Sea Cucumbers & Structure Plots

Eve198

Week 7

Maddie Armstrong & Rachael Bay

Received: 17 December 2017 | Revised: 13 March 2018 | Accepted: 19 March 2018
DOI: 10.1111/mec.14589

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*)

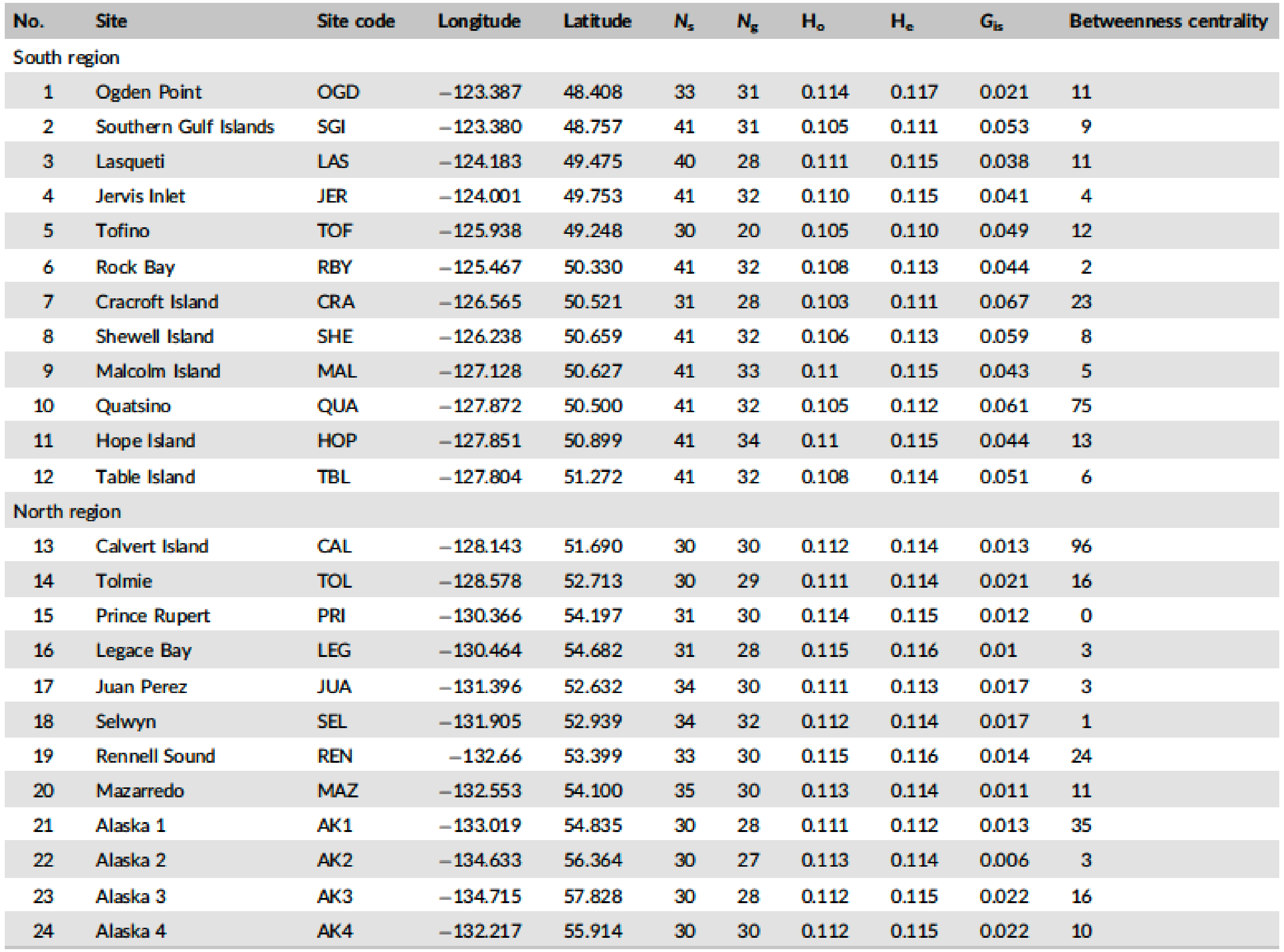
Amanda Xuereb[1] | Laura Benestan[2] | Éric Normandeau[2] | Rémi M. Daigle[1] |
Janelle M. R. Curtis[3] | Louis Bernatchez[2] | Marie-Josée Fortin[1]

© California Sea Grant

© Joel Sartore

N=30-41 individuals per site, 24 sites total
N=717 individuals total

TABLE 1 Geographic coordinates, number of samples collected (N_s) and successfully genotyped (N_g), observed and expected heterozygosity (H_o and H_e, respectively), the measured inbreeding coefficient (G_{IS}), and the betweenness centrality calculated from the Population Graph of genetic covariance across sampling locations



