

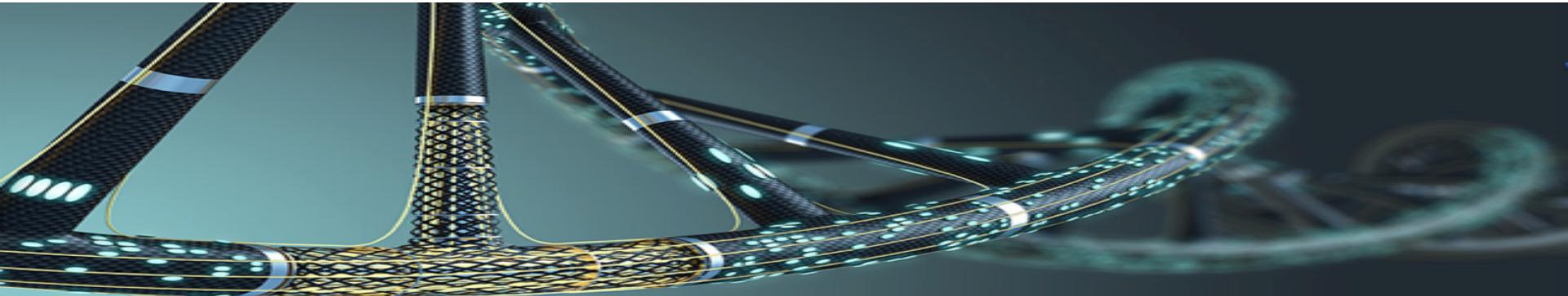


Майнер по биоинформатике

Семестр 2

Лекция 11

Мария Попцова



Chimpanzee (*Pan troglodytes*)

Chimpanzee (*Pan troglodytes*)

As recently as five million years ago, humans and chimps shared a common ancestor. A chimp possesses approximately 20,000 to 25,000 genes... approximately the same number as humans. Of that total, there are only 50 human genes that have no homolog in chimps. Differences between the two species are, therefore, due more to changes in gene regulation. Identifying genes that are divergent between humans and chimps should be helpful in understanding disease susceptibility. For example, chimps do not suffer from some human diseases such as AIDS and malaria. A comparison of the two genomes may reveal that there are genetic reasons for this and result in new ways of treating and preventing human disease.



iStockphoto.comaree

2005

First chimpanzee genome

the common chimpanzee (*Pan troglodytes*)

Article | Published: 01 September 2005

Initial sequence of the chimpanzee genome and comparison with the human genome

The Chimpanzee Sequencing and Analysis Consortium

Nature 437, 69–87(2005) | Cite this article

14k Accesses | 1348 Citations | 372 Altmetric | Metrics

Richard Dawkins : Comparing the Human and Chimpanzee Genomes



<https://www.youtube.com/watch?v=WMPlr4tD64A>

Figure 1: Human-chimpanzee divergence in 1-Mb segments across the genome.

