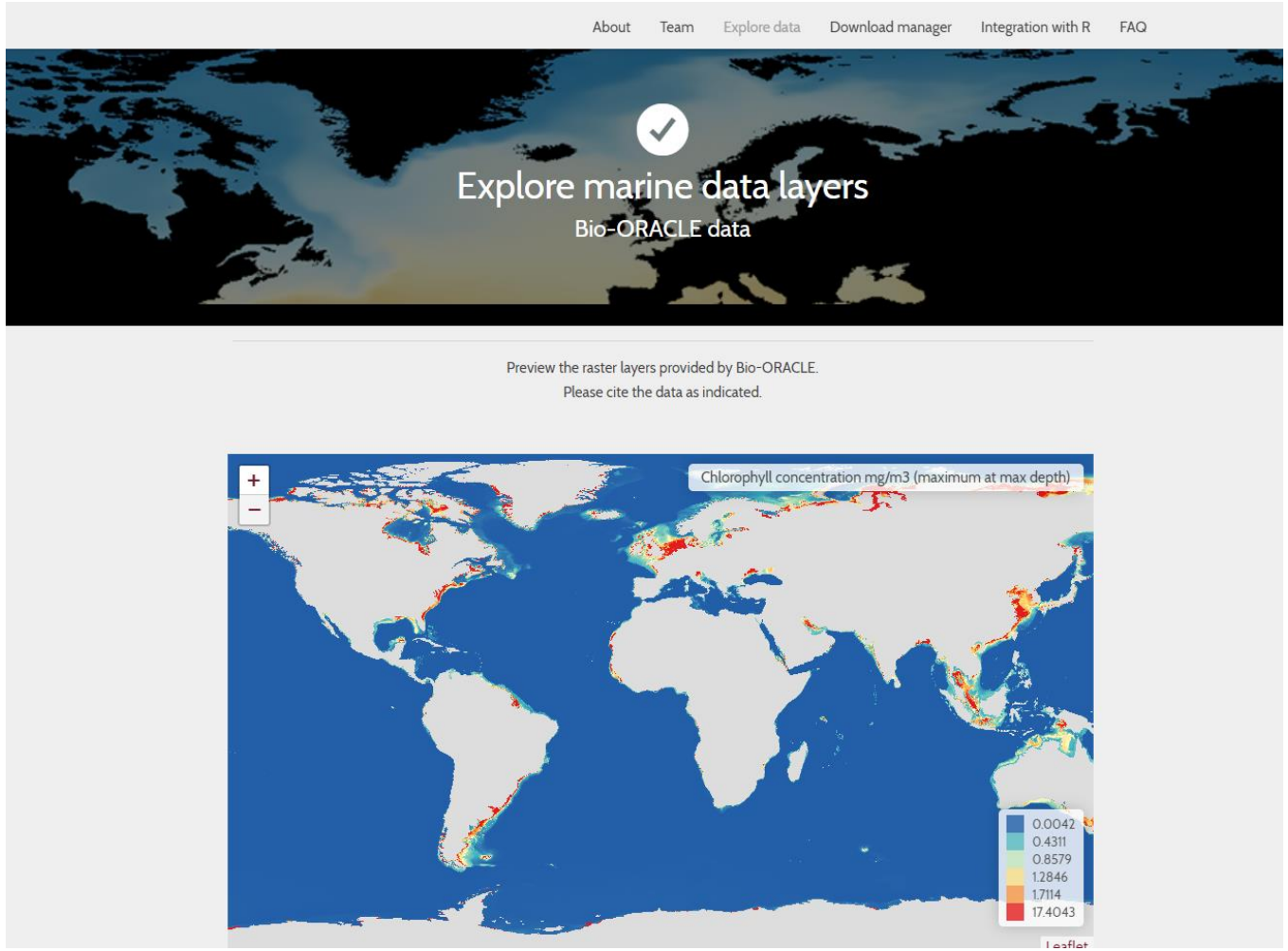
Genome-Environment Associations

November 1st 2021

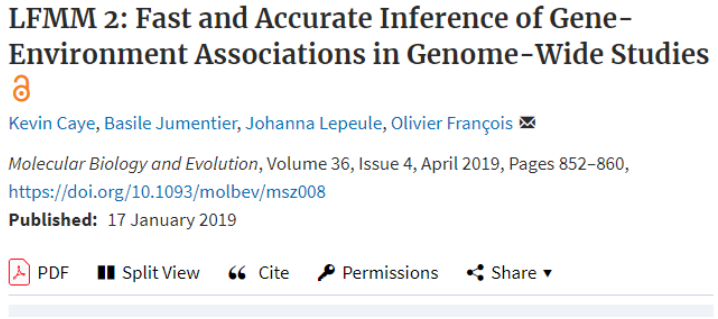
Cal State LA

Environmental data



https://www.bio-oracle.org/explore-data.php

GEA analysis



LFMM 2: Fast and Accurate Inference of Gene-Environment Associations in Genome-Wide Studies

Kevin Caye, Basile Jumentier, Johanna Lepeule, Olivier François

Molecular Biology and Evolution, Volume 36, Issue 4, April 2019, Pages 852–860, https://doi.org/10.1093/molbev/msz008

Published: 17 January 2019

PDF Split View Cite Permissions Share

Correlation between environmental data and genomic data

We'll be using the R package lfmm (latent factor mixed model)

Basically, a regression model with fixed and latent effects

$$Y=XB^T+W+E$$

Response variable (genomic variation)

