Clustered cis-regulatory elements underlie adaptive divergence in sticklebacks

Felicity Jones

Friedrich Miescher Laboratory of Max Planck Society

Tübingen, Germany

Male threespine stickleback

DFG
Deutsche Forschungsgemeinschaft

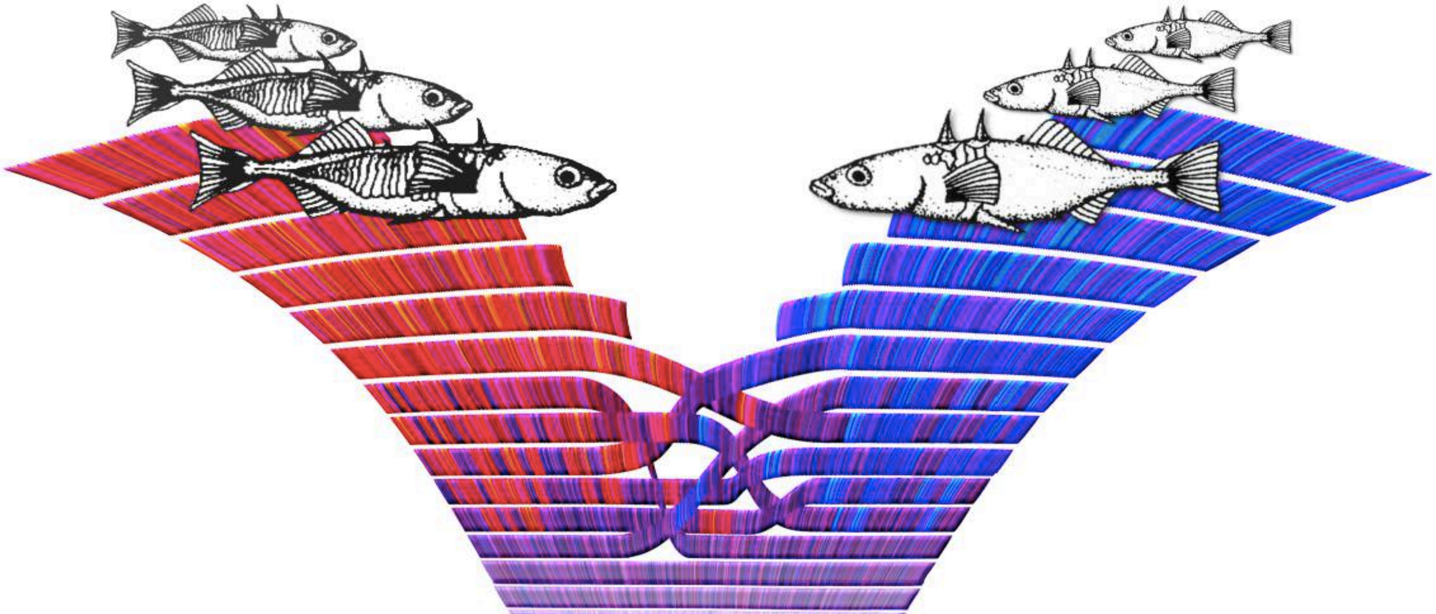
erc
European Research Council
Consolidator Grant

MAX-PLANCK-GESELLSCHAFT

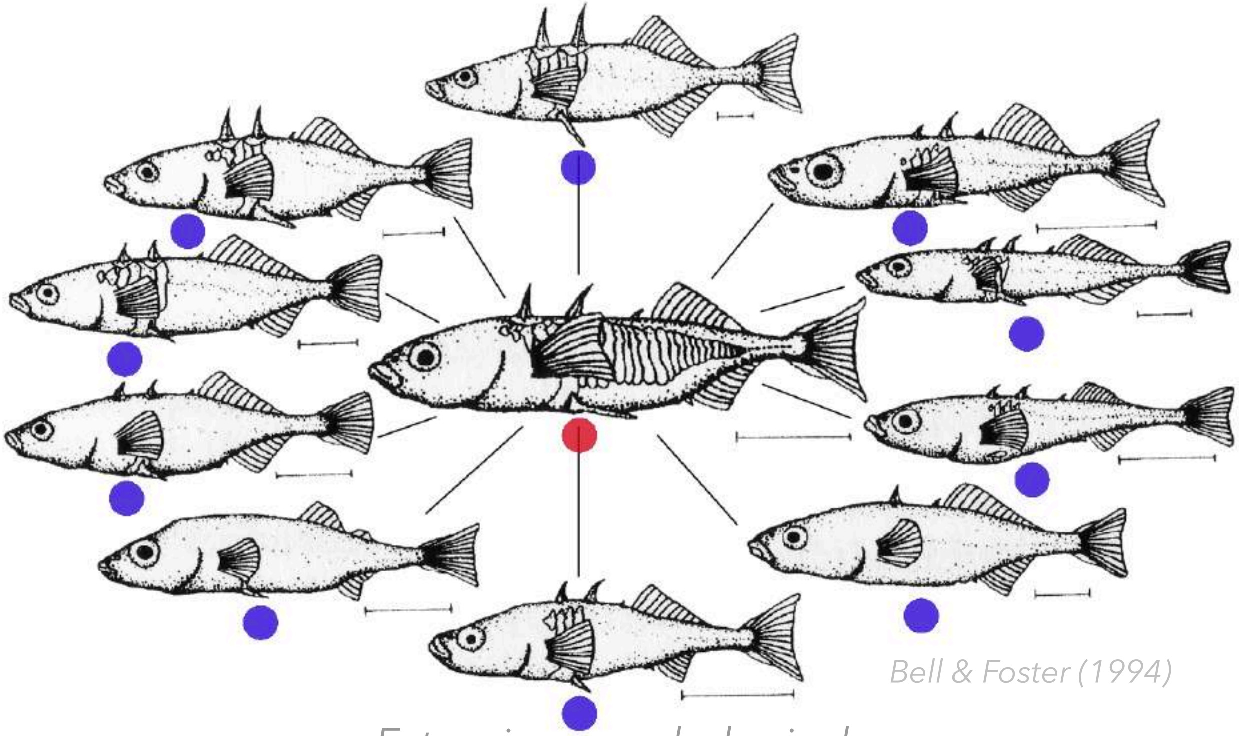
fml

Molecular mechanisms of adaptation & speciation

1. The regulatory control of adaptive divergence in gene expression
2. Functional dissection of adaptive regulatory elements
3. The evolution & role of recombination during adaptive divergence



Threespine sticklebacks have undergone a recent adaptive radiation



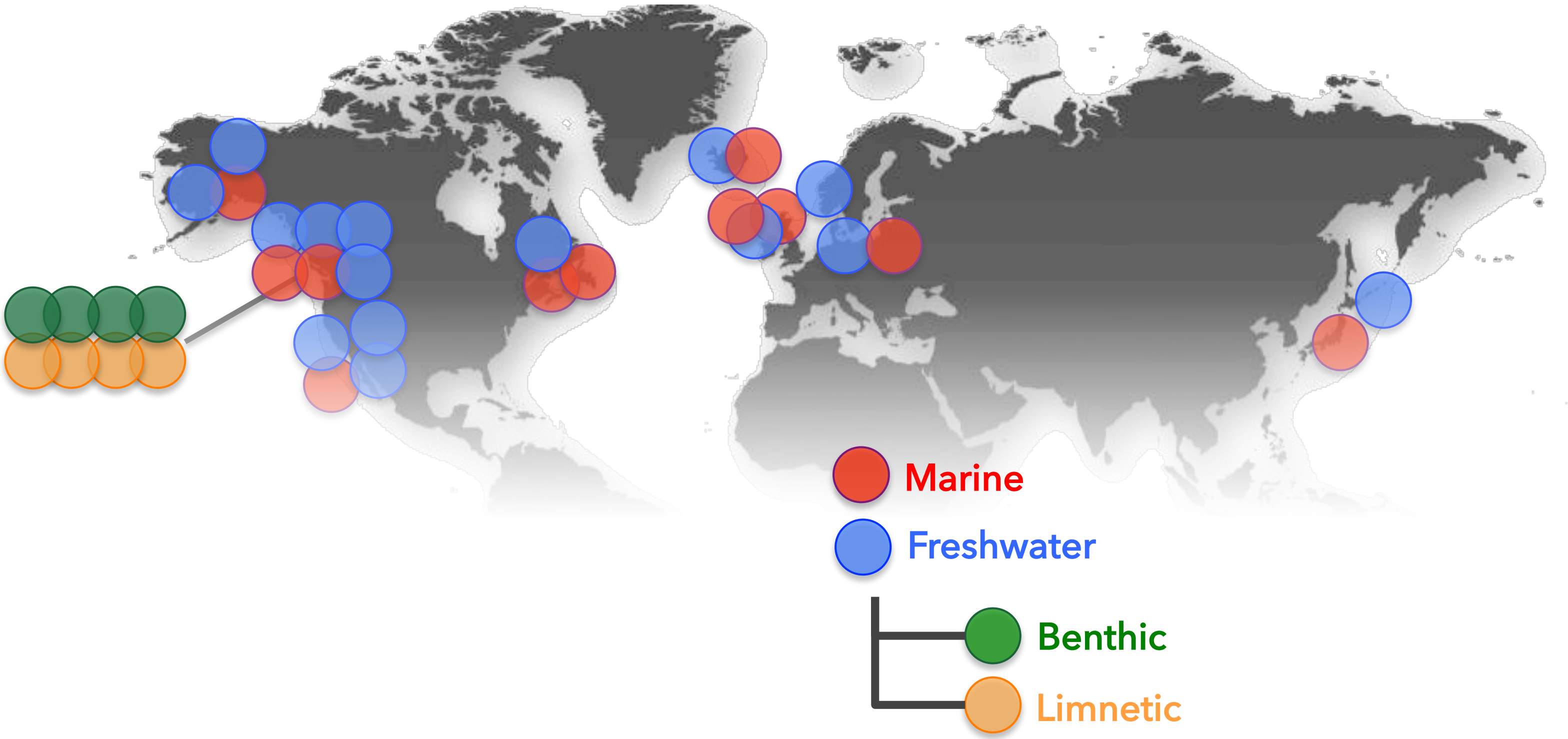
Bell & Foster (1994)

Extensive morphological, physiological, behavioural diversity

Marine

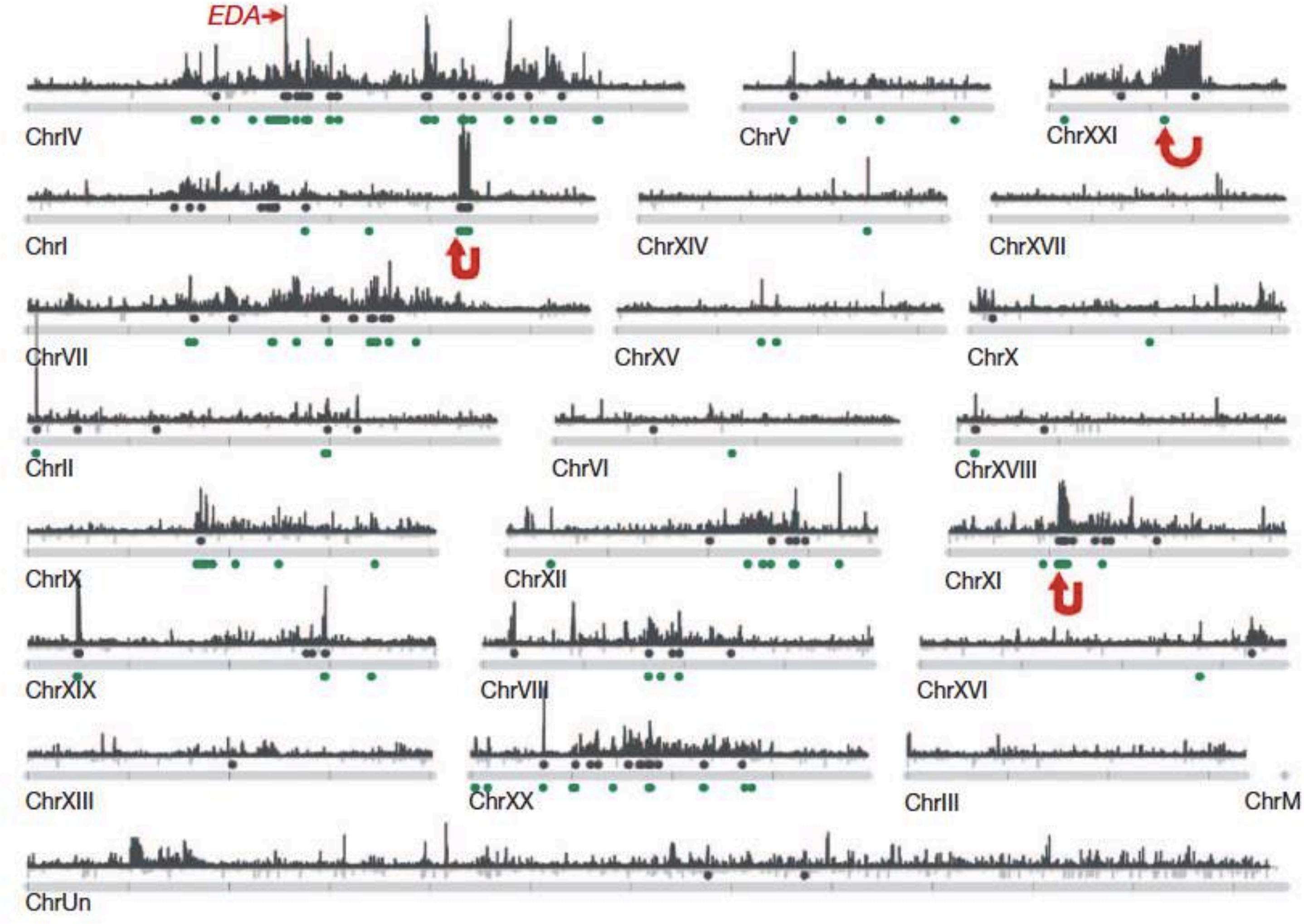
Freshwater

Repeated & independent evolution provides biological replicates of the evolutionary process



Whole genome sequencing of marine–freshwater species pairs revealed parallel adaptive divergence at ~81 predominantly intergenic genomic loci

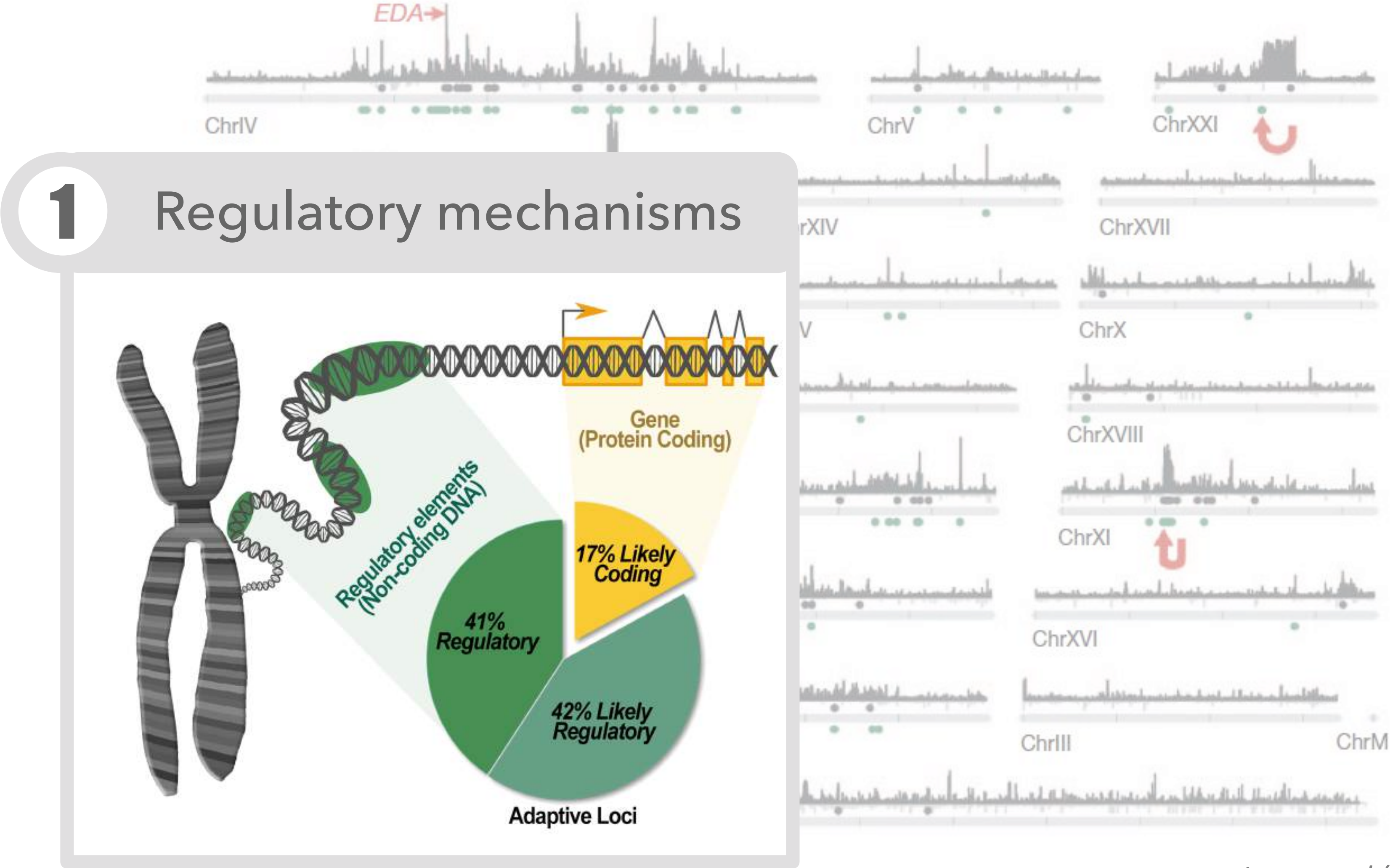
High resolution map of 81 parallel marine-freshwater adaptive loci (FDR 0.02)



Jones et al (2012) Nature

Suggests two molecular mechanisms play a role in parallel adaptive divergence of sticklebacks:

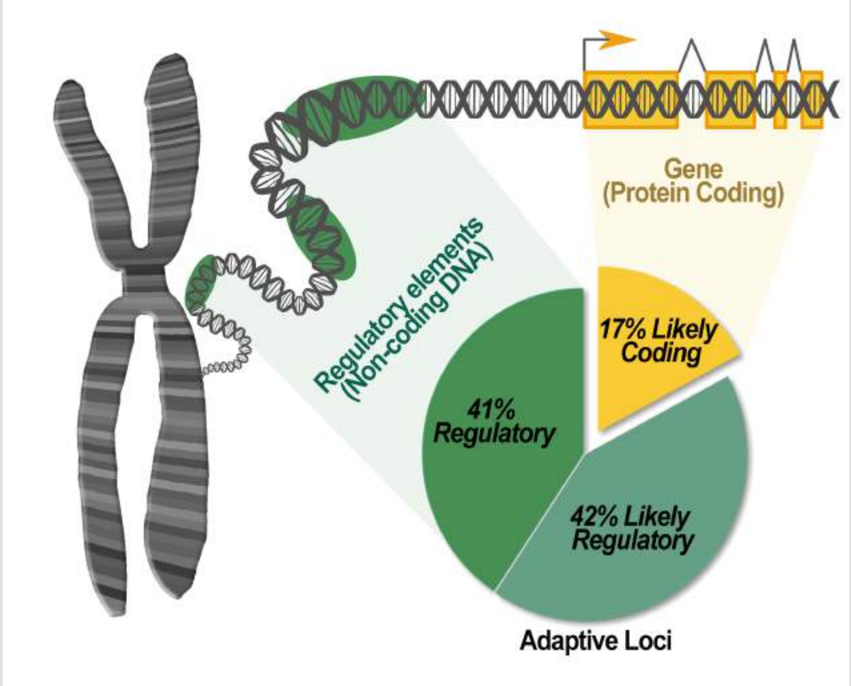
1 Regulatory mechanisms



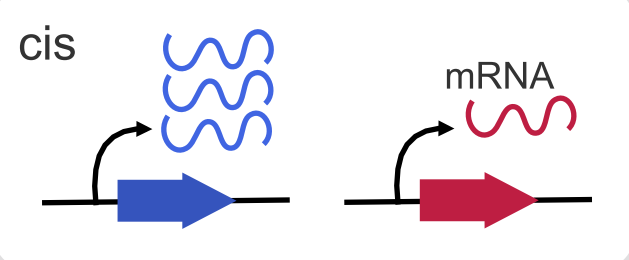
Jones et al (2012) Nature

Suggests two molecular mechanisms play a role in parallel adaptive divergence of sticklebacks:

