Human adaptation to high altitude

Emilia Huerta Sanchez
Postdoc
UC Berkeley
Elucidating evolution

Genetic Variation

Evolution

Population Genetics

Statistical Tools
Human migrations

- 195 KYA
- 40 KYA
- 47 KYA
- 39 KYA
- 15–25 KYA
- 50 KYA
What does genetics tell us about human adaptation?
So how different are we?
Humans are genetically very similar to each other

- 90% of variation within populations
- Large genetic differences are very rare
- Such differences can appear if they improve adaptation to local environment
  - E.g. climate, diet, microbes
Humans are genetically very similar to each other.
Positive natural selection
Positive natural selection
Positive natural selection
Positive natural selection
Positive natural selection
Selection on a de novo mutation
Selection on a de novo mutation
Selection on standing variation
Model of selection on a de novo mutation (SDN)
Model of selection on standing variation (SSV)
Changes in skin pigmentation

UV light intensity

Distribution of skin pigmentation

Human Skin Color Distribution
Changes in skin pigmentation

UV light intensity

Distribution of skin pigmentation

Human Skin Color Distribution

SLC24A5
Adaptation to high altitude

- At high altitude there are fewer oxygen molecules in a breath of air than at sea level
- Humans inhabit three regions of the world that are at extreme altitudes
Response to high altitude environments

• Hemoglobin concentration: the protein in red blood cells that carries oxygen
Response to high altitude environments in Andeans

- Hemoglobin: the protein in red blood cells that carries oxygen

The physiological response to hypoxia was first scientifically investigated in the late 19th century in South America.

- It was believed that increasing hemoglobin concentration was beneficial
Response to high altitude environments in Tibetans

In the 70s, studies showed that Tibetans had a different physiological response.
Contrasting Tibetans, Andeans and Ethiopian response

Beall Cynthia, 2006
Higher fertility and lower infant mortality rate in high altitude natives than in acclimatized low altitude natives
Exome Sequencing

- Exome (all the exons of the genome) – the coding part of the genome
- Technology: Exon capture & High-throughput sequencing
- 50 Tibetan Individuals living above 4000m altitude
Other data

• A closely related population: Han from Beijing

• We used an outgroup population: 200 Danish exomes
Mutation Frequencies
Identifying signatures of positive selection

Diagram:

- High
- Low
- Outgroup
Identifying signatures of positive selection

\[ PBS_{High} = \frac{\, T_{High,Low}}{} \]

Yi* X., Liang* Y., Huerta-Sanchez* E., Jin* X., Pool* J. et al. (2010) Science
Identifying signatures of positive selection

\[ PBS_{High} = T_{High,Low} + T_{High,Outgroup} \]
Identifying signatures of positive selection

\[ PBS_{High} = T_{High,Low} + T_{High,Outgroup} - T_{Low,Outgroup} \]
Identifying signatures of positive selection

\[ PBS_{High} = \frac{1}{2} \left[ T_{High,Low} + T_{High,Outgroup} - T_{Low,Outgroup} \right] \]
Under no positive selection

High  Low  Outgroup
Under positive selection

High

Low

Outgroup
Largest branch length: \textit{EPAS1}
Distribution of branch length values across genes

![Graph showing distribution of branch length values across genes with a focus on EPAS1.](image-url)
EPAS1: Hypoxia inducible factor 2

- Major Transcription factor that orchestrates response to low oxygen levels
- Regulates several genes involved in red blood cell production
- Mutations in *EPAS1* have been associated with super-athlete performances
- Highly expressed in the adult and fetal lung and placenta
EPAS1: large frequency differences.

<table>
<thead>
<tr>
<th></th>
<th>Tib</th>
<th>Han</th>
<th>Danes</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>0.87</td>
<td>0.09</td>
<td>0.0</td>
</tr>
</tbody>
</table>

