Biological Journal of the Linnean Society, 2019, **128**, 415–434. With 5 figures.

Genetic divergence outpaces phenotypic evolution among threespine stickleback populations in old freshwater habitats

MARK C. CURREY, SUSAN L. BASSHAM and WILLIAM A. CRESKO*

Institute of Ecology and Evolution, University of Oregon, Eugene, OR 97403-1254, USA

Received 22 April 2019; revised 18 June 2019; accepted for publication 19 June 2019

Species such as threespine stickleback (*Gasterosteus aculeatus*) that inhabit divergent selective environments and that have diversified on different time scales can be of value for understanding evolutionary processes. Here we synthesize high-resolution genotypic and phenotypic data to explore a largely unstudied distribution of threespine stickleback populations living in oceanic and freshwater habitats along coastal and inland regions of Oregon. Many inland aquatic habitats of Oregon remained unglaciated during the last ice age, meaning that some extant Oregon lake and river stickleback may have descended from freshwater populations established long before more well-studied, post-glacial freshwater populations. To address the degree of congruence between genetic and phenotypic divergence, we directly compared Oregon stickleback to much younger (post-glacial) Alaskan populations. We found phenotypic variation in Oregon stickleback to be primarily partitioned between oceanic and freshwater habitats, as has been documented in other stickleback systems. However, the main axis of genetic divergence was between coastal and inland regions regardless of habitat type. Furthermore, when comparing patterns between Oregon and Alaska we found similar levels of phenotypic divergence, but much greater genetic divergence among Oregon's populations. The Oregon stickleback system therefore appears well suited for future studies linking genotypic and phenotypic change, further extending the utility of this small fish to provide general insights into evolutionary processes.

ADDITIONAL KEYWORDS: *Gasterosteus aculeatus* – population structure – principal component analysis – RAD-seq.

INTRODUCTION

Studies that pair the partitioning of phenotypic and genetic variation within and among populations have long been used to infer ecological and evolutionary processes from patterns observed in nature (Endler, 1986; Storz, 2002; Nosil *et al.*, 2008; Nosil & Schluter, 2011). Integration of these data types is most effectual for biological inference when data are collected from the same set of individuals in populations that occupy diverse habitats and is especially advantageous in systems where parallel phenotypic evolution manifests across geographical locations and/or geological timescales (Endler, 1977, 1986; Schluter, 2000). Recent studies that unite population genomic and phenotypic data from hundreds or thousands of individuals have provided novel insights into the direction and pace of genetic and phenotypic evolution (Nosil *et al.*, 2002; Nadeau *et al.*, 2012; Soria-Carrasco *et al.*, 2014; Ferrero-Serrano & Assmann, 2019). Such combined data sets are ushering in a new age of natural history that includes the mapping of genetic diversity to geographical and phenotypic variation in areas such as conservation genetics (Funk *et al.*, 2012; Casillas & Barbadilla, 2017), population genomics (Andrews *et al.*, 2016), landscape genetics (Dionne *et al.*, 2008; McCairns & Bernatchez, 2008; Gaggiotti *et al.*, 2009) and genome-wide association studies (Pallares *et al.*, 2014; Gienapp *et al.*, 2017; Saltz *et al.*, 2017).

A powerful way to implement this new synthesis is to perform population genomic analyses in an organism for which extensive data on the ecology, evolution and biogeography have already been gathered. The threespine stickleback (*Gasterosteus aculeatus*) is one such system. This small fish is abundant in oceanic, estuarine and freshwater habitats, encompassing populations with diverse life history forms in nearly all parts of the coastline in the northern half of the Northern Hemisphere (Bell & Foster, 1994).

*Corresponding author. E-mail: wcresko@uoregon.edu

Consequently, the species has for decades been a rewarding focus of studies in behaviour, ecology, physiology, ecotoxicology and evolution (Tinbergen, 1950; Foster, 1995; Bernhardt *et al.*, 2006; Hohenlohe *et al.*, 2010; Jones *et al.*, 2012; Reimchen *et al.*, 2013; Spence *et al.*, 2013; Furin *et al.*, 2015; Teigen *et al.*, 2015; Divino *et al.*, 2016; Greenwood *et al.*, 2016; Bassham *et al.*, 2018; Hani *et al.*, 2018). The natural history of the species is therefore one of the best documented of any vertebrate, and the ecological and evolutionary processes that create and maintain its diversity are increasingly well understood (Colosimo *et al.*, 2004; Cresko *et al.*, 2004; Hohenlohe *et al.*, 2010; Schluter *et al.*, 2010; Jones *et al.*, 2012; Bassham *et al.*, 2018).

An exceptional attribute of this stickleback species is its repeated ecological transition across profoundly different environments. Anadromous stickleback populations have episodically colonized and adapted to freshwater habitats (Schluter & Conte, 2009; Bell & Aguirre, 2013), probably throughout time since the origin of the species (Bell & Foster, 1994; Bell, 2009; Bell *et al.*, 2009) and continuing to the present day as new habitats are created (Hagen & Gilbertson, 1972; Klepaker, 1993; Bell, 2001; Bell *et al.*, 2004; Bell & Aguirre, 2013; Lescak *et al.*, 2015). This long history has provided a rich set of replicate natural evolutionary experiments in which phenotypic diversification spans morphological, behavioural, physiological and life history traits. Among these, the best documented are morphological phenomena, including change in defensive armour (e.g. loss of lateral plates and reduction or loss of the pelvic structure), differences in body shape related to foraging or predation avoidance, and changes in craniofacial morphology linked to alternative foraging ecologies (Hagen & Gilbertson, 1972; Bell & Foster, 1994; Walker, 1997; Kimmel *et al.*, 2012; Wund *et al.*, 2015).

At least some of the parallel phenotypic divergence observed in many of these populations has a parallel genetic basis (Cresko *et al.*, 2004; Colosimo *et al.*, 2005; Hohenlohe *et al.*, 2010; Deagle *et al.*, 2012; Jones *et al.*, 2012; Bassham *et al.*, 2018; Miller *et al.*, 2019). Recent work shows that not only can this phenotypic and genetic change occur in the thousands of years since the end of the last glacial maximum (Bell & Orti, 1994; Rundle *et al.*, 2000; Cresko *et al.*, 2004; Hohenlohe *et al.*, 2010; Hendry *et al.*, 2011; Reimchen *et al.*, 2013), but significant change can occur in freshwater populations founded even just decades ago (Bell *et al.*, 2004; Kitano *et al.*, 2008; Gelmond *et al.*, 2009; Lucek *et al.*, 2014; Lescak *et al.*, 2015; Bassham *et al.*, 2018; Nelson & Cresko, 2018).

Although phenotypic and genetic partitioning in young post-glacial stickleback populations has been intensively pursued (Schluter, 1993; Bell & Orti, 1994; Cresko *et al.*, 2004; Shapiro *et al.*, 2004; Colosimo *et al.*, 2005; Kimmel *et al.*, 2005; Hohenlohe *et al.*, 2010; Jones *et al.*, 2012; Catchen *et al.*, 2013; Reimchen *et al.*, 2013; Bassham *et al.*, 2018), much less work has investigated older habitats at lower latitudes that were not subject to the most recent glacial events (but see Hagen & Gilbertson, 1972; Bell & Richkind, 1981; Baumgartner & Bell, 1984; Sanchez-Gonzales *et al.*, 2001; Morris *et al.*, 2018). An open question is how well these patterns extend to stickleback populations that have inhabited much older landscapes. To help answer this question we gathered extensive phenotypic and population genomic data from Oregon stickleback populations, which have been largely unexplored (but see Catchen *et al.*, 2013; Morris *et al.*, 2018), and compared these data to young freshwater stickleback systems such as in south-central Alaska (~12 000–15 000 years old) and Middleton Island (~55 years old).

The unique geography and geological history of Oregon has created aquatic habitats of varying ages and degrees of connectivity with the ocean, characterized by young freshwater habitats along the Oregon coast and much older inland habitats. There are 22 major estuaries along the Oregon coast, many of which were formed when rising sea level drowned river mouths during the last glacial maximum (Allen & Baldwin, 1944). The coastline is also dotted with many freshwater lakes, some of which were formed during the Holocene (0.1–7 kya) by growing sand dunes that blocked run-off (Peterson *et al.*, 2007). Often these coastal lakes outflow to the sea via short channels. In contrast, Oregon's inland stickleback habitats can be hundreds of kilometres from the sea, along the Willamette, Deschutes and Umpqua River basins, which have existed in roughly their present forms for millions of years (Booth *et al.*, 2003).

Previous research on stickleback divergence suggests that much of the genetic variation important for initial adaptation to freshwater is carried by – but not expressed in – oceanic populations and predates the last glacial maximum (Schluter & Conte, 2009; Hohenlohe *et al.*, 2010; McGuigan *et al.*, 2011; Terekhanova *et al.*, 2014; Bassham *et al.*, 2018; Nelson & Cresko, 2018). We therefore expected to find patterns of phenotypic change from the oceanic form in both young coastal and old inland freshwater populations broadly similar to what has been previously described in other stickleback systems. However, because of the existence of more isolated and potentially much older inland freshwater populations in Oregon, we hypothesized that levels of genetic differentiation would be more extreme than those in freshwater populations found along the coast in Oregon or in young populations in Alaska.

To test these predictions, we collected morphometric data (defensive, trophic and body size traits) and restriction site-associated DNA sequencing (RAD-seq) population genomic data from thousands of individual

stickleback from locations throughout north-western Oregon. Our aim was first to describe the patterns of phenotypic and genetic divergence among populations that inhabit the diversity of aquatic habitats found throughout Oregon. Then we compare these patterns to those observed in more recently colonized habitats in Alaska. Ours lays a foundation for targeted work on the genetics of parallel divergence and local adaptation, among other studies.

MATERIAL AND METHODS

Collection and processing of stickleback samples

Oceanic, coastal freshwater and inland freshwater populations were collected throughout north-west Oregon. Oceanic and freshwater Alaska populations were collected from Middleton Island, locations in the Matanuska-Sustina valley, and the Kenai peninsula. Some populations from all locations were used in previous studies (Catchen *et al.*, 2013; Lescak *et al.*, 2015; Bassham *et al.*, 2018). All Oregon and most Alaska collections were made between 2007 and 2017, with some of the Alaska samples collected in the mid-1990s. These efforts provided a total of 1419 individuals in 47 populations used in this study (Fig. 1, Table 1). Most stickleback samples were collected using minnow traps, as previously described (Cresko *et al.*, 2004; Catchen *et al.*, 2013). A subset were obtained from state and federal agencies, namely Oregon Department of Fish and Wildlife (ODFW) and National Oceanic and Atmospheric Administration (NOAA), which were conducting collections of other fish species in Oregon using various methods. GPS coordinates of collecting locations were obtained using Google Earth. Fish were killed with tricaine methanesulfonate (MS222) and immediately stored in 95% ethanol until they were fixed and stained for bone. For genetic analyses, tissue samples of both pectoral and caudal fins were collected fresh or after fixing in ethanol and stored at −20 or −80 °C or were immediately processed for DNA extraction and subsequent RAD-seq. For all samples, the soma of each individual was fixed in 4% paraformaldehyde overnight at room temperature and then stained with Alizarin Red for morphometric analysis (Cresko *et al.*, 2004; Lescak *et al.*, 2015). All research was approved by the Institutional Animal Care and Use Committees (IACUCs) of the University of Alaska Anchorage and the University of Oregon. Fish were collected under Alaska Department of Fish and Game permits SF-2011-153 and SF2014-035 and ODFW scientific taking permits OR2007-3495, 13920, 16933, 17664, 19122 and 20770.