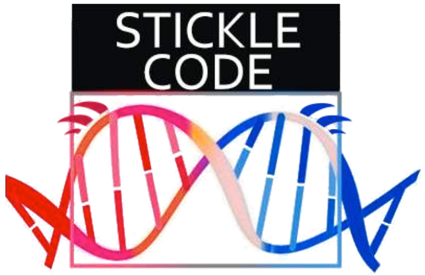
1 Regulatory mechanisms



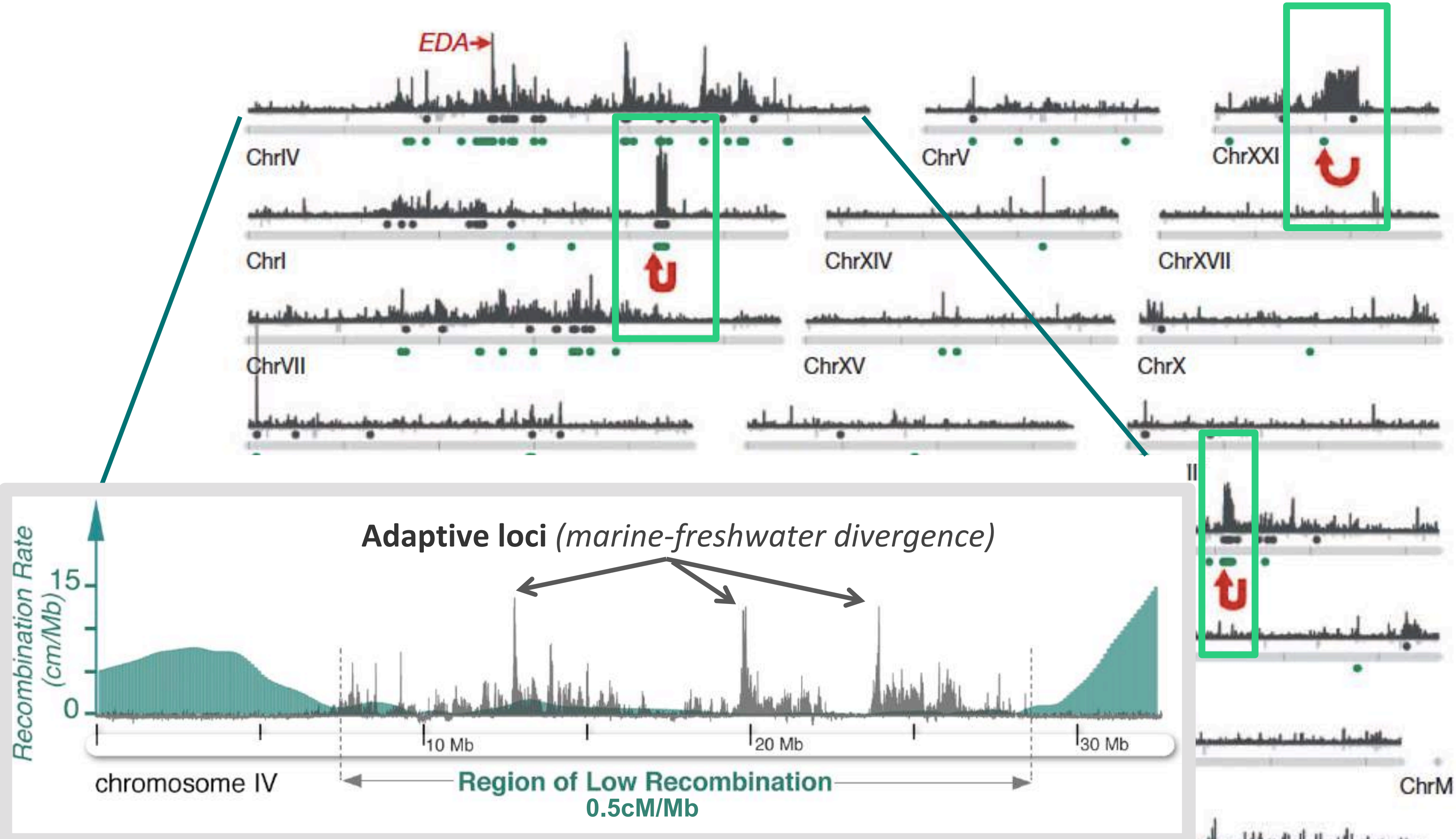
How is adaptive divergence in gene expression regulated?



What are the mutations & functional elements?



Suggests two molecular mechanisms play a role in parallel adaptive divergence of sticklebacks:



Adaptive loci *(marine-freshwater divergence)*

The genomic recombination landscape shapes the loci of adaptation

2

Jones et al (2012) Nature

How is adaptive divergence in gene expression regulated?

Identifying adaptive gene expression differences using comparative transcriptomics

Dr Jukka-Pekka Verta

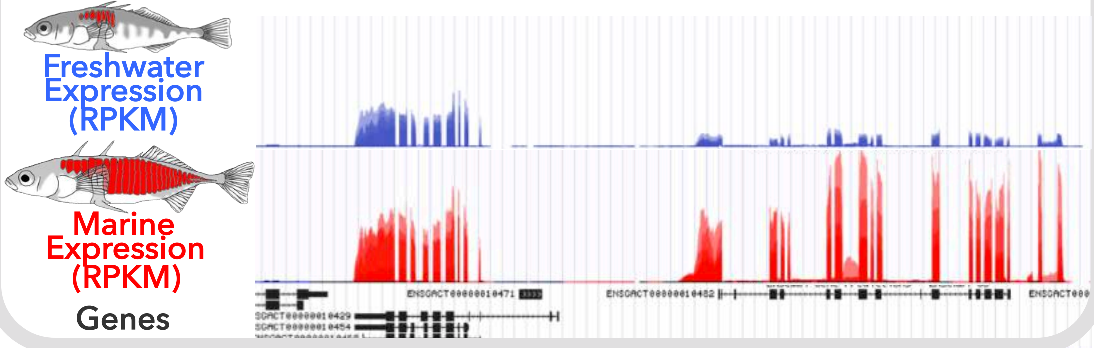
Verta & Jones (in review)
BioRxiv
doi: https://doi.org/10.1101/412932

How is adaptive divergence in gene expression regulated?

Identifying adaptive gene expression differences using comparative transcriptomics

Dr Jukka-Pekka Verta

RNAseq of Gill Tissue



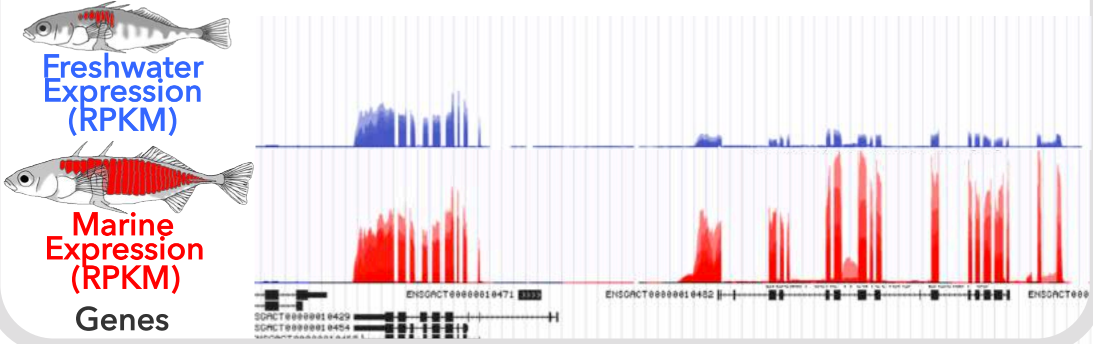
Verta & Jones (in review)
BioRxiv
doi: https://doi.org/10.1101/412932

How is adaptive divergence in gene expression regulated?

Identifying adaptive gene expression differences using comparative transcriptomics

Dr Jukka-Pekka Verta

RNAseq of Gill Tissue

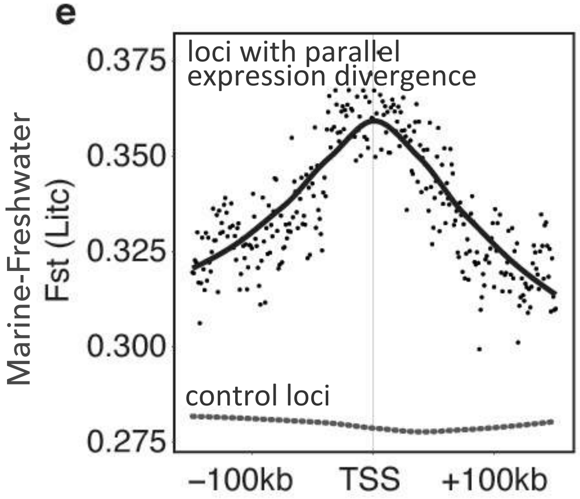


Adaptive gene expression differences
(parallel divergence in gene expression)

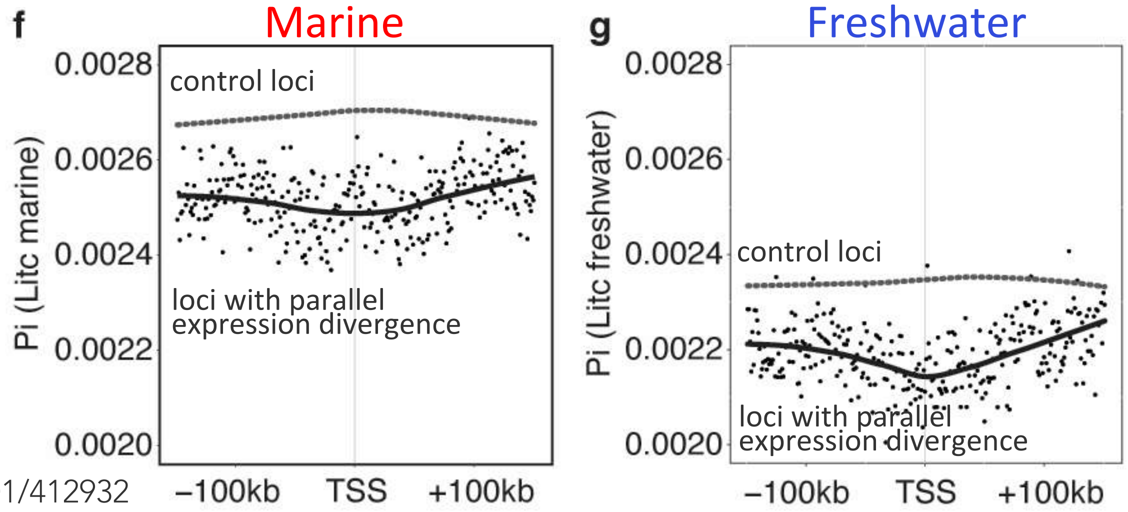
Verta & Jones (in review)
BioRxiv
doi: https://doi.org/10.1101/412932

Molecular signatures of natural selection acting on genes with parallel gene expression divergence

Elevated genetic divergence ($F_{ST\ MAR\ v\ FW}$) around TSS



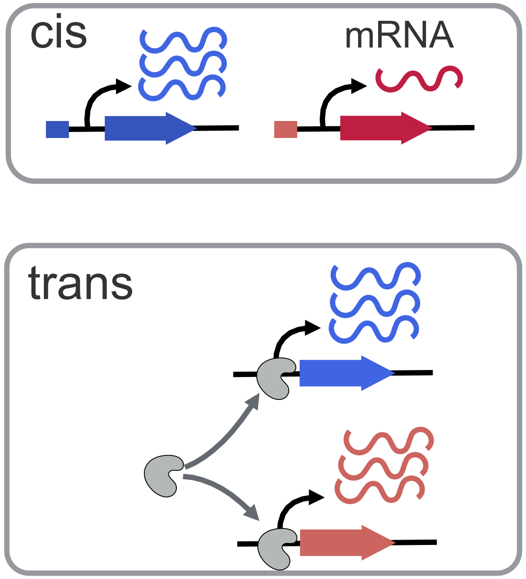
Reduced nucleotide diversity (pi) around TSS

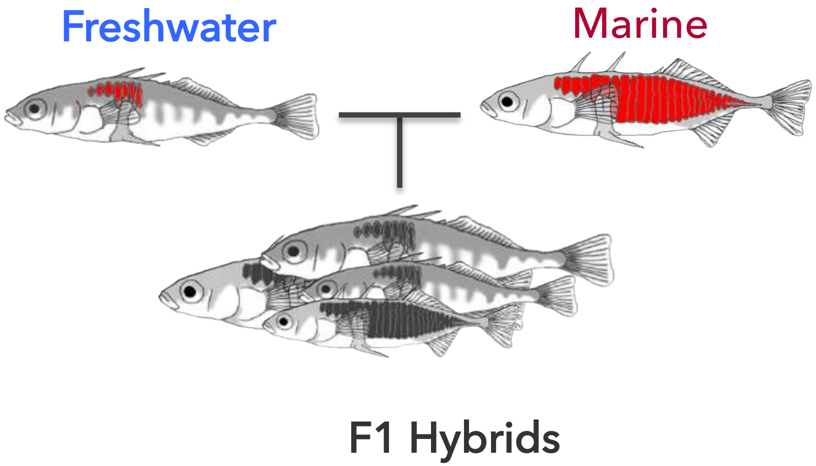


Verta & Jones (in review)
BioRxiv
doi: https://doi.org/10.1101/412932

How are adaptive expression differences controlled?

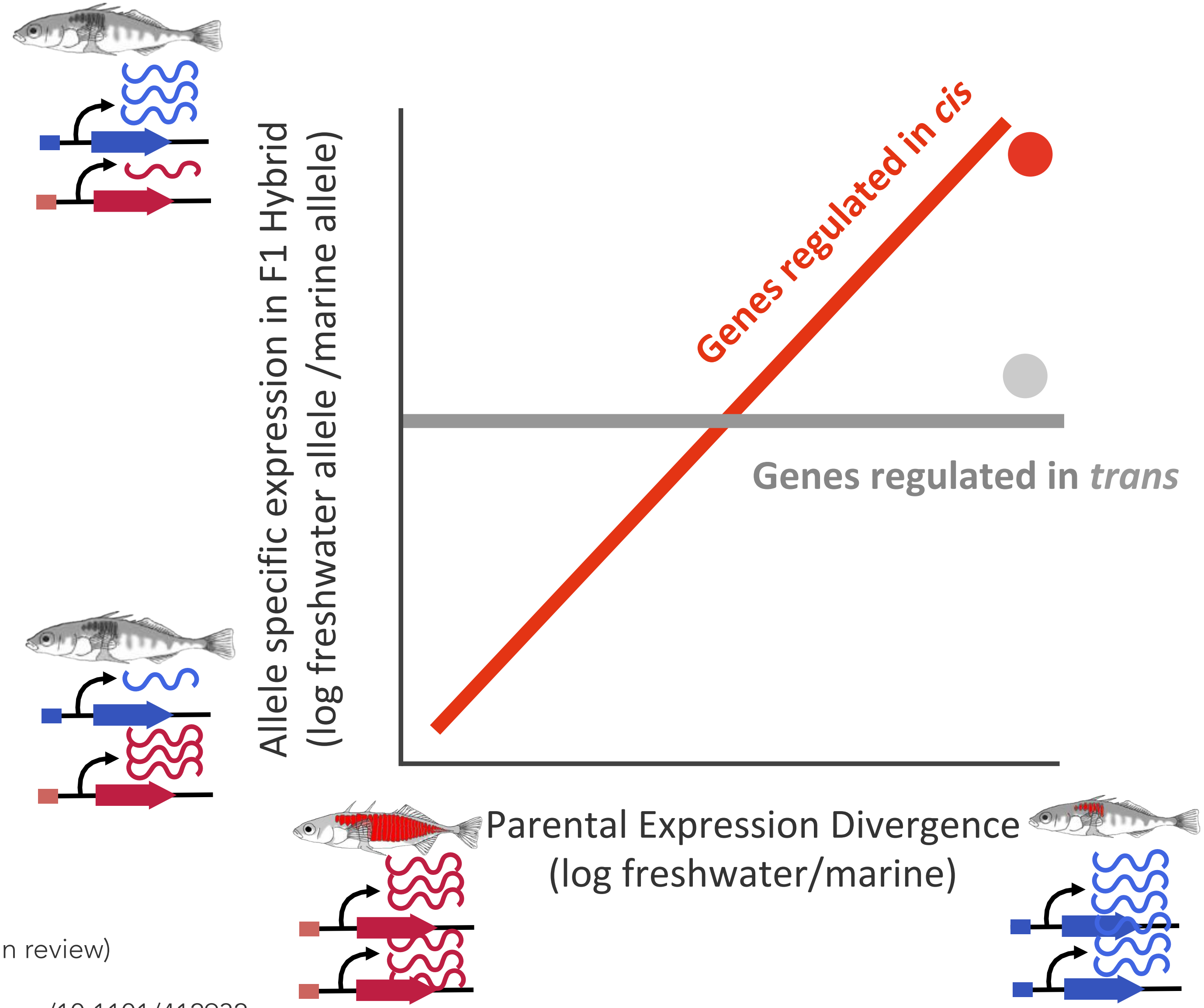
trans vs *cis*-regulation





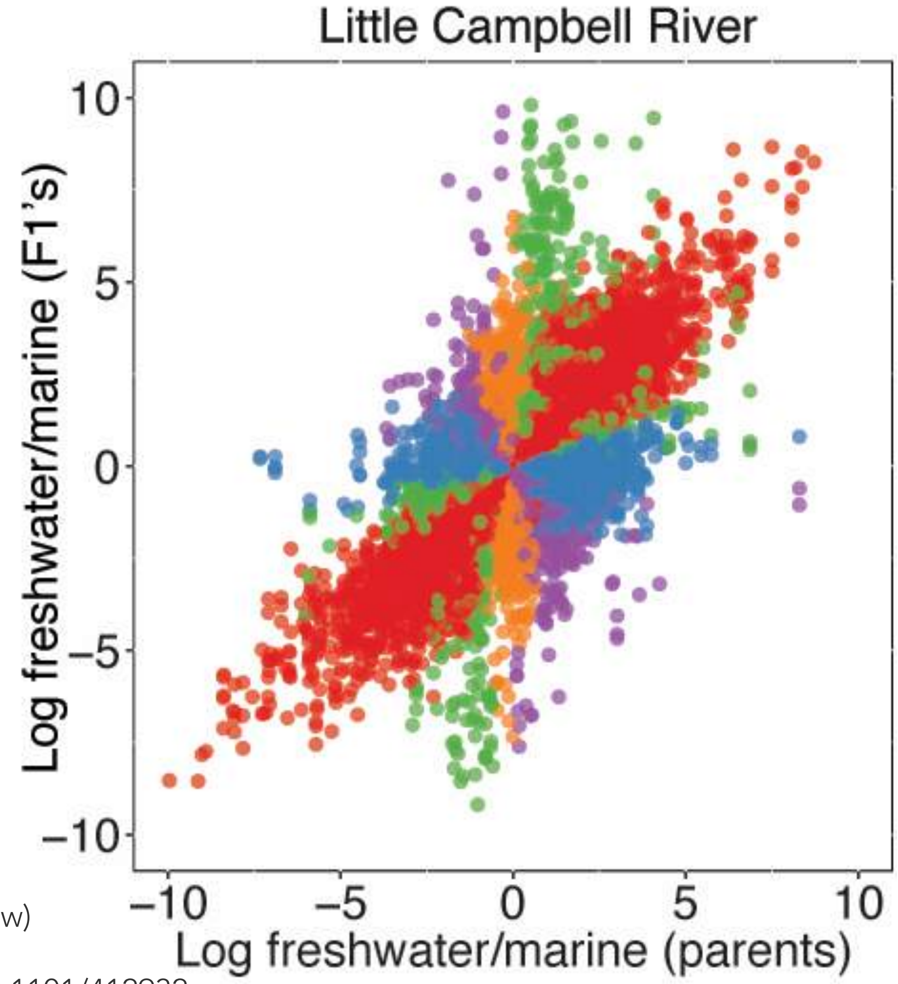
Verta & Jones (in review)
BioRxiv
doi: https://doi.org/10.1101/412932

Allele-Specific Expression assays:
RNAseq of F1 hybrids to determine *cis*- vs *trans*-regulation



Verta & Jones (in review)
BioRxiv
doi: https://doi.org/10.1101/412932

Cis-regulation is the predominant mechanism underlying gene expression divergence

