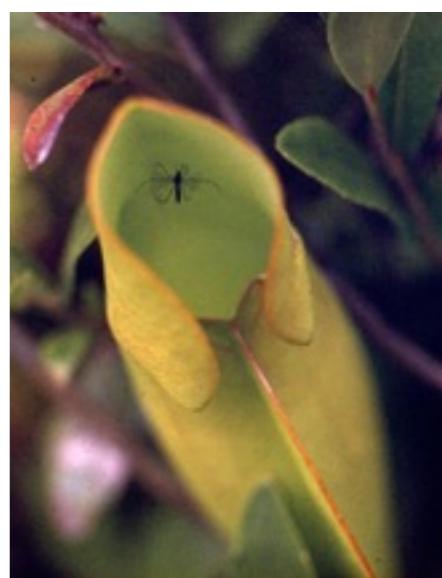
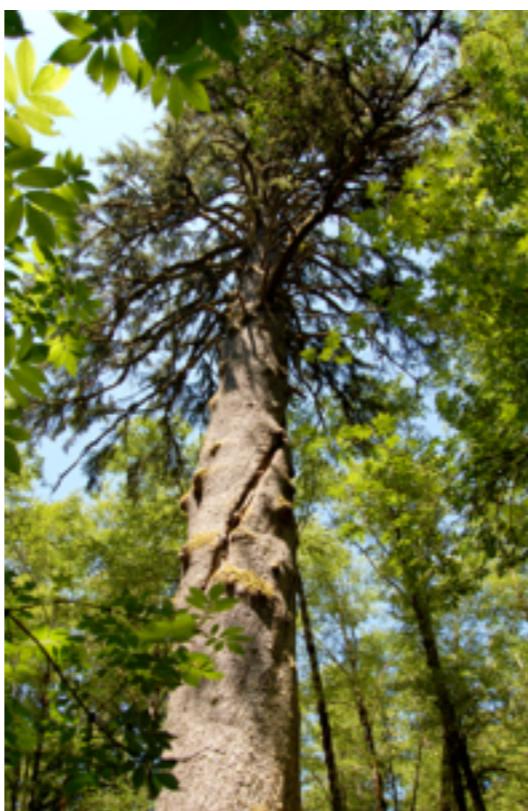
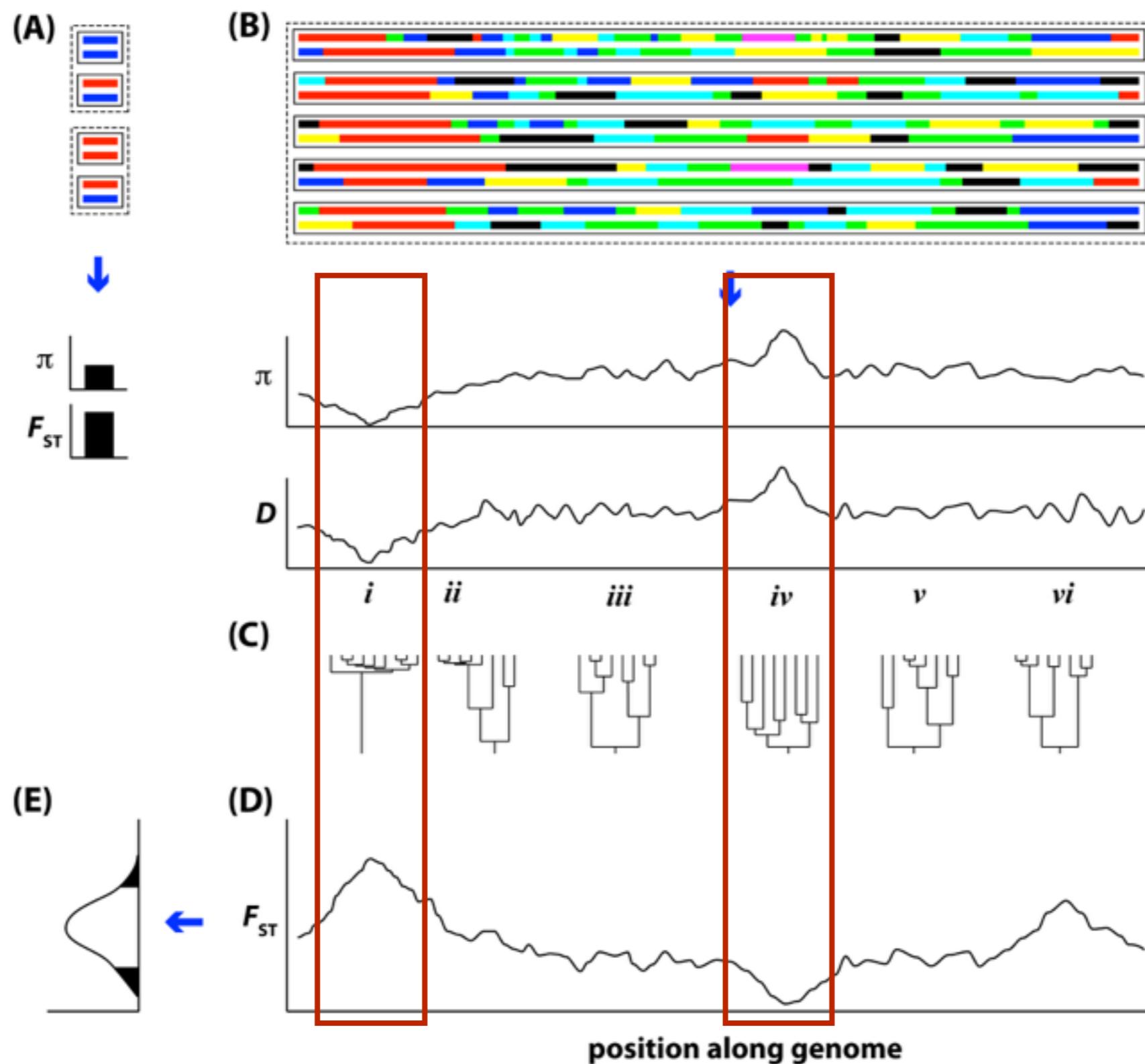


