Topic 9：Adaption

Group4：程心怡；黄甫保钱；陈婉桐；廖阔；申雪纯

2020.10.23

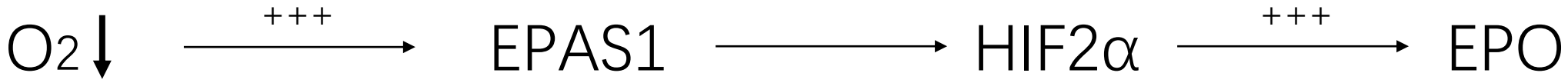
Part1

Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA

Background：Tibetans could adapt to high altitude (low oxygen)

- lower infant mortality
- higher fertility
- higher birth weight
- **limited increase in haemoglobin levels in low oxygen condition**

EPAS1：a transcription factor induced under hypoxic conditions



EPAS1 endothelial PAS domain protein 1 [*Homo sapiens* (human)]

Download Datasets

Gene ID: 2034, updated on 17-Oct-2021

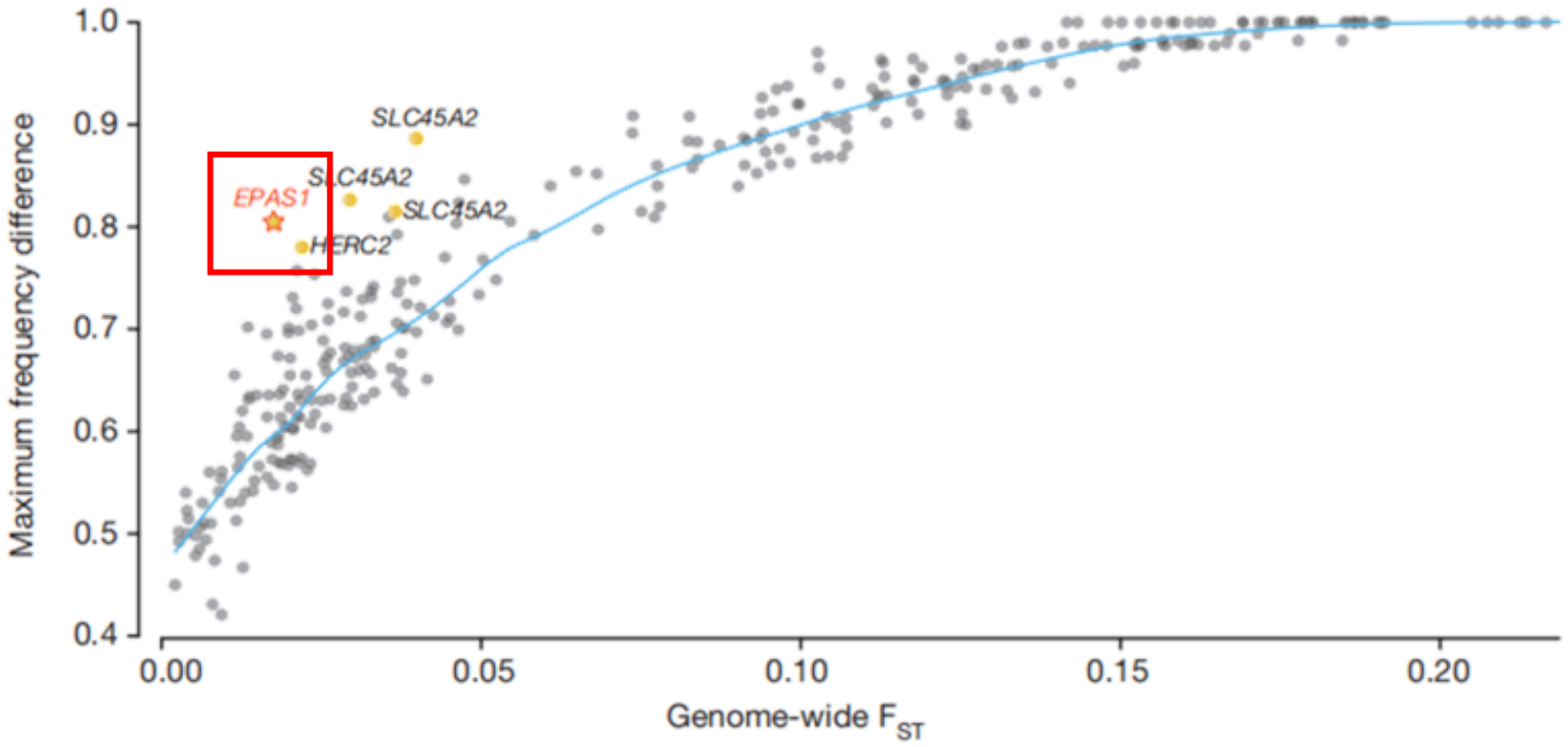
Summary

Official Symbol	EPAS1 provided by HGNC
Official Full Name	endothelial PAS domain protein 1 provided by HGNC
Primary source	HGNC:HGNC:3374
See related	Ensembl:ENSG00000116016 MIM:603349
Gene type	protein coding
RefSeq status	REVIEWED
Organism	Homo sapiens
Lineage	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo
Also known as	HLF; MOP2; ECYT4; HIF2A; PASD2; bHLHe73
Summary	This gene encodes a transcription factor involved in the induction of genes regulated by oxygen, which is induced as oxygen levels fall. The encoded protein contains a basic-helix-loop-helix domain protein dimerization domain as well as a domain found in proteins in signal transduction pathways which respond to oxygen levels. Mutations in this gene are associated with erythrocytosis familial type 4. [provided by RefSeq, Nov 2009]
Expression	Broad expression in lung (RPKM 304.3), placenta (RPKM 244.1) and 22 other tissues See more
Orthologs	mouse all

NEW **Try the new Gene table**

Try the new Transcript table

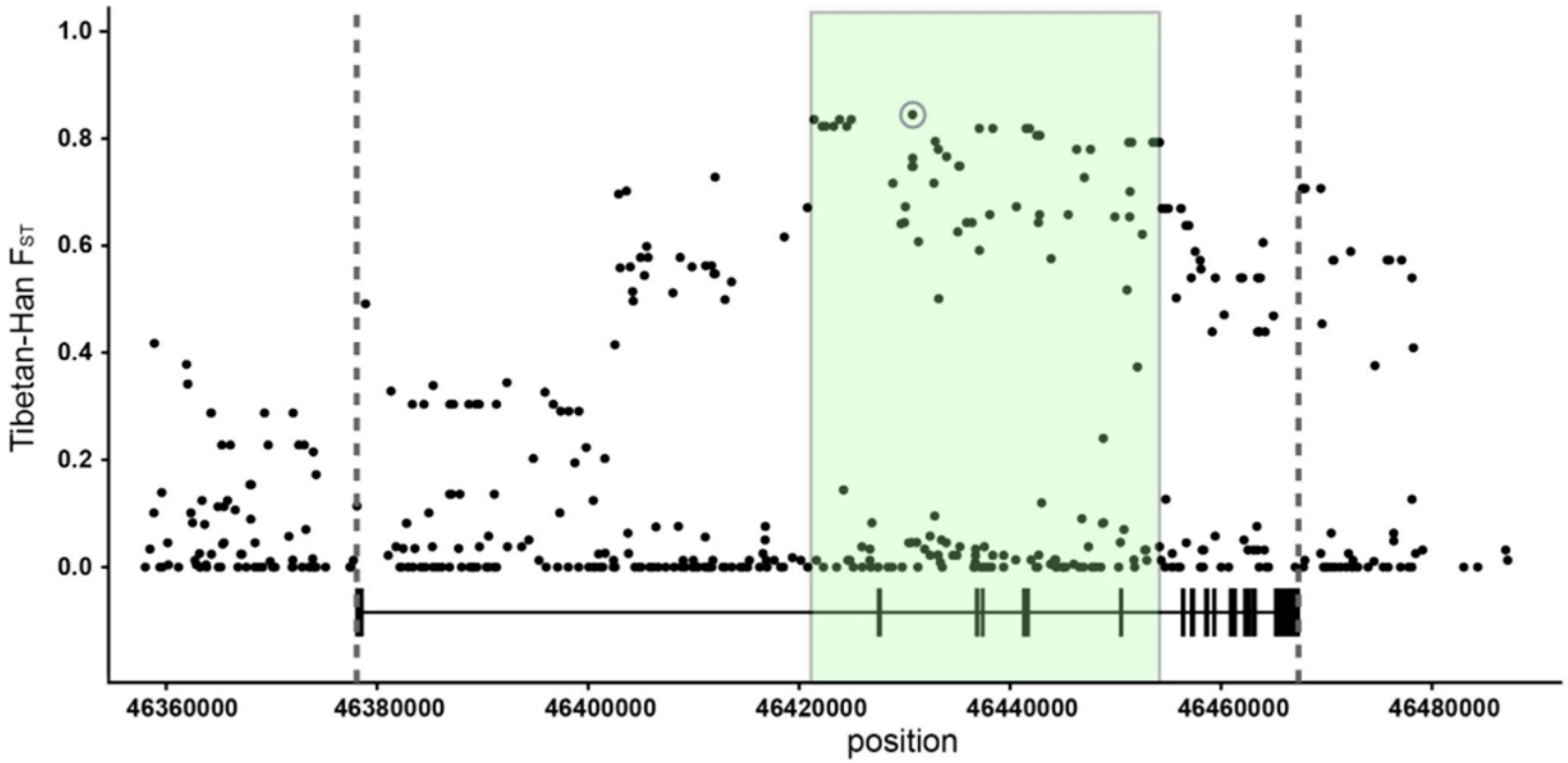
Fixation index: EPSA1 shows more differentiation between Tibetans and Han



Genome-wide F_{ST} versus maximal allele frequency difference

The relationship between genome-wide F_{ST}(x axis) computed for each pair of the 26 populations and maximal allele frequency difference (y axis)

A 32.7-kb region of EPSA1 contain much differentiated SNPs



F_{ST} calculated for each SNP between Tibetan and Han populations.

