

Objective:

Investigating the Role of Archaic Introgression in Genes Associated with Hypoxia Resistance and Climbing Advantages among Austrian Elite Climbers. Understanding the role of archaic introgression in the genetic makeup of Austrian Elite Climbers could provide valuable insights into the adaptive mechanisms that contribute to their exceptional abilities in climbing and resistance to hypoxia. By investigating specific genetic variants, we aim to uncover the intricate connections between ancient introgression events and the unique physiological traits observed in this population.

Introduction:

Archaic introgression in humans refers to the interbreeding events between anatomically modern humans and archaic hominins, such as Neanderthals and Denisovans. This phenomenon involves the transfer of genetic material between species through hybridization and backcrossing over generations. The resulting genetic diversity can influence various traits, including immunity, disease susceptibility, and potentially, adaptations to specific environments such as high-altitude regions.

In the context of Austrian Elite Climbers, we aim to explore the genetic variants associated with hypoxia resistance and climbing advantages, focusing on the potential influence of archaic introgression.

Key Points:

1. Modern Human-Archaic Populations Interbreeding:

- Modern humans, originating out of Africa, have interbred with archaic populations like Denisovans and Neanderthals.
- This interbreeding has led to the flow of genetic material from archaic populations into the gene pool of modern humans, known as archaic introgression.

2. Geographic Distribution of Introgressed DNA:

- Asian genomes carry introgressed DNA from both Denisovans and Neanderthals.
- East Asians exhibit evidence of introgression from two distinct Denisovan populations.
- South Asians and Oceanians carry introgression from one Denisovan population.
- European genomes primarily contain introgressed Neanderthal DNA.

4. Examples of Archaic Introgression Benefits:

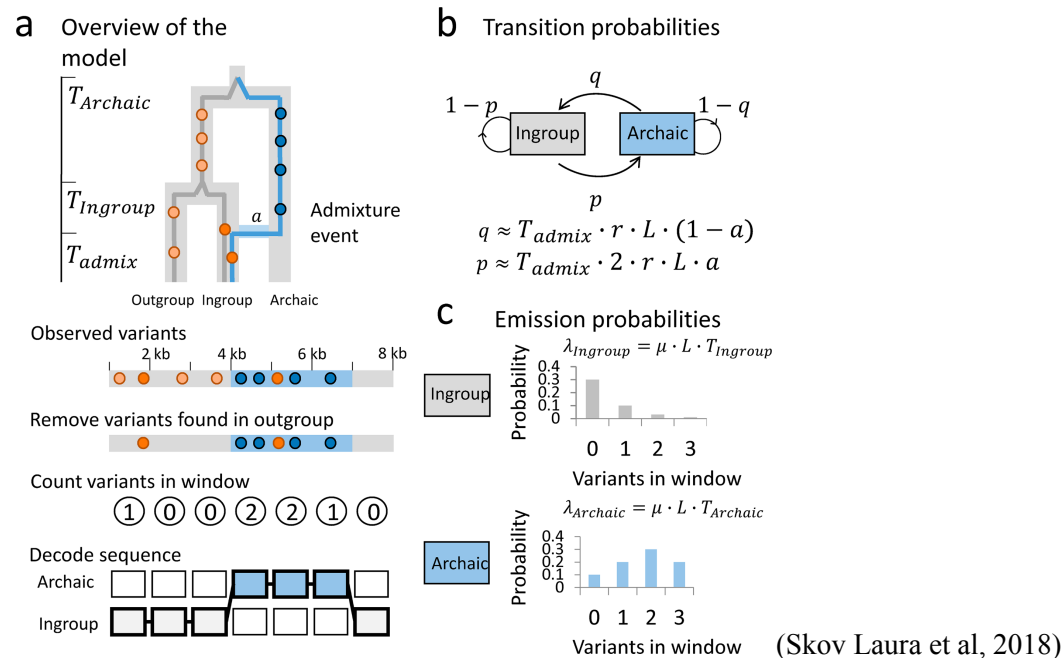
- Introgression has been identified as a contributing factor to the ability of certain populations, like Tibetans, to thrive in high-altitude environments.
- The EPAS1 gene, showing signatures of Denisovan introgression, may play a role in hypoxia resistance. (Huerta-Sánchez et al, 2014)

