



Člověk a šimpanz



Jak velké genetické rozdíly jsou důležité
(Týden vědy a techniky, 8. 11. 2006)

Radek Zíka

zika@img.cas.cz

Ústav molekulární genetiky AVČR

<http://www.img.cas.cz>



CZECH FOBIA

Syllabus

- Úvodní poznámky a termíny z molekulární biologie
- Genomové projekty a zdroje na internetu
- Rozdíly mezi šimpanzem a člověkem na úrovni DNA
- Několik zajímavých genů
 - FOXP2
 - imunitní systém
 - MYH16



AGCTTACCGTAATTTACCGCTTACCGTAATTTACCG
TGGCCTACTTACCGTAATTTACCGCTGACCTACCT
CTACGCTTACCGTAATTTACCGATTACCGTAGGCC
AATATACTCTTACCGTAATTTACCGACTCCTCCGG
CTTACCGTAATTTACCGCTTACCTATCCGTTTACT
ATCGGTCTTACCGTAATTTACCGATTACCGTAGGCC
CTCTACCGTAATTTACCGTAATTTACCGTAATTTACCG
AATACTACTTACCGTAATTTACCGCTTACCTCCGG
CTGACTACCTTACCGTAATTTACCGCTTACCTCCGG
CGGACTACCGTAATTTACCGTAATTTACCGTAATTTACCG
AGCTTACCGTAATTTACCGTAATTTACCGTAATTTACCG
TGGCCTACTTACCGTAATTTACCGCTGACCTACCT
CTACGCTTACCGTAATTTACCGTAATTTACCGTAGGCC
AATATACTCTTACCGTAATTTACCGACTCCTCCGG
CTTACCGTAATTTACCGCTTACCTATCCGTTTACT
ATCGGTCTTACCGTAATTTACCGATTACCGTAGGCC
CTCTACCGTAATTTACCGTAATTTACCGTAATTTACCG
AATACTACTTACCGTAATTTACCGCTTACCTCCGG
CTGACTACCTTACCGTAATTTACCGCTTACCTCCGG
CGGACTACCGTAATTTACCGTAATTTACCGTAATTTACCG
AGCTTACCGTAATTTACCGTAATTTACCGTAATTTACCG
TGGCCTACTTACCGTAATTTACCGCTGACCTACCT
CTACGCTTACCGTAATTTACCGTAATTTACCGTAGGCC
AATATACTCTTACCGTAATTTACCGACTCCTCCGG
CTTACCGTAATTTACCGCTTACCTATCCGTTTACT
ATCGGTCTTACCGTAATTTACCGATTACCGTAGGCC
CTCTACCGTAATTTACCGTAATTTACCGTAATTTACCG
AATACTACTTACCGTAATTTACCGCTTACCTCCGG

Úvod

Genomika

zkoumá genomy jednotlivých organismů

Bioinformatika

využívá informatické metody pro popis struktury a funkce biologických systémů

DNA zde není proto, aby sloužila organismům, naopak, organizmy jsou zde proto, aby sloužily DNA

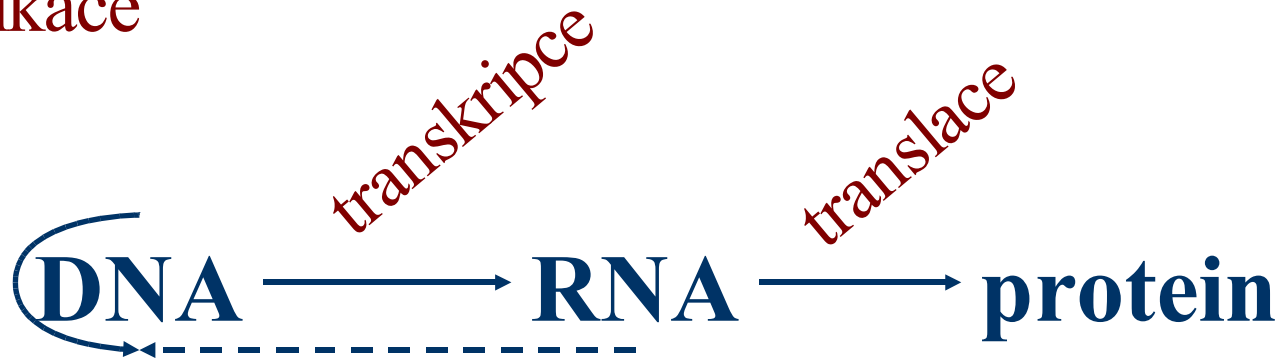
Richard Dawkins

Komparativní genetika se zabývá podobnostmi. Ale v případě genomu šimpanze hledáme především rozdíly.