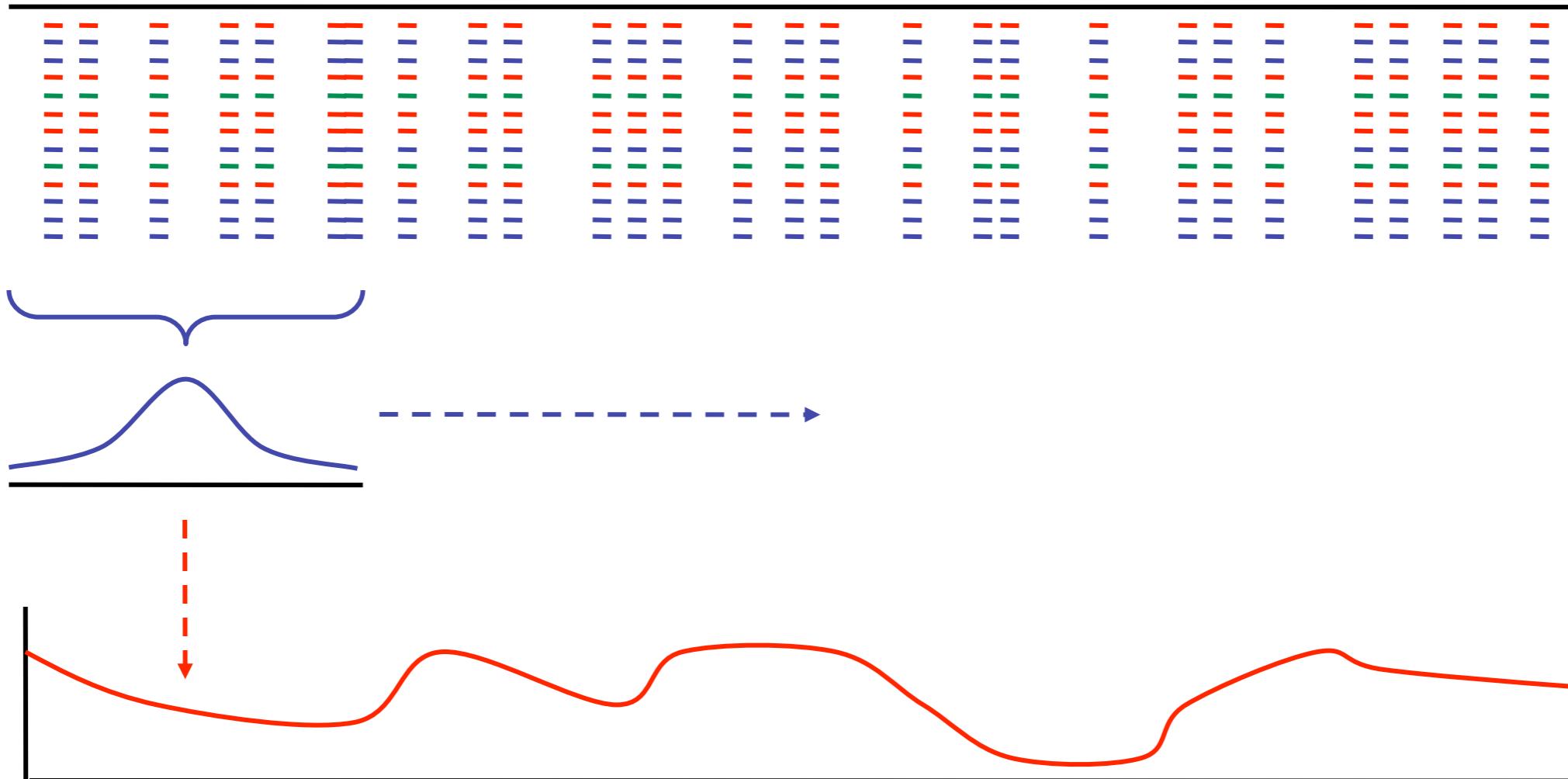
What if you have a reference genome?



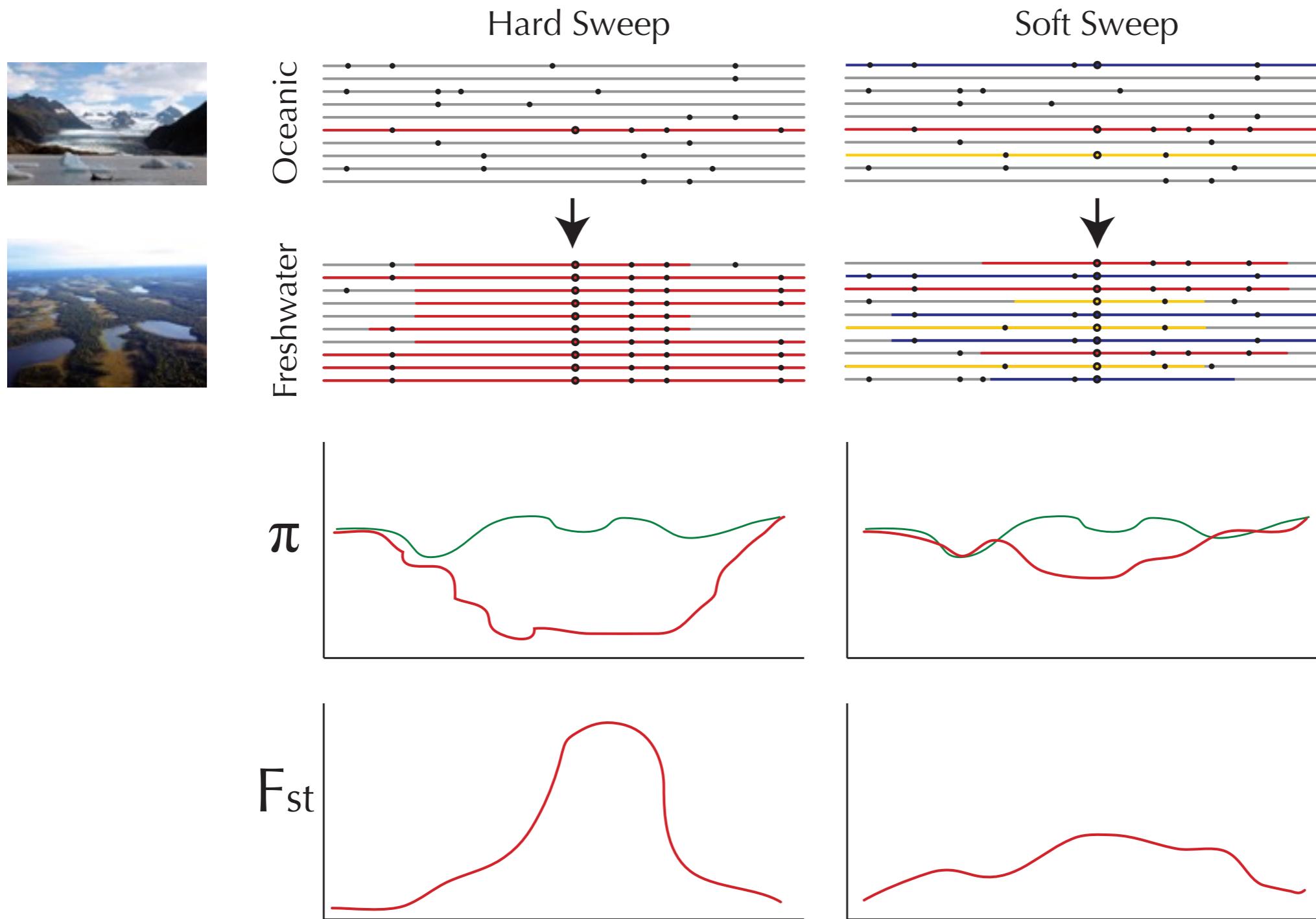
Population Genetics → Population Genomics