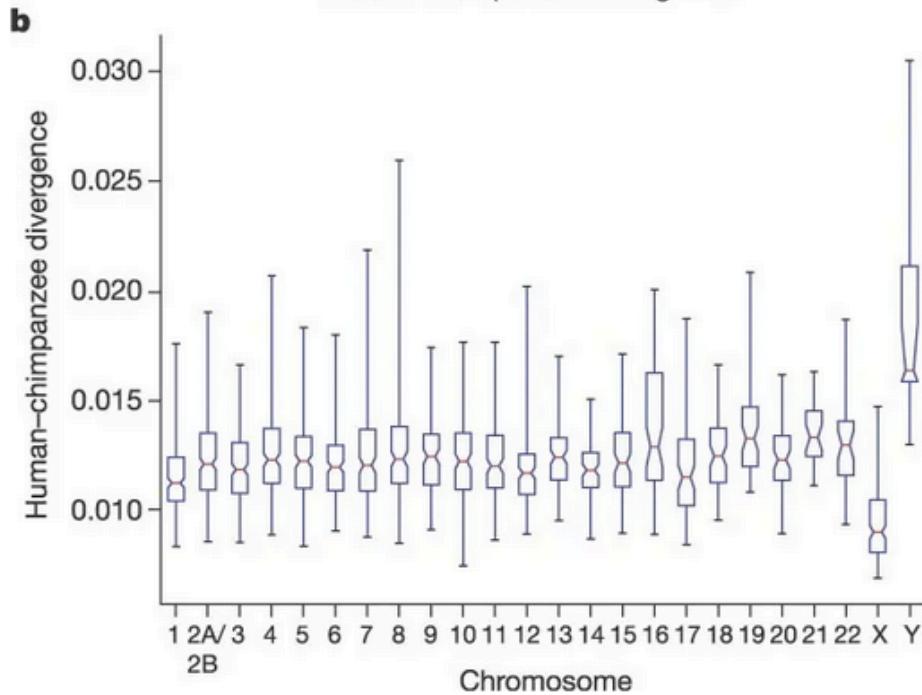
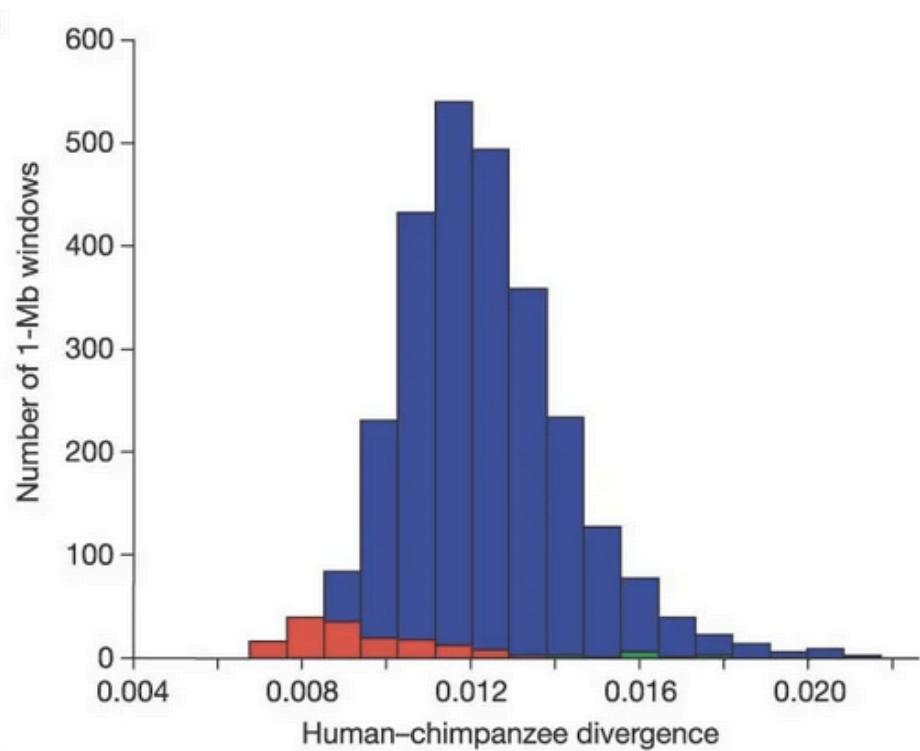
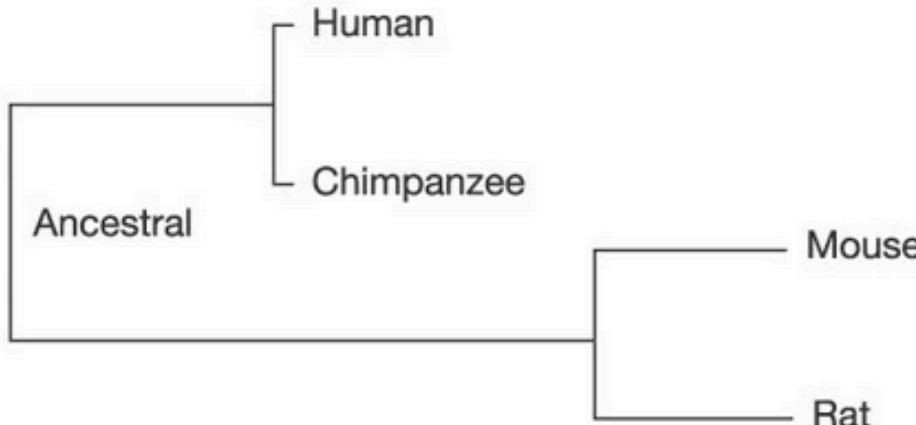


Figure 9: Human–chimpanzee–mouse–rat tree with branch-specific K_A/K_S (ω) values.

a



b

	Hominid		Ancestral	Murid	
	Human	Chimpanzee		Mouse	Rat
K_A	0.00128	0.00117	0.0629	0.0128	0.0133
K_S	0.00617	0.00602	0.427	0.0902	0.0972
K_A/K_S (ω)	0.208	0.194	0.147	0.142	0.137

Evolutionary tree. The branch lengths are proportional to the absolute rates of amino acid divergence. **b**, Maximum-likelihood estimates of the rates of evolution in protein-coding genes for humans, chimpanzees, mice and rats. In the text, ω_{hominid} is the K_A/K_S of the combined human and chimpanzee branches and ω_{murid} of the combined mouse and rat branches. The slight difference between ω_{human} and $\omega_{\text{chimpanzee}}$ is not statistically significant; masking of some heterozygous bases in the chimpanzee sequence may contribute to the observed difference (see Supplementary Information ‘Gene evolution’).

Table 4 Rapidly diverging gene clusters in human and chimpanzee

Location (human)	Cluster	Median K_A/K_I^*
1q21	Epidermal differentiation complex	1.46
6p22	Olfactory receptors and HLA-A	0.96
20p11	Cystatins	0.94
19q13	Pregnancy-specific glycoproteins	0.94
17q21	Hair keratins and keratin-associated proteins	0.93
19q13	CD33-related Siglecs	0.90
20q13	WAP domain protease inhibitors	0.90
22q11	Immunoglobulin-λ/breakpoint critical region	0.85
12p13	Taste receptors, type 2	0.81
17q12	Chemokine (C-C motif) ligands	0.81
19q13	Leukocyte-associated immunoglobulin-like receptors	0.80
5q31	Protocadherin-β	0.77
1q32	Complement component 4-binding proteins	0.76
21q22	Keratin-associated proteins and uncharacterized ORFs	0.76
1q23	CD1 antigens	0.72
4q13	Chemokine (C-X-C motif) ligands	0.70

*Maximum median K_A/K_I if the cluster stretched over more than one window of ten genes.

