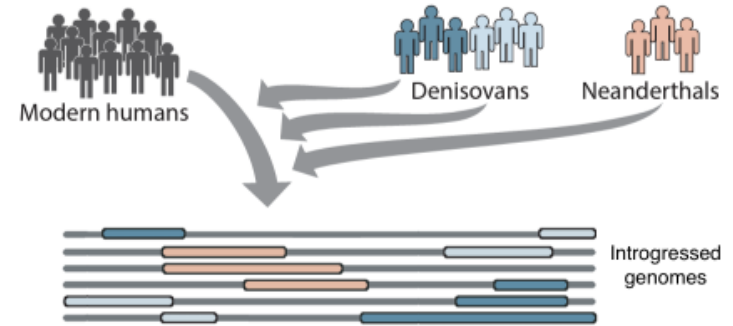
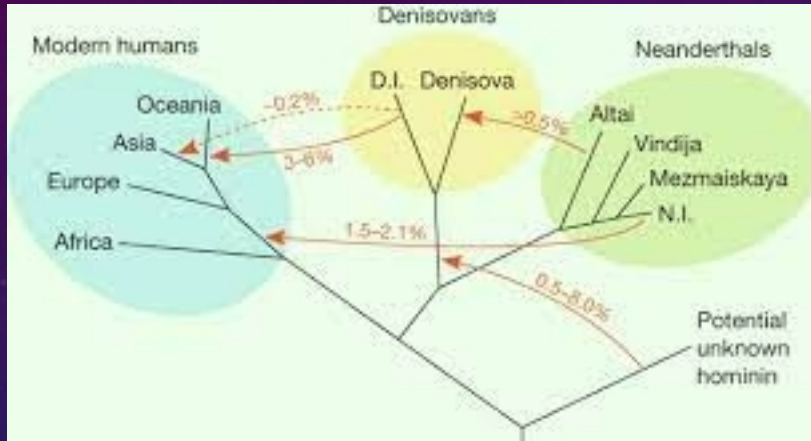




Investigating adaptive introgression in the climbing abilities in Austrian Elite Climbers