Collection and scaling of phenotypic data

Fourteen morphological traits representing aspects of body size and shape, defensive armour, and trophic structures were scored in all fish in this study. To capture linear measurements, each fish was photographed both laterally and ventrally along with a size standard, and measurements were made using ImageJ (Schneider *et al.*, 2012). Each measurement was performed twice during independent scoring events, and the average of the two scoring events was used for subsequent analysis. The 13 linear measurements that were taken from lateral photographs are represented in Supporting Information Fig. S1A and B, and include standard length, body depth, aspects of the head and jaw, defensive spine lengths, and aspects of the pelvic structure.

To standardize the linear measurements by size, we employed a method described by Reist (1985) that has been used in similar morphological studies (Reimchen & Nosil, 2006; Reimchen *et al.*, 2013). Size standardization of morphological traits was calculated according to: $\log y'_{ij} = \log y_{ij} - \beta(\log x_i - \log \bar{x})$, where $\log y'_{ij}$ is the size-standardized value for trait j in individual i, $\log y_{ij}$ is the original log-transformed value, β is the slope of the regression between trait j and standard length (SL) using a model II regression, $\log x_i$ is the SL of individual i, and $\log \bar{x}$ is the mean SL of all of the individuals in the dataset. The relationship between measurements of most of the individual traits and SL differed among populations (most, $P < 0.001$ trait × SL interaction, ANCOVA), so population-specific slopes for each trait were used for size standardization. In some populations, a significant interaction between some traits and SL was not found, namely between: jaw length and SL in Eel Creek, Middleton Freshwater and Rabbit Slough; dorsal spine length and SL in Buell-Miller, Finley Swamp, Green Island, Pudding Creek, Meadow Creek, Mud Lake, Bear Paw and Middleton Fresh; pelvic spine length and SL in Pudding Creek, Meadow Creek and Middleton Island; and eye orbit size, opercle height and width and SL in Rabbit Slough. These traits were size-standardized by omitting the population-specific slope term from the equation.

Two meristic traits were also captured: the number of lateral plates on the left side of each fish and the number of long gill rakers of the first gill arch on the right side (Supporting Information, Fig. S1C). These features were counted directly from stained fish on a stereomicroscope. Cranial facial bones were dissected to expose the gill rakers, leaving the left side of the head intact for other phenotypic analyses. Meristic traits were counted twice during independent scoring events and counts that differed were repeated until consistency was reached.

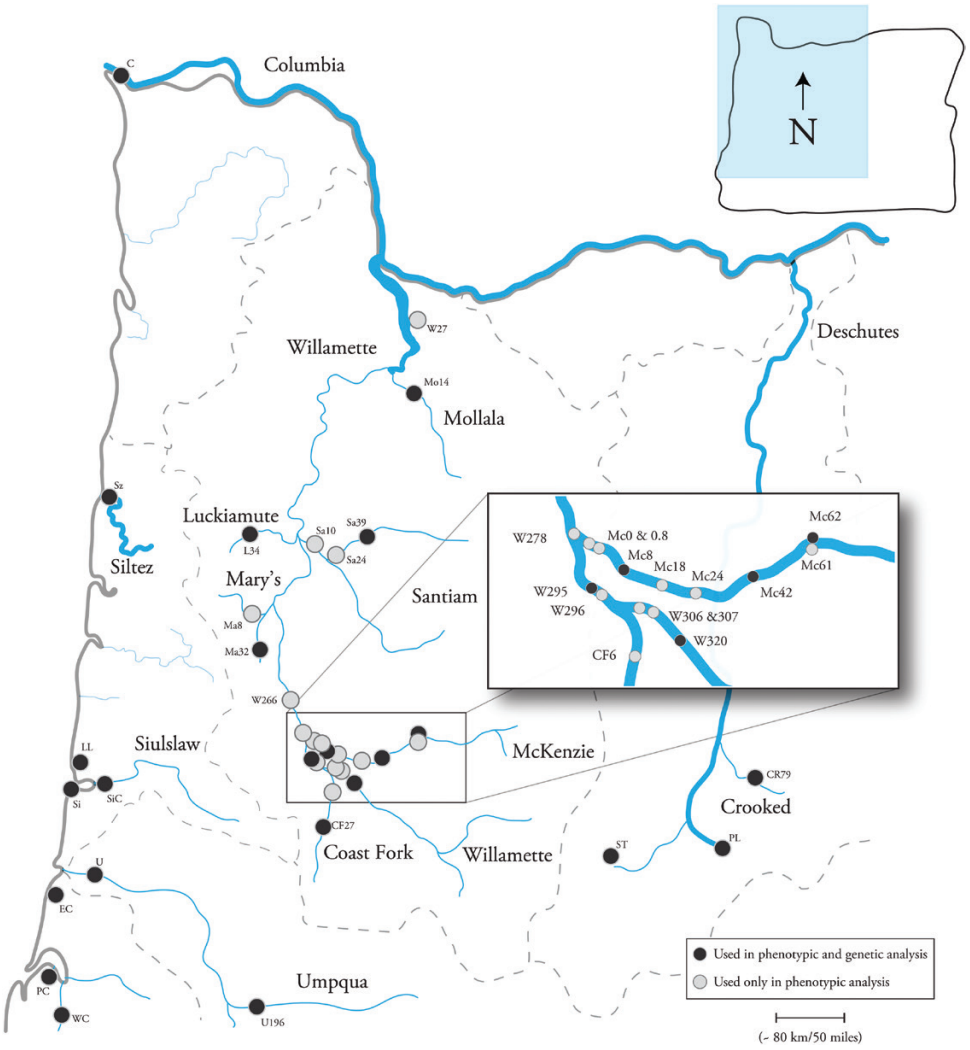

Figure 1. Regional distribution of Oregon populations of threespine stickleback used for phenotypic and genetic analysis. Black dots denote populations used in both analyses, and grey dots populations used in the phenotypic analysis only. Major rivers are shown in blue. Grey dashed lines are major watershed boundaries. Population abbreviations are given in Table 1. Numbers next to population abbreviations denote the distance in river kilometres from the mouth of the major river that each population is associated with according to USGS topo maps.

Analysis of phenotypic variation

To visualize how populations were distributed in phenotypic space, we used principal component analysis (PCA) of the size-corrected measurements of individual traits (described above) in R (R Development Core Team, 2013) using the package pcaMethods (Stacklies *et al.*, 2007). Within pcaMethods, variables were scaled with unit variance and mean-centred before ordination. Because lateral plates and gill rakers can have non-normal distributions, we performed an additional PCA with log-transformed values and obtained nearly identical loadings as when using raw values. Subsequently, raw values were used in all further analysis. To test hypotheses of the partitioning of phenotypic variation between different habitat and regional groupings, analyses of variance (ANOVAs) were performed with individual principal component (PC) scores separately using the R package lmne (Pinheiro *et al.*, 2019). In comparisons of Oregon populations, partitioning of phenotypic variation was tested using an ANOVA model that nested populations as a fixed effect within oceanic and freshwater groupings. In comparisons that included oceanic and freshwater populations from Oregon and Alaska, partitioning of variation was tested using an ANOVA model that nested populations as a fixed effect within regions. In comparisons that included Oregon and Alaska oceanic forms, phenotypic variation was tested using an ANOVA model that nested populations as a fixed effect within regional groupings. Effect size,

Table 1. Collection site details

Collection site	Abb.	Major basin/river	Elevation (m/ft)	Distance to ocean (km/miles)	GPS coordinates	Collection date(s)	Genotyped	Date/sequence length	Phenotyped
Columbia River	C	Columbia	0/0	11/7	46°14′36.04″N 123°54′9.11″W	Jul, 2012	48	2013/100 bp	60
Pony Creek Reservoir	PC	Coos	29/95	23/14	43°22′12.03″N 124°15′43.52″W	Mar, 2007	70	Catchen *et al.* (2013)	30
Winchester Creek	WC	Coos	2/8	10/6	43°16′37.39″N 124°19′8.57″W	2007	22	Catchen *et al.* (2013)	30
Crooked River	CR79	Crooked R.	1075/3530	589/366	44°17′25.57″N 120°50′47.07″W	Aug, 2008	30	Catchen *et al.* (2013)	30
Paulina Lake	PL	Deschutes	1947/6389	∞	43°42′48.66″N 121°16′21.48″W	Aug, 2008	21	Catchen *et al.* (2013)	30
South Twin Lake	ST	Deschutes	1323/4340	∞	43°42′38.38″N 121°46′4.53″W	Jun, 2007	50	Catchen *et al.* (2013)	30
Jont Creek	L34	Luckiamute	68/229	370/230	44°46′28.32″N 123°19′5.02″W	Oct, 2012	48	2013/100 bp	30
Dunawi Creek	Ma8	Mary's	73/239	383/238	44°32′55.04″N 123°18′9.66″W	Aug, 2013			30
Finley Grey Creek Swamp	Ma32	Mary's	95/312	407/253	44°23′44.09″N 123°20′35.42″W	Aug, 2013	48	2013/100 bp	29
Confluence Slough	Mc0	McKenzie	113/372	445/276	44°7′32.21″N 123°6′22.13″W	Jun, 2013			19
Confluence Island	Mc0.8	McKenzie	113/372	446/277	44°7′22.07″N 123°6′8.59″W	Jun, 2013			15
Wildish	Mc8	McKenzie	128/421	452/281	44°7′0.53″N 123°4′17.23″W	Jun, 2013			30
Riverbend	Mc18	McKenzie	132/432	461/287	44°4′40.97″N 123°1′34.92″W	2010–2013	140	Catchen *et al.* (2013)	41
Rainbow Water	Mc24	McKenzie	144/475	468/291	44°3′42.63″N 122°58′7.69″W	Aug, 2012			25
McKenzie Oxbow	Mc35	McKenzie	160/525	480/298	44°3′41.29″N 122°51′10.87″W	Oct, 2012			30
Walterville Slough	Mc42	McKenzie	184/605	486/302	44°4′13.64″N 122°47′54.68″W	Aug, 2012	24	2013/100 bp	50
Leaburg Fish Hatchery	Mc61	McKenzie	224/736	505/314	44°8′3.84″N 122°36′31.86″W	Oct, 2012			50
Leaburg Lake	Mc62	McKenzie	227/746	507/315	44°8′12.99″N 122°36′44.99″W	Oct, 2012	48	2013/100 bp	30

Table 1. Continued

Collection site	Abb.	Major basin/river	Elevation (m/ft)	Distance to ocean (km/miles)	GPS coordinates	Collection date(s)	Geno-typed	Date/se-quence length	Phenotyped
Milk Creek	Mo14	Mollala	48/159	235/146	45°14′12.82″N 122°37′54.38″W	Oct, 2013	32	2013/100 bp	30
I-5 Side Channel	Sa10	Santiam	58/189	346/215	44°44′11.80″N 123°2′56.67″W	Aug, 2013			23
Green's Bridge	Sa24	Santiam	76/248	360/224	44°42′32.23″N 122°58′12.74″W	Sep, 2013			30
Buell-Miller Slough	Sa39	Santiam	120/395	375/233	44°46′12.44″N 122°50′46.96″W	Aug, 2013	21	2013/100 bp	21
Millport Slough	Sz	Siletz	1/3	5/3	44°53′14.68″N 123°59′46.20″W	Apr, 2010	68	Catchen *et al.* (2013)	30
Lily Lake	LL	Siuslaw	4/14	1/0.5	44°5′33.69″N 124°06′58.81″W	Aug, 2015	24	2017/150 bp	30
South Jetty	Si	Siuslaw	1/3	2/1	44°0′7.76″N 124°7′59.26″W	Mar, 2009	86	Catchen *et al.* (2013)	30
Cushman Slough	SiC	Siuslaw	1.625	13/8	43°59′22.4″N 124°2′42.94″W	Mar, 2009	95	Catchen *et al.* (2013)	31
Eel Creek	EC	Ten Mile	5/17	8/5	43°35′13.06″N 124°11′9.96″W	2007–2017	24	2017/150 bp	27
Dean Creek	U	Umpqua	4/13	26/16	43°41′34.36″N 124°0′1.66″W	Jun, 2017	12	2017/150 bp	30
Page Road	U196	Umpqua	133/436	196/122	43°17′2.70″N 123°19′46.66″W	Sep, 2017	46	2018/150 bp	30
Mt Pisgah Lily Pond	CF6	Willamette	156/518	470/292	44°0′0.29″N 122°58′47.05″W	Aug, 2013			51
Lynx Hollow Slough	CF27	Willamette	176/576	491/305	43°51′35.07″N 123°1′24.98″W	Aug, 2013	43	2013/100 bp	30
Reed Canyon	W27	Willamette	29/95	190/118	45°28′54.66″N 122°37′48.04″W	Jul, 2009			29
Blue Ruin Island	W266	Willamette	105/345	428/266	44°13′35.77″N 123°9′15.27″W	Nov, 2012			28
Green Island	W278	Willamette	111/363	441/274	44°8′42.02″N 123°7′4.88″W	Oct, 2012	48	2013/100 bp	30
Mill Race	W295	Willamette	128/420	457/284	44°2′51.59″N 123°4′20.22″W	Prior to 2008			29
Science Factory	W296	Willamette	125/411	459/285	44°3′25.50″N 123°4′30.93″W	2012, 2013	48	2013/100 bp	62