Each dot represents the F_{ST} value for each SNP in EPAS1. The x axis is the physical position in the gene. Positions are based on the hg18 build of the human genome.

The SNP variations of Tibetans shows much similarity to Denisovan than Han Chinese within the 32.7 kb region in EPSA1.

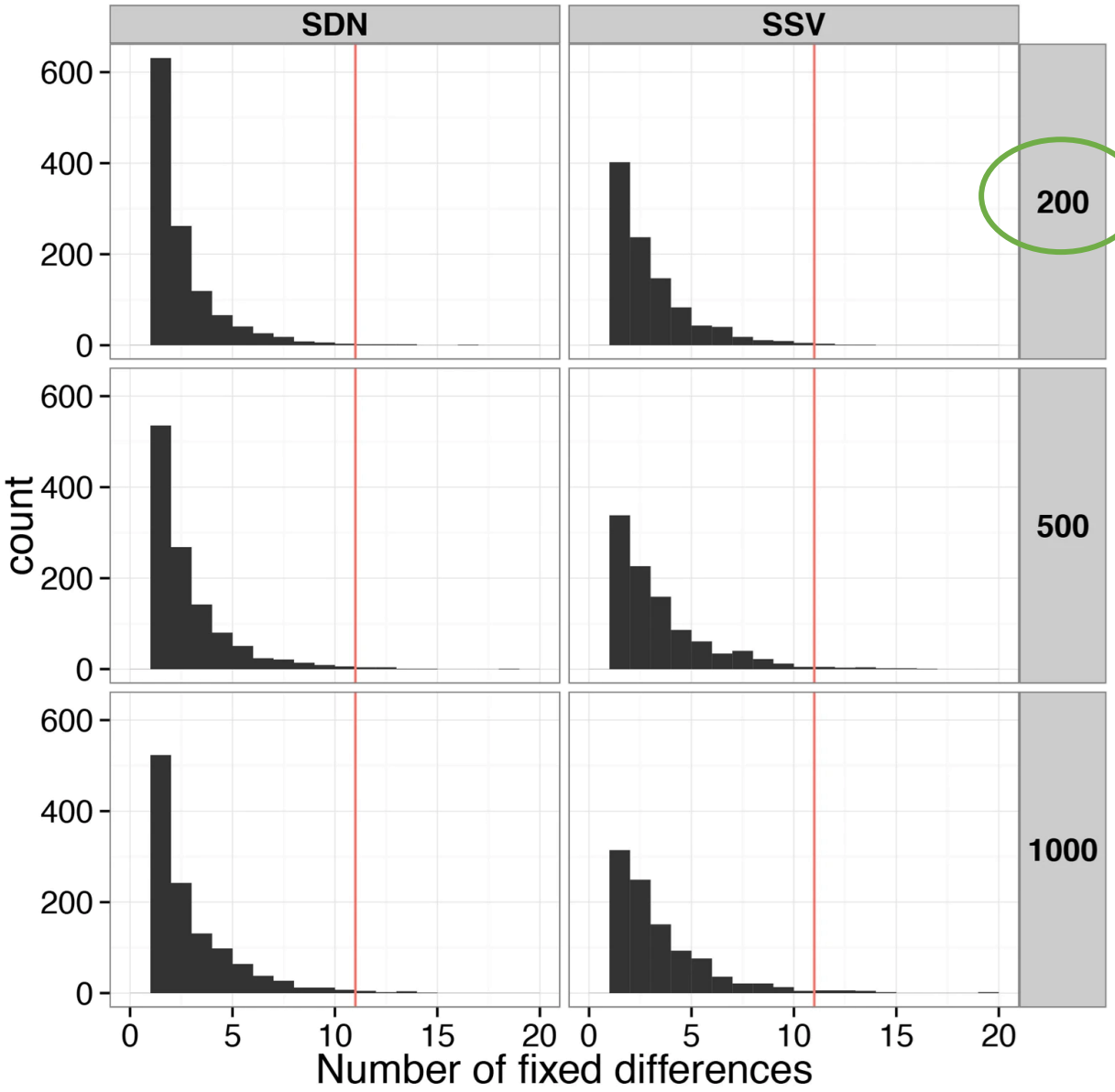
each column: one SNP
each row: one haplotype

green: Denisovan（another species live with us thousands years ago）
pink: Tibetans
yellow: Han Chinese
*：same SNP between Denisovan and Tibetans

Denisovan

natural mutation and selection?
gene introgression?

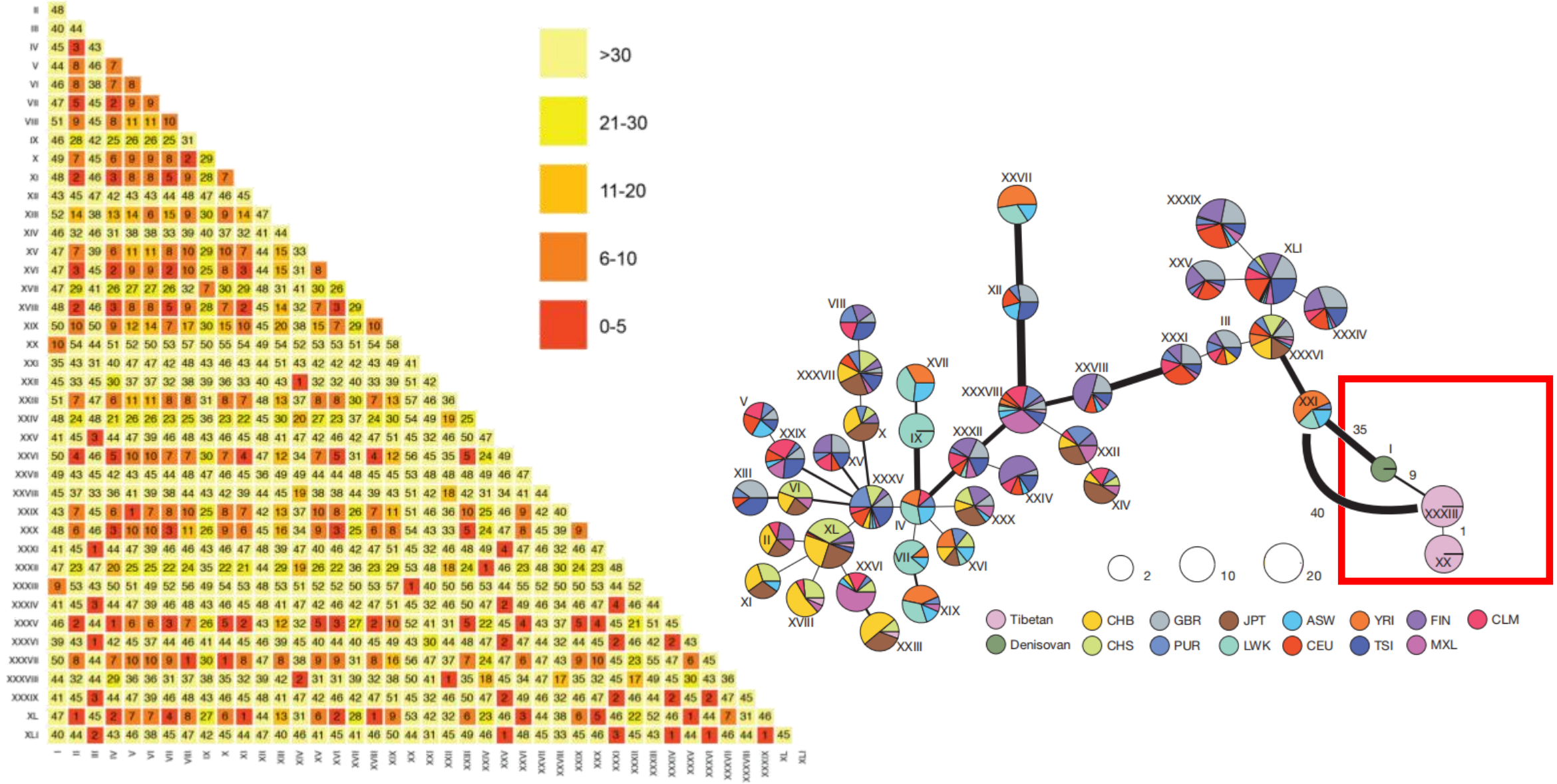
Differences in EPSA1 do not come from occasionally mutations and nature selection