Svante Pääbo

Centrální dogma molekulární genetiky

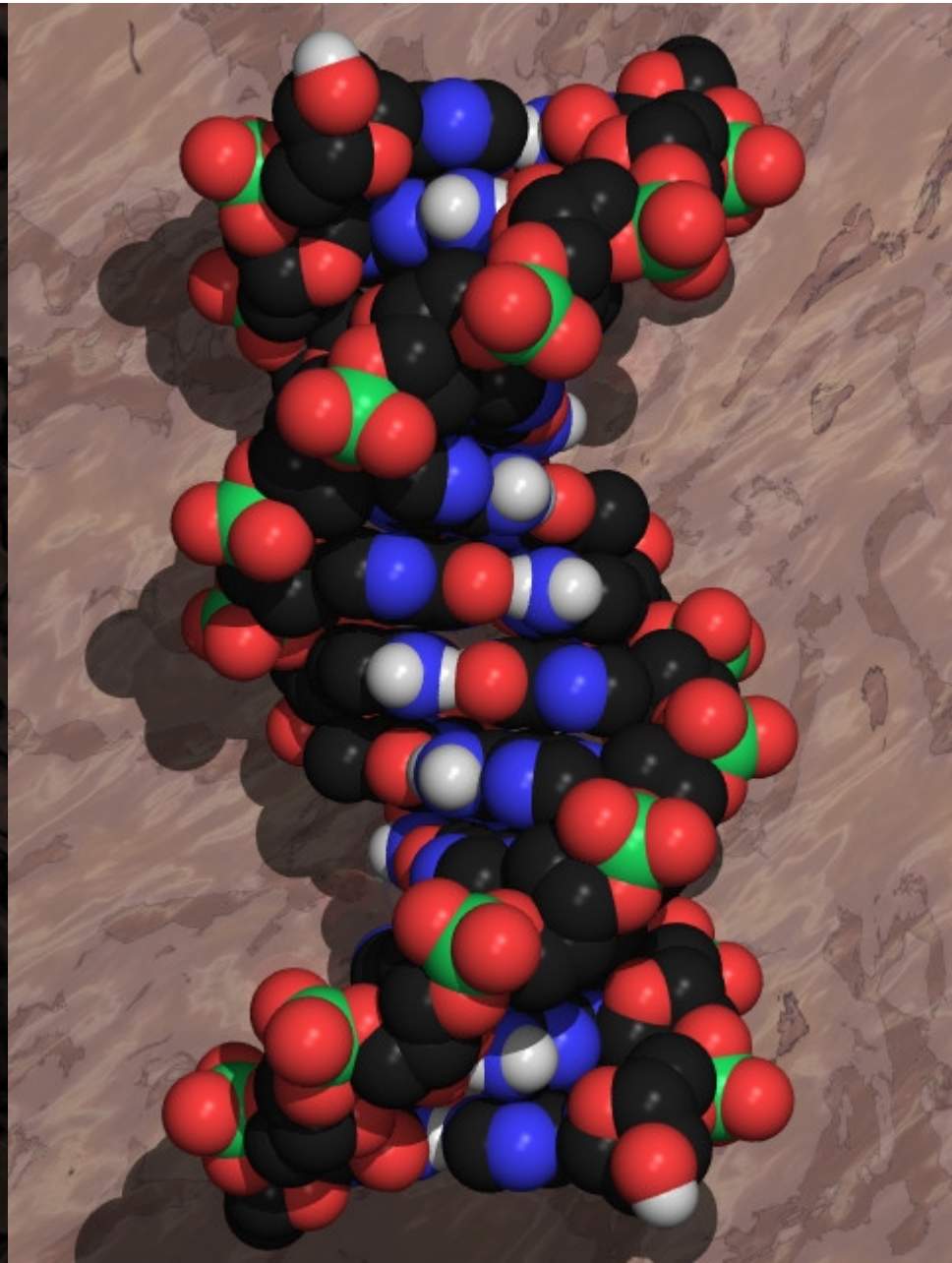
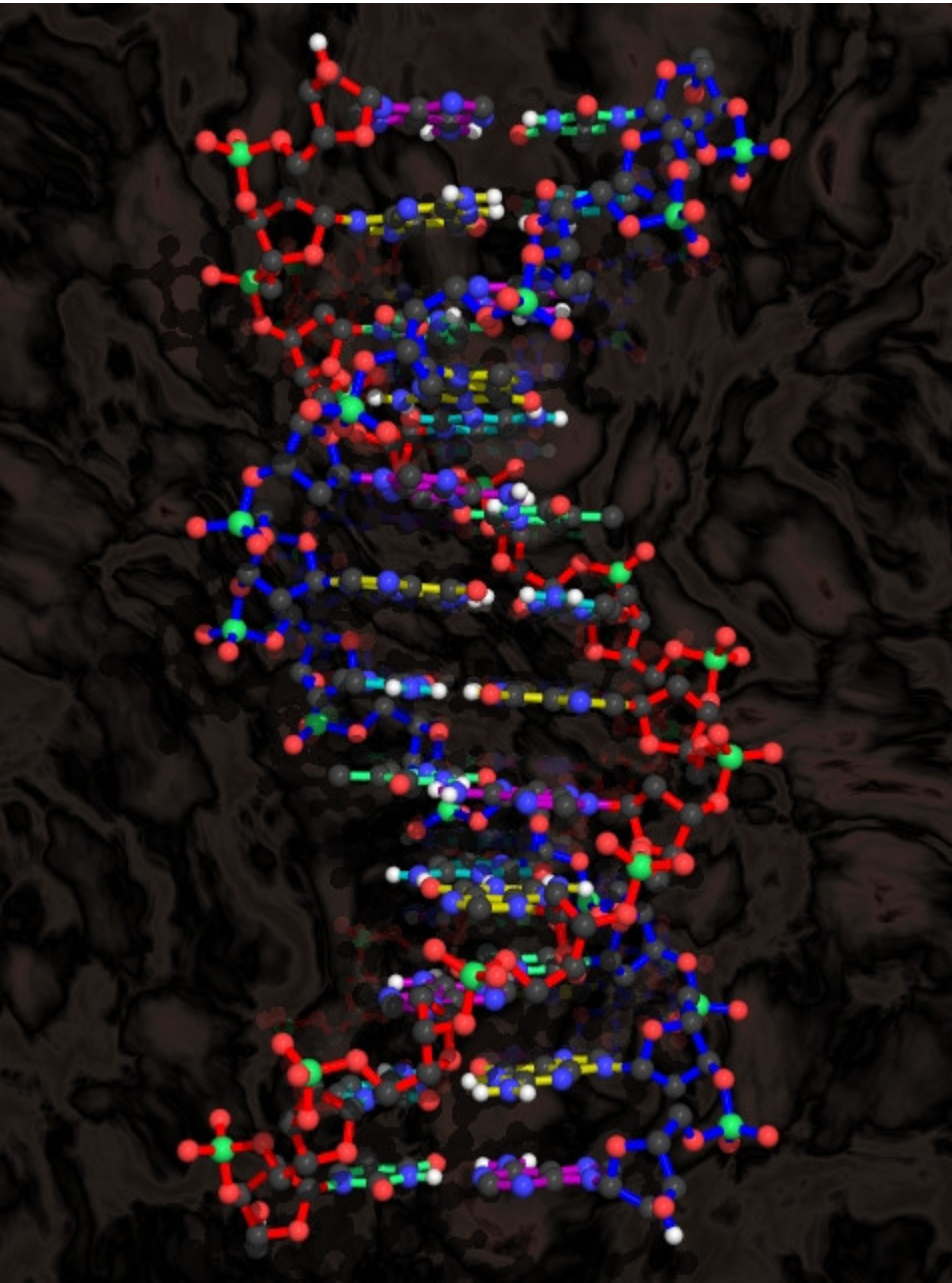
replikace