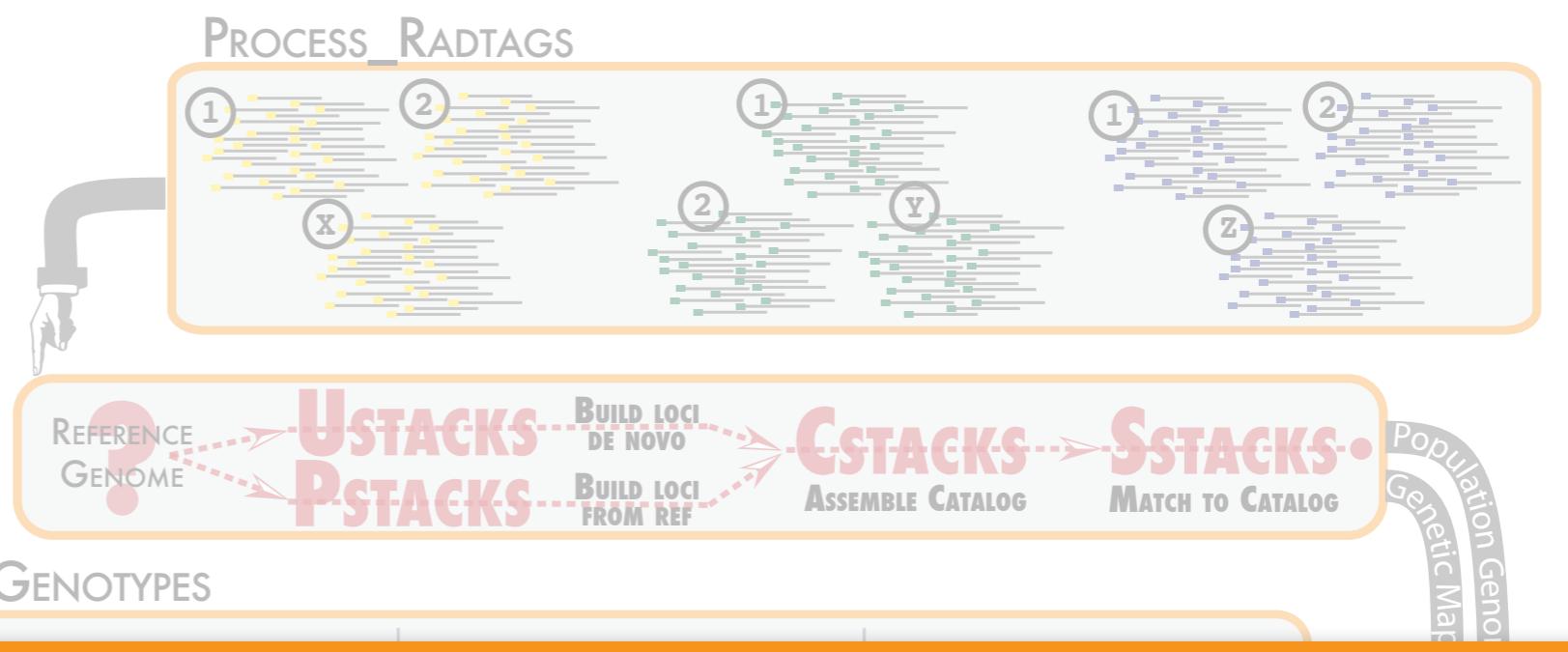
Making statistics continuous across the genome