Table 6 GO categories with accelerated divergence rates in hominids relative to murids

GO categories within 'biological process'	Number of orthologues	Amino acid divergence in hominids	Amino acid divergence in murids	K_A/K_S in hominids	K_A/K_S in murids
GO:0007283 spermatogenesis	43	0.0075	0.054	0.323	0.188
GO:0006869 lipid transport	22	0.0081	0.051	0.306	0.120
GO:0006865 amino acid transport	24	0.0058	0.033	0.218	0.084
GO:0015698 inorganic anion transport	29	0.0061	0.027	0.195	0.072
GO:0006486 protein amino acid glycosylation	50	0.0056	0.040	0.166	0.100
GO:0019932 second-messenger-mediated signalling	58	0.0049	0.036	0.159	0.083
GO:0007605 perception of sound	28	0.0052	0.033	0.158	0.085
GO:0016051 carbohydrate biosynthesis	27	0.0047	0.028	0.147	0.067
GO:0007268 synaptic transmission	93	0.0040	0.025	0.126	0.069
GO:0006813 potassium ion transport	65	0.0035	0.022	0.113	0.056

2012

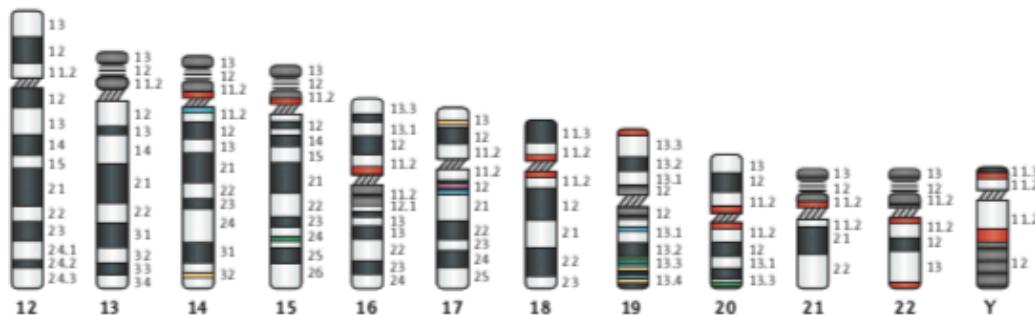
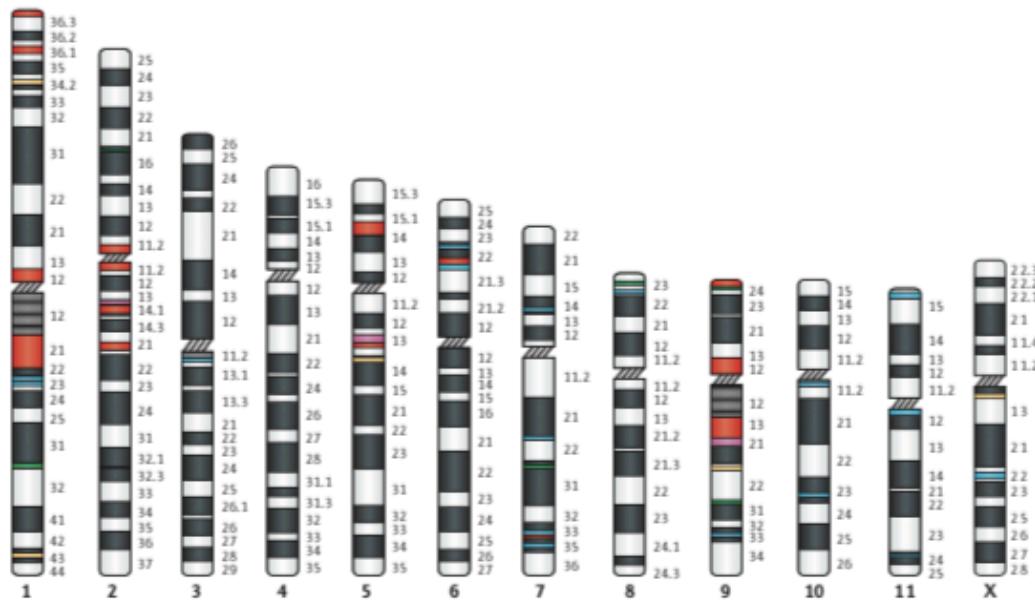
Review Article | Published: 16 November 2012

Evolution of genetic and genomic features unique to the human lineage

Majesta O'Bleness, Veronica B. Searles, Ajit Varki, Pascal Gagneux & James M. Sikela 

Nature Reviews Genetics 13, 853–866(2012) | Cite this article

600 Accesses | 77 Citations | 23 Altmetric | Metrics



■ Copy number change	■ Positive selection	■ Novel gene variant	■ Pseudogene	■ Expression change
---	---	--	--	---

Forces Shaping the Fastest Evolving Regions in the Human Genome

Katherine S Pollard , Sofie R Salama, Bryan King, Andrew D Kern, Tim Dreszer, Sol Katzman, Adam Siepel, Jakob S Pedersen, Gill Bejerano, Robert Baertsch, Kate R Rosenbloom, Jim Kent, David Haussler

Published: October 13, 2006 • <https://doi.org/10.1371/journal.pgen.0020168>

We found 202 genomic elements that are highly conserved in vertebrates but show evidence of significantly accelerated substitution rates in human. These are mostly in non-coding DNA, often near genes associated with transcription and DNA binding.

- HARs are short—on average just 227 base pairs long, much smaller than a gene.
- the combined list of identified HARs now includes nearly 3,000 genome segments.
- nearly all HARs are outside genes, some quite far away from any gene in the genome.