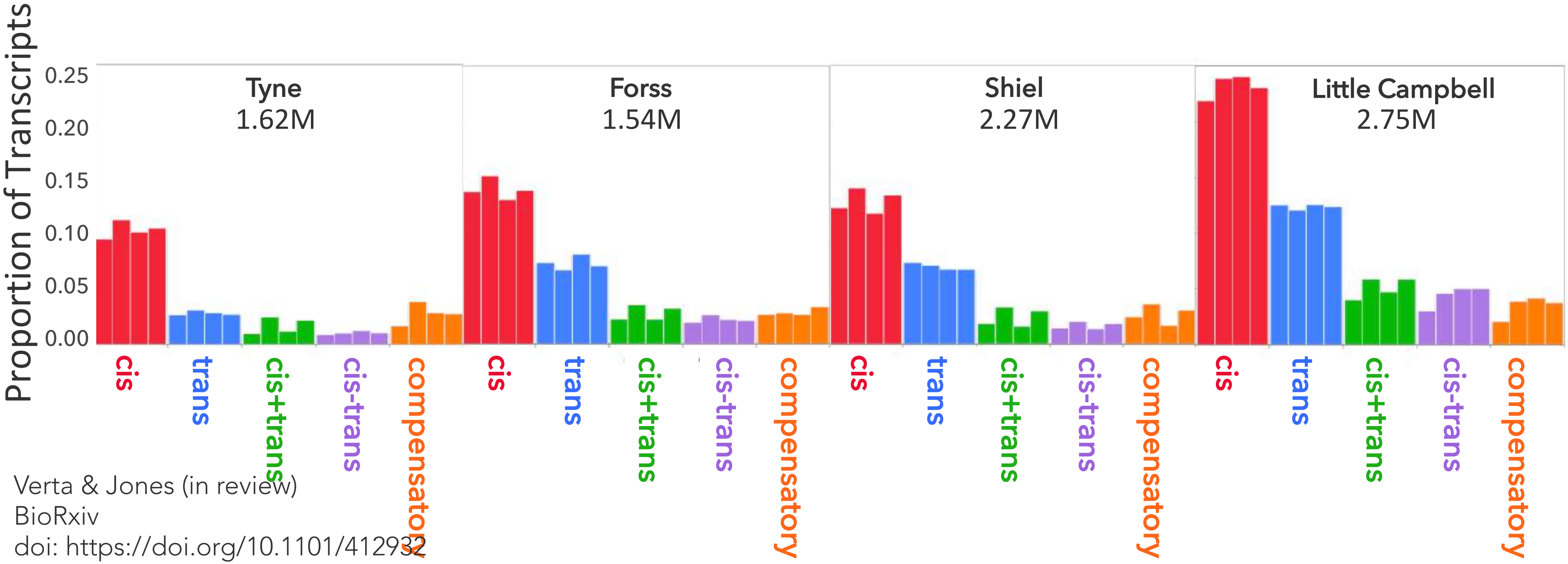


Genes regulated in *cis*

Genes regulated in *trans*

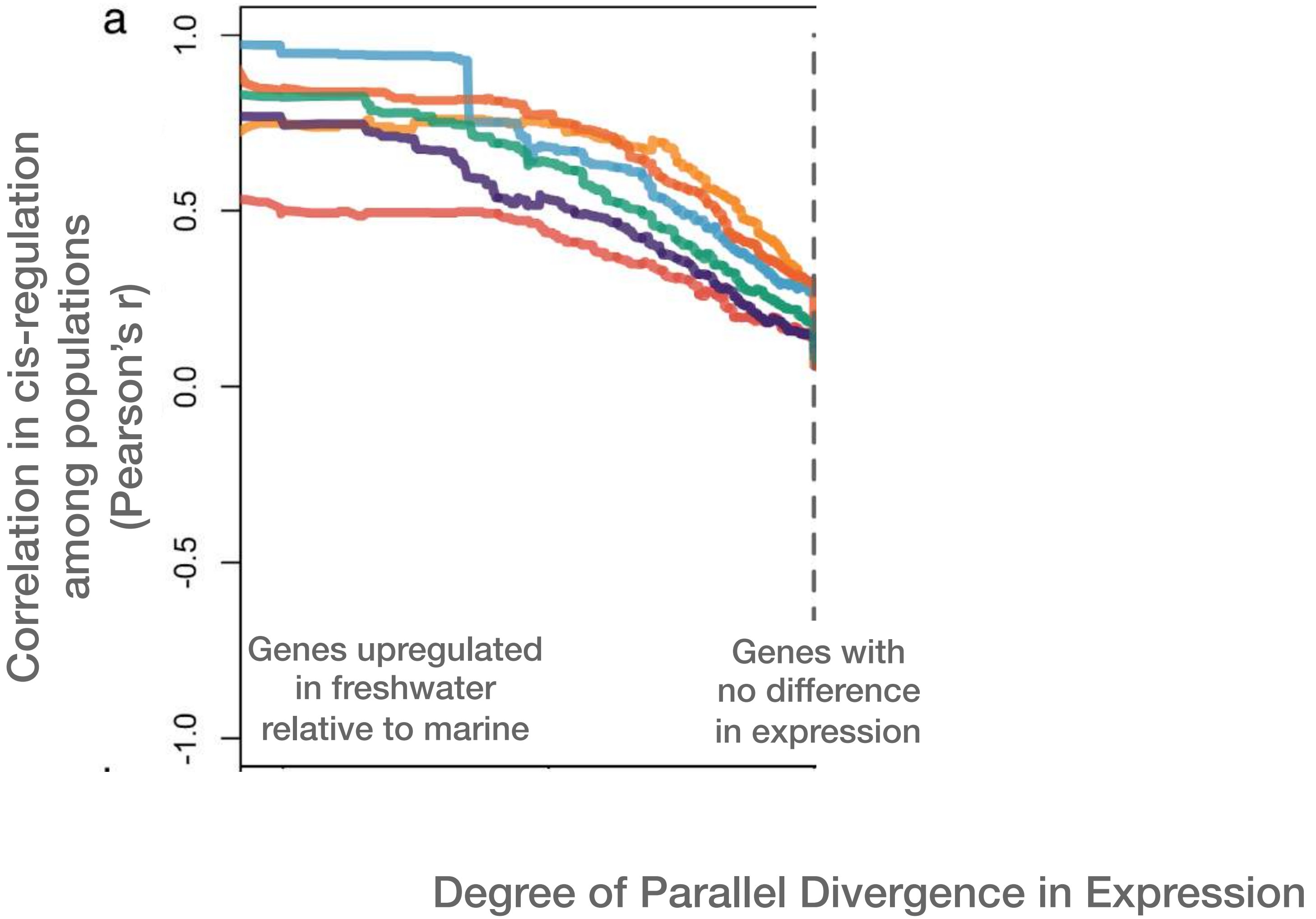
Verta & Jones (in review)
BioRxiv
doi: https://doi.org/10.1101/412932

Cis-regulation is the predominant in multiple mar-fw pairs (and is associated with the degree of genome-wide genetic divergence)



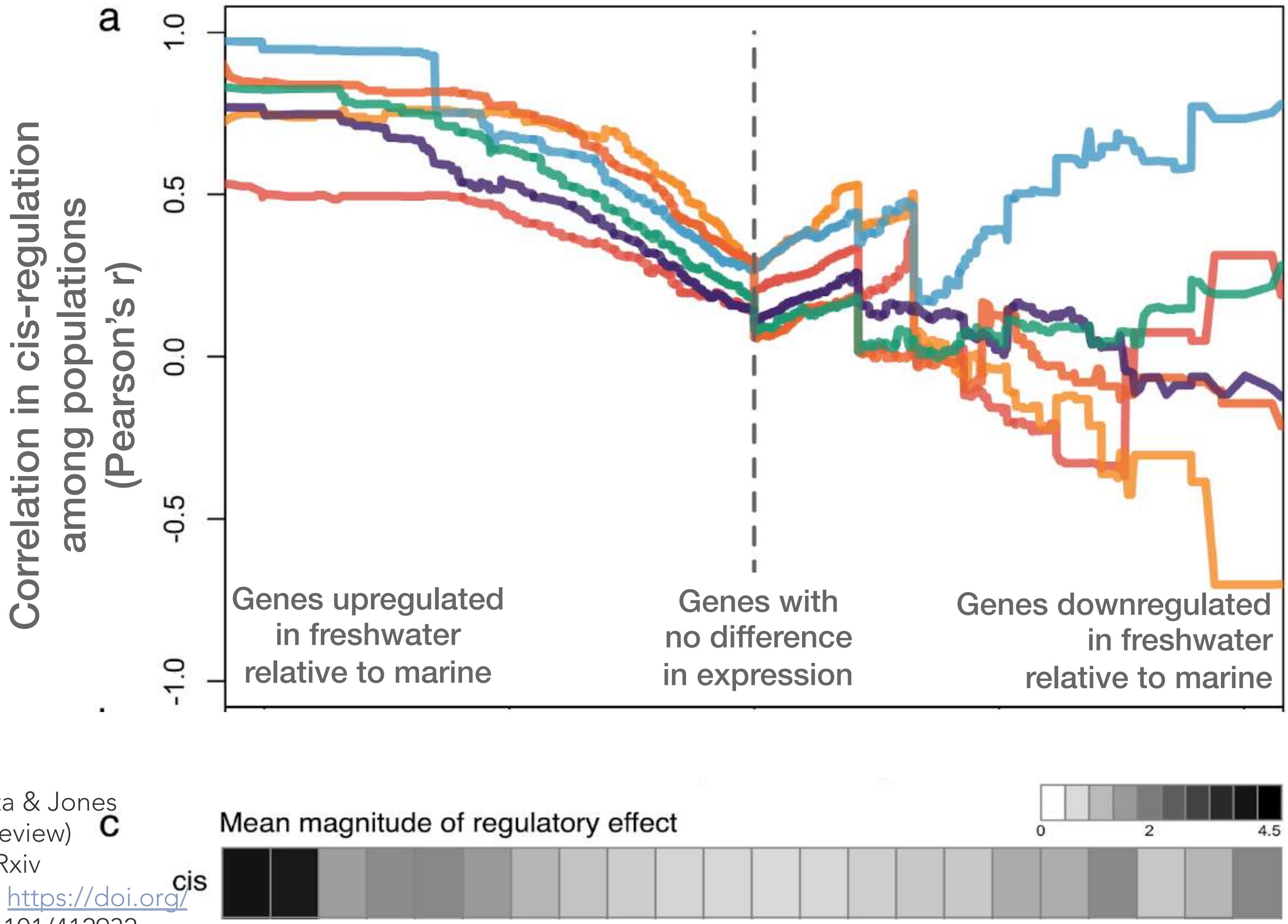
Verta & Jones (in review)
BioRxiv
doi: https://doi.org/10.1101/412932

Gene regulatory mechanisms and their magnitude evolve in parallel across populations



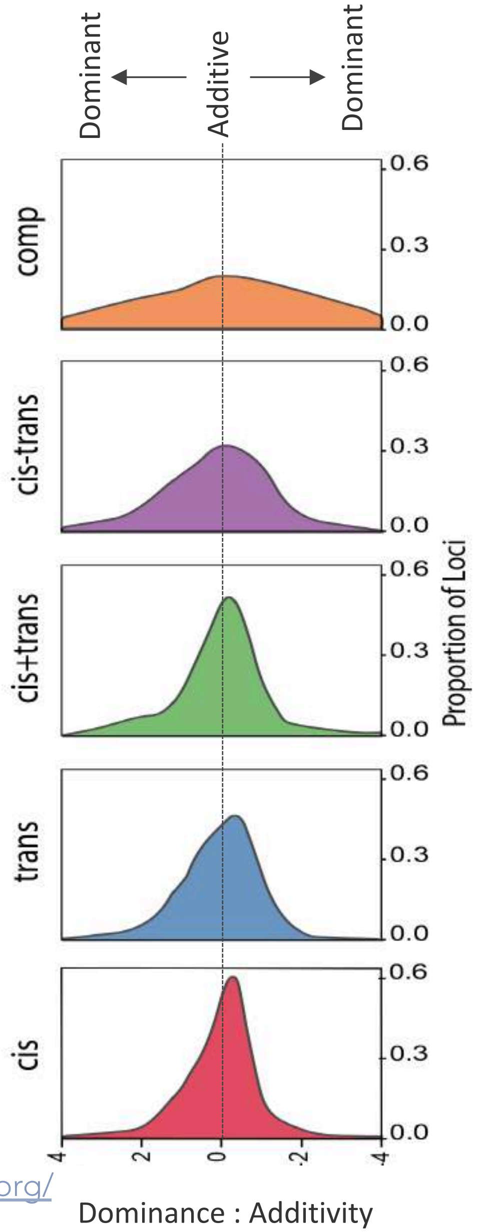
Verta & Jones (in review)
BioRxiv
doi: https://doi.org/10.1101/412932

The degree of parallelism in gene regulation appears to be associated with regulatory effect size



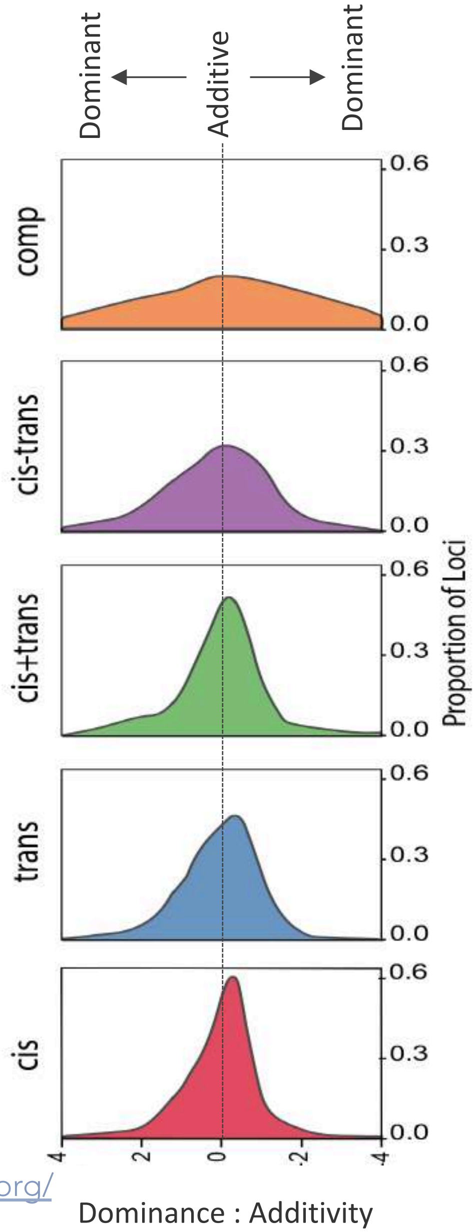
Verta & Jones
(in review)
BioRxiv
doi: https://doi.org/10.1101/412932

Cis-regulation is more additive than other forms of gene expression regulation – more visible to natural selection



Verta & Jones
(in review)
BioRxiv
doi: https://doi.org/10.1101/412932

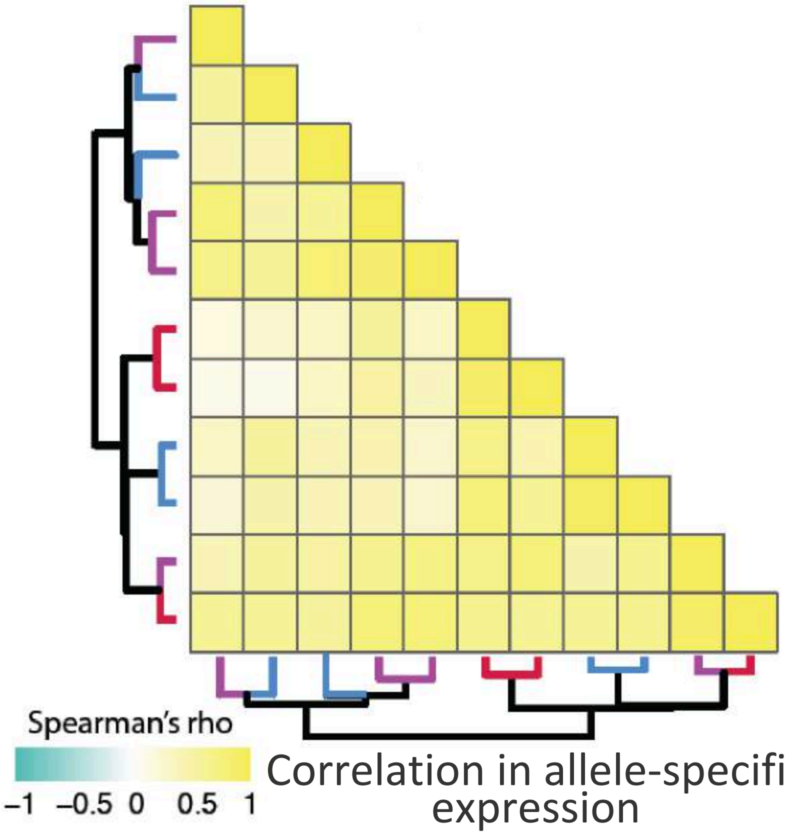
Cis-regulation is more additive than other forms of gene expression regulation – more visible to natural selection



Verta & Jones
(in review)
BioRxiv
doi: https://doi.org/10.1101/412932

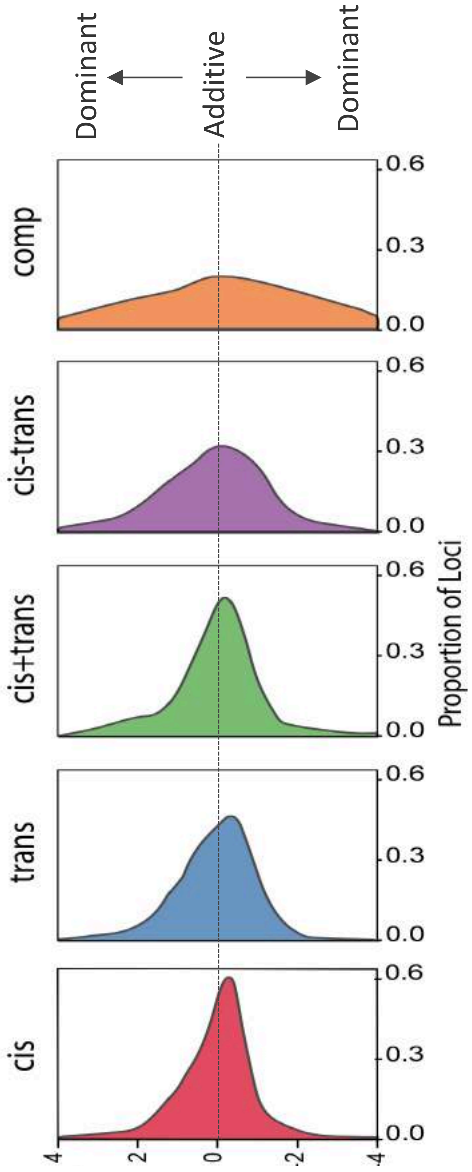
And stable under different environmental conditions

Siblings raised under different environmental conditions show strong correlation in cis-regulation



Correlation in allele-specific expression

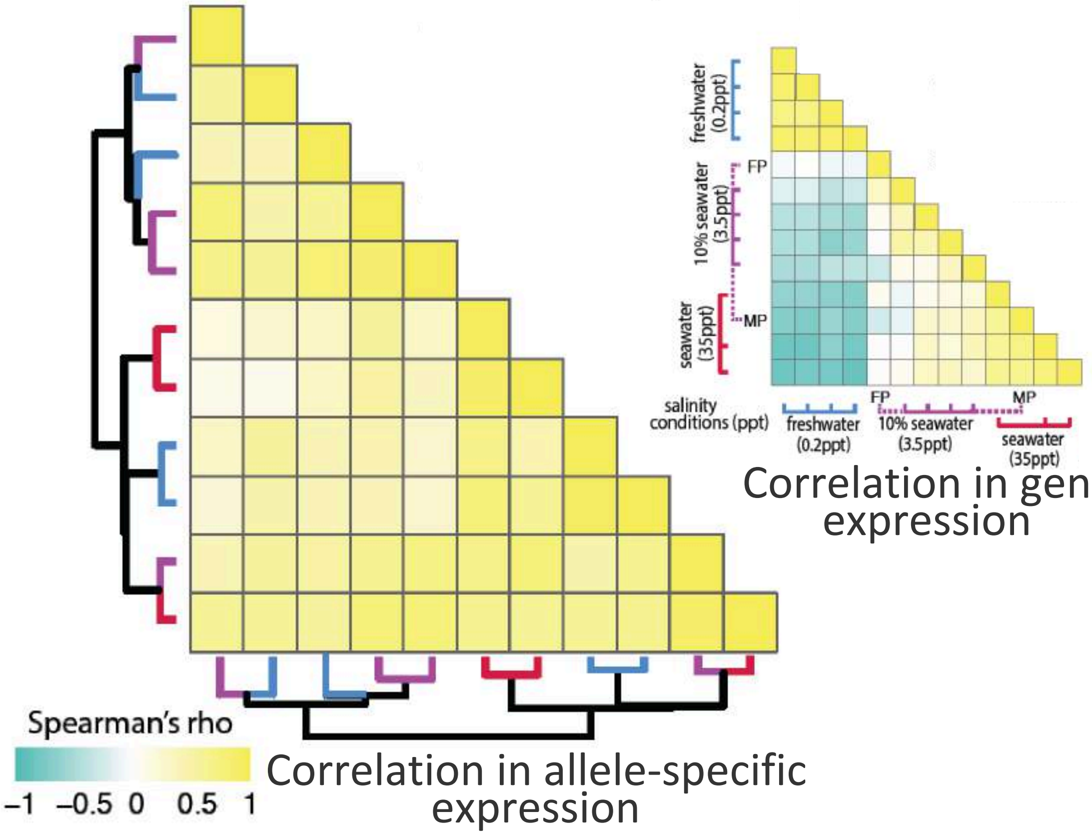
Cis-regulation is more additive than other forms of gene expression regulation – more visible to natural selection