Signatures of natural selection across the genome



Stacks

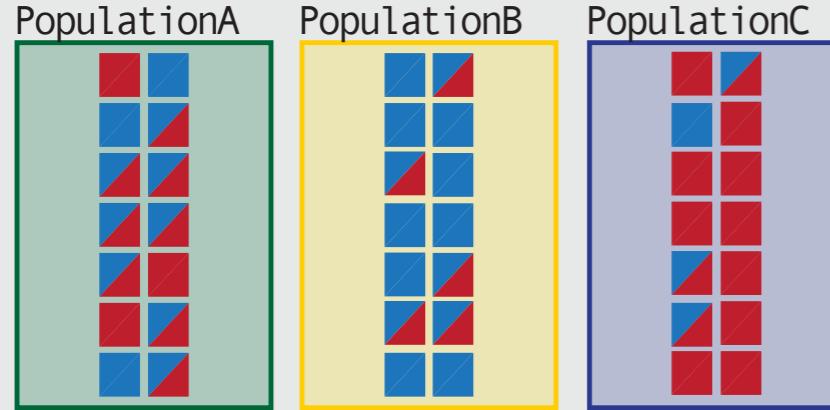


POPULATIONS

Locus 1

TGCAGGCAGGAGCTGAGCCCTGC^C_A GCTCCGACCAAACGTA

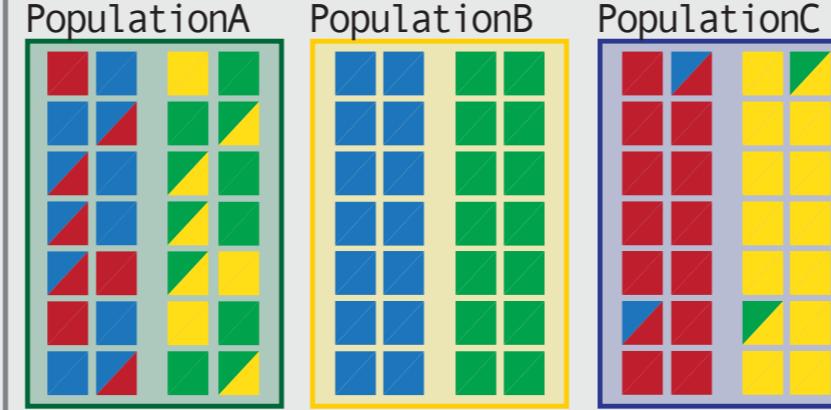
Haplotypes: **C** **A**



Locus 2

TGCAGGACACGGACTG^A_C AGCCCCTGCGGCACCA^C_T AACGTTTG

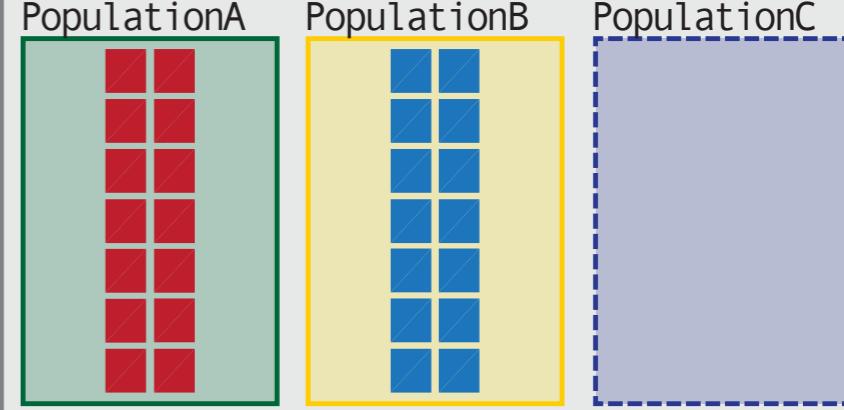
Haplotypes: **A C** **C T**



Locus N

TGCAGGACACACAGCTATTCCCTGCGGCCG^G_T CCAAACGTTTG

Haplotypes: **G** **T**



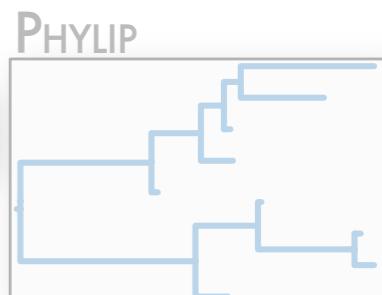
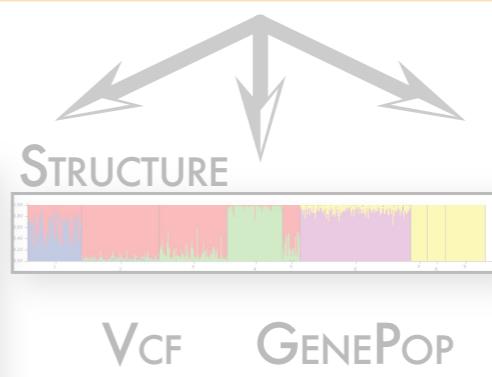
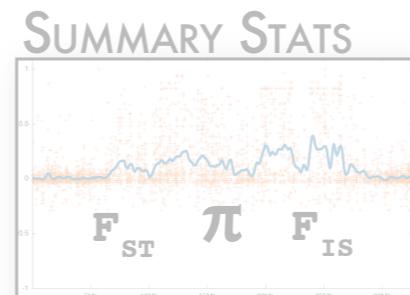
Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences

Julian M. Catchen,^{*} Angel Amores,[†] Paul Hohenlohe,^{*} William Cresko,^{*} and John H. Postlethwait,[‡]
^{*}Center for Ecology and Evolutionary Biology and [†]Institute of Neuroscience, University of Oregon, Eugene, Oregon 97403

Stacks: an analysis tool set for population genomics

JULIAN CATCHEN,^{*} PAUL A. HOHENLOHE,[†] SUSAN BASSHAM,^{*} ANGEL AMORES,[‡] and WILLIAM A. CRESKO^{*}
^{*}Institute of Ecology and Evolution, University of Oregon, Eugene, OR 97403-6289, USA, [†]Biological Sciences, University of Idaho, Moscow, ID 83844-3061, USA, [‡]Institute of Neuroscience, University of Oregon, Eugene, OR 97403-1254, USA

JOINMAP
R/QTL
ONEMAP
HAPLOTYPES



Differentiating SNPs from error

Restriction enzyme recognition site



TATAGCTGCATTTCATGTTCATTTATATTGAGTGAAGTTTCAGCCCTGCAGGTCTCCTGCTACCACCAATTCCACCTCTGCAGCTCATCCACAATGCAGCAGACCG
CTGCATTTCATGTTCATTTATATTGAGTGAAGTTTCAGCCCTGCA
CTGCATTTCATGTTCATTTATATTGAGTGAAGTATTCAAGCCCTGCA
CTGCATTTCATGTTCATTTATATTGAGTGAAGTTTCAGCCCTGCA
CTGCATTGCAAGTTCATTTATATTGAGTGAAGTTTCAGCCCTGCA
CTGCATTTCATGTTCATTTATATTGAGTGAAGTTTCAGCCCTGCA
CTGCATTTCATGTTCATTTATATTGAGTGAAGTATTCAAGCCCTGCA
CTGCATTTCATGTTCACTTCATTTATATTGAGTGAAGTTTCAGCCCTGCA
CTGCATTTCATGTTCATTTATATTGAGTGAAGTTTCAGCCCTGCA
CTGCATTTCATGTTCATTTATATTGAGTGAAGTTTCAGCCCTGCA

