

Gene regulation in primates evolves under  
tissue-specific selection pressures

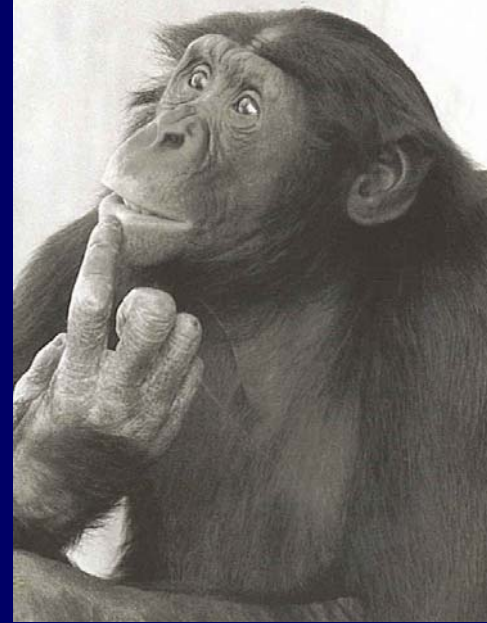
Yoav Gilad

Department of Human Genetics  
University of Chicago



*Gilad Lab*

Russ Bainer  
Ran Blekhman  
Carolyn Cain  
Claudia Garcia  
Athma Pai  
George Perry  
Luis Barreiro



*Former members*

Paola de Candia  
Adrien Chabot

*Collaborators*

Alicia Oshlack  
Gordon Smyth  
Terry Speed  
Matthew Stephens  
John Marioni  
Jonathan Pritchard

# Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty (Britten and Davidson 1971)

Recent experimental information on DNA sequence repetition is reviewed, and the significance of both repetitive and non-repetitive sequence considered. Included are a summary of data on the distribution of genome sizes in animals, new experiments on interspecific DNA homology, the distribution of sequence frequencies, and the interspersion of repetitive sequences within the genome. Aspects of the process of evolution are considered at the level of change in the DNA. the process by which novel structure and function could have arisen during evolution are considered speculatively in terms of the authors' gene regulation theory.

We suggest that evolutionary changes in anatomy and way of life are more often based on changes in the mechanisms controlling the expression of genes than on sequences changes in proteins. (King and Wilson 1975)

# The Central Dogma



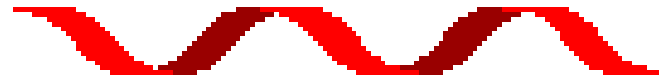
**DNA**

Rate of transcription

**Transcription:**  
the synthesis of  
an RNA copy of a  
segment of DNA



RNA modifications



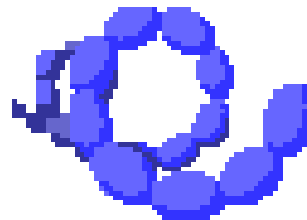
**RNA**

RNA stability

**Translation**



Rate of translation



Protein modifications

Protein stability

Protein interactions

**Protein**

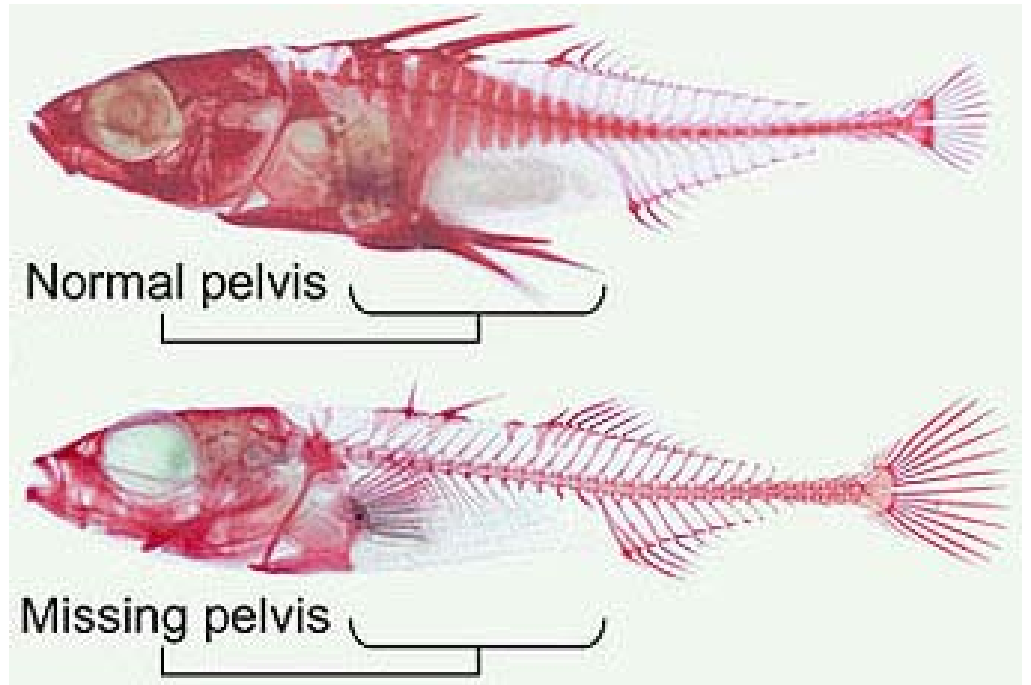
Steady state



**RNA**



## Sticklebacks





Like butterflies, different species of fruit flies decorate their wings with a great diversity of spots and patterns. (Credit for photo and caption: Nicolas Gompel and Benjamin Prud'homme, University of Wisconsin-Madison)



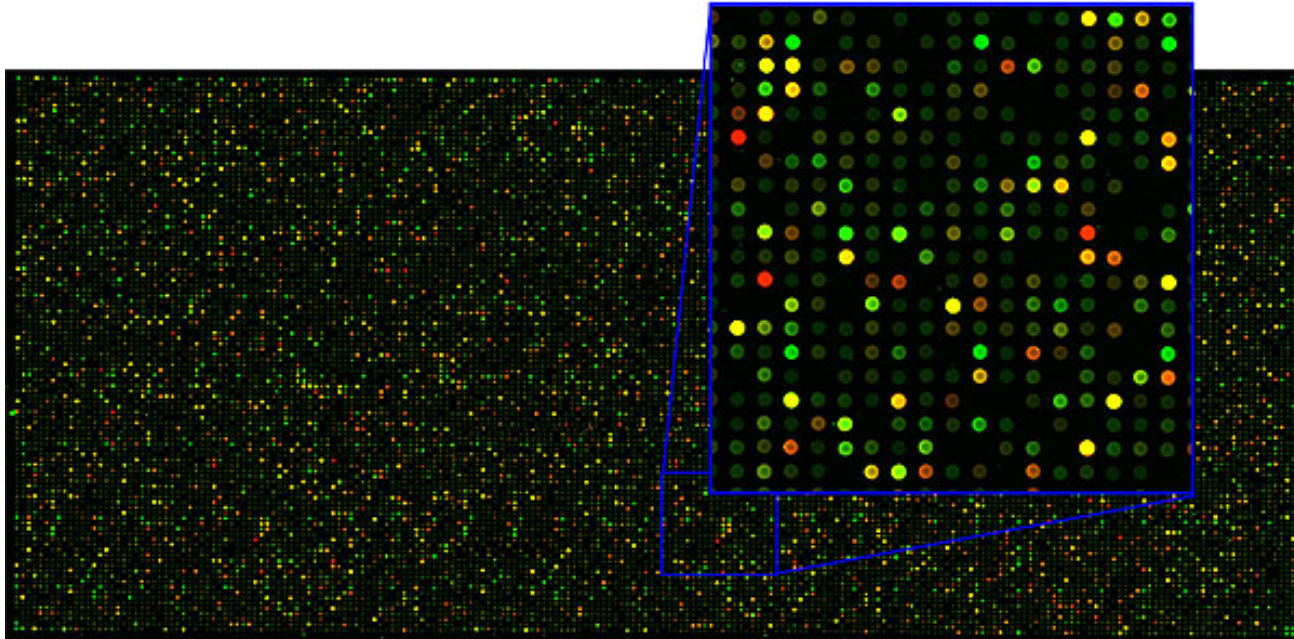
# What about primates?

Studying phenotypic evolution in primates is more difficult in this respect, because key experiments often cannot be performed to distinguish between competing hypotheses or to estimate important parameters (such as the neutral mutational variance). Moreover, material is often scarce, leading to largely unknown and uncontrolled environmental variance between samples. These limitations are particularly problematic for dynamic, environmentally sensitive traits like gene expression.

# Outline

- Inter-primate comparison of gene expression
- Identifying the direct regulatory targets of transcription factors
- Finding inter-species differences in regulatory pathways

## Multi-species microarray



The microarray contains probes from 18,109 genes, with the vast majority of genes represented by seven probes per gene per species, for a total of ~370,000 probes