Uniquely human gene regulators

- HAR regions are some of the most conserved sequences in the genomes of mammals.
- HAR1 -> non-coding RNA (будет на семинаре)
 - Шимпанзе-курица 2 замены
 - Шимпанзе-человек 18 замен
- HAR2 (*HACNS1*) -> enhancer in embryo morphogenesis

5 percent of HARs function as noncoding RNAs, while most are enhancers that control gene expression during embryonic development

UNDERSTANDING HUMAN ACCELERATED REGIONS

Sections of the genome that are largely conserved across mammals and even the entire animal kingdom, but differ in humans, are known as human accelerated regions (HARs). Deciphering their function may prove key to understanding what sets humans apart from other organisms. For example, 2xHAR.142 and 2xHAR.114, like many other HARs, function as enhancers, which increase or decrease the level of a gene's expression.



...GCGTAGAACATGAAAGATTCAAGATCAATGTACTCCCCATTTCATAG
ACGTCTATTAGTAACGAT...



...GCACAGAACATGAAAGATTCAAGATCAATGTACTCCCCAGTTCATAG
ATATGCTATTAGTAATAAT...

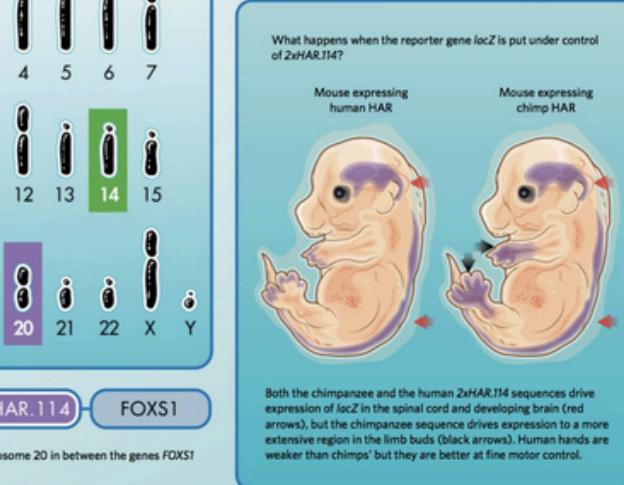
© NOVÉE STUDIO, 2006. ALL RIGHTS RESERVED

HAR FACTS

Location: Typically outside genes, though some HARs are found in gene introns

Size: Just 227 base pairs long, on average

Function: Most HARs studied so far are enhancers, sequences of DNA that increase or decrease the expression of a gene.

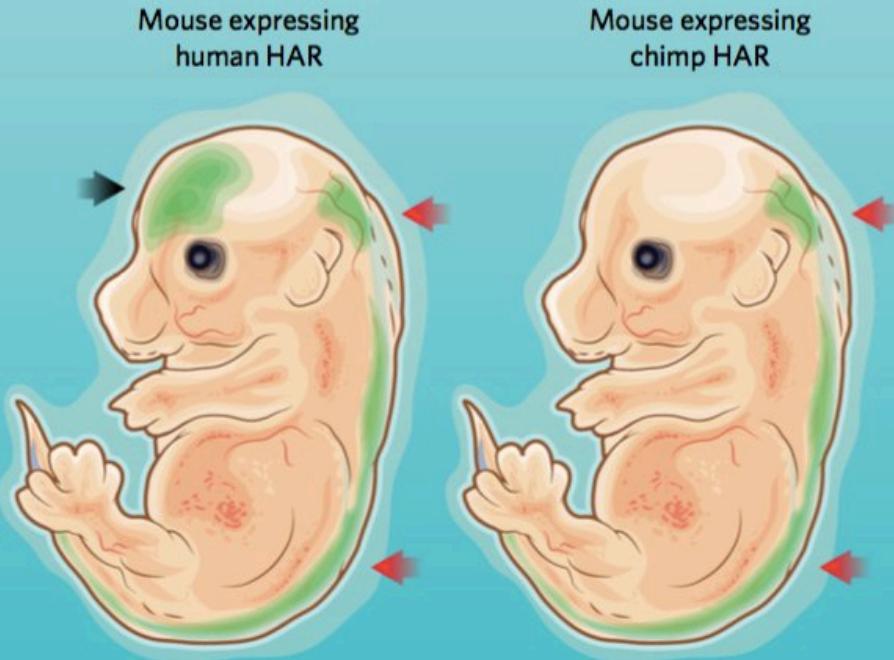


...CACAGGTCTGGAAGCCACTAAGCCCACATCTGGTTGGATTACA
TCAGGGCTGGTGACACTGCCTTCCCTTCTGGGTCCAGTGGCCTG
TATTCCACTGGCACCTGGACCCAATAA...



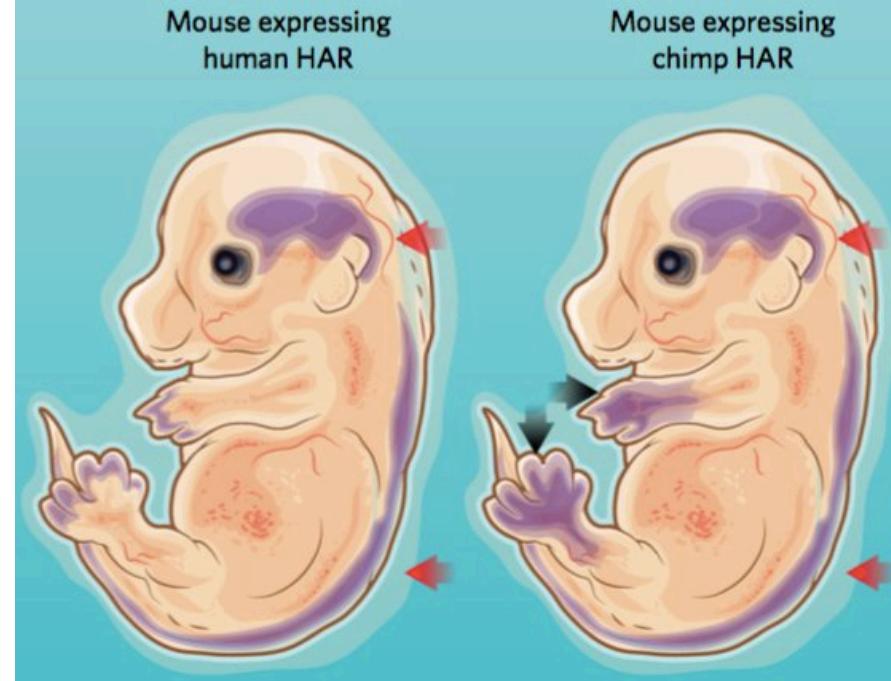
...CATGGGTCTGGAAGCCACTAAGGCCACATCTGGTTGGATTACAT
CAGGGCTGGTGACGCTGCCTTCCCTGCTGGGTCCAGTGGCCTG
ATTCCACTGGCACCTGGACCCAACA...