strength of selection

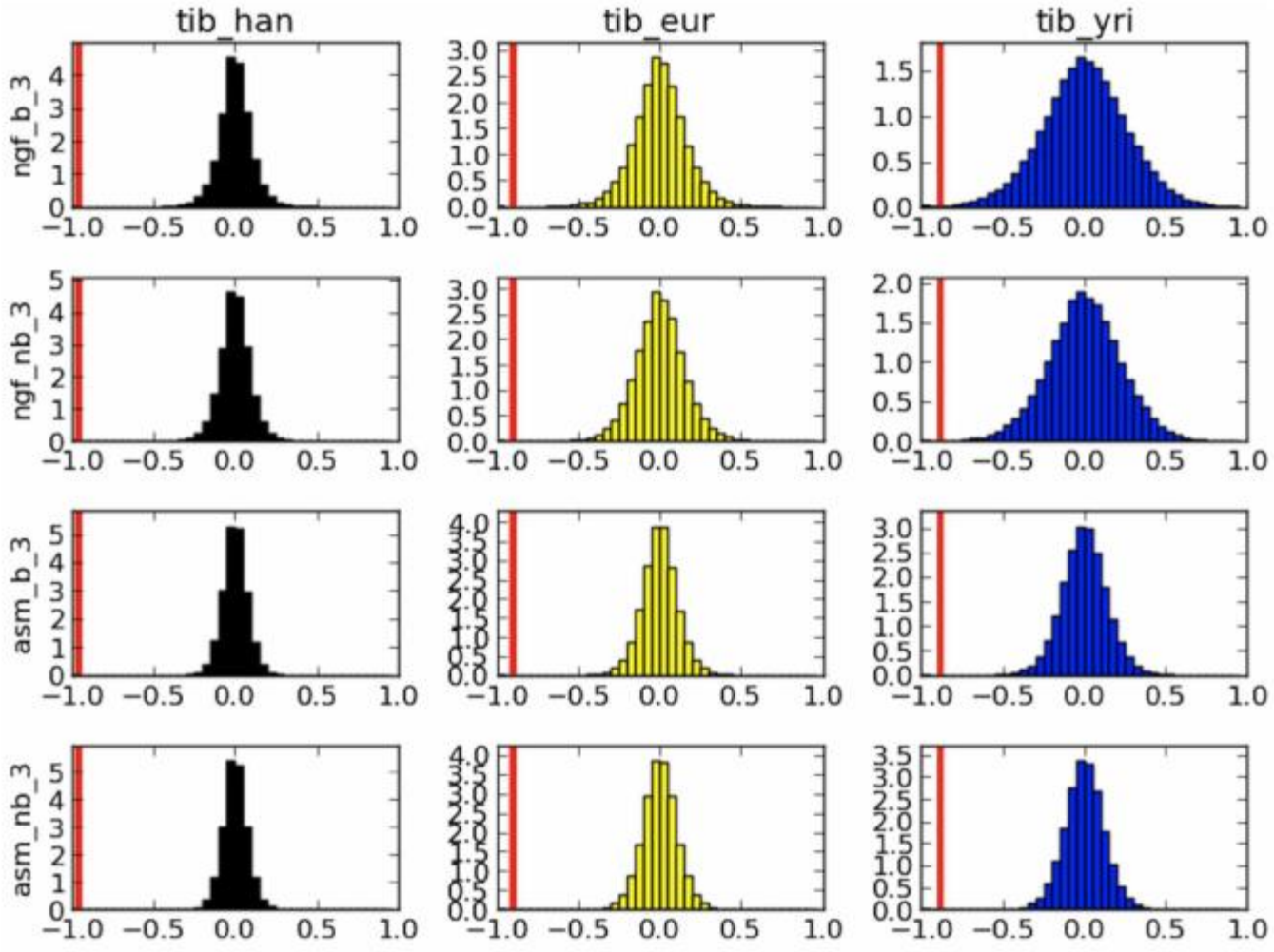
SDN: selection on de novo mutation mode
SSV: selection of standing variation
red line: mark of real data

Haplotype network shows Tibetan population much close to Denisovan than other population containing EPSA1



A haplotype network based on the number of pairwise differences between the 40 most common haplotypes.

P-values for D(TIB, CEU, Den, Chimp) under the four models simulated



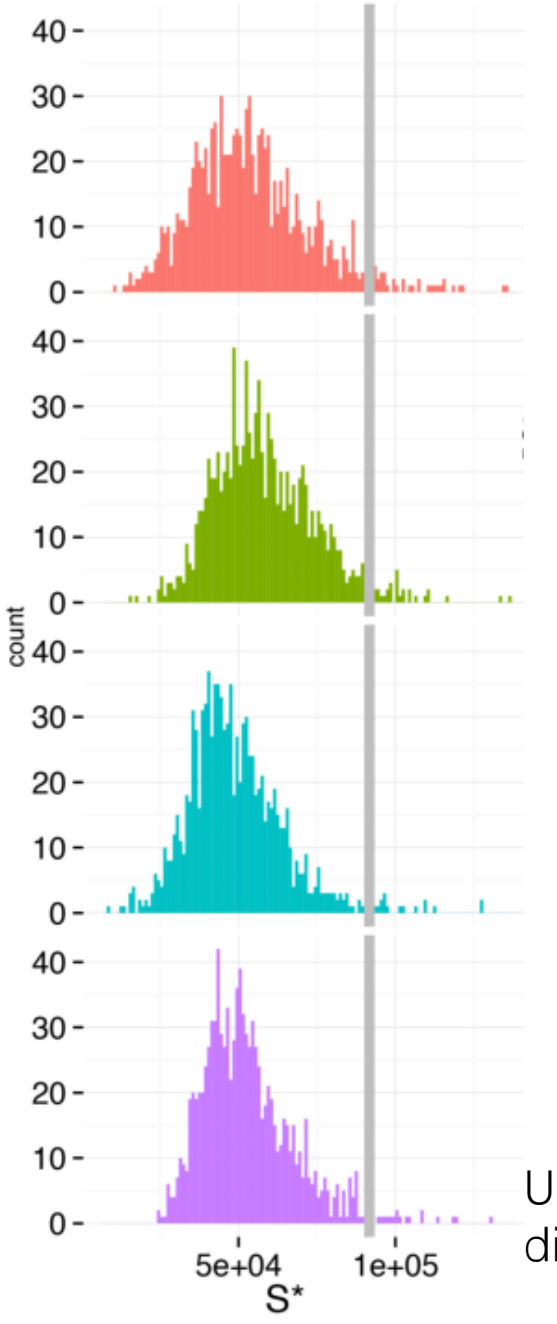
Model:
asm_b： ancestral structure and a bottleneck
asm_nb: ancestral structure and no bottleneck
ngf_b： no ancestral structure and with a bottleneck
ngf_nb: no ancestral structure and no bottleneck

assumed divergence time
of Tibetans and Han

3,000y divergence	TIB-CHB	TIB-CEU	TIB-YRI
Ngf_b	0.00016	0.00024	0.00088
Ngf_nb	0.00009	0.00014	0.00045
Asm_b	<1e-5	<1e-5	<1e-5
Asm_nb	<1e-5	<1e-5	<1e-5

The observed D values are significant ($P < 0.001$)

