Supplementary Material

Gene flow and selection interact to promote adaptive divergence in regions of low recombination

Author Information:

Samuk, Kieran (corresponding author, ksamuk@gmail.com)
Duke University, Biology
Durham, NC, USA

Owens, Greg L
University of British Columbia, Botany
Vancouver, BC, CAN

Delmore, Kira E
Max-Planck-Institut fur Evolutionsbiologie
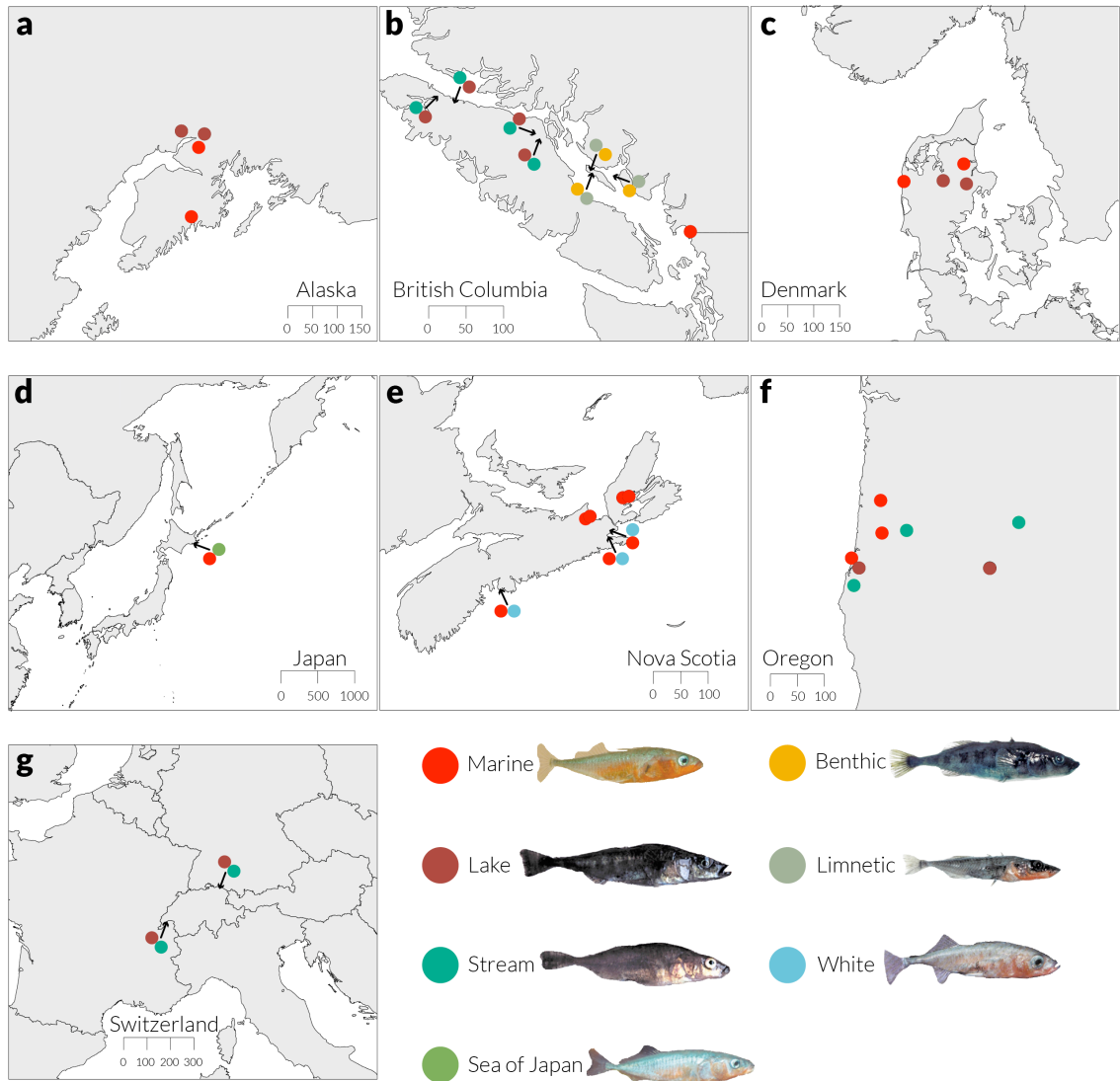Plon, Schleswig-Holstein, DE

Miller, Sara
Cornell University, Neurobiology and Behavior
Ithaca, NY, USA

Rennison, Diana J
Universitat Bern Institut fur Okologie und Evolution
Bern, BE, CH

Schluter, Dolph
University of British Columbia, Biodiversity Research Centre and Department of Zoology
Vancouver, BC, CAN