Reference Genome



Sequence Reads

Differentiating SNPs from error

TATAGCTGCATTCATGTT
CTGCATTCATGTT
CTGCATTCATGTT
CTGCATTCATGTT
CTGCATTGATGTT
CTGCATTCATGTT
CTGCATTCATGTT
CTGCATTCATGTT
CTGCATTCATGTT
CTGCATTCATGTT

Maximum likelihood genotyping based on multinomial distribution of nucleotide reads

$L(n_1 \text{ hom}) = P(n_1, n_2, n_3, n_4) = \frac{n!}{n_1! n_2! n_3! n_4!} \left(1 - \frac{3\epsilon}{4}\right)^{n_1} \left(\frac{\epsilon}{4}\right)^{n_2} \left(\frac{\epsilon}{4}\right)^{n_3} \left(\frac{\epsilon}{4}\right)^{n_4}$

$L(n_1 n_2 \text{ het}) = P(n_1, n_2, n_3, n_4) = \frac{n!}{n_1! n_2! n_3! n_4!} \left(0.5 - \frac{\epsilon}{4}\right)^{n_1} \left(0.5 - \frac{\epsilon}{4}\right)^{n_2} \left(\frac{\epsilon}{4}\right)^{n_3} \left(\frac{\epsilon}{4}\right)^{n_4}$

Maximum likelihood genotyping based on multinomial distribution of nucleotide reads

genome.uoregon.edu

1 (1 tags) tags per page 50

Id	SNP	Consensus	Matches	Ratio
v 76 annotate	Yes [2nuc]	TGCAGGAATCCAGTTGAAAGTCCTCACACGAGGCAAATATCGCCTGAGGGTCCGGAGA T CCAGTTGAGCACCACGCC ACTGAGACAT CG AA Chr: groupIII, 10.79Mb, +, LnL: -6.8567	75	aa: 42 (56.0%) ab: 4 (5.3%) ac: 4 (5.3%) bb: 25 (33.3%)

View: Haplotypes Allele Depths LnLs

SNPs

Column	Alleles
59	T / C / - / -
90	T / C / - / -

SUMMARY STATS

F_{ST} STATS

Population cs

cs 1335.01	cs 1335.02	cs 1335.03	cs 1335.04	cs 1335.05	cs 1335.06	cs 1335.07	cs 1335.08	cs 1335.09	cs 1335.10
CT / CC	CC	CC	CC	CC	CC	TC / CC	CC	CC	CT / CC
cs 1335.11	cs 1335.12	cs 1335.13	cs 1335.14	cs 1335.15	cs 1335.16	cs 1335.17	cs 1335.19	cs 1335.20	cs 1335.22
CC / TC	CC	CC	CC / TC	CC	CT / CC	CC	CC	CC	CC
cs 1335.23	cs 1335.24	cs 1335.25	cs 1335.26	cs 1335.27					
CC	CC	CC	CC / TC	CC					

Population pcr

pcr 1312.01	pcr 1312.02	pcr 1312.03	pcr 1312.04	pcr 1312.05	pcr 1312.06	pcr 1312.07	pcr 1312.08	pcr 1312.09	pcr 1312.10
CC	CC	CC	CC	CC	CC	CC	CC	CC	CC
pcr 1312.11	pcr 1312.12	pcr 1312.13	pcr 1312.14	pcr 1312.15	pcr 1210.05	pcr 1211.01	pcr 1211.02	pcr 1211.03	pcr 1211.04
CC	CC	CC	CC	CC	CC / CT	CC	CC	CC	CC
pcr 1211.05	pcr 1211.06	pcr 1211.07	pcr 1211.08	pcr 1211.09					
CC	CC	CC	CC	CC					

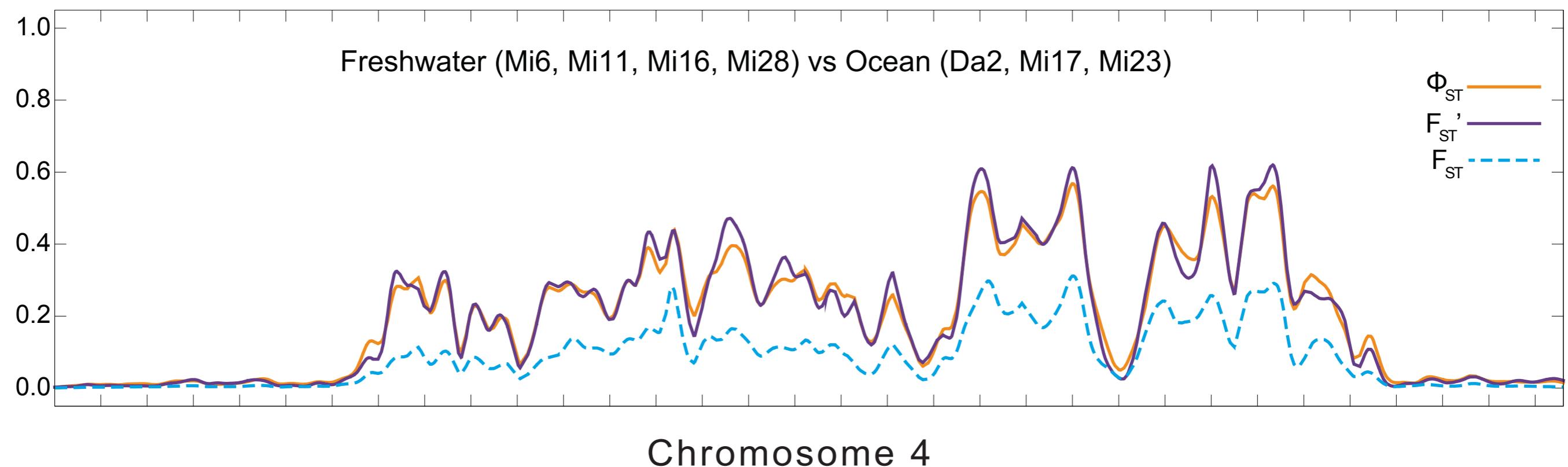
Population stl

stl 1274.31	stl 1274.32	stl 1274.33	stl 1274.34	stl 1274.35	stl 1274.36	stl 1274.37	stl 1274.38	stl 1274.39	stl 1274.40
TC									
stl 1274.41	stl 1274.42	stl 1274.43	stl 1274.44	stl 1274.45	stl 1274.46	stl 1274.47	stl 1274.48	stl 1274.49	stl 1274.50
TC									
stl 1274.51	stl 1274.52	stl 1274.53	stl 1274.54	stl 1274.55					
TC	TC	TC	TC	TC					