Null distributions of S* statistics under models of no gene flow



ancestral structure with bottleneck

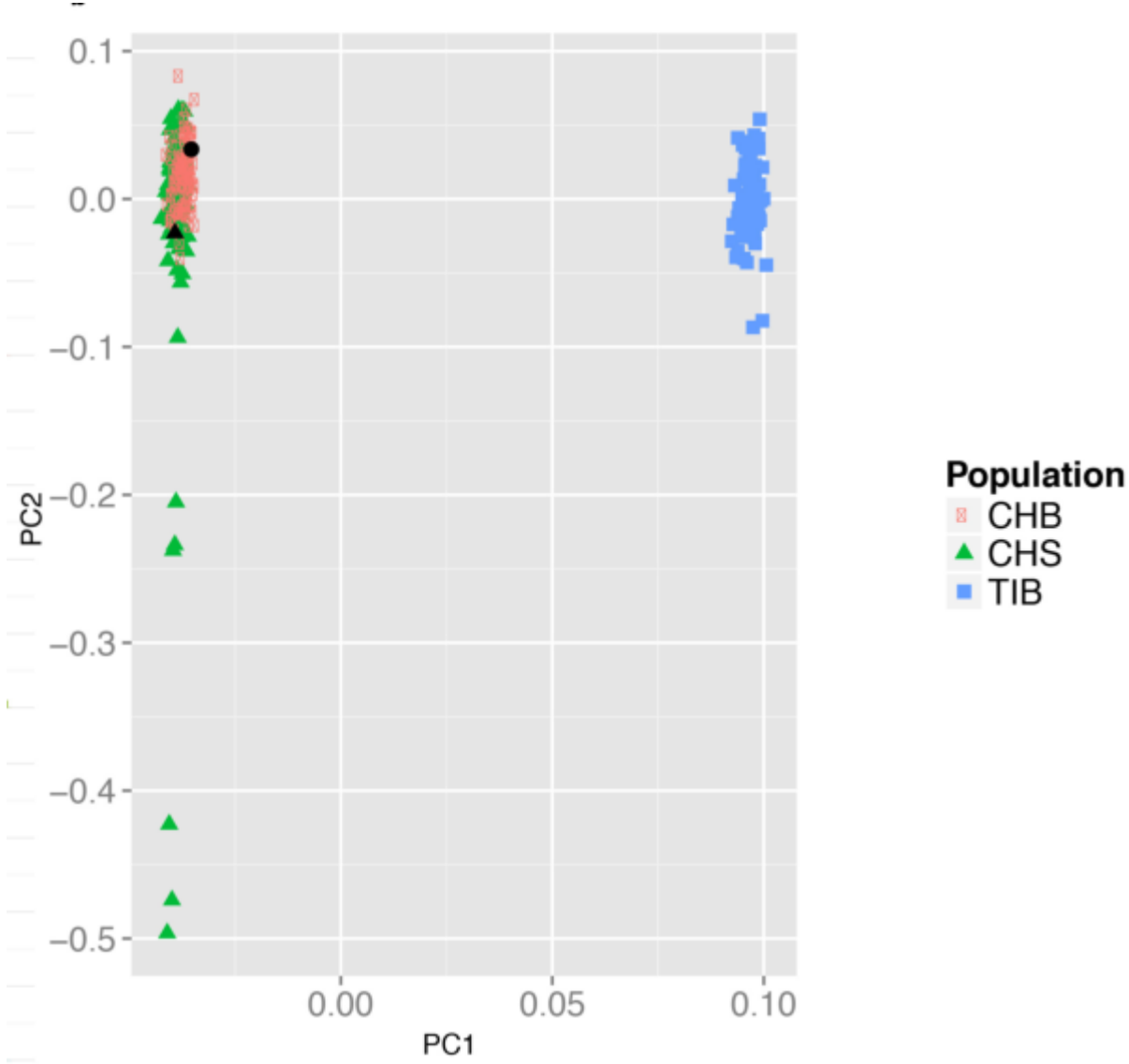
ancestral structure without bottleneck

with bottleneck

without bottleneck

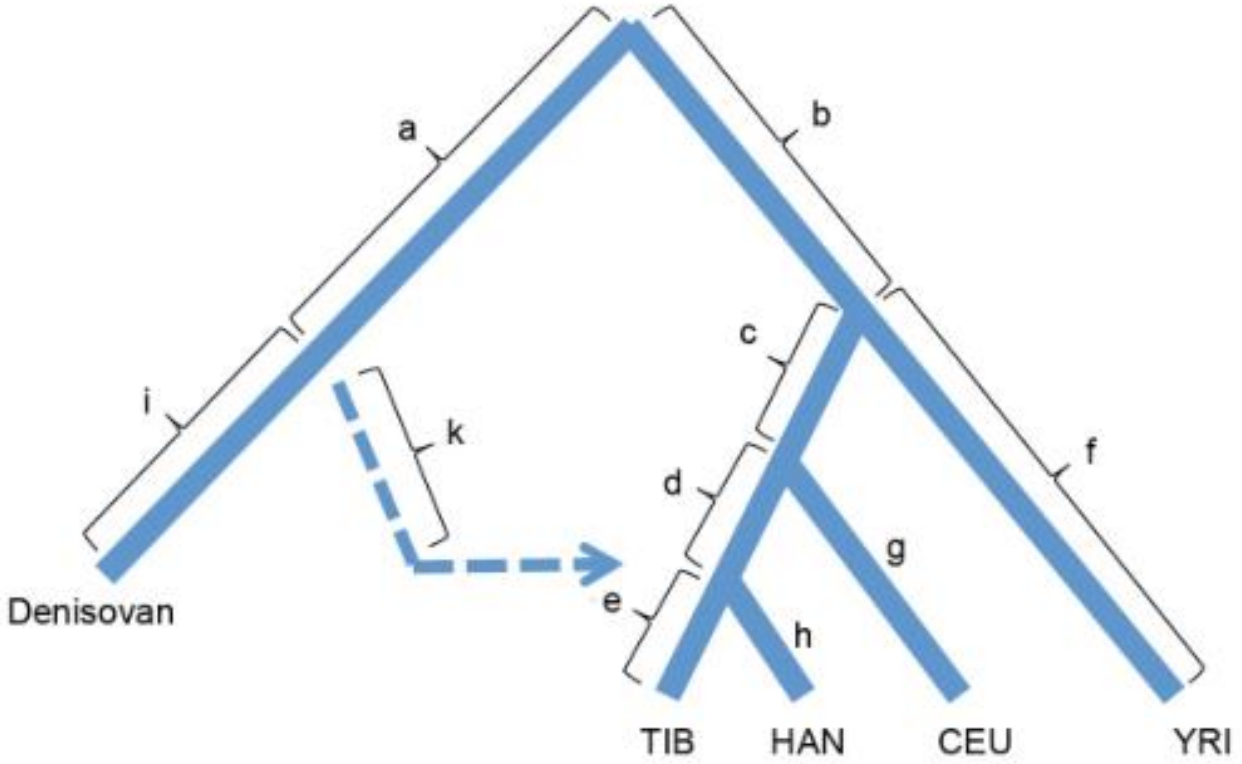
Under all four models, S* is significantly different from the null distribution with all the empirical p-values lying below 0.035.

The CHB individual with a Denisovan-like haplotype in EPAS1 is not a descendant of a recent immigrant from Tibet



All the CHB and the CHS individuals cluster together and principal component 1 clearly separates Tibetans from CHB and CHS individuals.

Conclusion：Introgression of Denisovan DNA into Tibetans cause adaptation to altitude



$$T^{Tib-YRI} = a + k + b + f + e$$

$$T^{Denisovan-YRI} = a + i + b + f$$

Illustration of the genealogical structure in a model with gene flow from Denisovans to Tibet

Part2 | The evolution of human haemoglobin in structure and function

nature

Article | Published: 20 May 2020

Origin of complexity in haemoglobin evolution

Arvind S. Pillai, Shane A. Chandler, Yang Liu, Anthony V. Signore, Carlos R. Cortez-Romero, Justin L. P. Benesch, Arthur Laganowsky, Jay F. Storz, Georg K. A. Hochberg & Joseph W. Thornton

THE UNIVERSITY OF CHICAGO

Department of Ecology & Evolution

Natural molecular complex

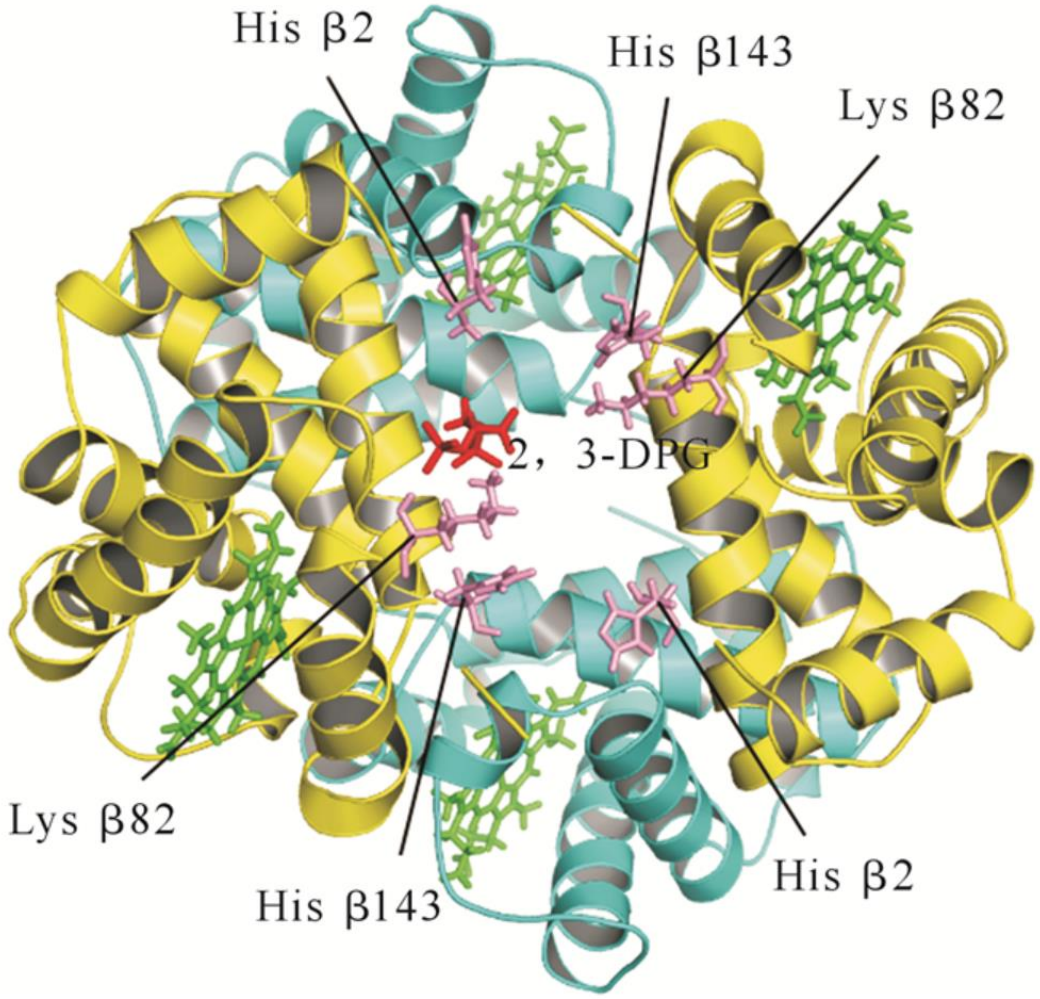
- interfaces <- sterically tight, electrostatically complementary interactions
- allostery & cooperativity <- residues that connect surfaces to active sites

How did it evolve?

classical explanation:

- multimerization (new functions)
- selection-driving accumulation of substitutions

There's no accurate description of the evolution of any natural molecular complex ever before, and a detailed reconstruction of the historical steps by which it evolved is lacked.