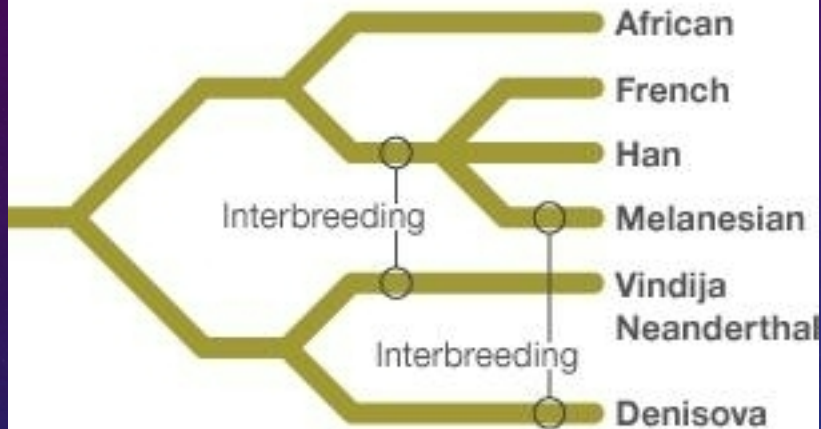


# Archaic Introgression

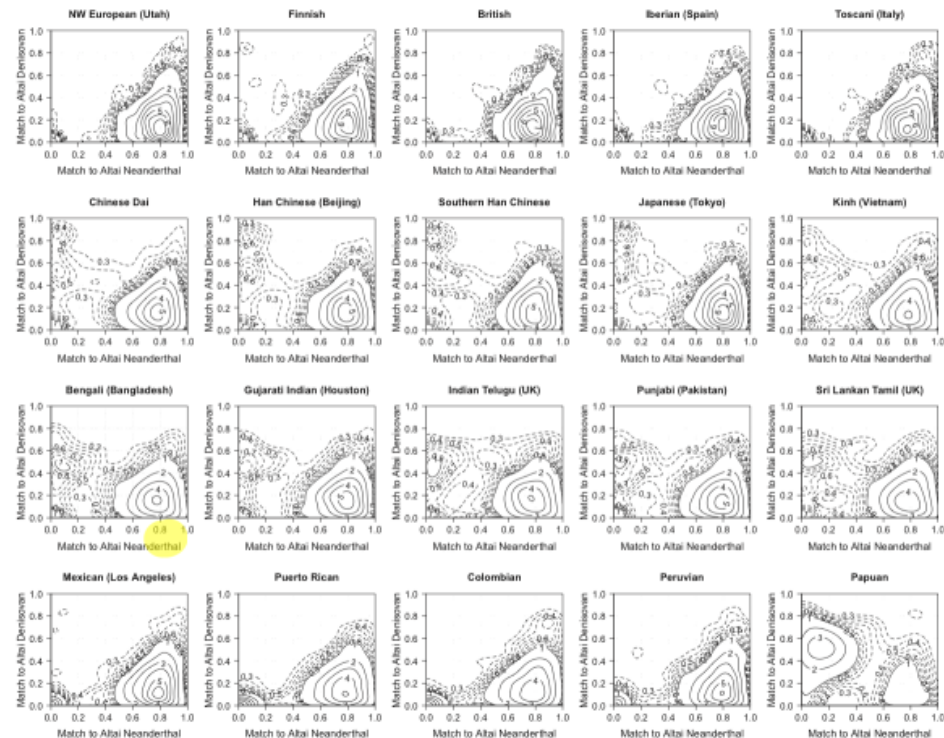


- the transfer of genetic information from one species to another as a result of hybridization and repeated backcrossing

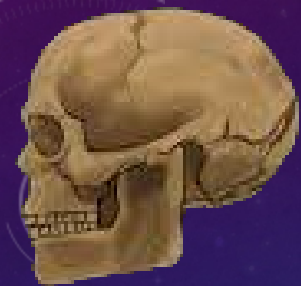
# Human family tree



Source: Nature

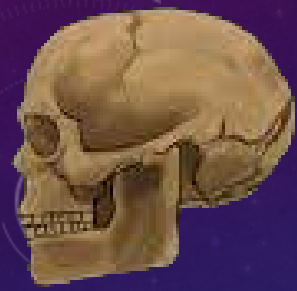
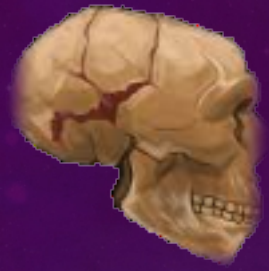






# Denisovans vs Neanderthals

	Denisovan	Neanderthal
<b>Era Alive</b>	Lower and Middle Paleolithic	Middle to Late Pleistocene
<b>Appearance</b>	The widest and flattest skull shape; large teeth and wide pelvises. Matured slower than Neanderthals	Wide skulls and pelvises, though more aligned with modern humans. Matured at a rate close to modern humans
<b>Geographical Location</b>	Eastern Asia	Europe and Western Asia
<b>Behavior</b>	Very little is known, but it is suspected that Denisovans bred frequently with Neanderthals and lived similar lives	Complex and industrious lives with complicated social structures. Created Neanderthal and Denisovan offspring, but only in certain locations
<b>Location Discovered</b>	Denisova Cave in Russia; 2008	Belgium in the 19th century



## Introduction and objective



- Goal: To find whether the genes responsible for aiding Elite Climbers have origins in archaic introgression (special interest in Denisovan introgression).
- Topics covered: Hypoxia resistance, climbing advantages, archaic introgression

# Genes associated with climbing advantages

- Investigation and record of genes which are associated with hypoxia-resistance
  - Recorded over 56600 genes

Database

## iHypoxia: An Integrative Database of Protein Expression Dynamics in Response to Hypoxia in Animals

Ze-Xian Liu<sup>1</sup>\*, Panqin Wang<sup>2</sup>\*, Qingfeng Zhang<sup>1</sup>\*, Shihua Li<sup>1 2</sup>, Yuxin Zhang<sup>2</sup>,  
Yutong Guo<sup>2</sup>, Chongchong Jia<sup>2</sup>, Tian Shao<sup>2</sup>, Lin Li<sup>1</sup>, Han Cheng<sup>2</sup>  ,  
Zhenlong Wang<sup>2</sup>  

<https://www.sciencedirect.com/science/article/pii/S1672022922001504>

#CHROM	TYPE	POS1	POS2	NAME
1	gene	14404	29570	WASH7P
1	gene	29554	31109	MIR1302-2HG
1	gene	34554	36081	FAM138A
1	gene	52473	53312	OR4G4P
1	gene	57598	64116	OR4G11P
1	gene	65419	71585	OR4F5
1	gene	89295	133723	AL627309.1
1	gene	89551	91105	AL627309.3
1	gene	131025	134836	CICP27
1	gene	134901	139379	AL627309.1

# Materials and Methods

- Sequenced data for the Elite Climbers + controls in gzip—compressed VCF Format (22 autosomal)
  - Samples in VCF:
    - 20 Austrian controls
    - 10 Austrian Elite Climbers(picked because of their extreme abilities)
    - 13 other European Climbers
- 1000 Genomes Project sequenced data
- Sequenced data for Altai Denisovan and Vindija Neanderthal genome populations in gzip-compressed VCF format

## Software and algorithms

SPrime	Brian Browning ( <a href="#">Browning et al. 2018</a> )	<a href="https://github.com/browning-lab/sprime">https://github.com/browning-lab/sprime</a> ↗
Bcftools	SAMtools	<a href="http://samtools.github.io/bcftools/bcftools.html">http://samtools.github.io/bcftools/bcftools.html</a> ↗