Figure S1 | Collection locations of stickleback populations used in the study.
Ecotypes are color-coded. Arrows indicate locations where two ecotypes/populations are found in near or complete sympatry. Scale bars indicate distances in kilometers. Populations shown as inland (e.g. the Oregon populations, sixth panel) are found in lake or streams. Photographs obtained from previously published figures, see Table S1 for further population sample information.
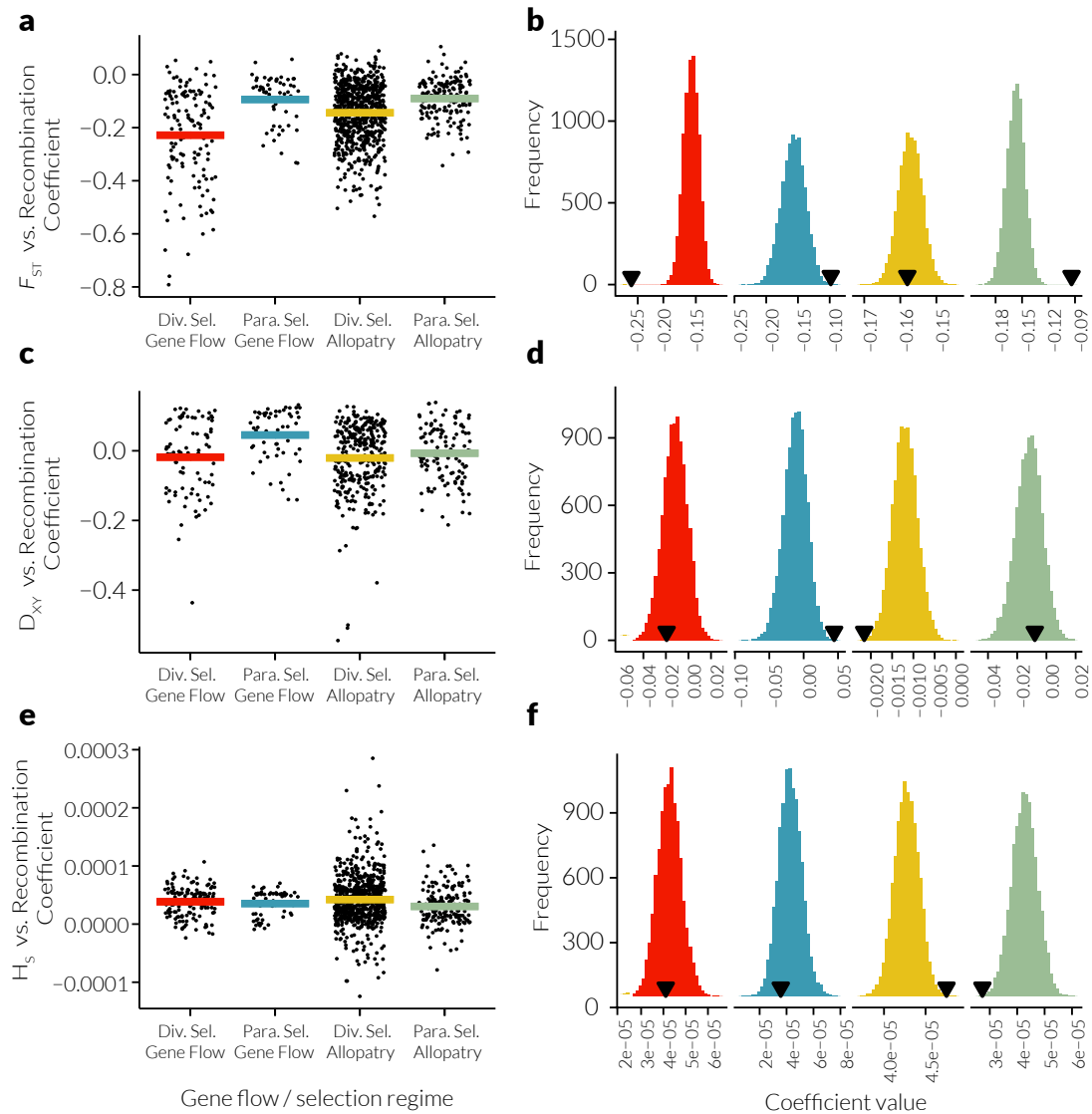
Figure S2 | Low recombination tendency estimates and permutation significance tests for three population genetic parameters across the four categories of population pairs, differing in gene flow and selection regimes. (a,c,e) Each dot represents a coefficient derived from a single pairwise comparison, measuring F_{ST}, D_{XY} or H_S (mean intra-population heterozygosity), with the colored line representing the category mean. (b,d,f) Null expectations (histograms) and observed values (black arrows) for permutation tests of the significance of the differences in mean recombination tendency seen in a, c, and e respectively. Observed means in the tails of the distributions indicate significance (see main text for P values).