# Study design

**3 Tissues**

Liver

Kidney

Heart

**3 Species**

Human

Chimp

Rhesus

**6 Samples**

H1 H2 H3 H4 H5 H6

C1 C2 C3 C4 C5 C6

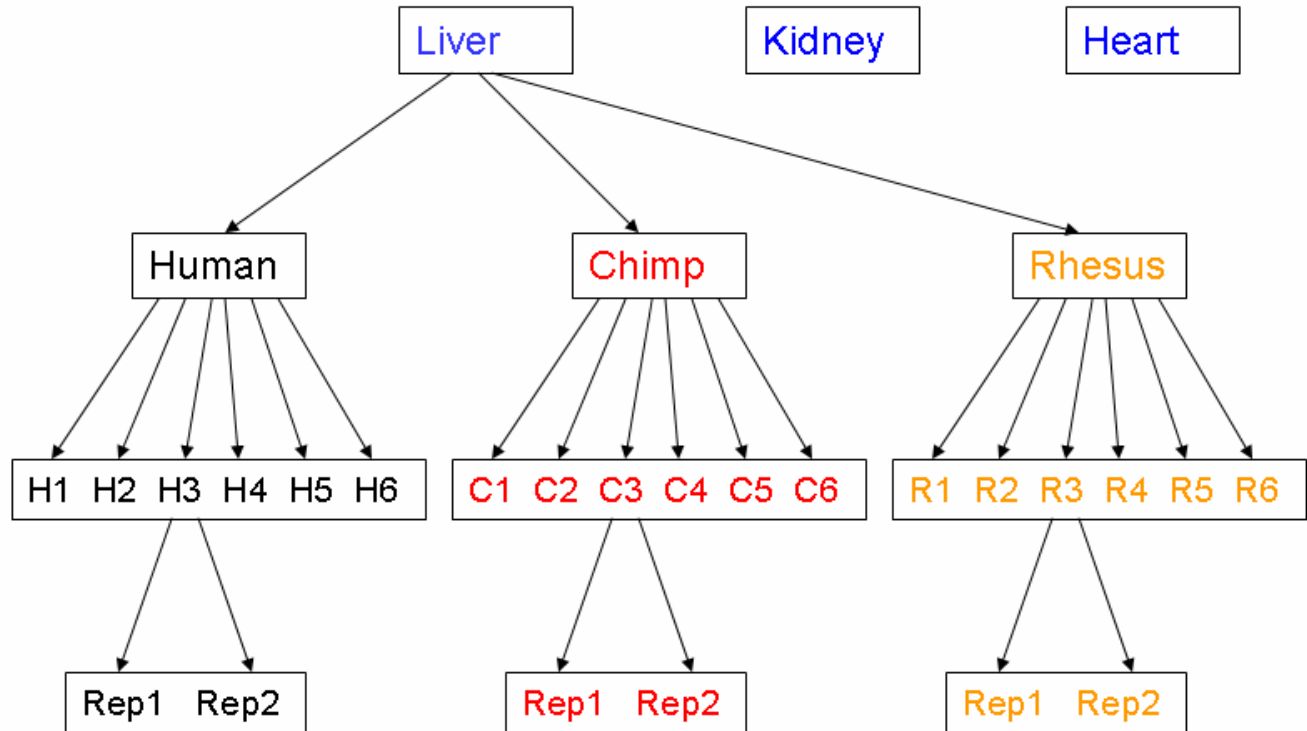
R1 R2 R3 R4 R5 R6

**2 Replicates**

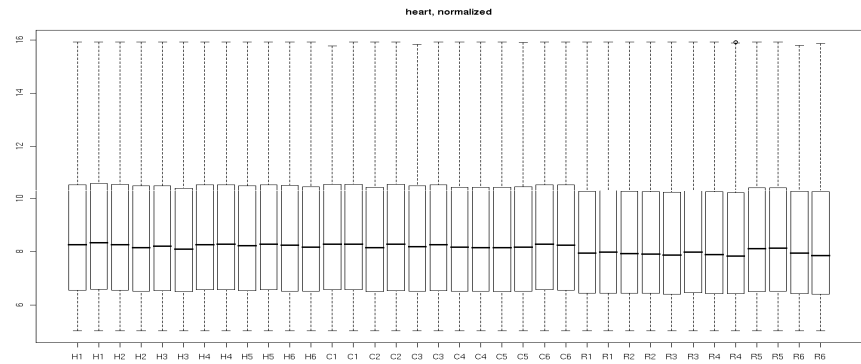
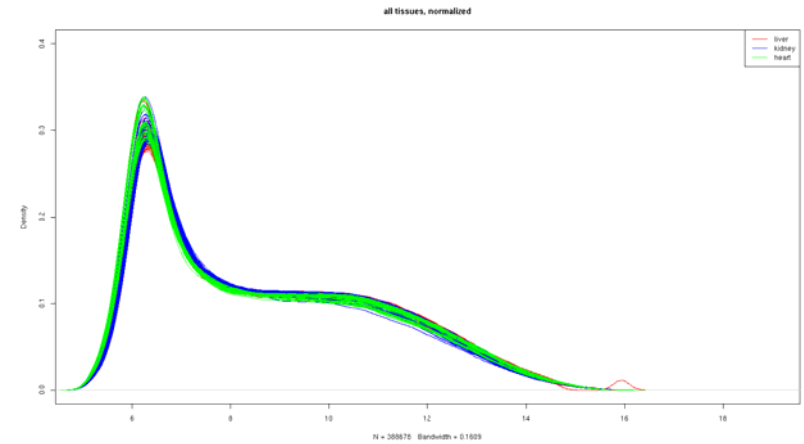
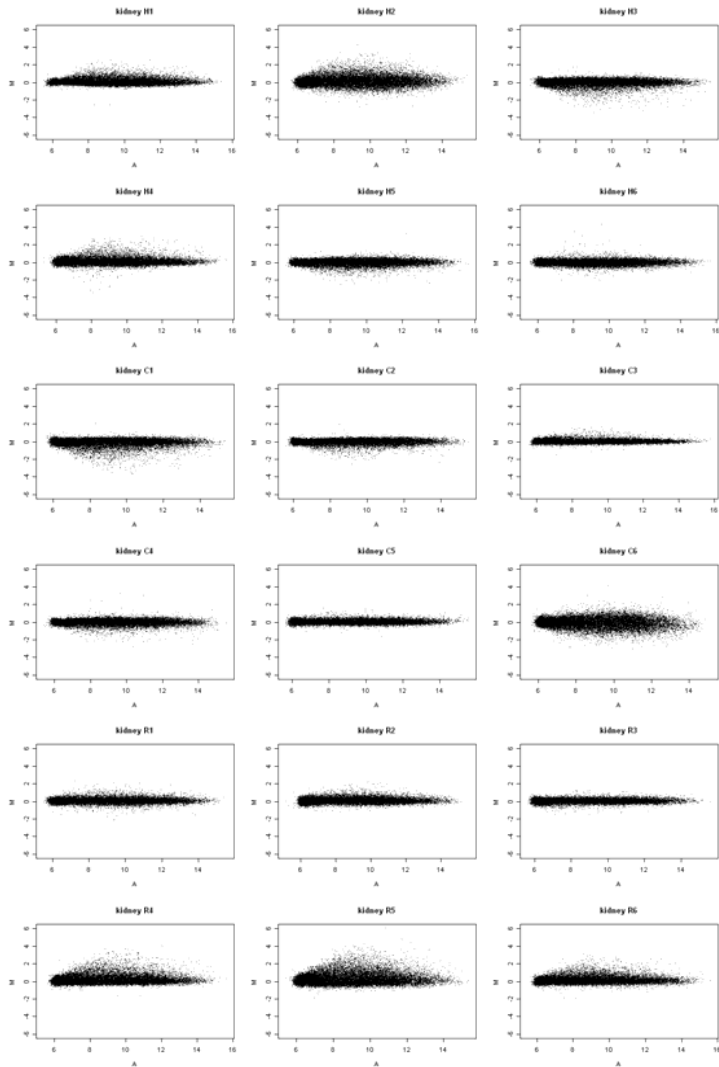
Rep1 Rep2

Rep1 Rep2

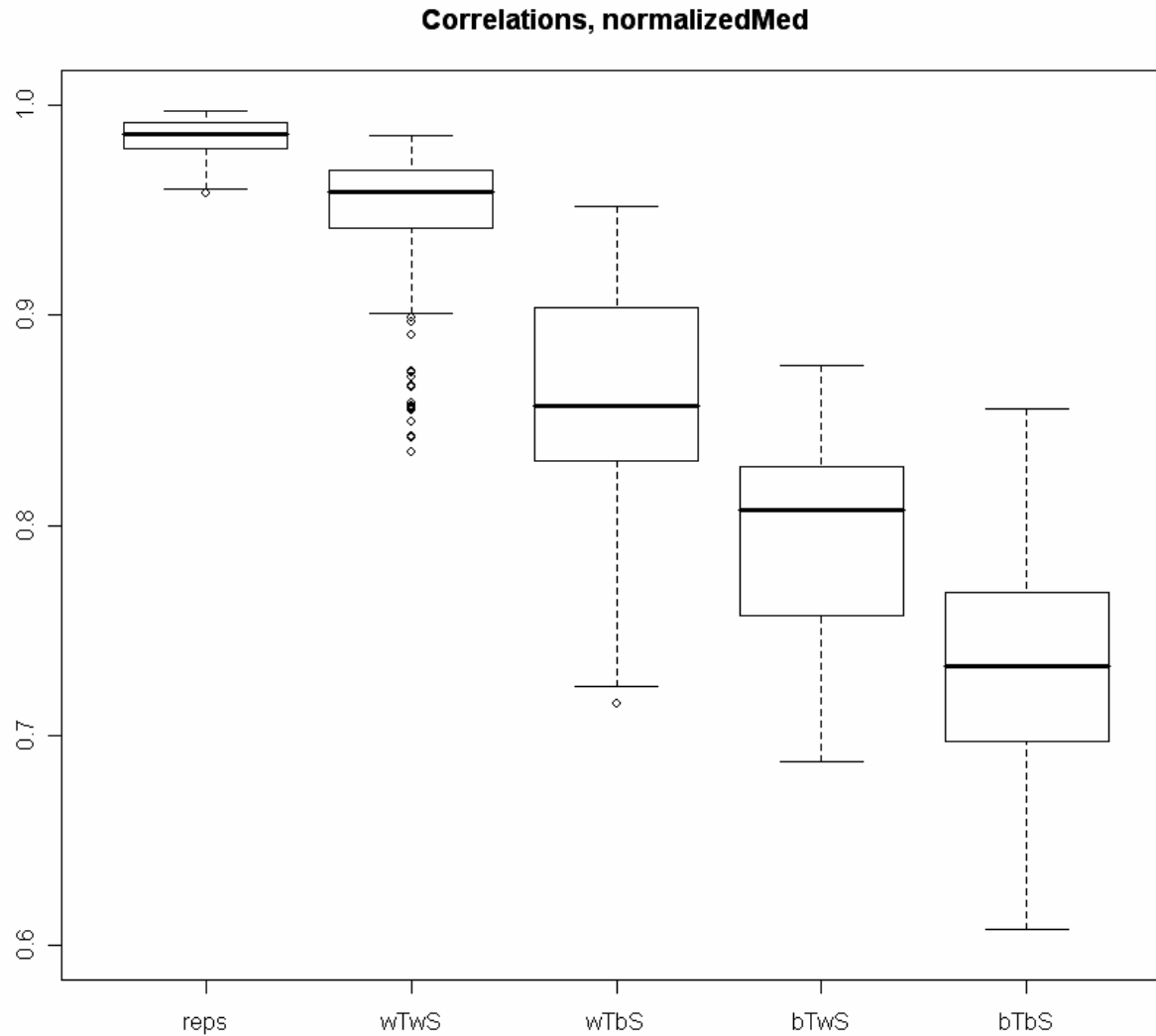
Rep1 Rep2



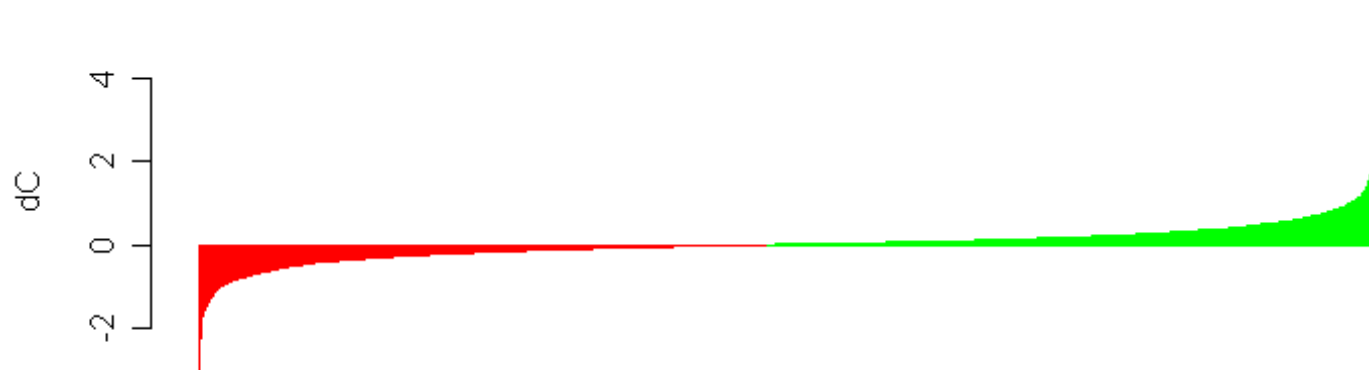
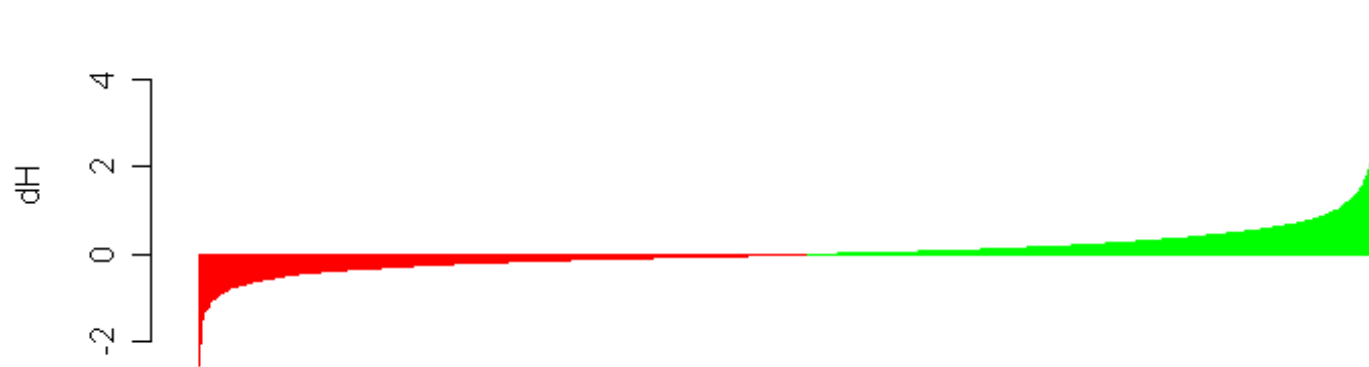
# Quality controls (normalized background corrected values)



# Quality controls (normalized background corrected values)



# Quality controls (normalized background corrected values)



# Linear model

$$Y_{stpi} = \mu_{st} + \kappa_{scp} + \pi_p + \rho_{ps} + \alpha_{sti} + \varepsilon_{stpi}$$

Expression estimate:  
gene+species+tissue

Sequence  
mismatch

Species and  
probe

Individuals



## Inter-species and inter-tissue differences in gene expression

<b>DE between tissues (within species)</b>	<b>liver-kidney</b>	<b>liver-heart</b>	<b>kidney-heart</b>
Human	2810	2662	2124
Chimpanzee	2590	2894	2222
Rhesus Macaque	2522	2768	2215
<b>DE between species</b>	<b>Liver</b>	<b>Kidney</b>	<b>Heart</b>
Human vs. Chimpanzee	2809	3368	3197
Human vs. Rhesus Macaque	5525	6250	5545
Chimpanzee vs. Rhesus Macaque	4871	6270	5021

