# 

### **Motivation**

Identifying regions with high amounts of consistent differences across populations has proven useful for identifying evolutionary divergence. Studies on drosophila melanogaster and marine-freshwater stickleback fish have been performed measuring such consistent differences across groups, but none have to my knowledge been made available online.

I implemented two algorithms for such measurements as tools in the Genomic HyperBrowser.

## Methods

Two algorithms measuring genomic divergence have been implemented:

- **Cluster separation scoring** of sliding windows based on two-dimensional scaling of the pairwise differences between individuals in the groups
- Performing *Fisher's exact test* on 2x2 contingency table on each SNP in a sliding window of the genome

The algorithms are visualized below. For the cluster separation scores significance is measured through random permutation testing. Scores for windows

based on Fisher's exact score is meanwhile found by taking the upper 5% treshold of all scores obtained for the window.

#### Implemented as Galaxy tools

Calculate cluster separation score						
Genome build: Stickleback	tab	Minor	Major		Minor	Major
Group 1 SNPs: Freshwater						
Group 2 SNPs: Marine	ge	6	5	Freshwater	11	0
Pairwise individual comparison metric: Count individual SNP-differences in window	) ti					
Sliding window size:		5	5	Marine	0	10
2500	× Z					
sliding window step:						
500						
/inimum significance score runs:						
10						
Aaximal significance scores:			V			V
20000						
			G			G G
Dutput format: Tabular ▼		9	G			G
Execute	ata vate					Ğ
						G G
	es mi		G			G





## Results

The figure shows the cluster separation scores and corresponding p-values for the EDA locus in chromosome IV of the stickleback genome, known for marine-freshwater divergence.

With the correct parameter choices the results from my tools reproduce earlier published results.

## **References and Acknowledgements**

Stickleback study which was reproduced and methods figure was adapted from is published in "The genomic basis of adaptive evolution in threespine sticklebacks". Jones et al.Nature 484.7392 (Apr. 2012), pp. 55–61.

The work presented is the result of a master's thesis supervised by Geir-Kjetil Sandve with input from Lex Nederbragt and Bastiaan Star. Poster designed by Eskil Frøyen Nybø.





**Torkil Vederhus**