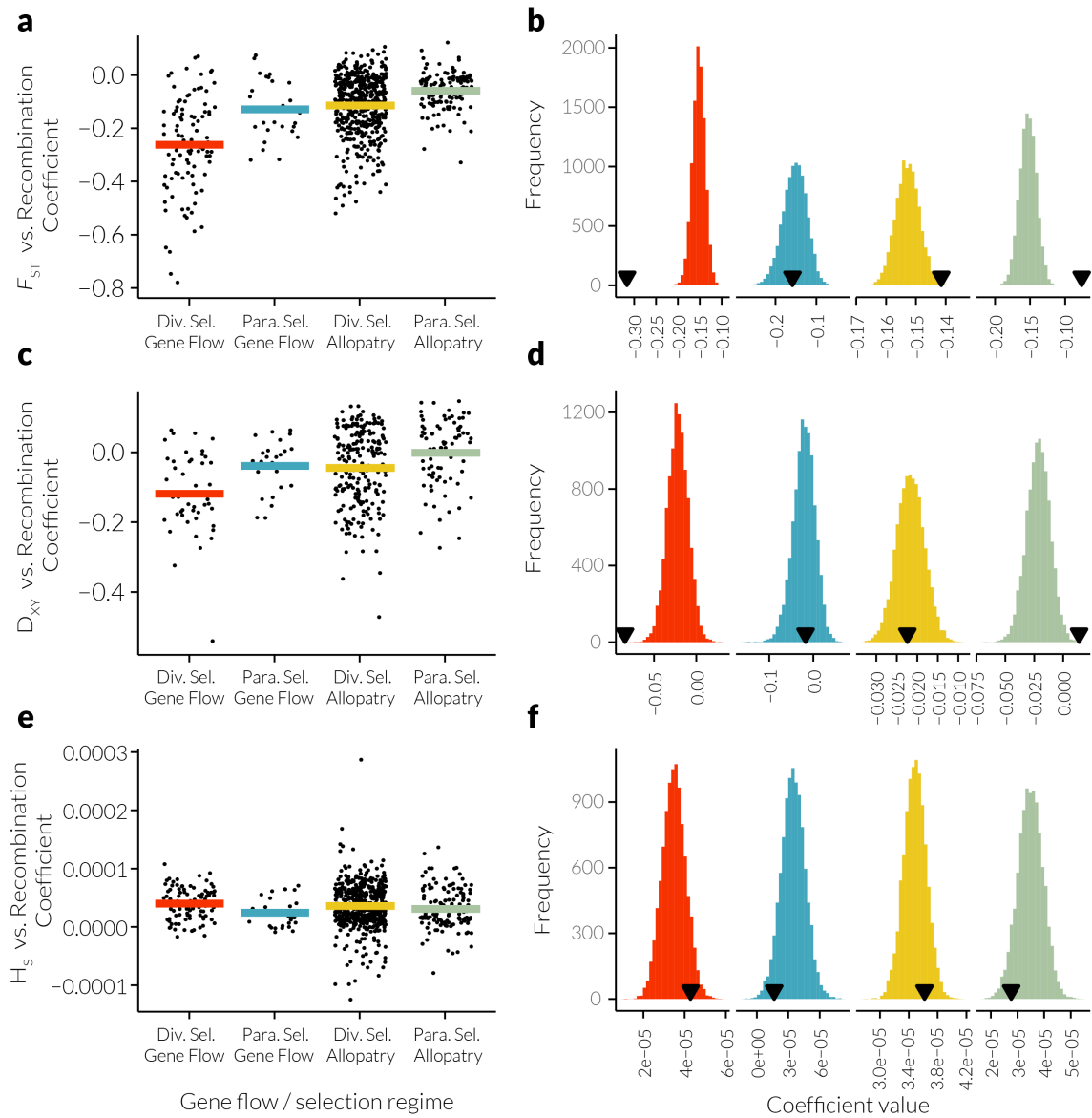
Figure S3 | Low recombination tendency estimates and permutation significance tests for three population genetic parameters across the four categories of population pairs, differing in gene flow and selection regimes. Comparisons were restricted to those with intermediate genome-wide mean FST (~0.18-0.67) (a,c,e) Each dot represents a coefficient derived from a single pairwise comparison, measuring F_{ST}, D_{XY} or H_S (mean intra-population heterozygosity), with the colored line representing the category mean. (b,d,f) Null expectations (histograms) and observed values (black arrows) for permutation tests of the significance of the differences in mean recombination tendency seen in a, c, and e respectively. Observed means in the tails of the distributions indicate significance (see main text for P values).
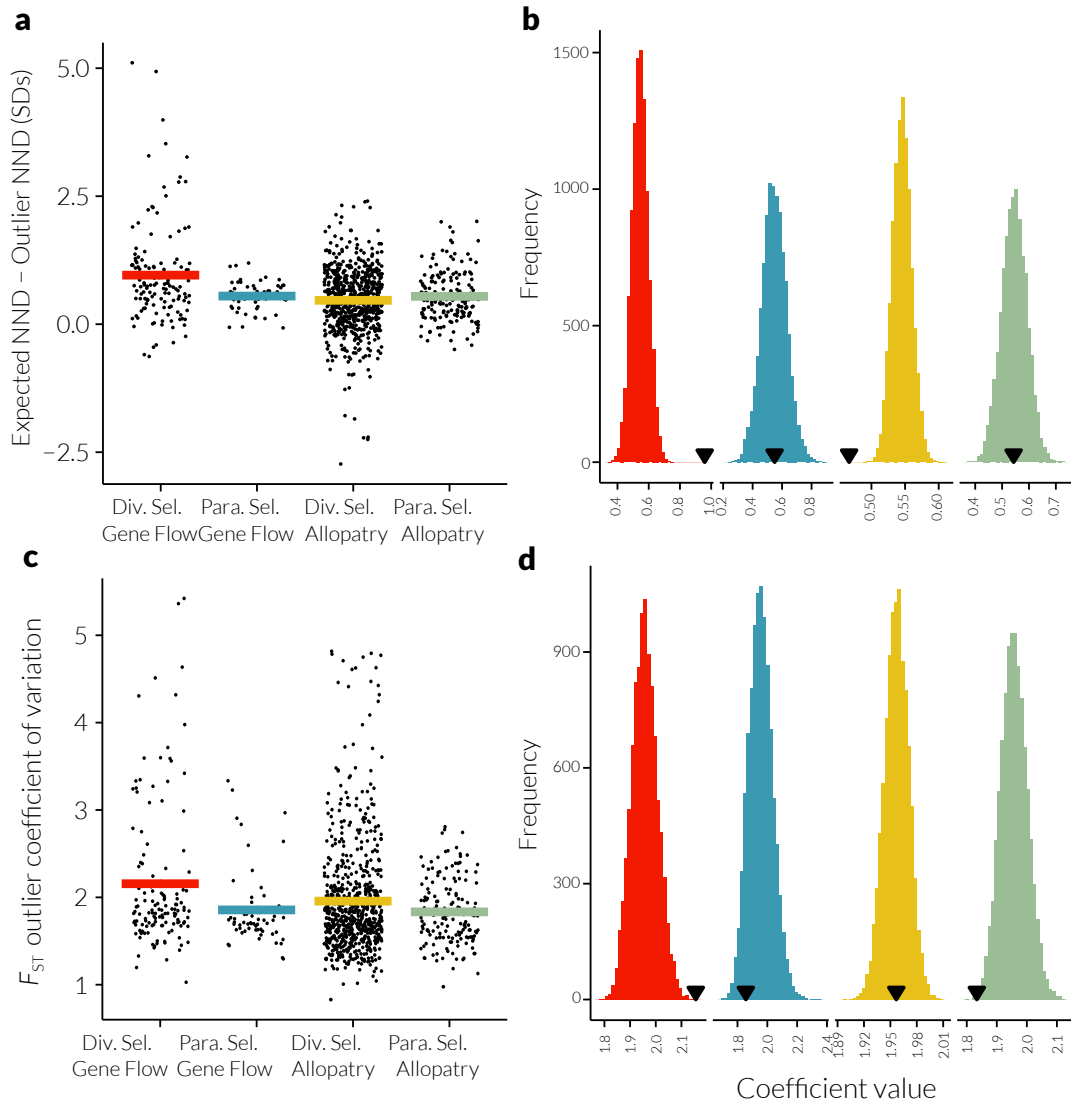