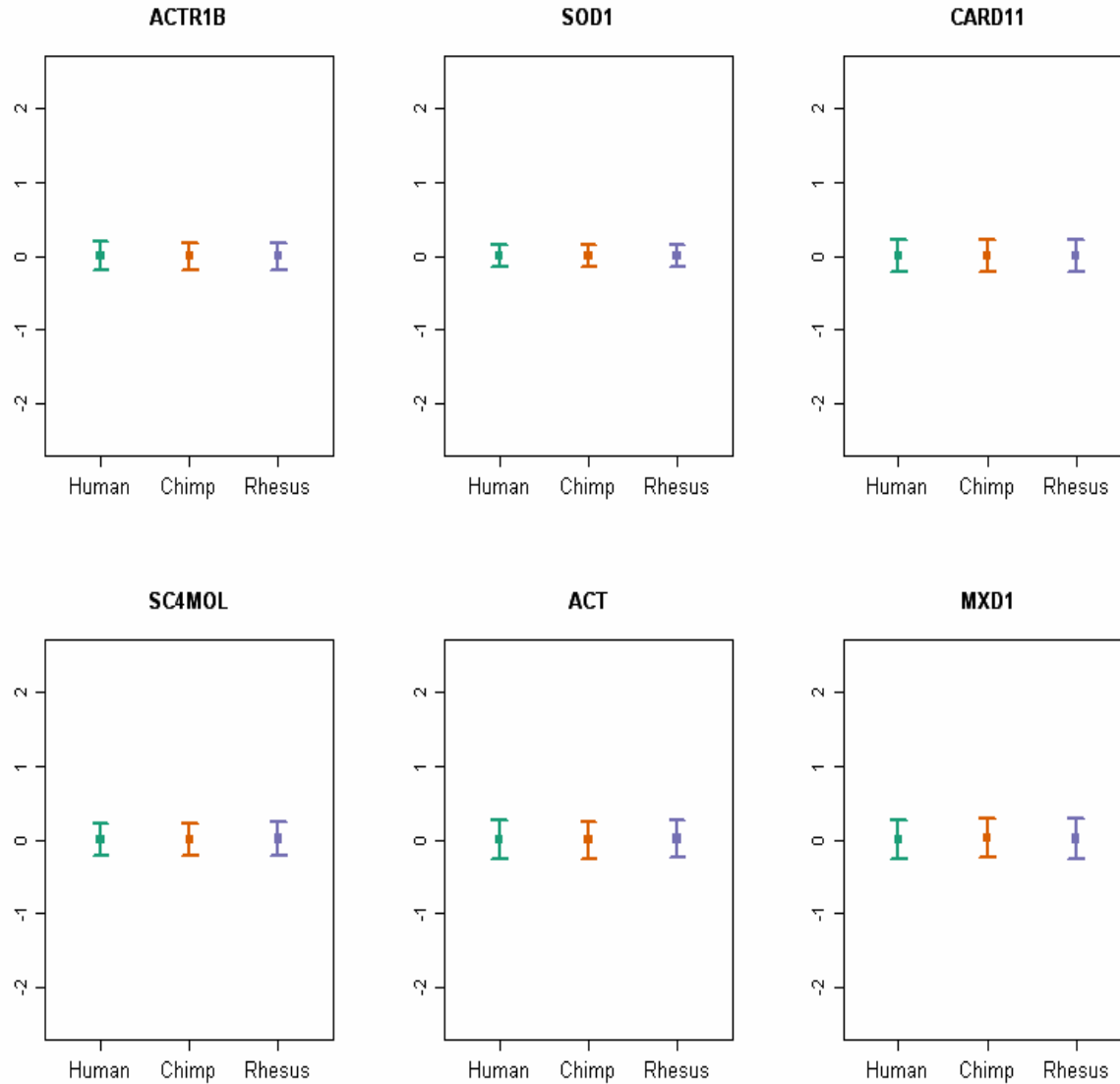
With this comparative expression data we can -

- Study differences in regulatory pathways between species
- Study differences in regulatory networks between species
- Study processes that shape regulatory evolution (such as duplications and chromosomal rearrangements)

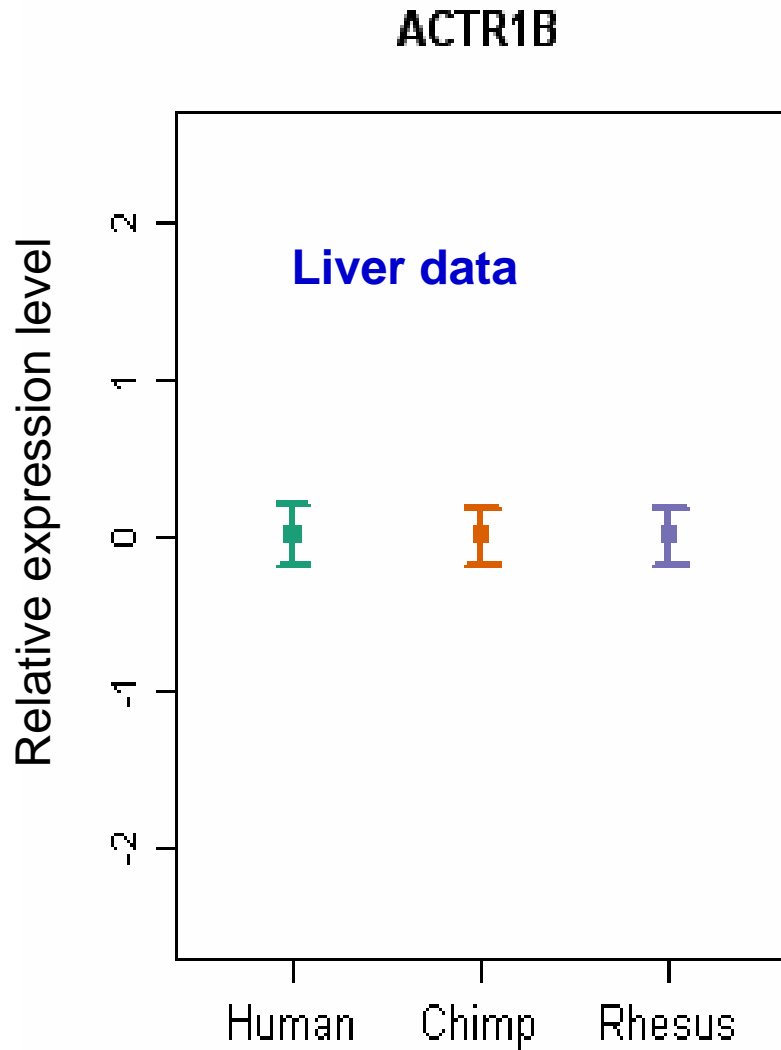
How to identify genes whose regulation evolves under natural selection?

Can we use test for departures from a null model of 'no selection?

# Stabilizing selection on gene regulation

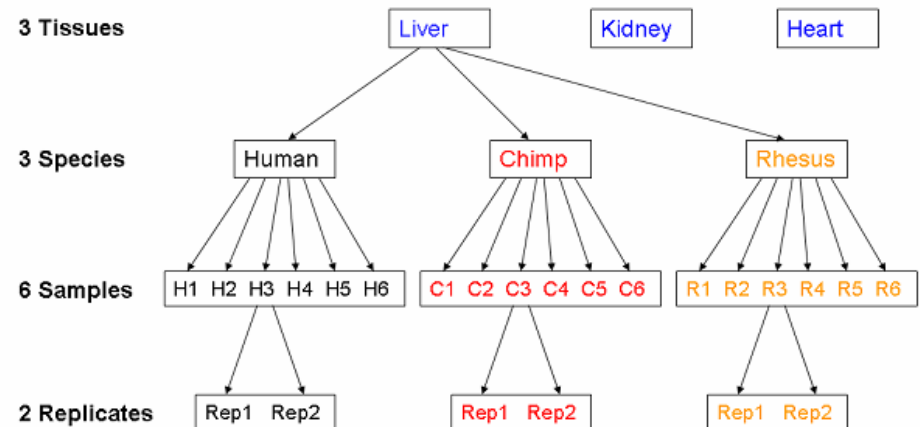


# Stabilizing selection on gene regulation

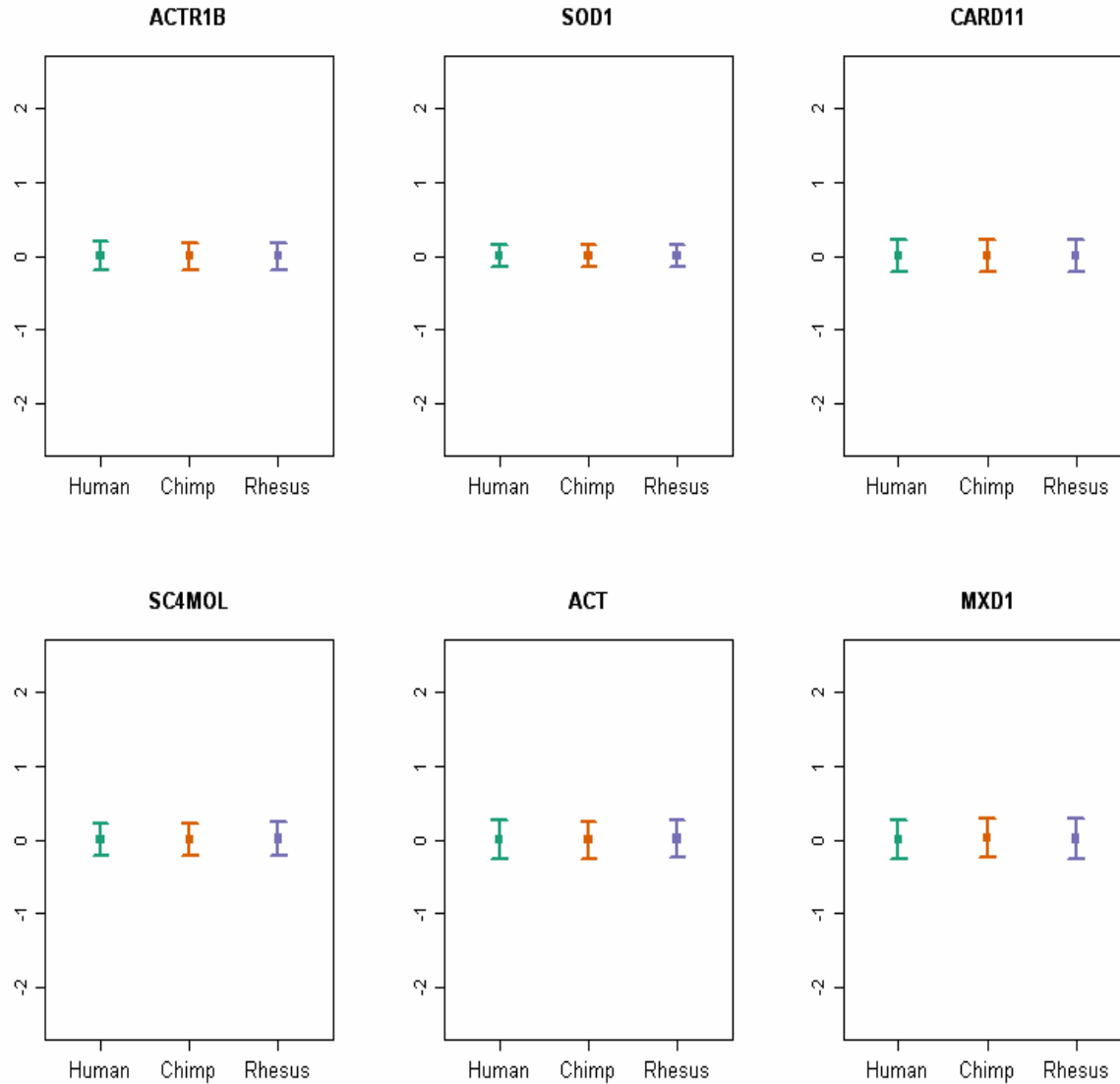


- Human
- Chimpanzee
- Rhesus

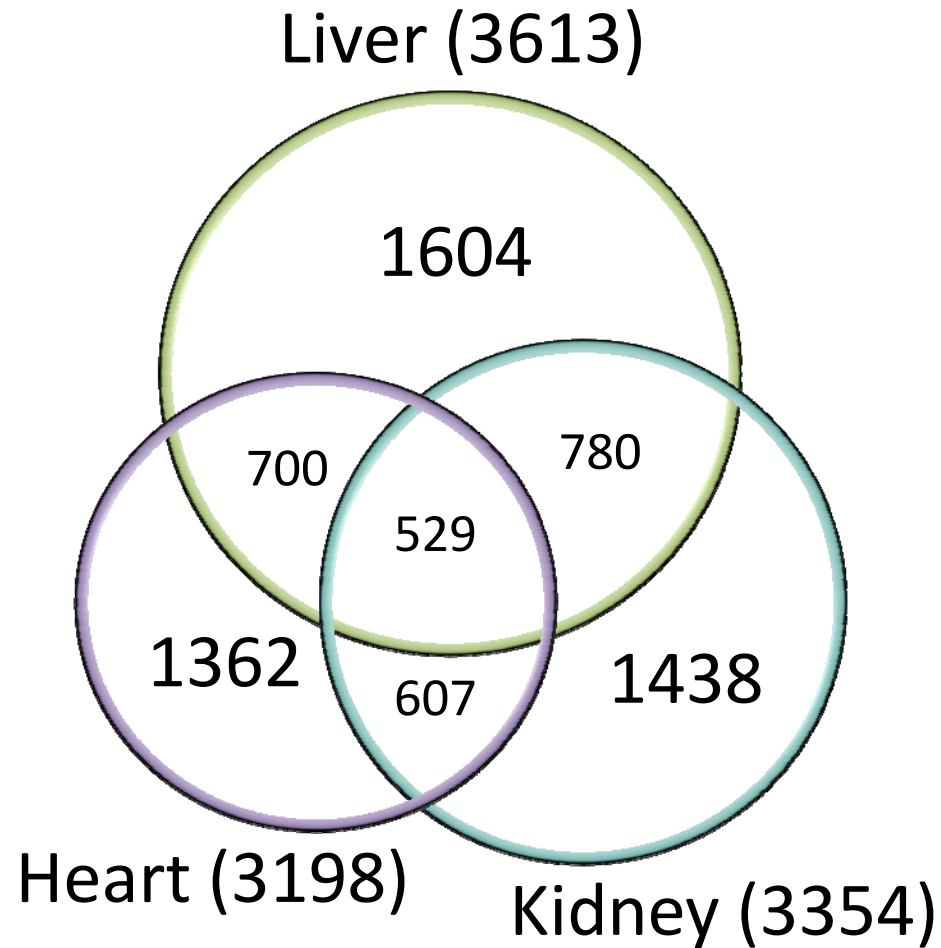
## Study design



# Stabilizing selection on gene regulation (examples from liver data)



## Stabilizing selection on gene regulation



# Stabilizing selection on gene regulation

<b>Tissue</b>	<b>Category</b>	<b><i>P</i>-value</b>
Liver (functional categories)	Housekeeping	$< 10^{-13}$
	Metabolic (GO)	$< 10^{-9}$
	Transcription factors (GO)	$< 10^{-4}$
	Transcription factors (validated)	$< 10^{-4}$
	Associated with cancer	$< 10^{-3}$
Liver (pathways)	Methionine metabolism	$< 10^{-3}$
	Complement and coagulation cascades	$< 10^{-3}$
	Adherens junction	$< 10^{-3}$
	Cell cycle	0.003
	TGF-beta signaling pathway	0.007
Kidney (functional categories)	Housekeeping	$< 10^{-7}$
	Transcription factors (GO)	0.002
	Transcription factors (validated)	0.013
	Metabolic (GO)	0.043
Kidney (pathways)	Amyotrophic lateral sclerosis (ALS)	$< 10^{-3}$
	Reductive carboxylate cycle (CO2 fixation)	0.020
	Ribosome	0.025
	Neurodegenerative Diseases	0.029
	Pathogenic Escherichia coli infection - EHEC	0.033
	Metabolic (GO)	$< 10^{-9}$