No.	Site	Site code	Longitude	Latitude	N_s	N_g	H_o	H_e	G_{IS}	Betweenness centrality
South region										
1	Ogden Point	OGD	−123.387	48.408	33	31	0.114	0.117	0.021	11
2	Southern Gulf Islands	SGI	−123.380	48.757	41	31	0.105	0.111	0.053	9
3	Lasqueti	LAS	−124.183	49.475	40	28	0.111	0.115	0.038	11
4	Jervis Inlet	JER	−124.001	49.753	41	32	0.110	0.115	0.041	4
5	Tofino	TOF	−125.938	49.248	30	20	0.105	0.110	0.049	12
6	Rock Bay	RBY	−125.467	50.330	41	32	0.108	0.113	0.044	2
7	Cracroft Island	CRA	−126.565	50.521	31	28	0.103	0.111	0.067	23
8	Shewell Island	SHE	−126.238	50.659	41	32	0.106	0.113	0.059	8
9	Malcolm Island	MAL	−127.128	50.627	41	33	0.11	0.115	0.043	5
10	Quatsino	QUA	−127.872	50.500	41	32	0.105	0.112	0.061	75
11	Hope Island	HOP	−127.851	50.899	41	34	0.11	0.115	0.044	13
12	Table Island	TBL	−127.804	51.272	41	32	0.108	0.114	0.051	6
North region										
13	Calvert Island	CAL	−128.143	51.690	30	30	0.112	0.114	0.013	96
14	Tolmie	TOL	−128.578	52.713	30	29	0.111	0.114	0.021	16
15	Prince Rupert	PRI	−130.366	54.197	31	30	0.114	0.115	0.012	0
16	Legace Bay	LEG	−130.464	54.682	31	28	0.115	0.116	0.01	3
17	Juan Perez	JUA	−131.396	52.632	34	30	0.111	0.113	0.017	3
18	Selwyn	SEL	−131.905	52.939	34	32	0.112	0.114	0.017	1
19	Rennell Sound	REN	−132.66	53.399	33	30	0.115	0.116	0.014	24
20	Mazarredo	MAZ	−132.553	54.100	35	30	0.113	0.114	0.011	11
21	Alaska 1	AK1	−133.019	54.835	30	28	0.111	0.112	0.013	35
22	Alaska 2	AK2	−134.633	56.364	30	27	0.113	0.114	0.006	3
23	Alaska 3	AK3	−134.715	57.828	30	28	0.112	0.115	0.022	16
24	Alaska 4	AK4	−132.217	55.914	30	30	0.112	0.115	0.022	10