Figure S4 | Clustering estimates and permutation significance tests for two metrics of clustering across different gene flow / selection regimes. (a,c) Each dot represents a clustering metric averaged across all chromosomes in a single comparison, with the colored line representing the mean estimate for that regime. (b,d) Null expectations (histograms) and observed values (black arrows) for permutation tests of the significance of the differences in mean recombination tendency seen in a and b respectively. Observed means in the tails of the distributions indicate significance (see main text for *P* values).
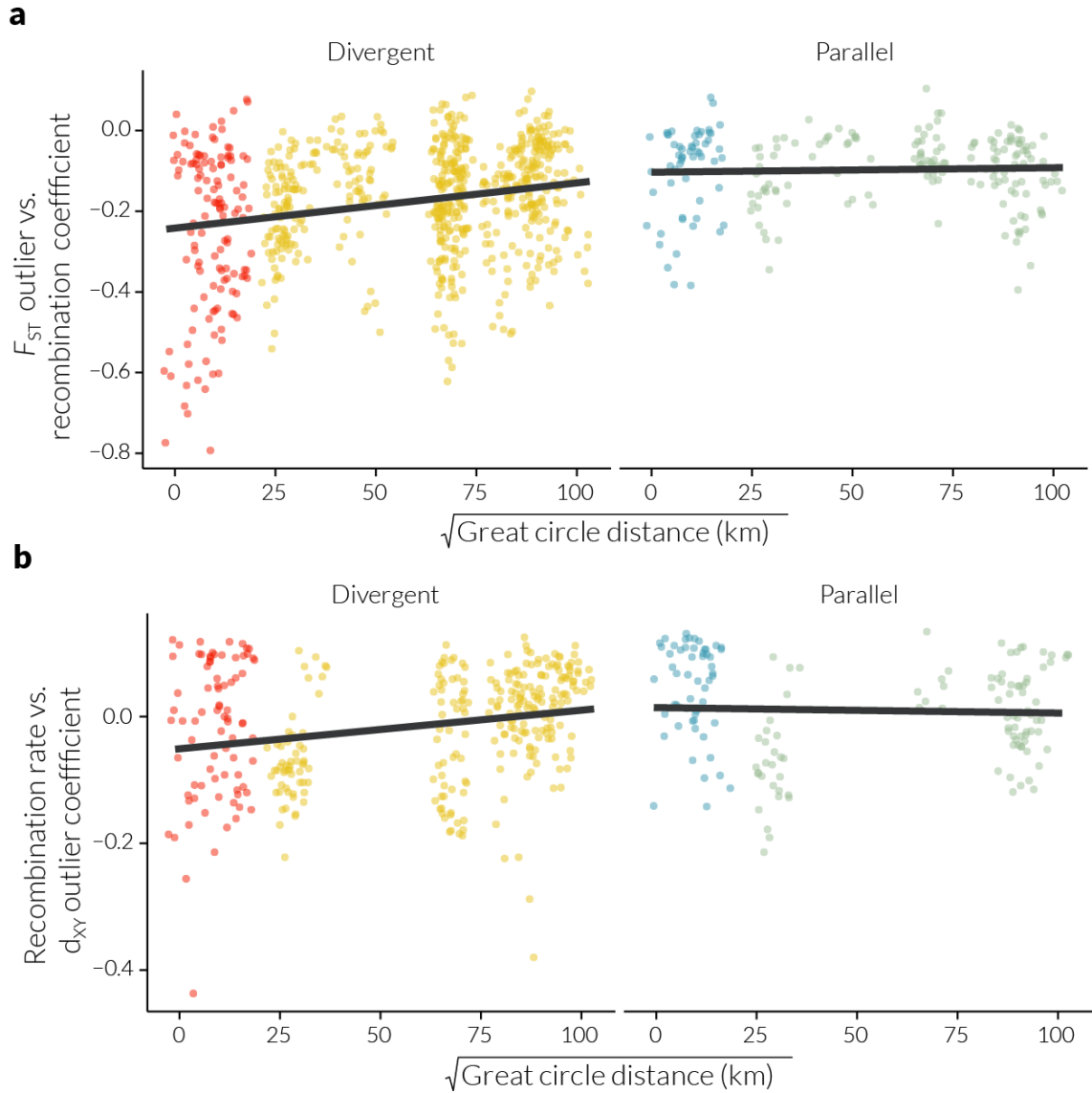
Figure S5 | The tendency for outliers to occur in regions of low recombination as a function of geographic distance between populations. (a) FST outliers and (b) dxy outliers. Individual dots represent the regression coefficients of a logistic regression of outlier tendency vs. recombination rate from single comparisons. Points are colored according to their qualitative gene flow / selection regimes. Great circle distance is the shortest geographic distance between two populations (equivalent to Euclidean distance, but on the surface of a sphere).