reverzní
transkripce

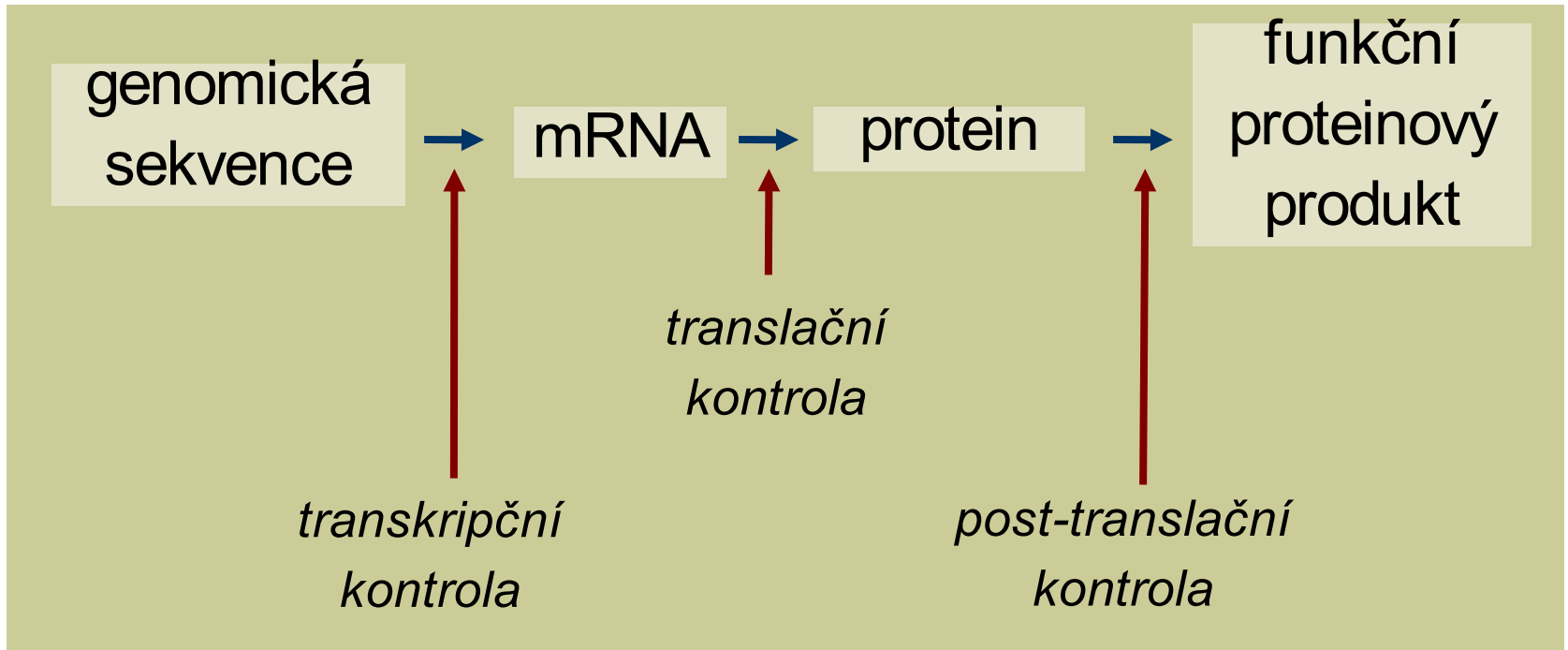
informace → funkce

DNA

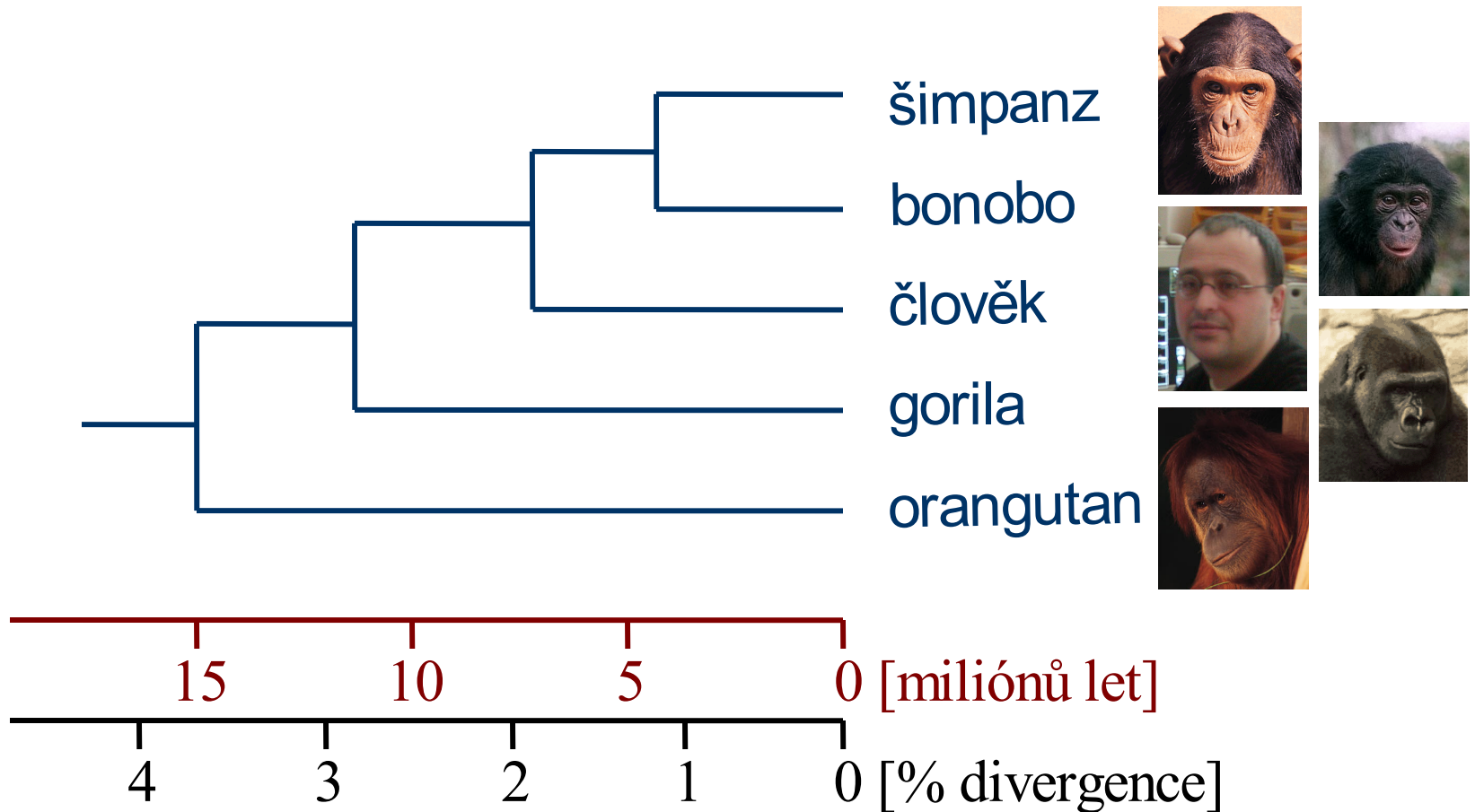


Úrovně kontroly

počet genů \neq úroveň mRNA \neq úroveň genové exprese \neq
množství a efektivní účinnost proteinu



Evolutione primátů



Proč zrovna genom šimpanze?

Medical Condition

Definite

HIV progression to AIDS

Influenza A symptomatology

Hepatitis B/C late complications

P. falciparum malaria

Menopause

Likely

E. coli K99 gastroenteritis

Alzheimer's disease pathology

Coronary atherosclerosis

Epithelial cancers

Speculative

Menstrual blood loss

Early fetal wastage

Humans

Common

Moderate to severe

Moderate to severe

Susceptible

Universal

Resistant

Complete

Common

Common

Variable

High

Great Apes

Very rare

Mild

Mild

Resistant

Rare

Sensitive?

Incomplete

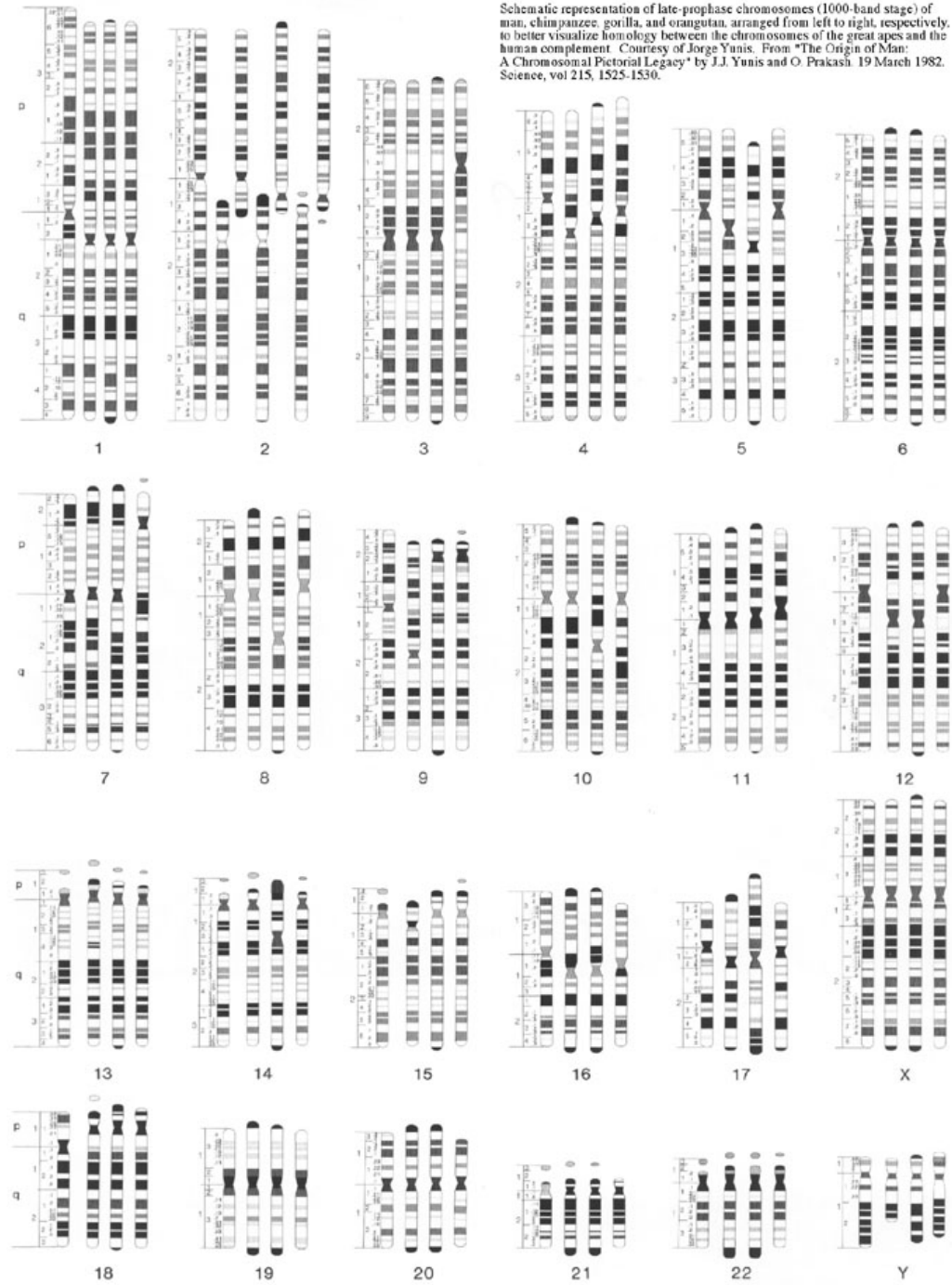
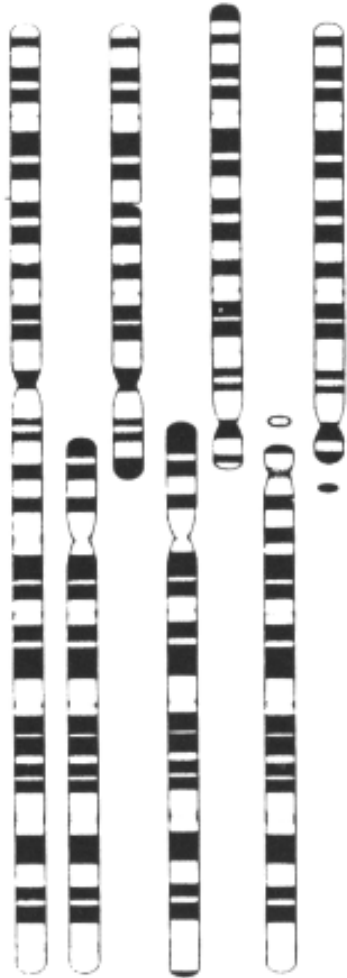
Uncommon