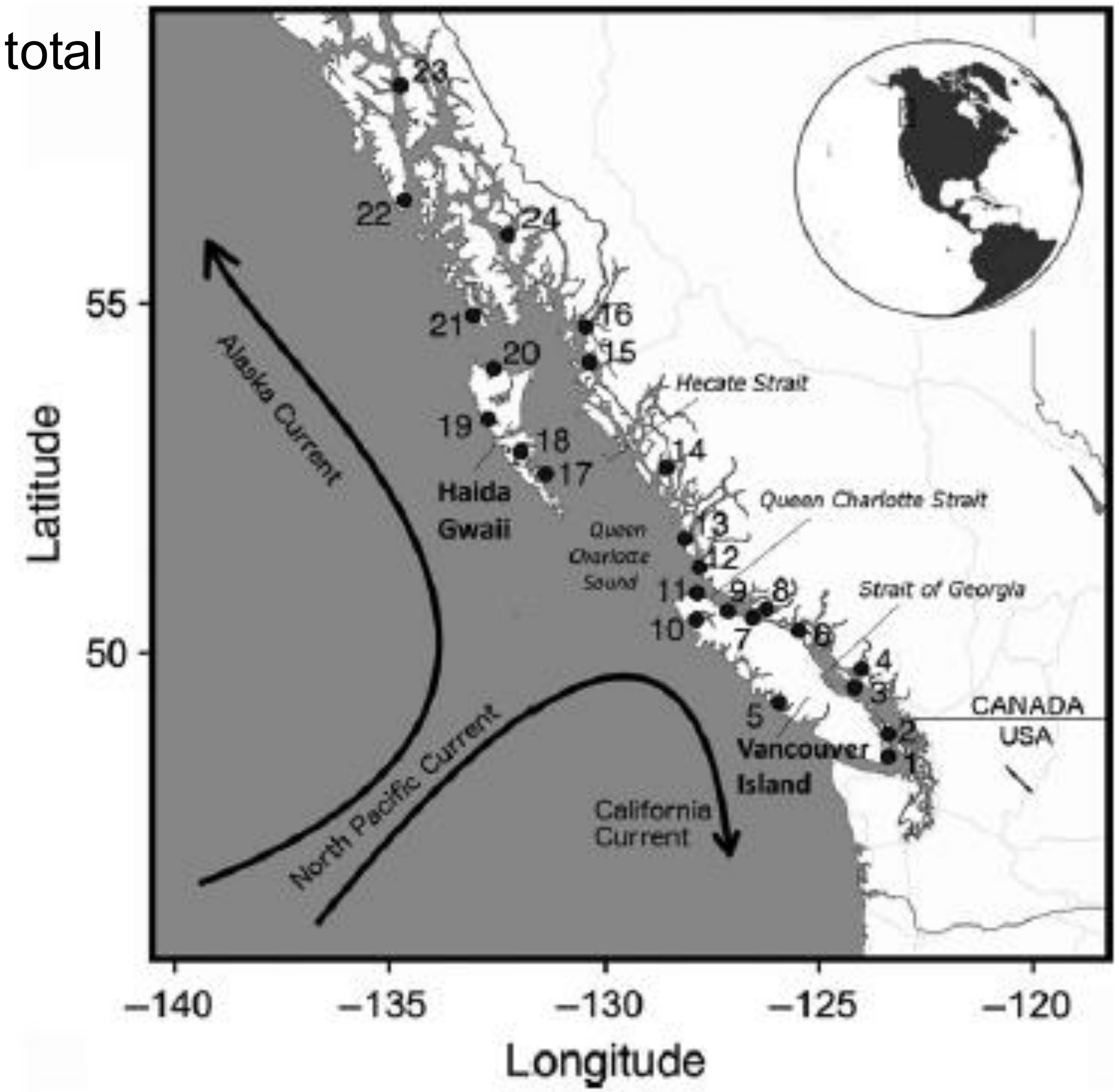


FIGURE 1 Map of sampling locations in coastal British Columbia (1–20) and southeastern Alaska (21–24). Site labels correspond with numbers in Table 1

Tested two hypotheses for spatial structure:

- Isolation by Distance (IBD) aka structure driven by distance
- Isolation by Resistance (IBR) aka structure driven by ocean circulation

Bioinformatic Pipeline:

Extracted DNA from tissue clips

RADseq via single end sequencing

Fastqc

Adapters trimmed with cutadapt

https://marineomics.github.io/RADseq.html

libraries demultiplexed with STACKS

Aligned to reference genome of closely related sea cucumber species

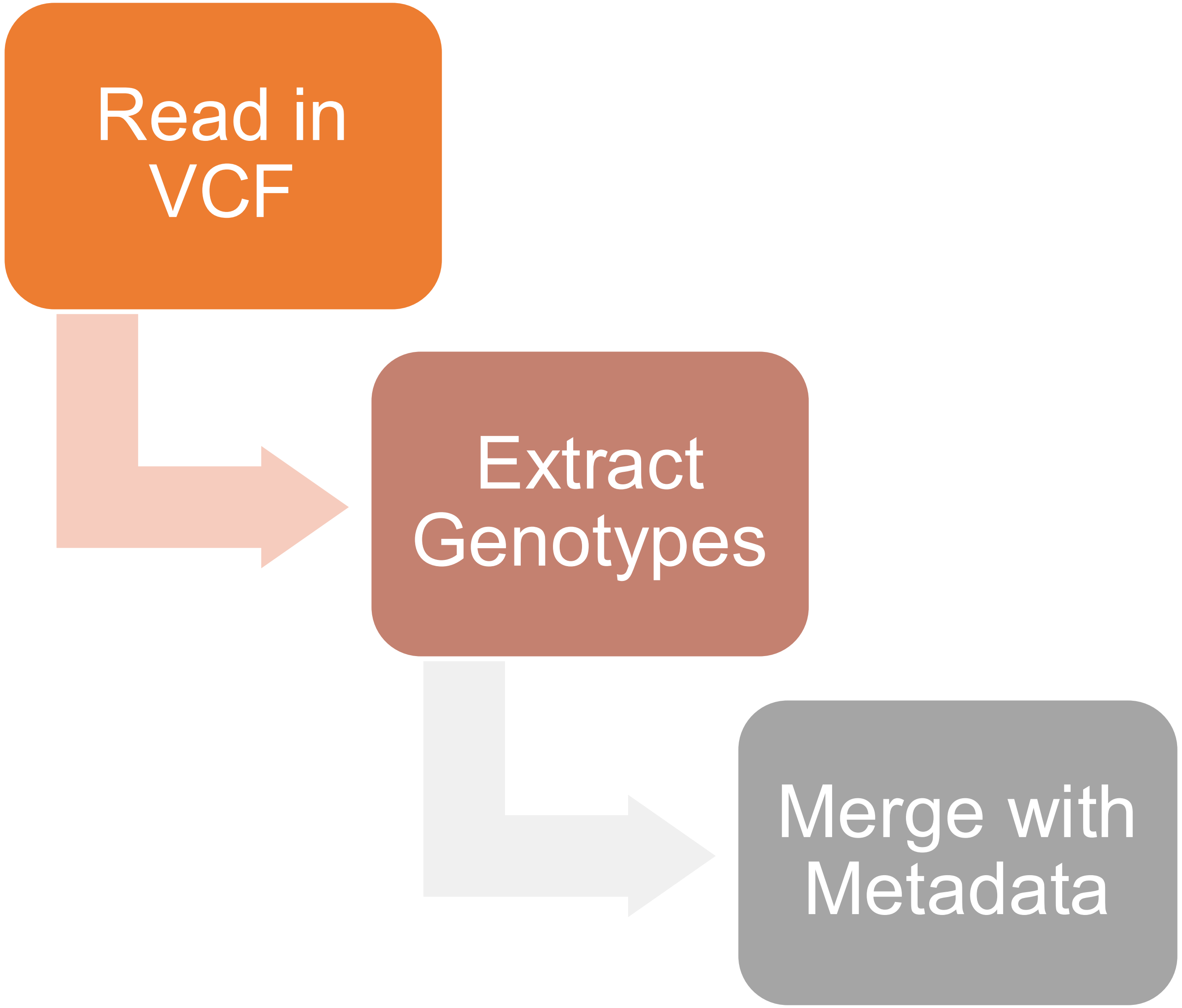
Filtered data to include only SNPs shared between regions/individuals