What happens when the reporter gene *lacZ* is put under control of *2xHAR.142*?



Both the chimpanzee and human forms of *2xHAR.142* enhance the expression of *lacZ* in the hindbrain and spinal cord (red arrows), but only the human sequence drives expression in the developing cortex (black arrow), a brain region that grew disproportionately in size after humans split from chimps.

What happens when the reporter gene *lacZ* is put under control of *2xHAR.114*?



Both the chimpanzee and the human *2xHAR.114* sequences drive expression of *lacZ* in the spinal cord and developing brain (red arrows), but the chimpanzee sequence drives expression to a more extensive region in the limb buds (black arrows). Human hands are weaker than chimps' but they are better at fine motor control.

Псевдогены

- Фермент аскорбиновой кислоты (и у человека, и у шимпанзе)
- Инактивированы рецепторы обоняния
- Инактивирован ген кератина (волосяной покров)
- FOXP2 – две аминокислотные замены

Ген речи FOXP2 (TF)

Letter | Published: 12 November 2009

Human-specific transcriptional regulation of CNS development genes by FOXP2

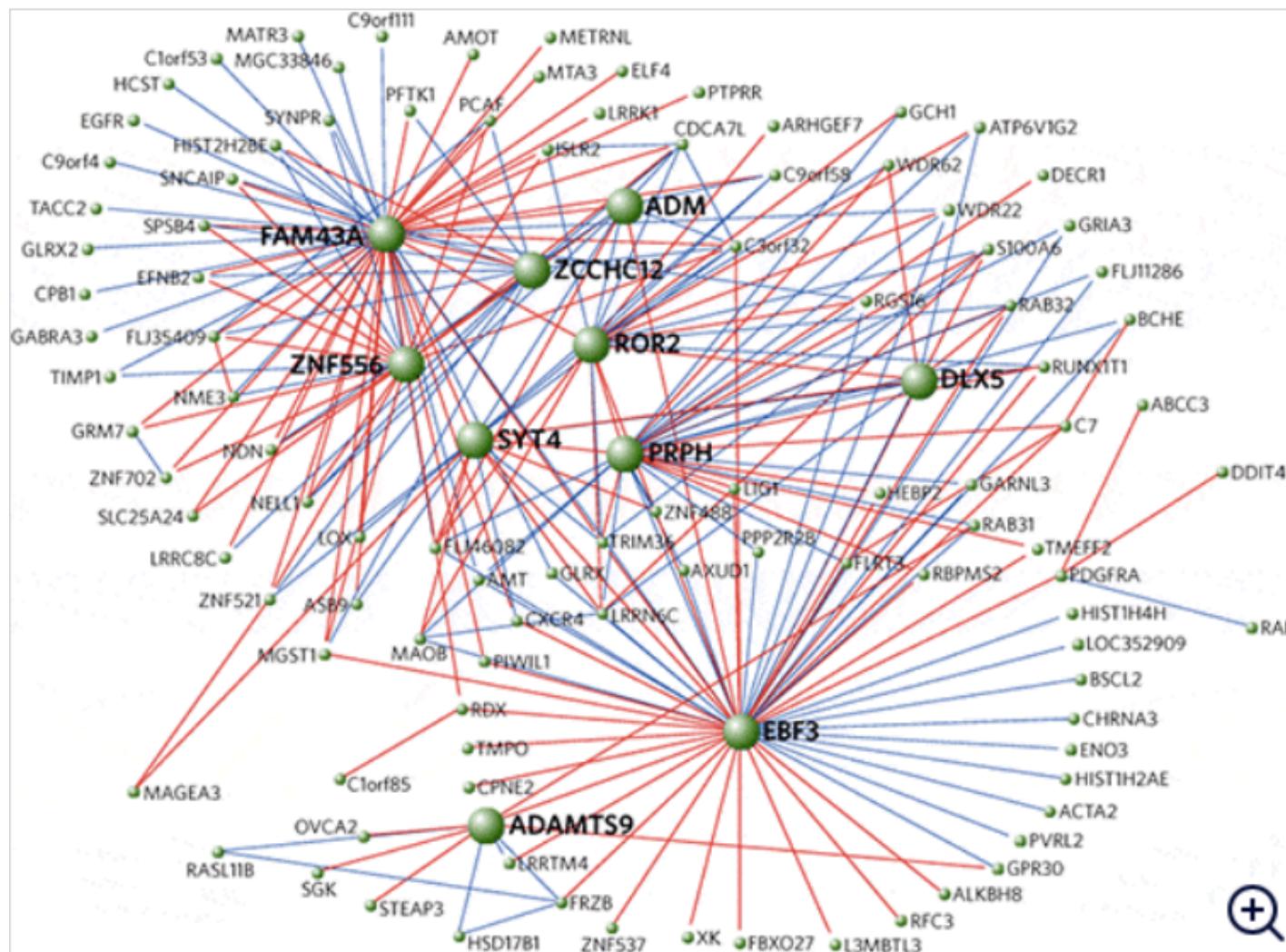
Genevieve Konopka , Jamee M. Bomar, Kellen Winden, Giovanni Coppola, Zophonias O. Jonsson, Fuying Gao, Sophia Peng, Todd M. Preuss, James A. Wohlschlegel & Daniel H. Geschwind 

