

The impact of genomic neighbourhood on the evolution of human and chimpanzee transcriptome

Subhajyoti De, Sarah A Teichmann and M. Madan Babui
MRC Laboratory of Molecular Biology, Cambridge, UK

Genome Res published online Febr 19 2009
doi:10.1101/gr.086165.108

Triinu Kõressaar
Bioinformatics Jclub,
10.03.2009

Introduction

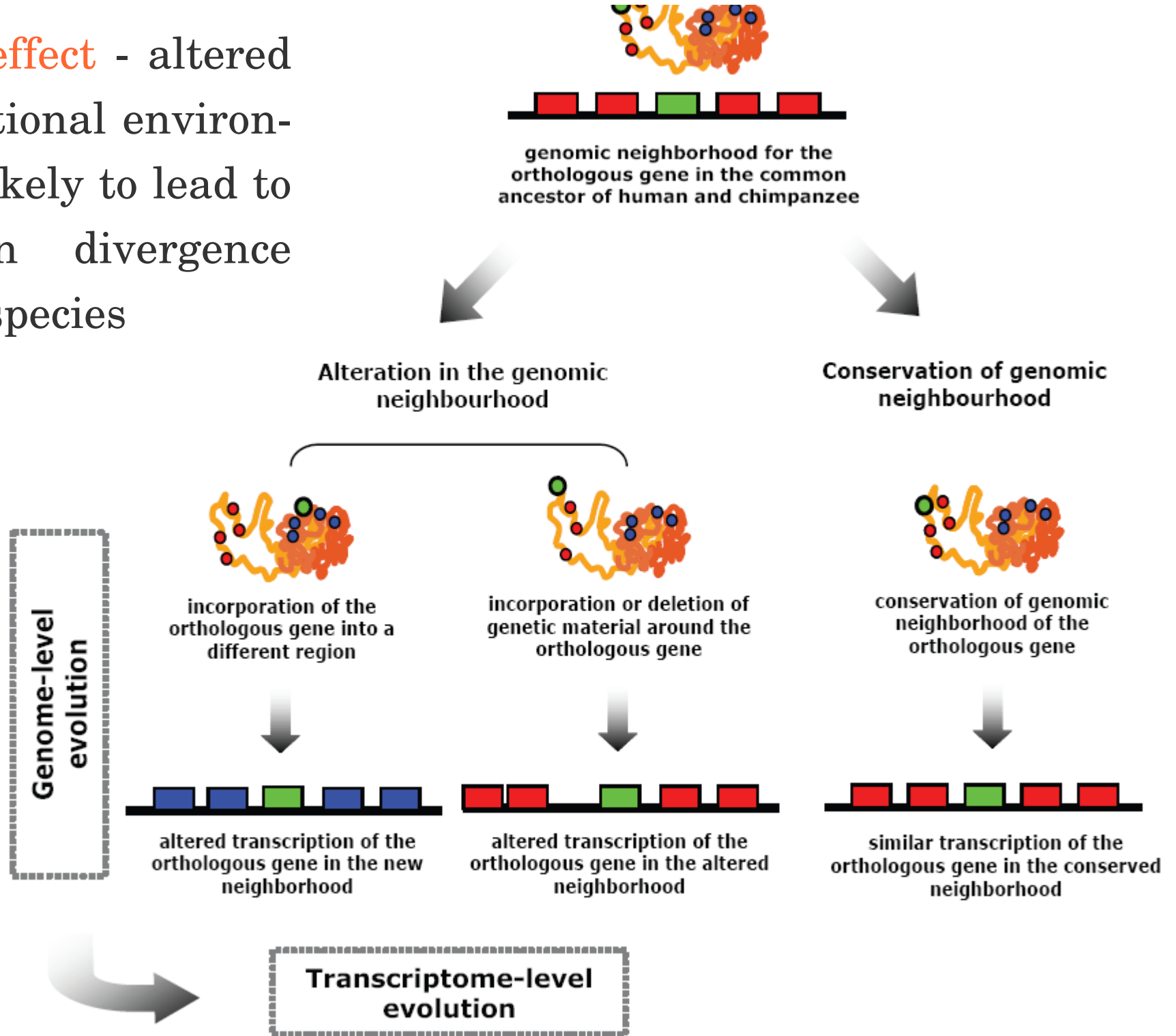
Interested in:

Phenotypic differences between humans and chimpanzee

..more precisely

- Expression level divergence between two species and
- Genomic neighbourhood of orthologous genes in human and chimpanzee

Position effect - altered transcriptional environment is likely to lead to expression divergence between species



1. **How** many genes show an alteration in their genomic neighbourhood in human and chimpanzee after the split from their common ancestor?

2. **Do** genes with altered neighbourhood display more gene expression divergence than other genes?

3. **In** which parts of the body are such genes expressed?

A

CGN - Conservation of Genomic Neighbourhood

A score calculated for every gene

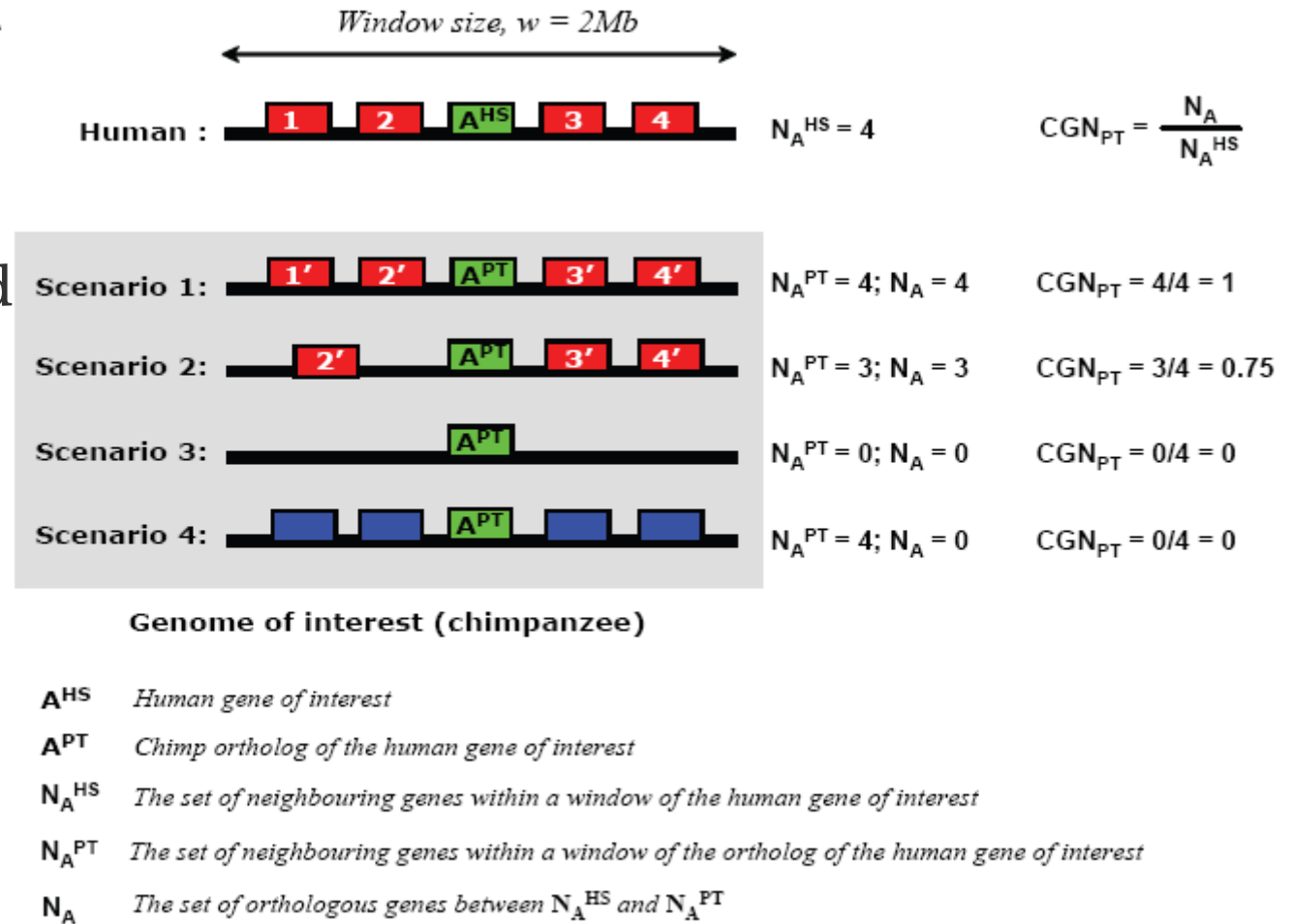
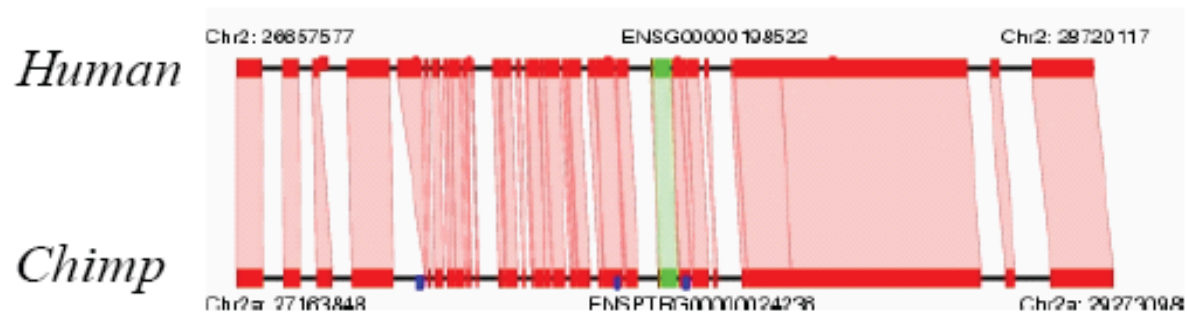
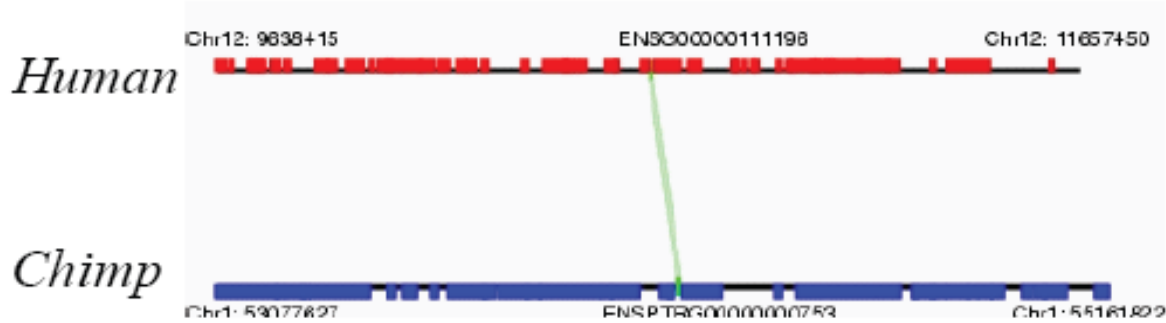