Methodology

Tools used

- The main tool chosen to do this analysis uses a hidden markov model to categorize genomic segments into states characterized by varying densities of genetic variants. It works best in a situation where introgression occurred solely within an ingroup, involving a deeply divergent archaic population, and not within the outgroup. This is applicable to this project since it is known from previous research that introgression occurred from Neanderthals into non-Africans 50,000 or so years ago, with Denisovan introgression being much older (Skvov L et al, 2018)
- Why this model?
 - It does not require reference genome
 - It works for both phased and unphased data
 - It works for data with missing bases
 - When compared with other popular introgression tools (such as Sprime) using data similar to the data in this project, it had the best tradeoff between precision and detection frequency

Overview of the algorithm:



Baum-Welch in detecting introgression (Skvov L et al (2018))

1. A HMM is defined with two hidden states, "Ingroup" and "Archaic". These states represent the source of a particular segment of the genome - either from the main population (Ingroup) or from an archaic population that interbred with the main population (Archaic).

2. The probability of changing from one state to another is represented by p (for Ingroup) and q (for Archaic). These transition probabilities can be related to the recombination rate (r), the time since admixture (T_{admix}), and the proportion of the genome that comes from the archaic population (a).
3. The genome is divided into windows of a certain length (L), typically 1000 base pairs. Within each window, the number of variants that are unique to the Ingroup or Archaic population (i.e., "private variants") is assumed to follow a Poisson distribution. The mean of this distribution (λ) depends on the mutation rate (μ), the length of the window (L), and the mean coalescence time (T), which is the time to the most recent common ancestor of the Ingroup and the outgroup, or the Archaic and the outgroup.
4. The Poisson rates are then adjusted to account for the number of missing bases in a window and the local mutation rate. This is done for each window i .
5. The Baum-Welch algorithm, an iterative method for estimating the parameters of a HMM, is used to find the values of p , q , λ_{Ingroup} , and λ_{Archaic} that maximize the likelihood of the observed data. It uses forward-backwards algorithm (which combines forward and backward probabilities to compute the posterior marginal, which is the probability of being in state i at time t given the entire observation sequence. This is done for each state at each time step.)
7. Once the optimal parameters are found, they can be used to "decode" the genome and identify candidate introgressed segments. This is done using posterior decoding, which calculates the probability of each state at each position in the genome given the observed data. Regions with a high posterior probability of being in the Archaic state are considered candidate introgressed segments.

Materials

- **Ingroup: 10 Elite Climbers genetic data in gzip—compressed VCF Format (22 autosomal)**
- **1000 Genomes Project sequenced data for outgroup (Yoruba)**
- **Sequenced data for Altai Denisovan and Vindija Neanderthal genome populations in gzip-compressed BCF format**
- **Text files indicating the names of all samples**
- **File with hypoxia_resistance gene information (Ze-Xian Liu et al, 2022)**

Results/Figures

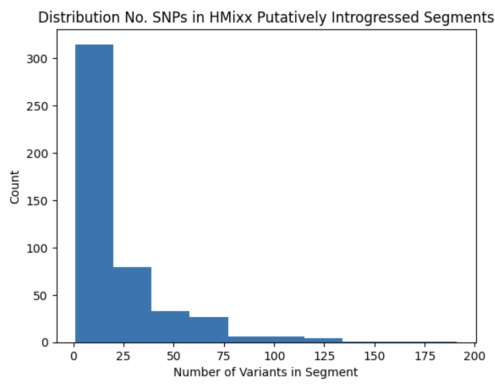
Steps

- After collecting and preprocessing the materials, the Hmixx algorithm was run on the data and the results were further analyzed
 - The window sizes were kept at 1000 bp

- ```
> state_names = ['Human', 'Archaic']
> starting_probabilities = [0.98, 0.02]
> transitions = [[1.0, 0.0], [0.02, 0.98]]
> emissions = [0.04, 0.4]
```