## Identifying genes whose regulation evolves under directional selection

*full:*  $\mu_s = \mu_H + \mu_C + \mu_R$

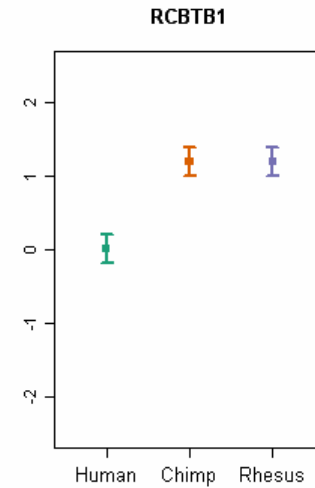
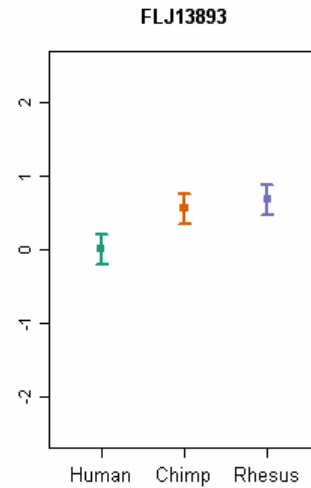
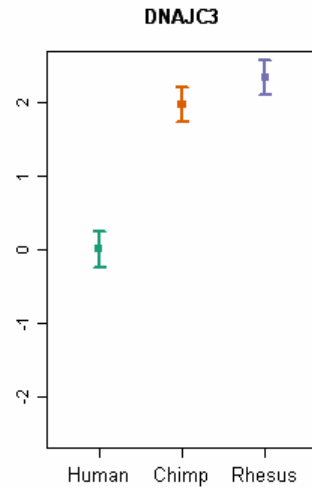
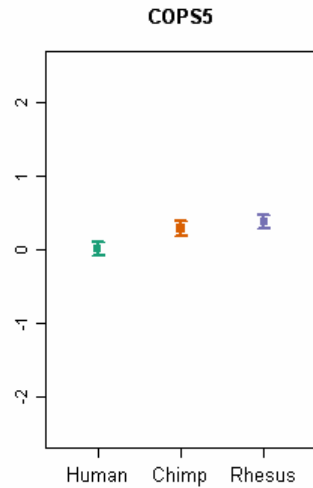
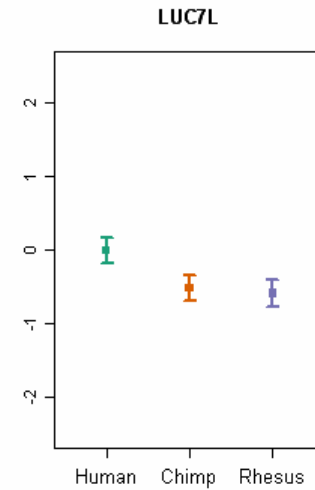
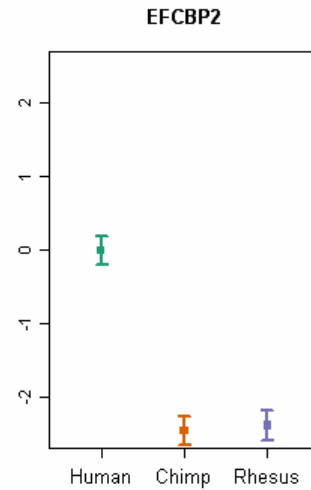
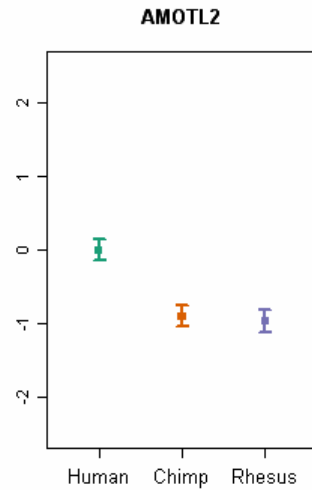
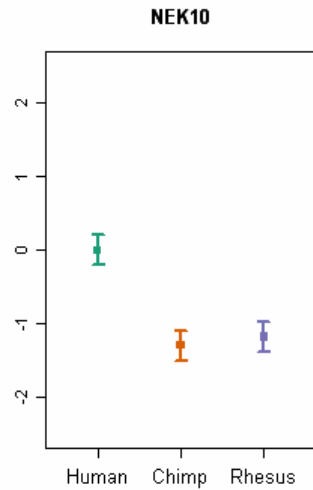
*reduced:*  $\mu_s = \mu_{HCR}$

*reducedCR:*  $\mu_s = \mu_H + \mu_{CR}$

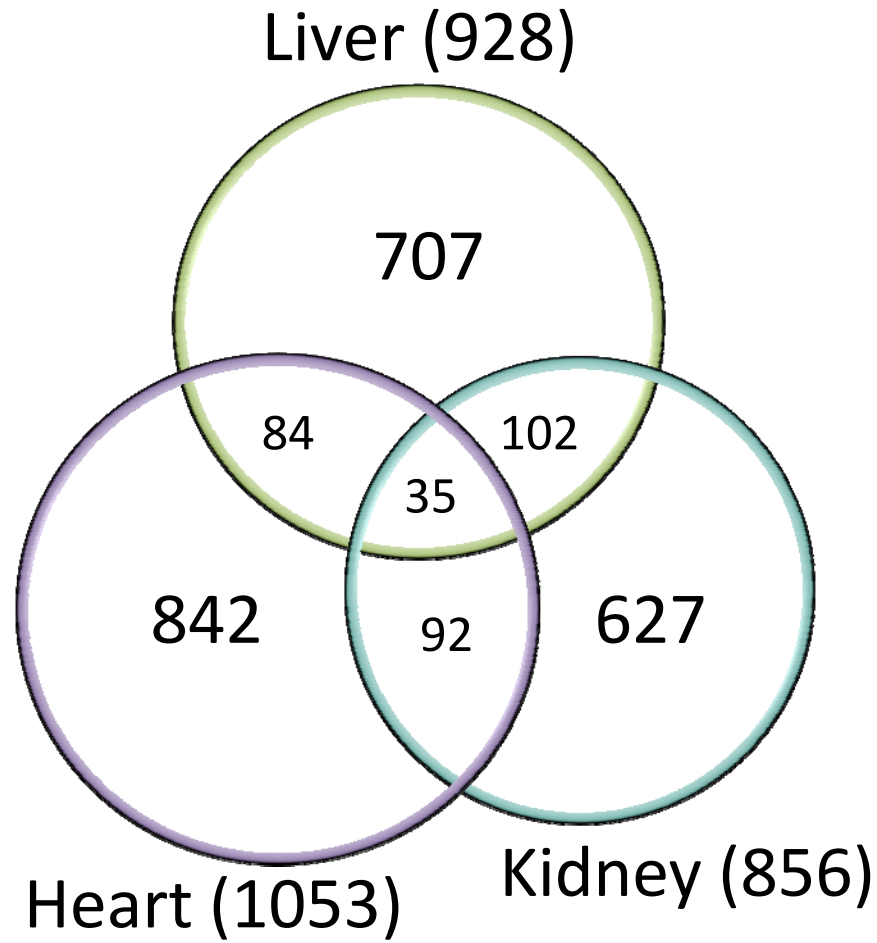
1.  $LR(x) = -\log_2 \left( \frac{L(\text{reducedCR}|x)}{L(\text{full}|x)} \right)$

2.  $LR(x) = -\log_2 \left( \frac{L(\text{reduced}|x)}{L(\text{reducedCR}|x)} \right)$

# Directional selection on gene regulation (examples from liver data)



## Directional selection on gene regulation



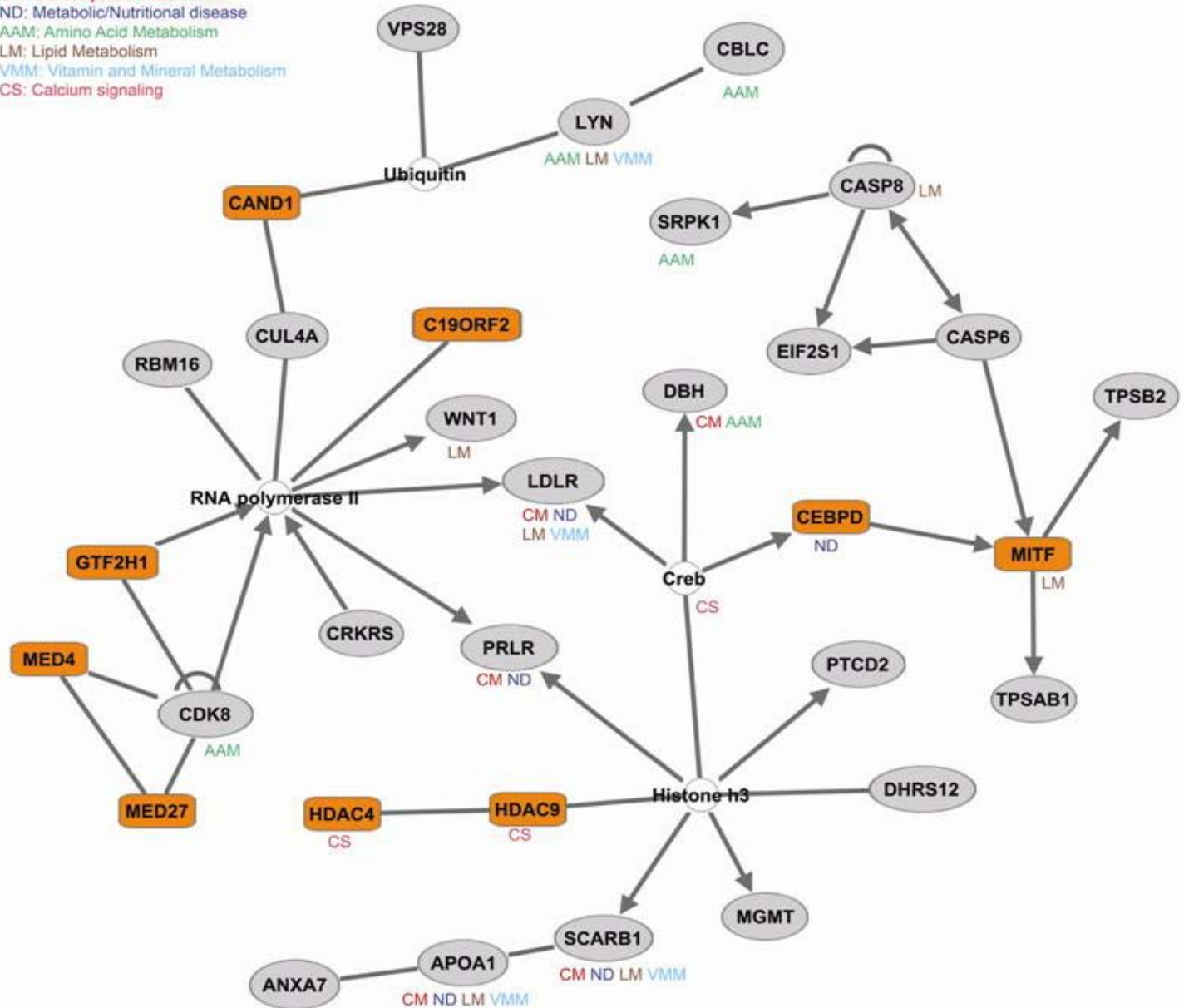
## Directional selection on gene regulation