Nature **462**, 213–217(2009) | Cite this article

876 Accesses | **239** Citations | **30** Altmetric | Metrics

Visualization of one of the modules containing FOXP2 and FOXP2 chimp differentially expressed genes.

Two-hundred-and-fifty pairs of genes with the greatest topological overlap are shown. Positive correlations are depicted in red and negative correlations are depicted in blue. The gene symbols for hub genes are accentuated in large, bold text.



SRGAP2

- Три копии у человека, одна копия у шимпанзе
- Увеличилось число нейронов, форма и плотность расположения синапсов

The Evolution of Human-Specific Genes by Duplication

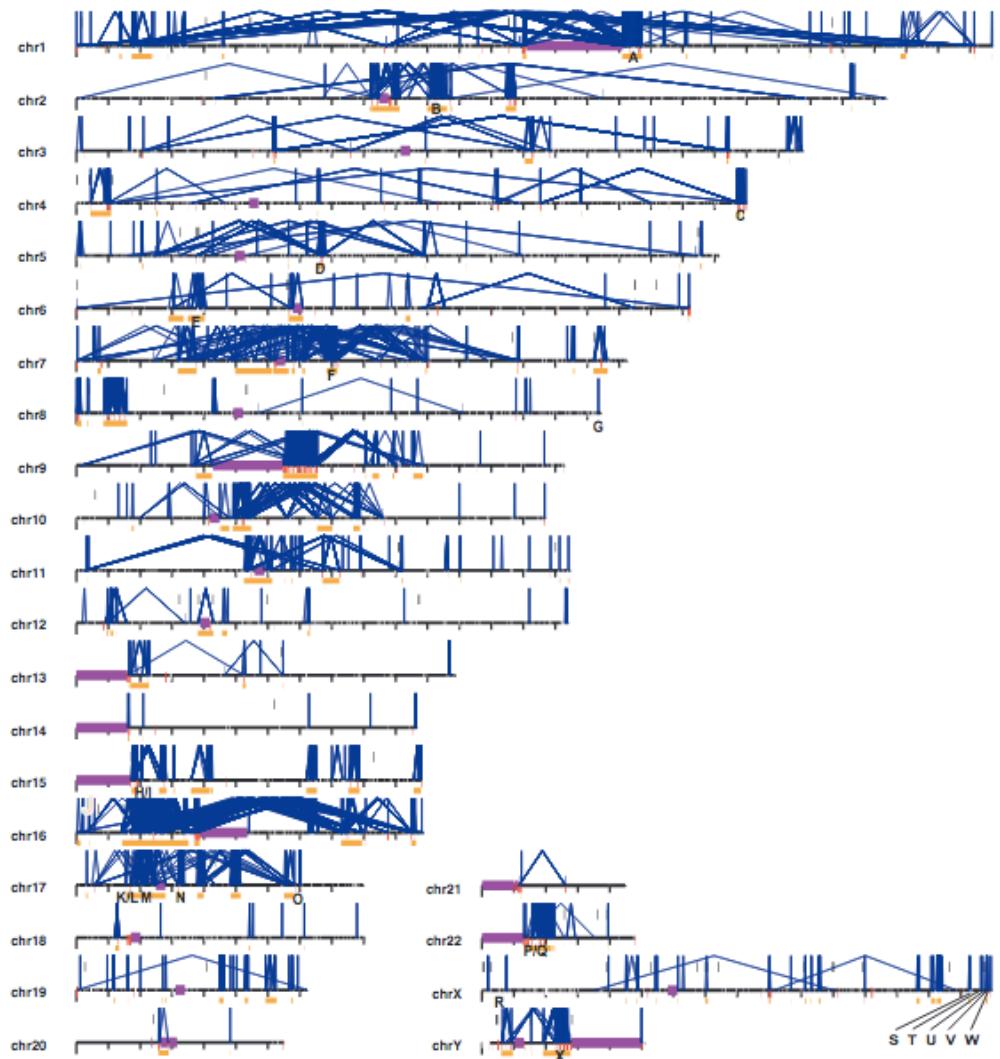
REPORT

Recent Segmental Duplications in the Human Genome

Jeffrey A. Bailey¹, Zhiping Gu², Royden A. Clark¹, Knut Reinert², Rhea V. Samonte¹, Stuart Sc...