Variant Call Format (VCF) file to interpret genetic data

converted VCF file with PLINK and PGDSPIDER for additional analyses

© Joel Sartore

How to get the data into a friendly format for population structure analyses:



Read in VCF

Extract Genotypes

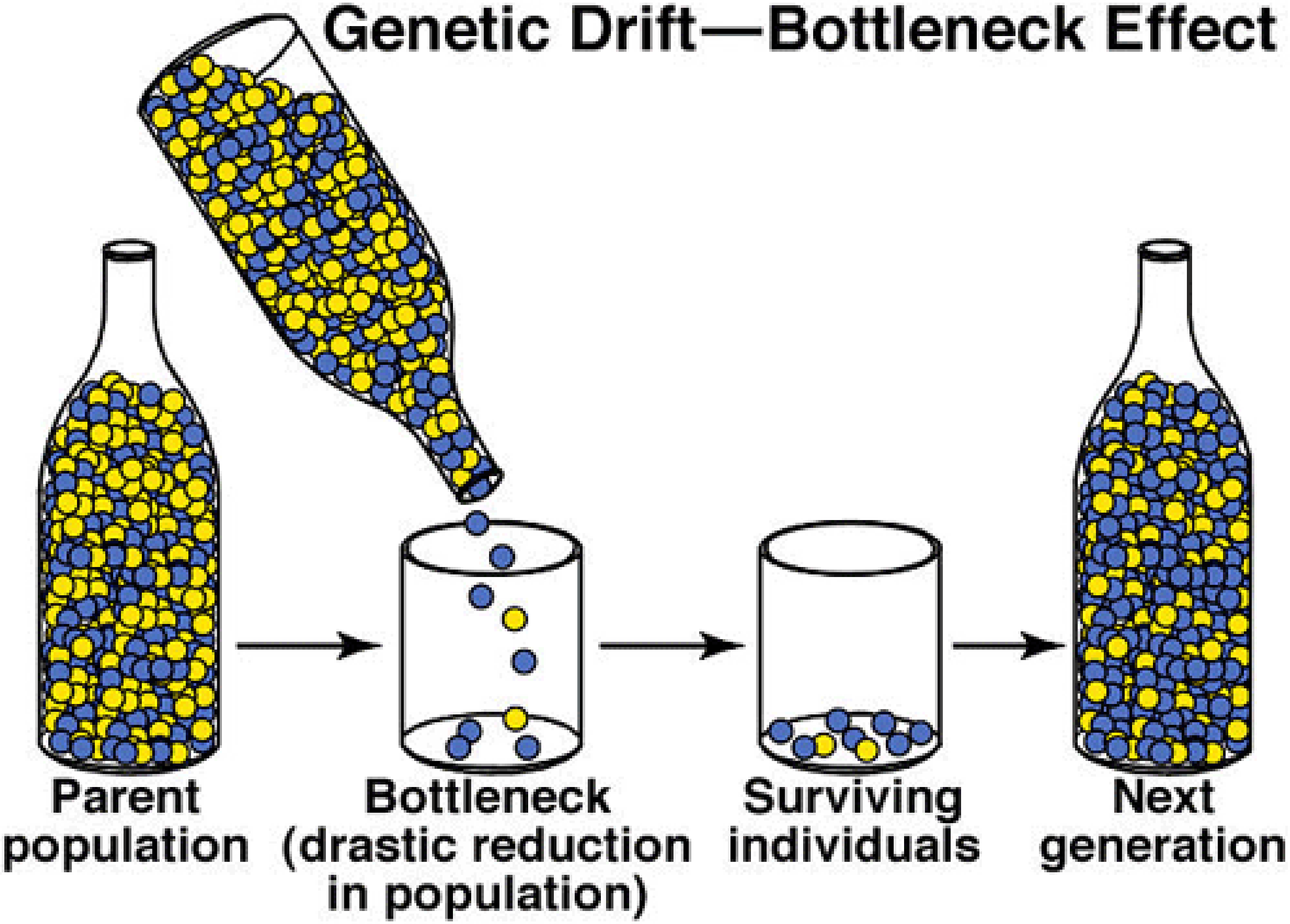
Merge with Metadata

What is population structure?