# Variant Call Format Files



#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	NA000001	NA000002	NA000003
20	14370	rs6054257	G	A	29	PASS	NS=3;DP=14;AF=0.5;DB;H2	GT:GQ:DP:HQ	0 0:48:1:51,51	1 0:48:8:51,51	1/1:43:5:.,.
20	17330	.	T	A	3	q10	NS=3;DP=11;AF=0.017	GT:GQ:DP:HQ	0 0:49:3:58,50	0 1:3:5:65,3	0/0:41:3
20	1110696	rs6040355	A	G,T	67	PASS	NS=2;DP=10;AF=0.333,0.667;AA=T;DB	GT:GQ:DP:HQ	1 2:21:6:23,27	2 1:2:0:18,2	2/2:35:4
20	1230237	.	T	.	47	PASS	NS=3;DP=13;AA=T	GT:GQ:DP:HQ	0 0:54:7:56,60	0 0:48:4:51,51	0/0:61:2
20	1234567	microsat1	GTC	G,GTCT	50	PASS	NS=3;DP=9;AA=G	GT:GQ:DP	0/1:35:4	0/2:17:2	1/1:40:3



# Material and Methods: Data Pre-processing

Elite Climbers VCF files

Phase Elite Climber files



Filter for variants with <0.99 GP from Elite Climber files



Outgroup (Yoruba)

Create file with all Yoruba sample names found in 1000 genome project site



Extract Yoruba samples from 1000 genome chromosome files



Merge the 22 filtered Elite Climber VCF files with the Yoruba sample files

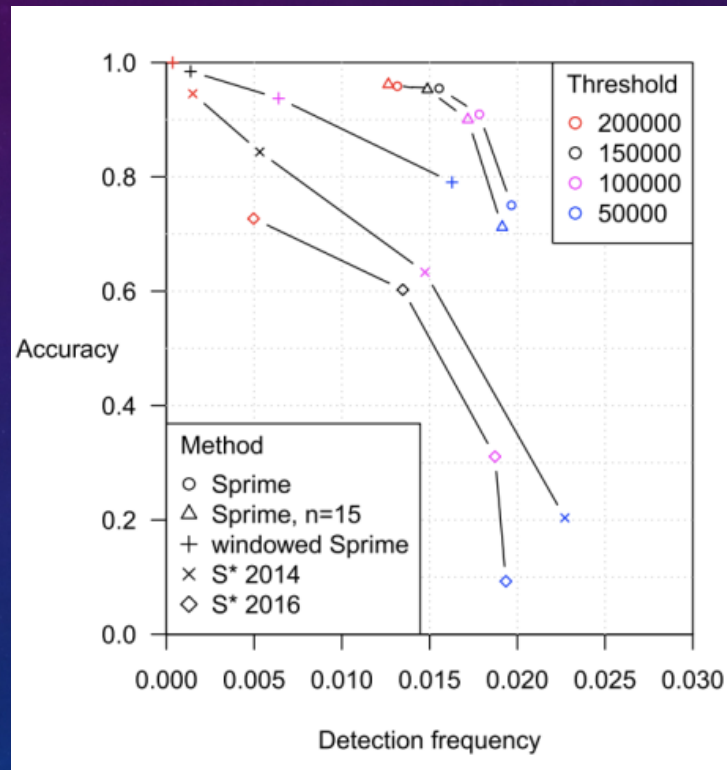
## Data Pre-processing of 22 Merged VCF files

Filter  
merged files  
for bi-allelic  
variants

Concatenate  
the resulting  
22 filtered  
merged VCF  
files

# TOOL: Sprime

- Method to detect introgressed archaic sequences without using a reference genome
  - optimized for detecting introgression from Neanderthals and Denisovans in modern humans
- **SPRIME parameters:**
  - `gt`= VCF file with target samples + outgroup samples
  - `outgroup`=A text file with outgroup sample names
  - `map`=Plink format recombination map
    - Sprime will use linear interpolation to estimate genetic positions between map positions.
  - `Out`=Directory to store outputs





## Install SPrime



1 min

Step 1

## Run SPrime

Score file for  
introgressed variants

1 h

Step 3

Find multiple sources of  
archaic introgression



1 min

Step 5

## Prepare input data for the SPrime

Recombination map  
Modern human  
genomes

5 h

Step 2

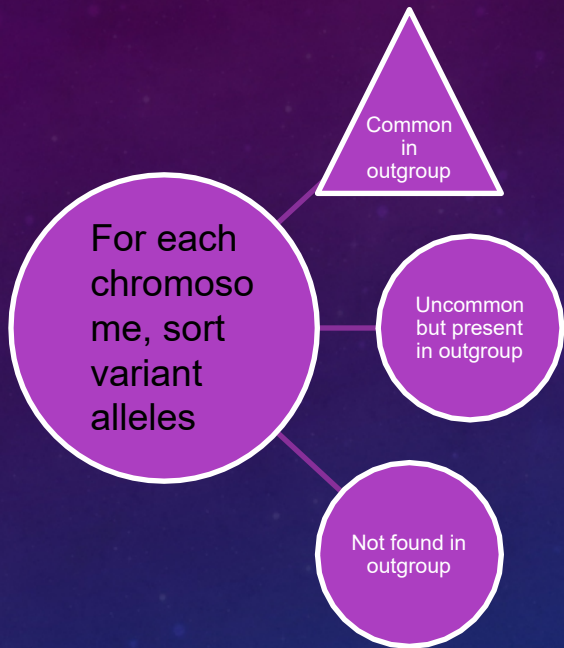
Calculate match rates to a  
known archaic genome