Fig 2: (A) Definition of conservation of genomic neighborhood (CGN). The gene of interest is shown in green and the neighbors of human genes are shown in red, labeled 1 through 4. The different evolutionary scenarios, resulting in a CGN score between 0 and 1, while comparing the neighborhood of orthologous chimpanzee and human genes are illustrated. Orthologs of the neighbors of the human gene in chimpanzee are shown in red and labeled 1' through 4'. Neighbors of the chimpanzee gene that are not neighbors of the human ortholog are shown in blue. Because of the way in which CGN score is calculated, the identification of genes with low CGN would include both the set of genes that have changed their neighbourhood either in the human lineage or in the chimpanzee lineage after the split from their common ancestor

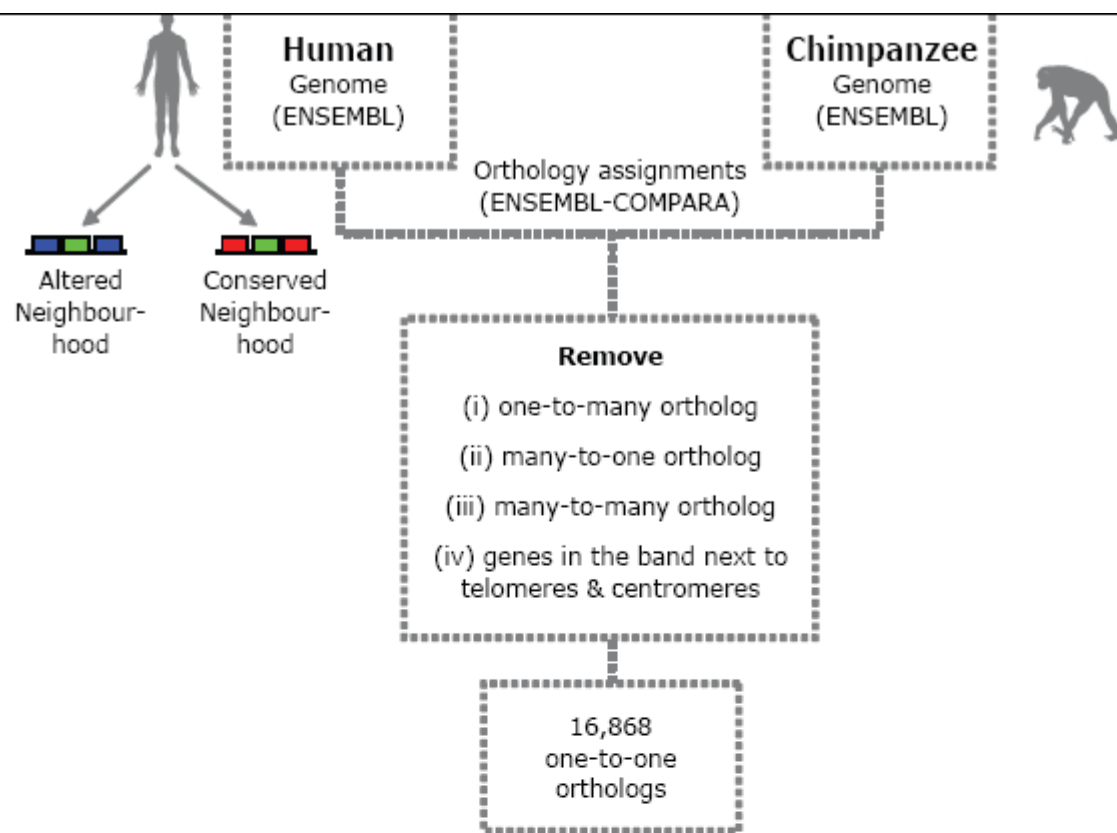
ENSG00000198522 (Znf512, CGN=0.88)