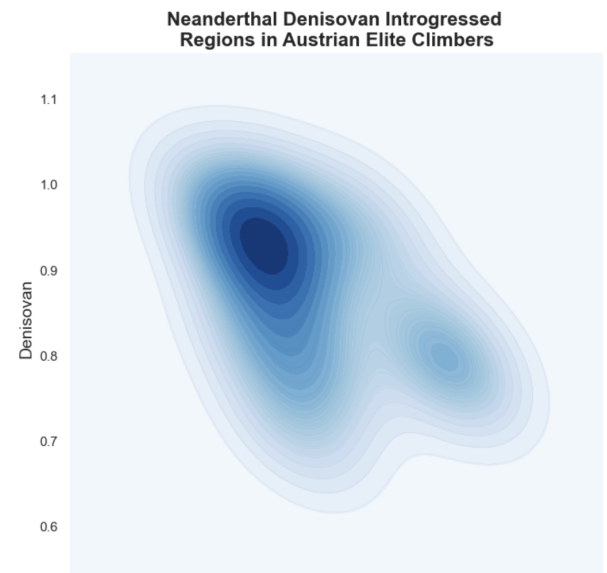
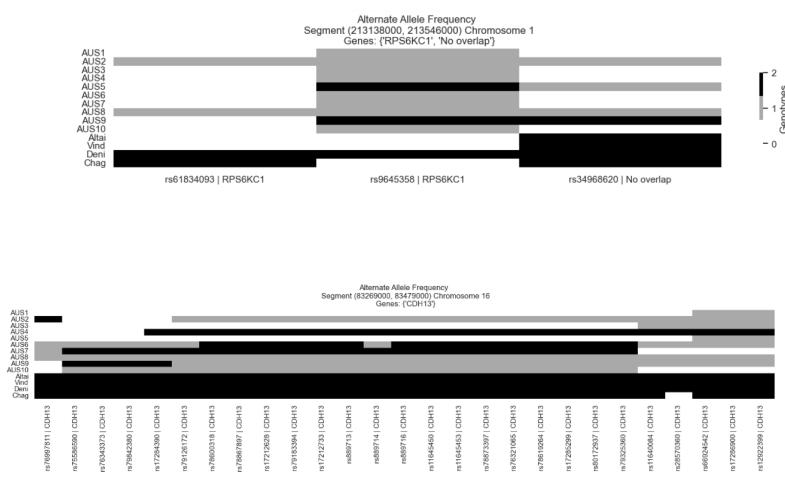
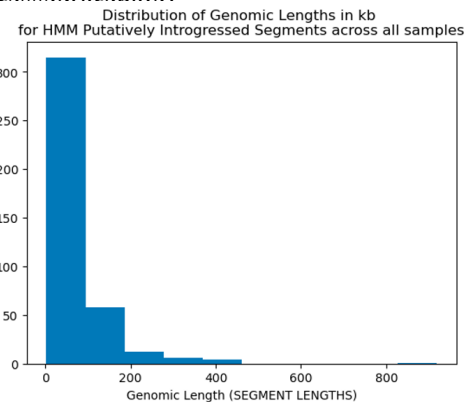
- Hmixx provides individual results per sample and shows which segments per chromosome were picked as introgressed(archaic) or not (human)

Proportion No. SNPs in HMixx Putatively Introgressed Segments



Distribution mean\_probs in HMM Putatively Introgressed Segments

| Number of mean_prob in Segment | Count |
|--------------------------------|-------|
| 0.5 - 0.6                      | 15    |
| 0.6 - 0.7                      | 35    |
| 0.7 - 0.8                      | 10    |
| 0.8 - 0.9                      | 15    |
| 0.9 - 1.0                      | 55    |
| 1.0 - 1.1                      | 45    |
| 1.1 - 1.2                      | 78    |
| 1.2 - 1.3                      | 105   |
| 1.3 - 1.4                      | 110   |



## Conclusion:

- There's evidence of some hypoxia resistance genes having origins in archaic introgression
  - Genes: 'ZNF385D', 'RPS6KC1' and CDH13 show the most evidence and should be investigated more
  - It is interesting to note that a gene like EPAS1 which is known for showing signs of introgression in other groups from other studies (Huerta-Sanchez, 2014) did not appear to be of much interest here.
- Need to take more genes/genetic material into consideration
  - It will be beneficial to find more literature on hypoxia-related genes as well as other climbing related genetic material. It might also help to increase sample size.
- It is interesting to see higher affinity for Denisovan Genome vs Neanderthal when the climbers are European. This needs to be further investigated.
- Some 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
  - To finalize results, it will still be helpful to check with the other tools such as Sprime to ensure our results also have the highest likelihood of being accurate
  - Performing statistical tests

## Bibliography

- Skov L, Hui R, Shchur V, Hobolth A, Scally A, Schierup MH, et al. (2018) Detecting archaic introgression using an unadmixed outgroup. PLoS Genet 14(9): e1007641. <https://doi.org/10.1371/journal.pgen.1007641>
- Huerta-Sánchez, Emilia, et al. "Altitude Adaptation in Tibetans Caused by Introgression of Denisovan-like DNA." Nature News, Nature Publishing Group, 2 July 2014, [www.nature.com/articles/nature13408](http://www.nature.com/articles/nature13408)
- 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>