Verta & Jones
(in review)
BioRxiv
doi: https://doi.org/10.1101/412932

And stable under different environmental conditions

Siblings raised under different environmental conditions show strong correlation in cis-regulation

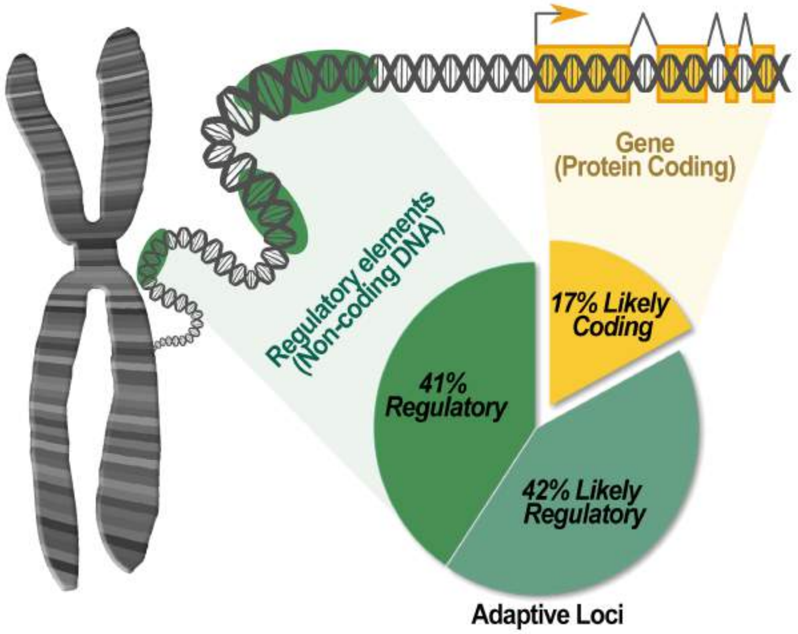


Correlation in gene expression

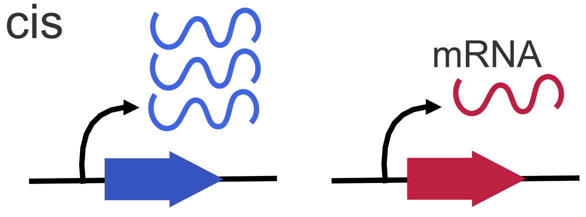
Correlation in allele-specific expression

Molecular mechanisms playing a role in parallel adaptive divergence of sticklebacks:

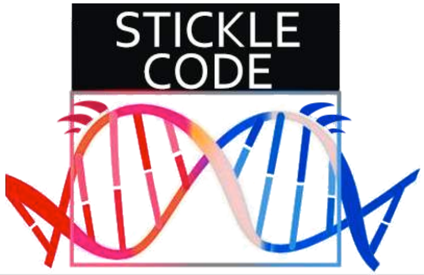
1 Regulatory mechanisms



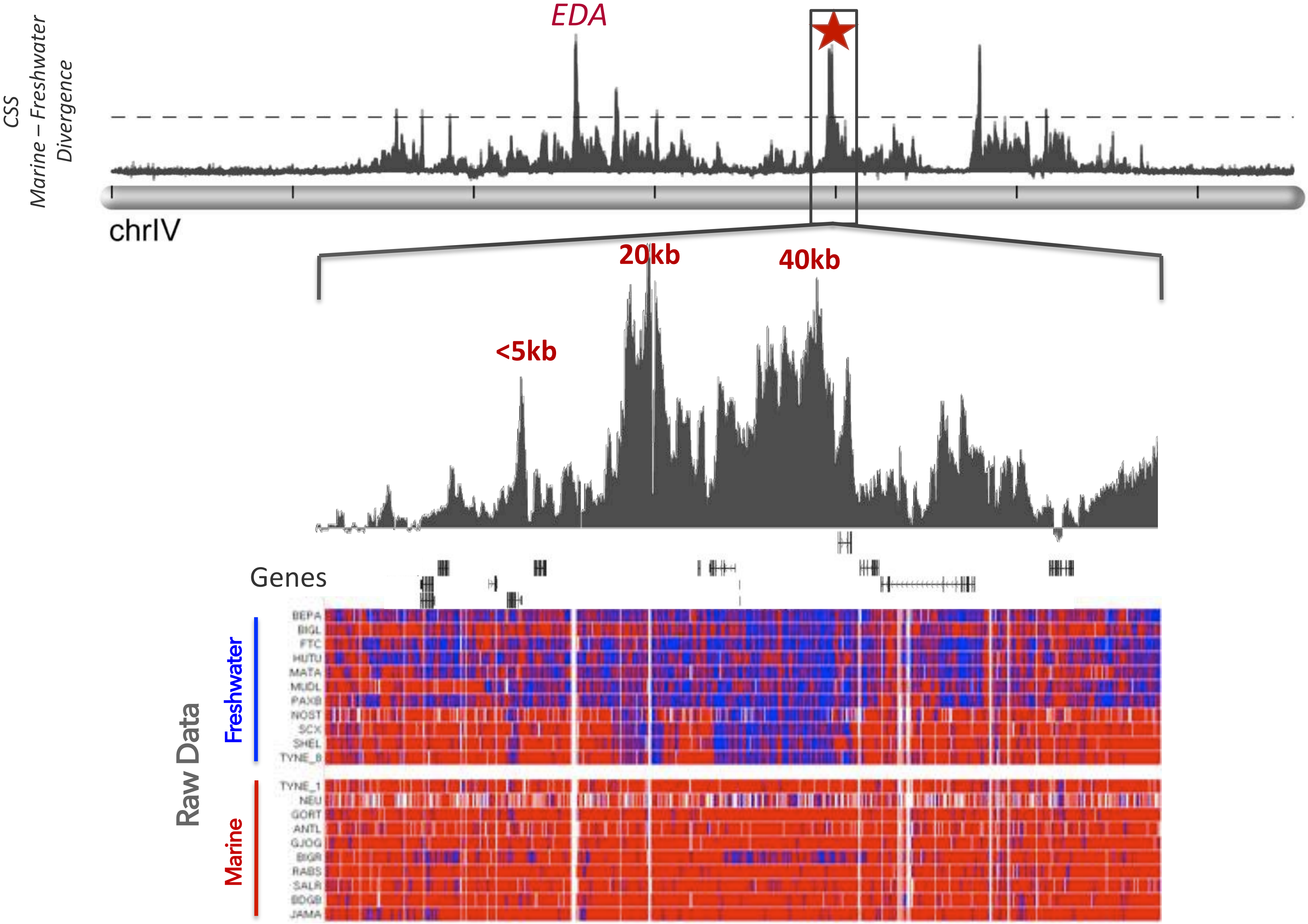
How is adaptive divergence in gene expression regulated?



What are the mutations & functional elements?



High resolution signals, good for functional follow-up … but ancient haplotypes carry many candidate mutations



Jones et al (2012) Nature

Identifying regulatory elements underlying divergent adaptation

Dr Stanley Neufeld

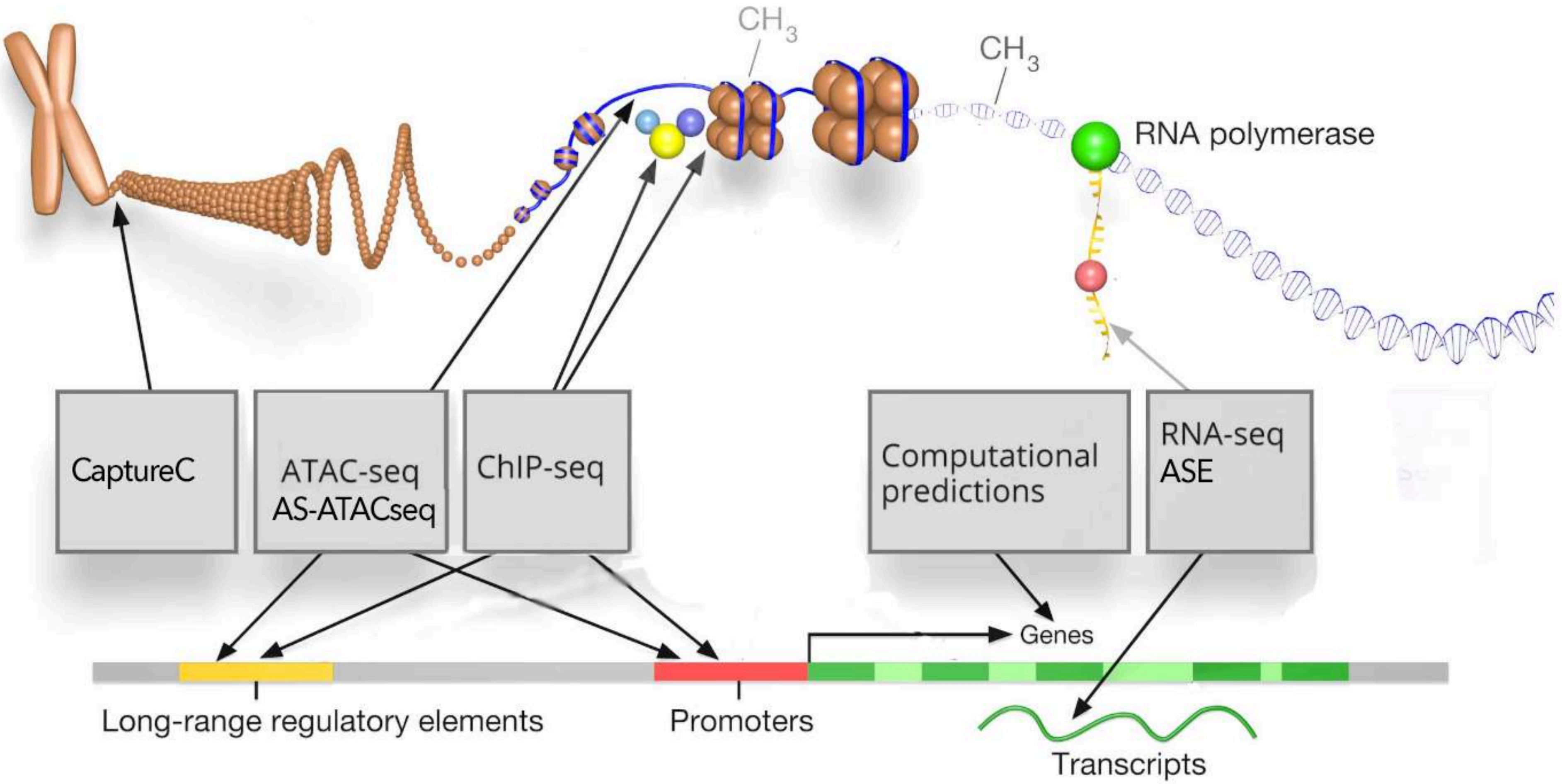
STICKLE CODE

Comparative functional annotations of adaptive loci using epigenomics

ChipSeq, ATACseq, RNAseq & CaptureC

Mapping poised & active enhancers, open chromatin
differential expression and regulatory interactions in marine and freshwater genomes

Open chromatin and epigenetic modifications to histones are predictive of functional and active regulatory elements



CaptureC

ATAC-seq
AS-ATACseq

ChIP-seq

Computational predictions

RNA-seq
ASE

Genes

Long-range regulatory elements

Promoters

Transcripts

STICKLE CODE

Modified from image from ENCODE - NHGRI & EBI

HMM epigenetic states

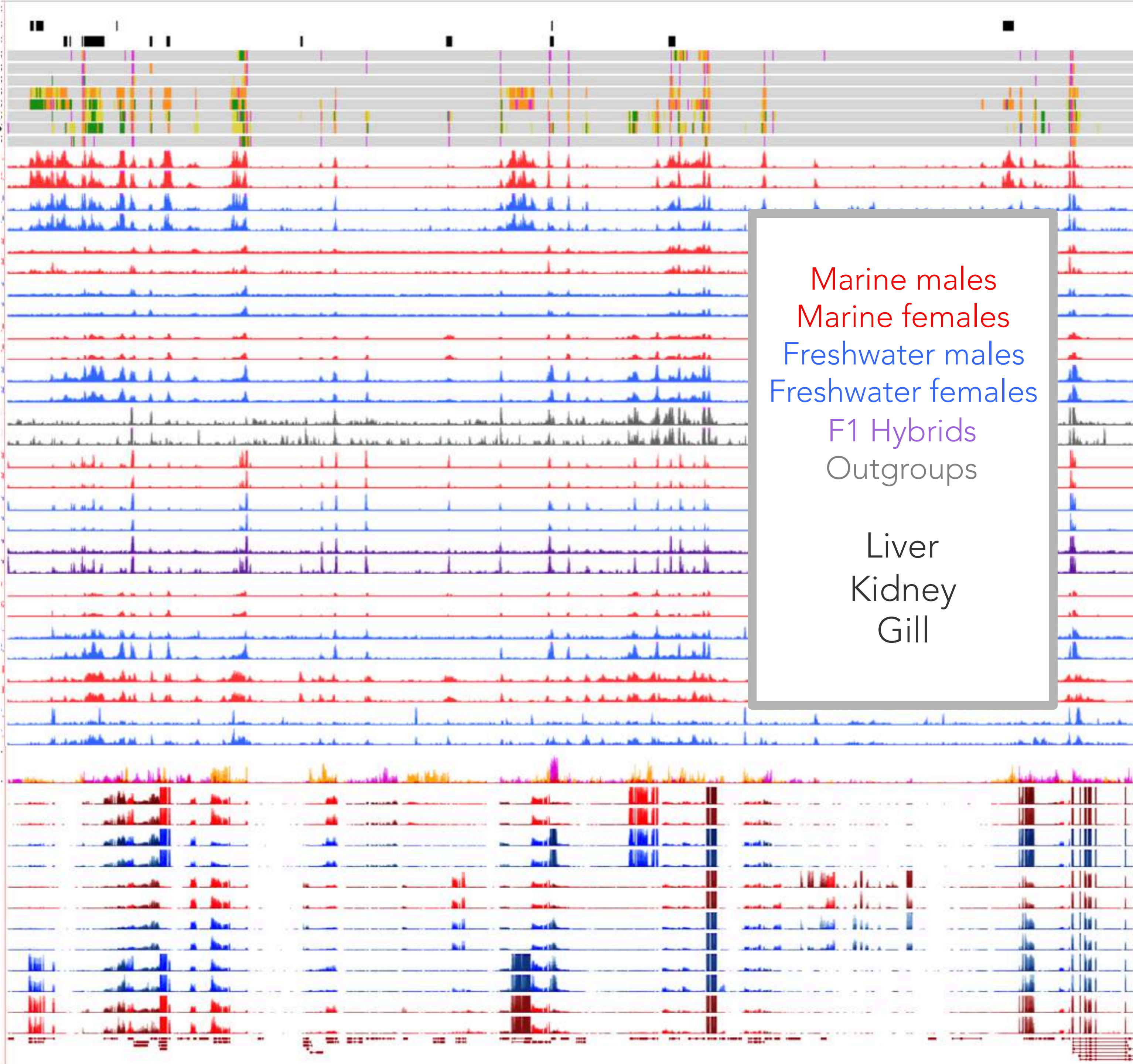
Histone modifications:

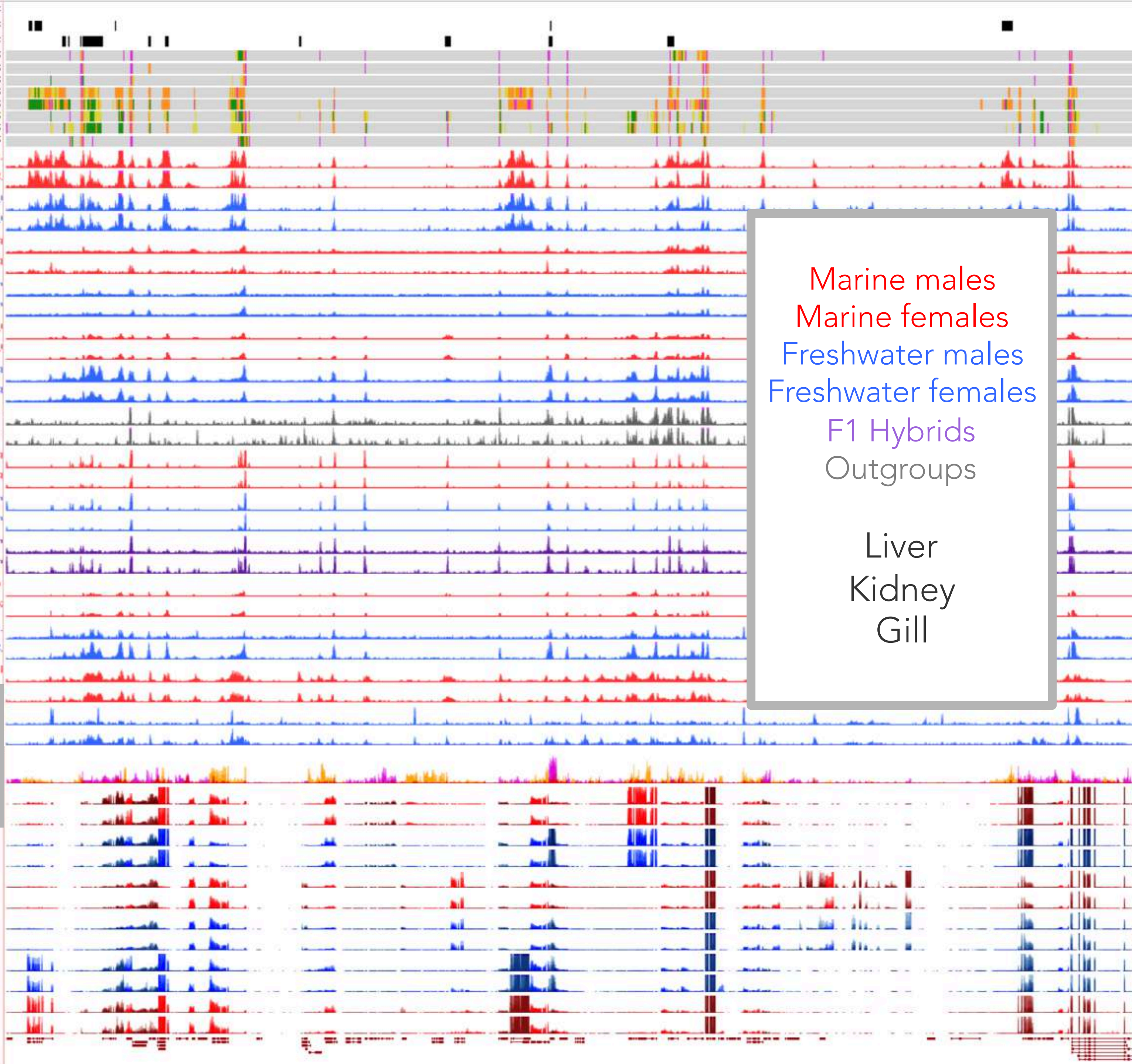
H3K27ac ChIPseq
H3K4me1 ChIPseq
H3K4me3 ChIPseq