Rare

Lower amount?

Low?

člověk
 šimpanz
 gorila
 orangutan



Schematic representation of late-prophase chromosomes (1000-band stage) of man, chimpanzee, gorilla, and orangutan, arranged from left to right, respectively, to better visualize homology between the chromosomes of the great apes and the human complement. Courtesy of Jorge Yunis. From "The Origin of Man: A Chromosomal Pictorial Legacy" by J.J. Yunis and O. Prakash. 19 March 1982. Science, vol 215, 1525-1530.

Koho jsme přečetli?

| | | |
|-----------------|------------------------|------|
| Člověk: | <i>Homo sapiens</i> | (hs) |
| Šimpanz: | <i>Pan troglodytes</i> | (pt) |

HGC: 9 neznámých lidí

- 5x mužská krev
- 3x spermie
- 1x 987SK buňky

Celera: 2 muži, 3 ženy

- Afroameričan
- Asiat - Číňan
- 2 Zakavkazané
- Hispánec - Mexičan

šimpanz: Clint (Yerkes National Primate Research Center)

Clint

He's tall, dark, and handsome, with a grin that turns heads, especially those of older women. Smart, playful, and a flirt, he's happiest when someone's scratching his back.

Clint died from heart failure,
but two cell lines from the primate
have been preserved.



January 8th, 2005: Clint was put down at the Yerkes National Primate Research Center in Atlanta. The cause of death was not immediately known. Clint, 24, was the living reference point for \$18 million worth of genetic code.

Rozdíly hs x pt - I

Počet genetických odlišností mezi člověkem a šimpanzem ...

- › ~ 60x méně než člověk vs. myš
- › ~ 10x méně než myš vs. krysa
- › ~ 10x více než mezi dvěma lidmi

Změny ...

- › 2.7% celkových rozdílů představují duplikace
- › 1.2% záměny jednotlivých nukleotidů
- › v “typickém” ortologu: 2 ak (1 na linii)!

Srovnání chr. 21 a 22 ...

1.44% rozdíl mezi DNA

68000 indels mezi hs chr. 21 a pt chr. 22

15% všech CpG je mutováno (23x více transicí a 7x více transverzí)

83% genů obsahuje rozdíl na úrovni aminokyselin

nejvíce jsou mutovány U3' oblasti

Rozdíly hs x pt - II

- U několika tříd genů se ukazuje vyšší rychlost evoluce u člověka než u šimpanze (*FOXP2*, *AH1*, *ASPM*, *MCPH1*). Jednou z nejvýznamějších skupin představují transkripční faktory.
- Více než 50 genů přítomných v lidském genomu schází nebo jsou částečně “vymazány” z genomu šimpanze.
- 3 klíčové geny účastníci se zánětlivých procesů jsou zřejmě “vymazány” z genomu šimpanze
- Lidé zřejmě ztratily funkci genu caspase-12, který produkuje enzym, který pomáhá jiným organismům proti Alzheimer's disease.

kde začít na internetu

Netscape: Ensembl Genome Server

File Edit View Go Communicator Help

Back Forward Reload Home Search Netscape Print Security Shop Stop

Location: http://www.ensembl.org/perl/contigview?chr=8&vc_start=87000000&vc_end=89000000

Ensembl ContigView

Home News BLAST Disease Browser Docs Download

Search for All I Lookup [e.g. U34879, AP000869]

Chr: 8 Nucleotides: 87000000 to 89000000 Band: q21.13 View Reset

Click anywhere on the red line below to reposition focus window

87.0Mb 89.0Mb

Contigs Markers

Genes

Feature View

scale (bp)
repeats
snp
Mus musculus sequence
snp
genscan
repeats
scale (bp)