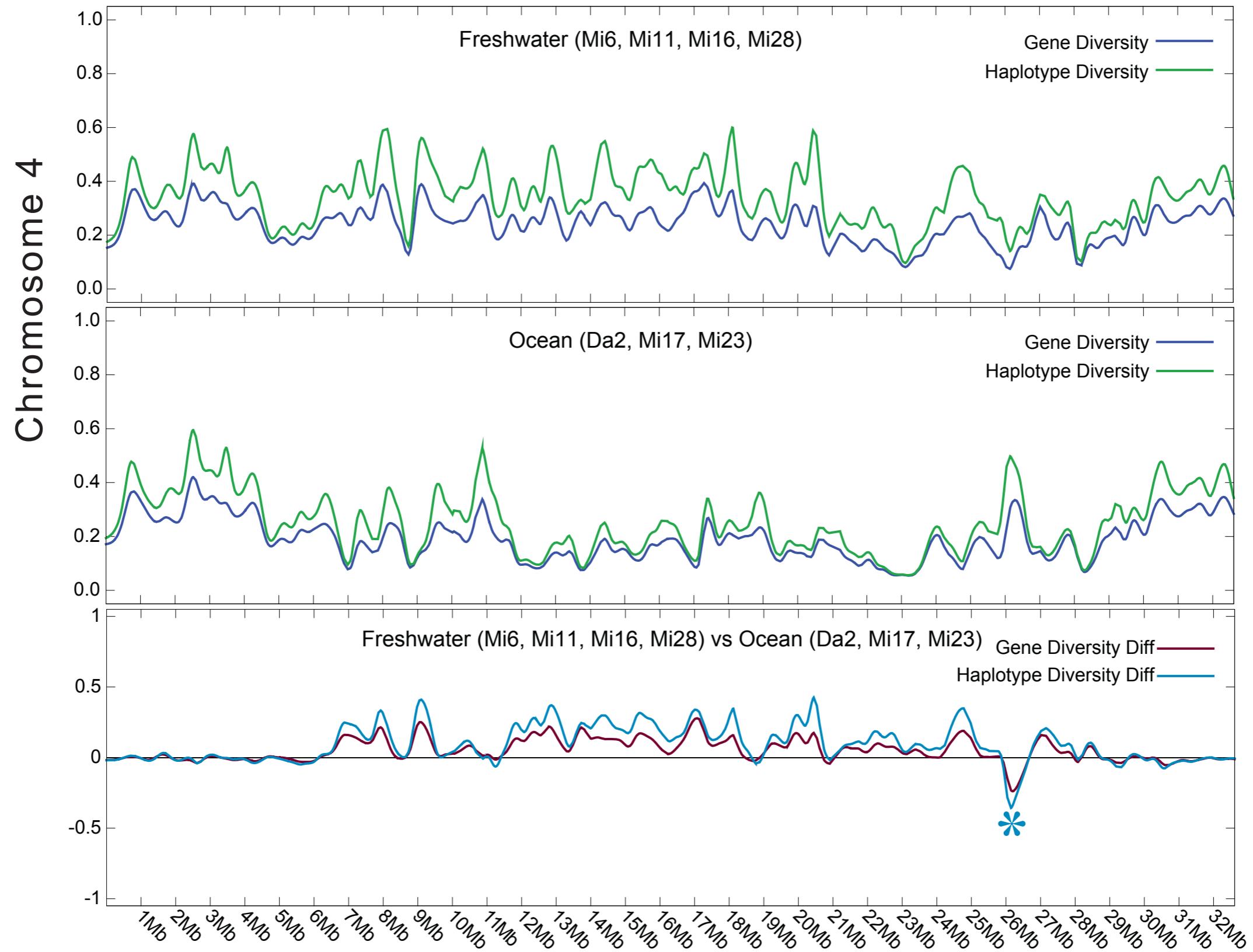
1 (1 tags) tags per page 50

haplotype stats versus snp stats

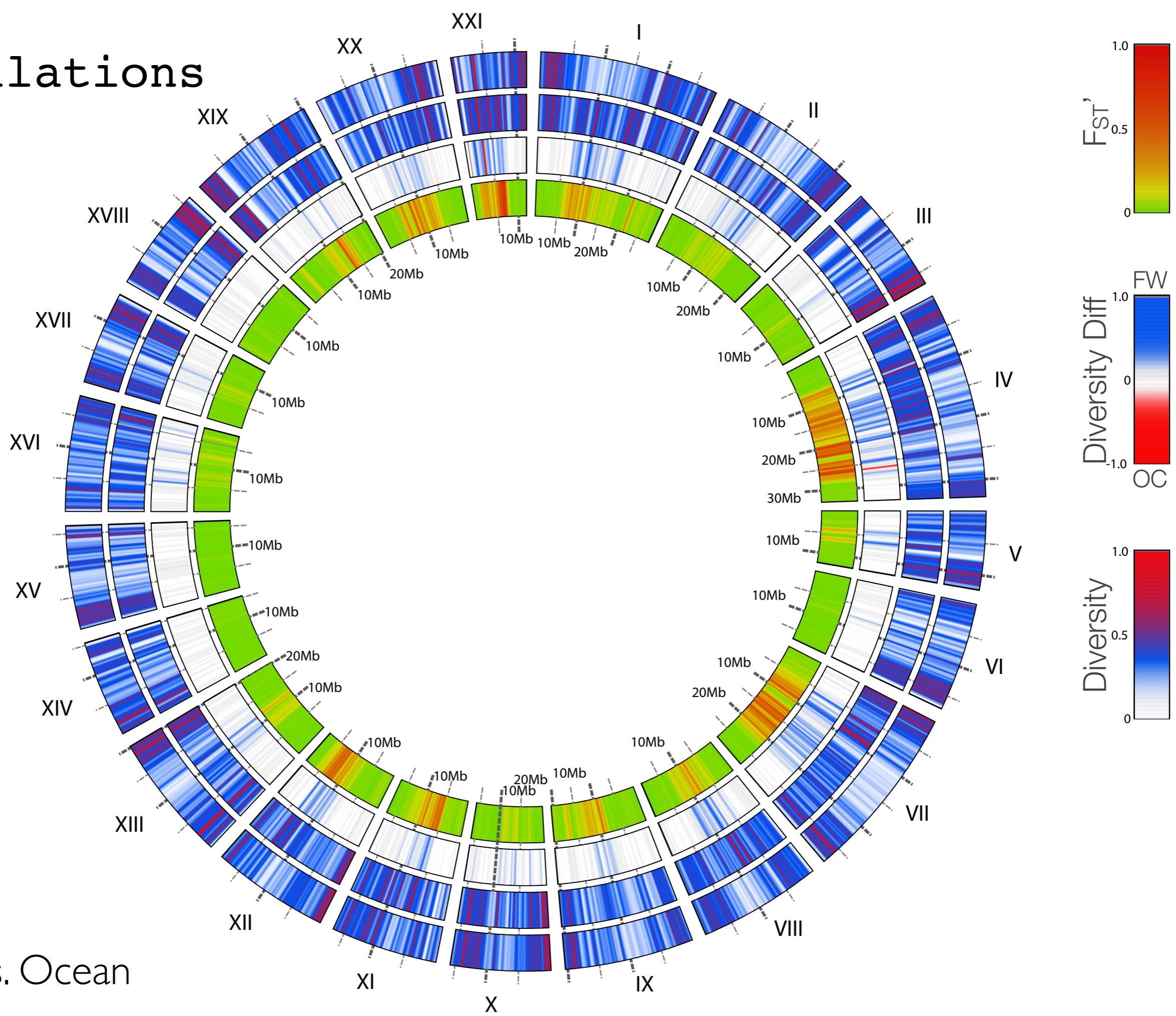
populations



populations