Table 1. Continued

Collection site	Abb.	Major basin/river	Elevation (m/ft)	Distance to ocean (km/miles)	GPS coordinates	Collection date(s)	Geno-typed	Date/se-quence length	Phenotyped
Chub Site	W306	Willamette	146/478	468/291	44°1′29.00″N 122°58′29.49″W	Jul, 2012			50
Pudding Creek	W307	Willamette	140/460	470/292	44°1′15.72″N 122°57′37.85″W	Jul, 2012			39
Dougren Slough	W320	Willamette	172/565	483/300	43°58′1.19″N 122°52′8.41″W	Sep, 2013	48	2013/100 bp	30
Bear Paw Lake	**BP**	**Mat-Su Ak**	**85/281**	**20/13**	**61°36′53.79″N 149°45′23.32″W**	**Jun, 2004**	**47**	**2018/150 bp**	**20**
Mud Lake	**Md45**	**Mat-Su Ak**	**30/97**	**82/51**	**61°56′8.27″N 150°58′44.01″W**	**1997**			**20**
Meadow Creek	**Me13**	**Mat-Su Ak**	**426/142**	**13/8**	**61°33′39.87″N 149°50′0.33″W**	**Jul, 1996**			**20**
Rabbit Slough	**Ra**	**Mat-Su Ak**	**18/5**	**8/5**	**61°32′10.85″N 149°15′30.01″W**	**2014**	**48**	**2017/150 bp**	**20**
Middleton Isl Fr (#15)	**MiF**	**Mid-Is Ak**	**0.3/1**	∞	**59°27′44.82″N 146°17′53.36″W**	**May, 2010**	**50**	Lescak *et al.* (2015)	**20**
Middleton Isl Oc (#23)	**MiO**	**Mid-Is Ak**	**0/0**	**0/0**	**59°27′39.60″N 146°17′45.60″W**	**May, 2010**	**32**	Lescak *et al.* (2015)	**20**
Resurrection Bay	**Re**	**Resur. Ak**	**0/0**	**0/0**	**60°6′40.98″N 149°24′24.93″W**	**1997**			**20**
Salmon Creek	**Re3**	**Resur. Ak**	**12/38**	**3.2/2**	**60°8′49.76″N 149°24′46.29″W**	**1995**			**20**
						Total	1321		1419

Details shown are: abbreviation (collection location plus approximate river location in kilometres) used in the figures, major basin or river location of the site located in, elevation (m/ft), distance to ocean (km/miles), GPS coordinates, collection dates, number of individuals used in genotypic analysis, date sequenced and length of read, and number of individuals from each population used in the phenotypic analysis. Collection sites are arranged in alphabetical order by major basin. Alaskan populations are separated at the bottom of the list and are shown in bold type. Infinity signs represent landlocked populations. Numbers associated with the Middleton Island collection site names reflect location numbers used in Lescak *et al.* (2015).

partial η^2, was calculated within the BaylorEdPsych package (Beaujean, 2012) by dividing the sum of squares treatment by the sum of squares residual.

RAD library construction and SNP genotyping

RAD-seq data from previous publications were included: all nine Oregon populations from Catchen *et al.* (2013) and two populations from Middleton Island, Alaska, from Lescak *et al.* (2015). From Middleton Island, we chose one phenotypically oceanic (Mi23) and one phenotypically freshwater (Mi15) population that are genetically divergent along the oceanic/freshwater spectrum (Lescak *et al.*, 2015). New RAD-seq libraries were generated for 17 additional populations (15 from Oregon and two from Alaska), using the restriction endonuclease *SbfI*-HF, and sequenced to 100 or 150 nucleotides on an Illumina HiSeq 2000 or 4000 platform (Table 1), as described in Catchen *et al.* (2013) and Lescak *et al.* (2015). Raw reads were de-multiplexed by quality score using the process_radtags program in the Stacks software pipeline, v.1.46 (Catchen *et al.*, 2011). This resulted in a total of more than 1.6 billion 100–150-bp sequences that passed several quality filters with an average of a little over 182 million sequences retained per population (Supporting Information, Table S1). Processed reads were then aligned against the stickleback genome using GSnap (Wu & Nacu, 2010), allowing for up to five nucleotide mismatches and gap lengths of two nucleotides. Only reads with unique alignments were retained. Previous RAD-seq data from 11 populations were included with the retained reads from above and processed through the Stacks pipeline (modules: pstacks, cstacks, sstacks and populations) to produce a catalogue of genotypes and to call genotypes for each individual. Further analyses include only loci that fulfilled the following criteria: present in all populations and present in at least 75% of the individuals in each population. This further stringent filtering resulted in 10 928 variant sites identified in all 28 populations (Table S2).

Analysis of genetic variation

The populations program within the Stacks framework was used to calculate genome-wide average F_{ST} by averaging individual F_{ST} values for each single nucleotide polymorphism (SNP). PCA and STRUCTURE analyses were performed to visualize population grouping and structure of genetic variation. To create a computationally manageable subset of markers from the thousands of RAD loci that were discovered, three sets of 1100 randomly sampled SNPs (restricted to one SNP per RAD tag to exclude tightly linked SNPs) were generated for each collection of individuals included in a comparison (e.g. all Oregon populations or all Willamette Basin populations). To investigate and visualize the axes of genetic variation, a PCA was performed on each set of 1100 randomly chosen polymorphic loci using the software Genodive (Meirmans & Van Tienderen, 2004). The mean and standard deviation of PC 1, PC 2 and PC 3 were visualized using R (R Development Core Team, 2013). For all PCAs, a covariance matrix was generated, and the significance of each PC was tested using a resampling method with 1000 permutations. Visual comparison of all three random SNP sets yielded consistent placement and relative position of populations to one another in each PCA plot (data not shown) so results from only one of the sets are presented here. With STRUCTURE (Pritchard *et al.*, 2000) we used the same set of 1100 randomly sampled loci used in the genetic PCA to identify patterns of the partitioning of genetic variation. Because we wanted to know how genetic variation was structured across the landscape and among populations at different levels of clustering (Rosenberg *et al.*, 2002), increasing values of K were analysed until the patterns of genetic partitioning were no longer in accordance with geographical region or population. In our runs this occurred at $K = 10$. Each STRUCTURE run was performed via ten runs for each value of K (the number of genetic groupings), with 40 000 burn-in steps and 40 000 replicates. If there was incongruence among the ten runs using these initial parameters, however, the number of burn-in steps and the number of replicates for each value of K were increased by 10 000 until a congruent answer was reached among all ten runs (Supporting Information, Table S3). Results were visualized using Clumpak (Kopelman *et al.*, 2015).

We used Genodive to perform an analysis of molecular variance (AMOVA) (Excoffier *et al.*, 1992) to test the partitioning of genetic variation between habitat and geographical groupings using the same set of 1100 randomly sampled loci. Coastal and inland populations were designated by habitat type and not by fish phenotype. All AMOVAs used an infinite allele model. Significance was determined by resampling with 1000 permutations.

RESULTS

Phenotypic variation in Oregon stickleback is partitioned between oceanic and freshwater habitats and not by geographical proximity

Phenotypic variation was predominantly partitioned among populations from oceanic vs. freshwater habitats. The first three axes together explained nearly 70% of the total variation. The major axis of phenotypic variation (PC 1) explained 29% of the total variation, and it significantly differentiated oceanic

and freshwater individuals ($F_{1,1244}$ = 3192.4, P < 2e-16, partial η^2 = 0.72; Fig. 2A; Supporting Information, Fig. S2A). As expected, variation in this PC was also significantly differentiated among populations within each habitat ($F_{37,1244}$ = 133.1, P < 2e-16, partial η^2 = 0.8). The partitioning of phenotypic variation along PC 1 was driven mainly by defensive traits (the number of lateral plates, pelvic girdle width and length, and dorsal and pelvic spine length), and by a trophic trait (gill raker number) (Fig. 2A; Table S4). It is of note that some coastal freshwater populations (e.g. Pony Creek and Lily Lake) are located at the edge of the freshwater cluster (Fig. 2A). This positioning along PC 1 is probably due to the presence of some fish in these populations that have intermediate lateral plate phenotypes. PC 2 explained nearly an equal amount (25%) of the phenotypic variation and also significantly differentiated oceanic and freshwater individuals, although with a very small effect size ($F_{1,1244}$ = 20.8, P < 0.0006, partial η^2 = 0.016; Fig. 2A; Fig. S2A). This PC significantly differentiates variation among populations within habitats ($F_{37,1244}$ = 23.79, P < 2e-16, partial η^2 = 0.41).Variation along PC 2 largely corresponds to craniofacial traits (jaw length, eye orbit diameter, dorsal cranial length and opercle bone size) (Fig. 2A; Table S4). PC 3, which explained 13% of the variation, also significantly partitioned variation between oceanic and freshwater individuals but with small effect size ($F_{1,1244}$ = 104.73, P < 2e-16, partial η^2 = 0.08; Fig. S2A). Again, there was significant partitioning of variation among populations within each habitat ($F_{37,1,244}$ = 38.11, P < 2e-16, partial η^2 = 0.53). Body depth, dorsal cranial length, and pelvic ascending process and girdle length all loaded highly on PC 3 (Table S4). Reductions in the number of lateral plates and gill rakers are evolutionary phenomena that have been found to be highly associated with the transition from oceanic to freshwater forms in stickleback (Bell & Foster, 1994) and are the primary drivers of differentiation along PC 1 in the above analysis (Table S4). Exclusion of these two traits had little qualitative effect on the ordination, aside from transposing the first two nearly equal principal components (Figs S2B, S3, Table S4).

Partitioning of genetic variation is qualitatively different from partitioning of phenotypic variation in Oregon stickleback

Unlike phenotypic variation, which we found to be partitioned chiefly by oceanic vs. freshwater habitat types, genotypic variation in Oregon is partitioned across geographical regions. Analysis of genetic variation differentiates coastal from inland populations, both in multivariate ordination using PCA (Fig. 2B) and in STRUCTURE analysis (Fig. 3). In our multivariate ordination, PC 1 accounts for the majority (~58%) of the total genetic variation and clearly separates coastal from inland populations (Fig. 2B; Table S5). In addition, some separation occurs between coastal freshwater and coastal oceanic populations along PC 2, although this outcome is not statistically significant. STRUCTURE analysis using the lowest level of clustering, K = 2, produces qualitatively similar genetic partitioning as seen in the genetic PCA. Clusters are anchored by the same two geographical regions, with both oceanic and freshwater coastal populations clustering to the exclusion of inland populations (Fig. 3). An interesting set of exceptions for both the genetic and the phenotypic analyses are the populations from (1) Milk Creek (a tributary of the Molalla River), (2) Buell-Miller Slough of the Santiam River, (3) Riverbend, Walterville and Leaburg (all three from the McKenzie River), and (4) Page Road from the Umpqua River (Fig. 2A, B). These exceptions will be described and discussed below.