Open chromatin
ATACseq

Gene expression
RNAseq

Transcript annotations





Marine v Fresh divergent enhancers

Histone modifications:
H3K27ac ChIPseq
H3K4me1 ChIPseq
H3K4me3 ChIPseq

Open chromatin
ATACseq

Marine v Fresh divergent gene expression

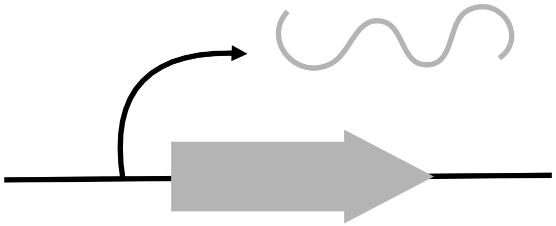
Gene expression
RNAseq

Transcript annotations

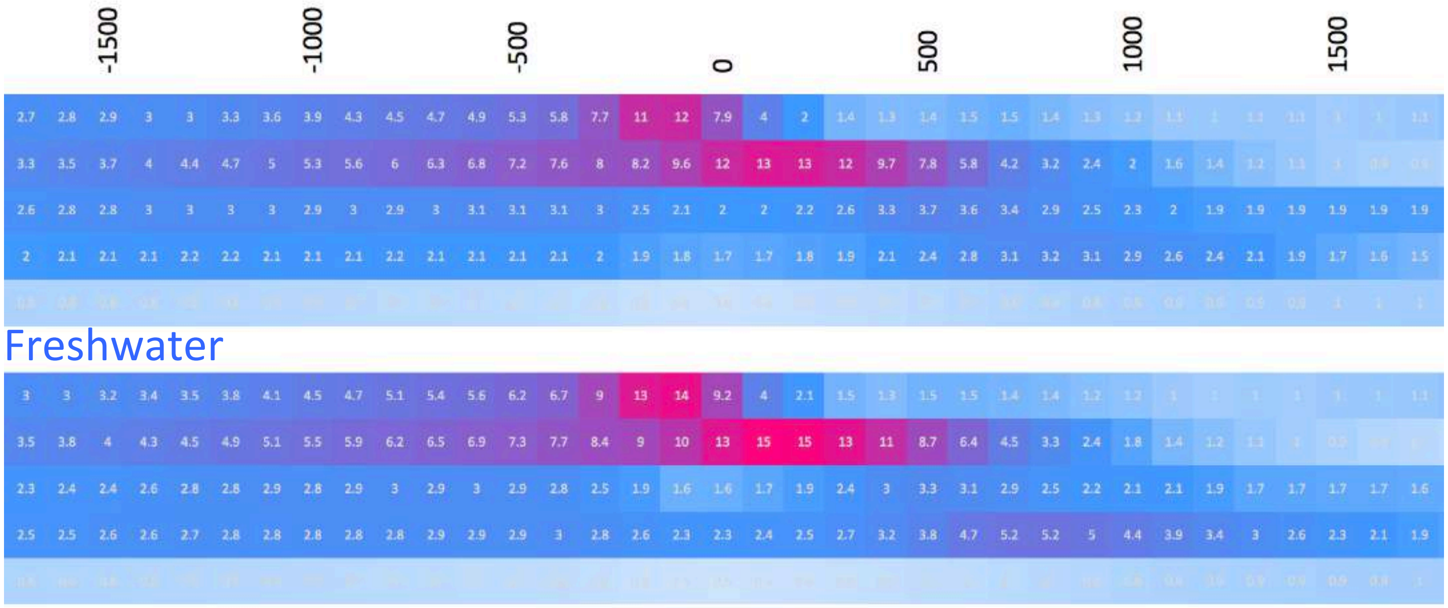
Marine males
Marine females
Freshwater males
Freshwater females
F1 Hybrids
Outgroups

Liver
Kidney
Gill

Enhancer marks are enriched around Transcription Start Sites



Chromatin State



			-1500					-1000					-500					0					500					1000					1500		
open chromatin	2.7	2.8	2.9	3	3	3.3	3.6	3.9	4.3	4.5	4.7	4.9	5.3	5.8	7.7	11	12	7.9	4	2	1.4	1.3	1.4	1.5	1.5	1.4	1.3	1.2	1.1	1	1.1	1.1	1	1	1.1
H3K27ac	3.3	3.5	3.7	4	4.4	4.7	5	5.3	5.6	6	6.3	6.8	7.2	7.6	8	8.2	9.6	12	13	13	12	9.7	7.8	5.8	4.2	3.2	2.4	2	1.6	1.4	1.2	1.1	1	0.9	0.9
H3K27ac & H3K4me1	2.6	2.8	2.8	3	3	3	3	2.9	3	2.9	3	3.1	3.1	3.1	3	2.5	2.1	2	2	2.2	2.6	3.3	3.7	3.6	3.4	2.9	2.5	2.3	2	1.9	1.9	1.9	1.9	1.9	1.9
H3K4me1	2	2.1	2.1	2.1	2.2	2.2	2.1	2.1	2.1	2.2	2.1	2.1	2.1	2.1	2	1.9	1.8	1.7	1.7	1.8	1.9	2.1	2.4	2.8	3.1	3.2	3.1	2.9	2.6	2.4	2.1	1.9	1.7	1.6	1.5
no mark	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.8	0.8	0.9	0.9	0.9	0.8	1	1	1

Freshwater

			-1500					-1000					-500					0					500					1000					1500		
open chromatin	3	3	3.2	3.4	3.5	3.8	4.1	4.5	4.7	5.1	5.4	5.6	6.2	6.7	9	13	14	9.2	4	2.1	1.5	1.3	1.5	1.5	1.4	1.4	1.2	1.2	1	1	1	1	1	1	1.1
H3K27ac	3.5	3.8	4	4.3	4.5	4.9	5.1	5.5	5.9	6.2	6.5	6.9	7.3	7.7	8.4	9	10	13	15	15	13	11	8.7	6.4	4.5	3.3	2.4	1.8	1.4	1.2	1.1	1	0.9	0.9	1
H3K27ac & H3K4me1	2.3	2.4	2.4	2.6	2.8	2.8	2.9	2.8	2.9	3	2.9	3	2.9	2.8	2.5	1.9	1.6	1.6	1.7	1.9	2.4	3	3.3	3.1	2.9	2.5	2.2	2.1	2.1	1.9	1.7	1.7	1.7	1.7	1.6
H3K4me1	2.5	2.5	2.6	2.6	2.7	2.8	2.8	2.8	2.8	2.8	2.9	2.9	2.9	3	2.8	2.6	2.3	2.3	2.4	2.5	2.7	3.2	3.8	4.7	5.2	5.2	5	4.4	3.9	3.4	3	2.6	2.3	2.1	1.9
no mark	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.8	0.8	0.8	0.8	0.9	0.9	0.9	0.8	1

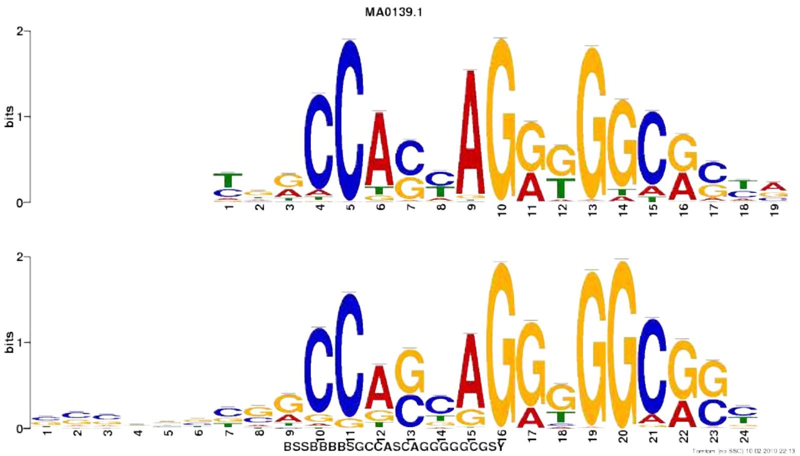
Marine

Fold enrichment

0	5	10	15

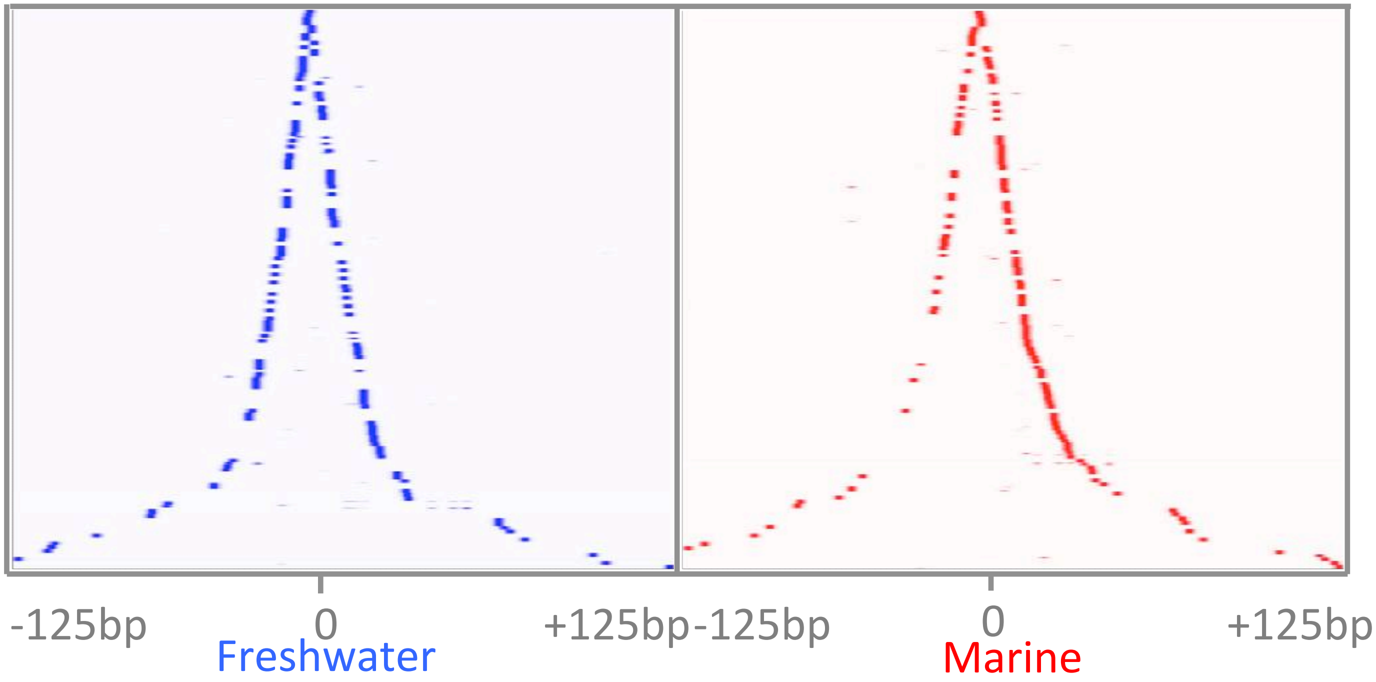
Open chromatin marks are enriched for CTCF binding sites

$p=1\times10^{-1679}$

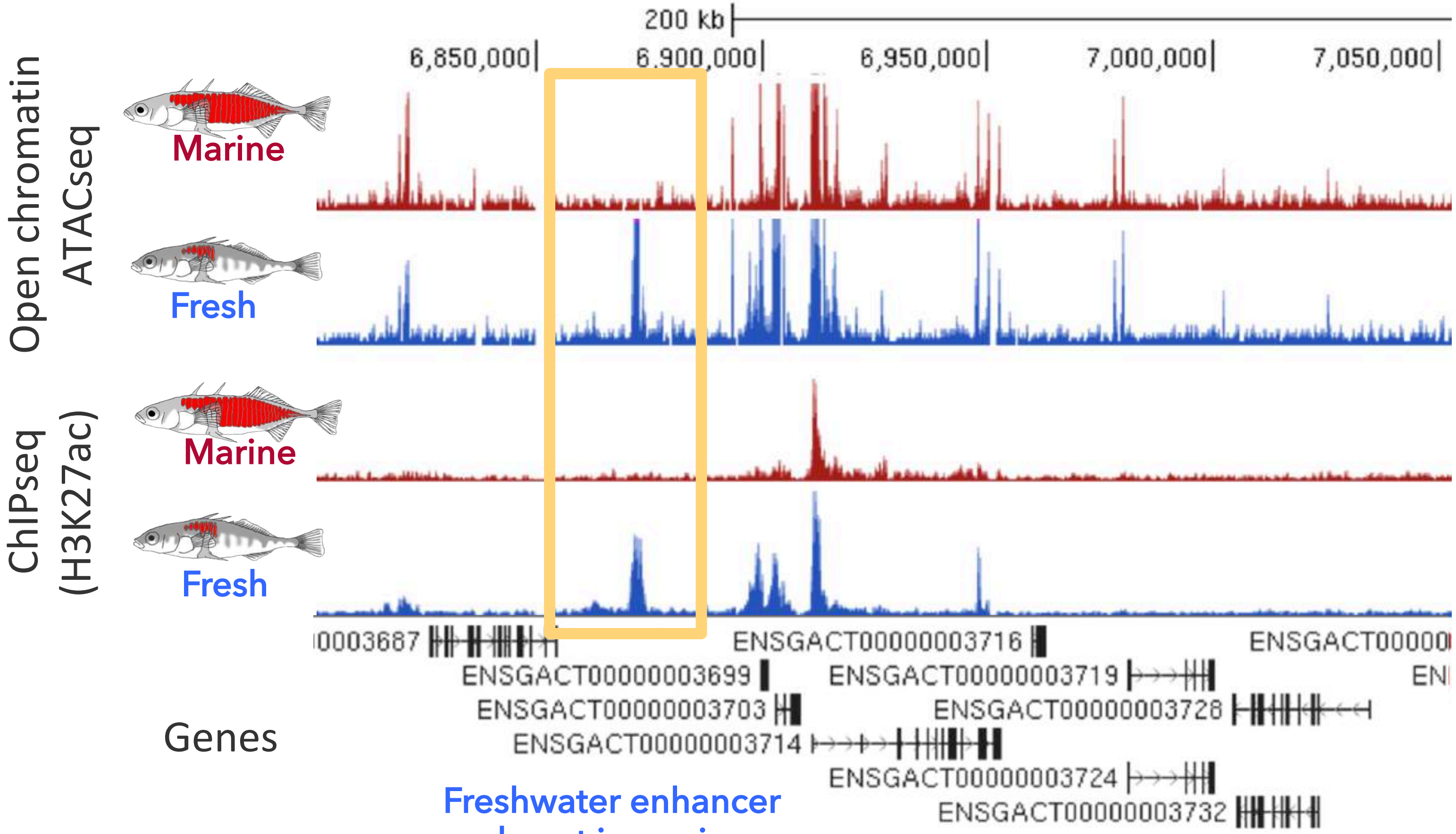


CTCF: 11zinc finger
DNA binding protein

- 3D chromatin structure
- Transcriptional regulation
- Insulation



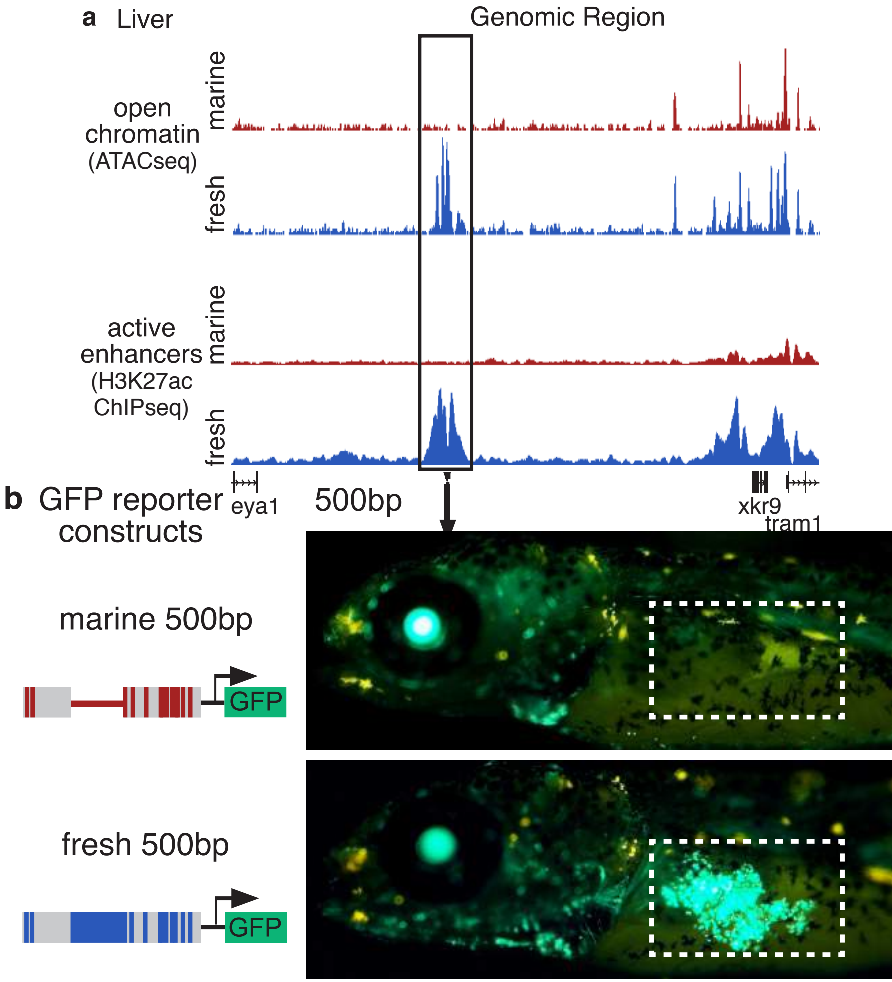
StickleCODE reveals high resolution signals of differential regulatory marks in marine & freshwater genomes



Freshwater enhancer absent in marine

Functional tests of divergent regulatory elements

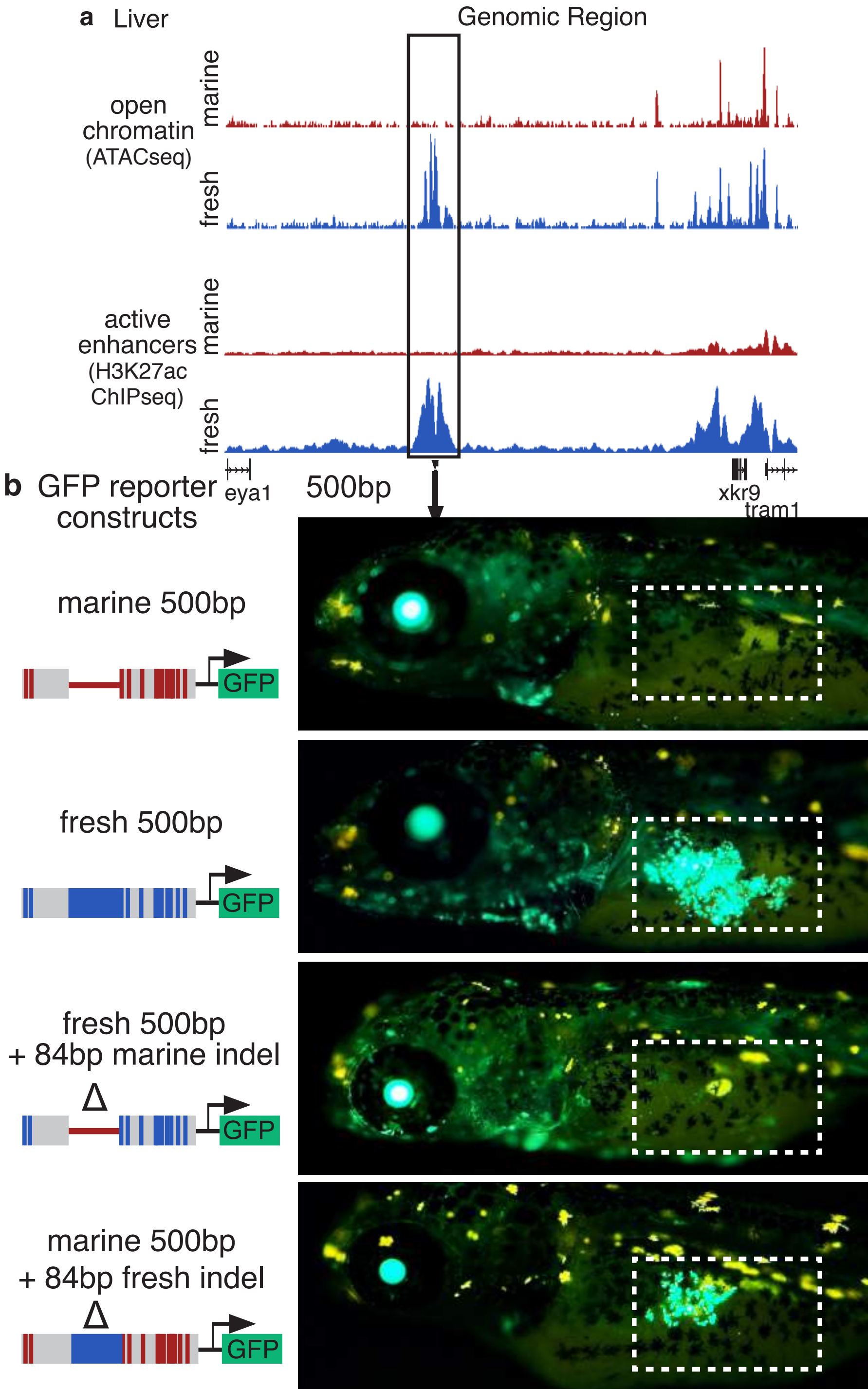
Divergent chromatin states to *cis*-controlled enhancers of divergent liver expression