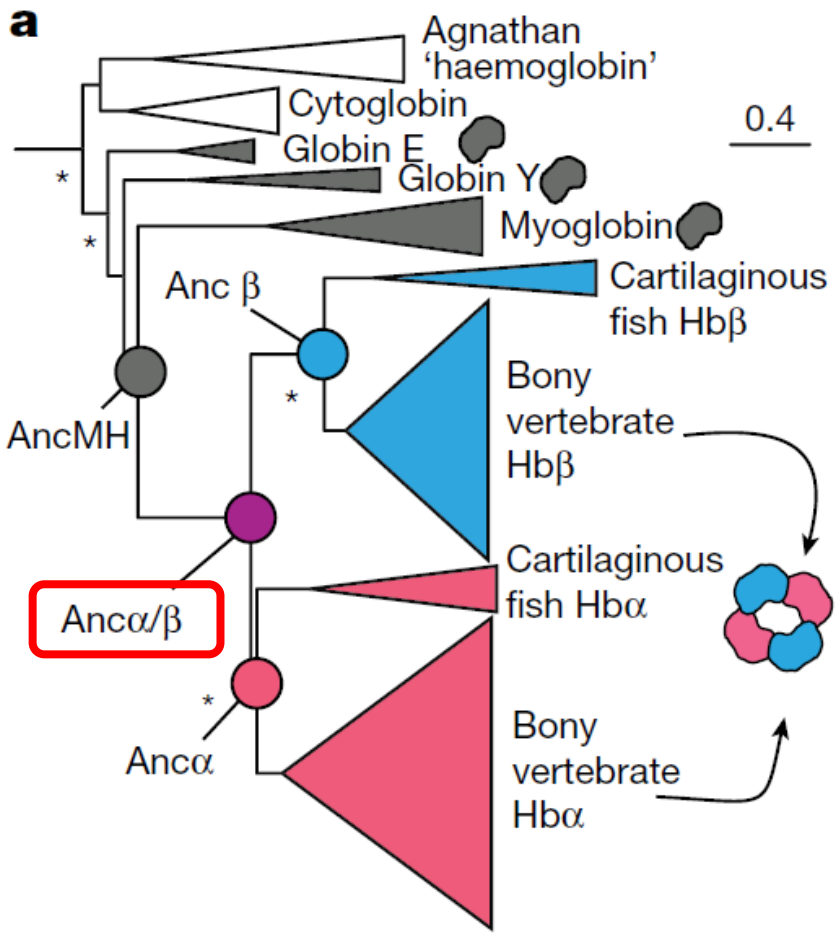


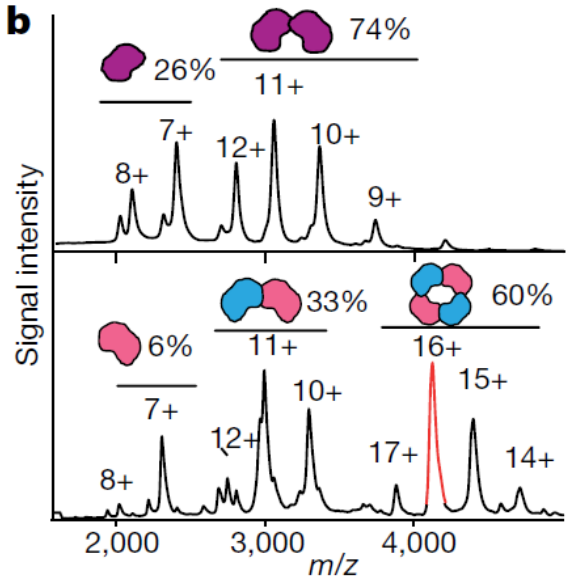
Why Haemoglobin (Hb)?

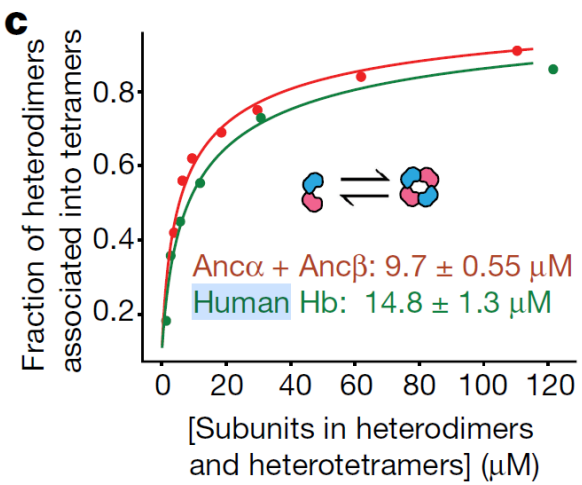
- The structural mechanisms that mediate Hb's multimeric assembly, cooperative oxygen binding, and allosteric regulation are **well established**.

- Hb's subunits descend by duplication and divergence from **the same ancestral proteins**, so their history can be reconstructed in a single analysis.

From monomer to homodimer —— the existence of the Ancα/β







Data: 177 annotated amino acid sequences of Hb and related paralogues in 72 species.

Methods: ML reconstruct + protein express & purify.

- AncMH: monomeric
- Ancα/β: homodimers
- Ancα: homodimers
- Ancβ: homotetramers
- Ancα+Ancβ: heterotetramers

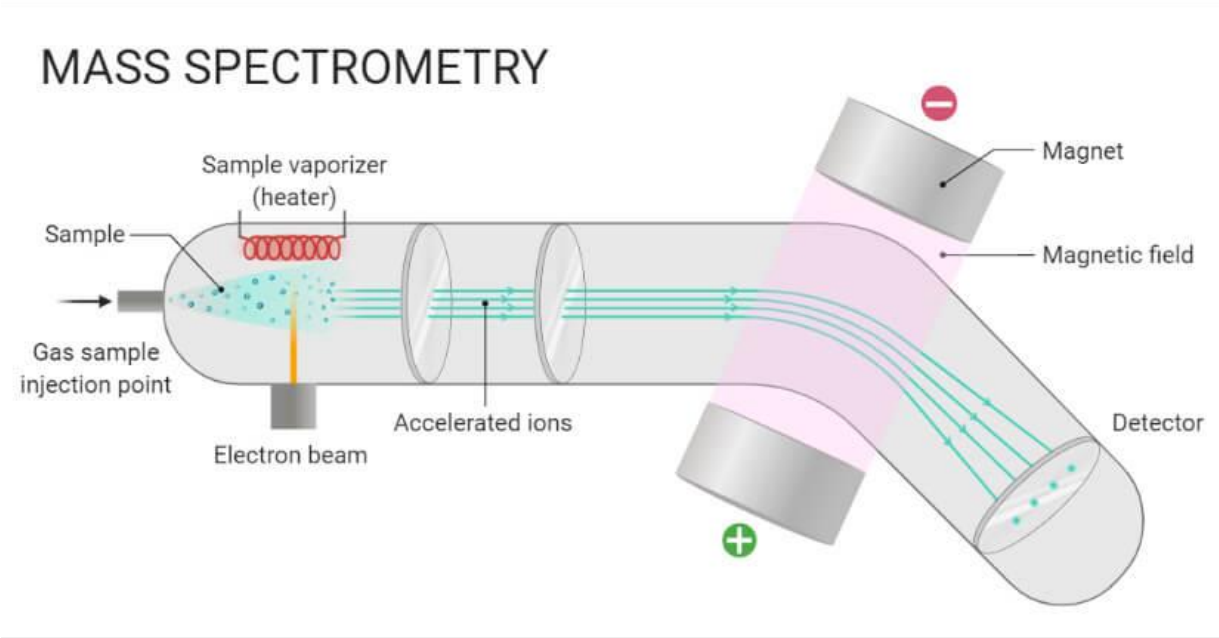
} Like extant Hb

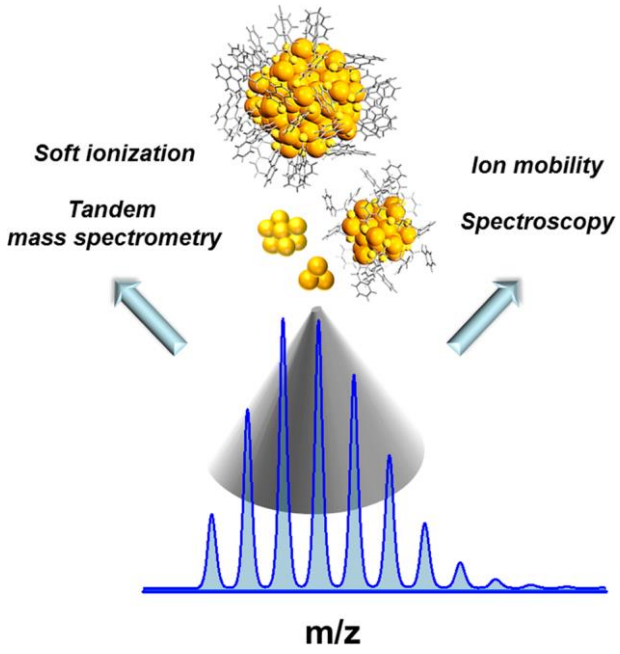
The Ancα/β homodimer is the evolutionary missing link between monomer and heterotetramer.

Supplementary: nMS theory

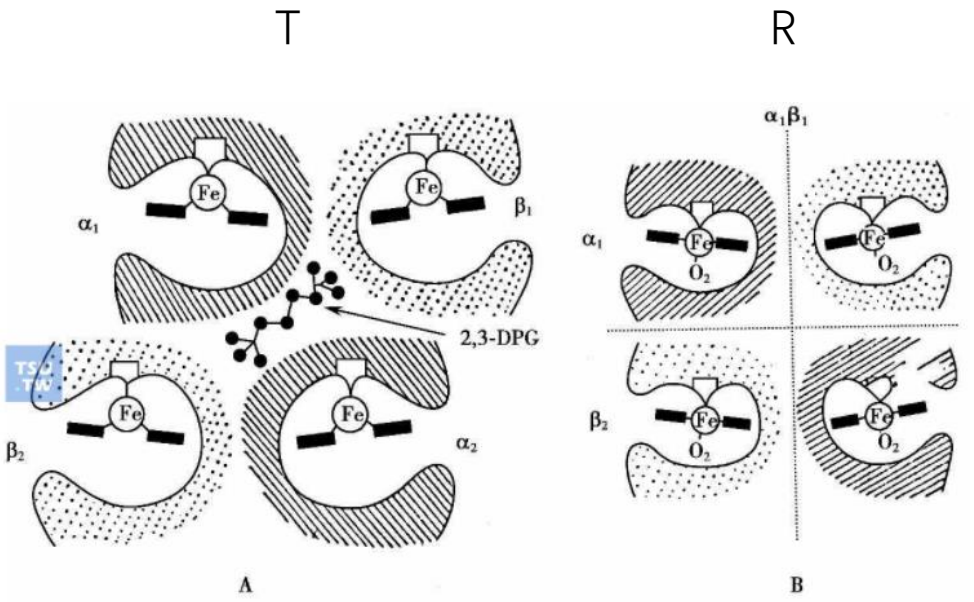
Native mass spectrometry (nMS) allows the topological investigation of intact protein complexes with high sensitivity and a theoretically unrestricted mass range. DOI: 10.1038/nmeth.1265.