Annotated  
introgressed variants

1.5 h

Step 4

# Sprime



- `java -Xmx50g -jar sprime.jar gt=[file path] outgroup=[file path] map=[file path] out=[file path]`
- **Cases used:**
  - **Case where Austrian Climbers are target samples**
  - **Case where other European Climbers are target samples**
  - **Case where the Austrian controls are target samples**

# Output

- The log file (.log) contains a summary of the analysis.
- The score file (.score) lists all introgressed variants with segment score greater than or equal to the minscore parameter(which is 100,000 by default).

```
out_file_1.log
sprime.jar(version: 20May22.855)
Copyright (C) 2017 Brian L. Browning
Enter "java -jar sprime.jar" to print a list of command line arguments

Start time: 07:39 PM EST on 21 Dec 2022

Command line: java -Xmx45511m -jar sprime.jar
gt=Concatenated_sorted.vcf.gz
outgroup=gqfs/data/ehuertas/data/2022_EliteClimbers/YORUBA_SAMPLENAME.txt
map=gqfs/data/ehuertas/data/2022_EliteClimbers/recombination_maps/plink_chr1_GRCh37.map
excludesamples=gqfs/data/ehuertas/data/2022_EliteClimbers/Exclude_Samples/Exclude_for_control.txt
chrom=1
out=gqfs/data/ehuertas/data/2022_EliteClimbers/Sprime_Outputs/Controls_Only/out_file_1

Number of outgroup samples: 108
Number of target samples: 20
Variant analyzed: 4951663
Segments detected: 160
Run time: 36 minutes 55 seconds

End time: 08:16 PM EST on 21 Dec 2022

sprime.jar finished
```

CHROM	POS	ID	REF	ALT	SEGMENT	ALLELE	SCORE
1	807512	rs10751454	A	G	137	0	107638
1	807761	rs4951932	C	A	137	0	107638
1	812267	rs7541694	A	G	137	0	107638
1	812284	rs7545372	C	G	137	0	107638
1	821887	rs6677354	A	G	137	0	107638
1	830807	rs6422669	G	C	137	0	107638
1	847250	rs7416129	G	A	137	0	107638
1	866920	rs2341361	A	G	137	0	107638
1	877831	rs6672356	T	C	137	0	107638
1	883625	rs4970378	A	G	137	0	107638
1	901652	rs2879814	A	G	137	0	107638
1	909768	rs2340593	A	G	137	0	107638
1	921716	rs13303278	C	A	137	1	107638
1	924528	rs34712273	C	A	137	1	107638
1	935222	rs2298214	C	A	137	1	107638
1	953952	rs9442612	G	A	137	1	107638
1	954777	rs61766299	C	A	137	1	107638
1	960409	rs4970392	G	C	137	1	107638
1	967658	rs4970349	C	T	137	1	107638
1	1047576	rs146157459	A	G	110	1	123007
1	1059199	rs537530630	C	T	110	1	123007
1	1066898	rs189892117	G	A	110	1	123007
1	1069421	rs554638733	G	A	110	1	123007
1	1072880	rs143580335	G	A	110	1	123007
1	1085424	rs145309131	T	C	110	1	123007
1	1105299	rs111997742	G	A	110	1	123007
1	1108951	rs76610658	G	T	110	1	123007
1	1109193	rs41300008	G	A	110	1	123007
1	1111147	rs112420268	A	C	110	1	123007