---

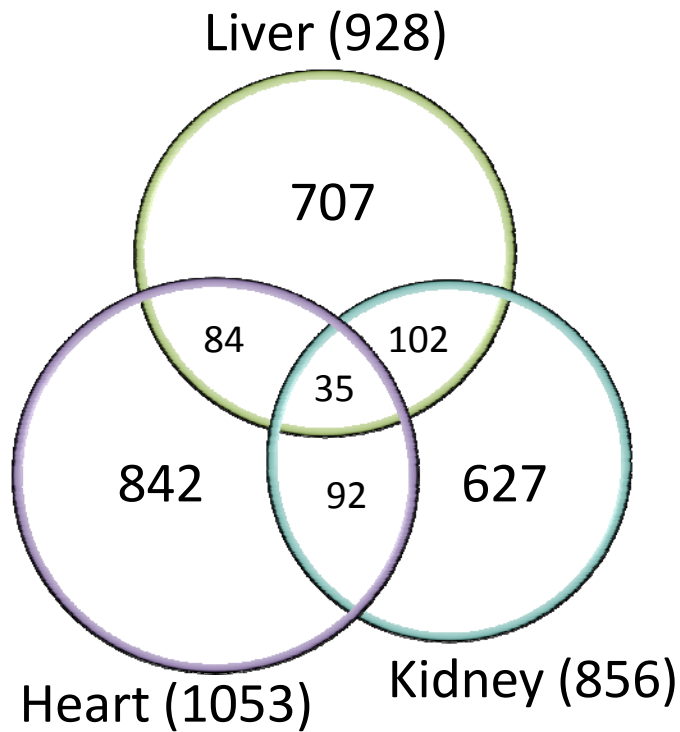
<b>Directional selection in any tissue (human higher or lower)</b>	Transcription factors (GO)	0.007
	Metabolic (GO)	0.012
	Tight junction	0.001
	Citrate cycle (TCA cycle)	0.002
	Adherens junction	0.014
	Glutamate metabolism	0.015
	Thyroid cancer	0.034

---

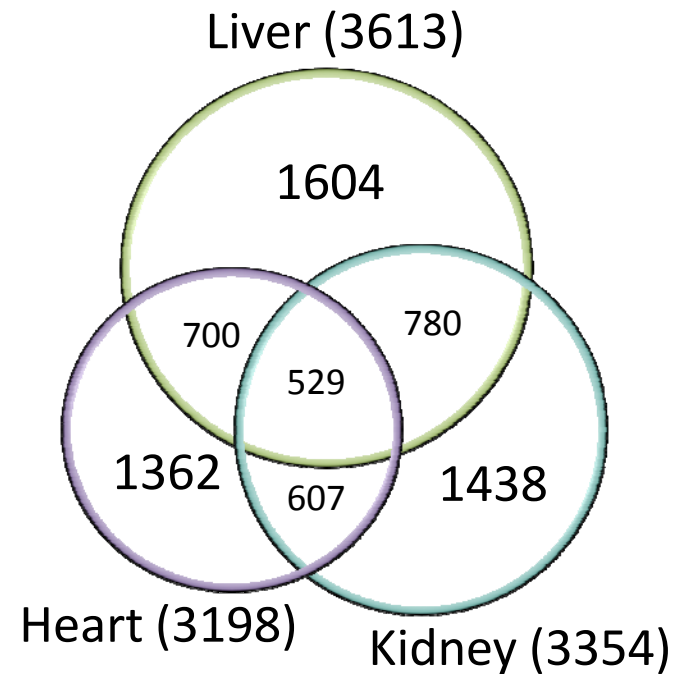
CM: Carbohydrate Metabolism  
 ND: Metabolic/Nutritional disease  
 AAM: Amino Acid Metabolism  
 LM: Lipid Metabolism  
 VMM: Vitamin and Mineral Metabolism  
 CS: Calcium signaling



# Tissue-specific selection on gene regulation

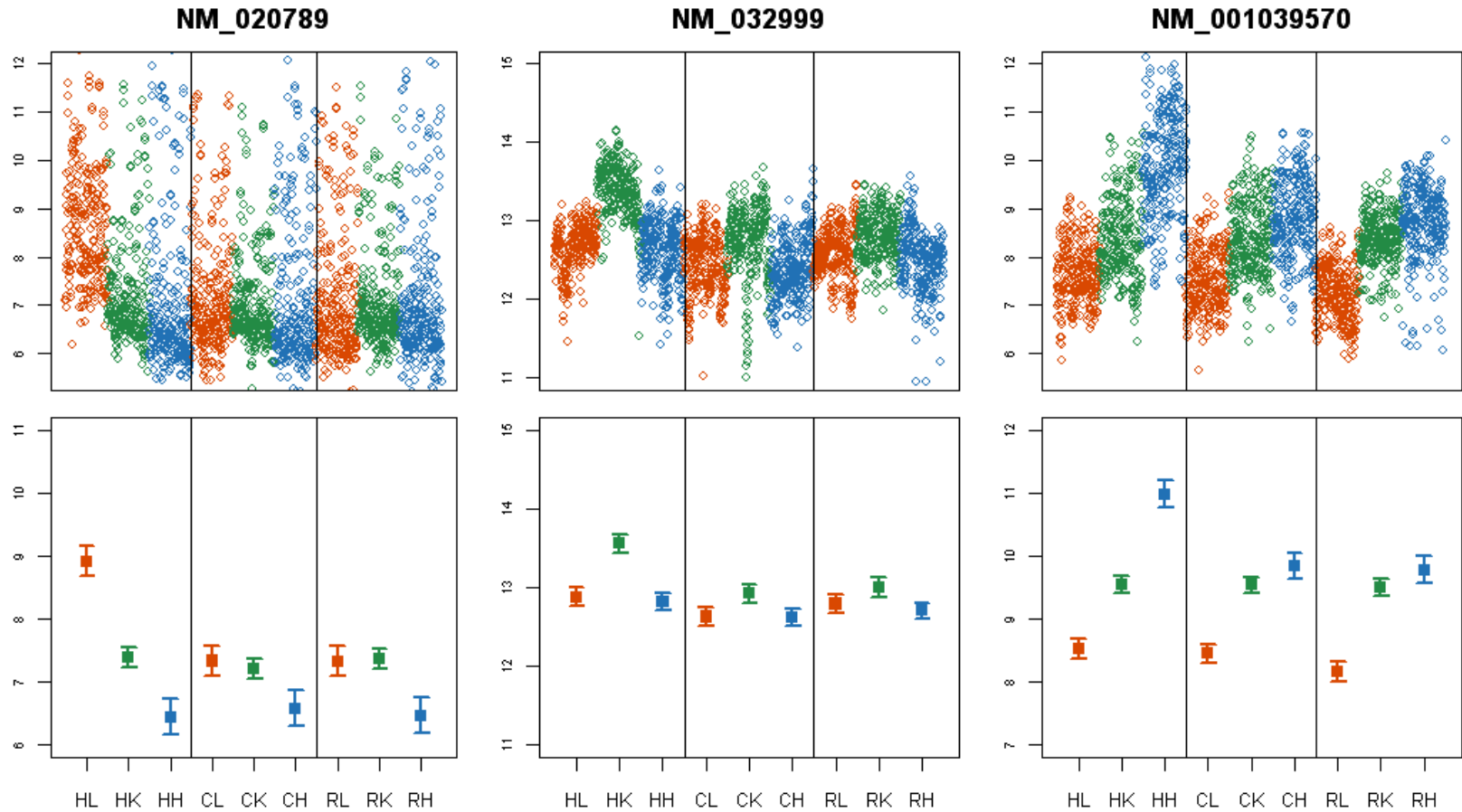


Stabilizing selection  
on gene regulation



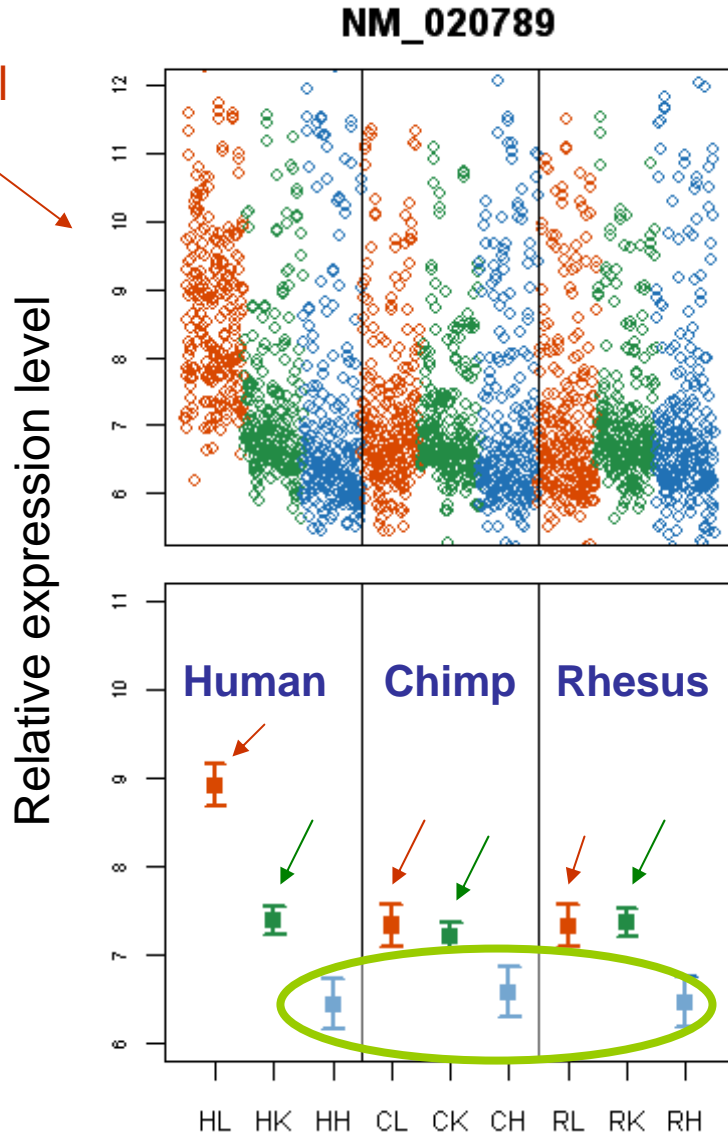
Directional selection  
on gene regulation

# Tissue-specific selection on gene regulation



# Tissue-specific selection on gene regulation

Probe level data

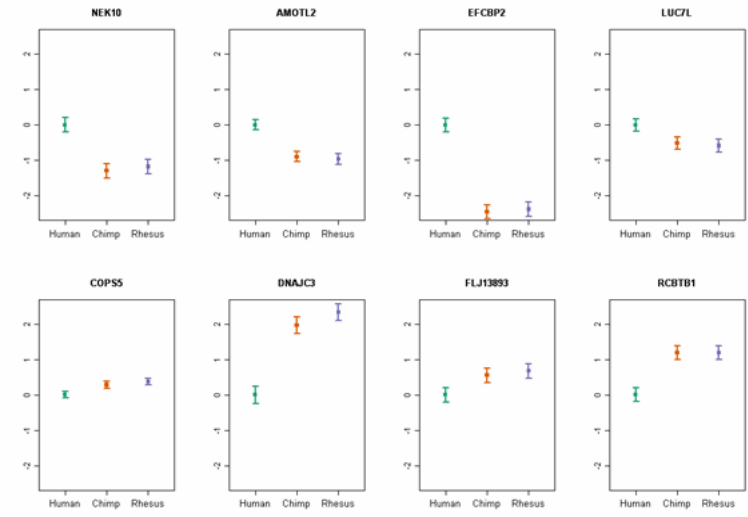


- Liver
- Kidney
- Heart

## Stabilizing selection on gene regulation

ACTR1B                      SOD1                      CARD11

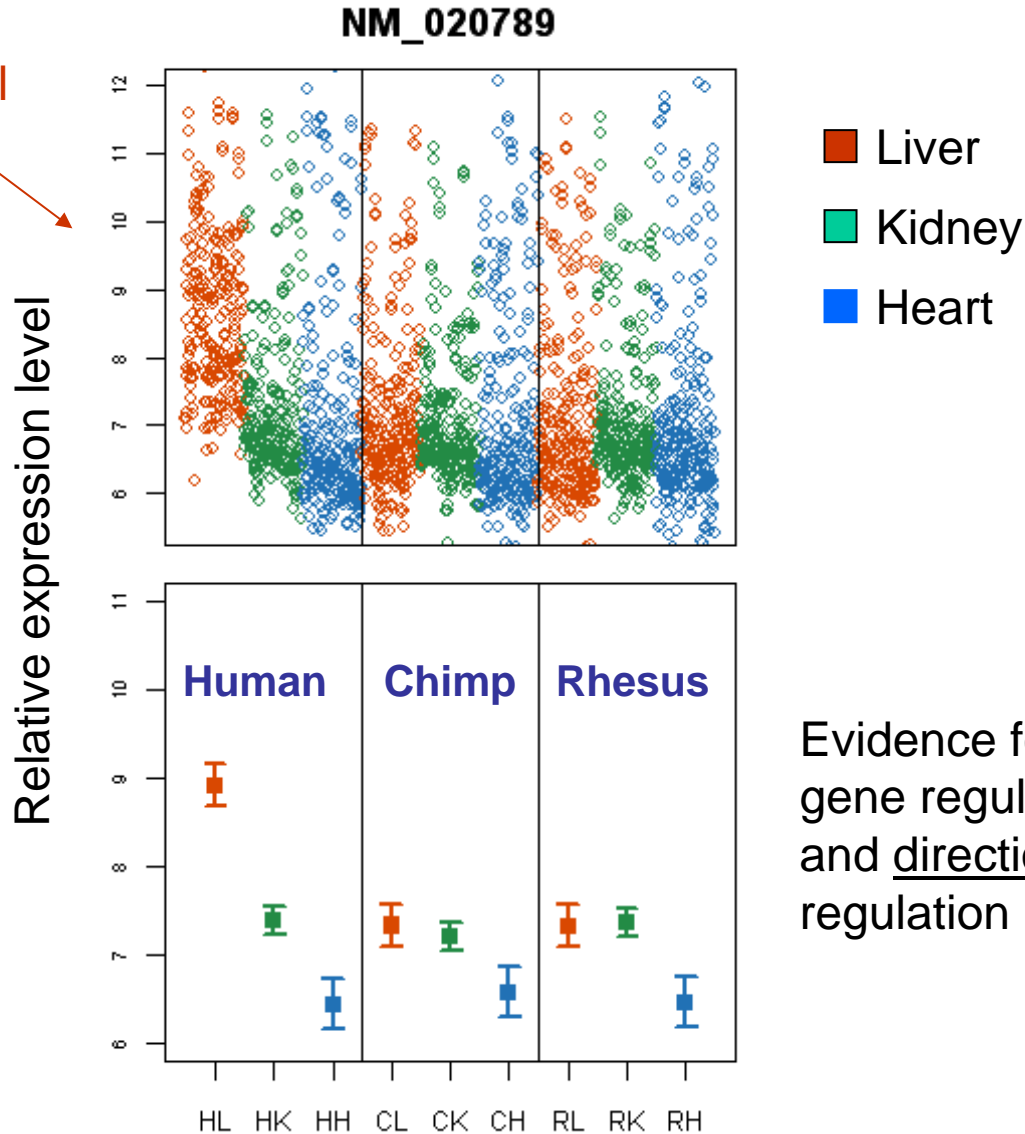
## Directional selection on gene regulation