Figure S6 | Genome wide divergence and diversity metrics for the comparison between Priest Lake benthic and limnetic populations. Each row depicts a single metric across the chromosomes of the stickleback genome (columns), excluding the sex chromosome. Black lines represent loess smooths of each metric, with the raw metric depicted in color in the background. "Hs" denotes average intrapopulation heterozygosity (similar to nucleotide diversity, pi, averaged across the two populations). Chromosomes are scaled to equal size to aid in visual comparison.

MILLPORT SLOUGH , OR, (MARINE) VS. SOUTH JETTY, OR, (MARINE) MEAN F_{ST} = 0.10

Figure S7|Genome wide divergence and diversity metrics for the comparison between Millport Slough marine and South Jetty marine populations. Each row depicts a single metric across the chromosomes of the stickleback genome (columns), excluding the sex chromosome. Black lines represent loess smooths of each metric, with the raw metric depicted in color in the background. "Hs" denotes the average intrapopulation heterozygosity (similar to nucleotide diversity, pi, averaged across the two populations). Chromosomes are scaled to equal size to aid in visual comparison.
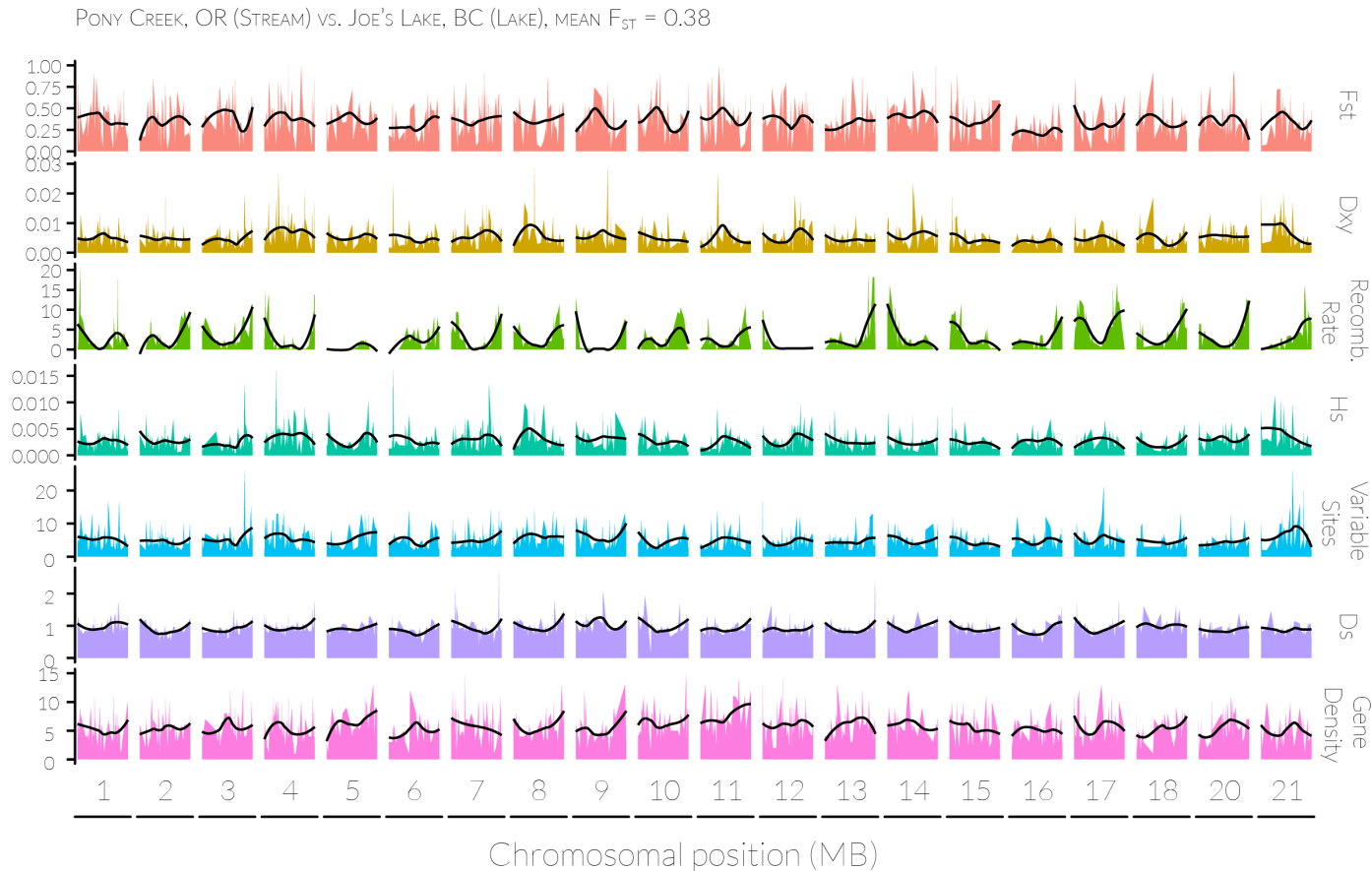
Figure S8|Genome wide divergence and diversity metrics for the comparison between Pony Creak stream and Joe's Lake lake populations. Each row depicts a single metric across the chromosomes of the stickleback genome (columns), excluding the sex chromosome. Black lines represent loess smooths of each metric, with the raw metric depicted in color in the background. "Hs" denotes the average intrapopulation heterozygosity (similar to nucleotide diversity, pi, averaged across the two populations). Chromosomes are scaled to equal size to aid in visual comparison.