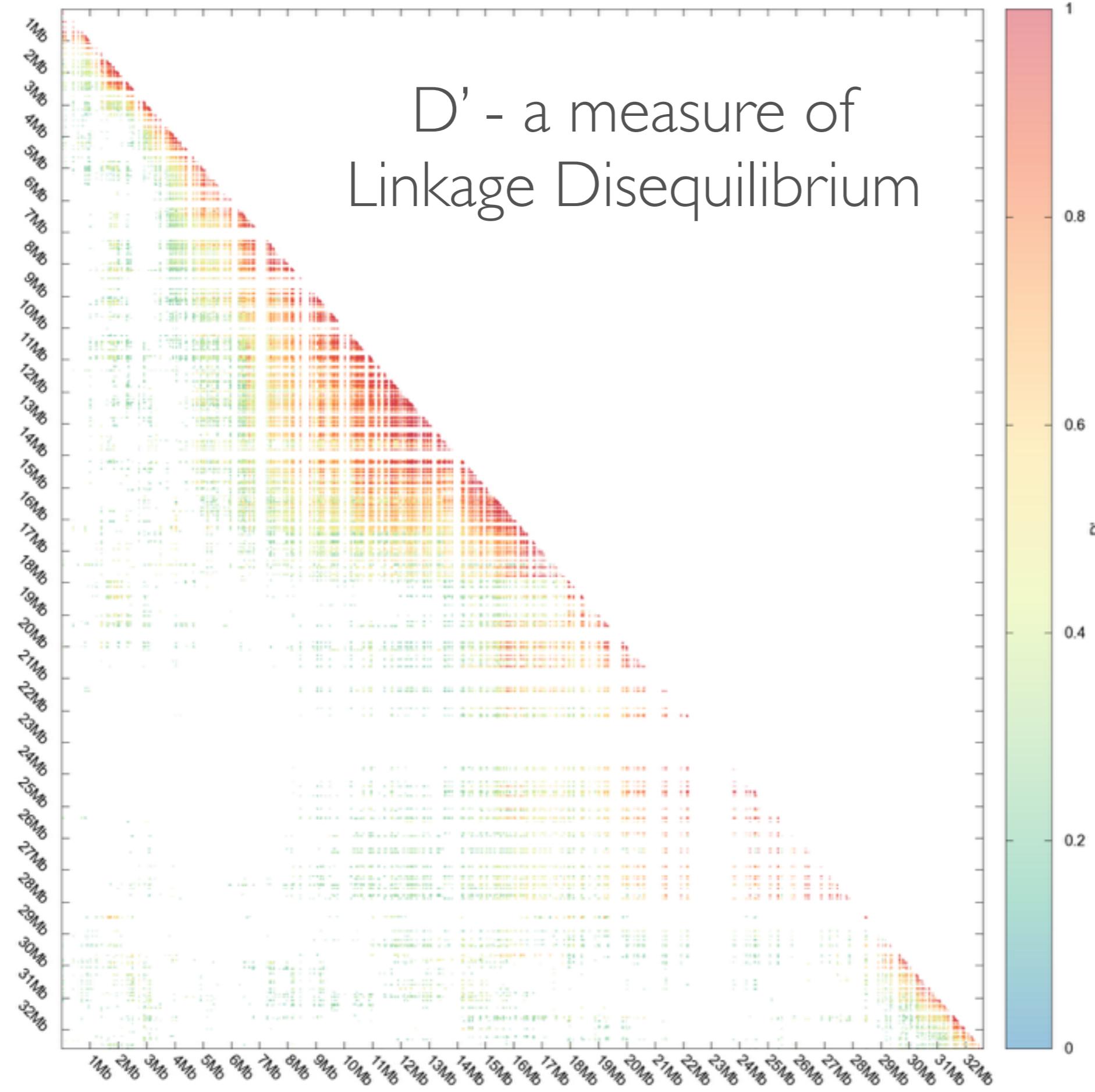


populations



FW2 vs. Ocean

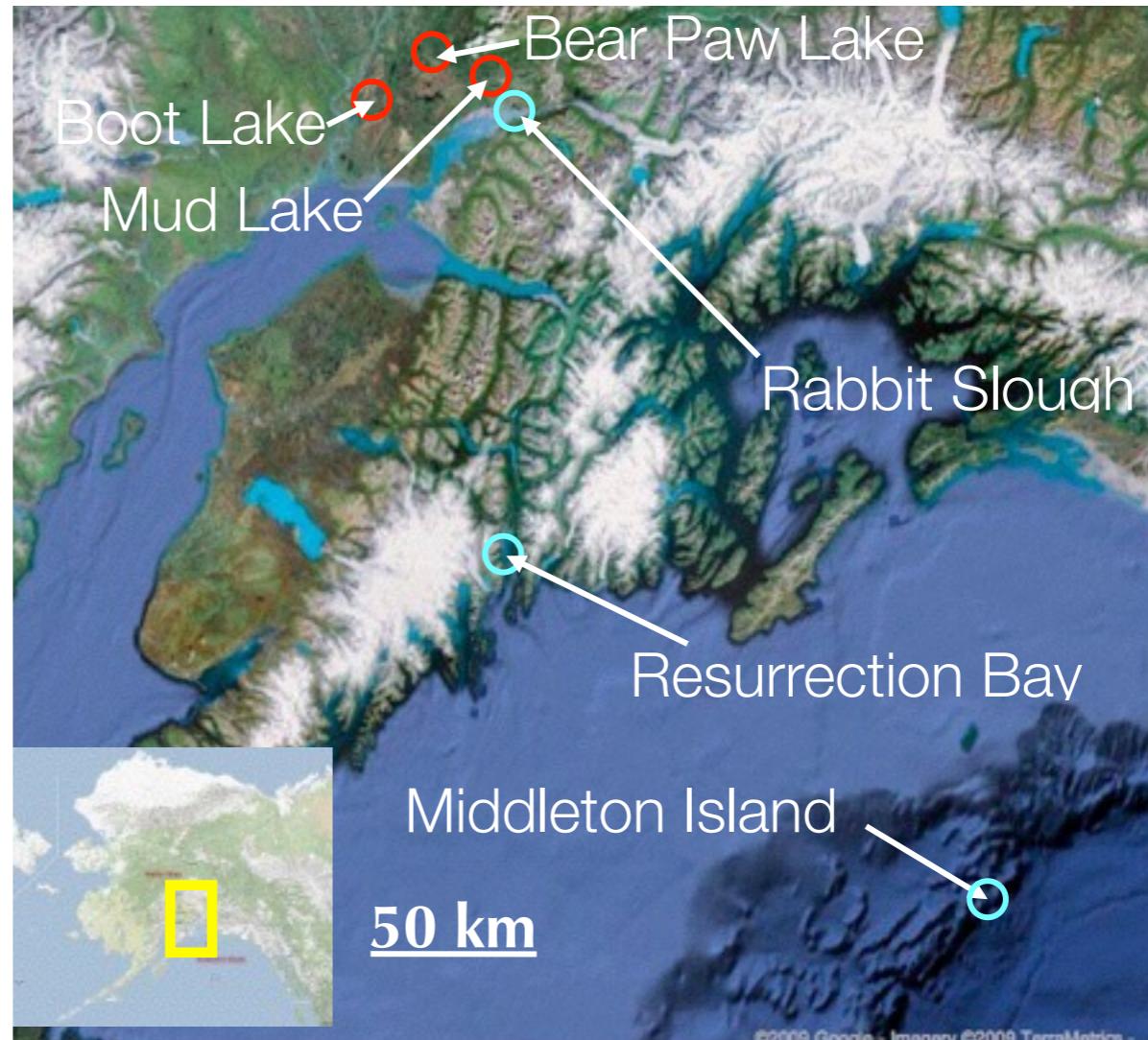
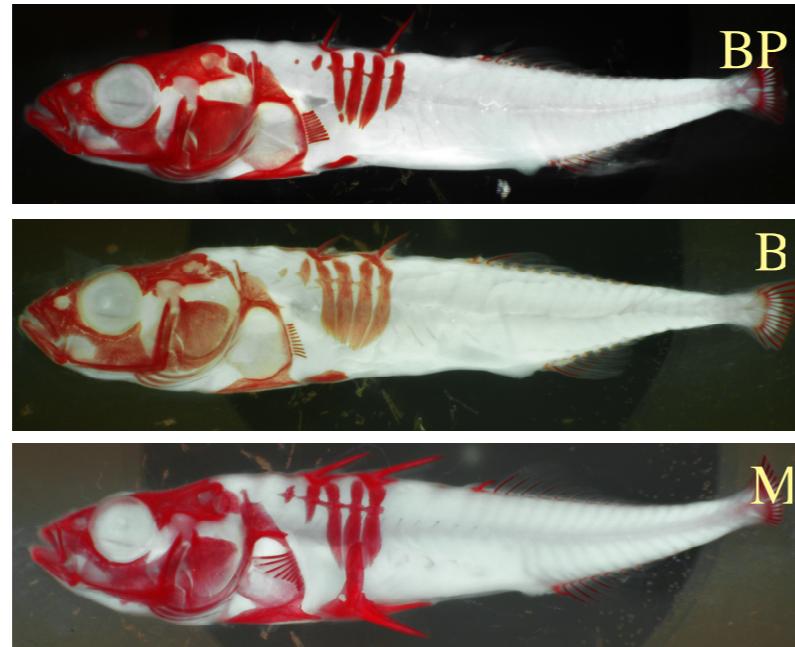
Beagle Mi12 groupIV



D' - a measure of
Linkage Disequilibrium

populations ->
Beagle ->
phasedstacks

Signatures of natural selection across the genome



- 20 individuals in each of 5 populations
- 2 Ocean & 3 Freshwater
- RAD-seq