At increasing levels of K (3–10), the two geographical regions that correspond to coastal and inland remain separate, although these two clusters become increasingly subdivided into groupings that correspond to habitat type (with coastal populations split into oceanic and freshwater groupings), whereas inland populations are largely separated by river basin (Fig. 3).

This pattern of the partitioning of genetic variation is also present when comparing average measures of population differentiation and AMOVA. On average, F_{ST} values between oceanic and inland populations (excluding the inland populations with outlying genetic affinity to oceanic populations) was ~0.16, ranging from 0.12 (Cushman Slough vs. Finley Grey Creek Swamp) to 0.25 (Dean Creek vs. Dougren Slough) (Supporting Information, Table S6). This is similar to previously reported levels of divergence between oceanic and Willamette Basin populations (Catchen *et al.*, 2013). In contrast, stickleback populations found in Oregon's coastal freshwater habitats demonstrated the lowest average genome-wide divergence from oceanic fish (average coastal freshwater compared to average oceanic, F_{ST} = 0.06). This measure of divergence is similar to what we found between young oceanic and freshwater pairs from south-central Alaska (F_{ST} = 0.056) and Middleton Island (F_{ST} = 0.057). Finally, we used AMOVA to quantify how genetic variation is partitioned within and between individuals, populations and geographical regions (coastal and inland). We found that most of the variation was partitioned between individuals (~46%), although there was little partitioning of variation between individuals within the populations (~4%). We found significant variation partitioned between the populations within each geographical region (~23%) and significant variation partitioned between

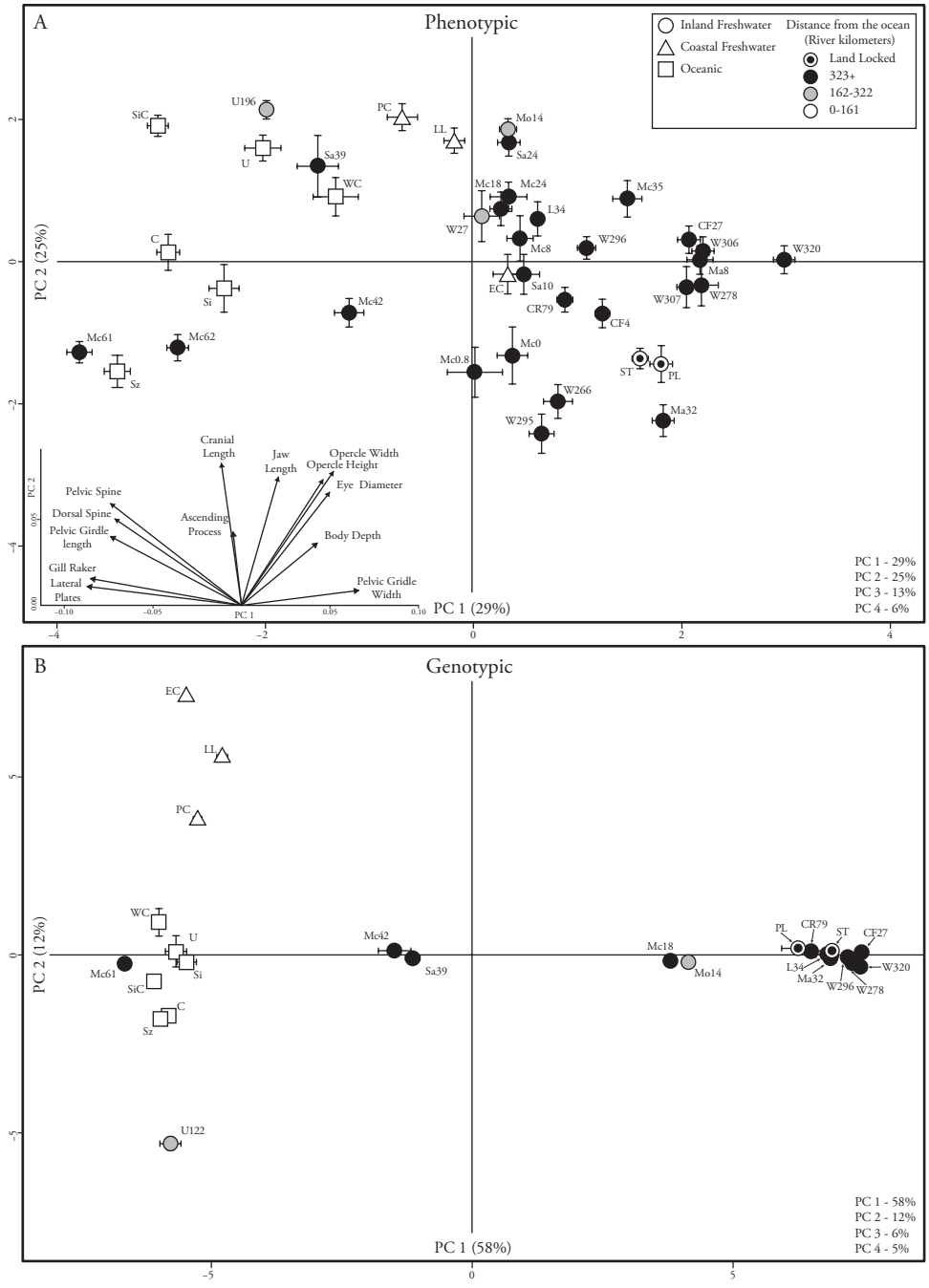

Figure 2. Phenotypic and genotypic PCA of Oregon populations. A, PCA partitions phenotypic variation between oceanic and freshwater stickleback populations. Each symbol represents population-level mean PC 1 and 2 scores with bars representing one standard error. PC 1 accounted for 29% of the total phenotypic variation and PC 2 accounted for 25% of the total phenotypic variation. Lower left subset represents trait loadings along PC 1 and PC 2. Arrows indicate the direction of a trait's loading and arrow length indicates its relative strength. B, genetic variation in Oregon stickleback is partitioned between regions and not by habitat types. PC 1 explains 63% of the genetic variation and separates coastal populations on the left and inland populations on the right. The average PC 1 and PC 2 score of each population is plotted, with bars representing one standard error. Population symbols are shaded according to the distance between collection site and the ocean, ranging from light (close) to dark (far) measured in river kilometres. Circles represent populations found in inland freshwater habitats, triangles represent populations found in coastal freshwater habitats, and squares represent populations found in oceanic habitats. Population names are given in Table 1. In the bottom right of each panel is the percentage variance explained for the first four PCs.

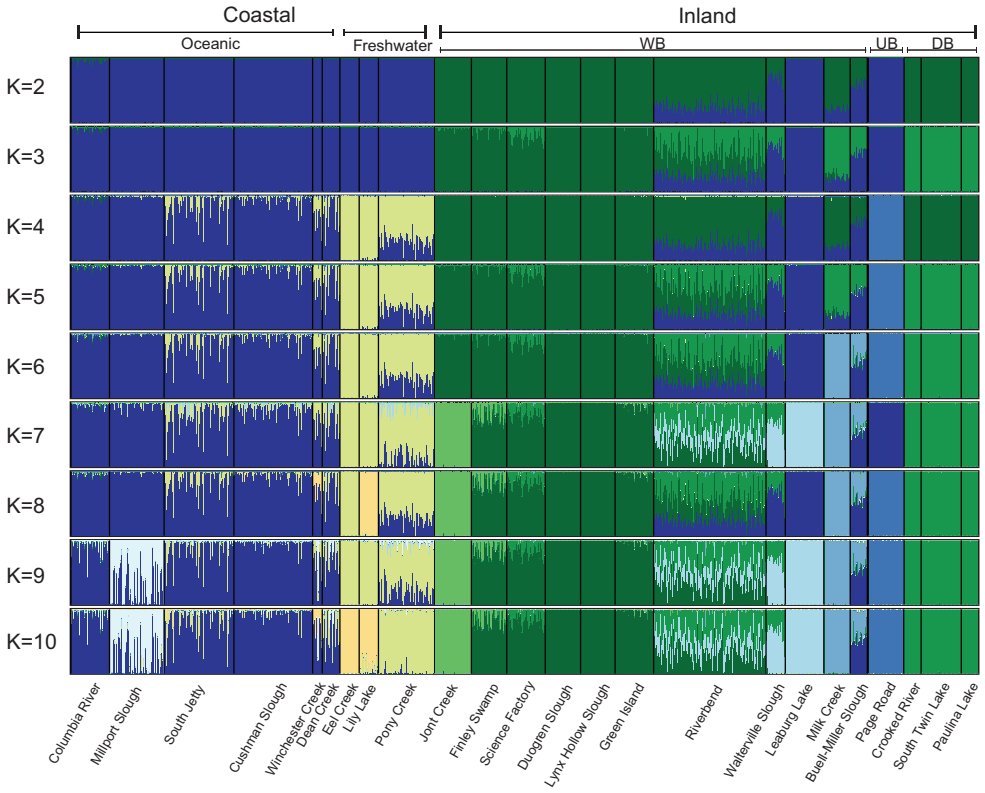

Figure 3. STRUCTURE plot of Oregon populations. Oregon populations of stickleback are genetically structured by region, with some exceptions. Each vertical bar represents an individual with colours representing the posterior probability of membership to that group, as estimated by STRUCTURE. Each level of *K* represents ten repeats that were run with increasing levels of burn-in and repetitions until congruence was reached. In general, at low levels of *K*, populations are partitioned between coastal and inland regions. With increasing levels of *K* the two regions are increasingly partitioned by location. However, some inland populations (e.g. Leaburg) do not strictly follow these patterns. WB = Willamette Basin, UB = Umpqua Basin, and DB = Deschutes Basin.

the coastal and inland geographical regions (~28%) (Table S7), findings broadly consistent with PCA and STRUCTURE.

Varying degrees of genetic divergence exist between populations within one of Oregon's major watersheds

Within the Willamette River Basin, Dougren Slough, Lynx Hollow and Jont Creek are all located far upstream in their individual tributary drainages, 19, 28 and 34 river kilometres, respectively. The Science Factory and Green Island sampling sites are downstream from Dougren Slough and Lynx Hollow along the valley floor, in the main stem of the Willamette near the confluence of these drainages (Fig. 1). Finley Grey Creek Swamp is also located on the valley floor, where it probably has high connectivity with the main stem via historical and present-day flooding (Benner & Sedell, 1997). There is strong genetic divergence among the upstream populations from different drainages, with F_{ST} values ranging from 0.33 (Jont creek vs. Lynx Hollow) to 0.37 (Dougren Slough vs. Lynx Hollow), and less divergence between these populations and their downstream counterparts, with F_{ST} values ranging from 0.09 (Dougren Slough vs. Science Factory) to 0.19 (Lynx Hollow vs. Finley Swamp) (Supporting Information, Table S6). This pattern is probably driven by an overall decrease in genetic variation in the upstream populations that can affect the relative degree of divergence measured by F_{ST} (Hernandez-Martich & Smith, 1990; Shaw *et al.*, 1991; Castric *et al.*, 2001; Noor & Bennett, 2009; Cruickshank & Hahn, 2014) (Table S2).