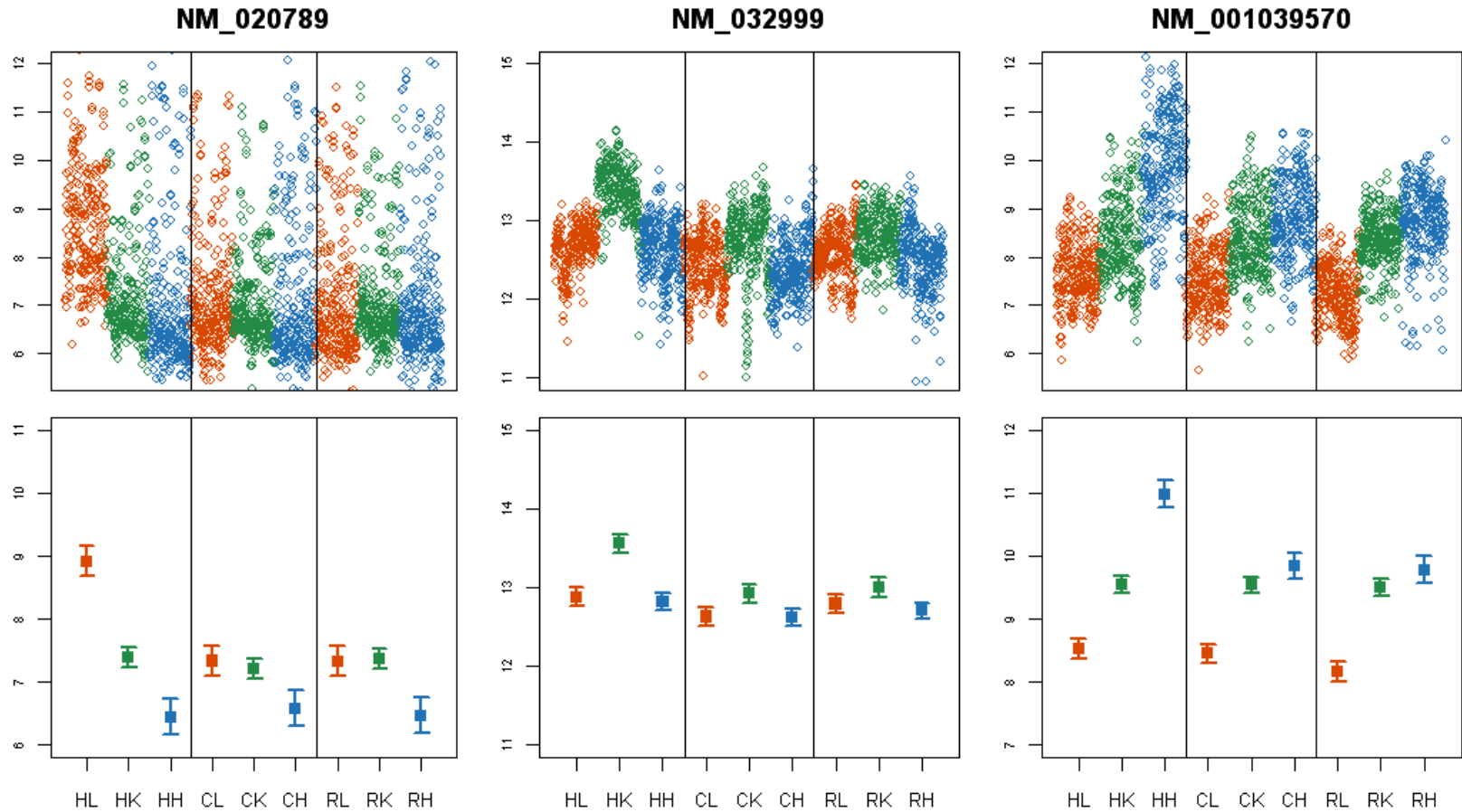
# Tissue-specific selection on gene regulation

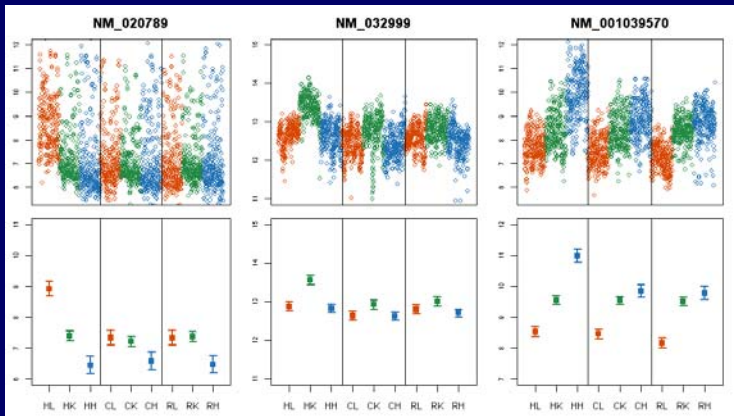
Probe level data



Evidence for stabilizing selection on gene regulation in **heart** and **kidney** and directional selection on gene regulation in the human **liver**

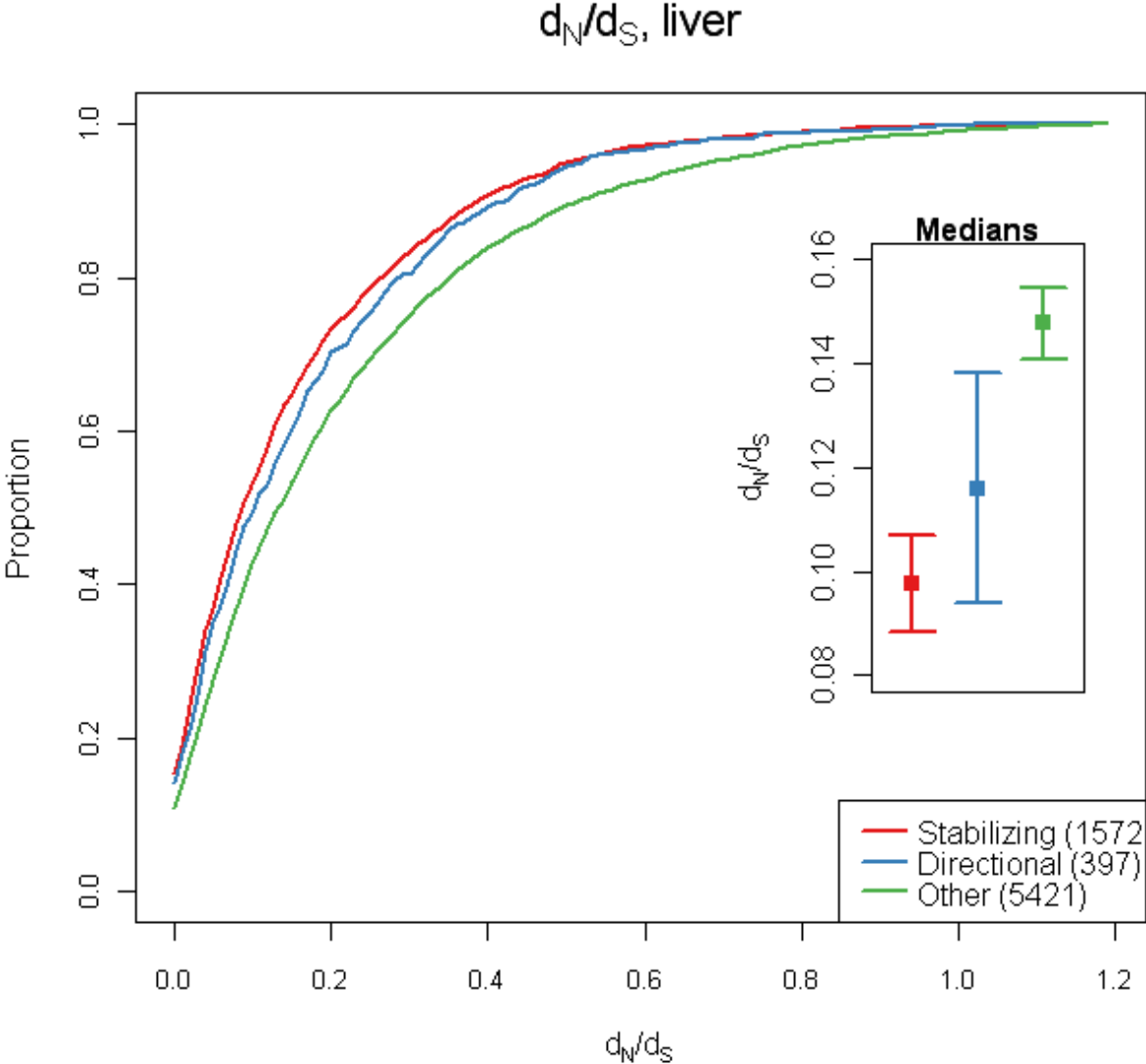
# Tissue-specific selection on gene regulation





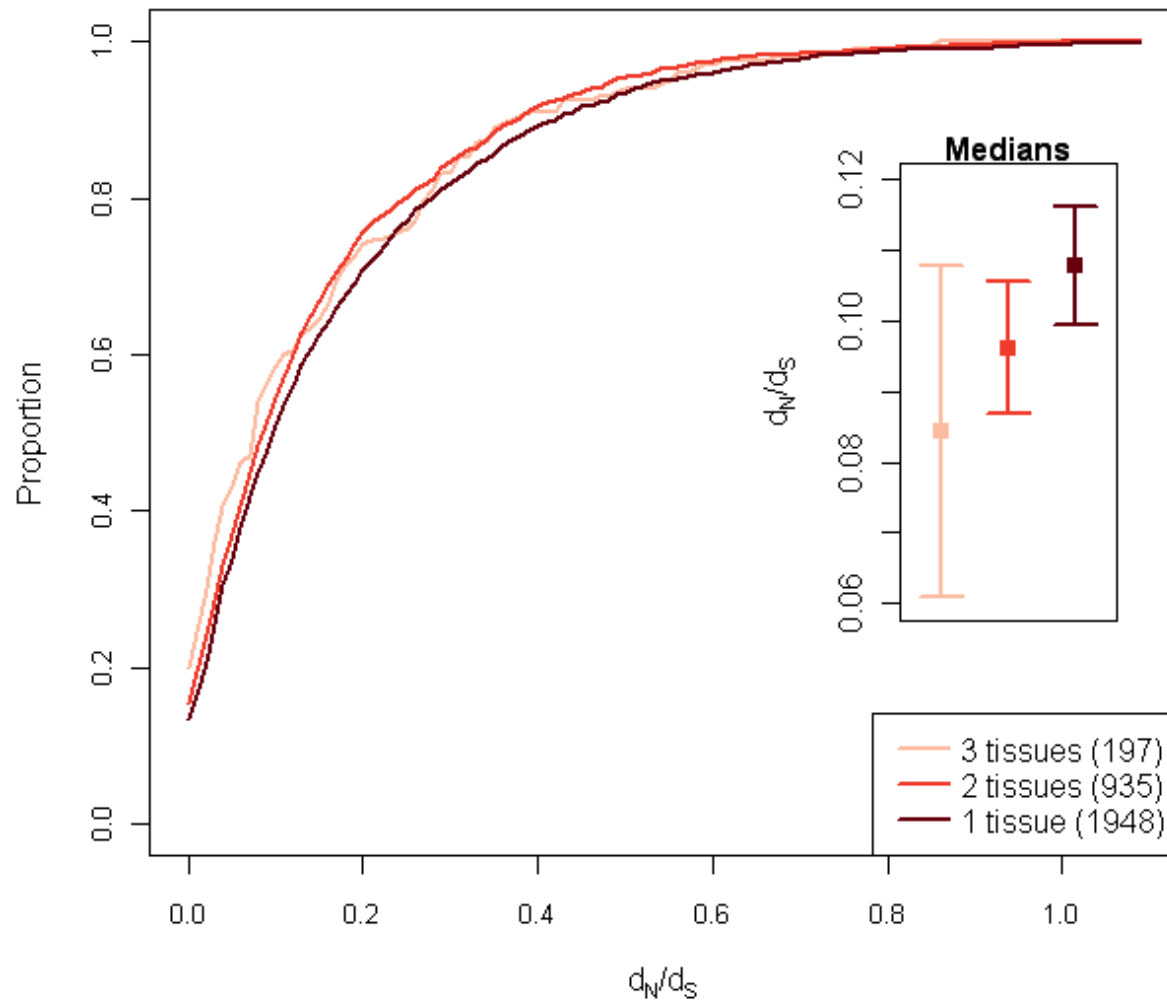
These observations are consistent with the hypothesis that adaptive circumscribed changes in gene regulation have fewer deleterious pleiotropic effects compared with changes at the protein sequence level.