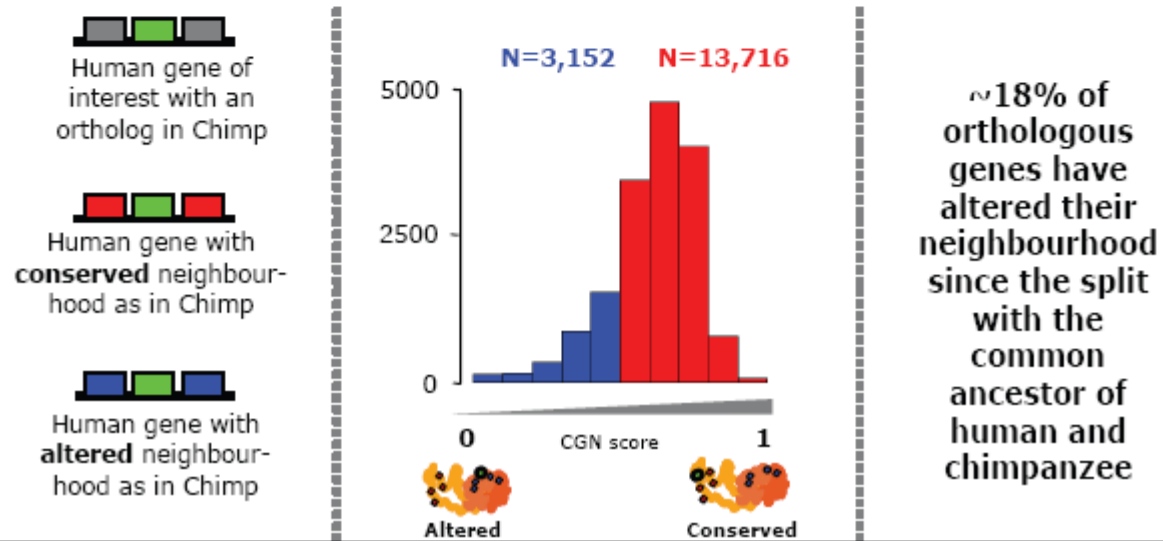
ENSG00000111196 (Mgn2, CGN=0)



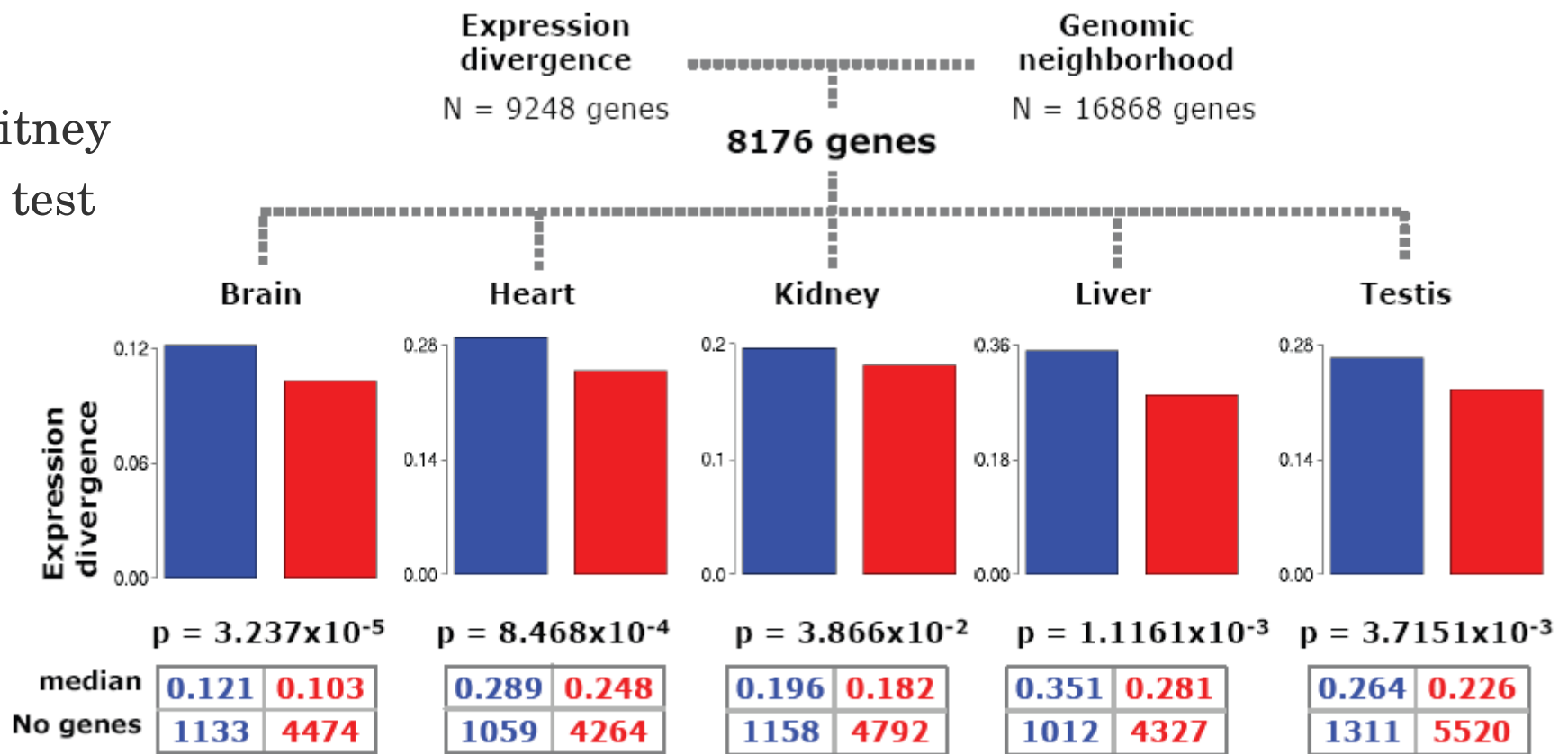
(B) An example of a gene with highly conserved genomic neighborhood (CGN=0.88, Znf512) and completely altered genomic neighborhood (CGN=0; Mgn2)



