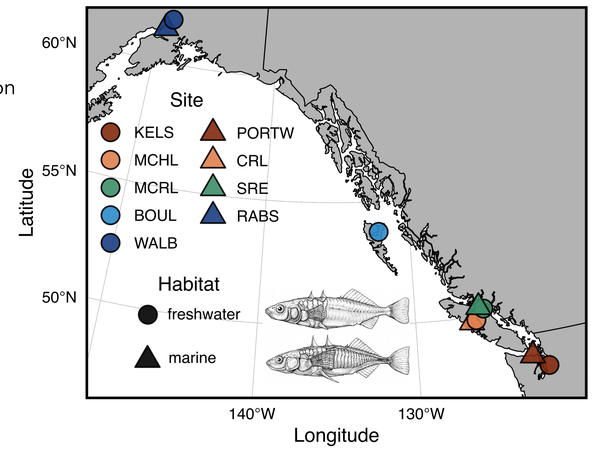
Cell-type-resolved regulatory evolution during marine-freshwater divergence in stickleback

Zachary Tobias[1], Benjamin Gastfriend[1], Matthew Farnitano[1], Martín Tresguerres[1,2], Richard Daneman[1] & Diana Rennison[1].

[1]University of California - San Diego, [2]Scripps Institution of Oceanography

Introduction

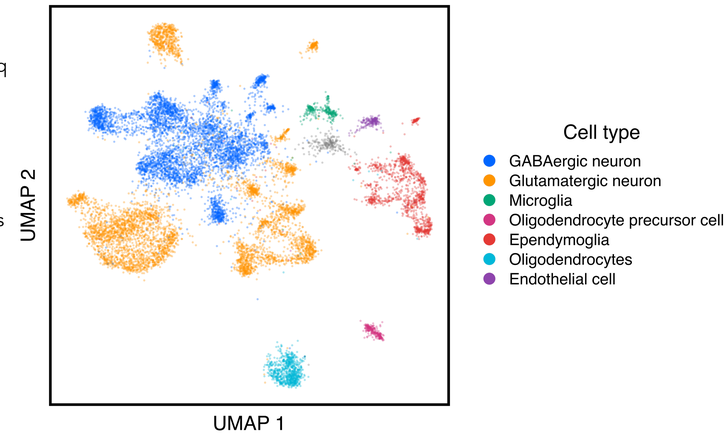
- The evolutionary importance of mutations in cis-regulatory elements (CREs) has been recognized since their discovery[1].
- The 'omics era has produced striking examples of CRE evolution in environmental adaptation, including recurrent deletion of a tissue-specific enhancer of the skeletal patterning gene *pitx1*, which underlies pelvic reduction in freshwater stickleback[2].
- Freshwater stickleback have undergone countless transitions from marine ancestors, involving parallel morphological shifts such as armor loss and spine reduction, as well as behavioral shifts in traits such as sociality and aggression[3,4].
- However, we have limited insight into the genetic basis and cellular context of mutations contributing to behavioral adaptation.
- Here, we examine whether repeated differentiation at loci associated with nervous-system CREs may underlie adaptive divergence.



Map of sampling locations. Fish illustrations from Bell, 1976.

Methods

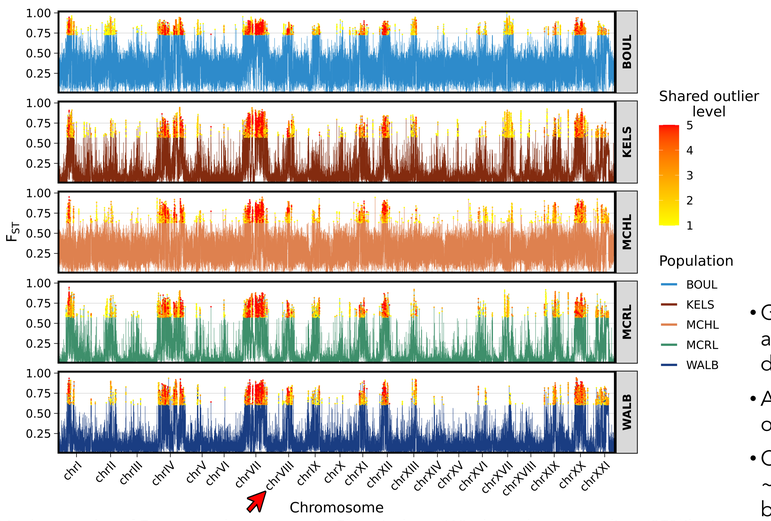
- Single-cell atlas construction
  - 10x Genomics Multiome paired single-nucleus RNAseq + ATACseq of pooled adult brains from lab-reared marine background.
  - SAMap[5] for transferring cell-type annotations from zebrafish.
  - Chromatin accessibility peaks called with MACS2[6] in Signac[7] and linked to expression of downstream genes with Seurat[8].
- Whole genome sequencing
  - Data generated by Rennison Lab or downloaded from
  - BWA[9], GATK3[10], bcftools[11] for mapping, calling, and filtering.
  - pixy[12] for windowed F_{ST} estimates of individual freshwater lakes against pooled marine.