87950000 88050001

Jump to UCSC dump

100%

enPath
me.ucsc.edu

links
cas.cz/links

Ensembl
http://www.ensembl.org/Homo_sapiens

n/Entrez/hum_srch

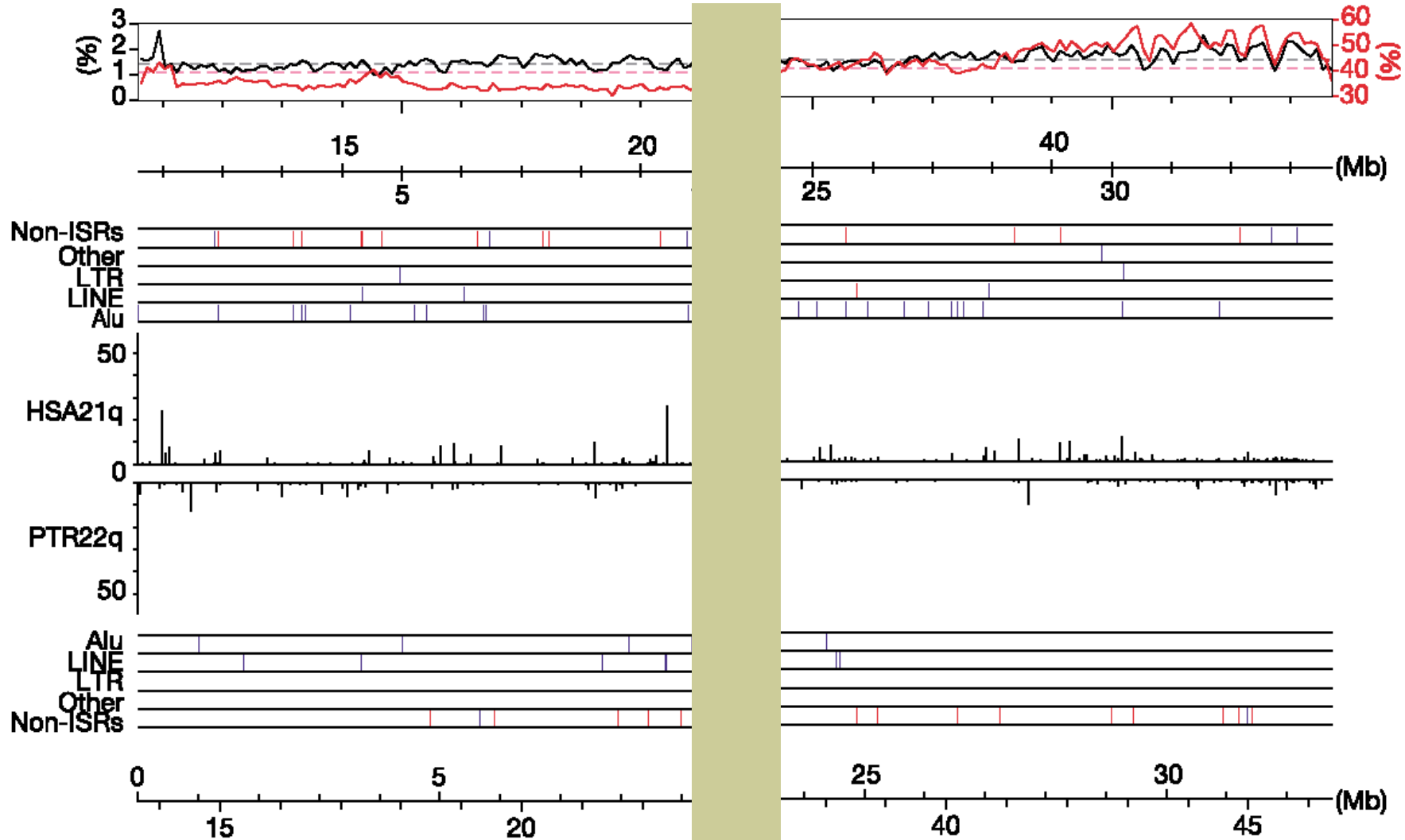
vybrané zdroje dat na internetu

SwissProt <http://www>
Entrez <http://www>

The screenshot shows the NCBI Entrez Nucleotide search interface. The browser window title is "Netscape: Entrez - Nucleotide". The address bar shows the URL: <http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?db=Nucleotide>. The page features a search bar with a "Go" button and a "Clear" button. Below the search bar, there are links for "Limits", "Preview/Index", "History", and "Clipboard". A yellow box contains the text: "Entrez Nucleotides, part of the Entrez search and retrieval system, is a collection of nucleotide entries from GenBank. The number of bases in GenBank grows at an exponential rate. Today's total is: **10527520791**". Below this, there are two sections: "Decoding the human genome" and "VecScreen". The "Decoding the human genome" section mentions that complete DNA sequences for human chromosomes 22 and 21 have been published. The "VecScreen" section mentions a new vector contamination detection tool. At the bottom, there is a "Top 10 organisms of the month" bar chart showing the number of new nucleotide sequences processed in GenBank for May, 2000. The chart shows human as the top organism with approximately 900 sequences, followed by HIV type 1 with approximately 700 sequences, and pig with approximately 400 sequences.

| Organism | Number of new nucleotide sequences |
|------------------------------------|------------------------------------|
| human | ~900 |
| HIV type 1 | ~700 |
| pig | ~400 |
| house mouse | ~300 |
| Drosophila simulans | ~250 |
| Botryllus schlosseri (sea squirts) | ~200 |
| rhesus monkey | ~150 |
| Theileria annulata | ~100 |
| Norway rat | ~100 |
| Brachionus plicatilis (rotifers) | ~100 |

hs ch21 vs pt ch22





Sledování evoluce

Ka/Ks (dn/ds) ratio: poměr nesynonymních
a synonymních substitucí

$\ll 1$: negativní (purifying) selekce

~ 1 : neutrální

$\gg 1$: pozitivní selekce

FOXP2

one major hope is that the differences between the sequences will reveal the genetic basis for our mental and linguistic capacities

- obsahuje forkhead doménu
- poškození způsobuje abnormality v řeči a chápání jazyka (KE rodina)
- silně konzervován - pouze 3 aminokyselinové změny mezi člověkem a myší
- dvě specifické záměny se rozšířily před cca 100 000 – 200 000 lety (vznik moderního člověka)
thr -> asp (233) a asp -> ser (325)