Functional tests of divergent regulatory elements

Divergent chromatin states to cis-controlled enhancers of divergent liver expression

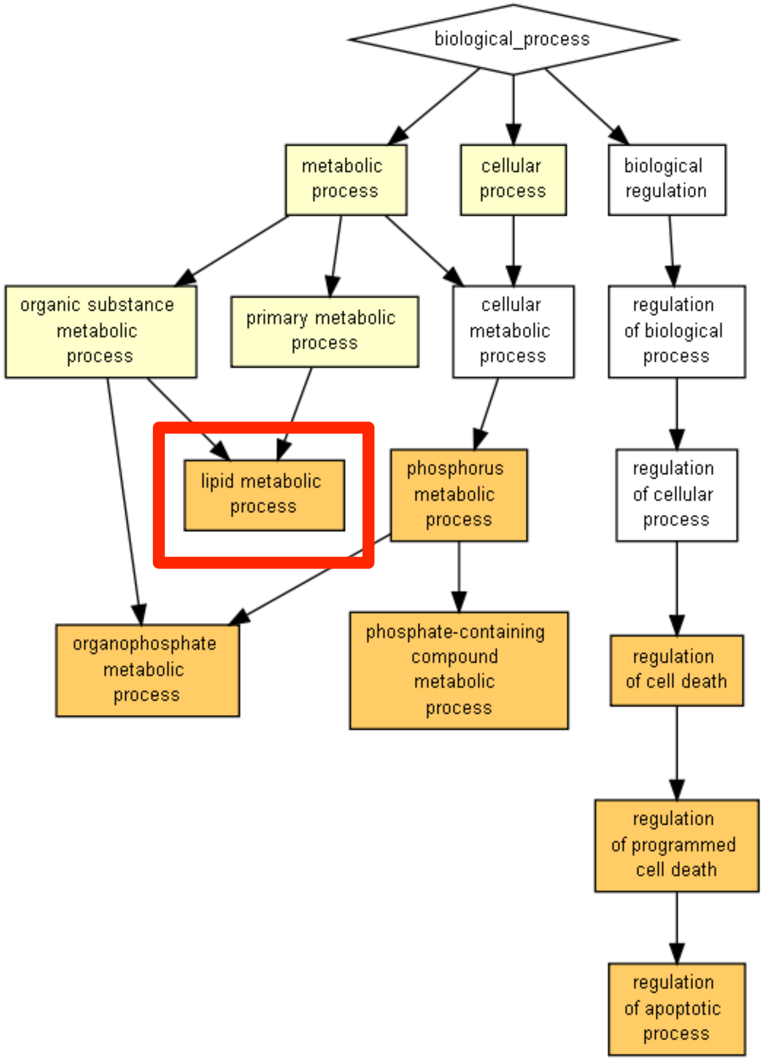
An 84bp insertion is necessary and sufficient for freshwater liver expression



The liver is particularly enriched for divergent regulatory marks between ecotypes

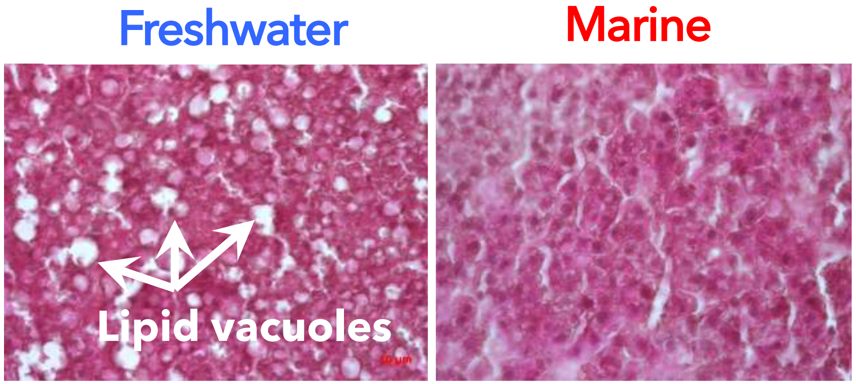
Tissue	**Total peaks**	**Divergent peaks (FDR 1%)**	**%**
Liver	23394	2148	9.1
Kidney	27638	458	1.7
Gill	17006	178	1.0

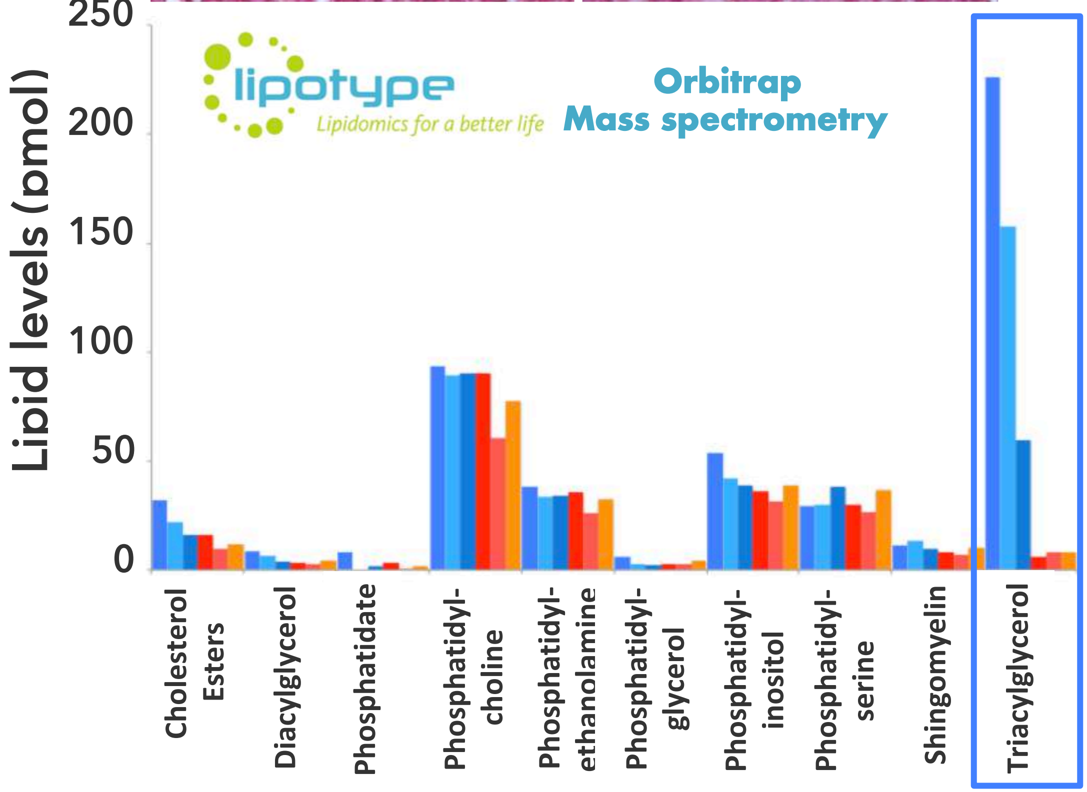
Genes proximal to marine-freshwater divergent liver enhancers are enriched for lipid metabolism



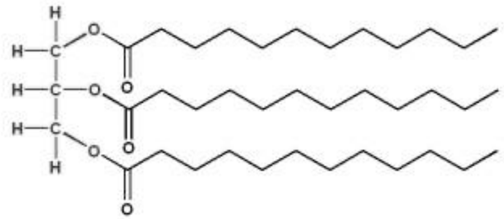
Description	P-value	Enrichment
organophosphate metabolic process	5.81E-07	1.85
lipid metabolic process	**9.02E-07**	**1.68**
regulation of programmed cell death	2.60E-06	1.89
phosphate-containing compound metabolic process	3.24E-06	1.49
regulation of apoptotic process	3.33E-06	1.89
phosphorus metabolic process	3.34E-06	1.48
regulation of cell death	5.61E-06	1.81

The livers of marine and freshwater sticklebacks differ considerably in size and lipid content





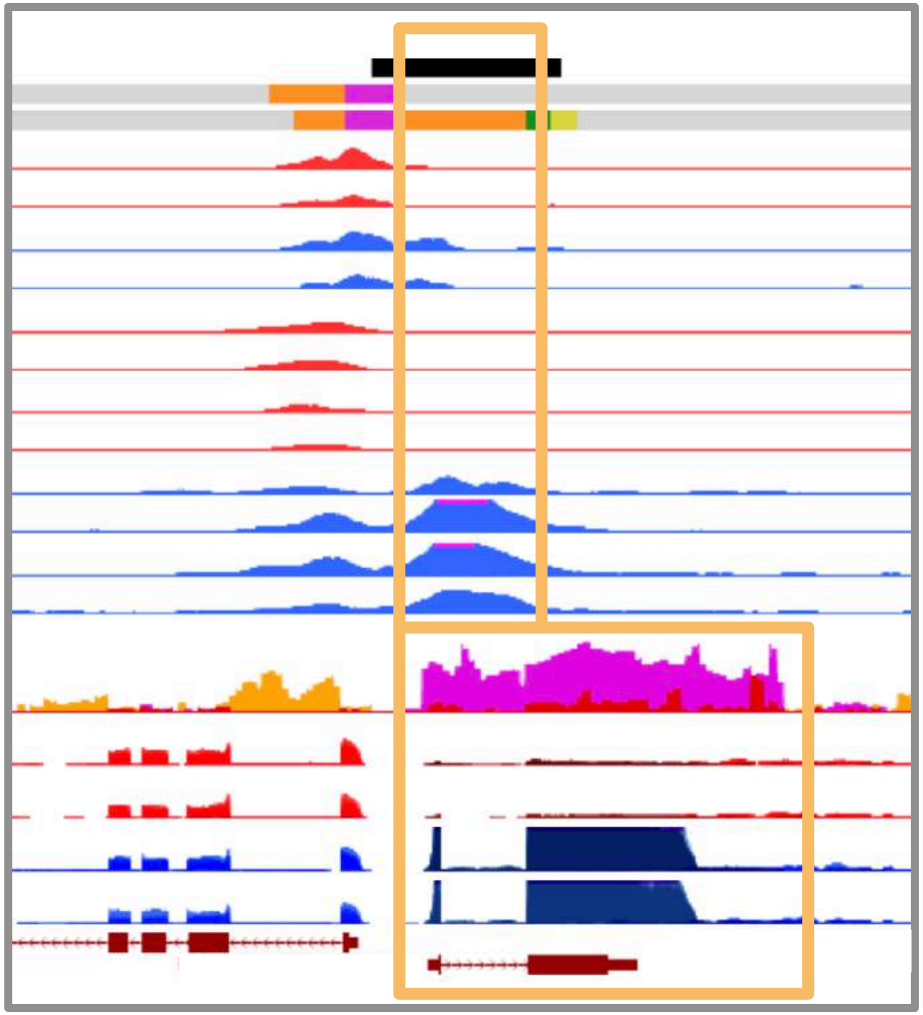
Under common lab conditions freshwater sticklebacks have >30x more triacylglycerides in their livers than marine sticklebacks



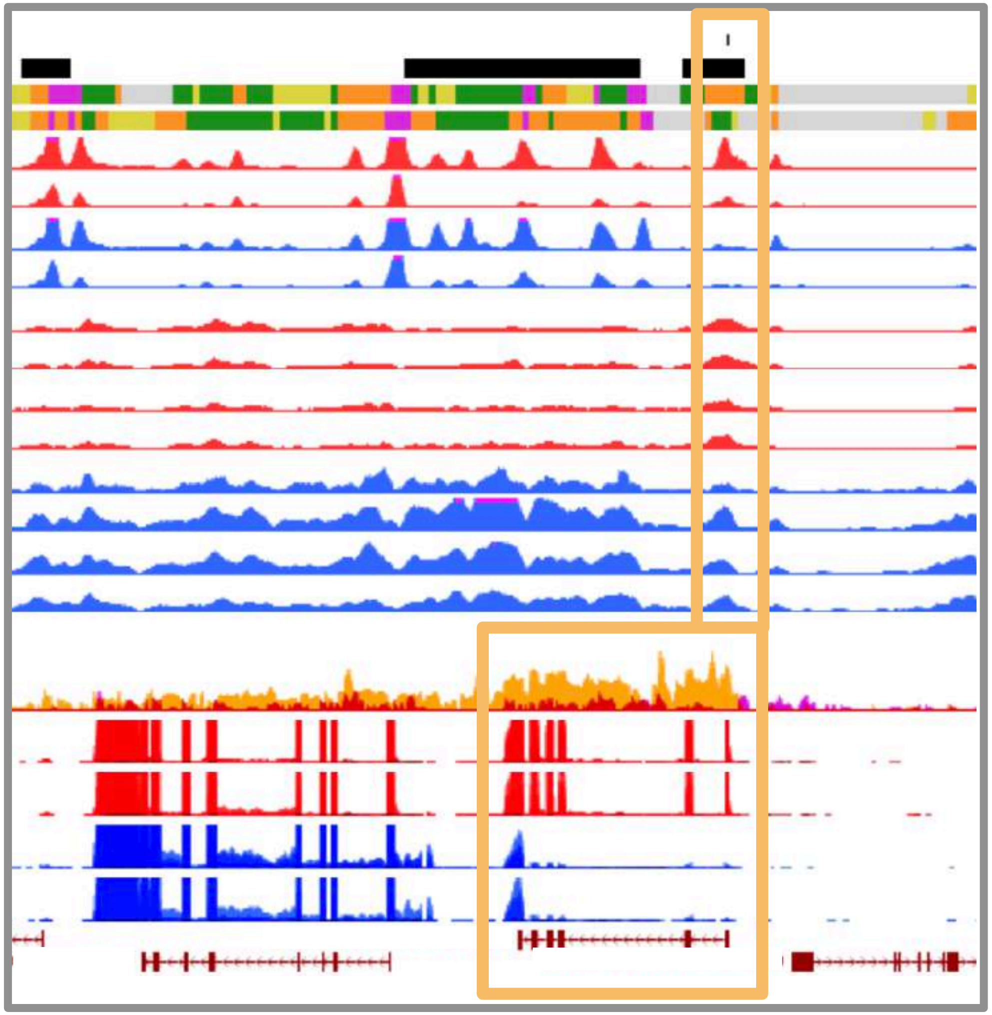
Major form of energy storage

Blood glucose ⟷ adipose

Divergent enhancers around genes that play key roles in obesity, fatty liver and insulin resistance

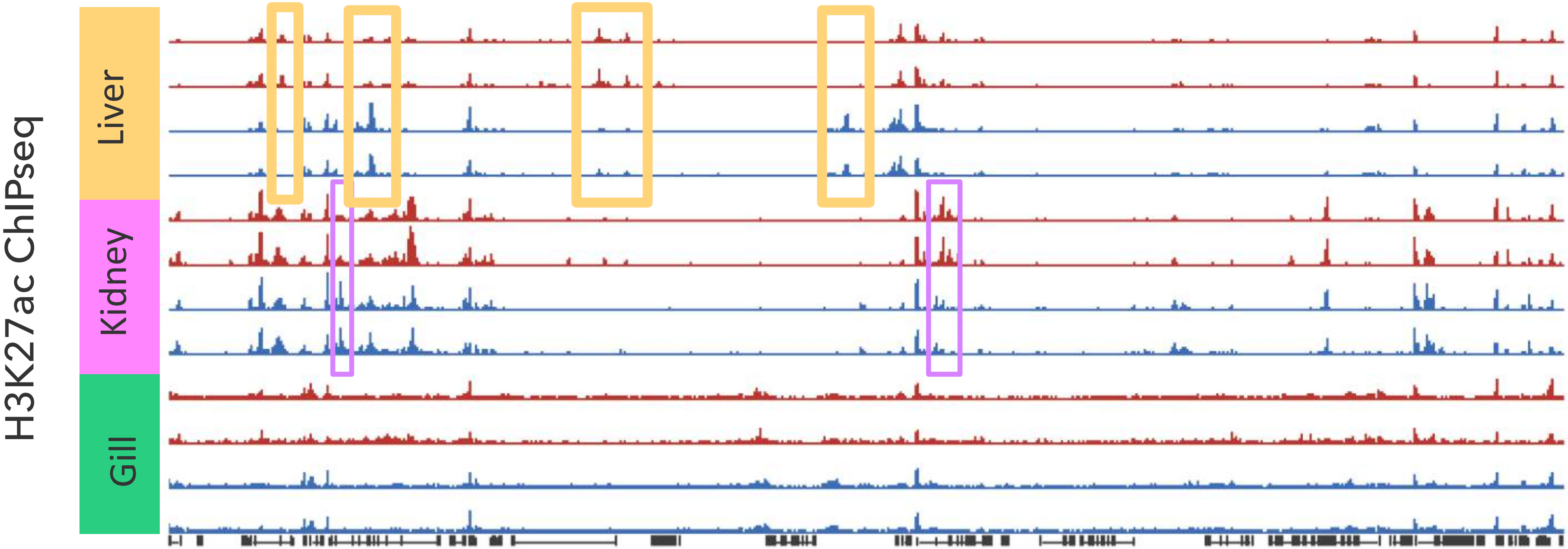


Overexpression in mice blocks high-fat-induced insulin resistance

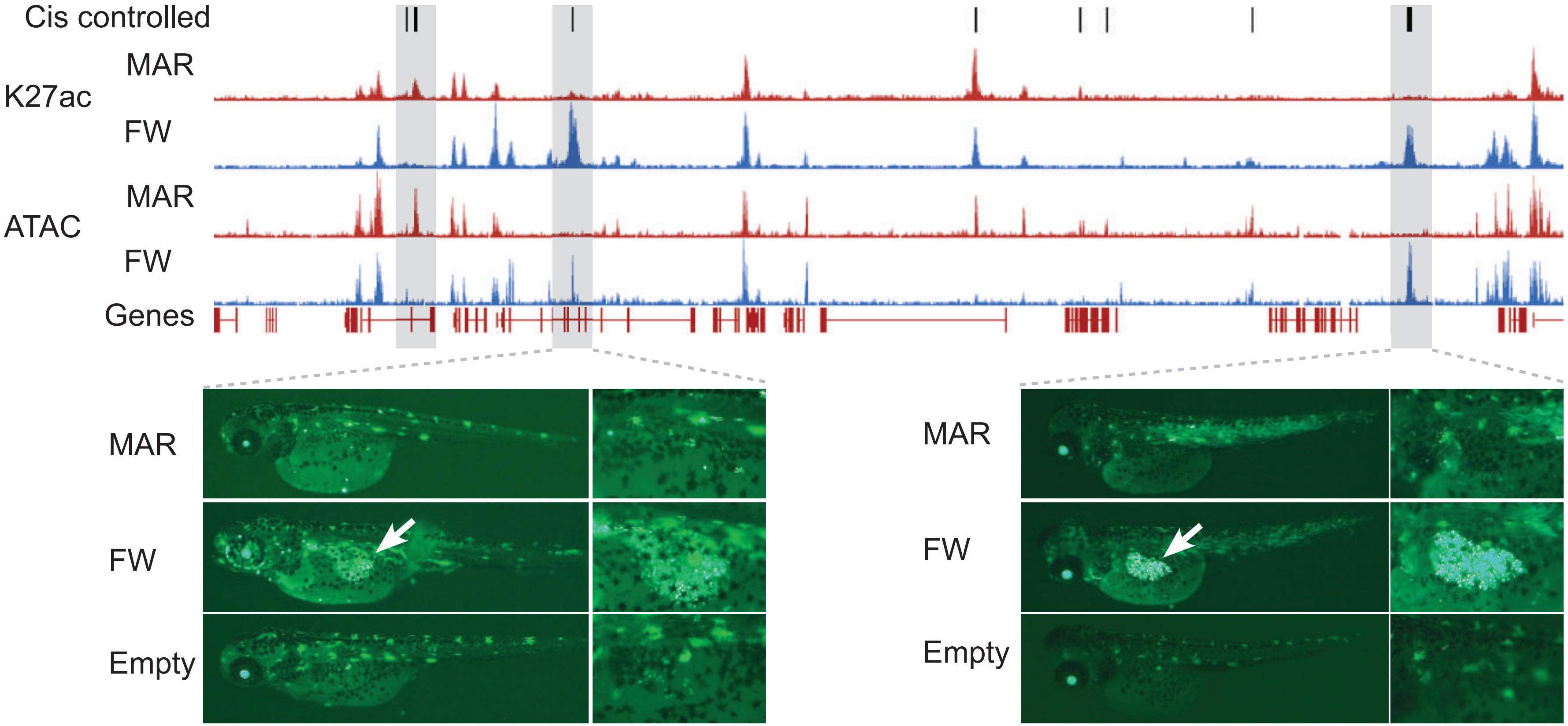


Knock-out improves glucose metabolism in obese mice

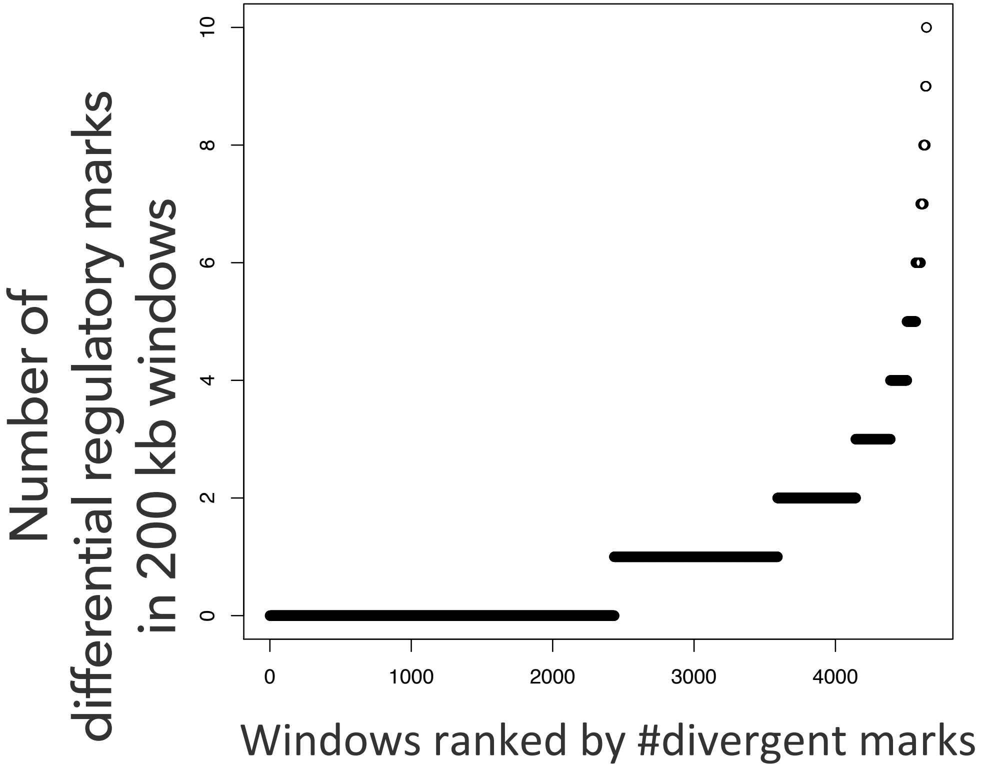
Divergent regulatory elements often appear in clusters across the genome in a tissue-specific manner