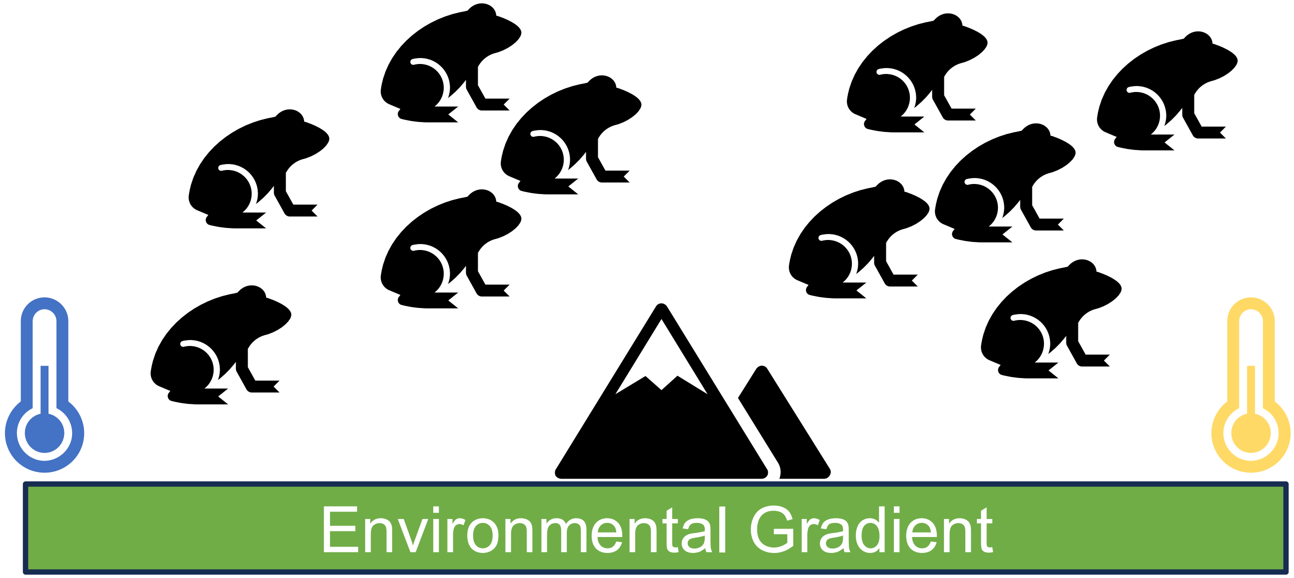
- Differences in allele/genotype frequencies between populations
- Genotypes share ancestry from some number (K) of ancestral populations
- No structure=little or no variation between populations
- What group are you= q-matrix
- Useful in identifying differences in ancestry between populations

http://galleries.neaq.org/2016/06/sea-cucumbers-east-meets-west.html

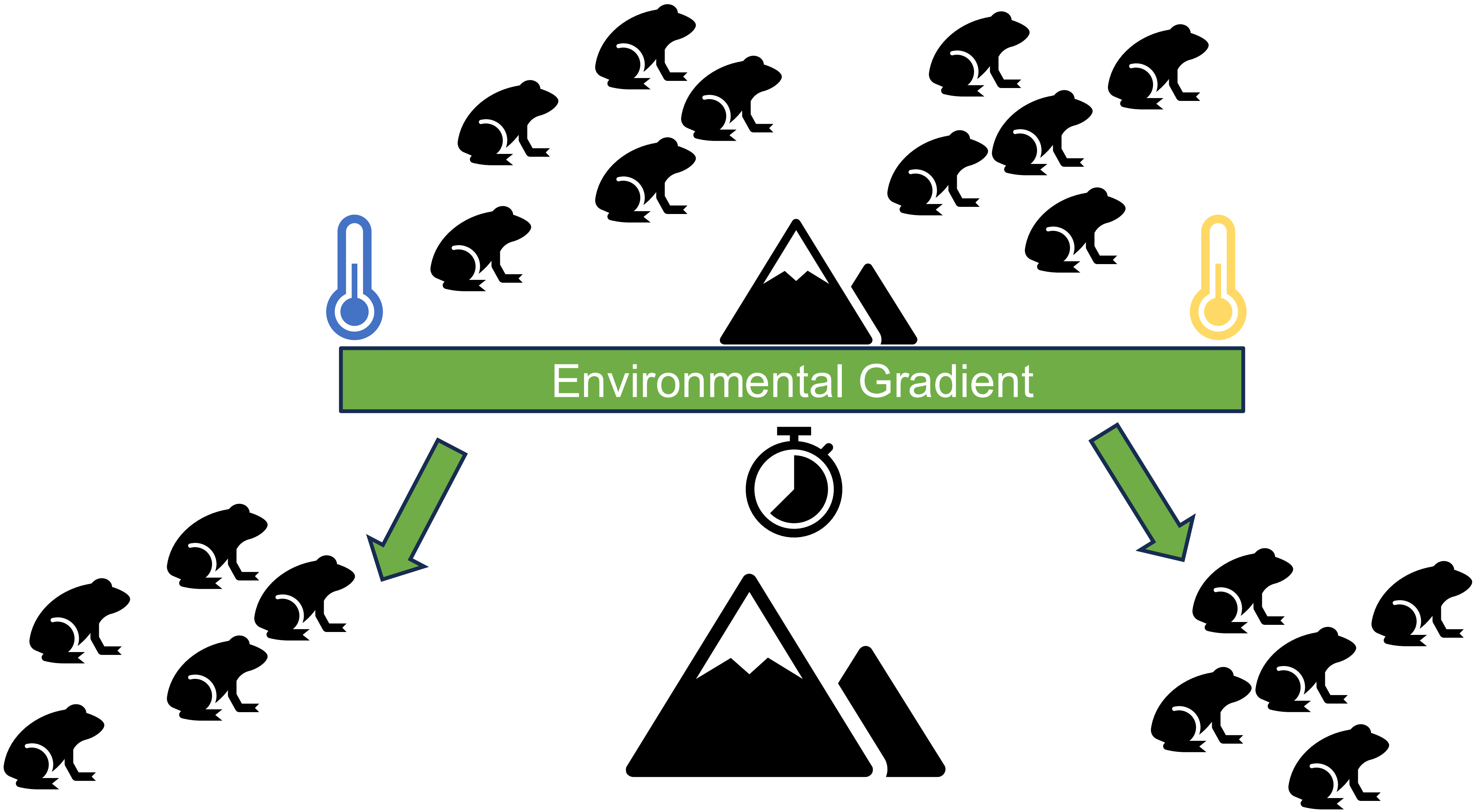
What can result in population structure?