Inland freshwater populations from the Willamette Basin are genetically structured by drainage, with exceptions

To better understand the patterns of genetic divergence within the Willamette Basin, we focused multivariate genetic analyses on populations collected

throughout this large network of tributaries. In an analysis including 11 populations from the Willamette River system, the first and the only significant genetic PC accounted for ~64% of the variation. Populations along PC 1 were differentiated into groups with low vs. higher than average lateral plate counts (Supporting Information, Fig. S4, Table S8). In STRUCTURE at K = 2, populations cluster into low- and extra-plated groupings, where extra-plated populations are defined as those with average lateral plate number >7. This manifests in lateral plates that extend to various degrees posteriorly of the supporting plates. A relatively high-elevation population in the McKenzie river (Leaburg) is separated in STRUCTURE from a cluster of valley floor stickleback (from Jont Creek, Finley Swamp, Science Factory, Dougren Slough, Lynx Hollow and Green Island). Fish collected from sites between Leaburg and the valley floor cluster are a mixture of both groups. At K = 3–5, the northern Willamette Basin extra-plated Milk Creek and Buell-Miller populations form an exclusive cluster, with the Buell-Miller population having partial membership coefficients in the Leaburg grouping (Fig. S5A). These findings were confirmed using AMOVA, whereby a significant amount of variation was partitioned between extra- and low-plated populations (~16 %) but was not partitioned among individuals (~54%) or among populations (~28%) (Table S9). When only the low-plated populations are considered, fish collected highest upstream in their drainages (Dougren Slough, Lynx Hollow Slough and Jont Creek) separate into nearly three discrete freshwater clusters at K = 3, while downstream locations in these drainages (Finley Grey Creek Swamp, Green Island and Science Factory) harbour mixtures of genotypes from these three discrete drainage clusters (Fig. S5B).

The structuring of the low plated populations and the pattern of divergence that we identified is consistent with observations in other riverine fish systems, which show strong unidirectional gene flow in upstream populations, presumably resulting from increasing river gradients with a gain in elevation (Hernandez-Martich & Smith, 1990; Shaw *et al.*, 1991; Castric *et al.*, 2001) (Fig. 4). However, the extra-plated populations did not fit this pattern. The discovery of populations very far from the coast that appear to have genetic and phenotypic affinity to oceanic populations was unexpected and warranted further investigation.

Some Oregon inland populations cluster both phenotypically and genetically with oceanic populations

One startling finding was the presence of extra-plated and high-plated stickleback in rivers deep within the inland Willamette and Umpqua Basins. Phenotypic analysis of multiple traits places these populations among the oceanic groups or among phenotypically intermediate groups (Fig. 2A; Supporting Information, Fig. S3A).

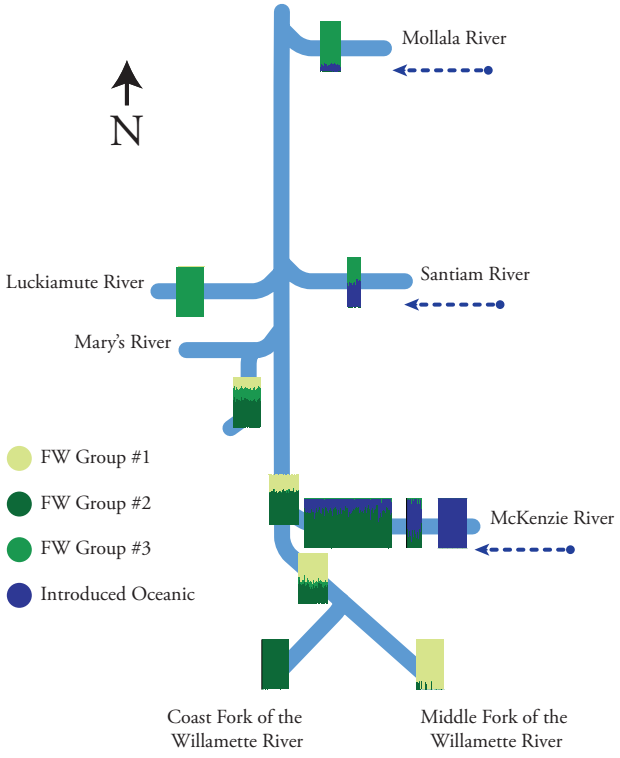

Figure 4. Genetic structure of Willamette Basin populations matches the pattern of distance, directional flow and sequence of confluences, with some exceptions. The genetic clusters found with STRUCTURE are plotted on a representation of the major rivers of the basin. Distinct clusters (groups 1 and 2) in the upper forks of the Willamette's main stem appear to be almost fully partitioned between Coast and Middle Forks. Downstream, these constituents grade into a third distinct genetic group (group 3), which appears to predominate lower in the watershed. Oceanic genotypes (blue) are differentially introgressed into freshwater genotypes along three rivers flowing down the west slope of the Cascade foothills.

Genetic analysis supports and clarifies these relationships. Along PC 1, Leaburg Lake (McKenzie River) and Page Road (Umpqua River) cluster distinctly with the coastal populations (and within the oceanic populations along PC 2) (Fig. 2B). Populations from Riverbend and Walterville (both of the McKenzie River), Buell-Miller Slough (Santiam River) and Milk Creek (Molalla River) are genetic intermediates between coastal populations and inland low-plated populations. At all levels of K tested in STRUCTURE analysis there is partial membership of

these ocean-like freshwater populations with oceanic populations. This is best exemplified at $K = 2–8$, where Leaburg, and to a lesser extent Page Road, Milk Creek and Buell-Miller, partially or fully cluster with oceanic populations (Fig. 3).

Consistent with STRUCTURE patterns, ocean-like populations in the Willamette Basin display less genome-wide divergence from oceanic fish (average F_{ST} ~ 0.08) than they do from other Willamette Basin populations (average F_{ST} ~ 0.21) (Supporting Information, Table S6). This modest degree of divergence between the oceanic-like Willamette Basin and actual oceanic populations is similar to genetic divergence reported between oceanic and very young freshwater populations on Middleton Island (average F_{ST} ~ 0.067; Lescak *et al.*, 2015), reflecting possible accidental introduction into these inland rivers since the 1980s of coastal fish via fish stocking (Zakel, 1984).

Phenotypic variation in Oregon stickleback is partitioned in a similar manner as in Alaska populations but is accompanied by much greater genetic divergence

Using our findings from the STRUCTURE and genetic PC analyses, we selected Oregon populations that exemplify strong oceanic vs. freshwater partitioning of genetic variation for phenotypic comparison with divergent pairs of Alaskan populations from the mainland (Rabbit Slough and Bear Paw) and from a marine island (Mi23 and Mi15). Using PCA we found similar patterns of phenotypic divergence in Oregon and Alaska populations along the major axis. PC 1, which explained 30% of the variation, groups oceanic populations in both Alaska and Oregon (Fig. 5A). This PC significantly differentiates oceanic and freshwater populations regardless of geographical region, with a very large effect size ($F_{1,730}$ = 1117.98, P < 2e-16, partial η^2 = 0.60; Fig. 5A; Supporting Information, Fig. S6). The phenotypic divergence between the oceanic and freshwater groupings was driven by the number of lateral plates and gill rakers, depth of the ascending process and body, eye orbit diameter, and size of the opercle (Fig. 5A; Table S4). Interestingly, dorsal and pelvic spine length and pelvic girdle length did not weigh as heavily on this axis of differentiation as they did in the analysis that included only the Oregon populations. Somewhat unexpectedly, we did find significant regional partitioning of phenotypic variation of PC 1 between Oregon and Alaska regardless of habitat type, with a moderate effect size ($F_{1,730}$ = 498.6, P < 2e-16, partial η^2 = 0.41) and significant differences among populations within each region ($F_{23,730}$ = 25.83, P < 2e-16, partial η^2 = 0.45).

PC 2 also partitioned phenotypic variation between habitat and regional groupings. PC 2, which explained 23% of the total variation, significantly partitioned variation by oceanic and freshwater groupings ($F_{1,730}$ = 1177.07, P < 2e-16, partial η^2 = 0.62) and region ($F_{1,730}$ = 393.67, P < 2e-16, partial η^2 = 0.35), with the habitat grouping having a larger effect size (Fig. 5A; Supporting Information, Fig. S6) and significant differences among populations within each region ($F_{23,730}$ = 41.99, P < 2e-16, partial η^2 = 0.57). As in PC 1, phenotypic variation is clearly partitioned between Oregon oceanic and freshwater populations along PC 2. However, unlike PC 1, there is less partitioning of phenotypic variation between Alaska habitat types along PC 2 (Fig. 5A). PC 3 does not significantly differentiate phenotypic variation between habitat but there is significant difference between regions with a very small effect size ($F_{1,730}$ = 0.136, P = 0.712, partial η^2 = 0.0002 and $F_{1,730}$ = 210.50, P < 2e-16, partial η^2 = 0.22; Fig. S6). There are also significant differences among populations within each region ($F_{23,730}$ = 45.47, P < 2e-16, partial η^2 = 0.59).

It is interesting to note that there is clear separation between Oregon and Alaska oceanic forms along PC 1 and 2 (Fig. 5A). To investigate this further, an additional PCA was performed with only the oceanic populations from these two regions. In this analysis, phenotypic variation is significantly partitioned between Oregon and Alaska along PC 1 ($F_{1,203}$ = 855.11, P < 2e-16, partial η^2 = 0.81) (Supporting Information, Fig. S7A) and significant differences among populations within Oregon and Alaska groupings ($F_{5,203}$ = 18.85, P = 2.24e-15, partial η^2 = 0.32). This is driven by differences in dorsal and pelvic spine length, pelvic width and length, and the length of the ascending process and body depth (Fig. S7B). Phenotypic differences in oceanic populations from different parts of their range have been documented before (Defaveri & Merila, 2013), although this is counter to the idea that oceanic stickleback consist largely of a single phenotypic form (Bell & Foster, 1994; Morris *et al.*, 2018), and implies that an underappreciated diversity of oceanic populations throughout this species' global range is likely.

Performing a genetic PCA, we found that PC 1, which accounts for ~51% of the partitioning of genetic variation and which is the only significant PC (Supporting Information, Table S10), separates Oregon populations between coastal and inland geographical regions regardless of habitat type, just as in the PCA including only Oregon populations (Fig. 5B). By contrast, all of the Alaska populations, both oceanic and freshwater,