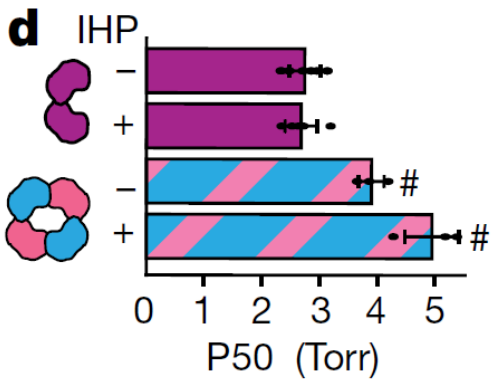
Components: an ion source, a mass analyzer, and a detector.





From monomer to homodimer —— the existence of the Ancα/β





e

	IHP	*n*	
α/β	–	1.02 ± 0.24	
	+	1.19 ± 0.14	
α+β	–	1.29 ± 0.08	*
	+	1.52 ± 0.07	*

T: low affinity

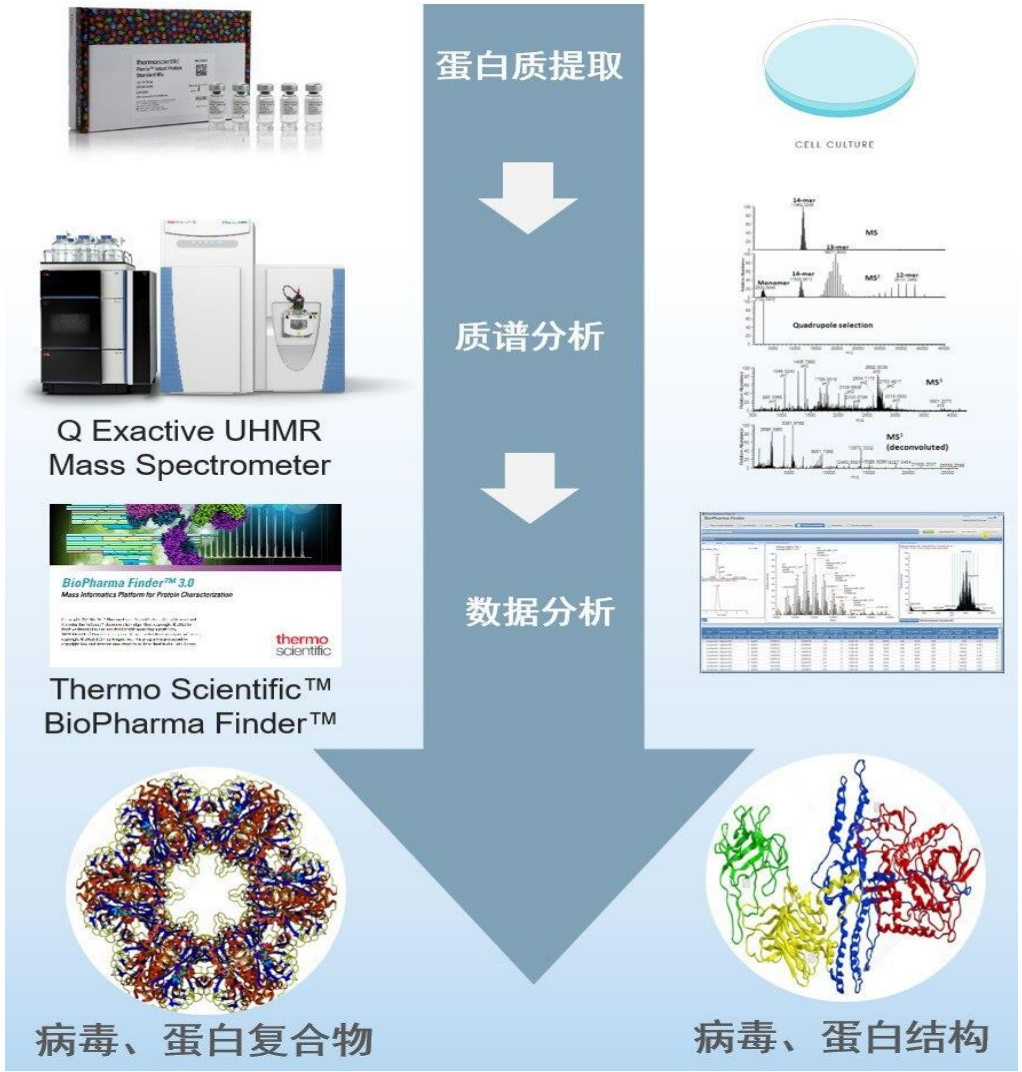
R: high affinity

CO_2 or IHP: R -> T, affinity ↓

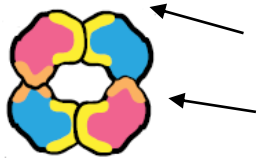
Ancα+Ancβ oxygen affinity (P50) and cooperativity (n) changed by IHP, indicating allosteric regulation.

How did the heterotetramers evolve?

Ancestral and derived interfaces



Two interfaces of Hb



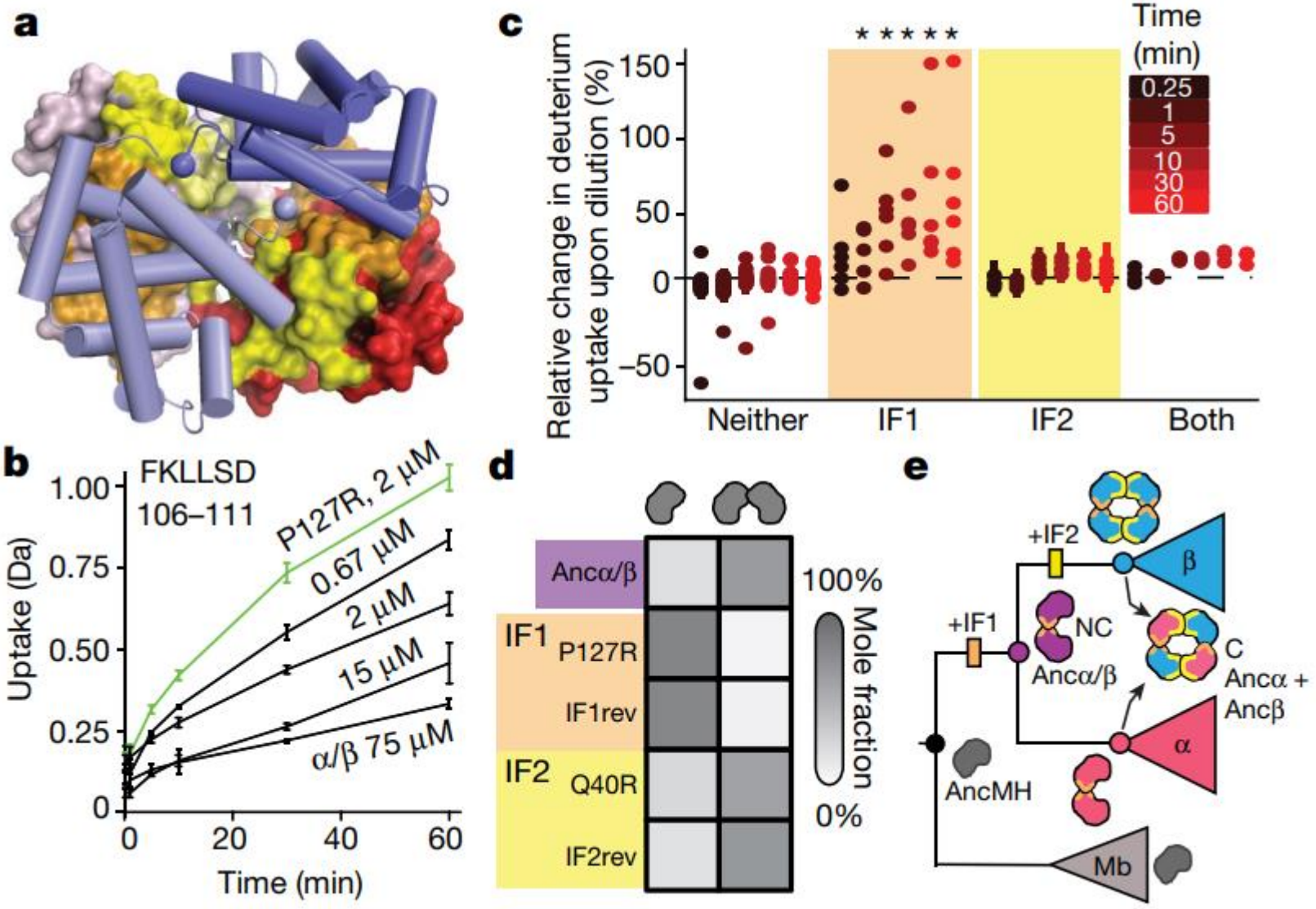
IF1: α1–β1, α2–β2

IF2: α1–β2, α2–β1

Technology

HDX MS (hydrogen deuterium exchange mass spectrometry) is a mass spectrometry technique for studying the spatial conformation of protein. Protein is immersed in heavy water solution, and hydrogen atoms of protein will be exchanged with deuterium atoms of heavy water.