What can result in population structure:

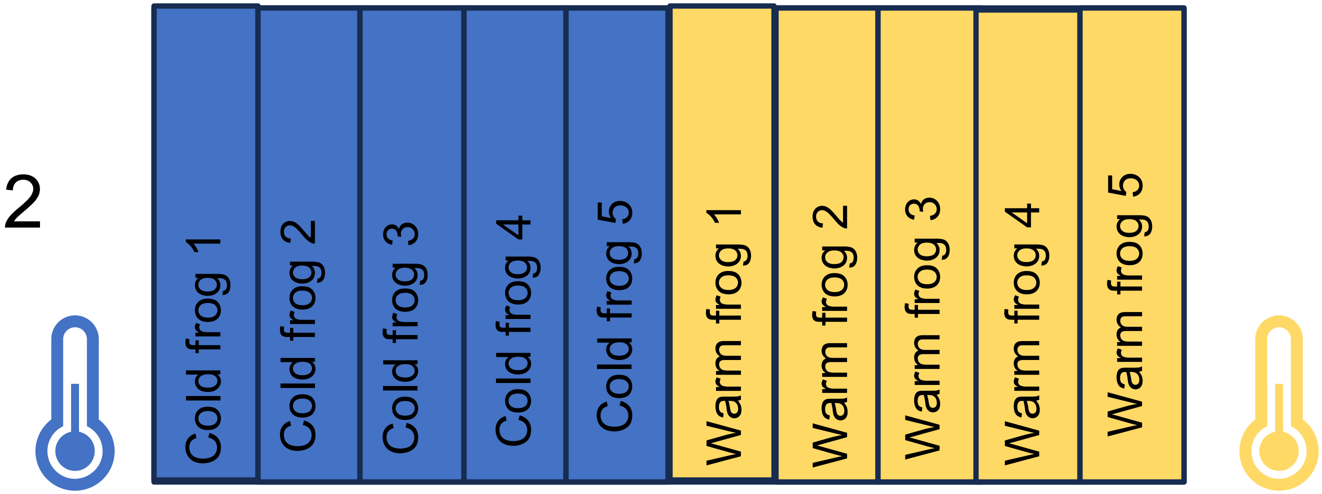


What can result in population structure:



What can result in population structure:

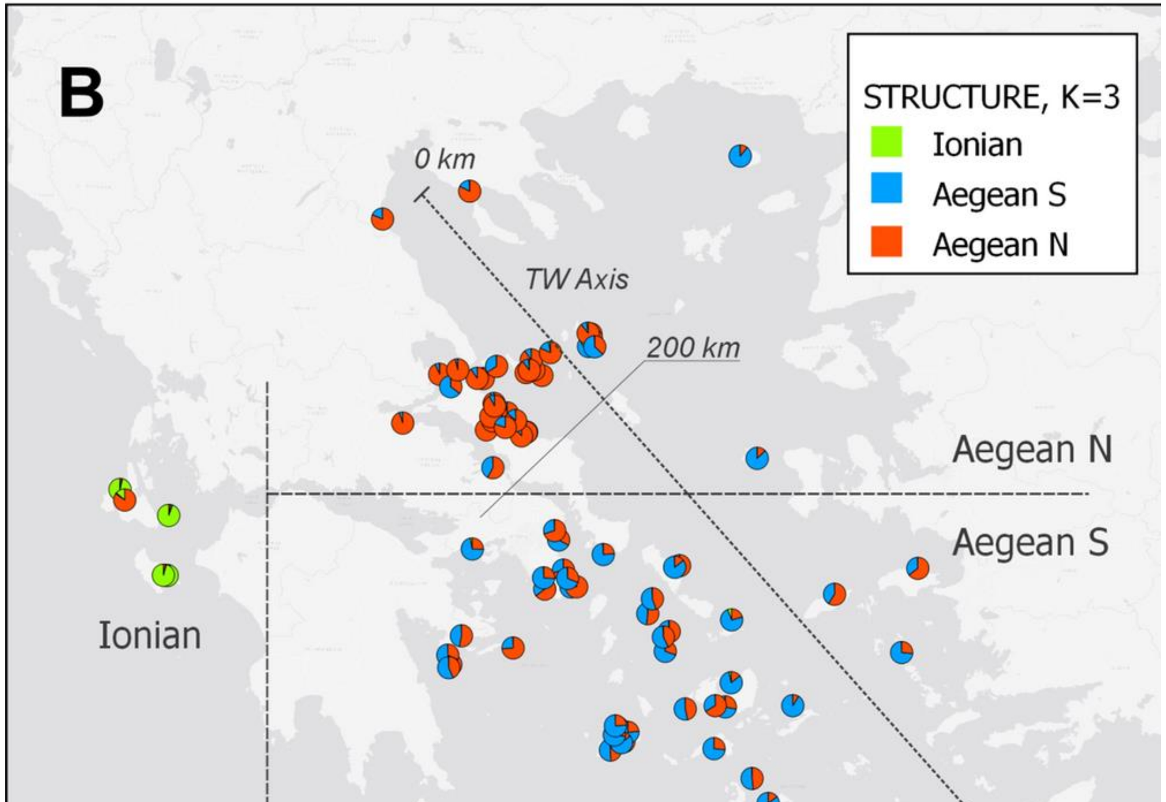
K=2

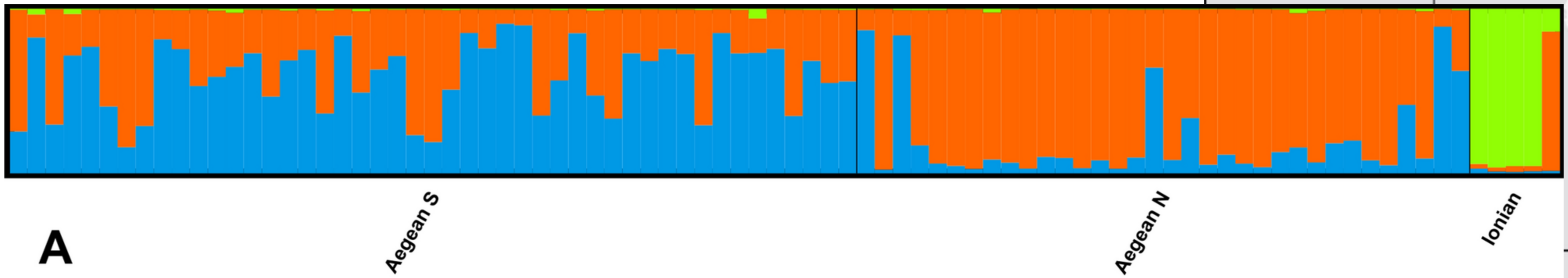


Use allele frequencies to assign likely membership for each individual to each population

What can result in population structure:

K=3



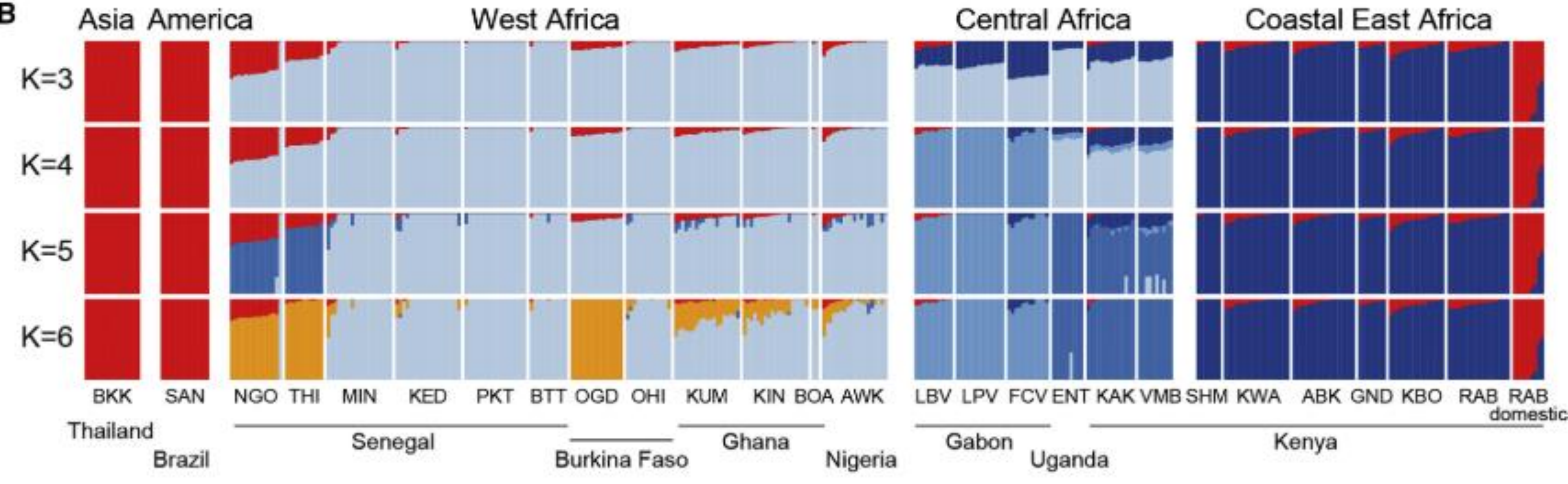


Karamanlidis et al. (2021) https://www.nature.com/articles/s41598-020-79712-1?fromPaywallRec=false


© Wikipedia

What can population structure tell us?