The coding regions of genes whose regulation evolves under natural selection are under more constraint



The coding regions of genes whose regulation evolves under natural selection in larger number of tissues are under more constraint

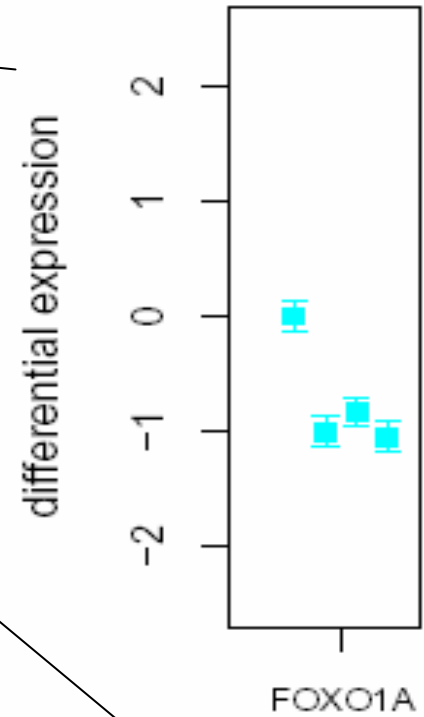
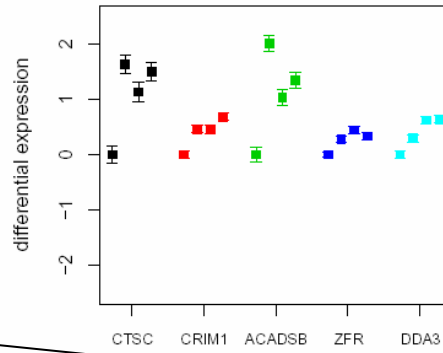
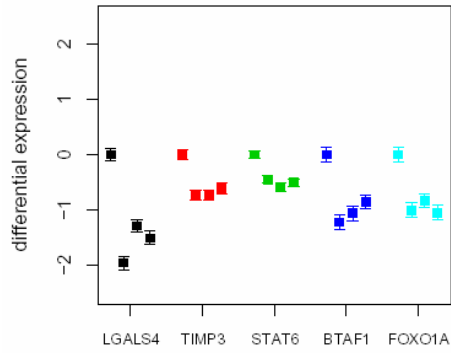
$d_N/d_S$ , stabilizing selection, by number of tissues



# Outline

- Inter-primate comparison of gene expression
- Identifying the direct regulatory targets of transcription factors
- Finding inter-species differences in regulatory pathways

# Genes whose regulation is under directional selection



*FOXO1a* is a target of the IGF/insulin pathway



In *C. elegans*, *Drosophila*, and rodents elevated levels of *DAF-16* increase the average and maximum life span

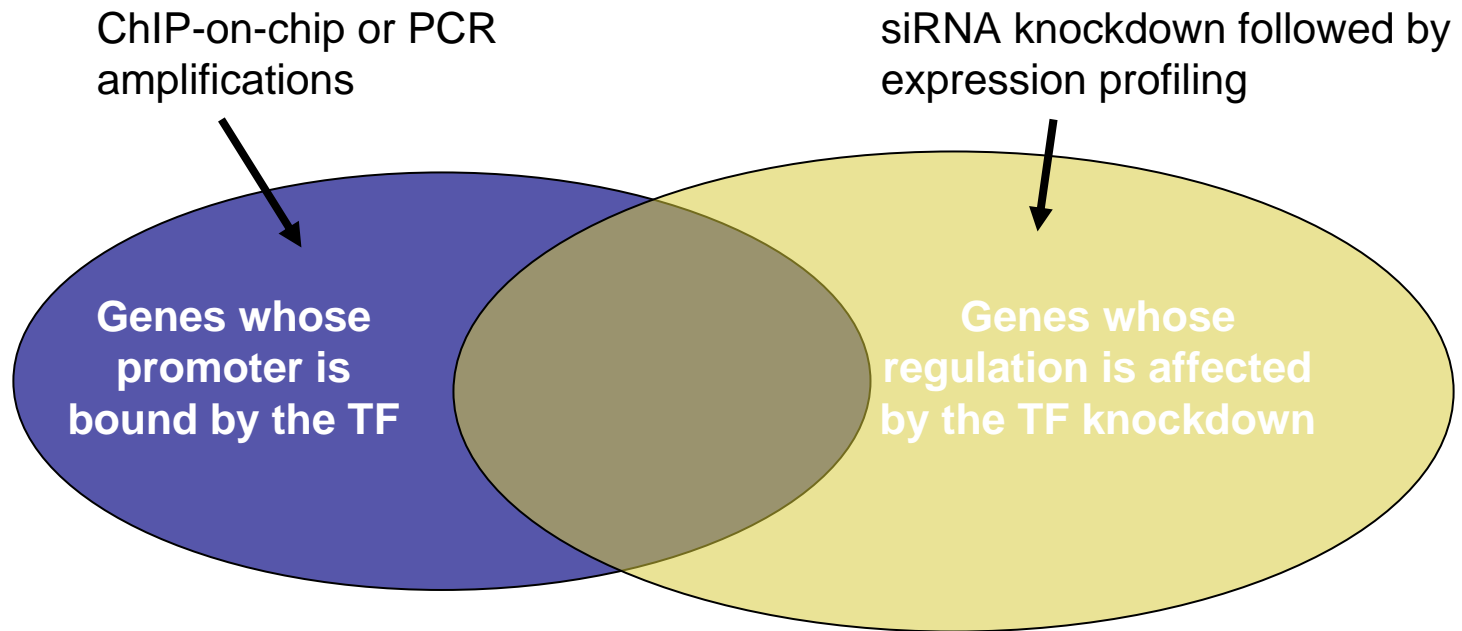
Lin et al., 1997; Giannakou et al., 2004; Furuyama et al., 2002