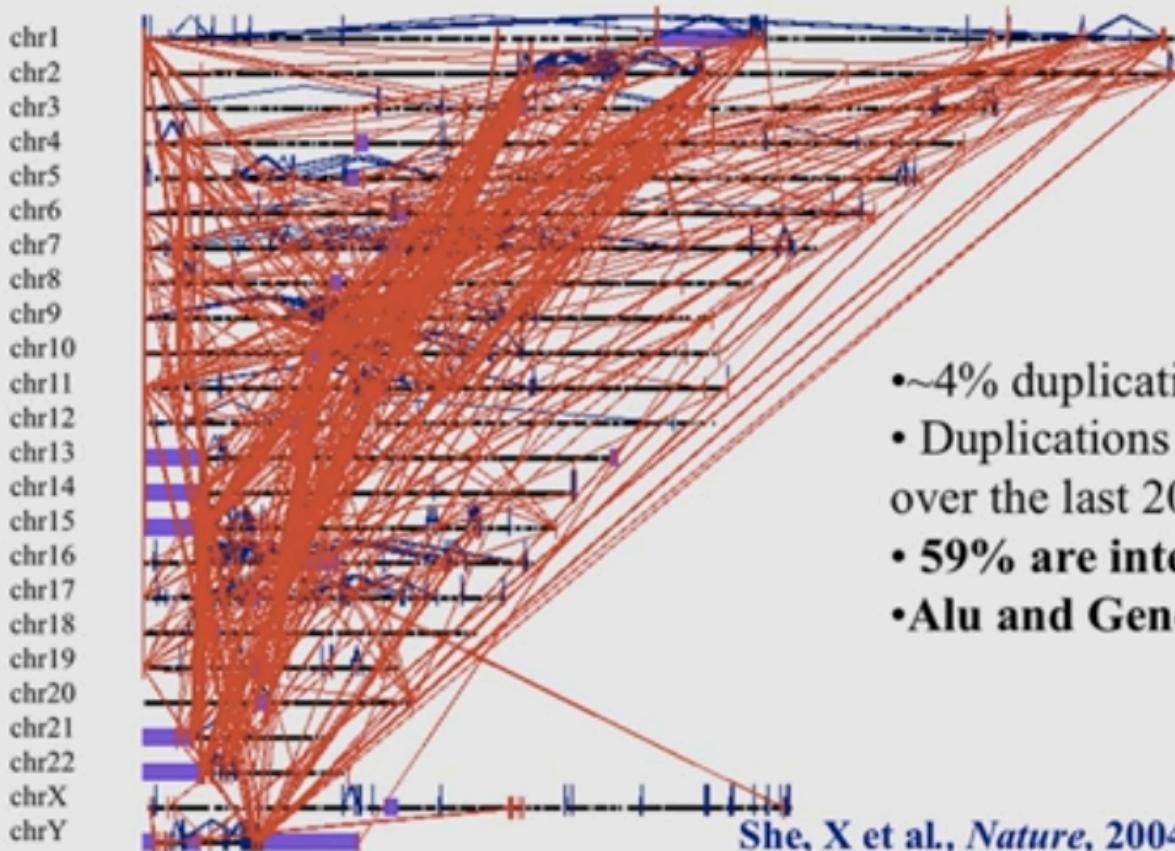
+ See all authors and affiliations

Science 09 Aug 2002:
Vol. 297, Issue 5583, pp. 1003-1007
DOI: 10.1126/science.1072047