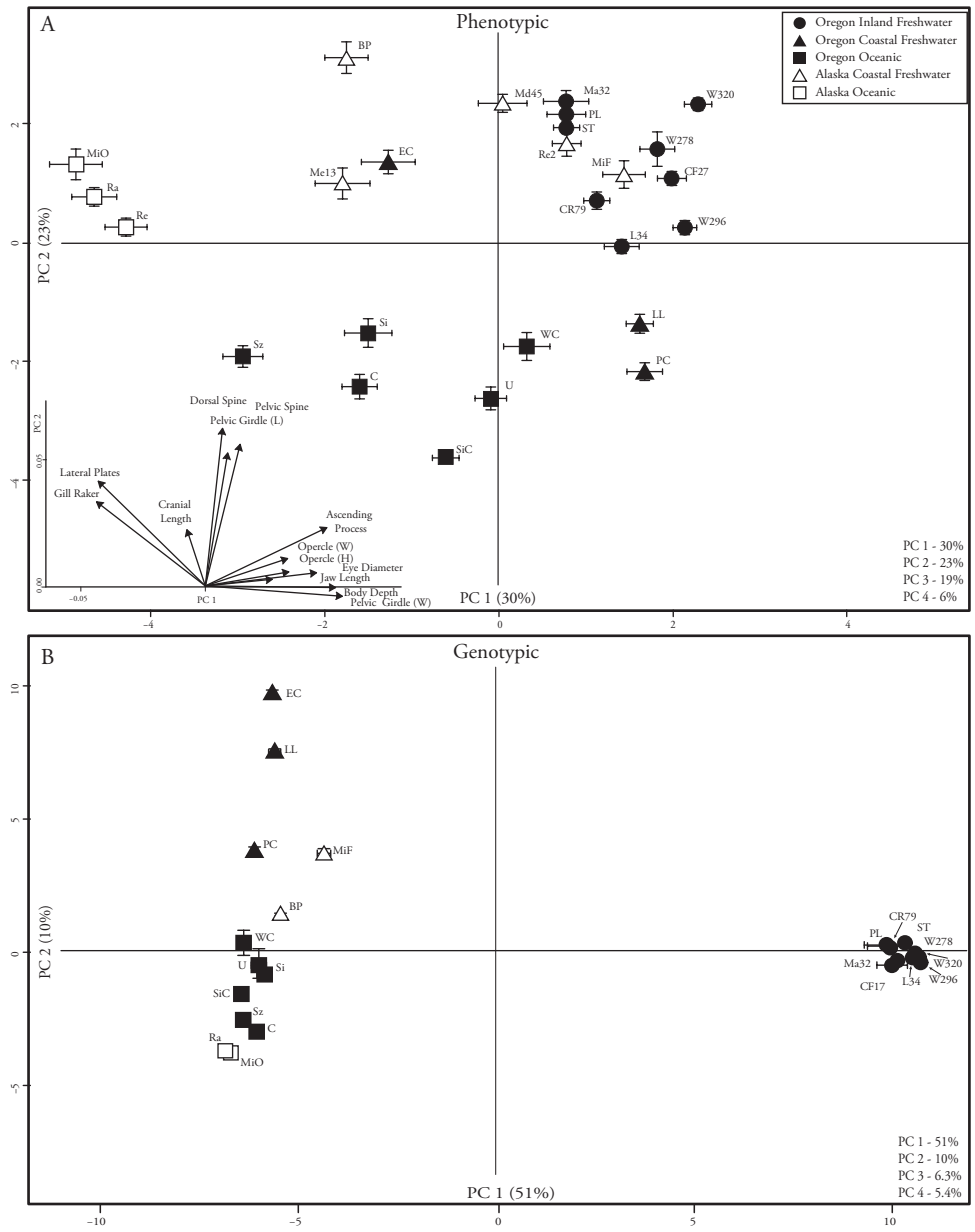

Figure 5. Phenotypic and genotypic PCA of Oregon including Alaska populations. A, Oregon and Alaska stickleback populations show similar phenotypic divergence between oceanic and freshwater forms. Phenotypic variation is partitioned between oceanic and freshwater forms along PC 1, which explains 30% of the variation, oceanic populations to the left and freshwater populations to the right, squares vs. circles. PC 2 explains 23% of the variation. The lower left subset represents trait loadings along PC 1 and PC 2. Arrows indicate the direction of a trait's loading and arrow length indicates its relative strength. B, Alaska populations group with Oregon coastal populations in a genetic PCA. The only significant principal component (PC 1) explained 51% of the genetic variation. Average PC 1 and PC 2 scores of each population are plotted, with bars representing one standard error of the mean. Circles represent populations found in inland freshwater habitats, triangles populations found in coastal freshwater habitats, and squares populations found in oceanic habitats. Population names are given in Table 1.

cluster only with Oregon coastal populations (Fig. 5B). Concordantly, genome-wide average F_{ST} between Alaska oceanic and freshwater populations is ~0.07, similar to the divergence between Oregon oceanic and Oregon coastal freshwater populations (F_{ST} ~ 0.06). By comparison, Oregon inland freshwater populations show more than twice as much divergence from Oregon oceanic populations, with an average F_{ST} of ~0.16.

DISCUSSION

Freshwater stickleback in Oregon differ in form from oceanic relatives to a similar degree as young, post-glacial populations, yet can harbour much greater genetic divergence

Freshwater stickleback across western Oregon display many aspects of the stereotypical phenotypic divergence from their oceanic relatives that has been described throughout the geographical range of this species (Bell & Foster, 1994; Conte *et al.*, 2015). This 'freshwater syndrome' includes reduction in the number of lateral plates and gill rakers, reduction of the pelvic structure and lengths of defensive spines. The well-studied loss of lateral plates and reduction of gill rakers are also major factors we observed in the partitioning of phenotypic space. We did not find, however, populations with characteristics that have been encountered more rarely, such as the partial or complete loss of a bony pelvic structure that has been documented in Alaska and California (Bell, 1987), nor did we find evidence for cases of the drastic phenotypic divergence that has been described between benthic and limnetic stickleback in British Columbia (Mcphail, 1984).

We found that morphological variation is partitioned between Oregon oceanic and freshwater populations to a similar extent as it is in comparisons of oceanic with freshwater populations that were founded just thousands or dozens of years ago (Bell & Foster, 1994; Cresko *et al.*, 2004; Lescak *et al.*, 2015). Similarities of divergence in Oregon sticklebacks to populations in post-glacial landscapes fade, however, when we explore population genomic metrics. Surprisingly, the major partitioning of genetic variation in Oregon stickleback is between coastal (both oceanic and freshwater) and inland populations, in stark contrast to other stickleback systems in which the major genetic and phenotypic axes divide oceanic from freshwater populations.

The discordance between the phenotypic and genetic data in the Oregon populations argues for several important predictions. First, the small amount of genetic divergence between coastal populations – regardless of habitat type – suggests that coastal freshwater populations were founded relatively recently and might be experiencing ongoing gene flow with oceanic fish. Second, the degree of genetic differentiation of inland vs. coastal fish is consistent with the hypothesis that Willamette Basin populations pre-date the last glacial maximum and/or are isolated by distance by the directional nature of river systems.

The observation of much more similar patterns of phenotypic divergence than of genetic divergence among very young (50 years on Middleton Island), post-glacial (south-central Alaska) and much older (inland Oregon) freshwater populations argues that the 'freshwater syndrome' evolves rapidly and hinges on reuse of standing genetic variation across particular genomic loci (Barrett & Schluter, 2008; Hohenlohe *et al.*, 2010; Feulner *et al.*, 2013; Terekhanova *et al.*, 2014; Marques *et al.*, 2016; Bassham *et al.*, 2018). However, this rapid burst of phenotypic and corresponding genetic change may then be followed by long periods of phenotypic evolutionary stasis underlain by continuing neutral genetic divergence. If some of the Oregon freshwater populations do indeed pre-date the last glacial maximum, a testable hypothesis is that these populations might harbour much older freshwater haplotypes at genomic loci implicated in re-use during recent transitions to freshwater (Bassham *et al.*, 2018; Nelson & Cresko, 2018).

An unexpected finding of oceanic-like fish far inland in major Oregon watersheds

Despite the overall pattern of oceanic and freshwater phenotypic divergence, and coastal and inland population genomic divergence, an anomalous phenotypic and genetic clustering of particular inland Oregon populations with coastal groups was noteworthy. These inland oceanic-like populations were not localized to a specific geographical region but were found in the higher elevation extent of the species' range in two major watersheds and in four major rivers within these watersheds. Others have reported extra-plated populations in and around Oregon (Rutter, 1896; Hagen & Gilbertson, 1972), northern California (Baumgartner & Bell, 1984), and in inland regions of Europe and British Columbia (Münzing, 1963; Reimchen *et al.*, 2013). The presence of these morphologies far inland presents different possible scenarios. These patterns could plausibly be the result of recent or episodic introductions by humans of coastal stickleback into these systems (Adachi *et al.*, 2012). Alternatively, or perhaps more likely in addition, selection could have favoured the maintenance of oceanic phenotypes in these clear water habitats, as has been reported in stickleback from British Columbia (Reimchen *et al.*, 2013) and Washington State (Kitano *et al.*, 2008).

A historical survey from 1896 found a small collection of stickleback approximately 193 river kilometres from the sea in the north Umpqua that were fully plated (Rutter, 1896). The fish we collected near this same site were the most plated of any of our Oregon freshwater collections and phenotypically resemble oceanic fish in other traits as well, suggesting this Umpqua River population has maintained a phenotypically oceanic shape for at least 120 years.

Suggestive of human introduction of oceanic fish, stickleback were not found in Leaburg Lake in the 1980s despite intensive sampling at Leaburg Dam by wildlife regulators (Zakel, 1984). Analysis of genetic

divergence shows that the current Leaburg population is more closely related to oceanic populations than it is to the low-plated fish found throughout the Willamette Basin. Strikingly, Leaburg is an order of magnitude more diverged from a population just 63 km away in the same river (F_{ST} = 0.38) than it is from oceanic fish ~773 river kilometres away (F_{ST} = 0.039). Human dispersal of stickleback can occur intentionally where they have been used as bait fish or as mosquito control or have been introduced unintentionally while being transported along with salmonids being released for sport (Adachi *et al.*, 2012). In this case it is possible coastal fish were inadvertently transported far inland to the Leaburg Fish Hatchery.

CONCLUSIONS

Here we present a comprehensive phenotypic and genetic analysis of dozens of populations of threespine stickleback from diverse habitats and geographical regions from across Oregon. Some aspects of the mosaic described here echo what has been reported in other stickleback systems, including a similar pattern of phenotypic oceanic–freshwater divergence as has been seen in younger stickleback systems. However, our findings of a deep coastal–inland geographical divide in population genetic partitioning, and the discovery of introduction in the Willamette Valley, afford an excellent opportunity to connect phenotype and older and more diverse genotypes than has previously been possible in stickleback. This work lays the foundation for future research that may provide a better understanding of the genetics of parallel divergence and local adaptation across spatial and temporal scales.

ACKNOWLEDGMENTS

We thank Steven James, Shira Mali, Savanah Olroyd, Erik Parker, Aimee Schultz, Roberta Torunsky, Jack Peplinski and Mikaeli Dirling for help with collection of phenotypic data. We thank Julian Catchen and Emily Lescak for helpful discussion and comments about this project. We thank Brian Bangs, Paul Olmstead, Sara Akins, Brian Cannon, Randy Wildman, Erik Withalm and Laurie Weitkamp from the Oregon Department of Fish and Wildlife (ODFW) and National Oceanic and Atmospheric Administration (NOAA) for help with procurement of many of the stickleback samples used. We thank the McKenzie River Trust for allowing us to collect samples from their properties. We also thank the three anonymous reviewers for their helpful comments. This research was supported primarily by National Science Foundation 0949053 and IOS 102728 (both to W.A.C.). Additional support came from National Institutes of Health grants 1R24GM079486-01A1 and OD011199 (both to W.A.C.).

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

REFERENCES

Adachi T, **Ishikawa A**, **Mori S**, **Makino W**, **Kume M**, **Kawata M**, **Kitano J. 2012.** Shifts in morphology and diet of non-native sticklebacks introduced into Japanese crater lakes. *Ecology and Evolution* **2:** 1083–1098.

Allen JE, **Baldwin EM. 1944.** Geology and coal resources of the Coos Bay quadrangle, Oregon. *AAPG Bulletin* **28:** 1779–1780.

Andrews K, **Good J**, **Miller M**, **Luikart G**, **Hohenlohe P. 2016.** Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics* **17:** 81–92.

Barrett RDH, **Schluter D. 2008.** Adaptation from standing genetic variation. *Trends in Ecology & Evolution* **23:** 38–44.

Bassham S, **Catchen J**, **Lescak E**, **von Hippel FA**, **Cresko WA. 2018.** Repeated selection of alternatively adapted haplotypes creates sweeping genomic remodeling in stickleback. *Genetics* **209**: 921–939.

Baumgartner JV, **Bell MA. 1984.** Lateral plate morph variation in California populations of the threespine stickleback, *Gasterosteus aculeatus*. *Evolution* **38:** 665–674.

Beaujean AA. 2012. *BaylorEdPsych: R package for Baylor University educational psychology quantitative courses. R package version 0.5*. Available at: http://CRAN. R-project.org/package=BaylorEdPsych

Bell MA. 1987. Interacting evolutionary constraints in pelvic reduction of threespine sticklebacks, *Gasterosteus aculeatus* (Pisces, Gasterosteidae). *Biological Journal of the Linnean Society* **31:** 347–382.

Bell MA. 2001. Lateral plate evolution in the threespine stickleback: getting nowhere fast. *Genetica* **112:** 445–461.

Bell MA. 2009. Implications of a fossil stickleback assemblage for Darwinian gradualism. *Journal of Fish Biology* **75:** 1977–1999.

Bell MA, **Aguirre WE. 2013.** Contemporary evolution, allelic recycling, and adaptive radiation of the threespine stickleback. *Evolutionary Ecology Research* **15:** 377–411.

