
Red Sea
ASEER
Khamees Mushait
Abha
JAZAN
Jizan
Other
DP3
DP4
DP1
DP2
DP5
DP6
SW1
SW2
SW3
SW4
SW5
0 50 100 200 Kilometers
Sources: Esri, HERE, DeLorme, Intermap, increment P Corp., GEBCO, USGS, FAO, NPS, NRCAN, GeoBase, IGN, Kadaster NL, Ordnance Survey, Esri Japan, METI, Esri China (Hong Kong), swisstopo, MapmyIndia, © OpenStreetMap contributors, and the GIS User Community