Ancestral and derived interfaces

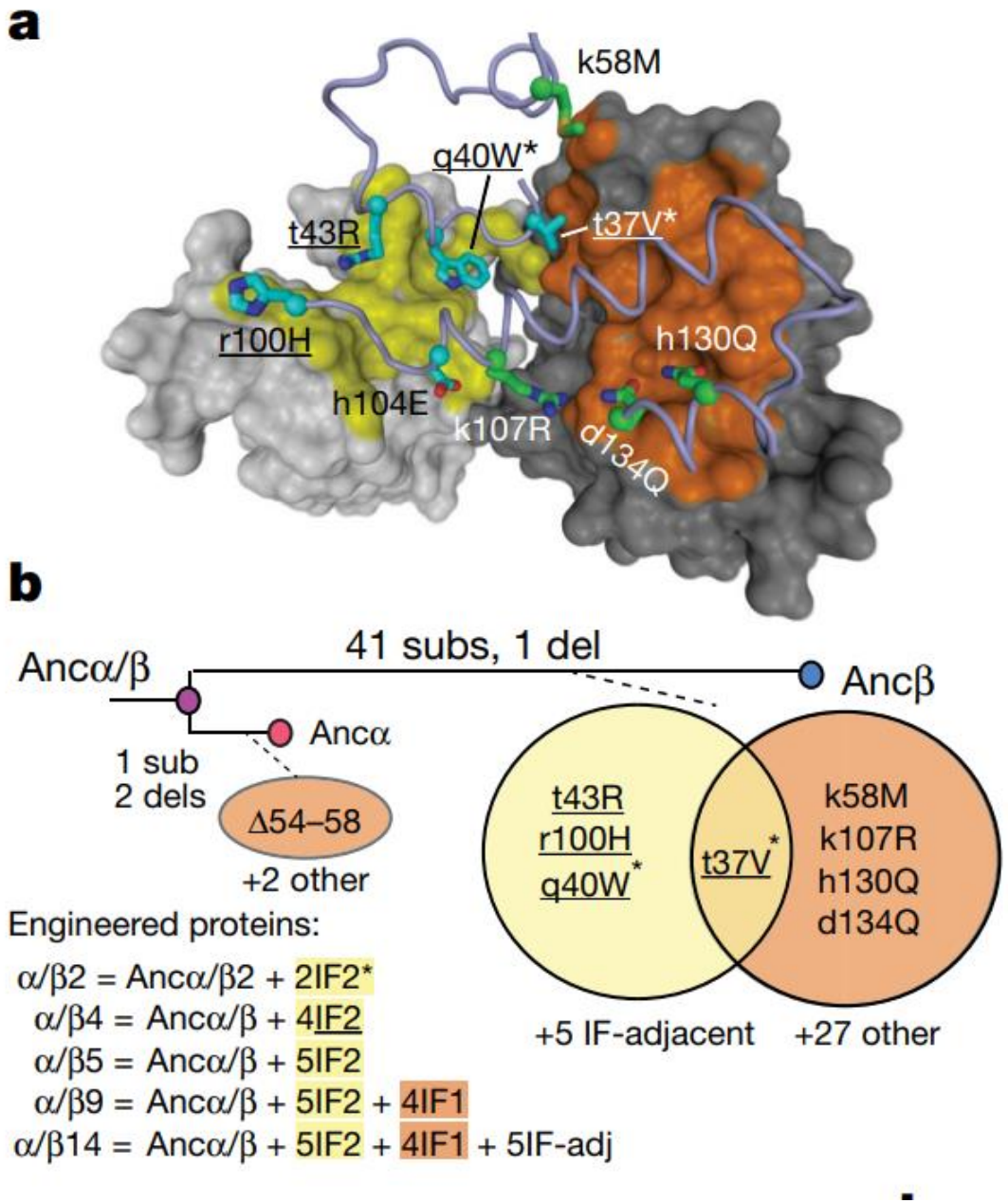


The interfaces in Ancα/β exists in a form of IF1-like.

Therefore, we can learn that the Hb first evolved IF1, then IF2.

Ancα/β homodimers therefore assembled via IF1. After duplication, IF2 evolved, enabling dimers to assemble into tetramers

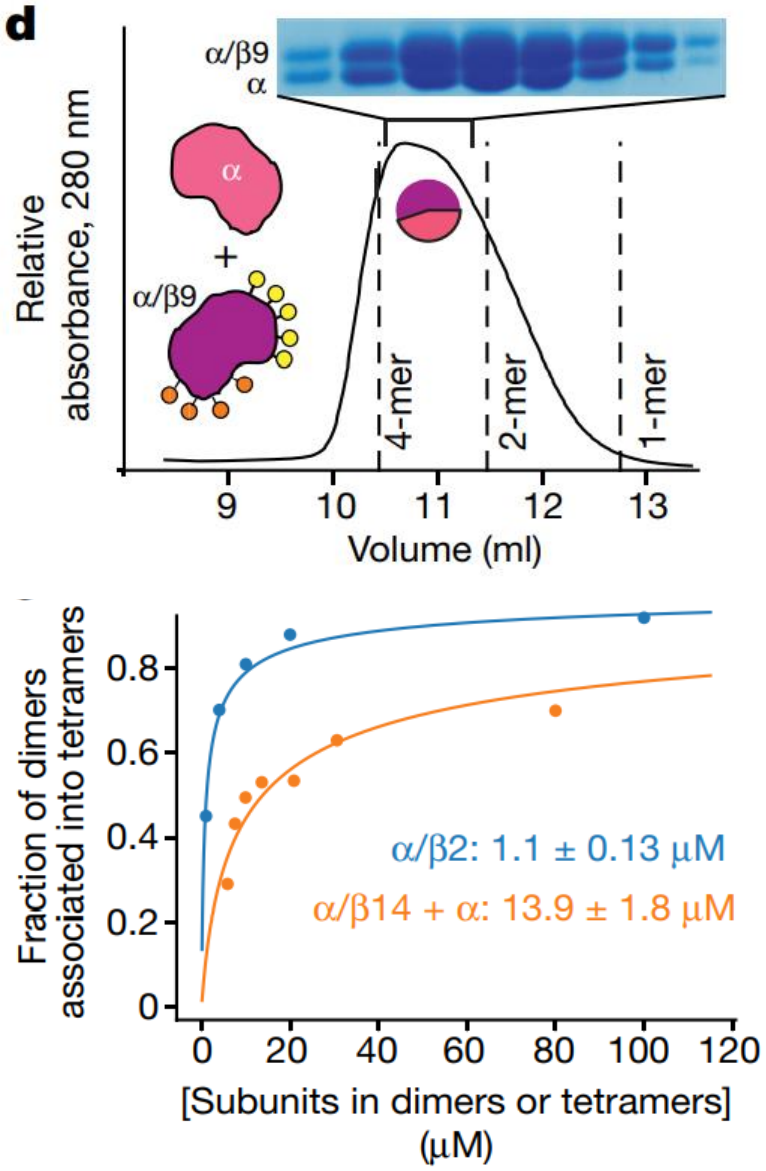
Genetic mechanisms for the new interface

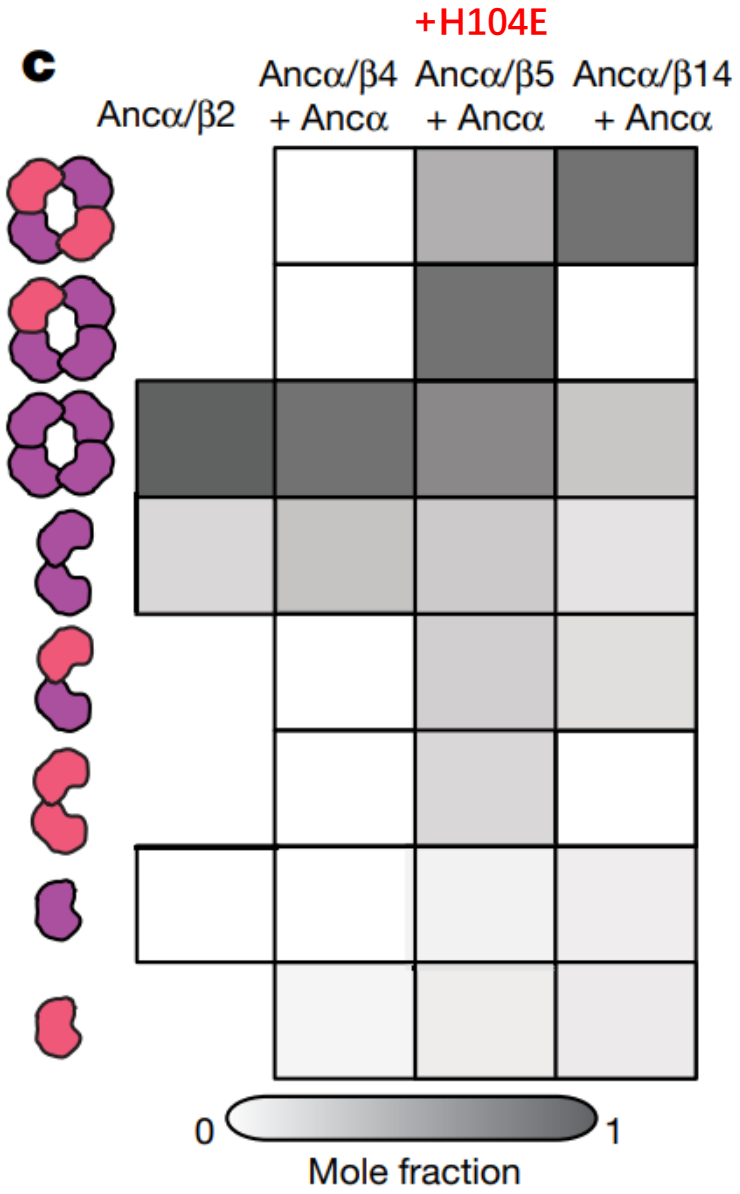


- Ancα branch: only 3 changes, of which none were at IF2.
- Ancβ branch: 42 changes, including 5 at IF2 and 4 others at IF1.