# Sprime Results

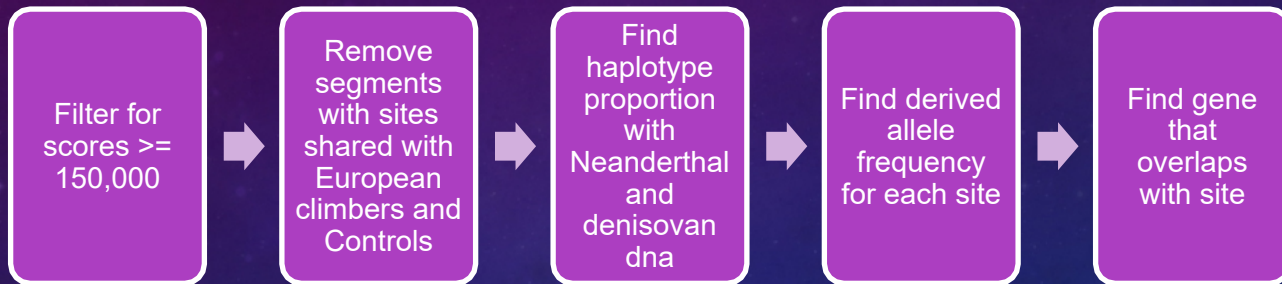
austrian\_control\_scores\_table

CHROM	POS	ID	REF	ALT	SEGMENT	ALLELE	SCORE	SPrime_Inference
1	167332504	rs534568998	G	T	48	1	195103	T
1	205342450	rs12122704	T	G	94	1	137916	G
1	205342652	rs12122771	T	G	94	1	137916	G
1	205343478	rs79344214	A	G	94	1	137916	G
1	205343628	rs79345034	T	C	94	1	137916	C
1	205344061	rs11240460	T	C	94	1	137916	C
1	205349507	rs77453816	G	T	94	1	137916	T
1	205353492	rs2274702	A	C	94	1	137916	C
1	205354695	rs7349182	C	T	94	1	137916	T
1	205355903	rs12135649	C	T	94	1	137916	T
1	205360764	rs12133217	G	A	94	1	137916	A
1	205360957	rs12140919	C	T	94	1	137916	T
1	205363898	rs140566229	C	T	94	1	137916	T
1	205363972	rs138003490	G	A	94	1	137916	A
1	205366677	rs12141590	C	G	94	1	137916	G
1	205369097	rs77594411	G	C	94	1	137916	C
1	205370469	rs143927182	C	A	94	1	137916	A
1	205378492	rs55780799	C	T	94	1	137916	T
1	205342351	rs12125326	G	A	94	1	137916	A
1	205341383	rs74398318	C	T	94	1	137916	T
1	205340765	rs144569305	G	A	94	1	137916	A
1	205339542	rs56300017	A	G	94	1	137916	G
1	203136361	rs200134555	T	G	148	1	105088	G
1	203139187	rs139323917	G	A	148	1	105088	A

austrian\_climbers\_scores\_table

#CHROM	POS	ID	REF	ALT	SEGMENT	ALLELE	SCORE	SPrime_Inference
1	192116126	rs80151391	A	G	129	1	103303	G
1	225200110	rs111862721	G	A	7	1	394707	A
1	225191880	rs78139584	C	G	7	1	394707	G
1	225191264	rs79257893	G	T	7	1	394707	T
1	225189138	rs80016617	G	A	7	1	394707	A
1	225188911	rs80023439	C	T	7	1	394707	T
1	225183870	rs79620384	A	G	7	1	394707	G
1	225182124	rs78638349	A	G	7	1	394707	G
1	225181214	rs75254122	C	A	7	1	394707	A
1	225176961	rs112489082	A	G	7	1	394707	G
1	225175625	rs112286465	G	A	7	1	394707	A
1	225175436	rs113493760	C	T	7	1	394707	T
1	225172387	rs111552187	C	G	7	1	394707	G
1	225202436	rs75044318	A	G	7	1	394707	G
1	225171163	rs76032339	A	G	7	1	394707	G
1	225165811	rs111669697	C	G	7	1	394707	G
1	225165768	rs78497820	A	G	7	1	394707	G
1	225164566	rs111676659	A	G	7	1	394707	G
1	225164379	rs541685284	G	A	7	1	394707	A
1	225164171	rs111896783	A	G	7	1	394707	G
1	225164150	rs75625634	C	T	7	1	394707	T
1	225162613	rs189673231	A	C	7	1	394707	C
1	225161877	rs72470459	A	G	7	1	394707	G
1	225160970	rs74699857	A	G	7	1	394707	G