- Modeling genetic clusters/variation in samples by frequency of genotypes
- K= the number of clusters/groups within our dataset

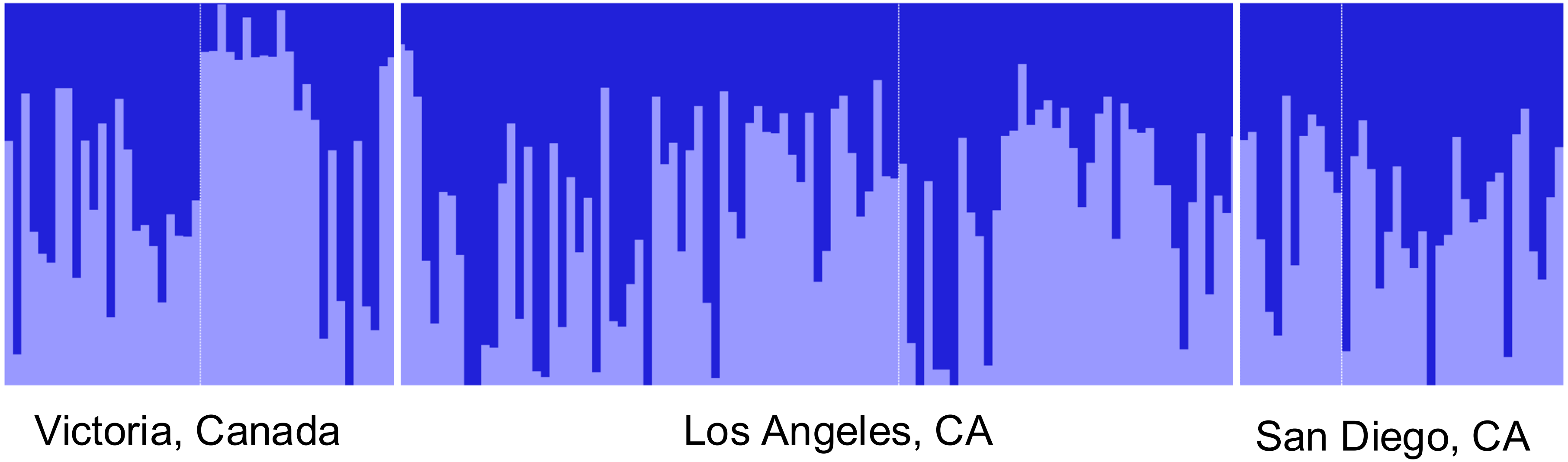


[Rose et al. (2020) https://www.cell.com/current-biology/fulltext/S0960-9822(20)30978-7?referringSource=articleShare](https://www.cell.com/current-biology/fulltext/S0960-9822(20)30978-7?referringSource=articleShare)

What can population structure tell us?

- Modeling genetic clusters/variation in samples by frequency of genotypes
- K= the number of clusters/groups within our dataset

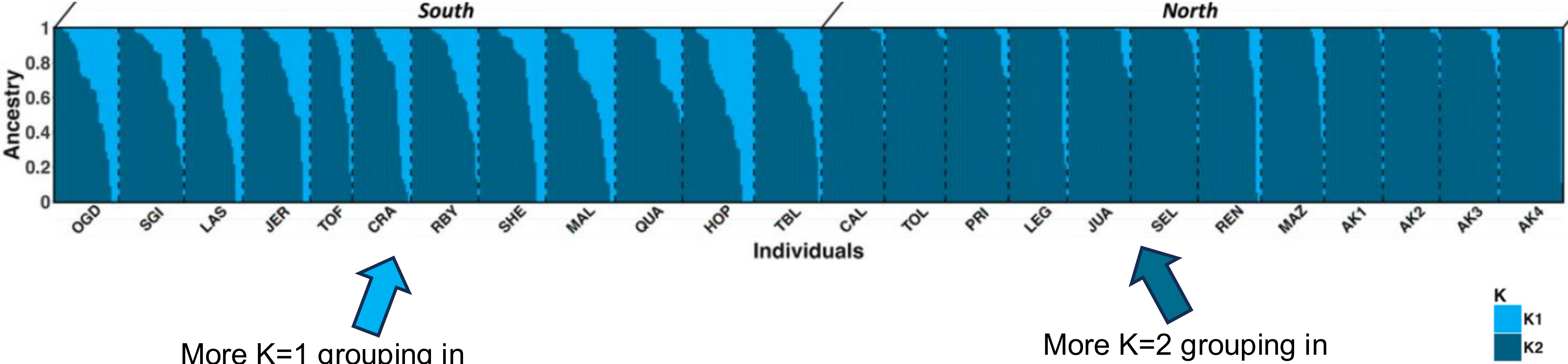
K=2 shown, but no structure between groups!




© CalPhotos

What can population structure tell us?

- Modeling genetic clusters/variation in samples by frequency of genotypes
- K= the number of clusters/groups within our dataset



More K=1 grouping in southern populations

More K=2 grouping in northern populations