Bell MA, **Aguirre WE**, **Buck NJ. 2004.** Twelve years of contemporary armor evolution in a threespine stickleback population. *Evolution* **58:** 814–824.

Bell MA, **Foster SA. 1994.** *The evolutionary biology of the threespine stickleback*. Oxford: Oxford University Press.

Bell MA, **Orti G. 1994.** Pelvic reduction in threespine stickleback from Cook Inlet lakes: geographical-distribution and intrapopulation variation. *Copeia* **1994:** 314–325.

Bell MA, **Richkind KE. 1981.** Clinal variation of lateral plates in threespine stickleback fish. *American Naturalist* **117:** 113–132.

Bell MA, **Stewart JD**, **Park PJ. 2009.** The world's oldest fossil threespine stickleback fish. *Copeia* **2009:** 256–265.

Benner PA, **Sedell JR. 1997.** Upper Willamette River landscape: a historic perspective. In: Laenen A, Dunnette D, eds. *River quality: dynamics and restoration*. Boca Raton: CRC Press, 23–47.

Bernhardt RR, **von Hippel FA**, **Cresko WA. 2006.** Perchlorate induces hermaphroditism in threespine sticklebacks. *Environmental Toxicology and Chemistry* **25:** 2087–2096.

Booth DB, **Troost KG**, **Clague JJ**, **Waitt RB. 2003.** The Cordilleran ice sheet. *Developments in Quaternary Sciences* **1:** 17–43.

Casillas S, **Barbadilla A. 2017.** Molecular population genetics. *Genetics* **205:** 1003–1035.

Castric V, **Bonney F**, **Bernatchez L. 2001.** Landscape structure and hierarchical genetic diversity in the brook charr, *Salvelinus fontinalis*. *Evolution* **55:** 1016–1028.

Catchen J, **Bassham S**, **Wilson T**, **Currey M**, **O'Brien C**, **Yeates Q**, **Cresko WA. 2013.** The population structure and recent colonization history of Oregon threespine stickleback determined using restriction-site associated DNA-sequencing. *Molecular Ecology* **22:** 2864–2883.

Catchen JM, **Amores A**, **Hohenlohe P**, **Cresko W**, **Postlethwait JH. 2011.** Stacks: building and genotyping loci de novo from short-read sequences. *G3-Genes Genomes Genetics* **1:** 171–182.

Colosimo PF, **Hosemann KE**, **Batabhadra S**, **Villarreal J**, **Dickson M**, **Grimwood J**, **Schmutz J**, **Myers RM**, **Schluter D**, **Kingsley DM. 2005.** Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* **307:** 1928–1933.

Colosimo PF, **Peichel CL**, **Nereng K**, **Blackman BK**, **Shapiro MD**, **Schluter D**, **Kingsley DM. 2004.** The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biology* **2:** 635–641.

Conte GL, **Arnegard ME**, **Best J**, **Chan YF**, **Jones FC**, **Kingsley DM**, **Schluter D**, **Peichel CL. 2015.** Extent of QTL reuse during repeated phenotypic divergence of sympatric threespine stickleback. *Genetics* **201:** 1189–1200.

Cresko WA, **Amores A**, **Wilson C**, **Murphy J**, **Currey M**, **Phillips P**, **Bell MA**, **Kimmel CB**, **Postlethwait JH. 2004.** Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. *Proceedings of the National Academy of Sciences of the United Staes of America* **101:** 6050–6055.

Cruickshank TE, **Hahn MW. 2014.** Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology* **23:** 3133–3157.

Deagle BE, **Jones FC**, **Chan YF**, **Absher DM**, **Kingsley DM**, **Reimchen TE. 2012.** Population genomics of parallel phenotypic evolution in stickleback across stream–lake ecological transitions. *Proceedings of the Royal Society B: Biological Sciences* **279:** 1277–1286.

Defaveri J, **Merila J. 2013.** Evidence for adaptive phenotypic differentiation in Baltic Sea sticklebacks. *Journal of Evolutionary Biology* **26:** 1700–1715.

Dionne M, **Caron F**, **Dodson JJ**, **Bernatchez L. 2008.** Landscape genetics and hierarchical genetic structure in Atlantic salmon: the interaction of gene flow and local adaptation. *Molecular Ecology* **17:** 2382–2396.

Divino JN, **Monette MY**, **McCormick SD**, **Yancey PH**, **Flannery KG**, **Bell MA**, **Rollins JL**, **von Hippel FA**, **Schultz ET. 2016.** Osmoregulatory physiology and rapid evolution of salinity tolerance in threespine stickleback recently introduced to fresh water. *Evolutionary Ecology Research* **17:** 179–201.

Endler JA. 1977. *Geographic variation, speciation, and clines*. Princeton: Princeton University Press.

Endler JA. 1986. *Natural selection in the wild*. Princeton: Princeton University Press.

Excoffier L, **Smouse PE**, **Quattro JM. 1992.** Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131:** 479–491.

Ferrero-Serrano A, **Assmann SM. 2019.** Phenotypic and genome-wide association with the local environment of *Arabidopsis*. *Nature Ecology & Evolution* **3:** 274–285.

Feulner PG, **Chain FJ**, **Panchal M**, **Eizaguirre C**, **Kalbe M**, **Lenz TL**, **Mundry M**, **Samonte IE**, **Stoll M**, **Milinski M**, **Reusch TB**, **Bornberg-Bauer E. 2013.** Genome-wide patterns of standing genetic variation in a marine population of three-spined sticklebacks. *Molecular Ecology* **22:** 635–649.

Foster SA. 1995. Understanding the evolution of behavior in threespine stickleback: the value of geographic variation. *Behaviour* **132:** 1107–1129.

Funk CW, **McKay JK**, **Hohenlohe PA**, **Allendorf FW. 2012.** Harnessing genomics for delineating conservation units. *Trends in Ecology & Evolution* **27:** 489–496.

Furin CG, **von Hippel FA**, **Postlethwait JH**, **Buck CL**, **Cresko WA**, **O'Hara TM. 2015.** Developmental timing of sodium perchlorate exposure alters angiogenesis, thyroid follicle proliferation and sexual maturation in stickleback. *General and Comparative Endocrinology* **219:** 24–35.

Gaggiotti OE, **Bekkevold D**, **Jørgensen HBH**, **Foll M**, **Carvalho GR**, **Andre C**, **Ruzzante DE. 2009.** Disentangling the effects of evolutionary, demographic, and environmental factors influencing genetic structure of natural populations: Atlantic herring as a case study. *Evolution* **63:** 2939–2951.

Gelmond O, **von Hippel FA**, **Christy MS. 2009**. Rapid ecological speciation in three-spined stickleback *Gasterosteus aculeatus* from Middleton Island, Alaska: the roles of selection and geographic isolation. *Journal of Fish Biology* **75:** 2037–2051.

Gienapp P, **Fior S**, **Guillaume F**, **Lasky J**, **Sork V**, **Csilléry K. 2017.** Genomic quantitative genetics to study evolution in the wild. *Trends in Ecology & Evolution* **32:** 897–908.

Greenwood AK, **Mills MG**, **Wark AR**, **Archambeault SL**, **Peichel CL. 2016**. Evolution of schooling behavior in threespine sticklebacks Is shaped by the *Eda* gene. *Genetics* **203:** 677–681.

Hagen DW, **Gilbertson LG. 1972.** Geographic variation and environmental selection in *Gasterosteus aculeatus* L. in the Pacific Northwest, America. *Evolution* **26:** 32–51.

Hani YMI, **Marchand A**, **Turies C**, **Kerambrun E**, **Palluel O**, Bado-**Nilles A**, **Beaudouin R**, **Porcher JM**, **Geffard A**, **Dedourge-Geffard O. 2018.** Digestive enzymes and gut morphometric parameters of threespine stickleback (*Gasterosteus aculeatus*): influence of body size and temperature. *PLoS One* **13:** e0194932.

Hendry AP, **Hudson K**, **Walker JA**, **Rasanen K**, **Chapman LJ. 2011.** Genetic divergence in morphology-performance mapping between Misty Lake and inlet stickleback. *Journal of Evolutionary Biology* **24:** 23–35.

Hernandez-Martich JD, **Smith MH. 1990.** Patterns of genetic-variation in eastern mosquitofish (*Gambusia holbrooki* Girard) from the piedmont and coastal plain of three drainages. *Copeia* **1990:** 619–630.

Hohenlohe PA, **Bassham S**, **Etter PD**, **Stiffler N**, **Johnson EA**, **Cresko WA. 2010.** Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics* **6:** e1000862.

Jones FC, **Grabherr MG**, **Chan YF**, **Russell P**, **Mauceli E**, **Johnson J**, **Swofford R**, **Pirun M**, **Zody MC**, **White S**, **Birney E**, **Searle S**, **Schmutz J**, **Grimwood J**, **Dickson MC**, **Myers RM**, **Miller CT**, **Summers BR**, **Knecht AK**, **Brady SD**, **Zhang H**, **Pollen AA**, **Howes T**, **Amemiya C**, **Baldwin J**, **Bloom T**, **Jaffe DB**, **Nicol R**, **Wilkinson J**, **Lander ES**, **Di Palma F**, **Lindblad-Toh K**, **Kingsley DM**; **Broad Institute Genome Sequencing P, Whole Genome Assembly T**. **2012.** The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **484:** 55–61.

Kimmel CB, **Hohenlohe PA**, **Ullmann B**, **Currey M**, **Cresko WA. 2012.** Developmental dissociation in morphological evolution of the stickleback opercle. *Evolution & Development* **14:** 326–337.

Kimmel CB, **Ullmann B**, **Walker C**, **Wilson C**, **Currey M**, **Phillips PC**, **Bell MA**, **Postlethwait JH**, **Cresko WA. 2005.** Evolution and development of facial bone morphology in threespine sticklebacks. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 5791–5796.

Kitano J, **Bolnick DI**, **Beauchamp DA**, **Mazur MM**, **Mori S**, **Nakano T**, **Peichel CL. 2008.** Reverse evolution of armor plates in the threespine stickleback. *Current Biology* **18:** 769–774.

Klepaker T. 1993. Morphological changes in a marine population of threespined stickleback, *Gasterosteus aculeatus*, recently isolated in fresh water. *Canadian Journal of Zoology* **71:** 1251–1258.

Kopelman NM, **Mayzel J**, **Jakobsson M**, **Rosenberg NA**, **Mayrose I. 2015.** Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources* **15:** 1179–1191.

Lescak EA, **Bassham SL**, **Catchen J**, **Gelmond O**, **Sherbick ML**, **von Hippel FA**, **Cresko WA. 2015.** Evolution of stickleback in 50 years on earthquake-uplifted islands. *Proceedings of the National Academy of Sciences of the United States of America* **112:** E7204–E7212.

Lucek K, **Sivasundar A**, **Kristjansson BK**, **Skulason S**, **Seehausen O. 2014.** Quick divergence but slow convergence during ecotype formation in lake and stream stickleback pairs of variable age. *Journal of Evolutionary Biology* **27:** 1878–1892.

Marques DA, **Lucek K**, **Meier JI**, **Mwaiko S**, **Wagner CE**, **Excoffier L**, **Seehausen O. 2016.** Genomics of rapid incipient speciation in sympatric threespine stickleback. *PLoS Genetics* **12:** e1005887.

McCairns RJ, **Bernatchez L. 2008.** Landscape genetic analyses reveal cryptic population structure and putative selection gradients in a large-scale estuarine environment. *Molecular Ecology* **17:** 3901–3916.

McGuigan K, **Nishimura N**, **Currey M**, **Hurwit D**, **Cresko WA. 2011.** Cryptic genetic variation and body size evolution in threespine stickleback. *Evolution* **65:** 1203–1211.

Mcphail JD. 1984. Ecology and evolution of sympatric sticklebacks (*Gasterosteus*): morphological and genetic-evidence for a species pair in Enos Lake, British Columbia. *Canadian Journal of Zoology* **62:** 1402–1408.