Figure S9|Genome wide divergence and diversity metrics for the comparison between Akkeshi maine and Lemvig marine populations. Each row depicts a single metric across the chromosomes of the stickleback genome (columns), excluding the sex chromosome. Black lines represent loess smooths of each metric, with the raw metric depicted in color in the background. "Hs" denotes the average intrapopulation heterozygosity (similar to nucleotide diversity, pi, averaged across the two populations). Chromosomes are scaled to equal size to aid in visual comparison.
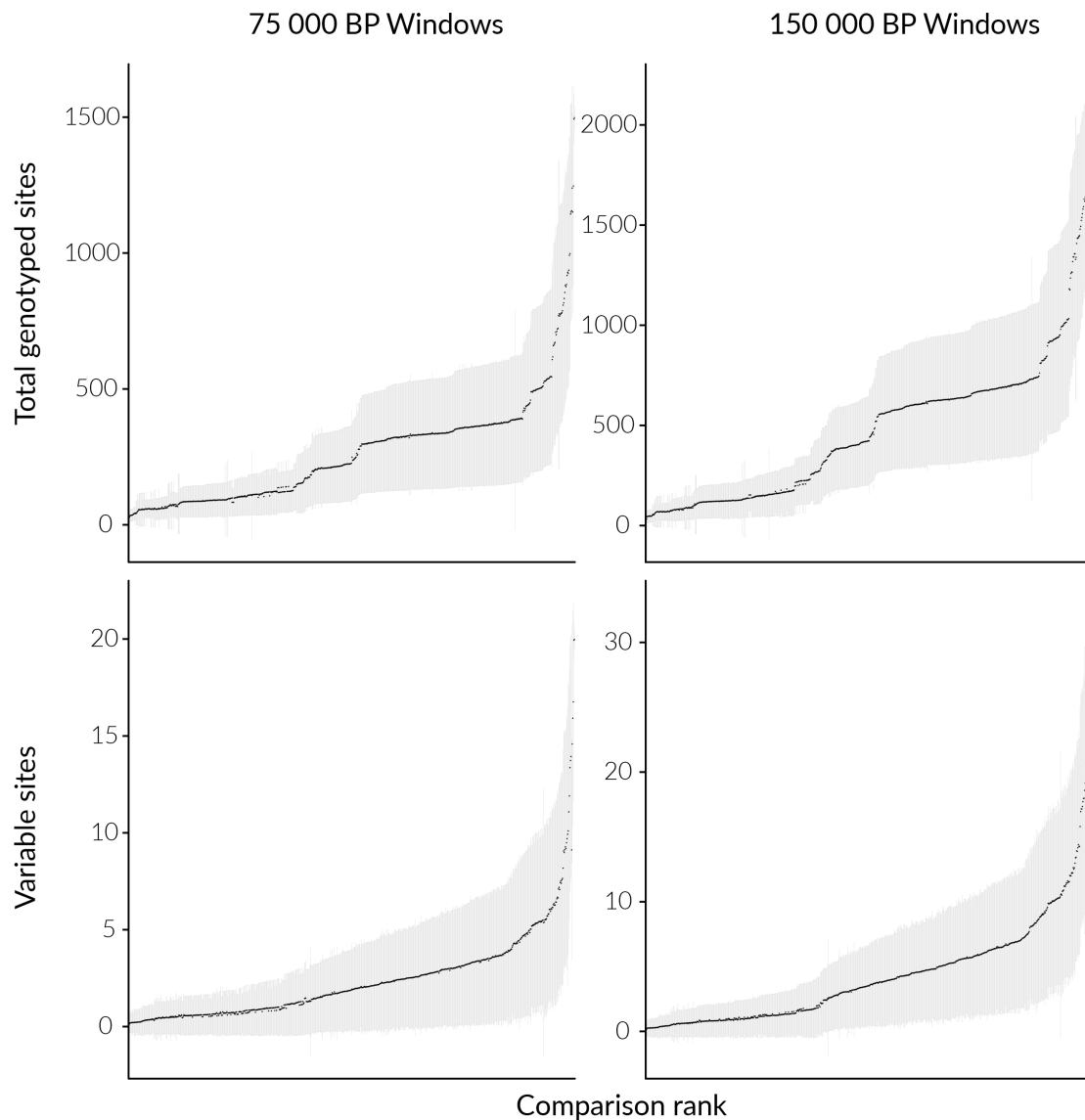
Figure S10 | Variation in total genotyped sites (top row) and variable sites (bottom row) for the stickleback comparisons, for 75kbp (left column) and 150kbp windows (right column). Each point corresponds to a single population comparison, and represents either the mean total genotyped or variable sites <u>shared</u> between the populations across all windows in the genome. Grey bars represent standard deviation. We required at least 3 variable sites in order to include a window in our analysis.
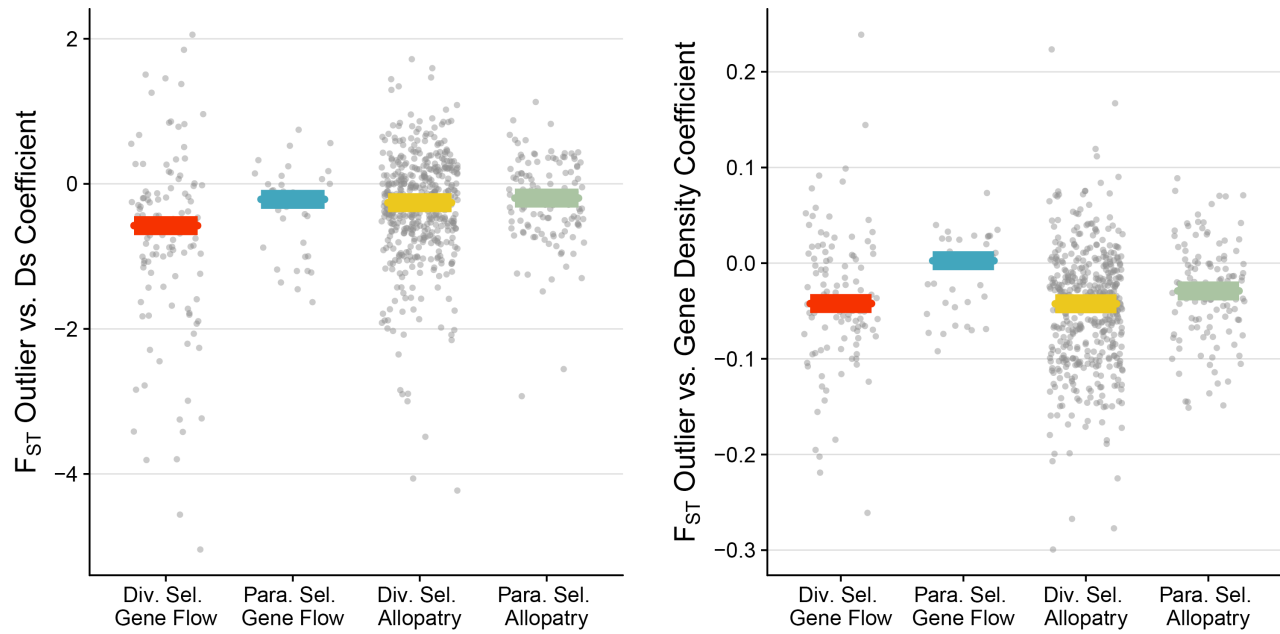
Figure S11| The distribution of GLM coefficients for Ds (left panel) and gene density (right panel) among the four gene flow selection regimes. DS-GF was not significantly different in mean GLM coefficient from the other comparisons for either Ds or Gene density (permutation test, DS: P = 0.12, Gene density: P = 0.332).