Calculate CGN



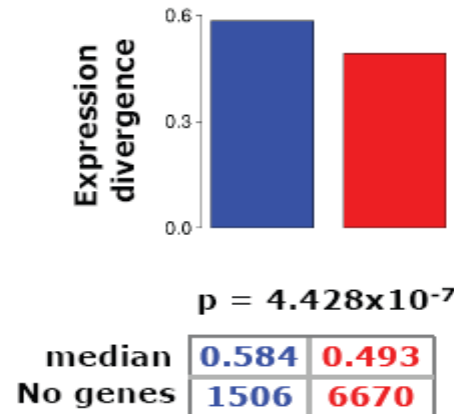
Mann-Whitney
two-tailed test



Key:

CGN ≤ 0.5
 CGN > 0.5

Maximal divergence across tissues



Genes with altered neighbourhood (low CGN) display higher expression divergence

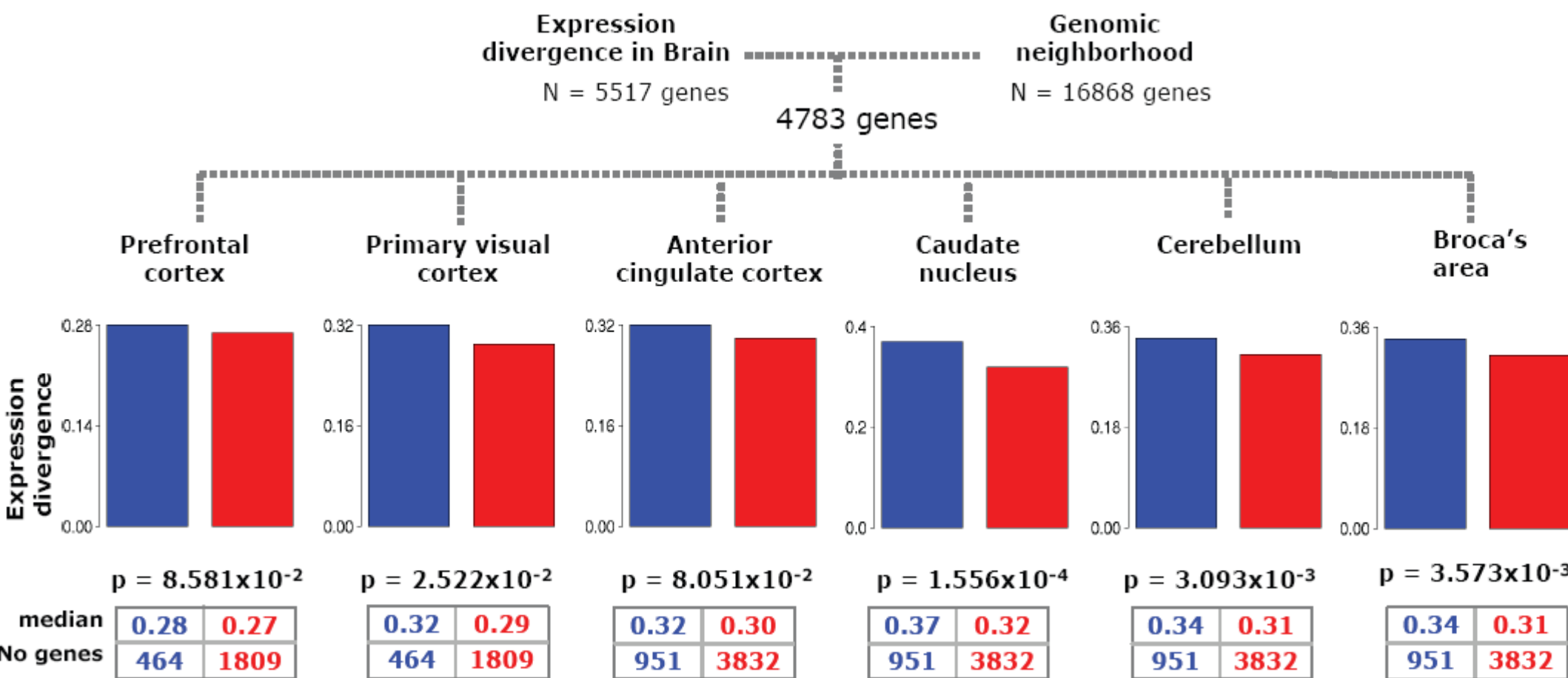


Fig 4: Investigating the impact of alteration in gene neighborhood on gene expression divergence in different brain parts. Distributions of the expression divergence value for genes with conserved (red) and altered (blue) neighborhood are compared using the Mann-Whitney test for six different brain parts. For visual clarity, median expression and sample size (*i.e.*, number of genes) for genes with conserved and altered neighborhood categories are shown for each brain part. Please see **SI-8** for control calculations.

Expression level
across 72 tissues
N = 17,185 genes
(33,459 transcripts)