UMAP of single nucleus expression profiles. Cell-type annotations from SAMap.

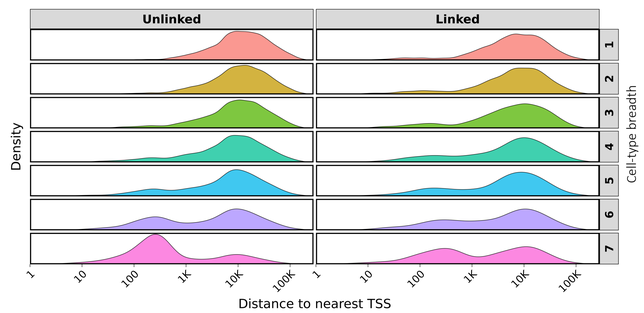
Results

Genome-wide signatures of parallel differentiation



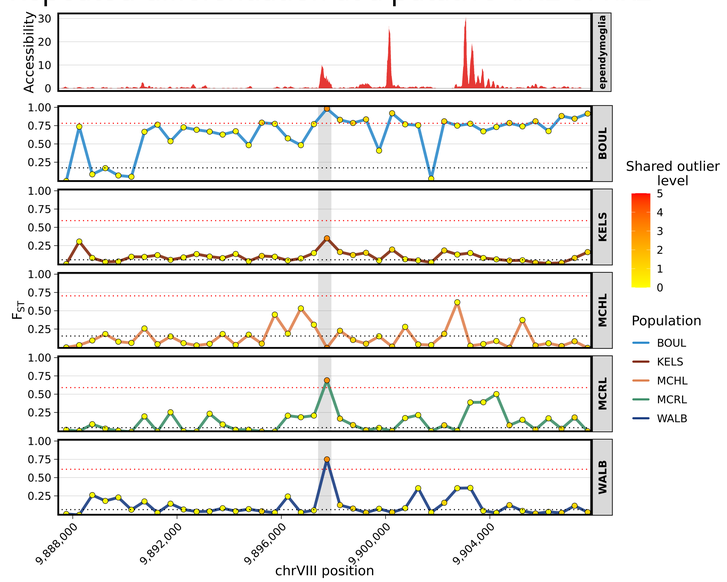
Manhattan plot of F_{ST} across the genome in 5kb windows. Windows that are outliers (95%ile) in at least one population are indicated with points, with increasing levels of parallelism indicated by warmer colors. Differentiation is concentrated with the canonical stickleback "adaptation hotspot" chromosomes (IV, VII, XVI, XX, and XXI; from Peichel & Marques, 2017[13]) but is also distributed elsewhere. Red arrow indicates chromosome VIII (see far right panel).

Cell-type-specific ATAC peaks are enriched for distal elements



- Genome-wide patterns of differentiation reveal a large proportion of shared outlier windows. These are concentrated on chromosomes enriched for QTL for traits involved in marine-freshwater divergence[13], but also appear elsewhere.
- Accessible chromatin (as identified by ATACseq) tends to be located further from nearby TSS when open in fewer cell types, likely representing putative distal CREs.
- On chrVIII, a ATAC peak unique to ependymoglia is associated with expression of the gene *rfx2*. This ~511 bp region overlaps a highly parallel outlier window with strong elevation above local F_{ST} background.
- Ependymal cells have been demonstrated to be involved adult neuroregeneration in early-diverging vertebrates[14,15].
- *rfx2* promotes ciliogenesis in ependymal cells[16], which is thought to mediate regenerative processes[17].

Repeated differentiation at a putative CRE of rfx2



Top track represents chromatin accessibility in ependymoglia. Bottom five tracks are F_{ST} of 500bp windows. Black dotted lines are genome-wide median, red dotted lines are 95%ile. Grey rectangle indicates bound of called ATAC peak. Point color indicates degree of parallelism.

Next Steps

- This pilot study demonstrates that we can identify repeated differentiation at putative CREs in individual nervous system cell types.
- However, disentangling direct selection on putative CREs from linked selection to genes or other features remains a challenge.
- We plan to phase the genomic data using our existing reference panel. This will enable haplotype-aware methods to more precisely identify selection on regulatory elements, but also allow us to ask if outlier regions have a shared ancestral origin or are mediated by *de novo* mutations.
- We will formally test if parallel freshwater adaptation has preferentially targeted regulatory regions relative to baseline.
- Incorporate other lines of evidence (e.g. ATAC footprinting, TF motif analysis) to more definitively identify CREs.

Acknowledgements

- Joshua Dominguez
- David Ferranti
- Andreas Härer
- Thomas Reimchen
- Troy Weldon
- Caitlin Wise
- NOAA Northwest Fisheries Science Center

Refs

Website

UC San Diego | Scripps Institution of Oceanography | Allen Discovery Center for Neurobiology in Changing Environments | NIH | The Kavli Foundation | Allen Family Philanthropies