An inversion on chromosome XXI contains multiple freshwater enhancers driving expression in the liver



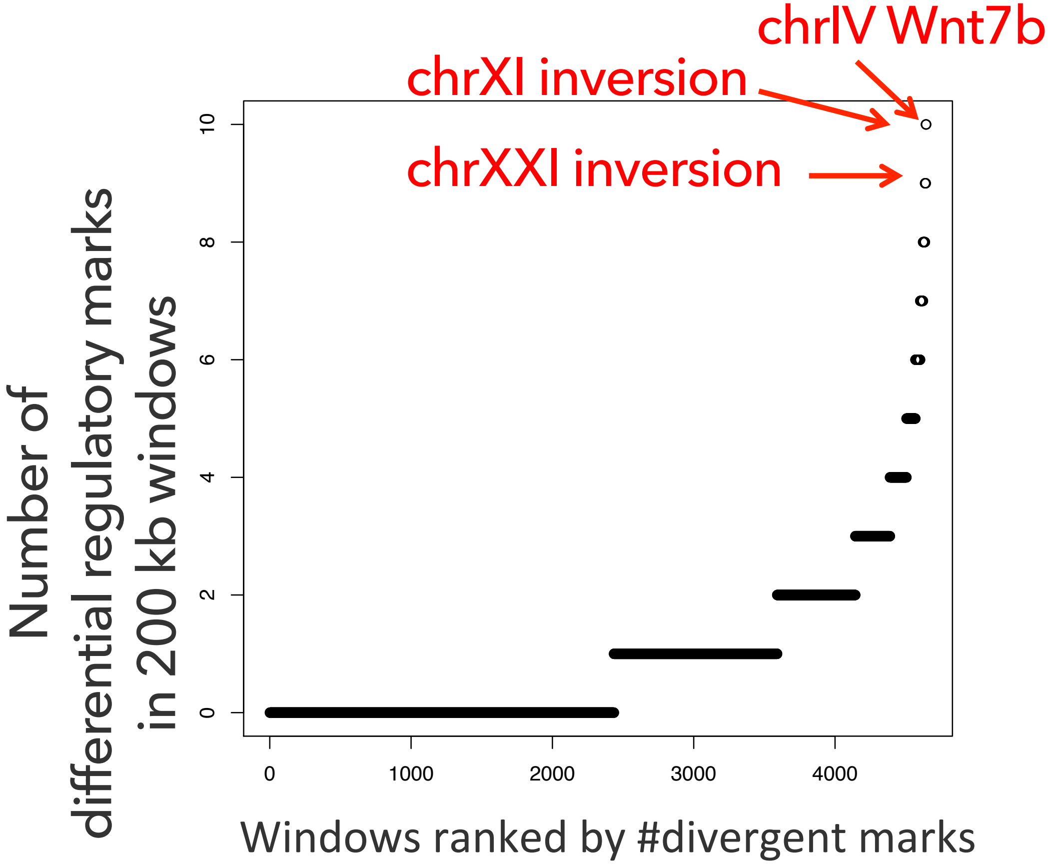
Genome-wide, a small number of genomic windows contain many divergent regulatory marks



Liver

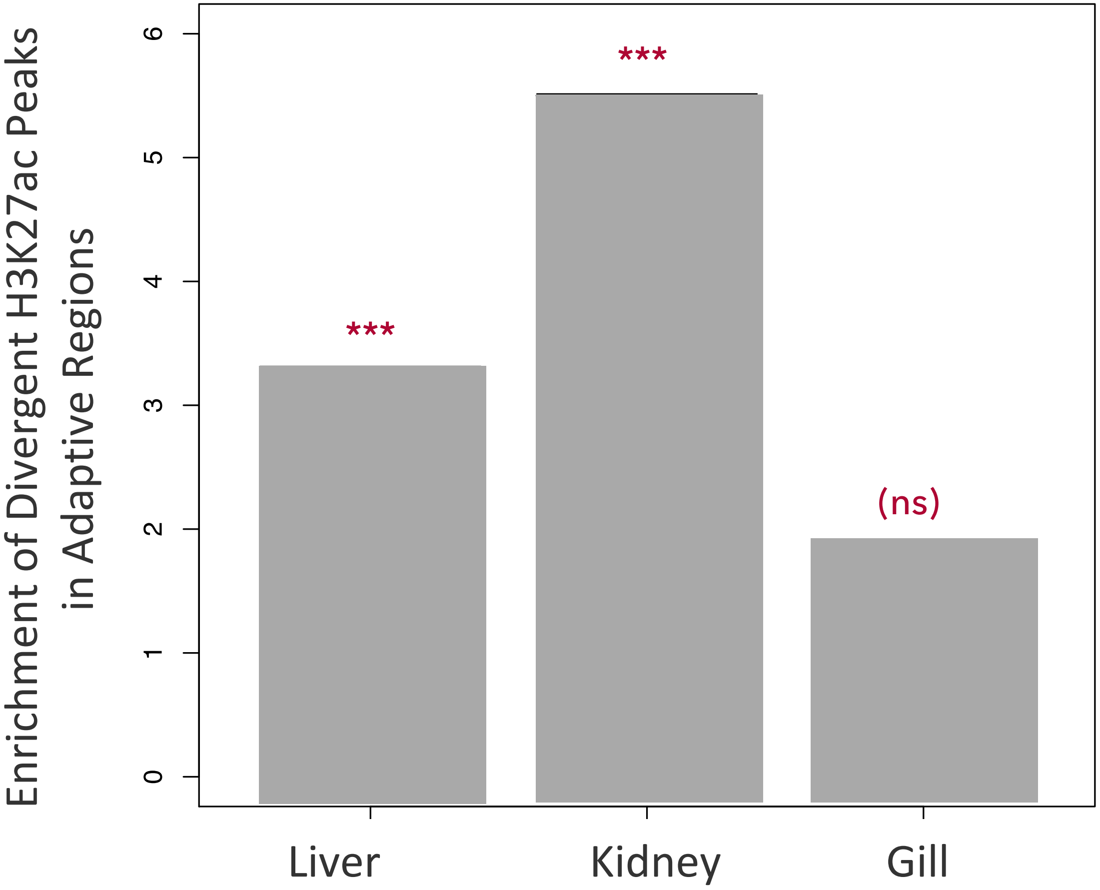
Genome-wide, a small number of genomic windows contain many divergent regulatory marks

(and are "old friends" in the lab)



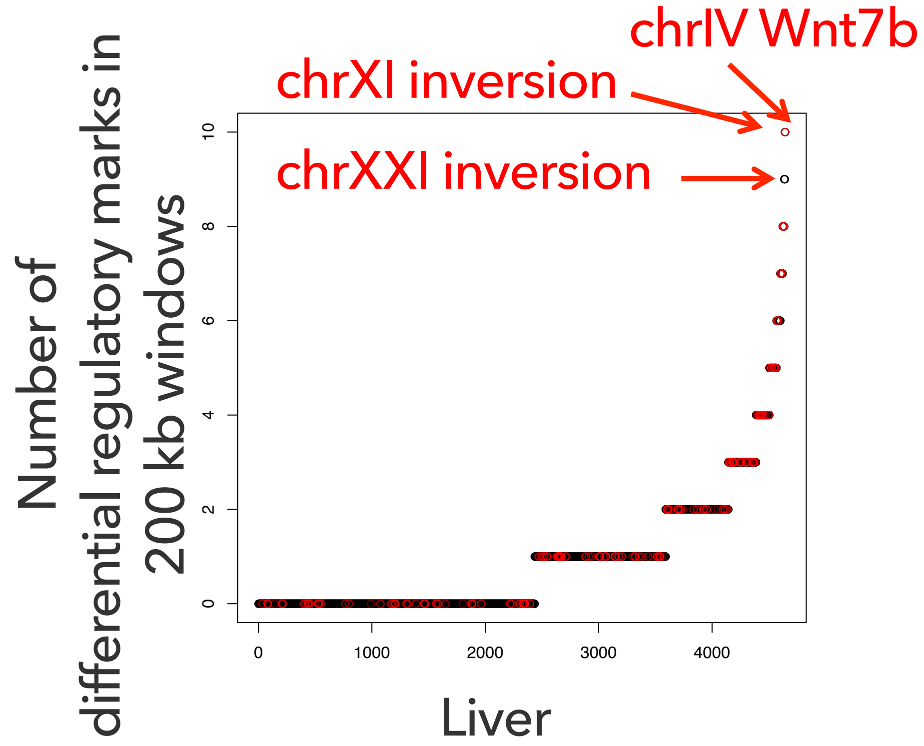
Liver

Adaptive genomic regions identified in previous study are enriched in divergent enhancers



These marine-freshwater adaptive loci are enriched for clusters of divergent regulatory marks

(Fisher's exact test: $p<1x10^{-51}$)



Clusters of divergent enhancers drive adaptive evolution in sticklebacks

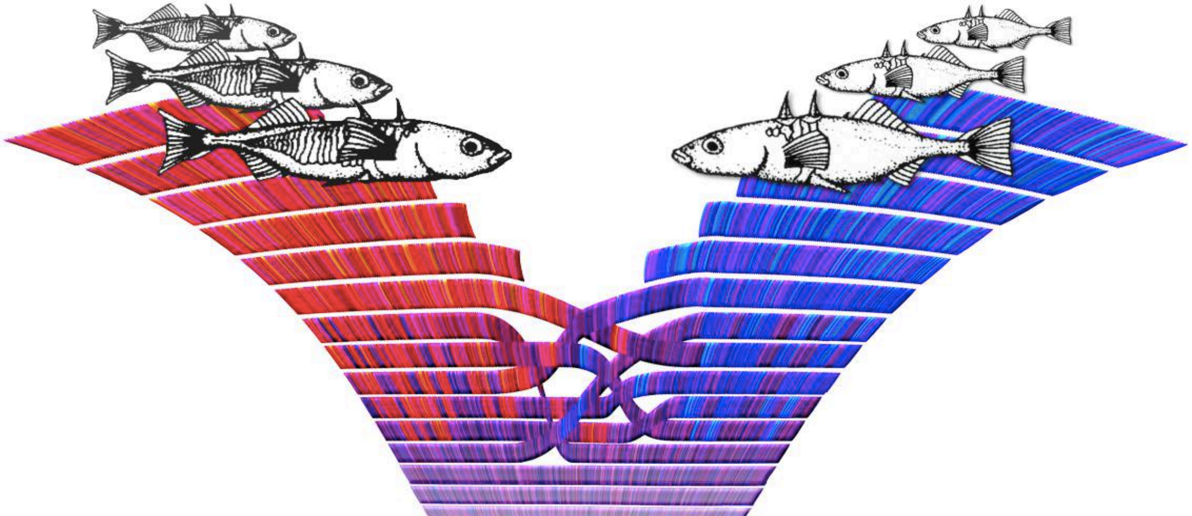
Molecular mechanisms of adaptation & speciation

1 The regulatory control of adaptive divergence in gene expression

cis-regulation is predominant

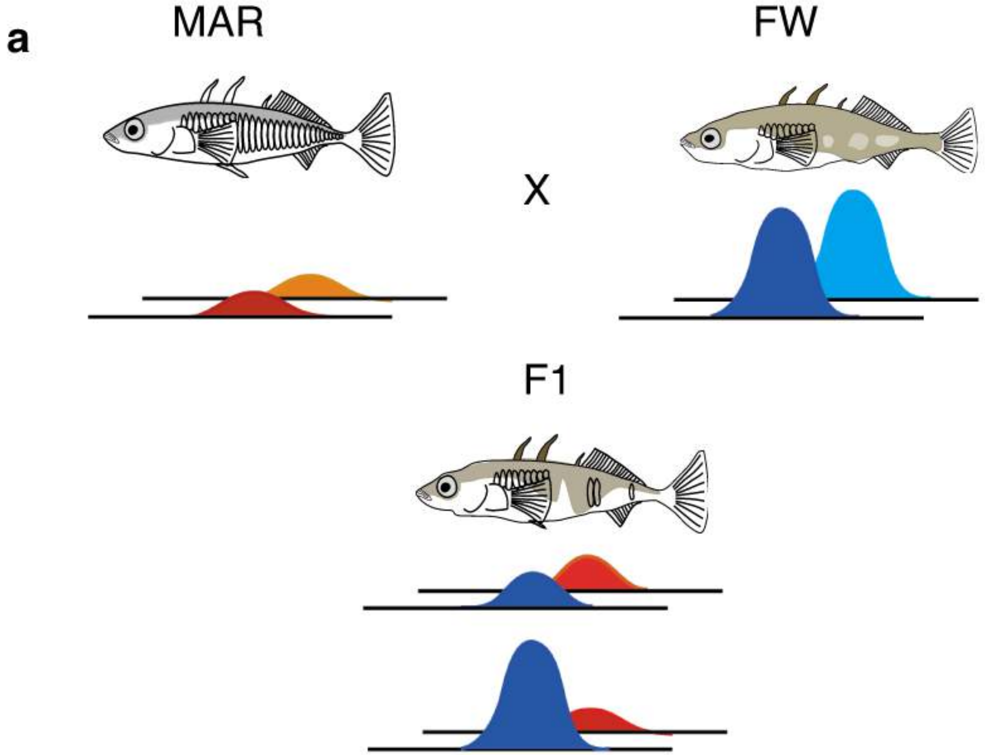
large effect size, additive, stable…

2 Comparative epigenomics reveals clusters of enhancer elements in adaptive loci.

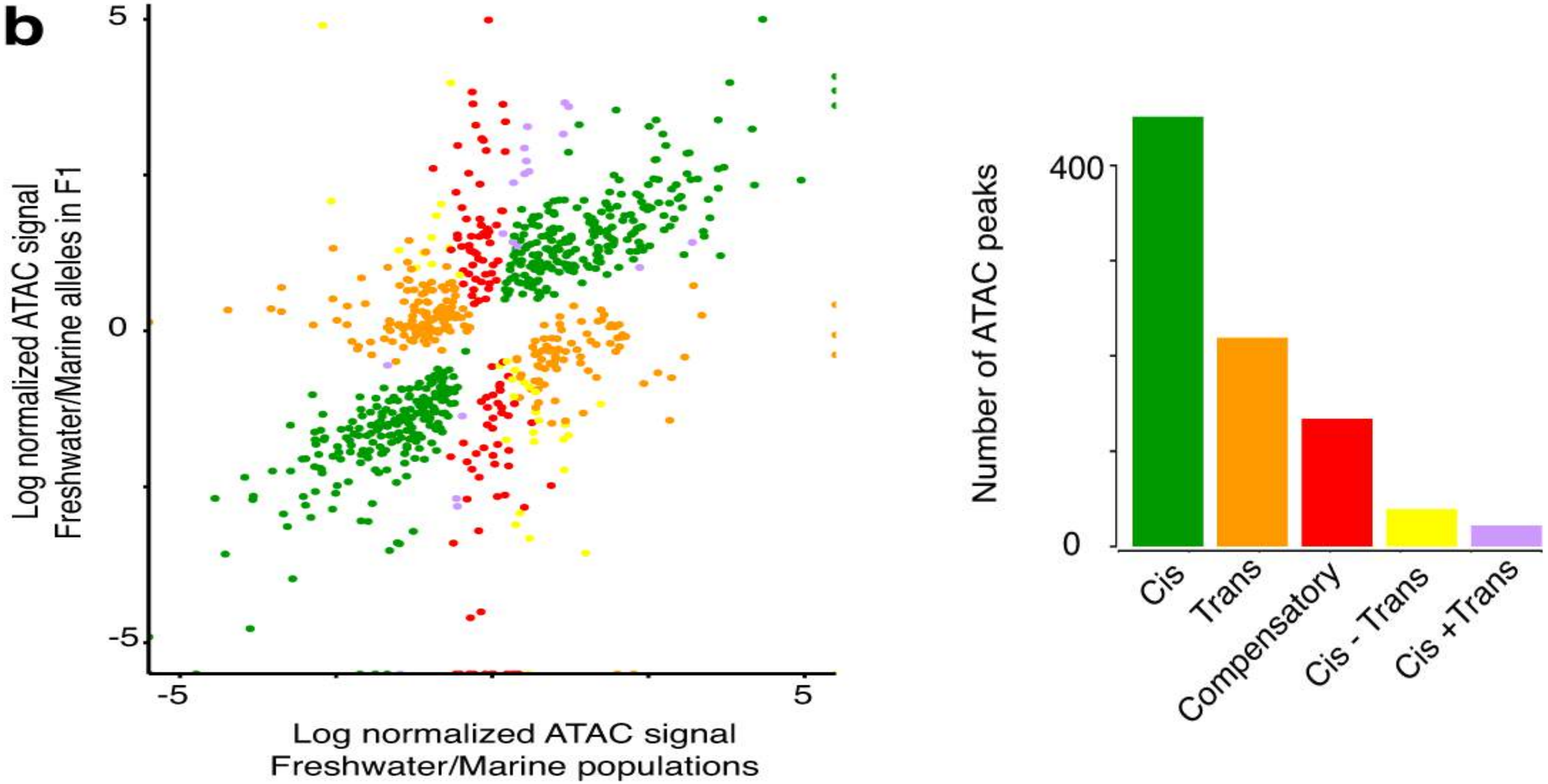


Allele-specific ATACseq to identify cis-control of open chromatin

Dr Stanley Neufeld



Allele-specific ATAC reveals that marine-freshwater divergence in chromatin state is mostly due to cis-regulatory elements



Molecular mechanisms of adaptation & speciation

1 The regulatory control of adaptive divergence in gene expression

- cis-regulation is predominant

large effect size, additive, stable…

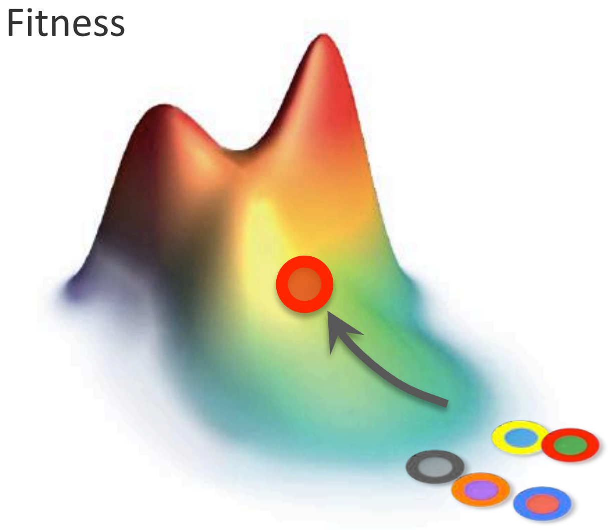
2 Comparative epigenomics reveals clusters of enhancer elements in adaptive loci.