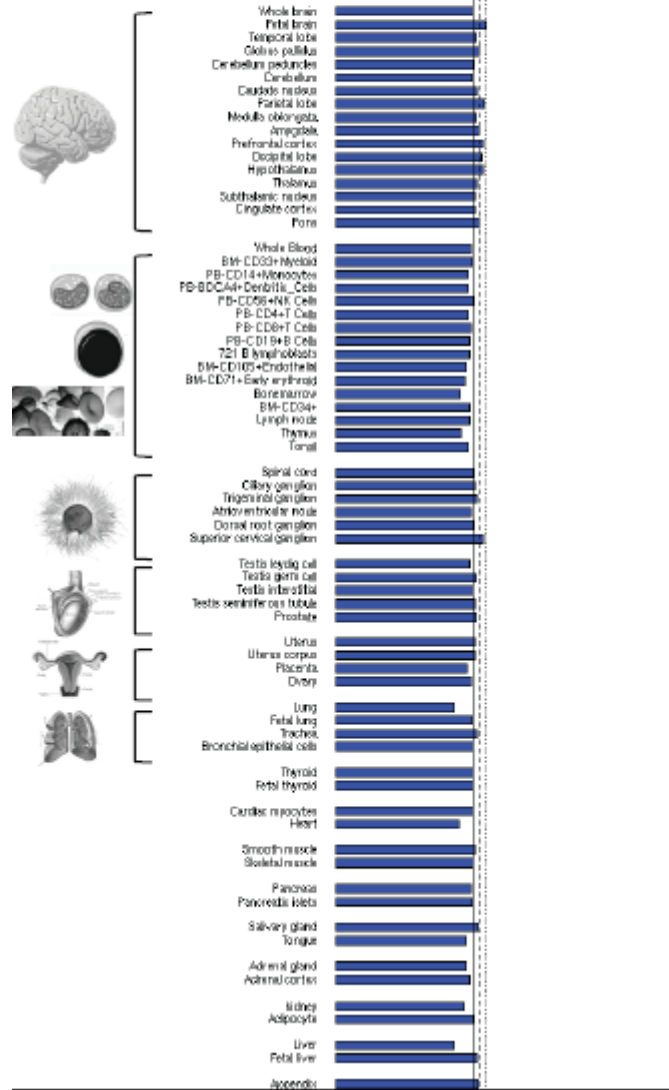
13,963 genes
(26,038 transcripts)

Genomic
neighborhood
N = 16,868 genes

A

Fraction of expressed
transcripts that map to genes
with low CGN score

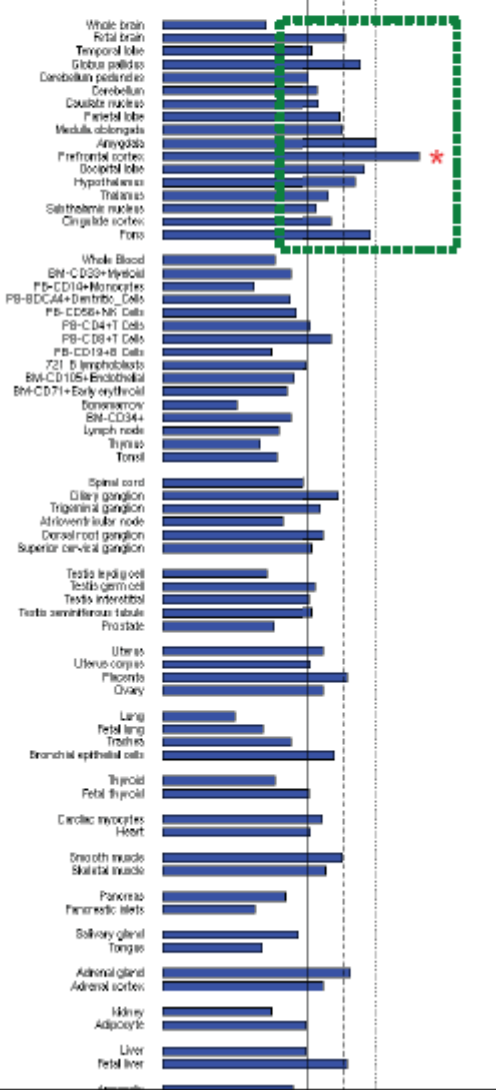
0.0 0.1 0.2 0.3 0.4



B

Fraction of tissue specifically expressed
transcripts that map to genes with
low CGN score

0.0 0.1 0.2 0.3 0.4



Conclusions

Considerable fraction of genes have **changed** their **genomic neighbourhood** in human and chimpanzee after the split from the common ancestor

Genes with **altered neighbourhood** are more likely to undergo expression divergence than genes with **conserved neighbourhood**

The same trend in six different brain parts that are equivalent between human and chimpanzee

The human brain shows enrichment to express genes in a **tissue-specific manner**

Expression divergence induced by this mechanism could have contributed to the phenotypic differences between humans and chimpanzee