nMS test

Genetic mechanisms for the new interface

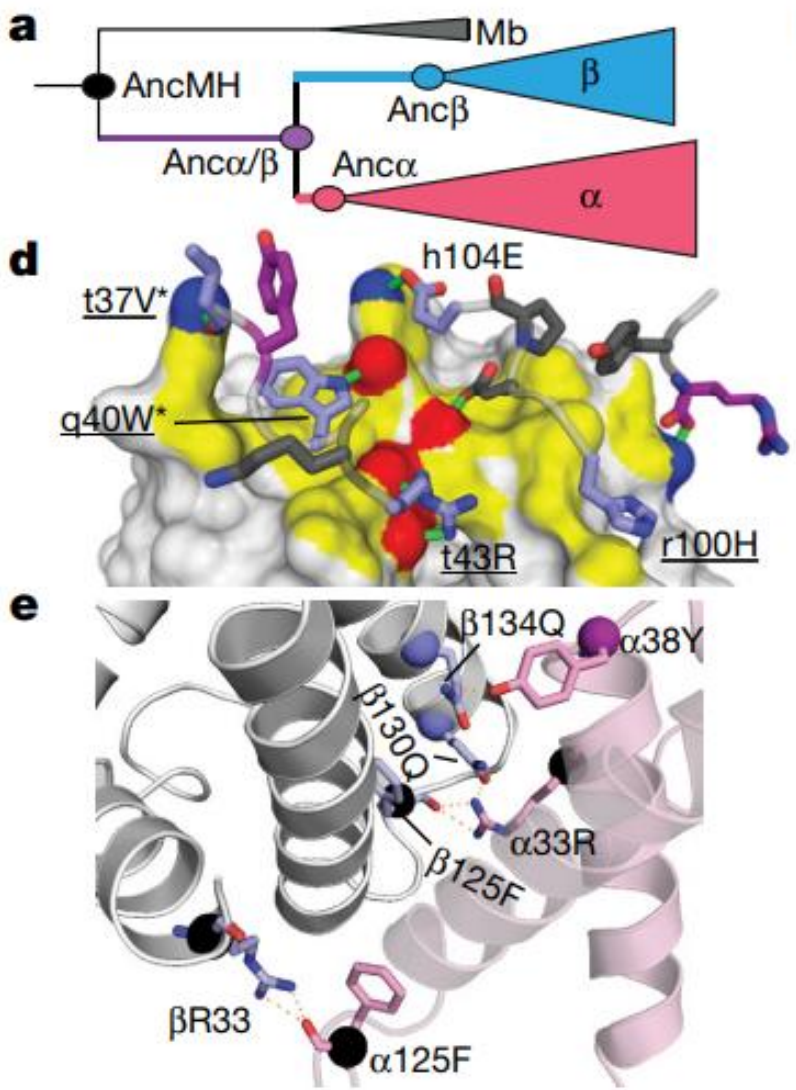


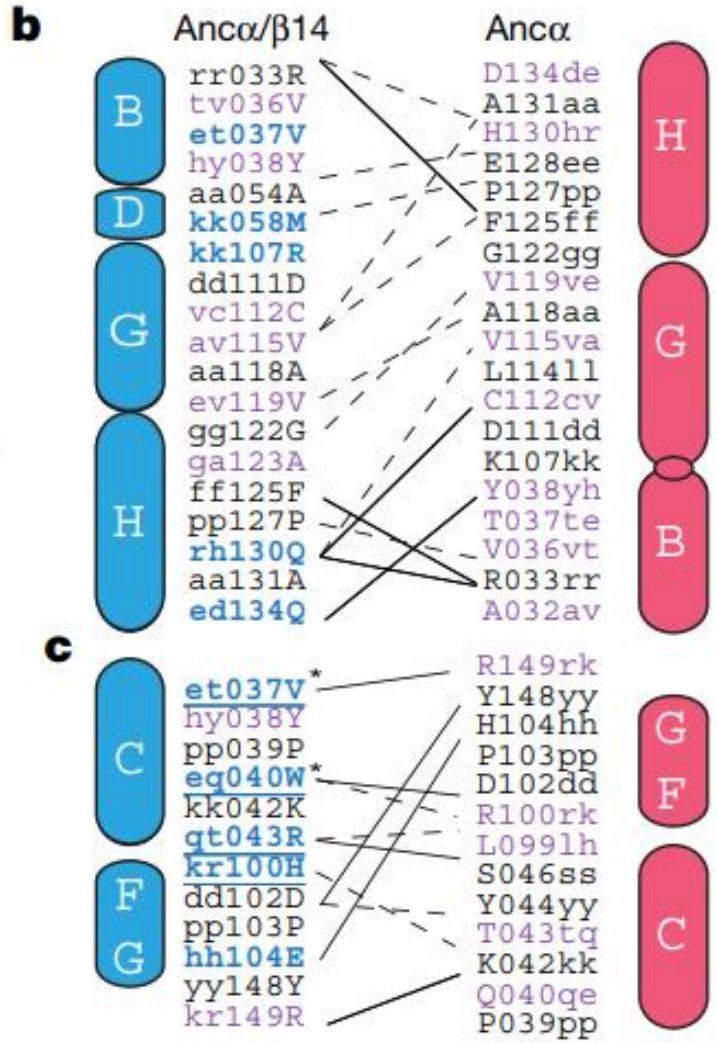


α/β2 = Ancα/β2 + 2IF2*
α/β4 = Ancα/β + 4IF2
α/β5 = Ancα/β + 5IF2
α/β9 = Ancα/β + 5IF2 + 4IF1
α/β14 = Ancα/β + 5IF2 + 4IF1 + 5IF-adj

- Changes at **IF2** created a strong new interface that conferred **tetramerization**;
- Changes at **IF1** yielded **heterospecificity**.
- In both cases, only a few substitutions were required.

Structural mechanisms for the new interface





Contact maps for residues buried at IF1 and IF2 of Ancα + Ancβ

Amino acids

black: conservative from AncMH to Ancα or Ancα/β14.

red: started to change from Ancα/β period

blue: changed only recently.

Interactions

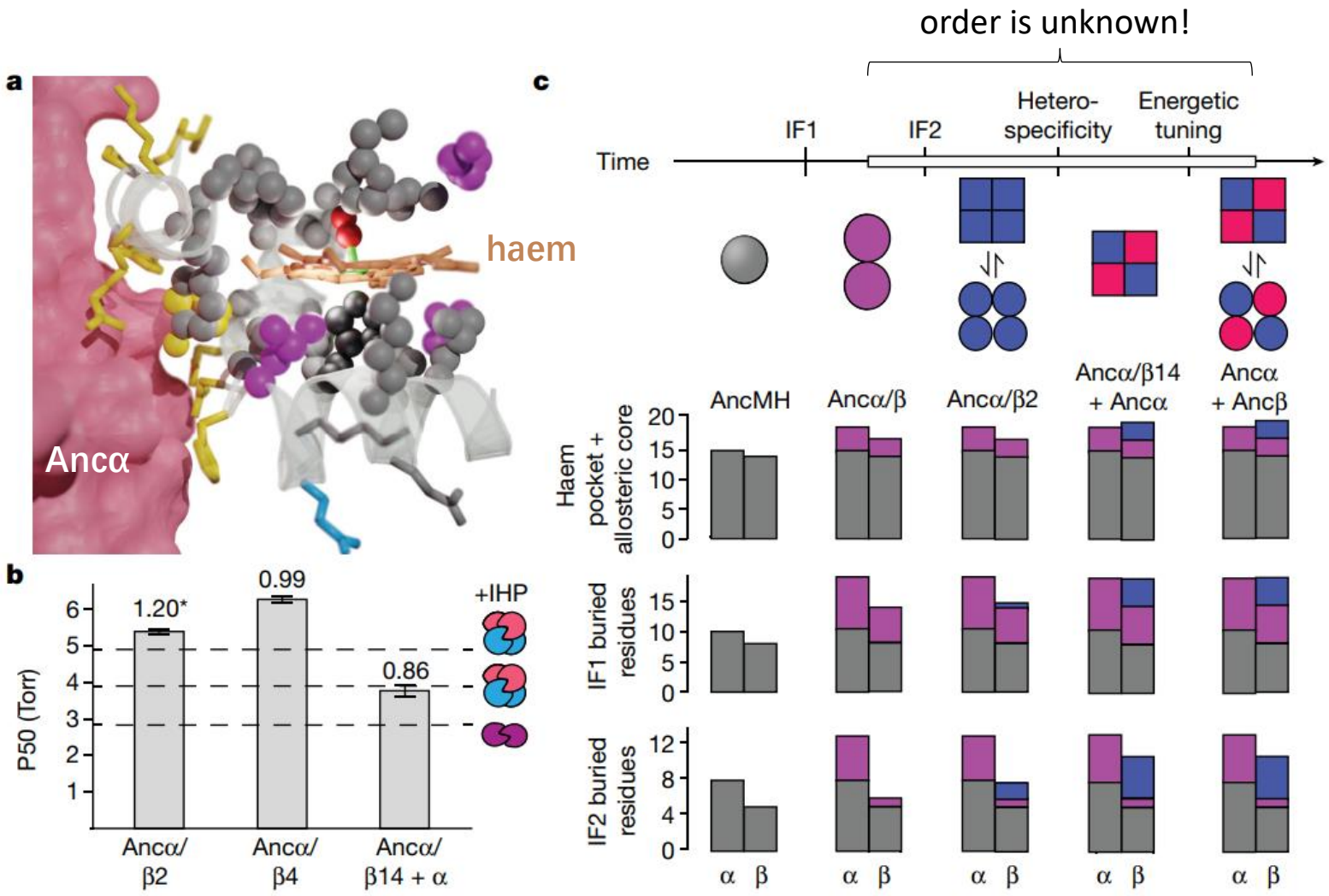
solid line: hydrogen bond

dotted line: van der Waals force.

(containing old & new substitutions!)

Mechanisms of cooperativity

Examined the phylogenetic history of residues in the **haem pocket** and **allosteric core**.



The structural properties that mediate the allosteric linkage already existed in Ancα/β.

When did the cooperativity emerge?

May have immediately generated (Ancα/β2);

May have evolved via a low-affinity tetrameric intermediate (Ancα/β4).

Conclusions

Evolution of natural molecular complex

Traditional view：

- Long history of functional optimization
- Natural selection

New view showed by Hb reconstruction：

- Just a few substitutions can emerge new complex and functions
- Neutral mutation

Thank you!

Q & A