KE family

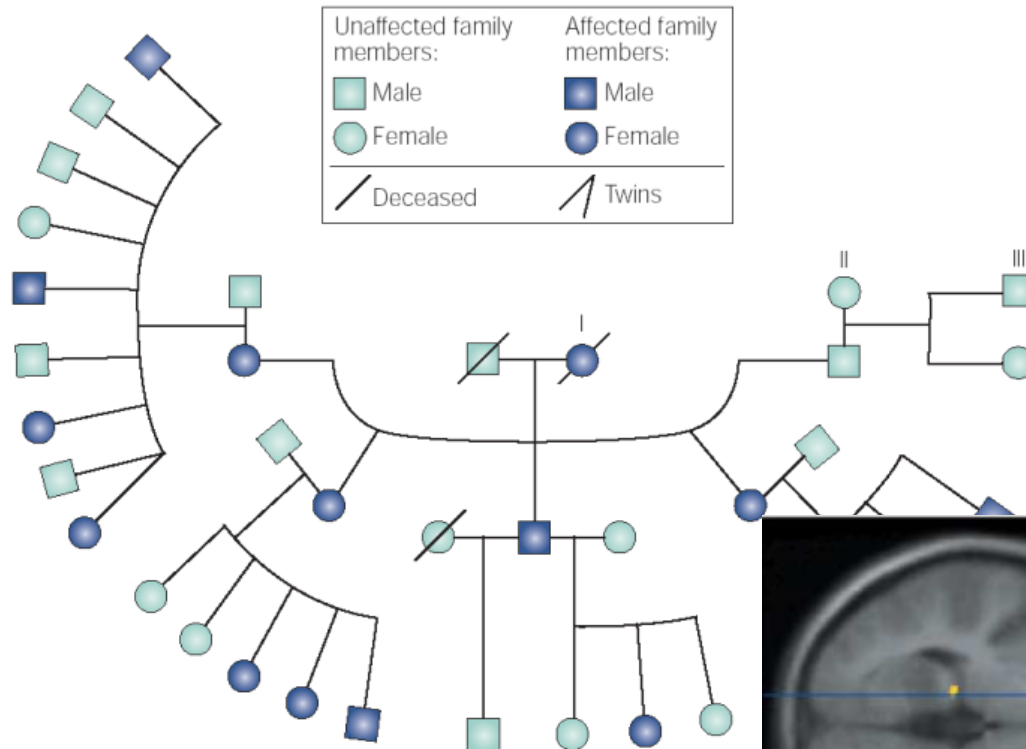
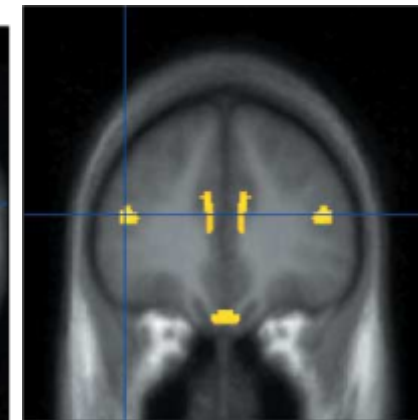
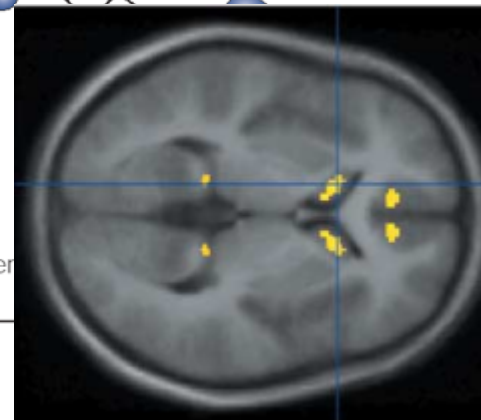
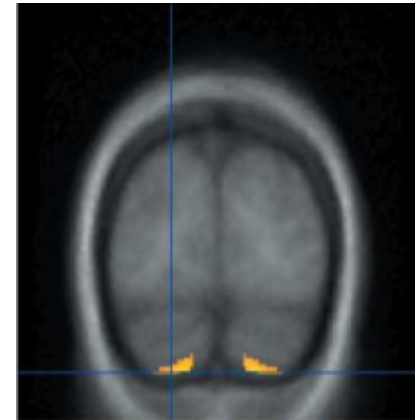
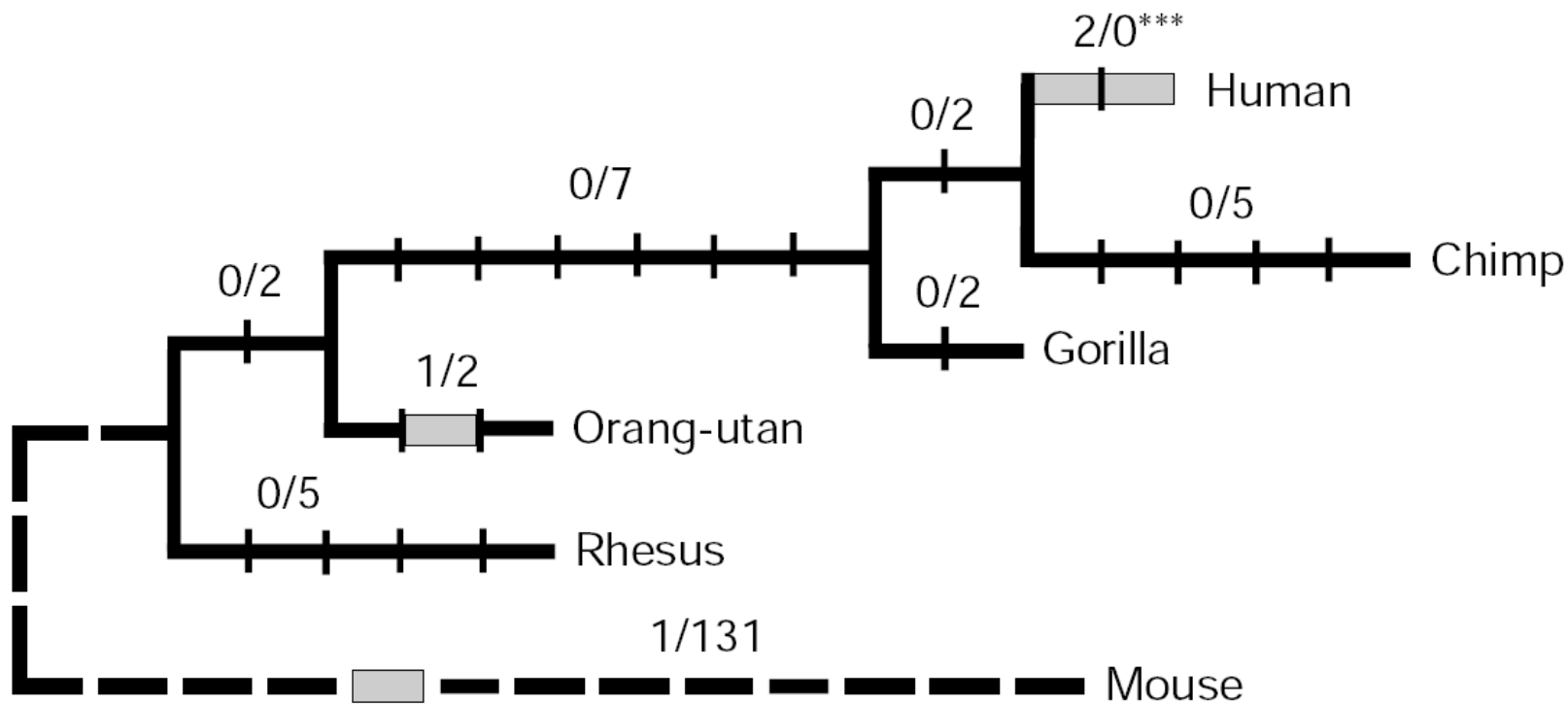


Figure 1 | **Pedigree of the KE family.** I, II and III represent the generation, from REF. 14 © (2002) Oxford University Press.