Table S1 Collection locations, names and metadata for all samples included in the study. Citations for each study are noted for the first occurrence of the study only. Data source abbreviates: Short Read Archive (SRA), the European Nucleotide Archive (ENA) and the Databank of Japan Sequence Read Archive (DRA). Sequencing technology abbreviations: Whole Genome Sequencing (WGS), Restriction Amplified Digest (RAD), Genotyping-by-Sequencing (GBS). BC refers to British Columbia, Canada.

Study	Latitude	Longitude	Region	Population	Ecotype	Data Source	Accession No.	Technology	No. Individuals
Yoshida	43.054	144.894	Japan	Japan	Sea of Japan	DRA	DRA001136	WGS	8
Yoshida	43.054	144.894	Japan	Japan	Marine	DRA	DRA001136	WGS	8
Catchen	43.145	-124.190	Oregon	Winchester Creek	Stream	SRA	SRA070979	RAD	22
Catchen	43.424	-121.153	Oregon	Pony Creek Reservoir	Lake	SRA	SRA070979	RAD	68
Catchen	43.427	-121.153	Oregon	Paulina Lake	Lake	SRA	SRA070979	RAD	22
Catchen	43.430	-124.076	Oregon	South Twin Lake	Lake	SRA	SRA070979	RAD	50
Catchen	43.592	-124.243	Oregon	Cushman Slough	Marine	SRA	SRA070979	RAD	98
Catchen	44.000	-123.563	Oregon	South Jetty	Marine	SRA	SRA070979	RAD	96
Catchen	44.043	-123.012	Oregon	Riverbend	Stream	SRA	SRA070979	RAD	140
Catchen	44.172	-120.504	Oregon	Crooked River	Stream	SRA	SRA070979	RAD	24
Catchen	44.531	-123.593	Oregon	Millport_Slough	Marine	SRA	SRA070979	RAD	68
Samuk	44.499	-63.903	Nova Scotia	Canal Lake	Marine	UBC	PRJNA388325	GBS	12
Samuk	44.499	-63.903	Nova Scotia	Canal Lake	White	UBC	PRJNA388325	GBS	15
Samuk	45.353	-61.473	Nova Scotia	Salmon River Estuary	Marine	UBC	PRJNA388325	GBS	14
Samuk	45.353	-61.473	Nova Scotia	Salmon River Estuary	White	UBC	PRJNA388325	GBS	17
Samuk	45.458	-61.612	Nova Scotia	Milford Haven	Marine	UBC	PRJNA388325	GBS	9
Samuk	45.458	-61.612	Nova Scotia	Milford Haven	White	UBC	PRJNA388325	GBS	7
Samuk	45.632	-61.960	Nova Scotia	Antigonish Landing	Marine	UBC	PRJNA388325	GBS	16
Samuk	45.672	-61.861	Nova Scotia	Captain's Pond	Marine	UBC	PRJNA388325	GBS	30
Samuk	45.970	-61.119	Nova Scotia	Skye River	Marine	UBC	PRJNA388325	GBS	15
Samuk	45.992	-60.985	Nova Scotia	Little Narrows	Marine	UBC	PRJNA388325	GBS	25
Jones	49.013	-122.778	BC	Little Campbell River	Marine	SRA	PRJNA247503	WGS	5
Rennison	49.663	-124.109	BC	Little Quarry Lake	Benthic	UBC	SRP107890	GBS	20
Rennison	49.663	-124.109	BC	Little Quarry Lake	Limnetic	UBC	SRP107890	GBS	10
Rennison	49.709	-124.525	BC	Paxton Lake	Limnetic	UBC	SRP107890	GBS	20