Meirmans PG, **Van Tienderen PH. 2004.** GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes* **4:** 792–794.

Miller SE, **Roesti M**, **Schluter D. 2019.** A single interacting species leads to widespread parallel evolution of the stickleback genome. *Current Biology* **29:** 530–537.

Morris MRJ, **Bowles E**, **Allen BE**, **Jamniczky HA**, **Rogers SM. 2018.** Contemporary ancestor? Adaptive divergence from standing genetic variation in Pacific marine threespine stickleback. *BMC Evolutionary Biology* **18:** 113.

Münzing J. 1963. The evolution of variation and distributional patterns in European populations of the three-spined stickleback, *Gasterosteus aculeatus*. *Evolution* **17:** 320–332.

Nadeau NJ, **Whibley A**, **Jones RT**, **Davey JW**, **Dasmahapatra KK**, **Baxter SW**, **Quail MA**, **Joron M**, **ffrench-Constant RH**, **Blaxter ML**, **Mallet J**, **Jiggins CD. 2012.** Genomic islands of divergence in hybridizing *Heliconius* butterflies identified by large-scale targeted sequencing. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367:** 343–353.

Nelson TC, **Cresko WA. 2018.** Ancient genomic variation underlies repeated ecological adaptation in young stickleback populations. *Evolution Letters* **2:** 9–21.

Noor MA, **Bennett SM. 2009.** Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. *Heredity* **103:** 439–444.

Nosil P, **Crespi BJ**, **Sandoval CP. 2002.** Host-plant adaptation drives the parallel evolution of reproductive isolation. *Nature* **417:** 440–443.

Nosil P, **Egan SP**, **Funk DJ. 2008.** Heterogeneous genomic differentiation between walking-stick ecotypes: "isolation by adaptation" and multiple roles for divergent selection. *Evolution* **62:** 316–336.

Nosil P, **Schluter D. 2011.** The genes underlying the process of speciation. *Trends in Ecology & Evolution* **26:** 160–167.

Pallares LF, **Harr B**, **Turner LM**, **Tautz D. 2014.** Use of a natural hybrid zone for genomewide association mapping of craniofacial traits in the house mouse. *Molecular Ecology* **23:** 5756–5770.

Peterson CD, **Stock E**, **Price DM**, **Hart R**, **Reckendorf F**, **Erlandson JM**, **Hostetler SW. 2007.** Ages, distributions, and origins of upland coastal dune sheets in Oregon, USA. *Geomorphology* **91:** 80–102.

Pinheiro J, **Bates D**, **DebRoy S**, **Sarkar D**, **authors E**, **Heisterkamp S**, **Van Willigen B**, **R Core Team**. **2019.** *nlme: linear and nonlinear mixed effects models. R package version 3*. Available at: https://cran.r-project.org/web/packages/nlme/index.html

Pritchard JK, **Stephens M**, **Donnelly P. 2000.** Inference of population structure using multilocus genotype data. *Genetics* **155:** 945–959.

R Development Core Team. **2013.** *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Reimchen TE, **Bergstrom C**, **Nosil P. 2013.** Natural selection and the adaptive radiation of Haida Gwaii stickleback. *Evolutionary Ecology Research* **15:** 241–269.

Reimchen TE, **Nosil P. 2006.** Replicated ecological landscapes and the evolution of morphological diversity among *Gasterosteus* populations from an archipelago on the west coast of Canada. *Canadian Journal of Zoology* **84:** 643–654.

Reist JD. 1985. An empirical evaluation of several univariate methods that adjust for size variation in morphometric data. *Canadian Journal of Zoology* **63:** 1429–1439.

Rosenberg NA, **Pritchard JK**, **Weber JL**, **Cann HM**, **Kidd KK**, **Zhivotovsky LA**, **Feldman MW. 2002.** Genetic structure of human populations. *Science* **298:** 2381–2385.

Rundle HD, **Nagel L**, **Boughman JW**, **Schluter D. 2000.** Natural selection and parallel speciation in sympatric sticklebacks. *Science* **287:** 306–308.

Rutter C. 1896. *Notes on fresh water fishes on the Pacific slope of North America*. Palo Alto: Leland Stanford Jr. University.

Saltz J, **Hessel F**, **Kelly M. 2017.** Trait correlations in the genomics era. *Trends in Ecology & Evolution* **32:** 279–290.

Sanchez-Gonzales S, **Ruiz-Campos G**, **Contreras-Balderas S. 2001.** Feeding ecology and habitat of the threespine stickleback, *Gasterosteus aculeatus microcephalus*, in a remnant population of northwestern Baja California, Mexico. *Ecology of Freshwater Fish* **10:** 191–197.

Schluter D. 1993. Adaptive radiation in sticklebacks: size, shape, and habitat use efficiency. *Ecology* **74:** 699–709.

Schluter D. 2000. *The ecology of adaptive radiation*. Oxford: Oxford University Press.

Schluter D, **Conte GL. 2009.** Genetics and ecological speciation. *Proceedings of the National Academy of Sciences of the United States of America* **106:** 9955–9962.

Schluter D, **Marchinko KB**, **Barrett RD**, **Rogers SM. 2010.** Natural selection and the genetics of adaptation in threespine stickleback. *Philosophical Transactions of the Royal Society B-Biological Sciences* **365:** 2479–2486.

Schneider CA, **Rasband WS**, **Eliceiri KW. 2012.** NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9:** 671–675.

Shapiro MD, **Marks ME**, **Peichel CL**, **Blackman BK**, **Nereng KS**, **Jonsson B**, **Schluter D**, **Kingsley DM. 2004.** Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature* **428:** 717–723.

Shaw PW, **Carvalho GR**, **Magurran AE**, **Seghers BH. 1991.** Population differentiation in Trinidadian guppies (*Poecilia reticulata*): patterns and problems. *Journal of Fish Biology* **39:** 203–209.

Soria-Carrasco V, **Gompert Z**, **Comeault AA**, **Farkas TE**, **Parchman TL**, **Johnston JS**, **Buerkle CA**, **Feder JL**, **Bast J**, **Schwander T**, **Egan SP**, **Crespi BJ**, **Nosil P. 2014.** Stick insect genomes reveal natural selection's role in parallel speciation. *Science* **344:** 738–742.

Spence R, **Wootton RJ**, **Barber I**, **Przybylski M**, **Smith C. 2013.** Ecological causes of morphological evolution in the three-spined stickleback. *Ecology and Evolution* **3:** 1717–1726.

Stacklies W, **Redestig H**, **Scholz M**, **Walther D**, **Selbig J. 2007.** pcaMethods – a bioconductor package providing PCA methods for incomplete data. *Bioinformatics* **23:** 1164–1167.

Storz JF. 2002. Contrasting patterns of divergence in quantitative traits and neutral DNA markers: analysis of clinal variation. *Molecular Ecology* **11:** 2537–2551.

Teigen LE, **Orczewska JI**, **McLaughlin J**, **O'Brien KM. 2015.** Cold acclimation increases levels of some heat shock protein and sirtuin isoforms in threespine stickleback. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* **188:** 139–147.

Terekhanova NV, **Logacheva MD**, **Penin AA**, **Neretina TV**, **Barmintseva AE**, **Bazykin GA**, **Kondrashov AS**, **Mugue NS. 2014.** Fast evolution from precast bricks: genomics of young freshwater populations of threespine stickleback *Gasterosteus aculeatus*. *PLoS Genetics* **10:** e1004696.

Tinbergen N. 1950. The hierarchical organization of nervous mechanisms underlying instinctive behaviour. In: Houck LD, Drickamer LC, eds. *Symposium for the Society for Experimental Biology* 4. Chicago: The University of Chicago Press, 406–413.

Walker JA. 1997. Ecological morphology of lacustrine threespine stickleback *Gasterosteus aculeatus* L (Gasterosteidae) body shape. *Biological Journal of the Linnean Society* **61:** 3–50.

Wu TD, **Nacu S. 2010.** Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics* **26:** 873–881.

Wund MA, **Baker JA**, **Golub JL**, **Foster SA. 2015.** The evolution of antipredator behaviour following relaxed and reversed selection in Alaskan threespine stickleback fish. *Animal Behaviour* **106:** 181–189.

Zakel JC. 1984. *Downstream migration of fish at Leaburg Dam, McKenzie River, Oregon, 1980 to 1983*. Salem: Information reports / Fish Division, Oregon Department of Fish and Wildlife.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's website.

Figure S1. Morphometric measurements covering aspects of trophic, defensive and size traits. A, lateral linear measurements: jaw length, eye orbit diameter, opercle height, opercle width, cranial length, body depth, ascending process, pelvic spine length, standard length and head angle. B, linear measurements of the pelvic structure: pelvic structure length and width, and pelvic spine length. The pelvic structure in A and C has been shaded slightly darker for easier identification. C, counts of long gill rakers. Standard length (SL) was measured from the rostral-most extent of the upper lip to the posterior end of the caudal peduncle where it meets the caudal fin rays. Body depth was measured from the mid-point of the pelvic joint articulating element to the dorsal outline of the fish. Length of the second dorsal and the left pelvic spine was measured from the tip of the spine to the flange at the base of the spine, next to its articulation with basal elements. The height of the ascending process, which extends from the ventral pelvic structure to the lateral plates, was measured from the mid-point of the articulation between the spine and the pelvic plate to the most dorsal tip of the process. In cases where the ascending process ended in multiple cusps, the posterior-most cusp was used. Pelvic structure length was measured from the anterior point of the left anterior process to the caudal tip of the left posterior process. Pelvic structure width was taken by measuring the distance between the inner edges of the two pelvic spine joints. Eye orbit diameter was measured across its widest diameter from a ventral point at the suture between the two suborbital bones. Jaw length was measured from the anterior-most tip of the premaxilla to where the premaxilla and the maxilla meet posteriorly. Dorsal cranial length was measured from the anterior-most tip of the nasal bone to the posterior-most tip of the frontal. Opercle height was measured from the dorso-anterior joint of the opercle to the ventral-most point, and width was measured from the joint to the posterior-most point.

Figure S2. Phenotypic differences between oceanic and freshwater forms in Oregon threespine stickleback shown by box plots of PCs 1–3. A, all phenotypic traits included. B, all traits except the number of lateral plates and gill rakers. Asterisks denote significant differences found using ANOVA.

Figure S3. Phenotypic PCA of Oregon populations excluding lateral plate and gill raker counts. A, phenotypic PCA partitions phenotypic variation between oceanic and freshwater stickleback populations along PC 2. B, dorsal and pelvic spine length and pelvic girdle length and width drive these differences.

Figure S4. Genetic PCA of Willamette Valley populations. Genetic variation is partitioned between low- and extra-plated populations. The only significant principal component (PC 1) explained 64% of the genetic variation and partitions variation between fish with few lateral plates and those with average lateral plate counts between eight and 33. Average PC 1 and PC 2 scores of each population are plotted, with bars showing one standard error. Symbols are colour coded for the average number of plates of each population.

Figure S5. Willamette Basin populations are genetically structured by drainage, despite connectivity of habitats. STRUCTURE groupings are plotted for multiple levels of *K*. Each vertical bar represents an individual colour-coded by its posterior probability of membership to a group. A, extra- and low-plated populations. At low levels of *K* populations are partitioned between average low- and extra-plated populations with increased partitioning with increasing levels of *K.* B, low-plated populations. At $K = 3$, Jont Creek, Dougren Slough and Lynx Hollow all group independently while populations downstream display individuals with mixed memberships.

Figure S6. Box plots demonstrating significant differences between oceanic and freshwater forms and between regions among Oregon and Alaska threespine stickleback populations in PC 1–3.

Figure S7. Oregon and Alaska oceanic populations demonstrate phenotypic differences. PCA of phenotypic traits comparing Oregon and Alaska oceanic populations.