---

- EPAS1: large frequency differences.
Other relevant genes

<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
<th>Nearby candidate</th>
<th>Tibetan PBS</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPAS1</td>
<td>endothelial PAS domain protein 1</td>
<td>(self)</td>
<td>0.514</td>
</tr>
<tr>
<td>C1orf124</td>
<td>hypothetical protein LOC83932 isoform a</td>
<td>EGLN1</td>
<td>0.277</td>
</tr>
<tr>
<td>DISC1</td>
<td>disrupted in schizophrenia 1 isoform L</td>
<td>EGLN1</td>
<td>0.251</td>
</tr>
<tr>
<td>ATP6V1E2</td>
<td>ATPase, H+ transporting, lysosomal 31kDa, V1</td>
<td>EPAS1</td>
<td>0.246</td>
</tr>
<tr>
<td>SPP1</td>
<td>secreted phosphoprotein 1 isoform c</td>
<td>(self)</td>
<td>0.238</td>
</tr>
<tr>
<td>PKLR</td>
<td>pyruvate kinase, liver and RBC isoform 1</td>
<td>(self)</td>
<td>0.230</td>
</tr>
<tr>
<td>C4orf7</td>
<td>chromosome 4 open reading frame 7</td>
<td></td>
<td>0.227</td>
</tr>
<tr>
<td>PSME2</td>
<td>proteasome activator subunit 2</td>
<td></td>
<td>0.222</td>
</tr>
<tr>
<td>OR10X1</td>
<td>olfactory receptor, family 10, subfamily X,</td>
<td>SPTA1</td>
<td>0.218</td>
</tr>
<tr>
<td>FAM9C</td>
<td>family with sequence similarity 9, member C</td>
<td>TMSB4X</td>
<td>0.216</td>
</tr>
<tr>
<td>LRRC3B</td>
<td>leucine rich repeat containing 3B</td>
<td></td>
<td>0.215</td>
</tr>
<tr>
<td>KRTAP21-2</td>
<td>keratin associated protein 21-2</td>
<td></td>
<td>0.213</td>
</tr>
<tr>
<td>HIST1H2BE</td>
<td>histone cluster 1, H2be</td>
<td>HFE</td>
<td>0.212</td>
</tr>
<tr>
<td>TTL3</td>
<td>tubulin tyrosine ligase-like family, member 3</td>
<td>HFE</td>
<td>0.206</td>
</tr>
<tr>
<td>HIST1H4B</td>
<td>histone cluster 1, H4b</td>
<td></td>
<td>0.204</td>
</tr>
<tr>
<td>ACVR1B</td>
<td>activin A type 1B receptor isoform a precursor</td>
<td></td>
<td>0.198</td>
</tr>
<tr>
<td>FXYD6</td>
<td>FXYD domain-containing ion transport regulator</td>
<td></td>
<td>0.192</td>
</tr>
<tr>
<td>NAGLU</td>
<td>alpha-N-acetylgalactosaminidase precursor</td>
<td></td>
<td>0.186</td>
</tr>
<tr>
<td>MDH1B</td>
<td>malate dehydrogenase 1B, NAD (soluble)</td>
<td></td>
<td>0.184</td>
</tr>
<tr>
<td>OR6Y1</td>
<td>olfactory receptor, family 6, subfamily Y,</td>
<td>SPTA1</td>
<td>0.183</td>
</tr>
<tr>
<td>HBB</td>
<td>beta globin</td>
<td>(self), HBG2</td>
<td>0.182</td>
</tr>
<tr>
<td>OTX1</td>
<td>orthodenticle homeobox 1</td>
<td></td>
<td>0.181</td>
</tr>
<tr>
<td>MBNL1</td>
<td>muscleblind-like 1 isoform b</td>
<td></td>
<td>0.179</td>
</tr>
<tr>
<td>IFI27L1</td>
<td>interferon, alpha-inducible protein 27-like 1</td>
<td></td>
<td>0.179</td>
</tr>
<tr>
<td>C18orf55</td>
<td>hypothetical protein LOC29090</td>
<td></td>
<td>0.178</td>
</tr>
<tr>
<td>RFX3</td>
<td>regulatory factor X3 isoform b</td>
<td></td>
<td>0.176</td>
</tr>
<tr>
<td>HBG2</td>
<td>G-gamma globin</td>
<td>(self), HBB</td>
<td>0.170</td>
</tr>
<tr>
<td>FANCA</td>
<td>Fanconi anemia, complementation group A isoform</td>
<td>(self)</td>
<td>0.169</td>
</tr>
<tr>
<td>HIST1H3C</td>
<td>histone cluster 1, H3c</td>
<td>HFE</td>
<td>0.168</td>
</tr>
<tr>
<td>TMEM206</td>
<td>transmembrane protein 206</td>
<td></td>
<td>0.166</td>
</tr>
</tbody>
</table>
Significant association with phenotype?

- Individuals with GG genotypes have **lower** hemoglobin concentration
Other studies have identified EPAS1

Sequencing of 50 Human Exomes Reveals Adaptation to High Altitude

Xin Yi,2,2a Yu Liang,2,2a Emilia Huerta-Sanchez,2a Xin Jin,3,4a Zha Xi Ping Guo,2a John E. Pool,2,2a Xun Xu,2a Hui Jiang,2a Nicolas Vinckenbosch,7 Thorfinn Sand Korneliusen,7 Hancheng Zheng,2,2a Tao Liu,2a Weiming He,2,2a Kui Li,2,2a Ruihang Luo,2a Xifang Nie,2a Honglong Wu,3,1a Meitu Zhao,2a Hongzhi Cao,2a Jinfou Zhou,2a Ying Shang,2a Shuzheng Li,2a Yi Yang,3a. A given, Pei Xiang Mi,2a, Gong Tian,2a, Jumin Xu,2a, Tao Jiang,2a, Renhua Wu,2a, Guangyu Zhou,2a, Meifang Tang,2a, Junjie Qin,2a, Tong Wang,3,4b Shijian Feng,3,4b Guohong Li,3,4b Huizing Li,3,4b Jiangbei Lu,3,4b Wei Wang,3,4b Feng Chen,3,4b Yiqing Wang,3,4b Xiaoguang Zheng,3,4b Zhouni Li,3,4b Zhizhou Buanbu,2a Ge Yang,3,4b Xinpeng Wang,3,4b Shuhui Tang,2a Guoyu Guo,2a Yong Chen,2a Zhen Luo,2a Lamu Guan,2a Zheng Cao,2a, Qinghui Zhang,2a, Weihan Ouyang,2a Xiaoli Ren,2a Huqing Liang,2a HuaSong Zheng,2a, Yebo Huang,2a, Jingxian Li,2a Lars Bolund,2a, Kirsten Kristiansen,2a, Yingu Li,2a, Yong Zhang,2a, Xingqiang Zhang,2a, Ruihang Li,2a, Songgang Li,2a, Huaming Yang,2a, Rasmus Nielsen,3,4b, Jun Wang,2a, Jian Wang,2a

