

# Evolutionary differences in binding sites of a transcription factor involved in behavior

- Humans and chimpanzees differ in their behavior, e.g. social behavior:



Stable communities of multiple males and females  
Altruistic behavior only toward related individuals



Monogamous / family  
Stronger pro-social and cooperative

- EGR1 is a transcription factor that is regulating behavior, e.g. social behavior, curiosity etc. It does so by binding to specific DNA sequence motifs in the genome, which leads to an activation of the genes next to these motifs  
These genes are then important for e.g. synaptic activities that control behavior

→ Does EGR1 regulate different genes in humans and chimpanzees?

## Exercises:

### 1. Find all genes in the human genome that have at least one binding site for EGR1

- What is the binding motif for EGR1? Is there more than one binding motif for EGR1?
- Plot the EGR1 motif(s)
- Download the promoters for all human transcripts (defined here as 5000 bp upstream of the transcription start site)
- Search for EGR1 motifs in the human promoters (Note: position frequency matrix needs to be converted to position count matrix by multiplying with 100)
- How many human promoters have at least one EGR1 motif? How many EGR1 motifs do human promoters have on average?
- Make a table or data frame containing the gene names and numbers of EGR1 motifs for each promoter

## **2. Find differences in EGR1 motifs between human and chimpanzee promoters**

- Search for EGR1 motifs in the promoters of all chimpanzee transcripts (defined again as 5000 bp upstream of the transcription start site)
- Which genes have more/fewer EGR1 binding sites in humans than in chimpanzees? (Note: You first need to find out which human and chimpanzee genes are orthologous)

## **3. Which EGR1 binding site changes are specific to humans?**

- Search for EGR1 motifs in all orang utan promoters
- Which genes have human specifically gained EGR1 binding sites? (Note: binding site should be present in humans but not in chimpanzees and not in orang utans)
- Which genes have human specifically lost EGR1 binding sites? (Note: binding site should be present in chimpanzees and orang utans but not in humans)
- What is the function of the genes with human specific changes in EGR1 binding sites?

### **Some libraries that should come in handy:**

- MotifDb
- Genomic Features
- biomaRt
- ...