Environmental variable

Fixed effects in matrix B

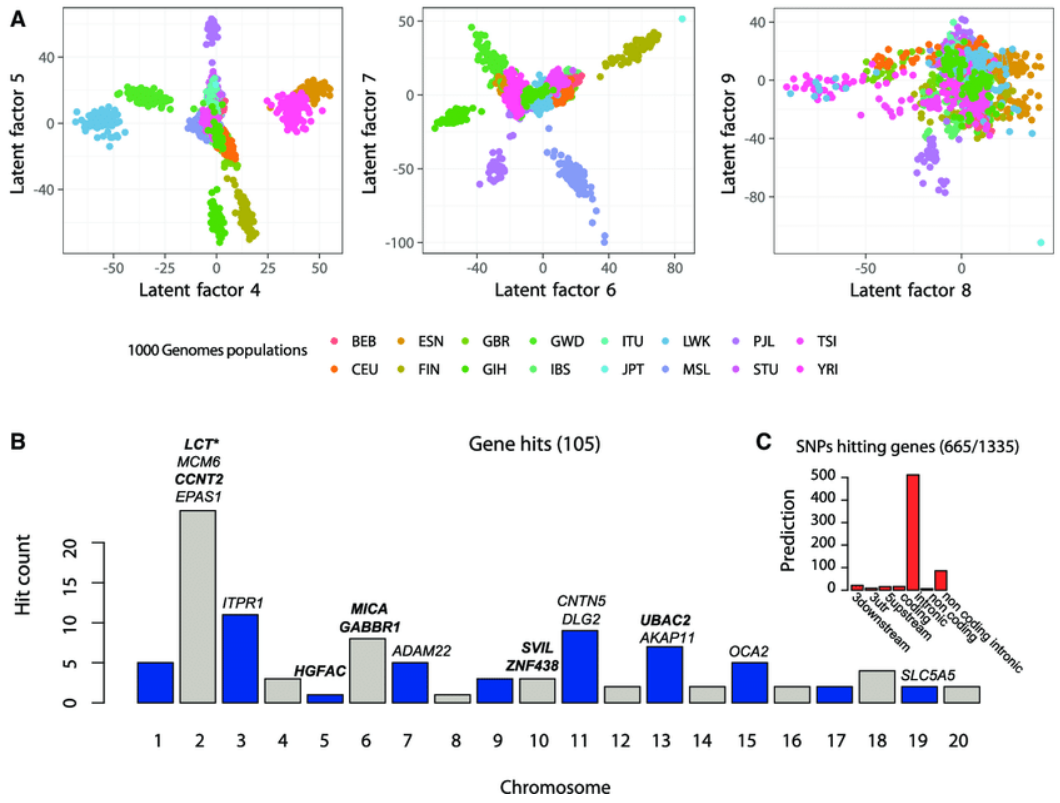
Latent effects in matrix W

Residual errors in E

https://academic.oup.com/mbe/article/36/4/852/5290100#138397544

Examples of genome-environment associations

Humans and climate



Caye et al. 2019 https://academic.oup.com/mbe/article/36/4/852/5290100#138397544

Population structure

Again this is a problem!

Population structure confounds GEA

Data for class today

3966 SNPs from 685 pacific sea cucumbers

ORIGINAL ARTICLE

WILEY MOLECULAR ECOLOGY

Asymmetric oceanographic processes mediate connectivity and population genetic structure, as revealed by RADseq, in a highly dispersive marine invertebrate (*Parastichopus californicus*)

Amanda Xuereb[1] | Laura Benestan[2] | Éric Normandeau[2] | Rémi M. Daigle[1] | Janelle M. R. Curtis[3] | Louis Bernatchez[2] | Marie-Josée Fortin[1]

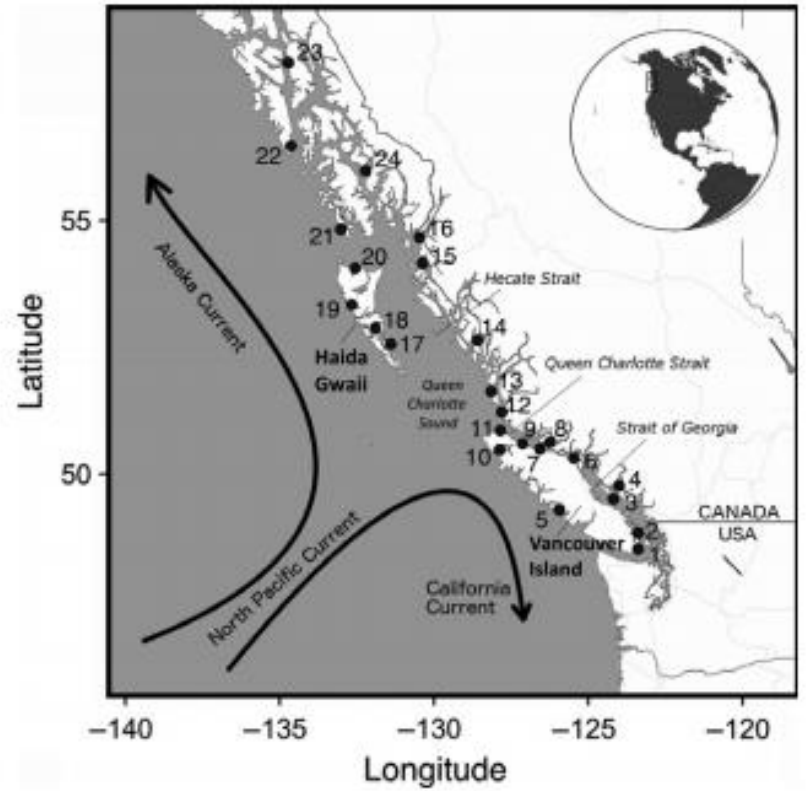


FIGURE 1 Map of sampling locations in coastal British Columbia (1–20) and southeastern Alaska (21–24). Site labels correspond with numbers in Table 1