- divergence in chromatin state is regulated in cis

Why should we care about *cis*-regulation in adaptation?

Recombination may be beneficial …. or deleterious and its modifiers subject to selection

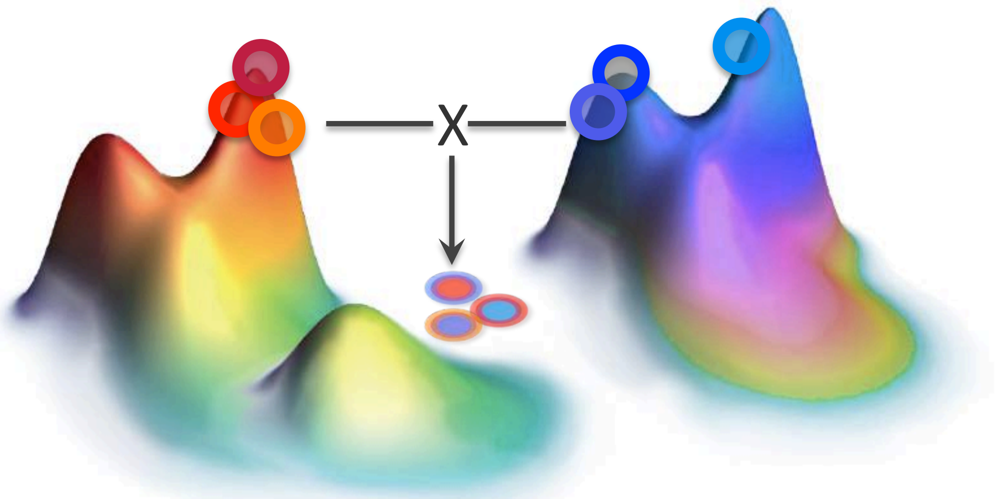
Beneficial recombination



Helps adaptation
Reduces background selection

Hill & Robertson (1966) Genetical Res. 8:269-294
Felsenstein (1974) Genetics 78:737-756

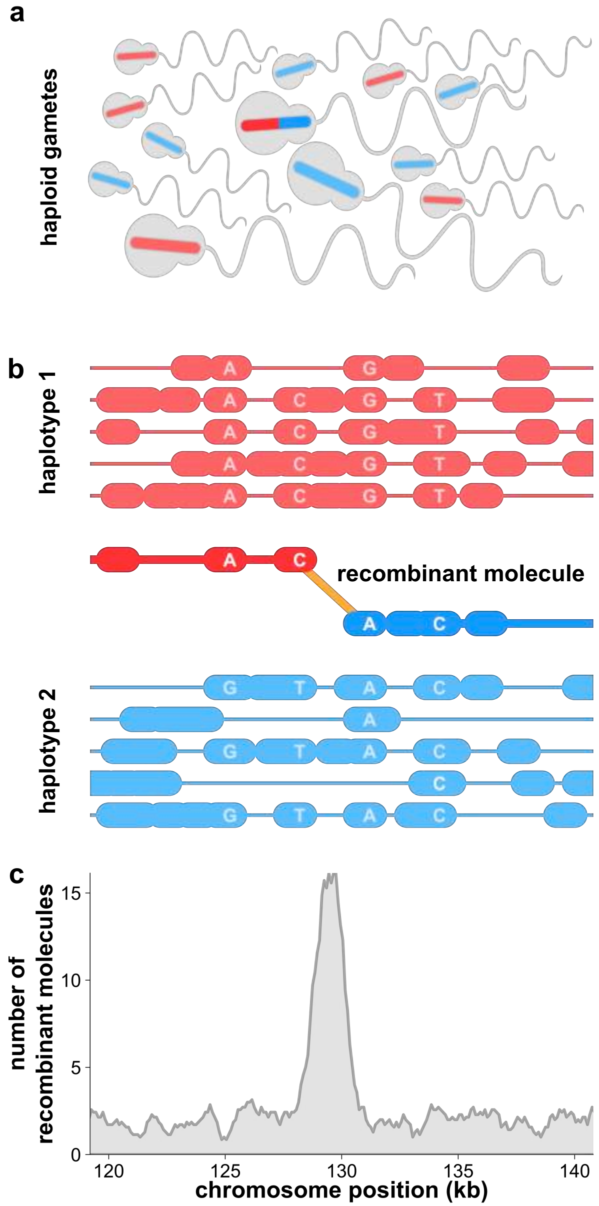
Deleterious recombination



Gene flow between divergently adapted ecotypes produces maladapted recombinants

Charlesworth & Charlesworth (1979) Genetics 91:581-589
Kirkpatrick & Barton (2006) Genetics 173:419-434

ReMIX: A novel pipeline to build high-resolution individualized cross over maps from gamete linked-read sequencing



Dr. Andreea Dreau

Dreau et al (in review)
BioRxiv
https://doi.org/10.1101/489989

Proof of principle: Can ReMIX detect known hotspots in mice? (*hlx*, a well known hotspot on mouse chr1)

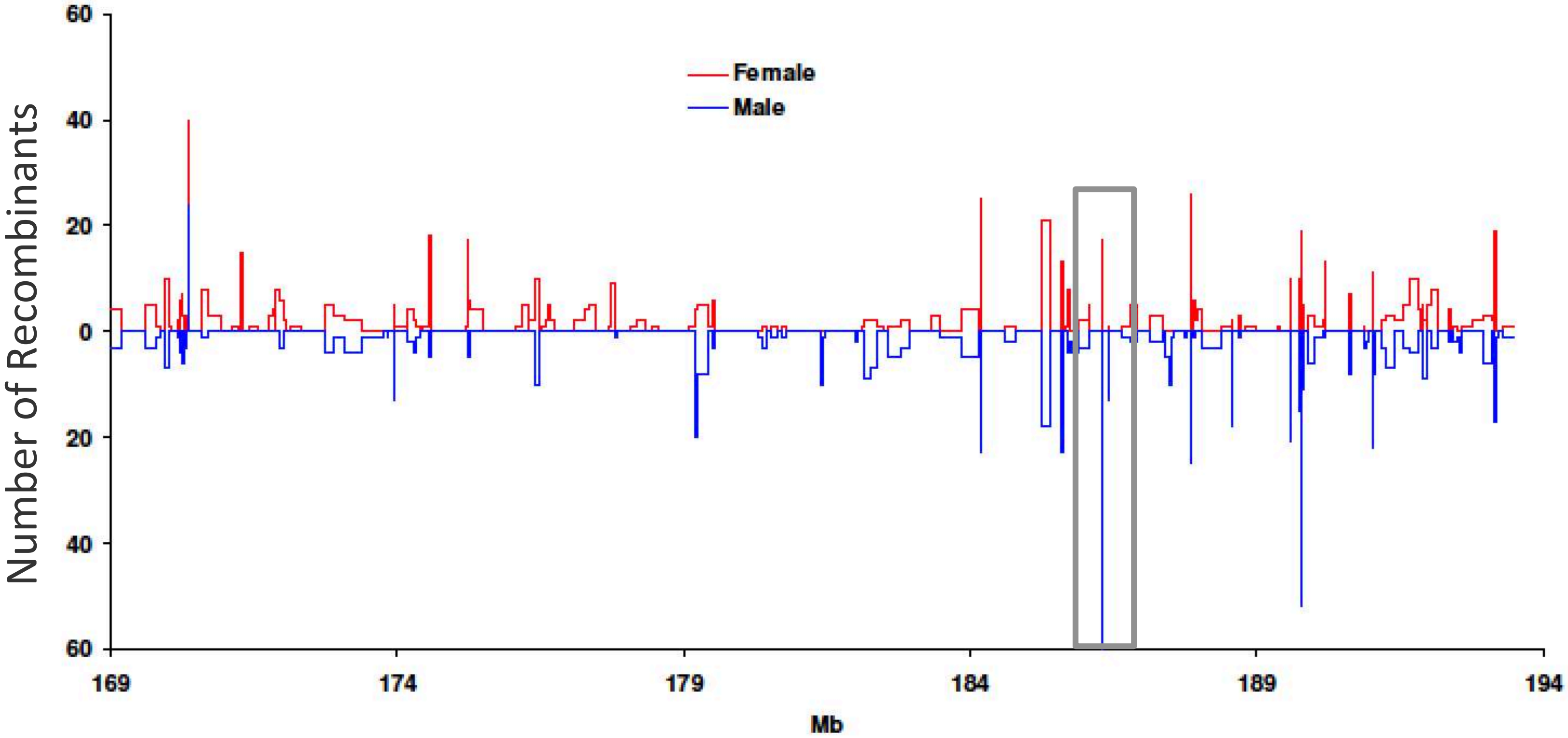


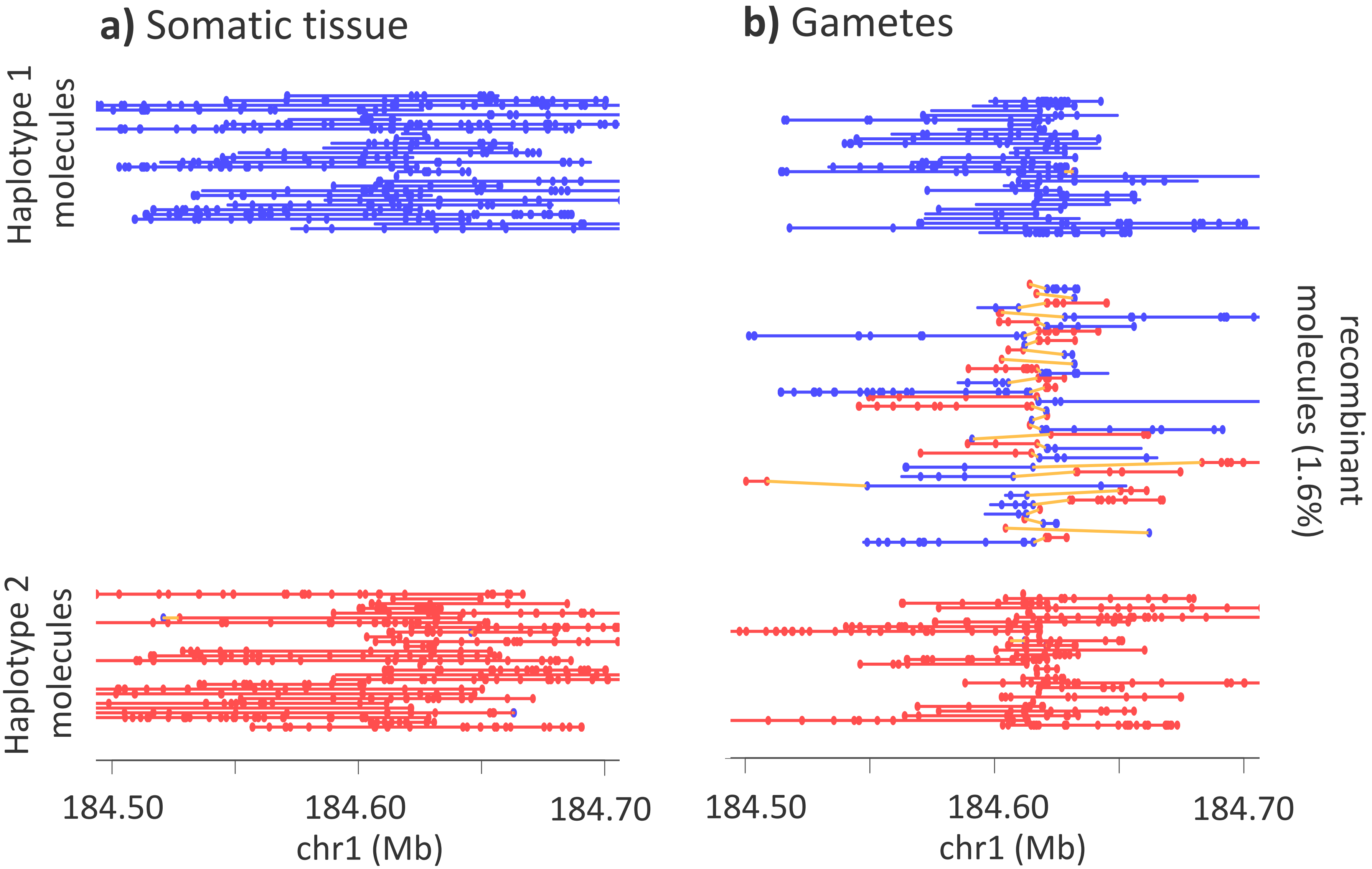
Fig S3. From Paigen et al (2008) PLoS Genetics

A strong male-specific recombination hotspot on mouse chr 1

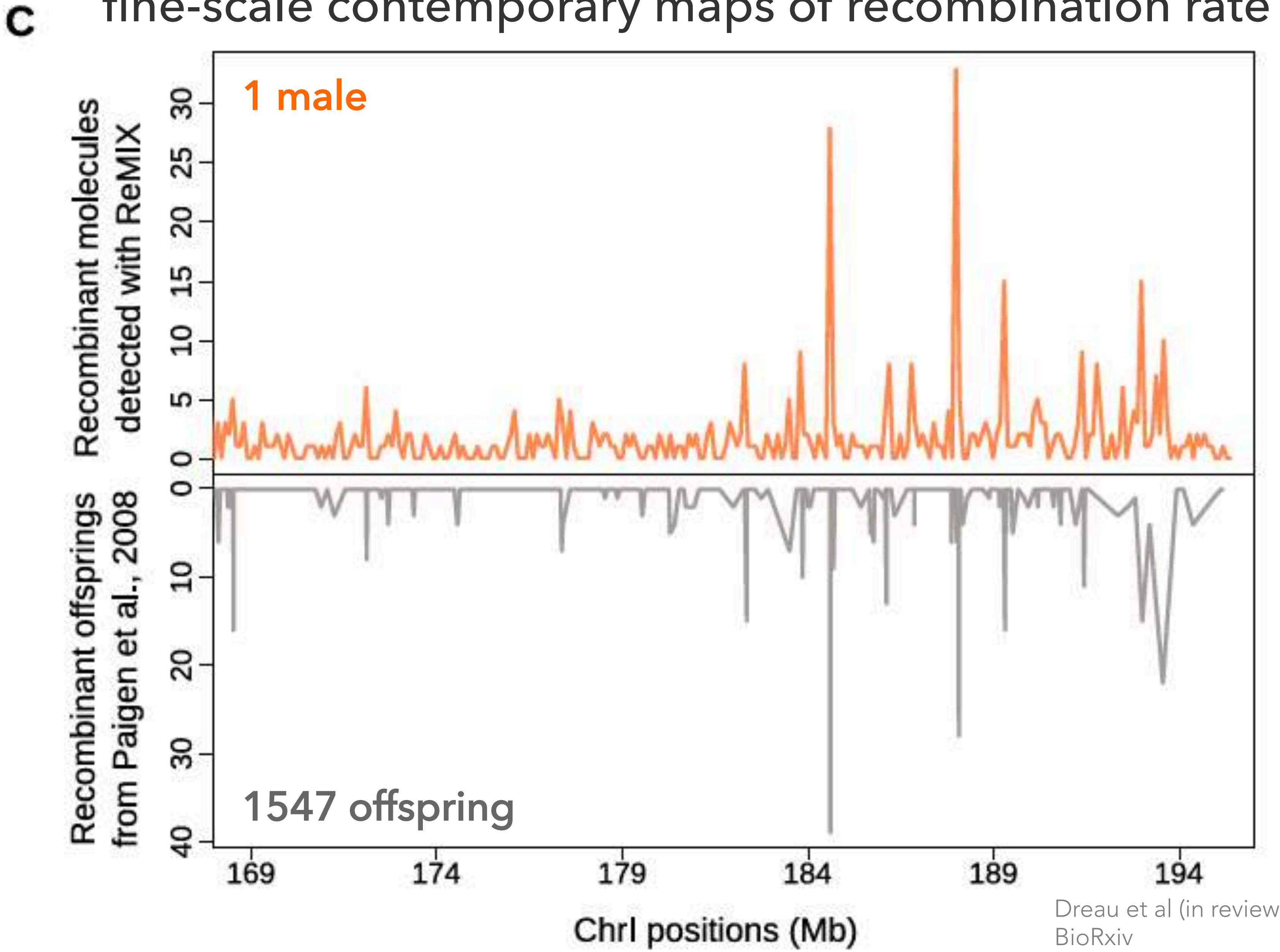
52 recombinant offspring out of 1547 total (3.3%)

Dreau et al (in review)
BioRxiv
https://doi.org/10.1101/489989

ReMIX successfully detects known recombination hotspots in mice

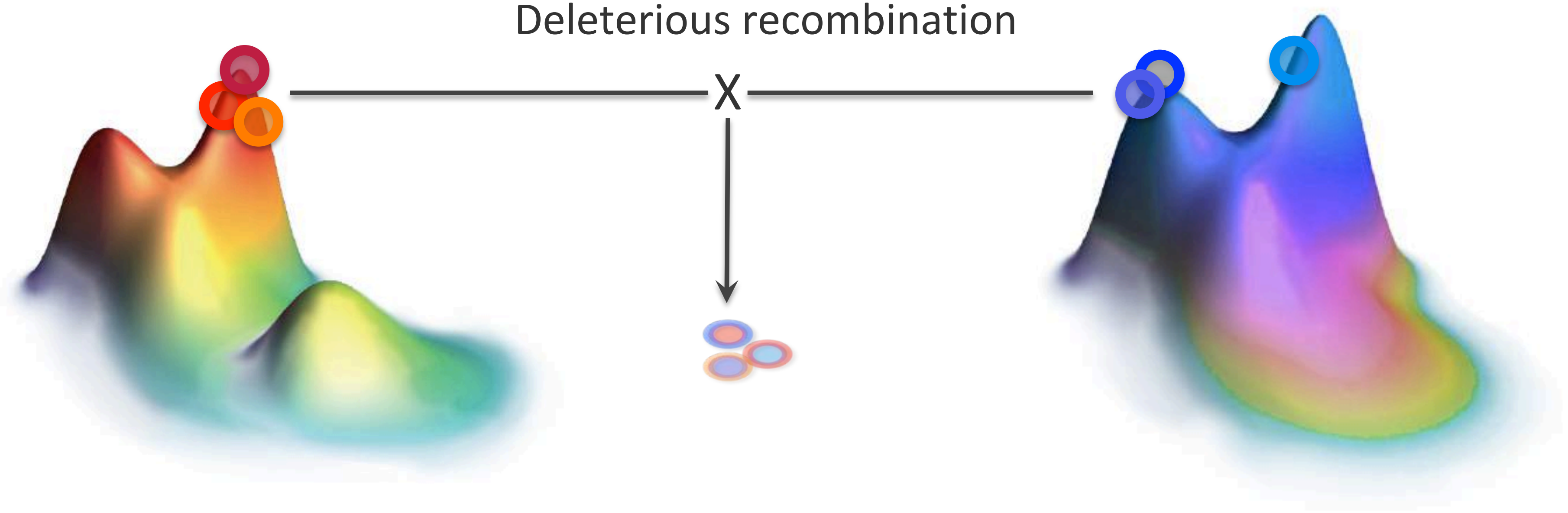


ReMIX: a fast and cost effective method to build fine-scale contemporary maps of recombination rate



Dreau et al (in review)
BioRxiv
https://doi.org/10.1101/489989

What is the molecular basis of recombination variation?
Is it subject to selection?



Deleterious recombination

Marine

Freshwater

What molecular mechanisms underlie adaptation and speciation?

Dr Jukka-Pekka Verta

Cis-regulation plays an important role in adaptive gene expression divergence

Dr Stanley Neufeld

Clusters of divergent enhancers drive adaptive divergence in sticklebacks

Dr. Andreea Dreau

Linked-read sequencing recombination hotspots and de novo diploid assemblies

Jones Lab, FML
Max Planck, Tuebingen

Deutsche Forschungsgemeinschaft

European Research Council Consolidator Grant

MAX-PLANCK-GESELLSCHAFT

Molecular mechanisms of adaptive divergence and speciation


Frank Chan