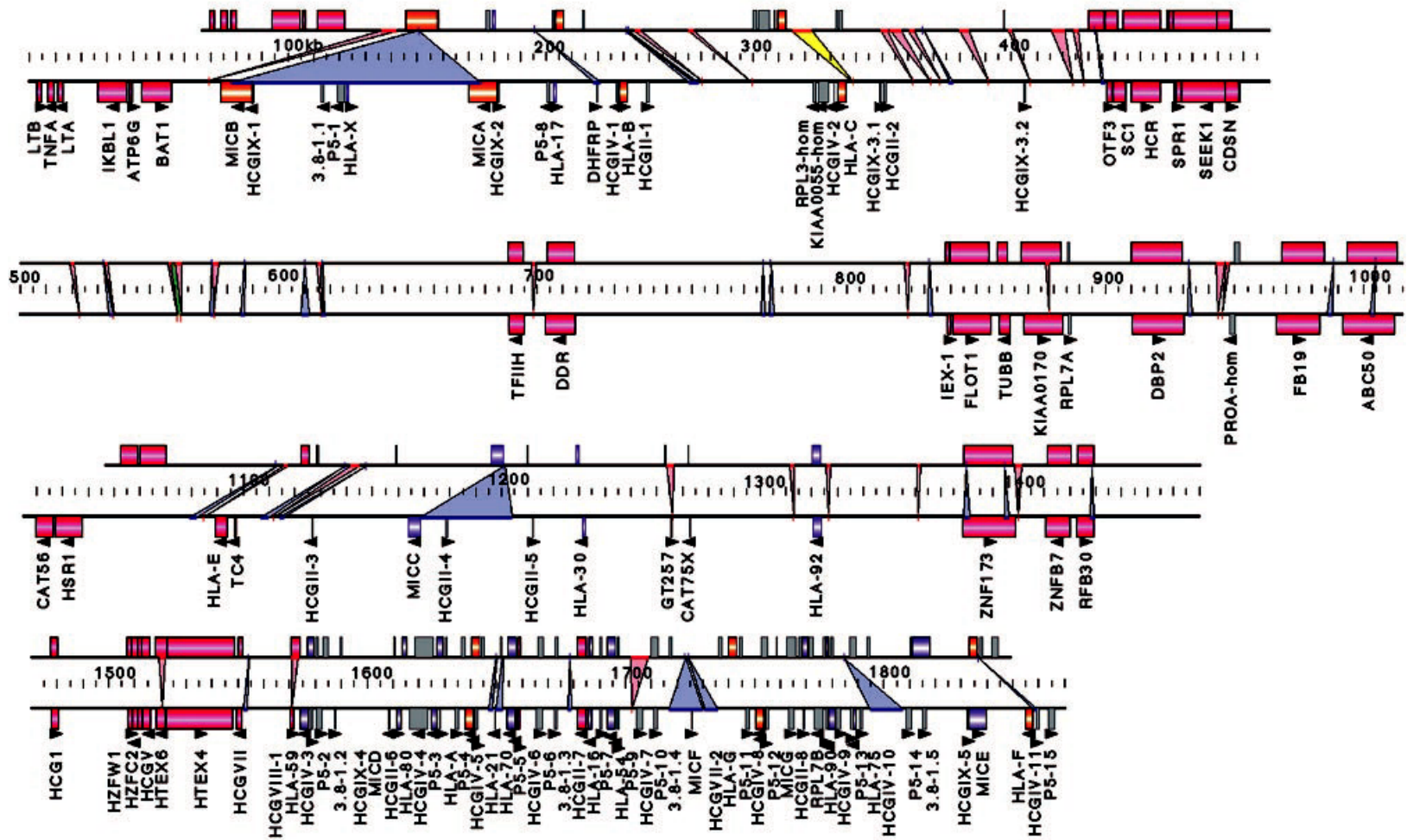
evoluce FOXP2



— nukleotidová záměna

■ aminokyselinová záměna

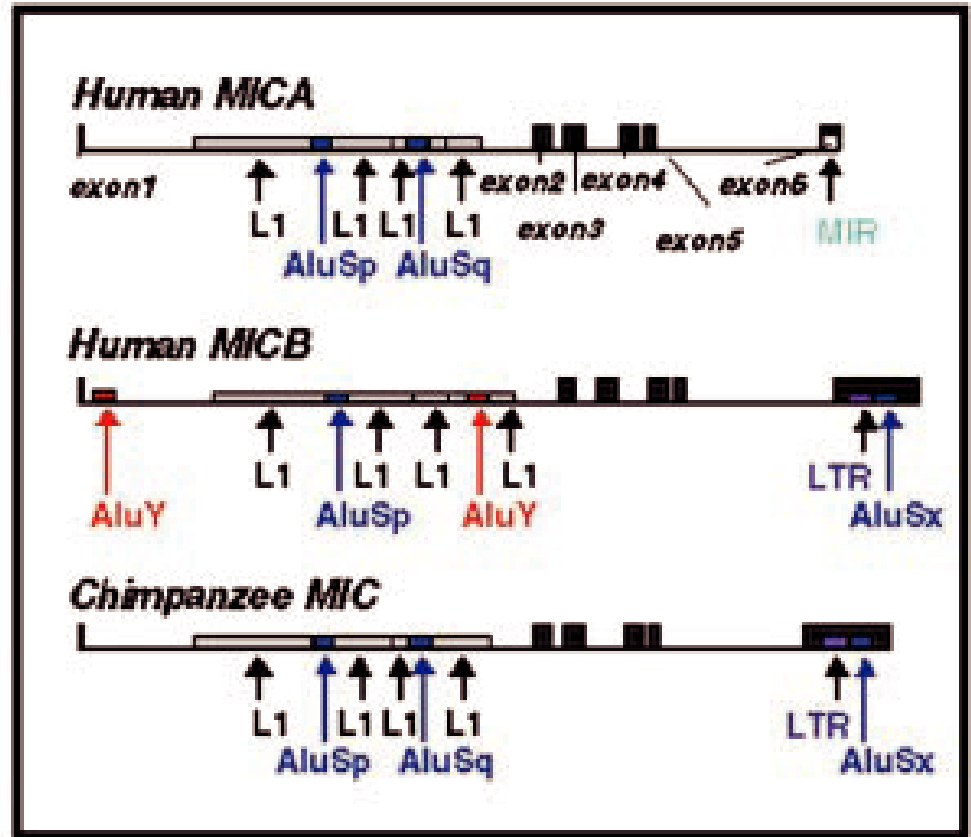
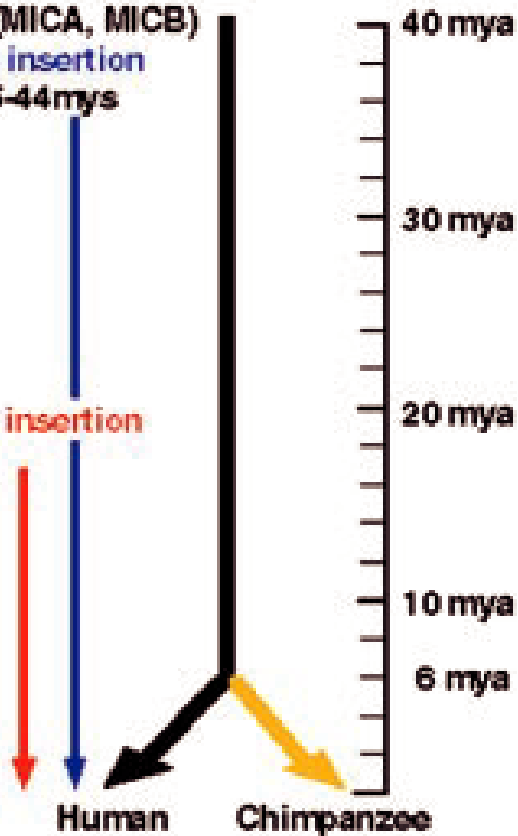
MHC I