Rennison	49.709	-124.525	BC	Paxton Lake	Benthic	UBC	SRP107890	GBS	20
Rennison	49.745	-124.566	BC	Priest Lake	Limnetic	UBC	SRP107890	GBS	20
Rennison	49.745	-124.566	BC	Priest Lake	Benthic	UBC	SRP107890	GBS	20
Roesti[23]	46.205	6.544	Switzerland	Lake Geneva	Stream	SRA	SRP007695	RAD	27
Roesti	46.313	6.344	Switzerland	Lake Geneva	Lake	SRA	SRP007695	RAD	27
Roesti	47.332	9.225	Switzerland	Lake Constance	Lake	SRA	SRP007695	RAD	27
Roesti	47.333	9.164	Switzerland	Lake Constance	Stream	SRA	SRP007695	RAD	27
Roesti	50.022	-125.336	BC	Boot Lake	Stream	SRA	SRP007695	RAD	27
Roesti	50.030	-125.323	BC	Boot Lake	Lake	SRA	SRP007695	RAD	26
Roesti	50.134	-125.331	BC	Roberts Lake	Lake	SRA	SRP007695	RAD	27
Roesti	50.143	-125.352	BC	Roberts Lake	Stream	SRA	SRP007695	RAD	27
Roesti	50.363	-127.156	BC	Misty Lake	Lake	SRA	SRP007695	RAD	27
Roesti	50.365	-127.322	BC	Joes Lake	Stream	SRA	SRP007695	RAD	26
Roesti	50.366	-127.170	BC	Misty Lake	Stream	SRA	SRP007695	RAD	27
Roesti	50.373	-127.291	BC	Joes Lake	Lake	SRA	SRP007695	RAD	27
Schluter	49.709	-124.525	BC	Paxton Lake	Limnetic	UBC	SRP107890	WGS	1
Schluter	49.709	-124.525	BC	Paxton Lake	Benthic	UBC	SRP107890	WGS	1
Schluter	49.745	-124.566	BC	Priest Lake	Limnetic	UBC	SRP107890	WGS	1
Schluter	49.745	-124.566	BC	Priest Lake	Benthic	UBC	SRP107890	WGS	1
Chain	56.369	8.182	Denmark	Lemvig	Marine	ENA	PRJEB2954	WGS	6
Ferchaud	56.330	10.048	Denmark	Hadsten Lake	Lake	SRA	SRX437379	RAD	20
Ferchaud	56.383	9.354	Denmark	Hald Lake	Lake	SRA	SRX437379	RAD	20
Ferchaud	56.663	9.969	Denmark	Mariager Inlet	Marine	SRA	SRX437379	RAD	20
Hohenlohe	60.127	-149.406	Alaska	Resurrection Bay	Marine	SRA	SRP001747	RAD	20
Hohenlohe	61.330	-149.151	Alaska	Rabbit Slough	Marine	SRA	SRP001747	RAD	16
Hohenlohe	61.563	-148.949	Alaska	Mud Lake	Lake	SRA	SRP001747	RAD	19
Hohenlohe	61.614	-149.756	Alaska	Bear Paw Lake	Lake	SRA	SRP001747	RAD	28
Miller	48.874	-123.509	BC	Bullock Lake	Lake	UBC	PRJNA387728	WGS	1
Miller	49.153	-123.828	BC	Hoggan Lake	Lake	UBC	PRJNA387728	WGS	1
Miller	48.782	-123.444	BC	Stowell Lake	Lake	UBC	PRJNA387728	WGS	1
Miller	49.508	-123.876	BC	Trout Lake	Lake	UBC	PRJNA387728	WGS	1

Miller	48.874	-123.509	BC	Bullock Lake	Lake	UBC	PRJNA387728	WGS	1
Miller	49.696	-124.506	BC	Cranby Lake	Lake	UBC	PRJNA387728	WGS	1
Miller	49.191	-122.655	BC	West Creek	Marine	UBC	PRJNA387728	WGS	1
Miller	50.373	-125.929	BC	Seyward	Marine	UBC	PRJNA387728	WGS	1
Miller	49.014	-122.772	BC	Little Campbell	Marine	UBC	PRJNA387728	WGS	1
Miller	49.612	-124.060	BC	Oyster Lagoon	Marine	UBC	PRJNA387728	WGS	1
Miller	48.824	-125.136	BC	Bamfield	Marine	UBC	PRJNA387728	WGS	1