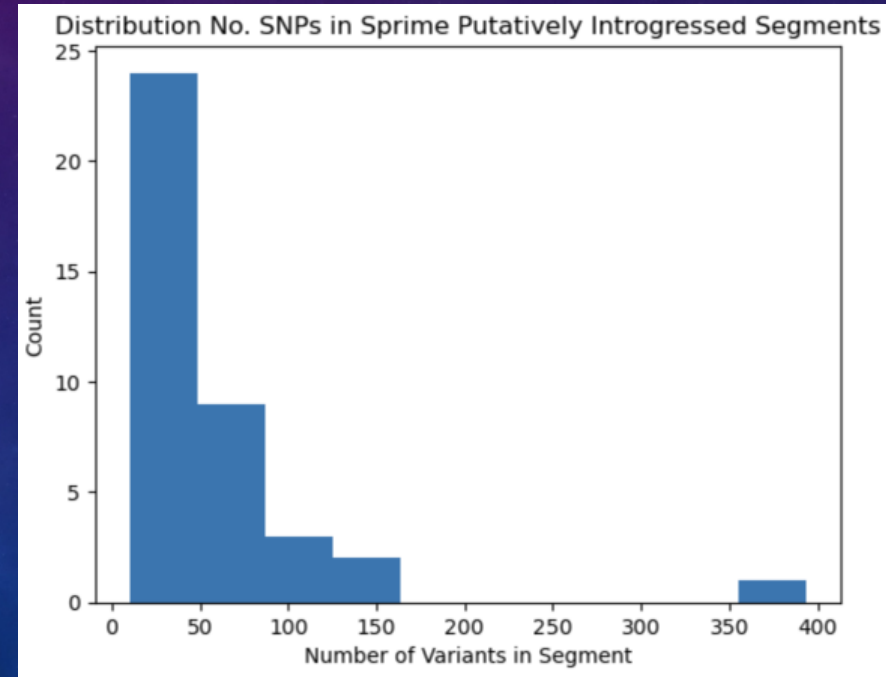
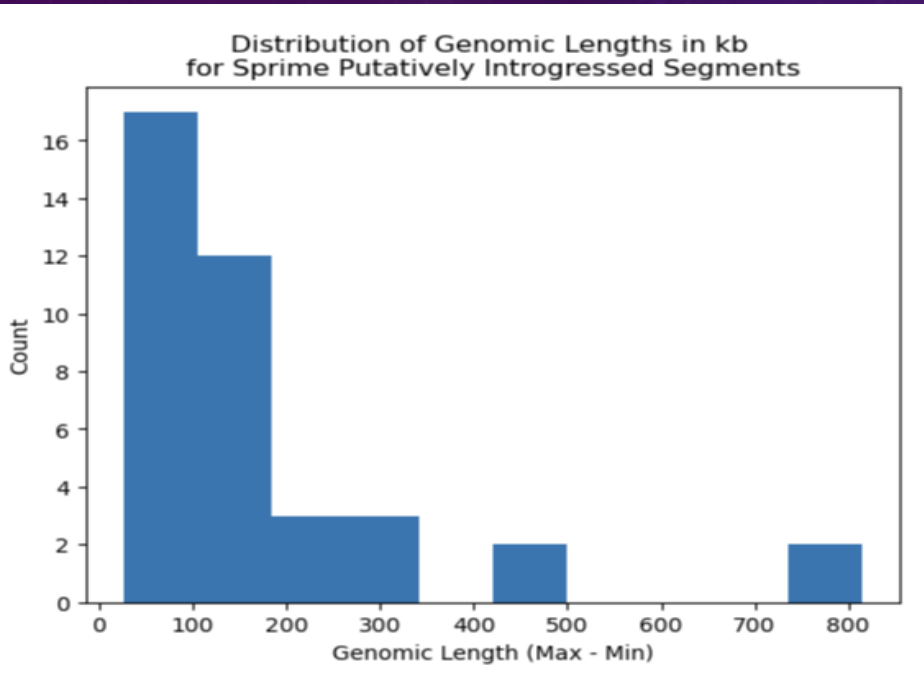
# More filtering