Human Genome Segmental Duplication Pattern

Human
Chimp
Gorillas
All share
this

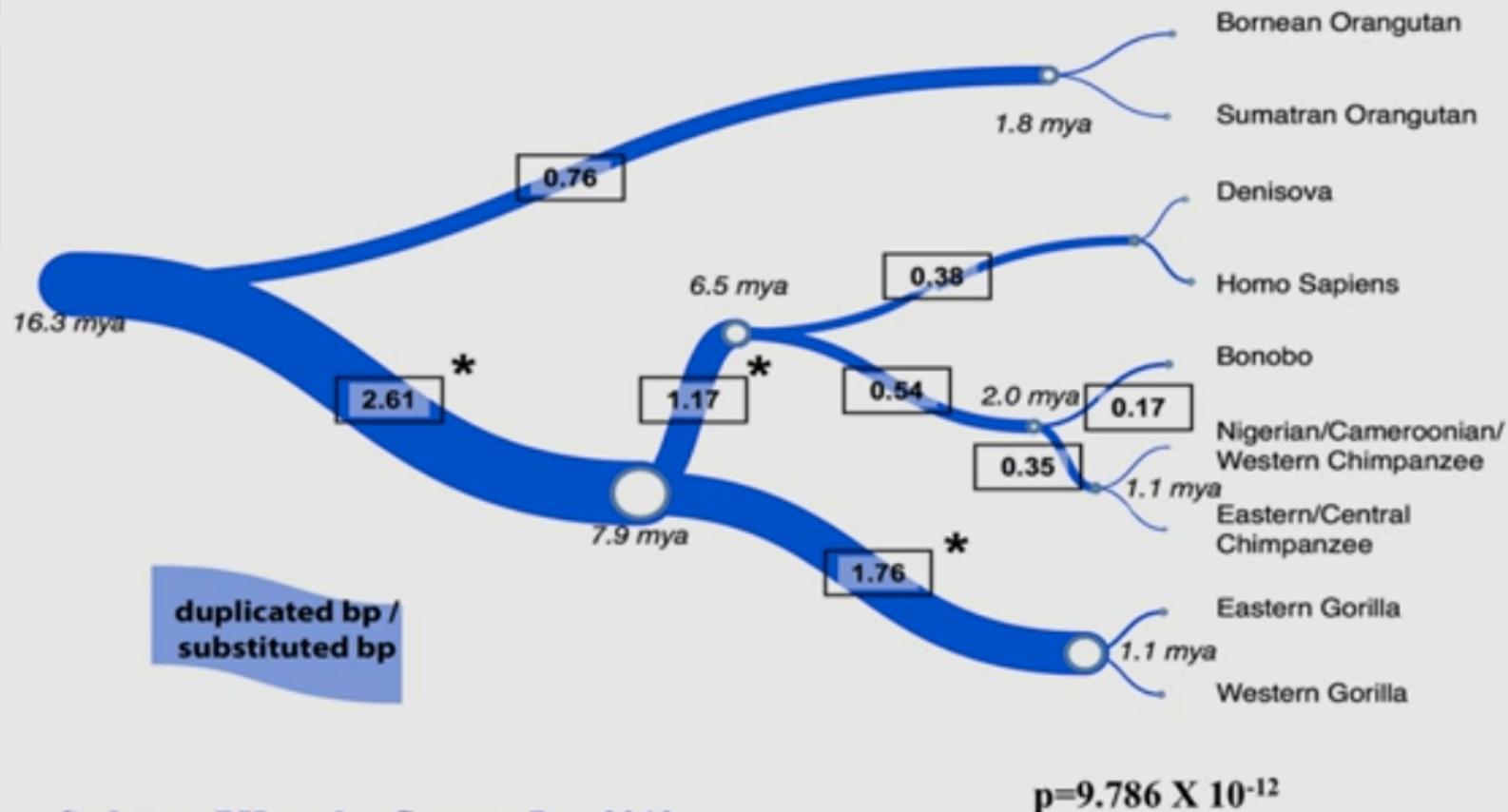


- ~4% duplication (125 Mb)
- Duplications pattern arose over the last 20 million years
- **59% are interspersed**
- **Alu and Gene-rich**

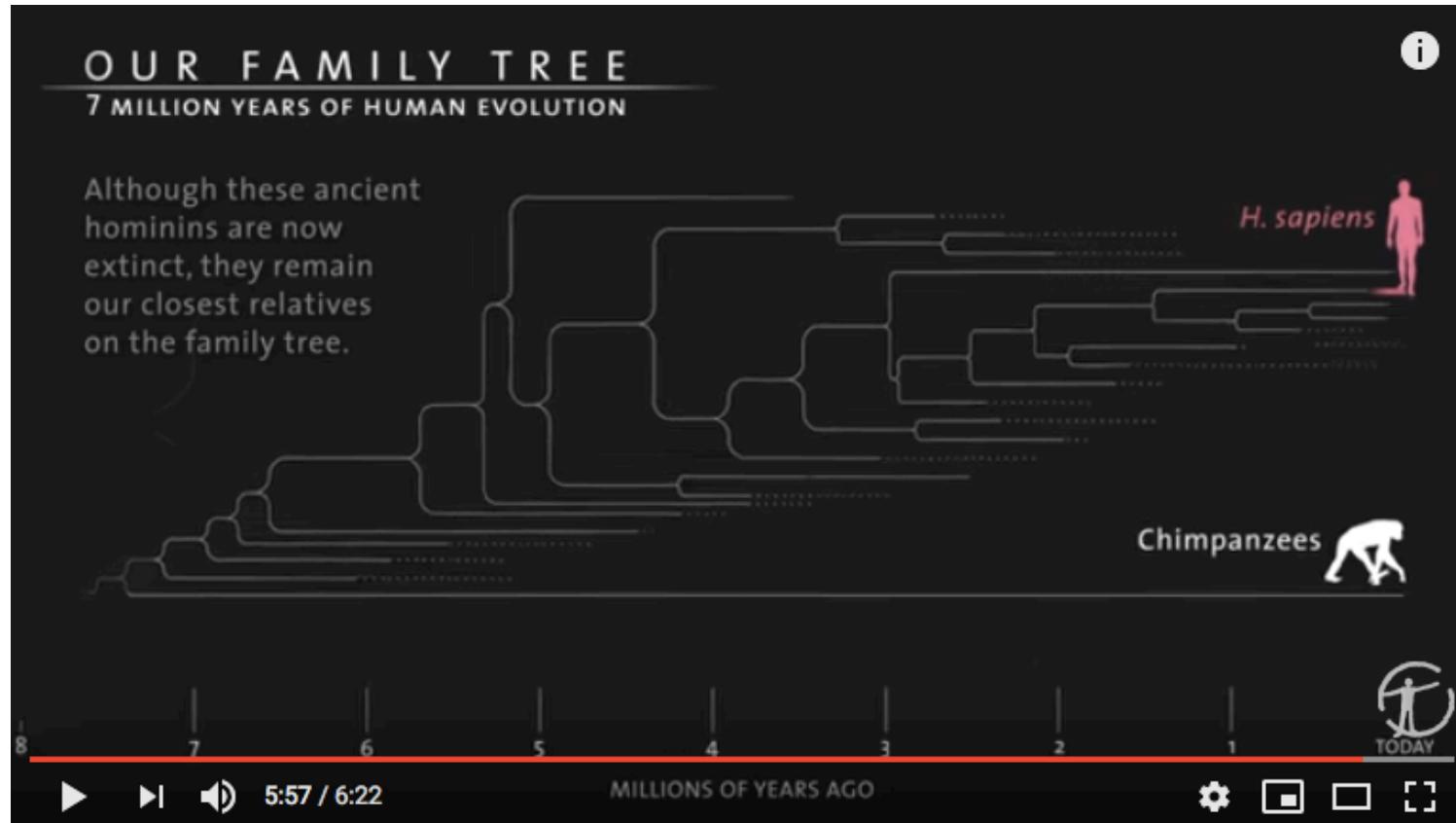
She, X et al., *Nature*, 2004

<http://humanparalogy.gs.washington.edu>

Great Ape “Burst” of Segmental Duplications



Seven Million Years of Human Evolution



<https://www.youtube.com/watch?v=DZv8VylQ7YU>