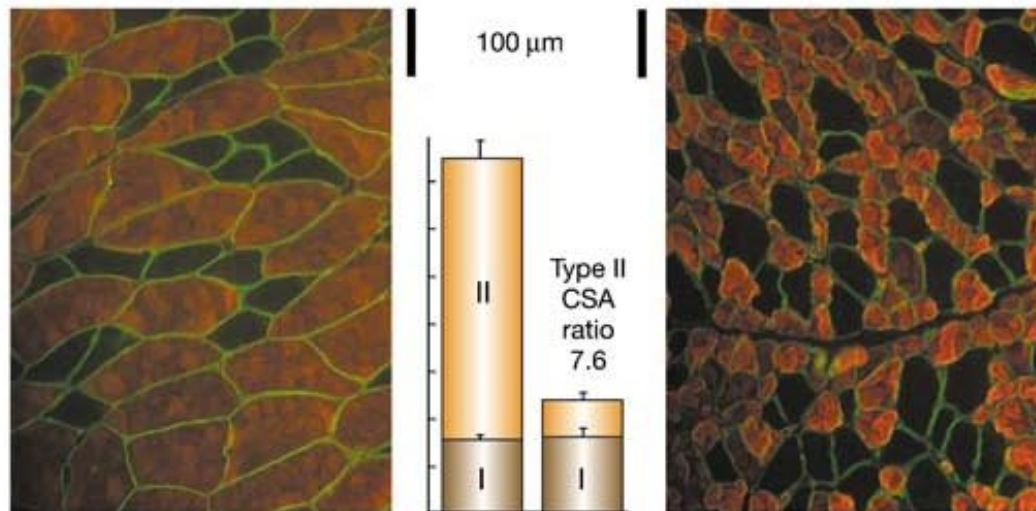
MIC A/B

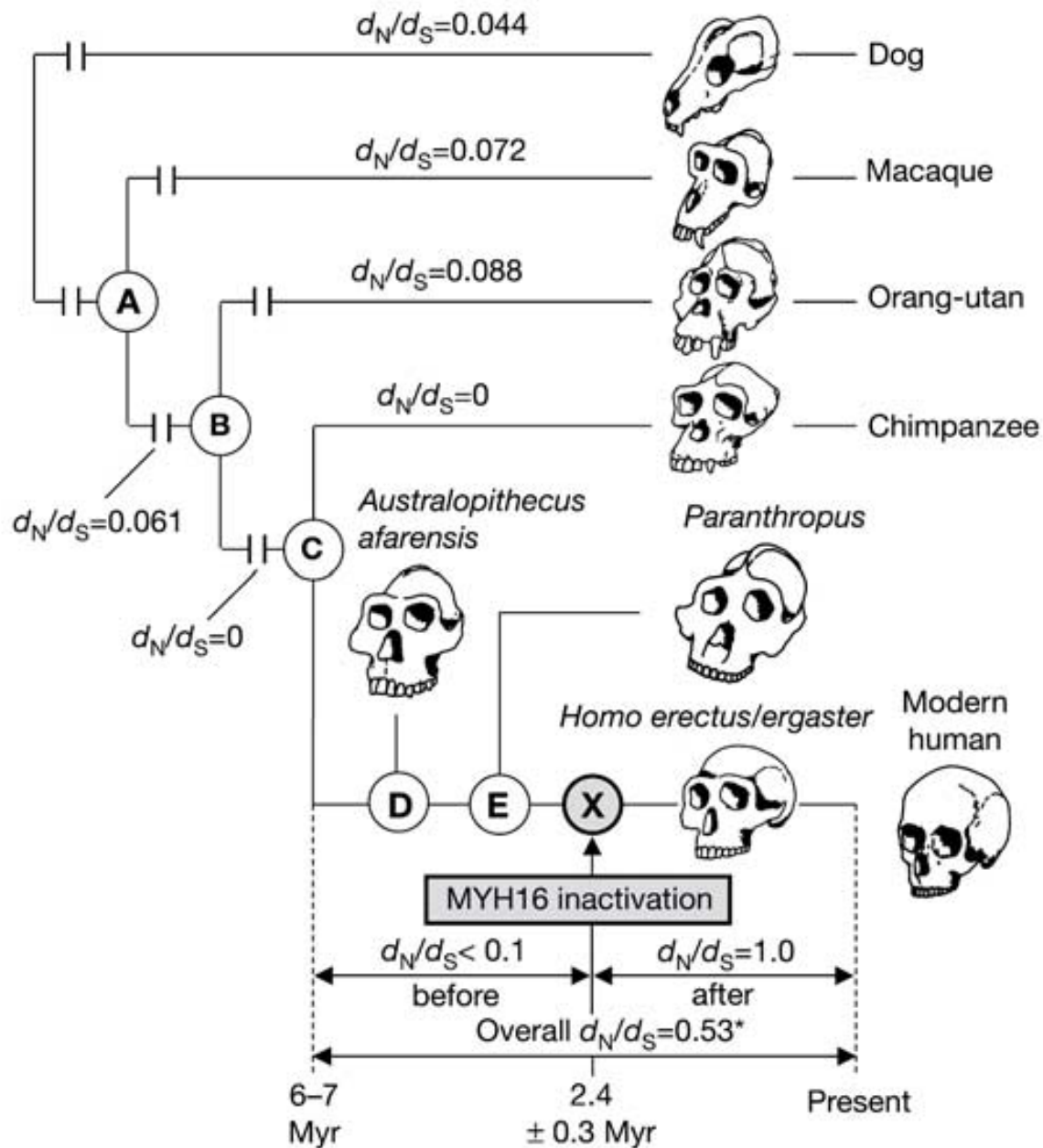
Duplication of MIC
gene (MICA, MICB)
AluS insertion
35-44mys

AluY insertion



MYH16 inactivation

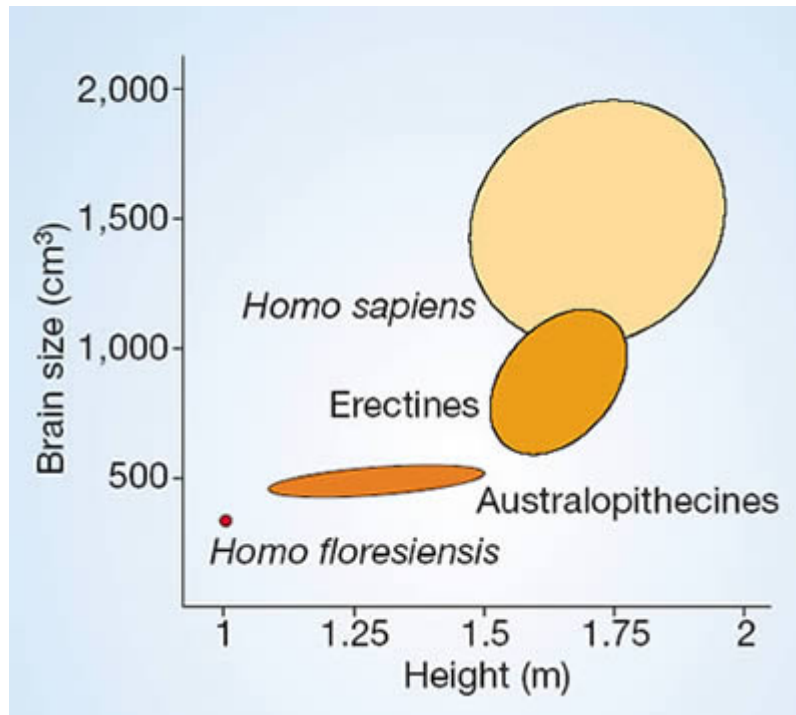




Stedman HH, Kozyak BW, Nelson A, Thesier DM, Su LT, Low DW, Bridges CR, Shrager JB, Minugh-Purvis N, Mitchell MA. Myosin gene mutation correlates with anatomical changes in the human lineage. *Nature*. 2004 Mar 25;428(6981):415-8.

Homo floresiensis

Velikost lebky (mozku)



Závěr? Ne!

Evan Eichler

“We're in a very nice intermediate stage of understanding human-chimp differences, we can't say, This is the difference that makes us human, but we can say, These are the regions of the genome that show a lot of potential and are excellent candidates to do further work on.”

Carina Dennis

“The chimp was a great start. But the genomes of our other primate relatives will help to reveal a whole lot more.”

kontakt

Jan Pačes

www: <http://bio.img.cas.cz>

email: hpaces@img.cas.cz

icq: #110872370

irc: efnet #hpaces

tel: +420 220183446

Radek Zíka

www: <http://bio.img.cas.cz/zikar>

email: zikar@img.cas.cz

icq: #111030547

jabber: [zikar@jabber.cz](jabber:zikar@jabber.cz)

tel: +420 220183446





Large-Scale Sequencing Center



GenBank (April 2006): ~130 000 000 000 nt
Swiss-prot (Rel. 51.0; 31-Oct-06): 88 541 632 aa

biolinks

http://bio.img.cas.cz/links

The screenshot shows a Netscape browser window titled "BioInformatic Links @ IMG - Homepage - Netscape". The address bar displays "http://bio.img.cas.cz/links/". The website content is organized into three columns: Databases, Programs, and Other Links.

| DATABASES | PROGRAMS | OTHER LINKS |
|--|--|--|
| Sequences [38] <ul style="list-style-type: none">DNA [2]Protein [3]RNA [5]Human Genome [1 2]Eucaryotic Genomes [4]Bacterial Genomes [5]Repeated sequences [4] | Alignment & DB search [14] <ul style="list-style-type: none">Pairwise Alignment [4]Multiple Alignment [9] DNA Sequence Analysis [12] <ul style="list-style-type: none">Gene prediction [3]Repeated sequences [4] Protein Analysis [8] <ul style="list-style-type: none">Folding [1]Structure Prediction [1]Structure alignment [1] Phylogenetic Analysis [9] <ul style="list-style-type: none">Distance Methods [1]Parsimony Methods [0]Maximum Likelihood Methods [4]Tree Inference, Tree Distance [3]Bootstrap, Other Tests [0]RFLP, RAPD Data Analysis [1] Visualisation [9] <ul style="list-style-type: none">Alignment [4]Molecular Structures [2] Expression Evaluation [7] <ul style="list-style-type: none">Chips [6] Other Tools [1] | Courses [8] Link Pages [23] Unsorted [4] <ul style="list-style-type: none">Distributed applications [2]Old [0] Multifunctional Servers [2] Search Engines [2] Literature [12] <ul style="list-style-type: none">Preprints [2]Journals [6] Our Favorite Prague Pubs [1] New Inserted Links <hr/> Link Search <input type="text"/> |