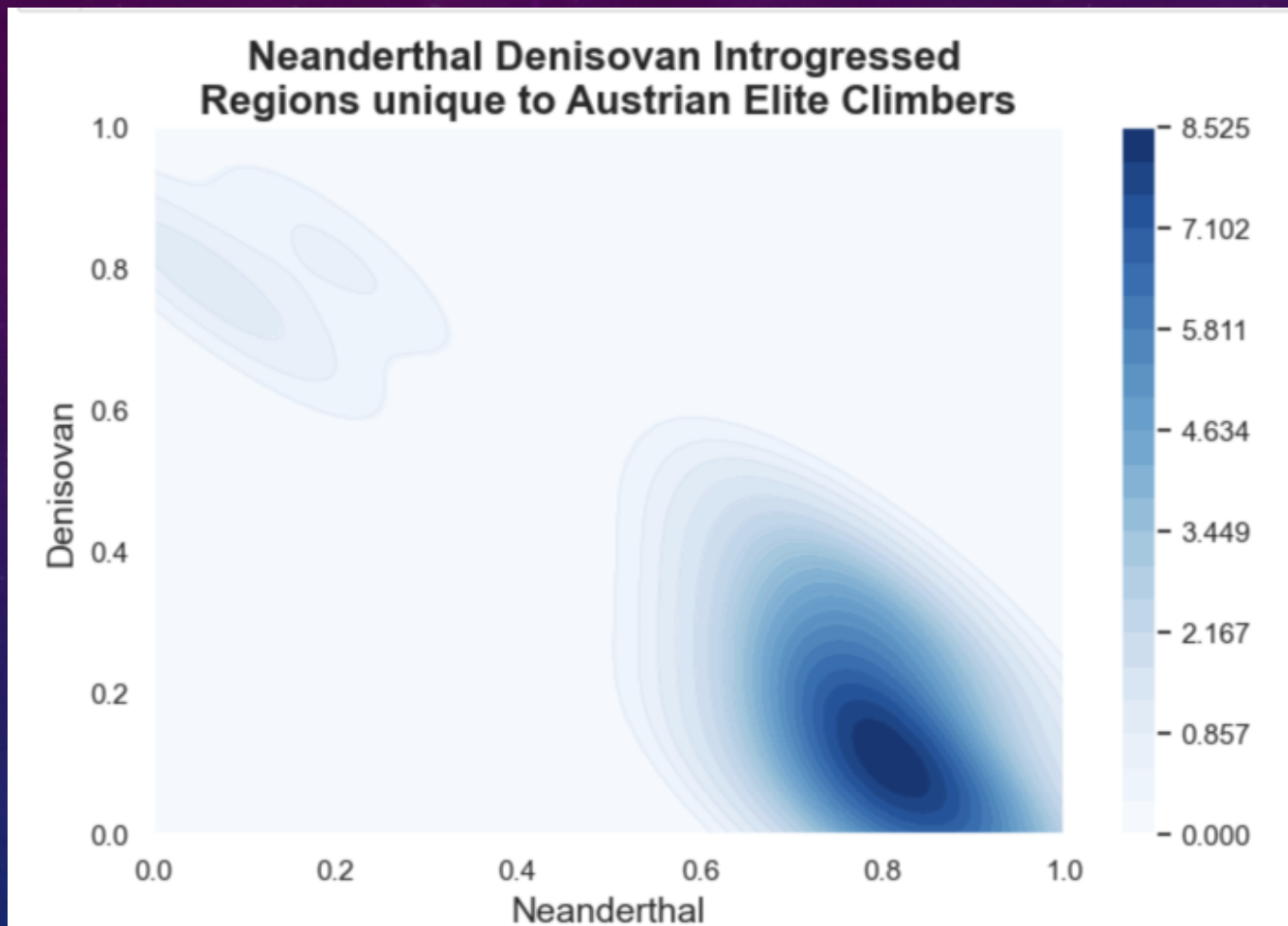
#CHROM	POS	ID	REF	ALT_x	AUS1	AUS2	AUS3	AUS4	AUS5	...	AUS8	AUS9	AUS10	Nean_genot	Deni_genot	SEGMENT	match_nean	match_deni	samp_der_freq	GENE
22	17810053	rs17808489	G	A	1 0	0 0	0 0	0 0	0 0	...	1 0	0 0	0 0	1/1	0/0	12	0.2	1.0	0.10	CECR9
22	17815187	rs78563168	C	T	0 0	0 0	0 0	0 0	0 0	...	1 0	0 0	0 0	1/1	0/0	12	0.1	1.0	0.05	None
22	17817037	rs4819575	C	T	0 0	0 0	0 0	0 0	0 0	...	1 0	0 0	0 0	1/1	0/0	12	0.1	1.0	0.05	None
22	17819348	rs2189076	T	C	0 0	0 0	0 0	0 0	0 0	...	1 0	0 0	0 0	1/1	0/0	12	0.1	1.0	0.05	None
22	17834918	rs5749051	G	C	0 0	0 0	0 0	0 0	0 0	...	0 1	0 0	0 0	1/1	1/1	12	0.1	0.1	0.05	None

# Results and Figures

- 28 putatively introgressed region/ segments found that are unique to Austrian Climbers







Most segments unique to Austrian Climbers present high affinity to Neanderthal Genome

AUS1  
AUS2  
AUS3  
AUS4  
AUS5  
AUS6  
AUS7  
AUS8  
AUS9  
AUS10  
Nea  
Deni

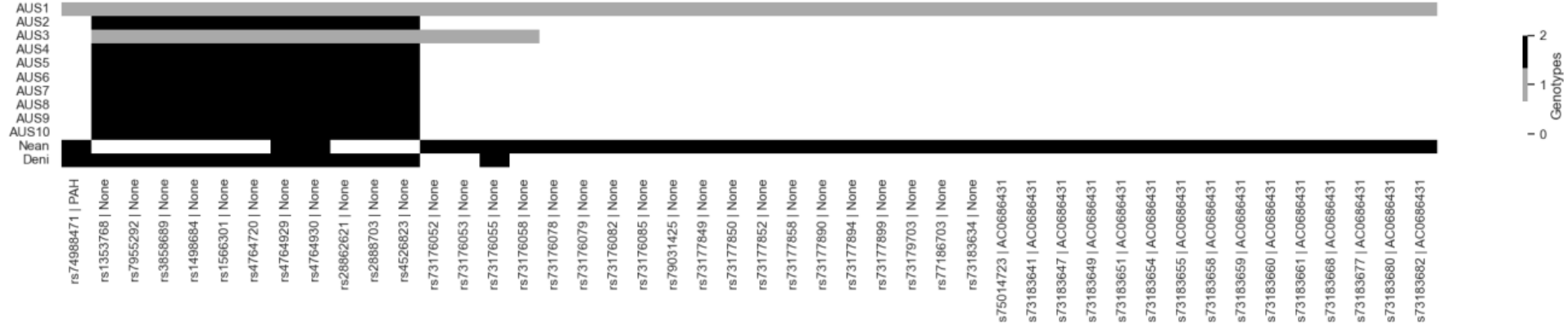


Alternate Allele Frequency  
Segment 70 Chromosome 1  
Genes: [None, 'TNFRSF1B']