Identifying Signatures of Natural Selection in Tibetan and Andean Populations Using Dense Genome Scan Data

Abigail Bigham1a, Marc Bauchet2a, Dallia Pintó3, Xianyu Mao4, Joshua M. Akey5, Rui Mei6, Stephen W. Scherer4a,1, Colleen G. Julian4, Megan J. Wilson6, David Lopez Herraez7, Tom Brutsaert8, Esteban J. Parra1, Lorna G. Moore9,10, Mark D. Shriver1

On the Origin of Tibetans and Their Genetic Basis in Adapting High-Altitude Environments

Bingbin Wang1a,2a, Yong-Biao Zhang1a, Feng Zhang1a, Hongbin Lin1a, Xumin Wang1a, Ning Wan1a, Zhengqin Ye1a, Hailu Weng1a, Lili Zhang1a, Xin Li1a, Jiangwei Yan1a, Panpan Wang2a, Tingting Wu2a, Longfei Cheng1a,2a, Jing Wang1a,2a, Duen-Mei Wang1a, Xu Ma1a,2a, Jun Yu1a

• Different data sets and different methods

Genetic Evidence for High-Altitude Adaptation in Tibet

Tamara S. Simonson,1 Yingzhong Yang,2,4 Chad D. Huff,2 Haixia Yun,2,4 Ga Qin,4 David J. Witherspoon,2 Zhendong Bai,4 Felipe R. Lorenzo,2 Jinchuan Xing,2

Lynn B. Jorde3,4, Josef T. Prchal1,2, Rilli Ge4,5

Natural selection on EPAS1 (HIF2α) associated with low hemoglobin concentration in Tibetan highlanders

Cynthia M. Beall1a,1b, Gianpiero L. Cavalleri1b, Lilibin Deng1,2c, Robert C. Elston3, Yang Gao,4 Jo Knight5,6, Chaohua Li,7, Jiang Chuan Li,8, Yu Liang4, Mark McCormack7, Hugh E. Montgomery1, Hao Pan1, Peter A. Robbins1,7, Kevin V. Shenass1, Sui Cheung Tam,1 Ngodop Tsering8, Krishna R. Veeramah9, Wei Wang4, Puchang Wang11, Michael E. Weale12, Yaomin Xu12, Zhe Xu12, Ling Yang12, M. Justin Zaman13, Changqing Zeng13,14, Li Zhang13,14, Xianglong Zhang13, Pingcuo Xia12,13, and Yong Tang Zheng13

Genetic Variations in Tibetan Populations and High-Altitude Adaptation at the Himalayas

Yi Peng1,2a, Zhuhui Yang1,2c, Hui Zhang1,2a, Chaoying Cui1,3a, Xuebin Qi1, Xiongjian Luo1, Xiang Tao4, Tianyi Wu1, Ouzhuluobu5, Basang6, Ciwangshangbu6, Danzengduoije6, Hua Chen6, Hong Shi7, and Bing Su8

A Genome-Wide Search for Signals of High-Altitude Adaptation in Tibetans

Shuhua Xu1,2a, Qilin Li3, Yaou Yang3, Jingke Tan3, Haiyi Lou3, Wenfei Jin3, Ling Yang3, Xuedong Pan3, Jiucun Wang3, Yiping Shen3, Bailin Wu3, Hongyan Wang3, and Li Jin1,2,4
What about Ethiopians?

Genotypes from 4 populations:

1. Ahmara (HA)
2. Tigrayan (HA)
3. Afar (LA)
4. Anuak (LA, outgroup)
What about Ethiopians?

Genotypes from 4 populations:

1. Ahmara (HA)
2. Tigrayan (HA)
3. Afar (LA)
4. Anuak (LA, outgroup)
What about Ethiopians?

High Altitude

Low altitude

Outgroup
A more challenging problem

1. More complex demographic history
2. Probably some admixture from a non-African group
Acknowledgements

UC Berkeley:
Benjamin Peter, Michael DiGiorgio, Nicolas Vinckenbosh, John Pool, Thorfinn Kornielsen and Rasmus Nielsen

BGI (Beijing Genomics Institute): Xin Jin, Yu Liang, Xin Yi, Zha Xi Ping Cuo, …, Jun Wang, Jian Wang and many others at

Ethiopian study:
Neil Bradman, Endashaw Bekele, Ayele Tarekegn, Luca Pagani, Peter Robbins, Mike Weale and Toomas Kivisild