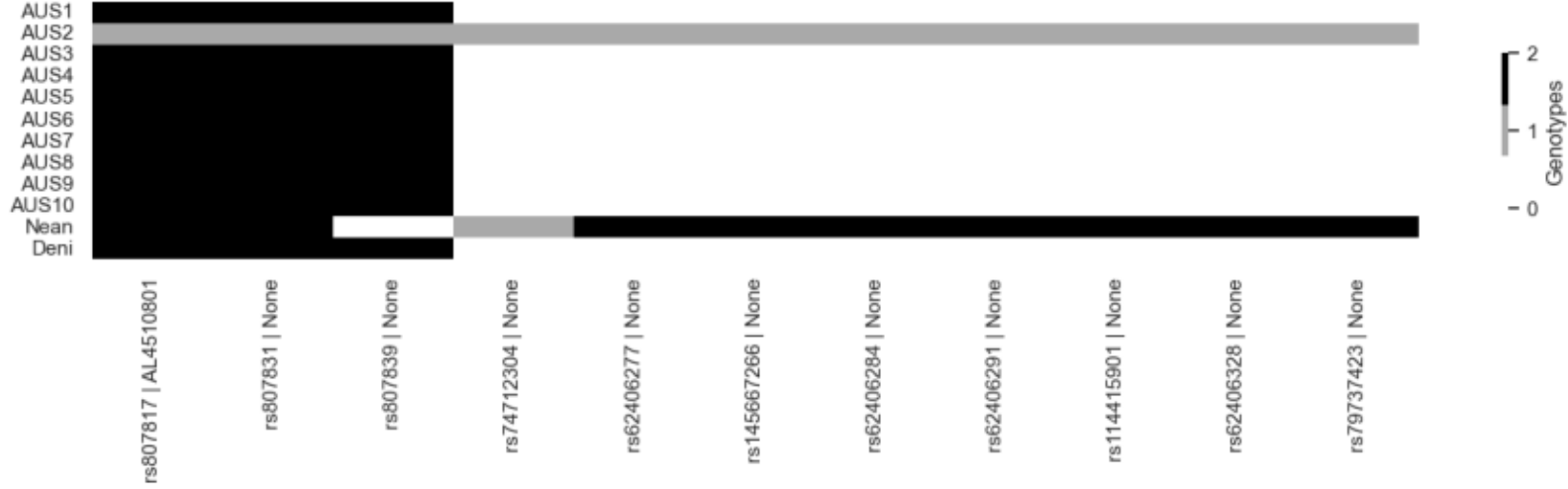
AUS1  
AUS2  
AUS3  
AUS4  
AUS5  
AUS6  
AUS7  
AUS8  
AUS9  
AUS10  
Nea  
Deni



Alternate Allele Frequency  
Segment 24 Chromosome 12  
Genes: {'AC0686431', 'PAH'}



Alternate Allele Frequency  
Segment 57 Chromosome 6  
Genes: {'AL4510801'}





# Conclusions (for now)

- There's some evidence of some of the hypoxia\_resistance genes being a result of archaic introgression
- There's some evidence of recent introgression with Denisivans
  - This could be as a result of direct interbreeding with Denisovans or interbreeding with Neanderthals who have some Denisovan DNA (since it is known that Neanderthal interbred with Denisovans)
- Need to take more genes/genetic material into consideration
  - The areas of interest do not show overlap perhaps due to the limited amount of genes considered
- Further analysis need to be done to be more certain of results
  - The austrian climbers had less variants picked as introgressed compared to other European and the control so will need more analysis.
  - Look more into Chromosome 12

# Works Cited

- Huerta-Sánchez E, Jin X, Asan, Bianba Z, Peter BM, Vinckenbosch N, Liang Y, Yi X, He M, Somel M, Ni P, Wang B, Ou X, Huasang, Luosang J, Cuo ZX, Li K, Gao G, Yin Y, Wang W, Zhang X, Xu X, Yang H, Li Y, Wang J, Wang J, Nielsen R. Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature*. 2014 Aug 14;512(7513):194-7. doi: 10.1038/nature13408. Epub 2014 Jul 2. PMID: 25043035; PMCID: PMC4134395.
- Ze-Xian Liu, Panqin Wang, Qingfeng Zhang, Shihua Li, Yuxin Zhang, Yutong Guo, Chongchong Jia, Tian Shao, Lin Li, Han Cheng, Zhenlong Wang, iHypoxia: An Integrative Database of Protein Expression Dynamics in Response to Hypoxia in Animals, Genomics, Proteomics & Bioinformatics, 2022, <https://doi.org/10.1016/j.gpb.2022.12.001>.