*FOXO* up-regulation

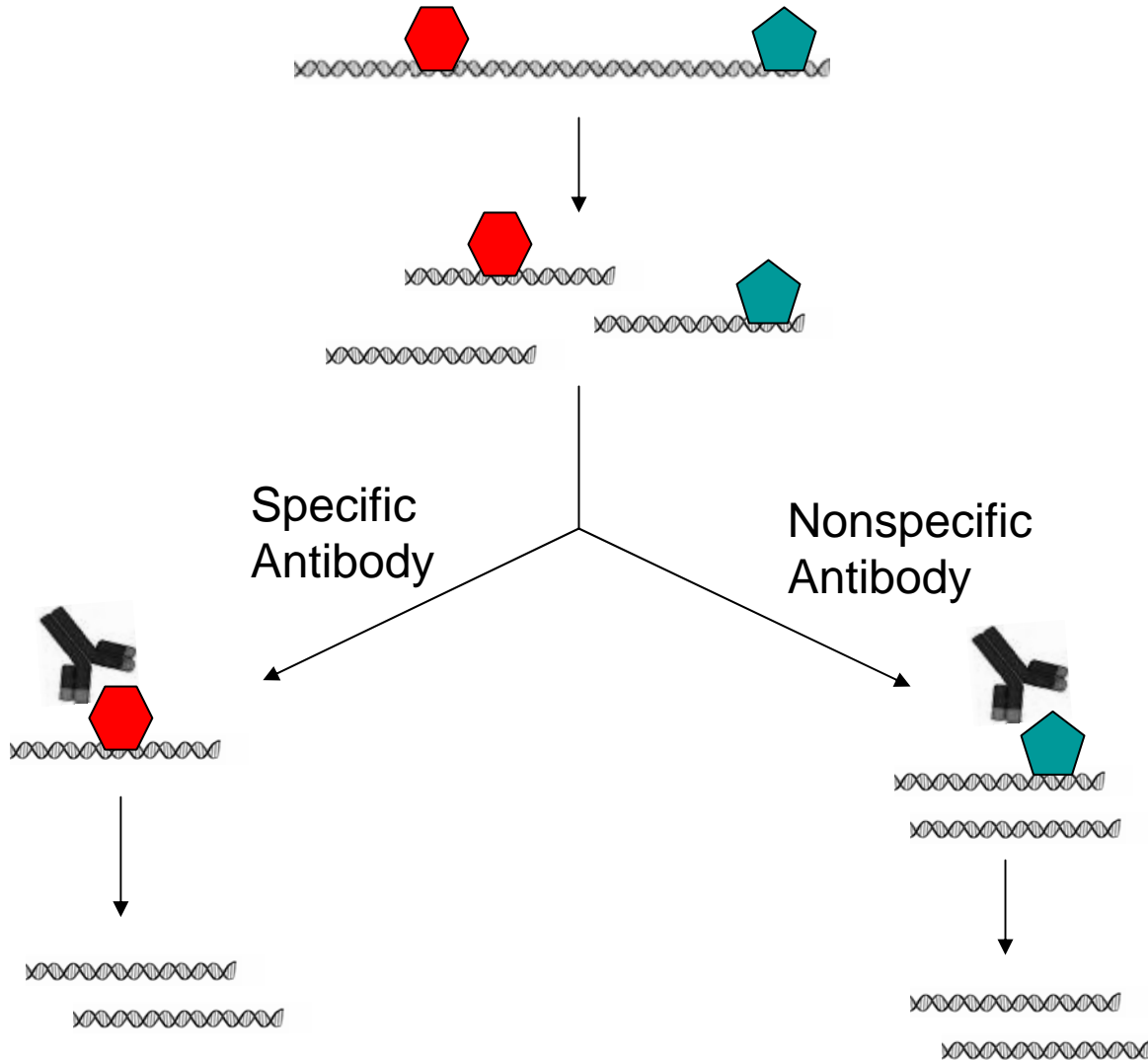


Increased life span

# Searching for target genes of transcription factors that are up-regulated specifically in human



# ChIP-chip



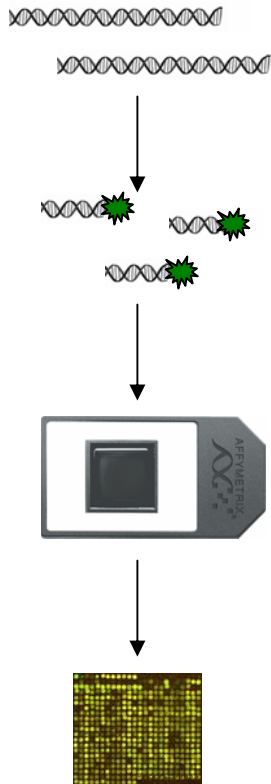
1. Crosslink & Sonicate

2. Immunoprecipitate

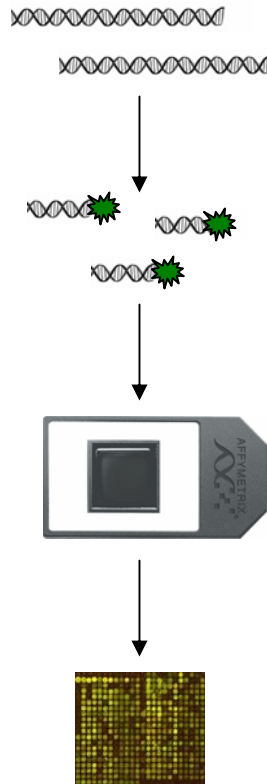
3. Decrosslink

# ChIP-chip

From Specific  
Antibody IP



From Nonspecific  
Antibody IP



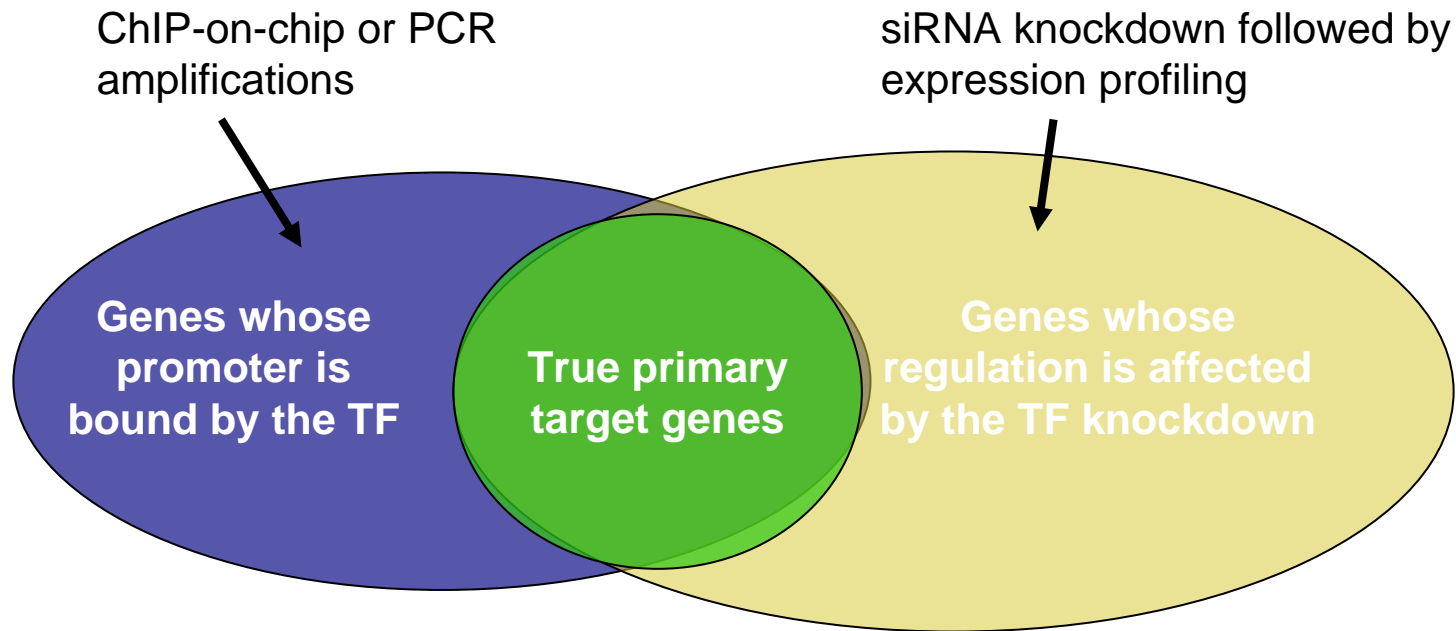
4. Amplify, fragment,  
and label DNA

5. Hybridize to Array  
(Affymetrix human promoter  
tiling array)

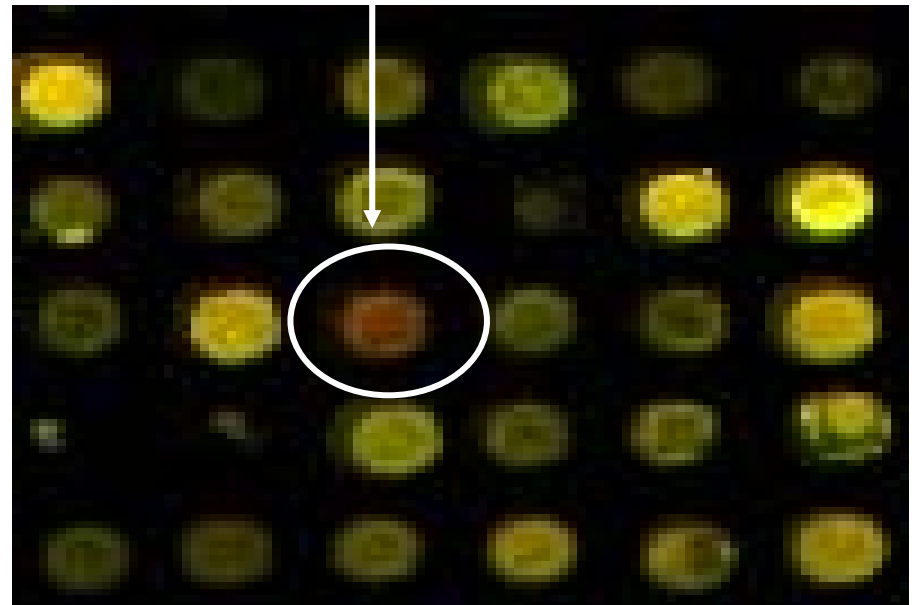
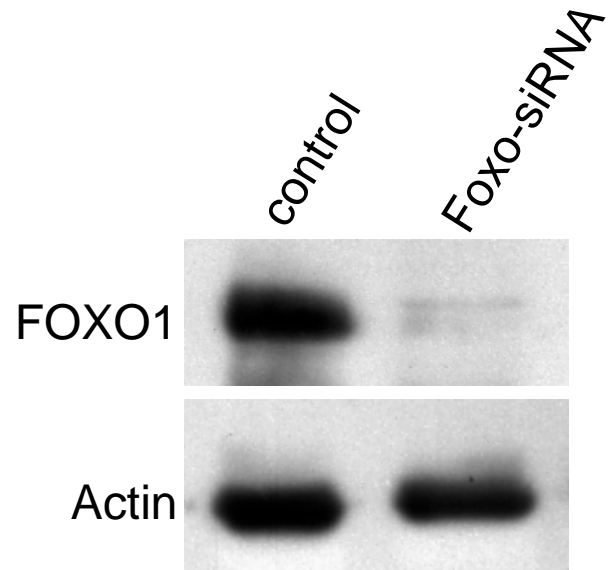
6. Scan

7. Normalize, Analyze

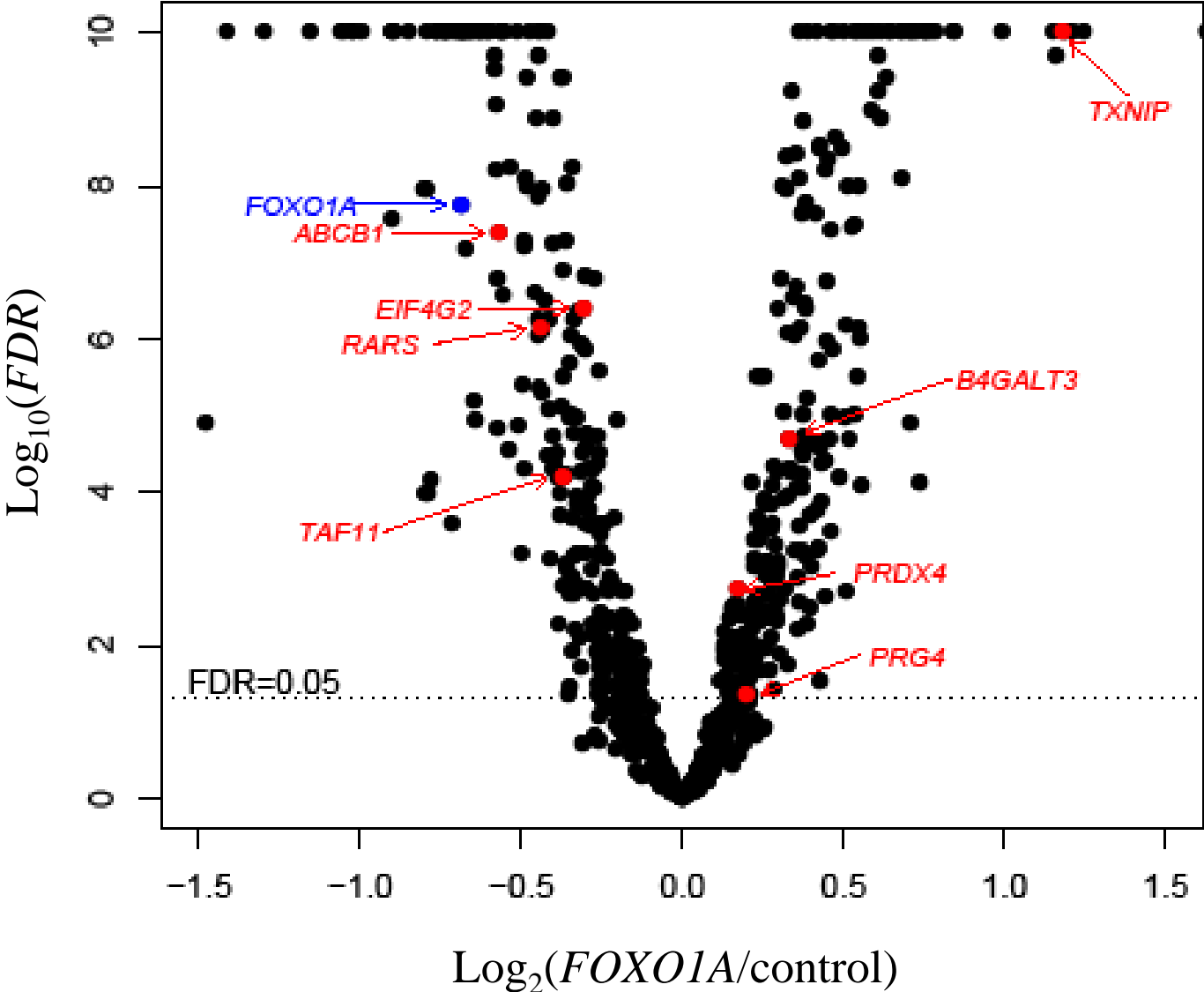
# Searching for target genes of transcription factors that are up-regulated specifically in human



# siRNA protein knock-down in liver c3a cells



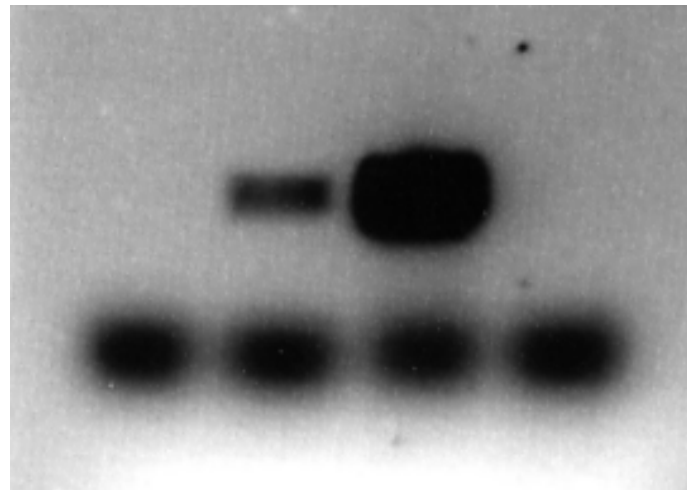
# Gene expression changes following *FOXO1A* knockdown



# Direct transcriptional targets of *FOXO1A*

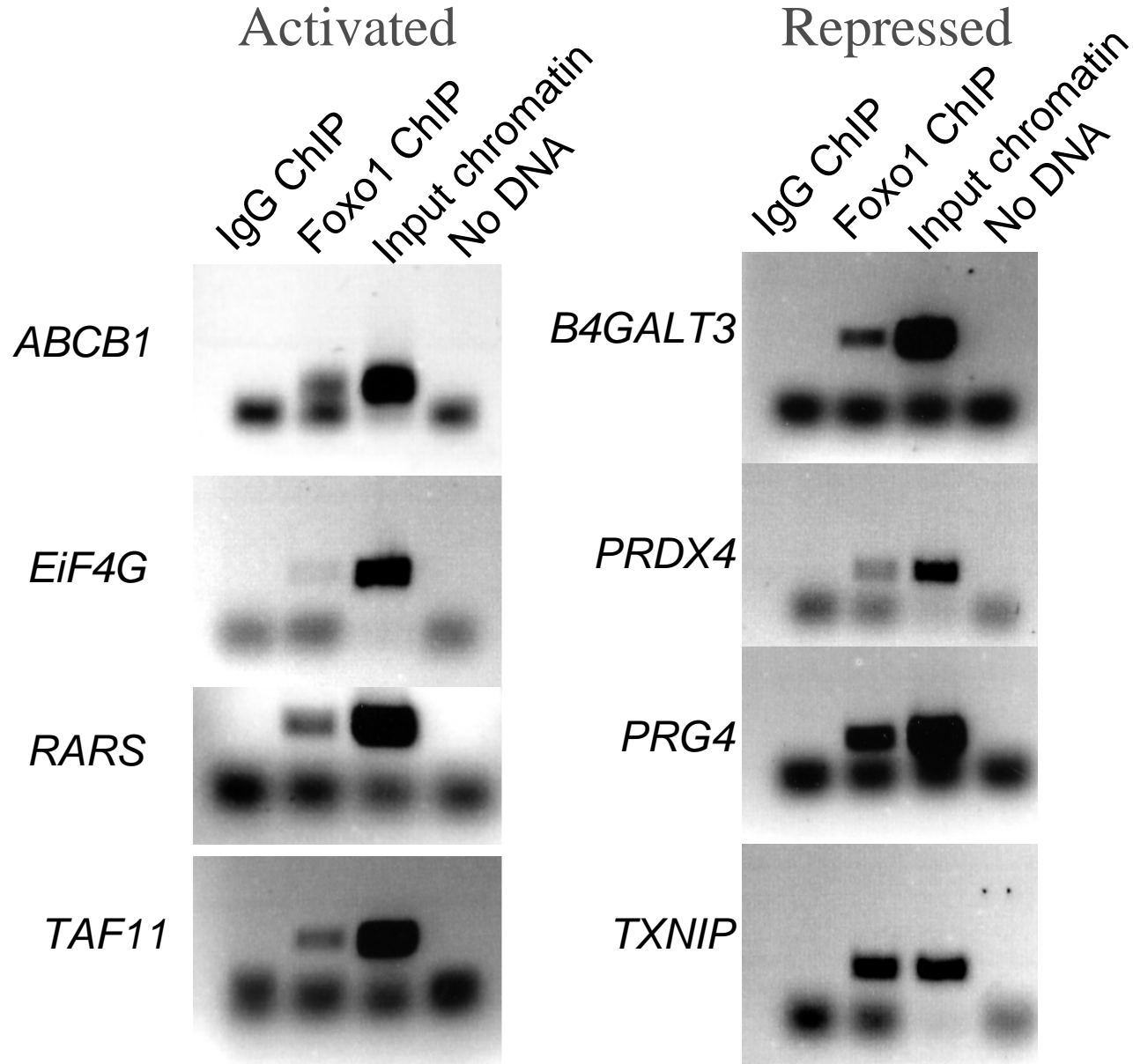
IgG ChIP      Foxo1 ChIP      Input chromatin  
No DNA

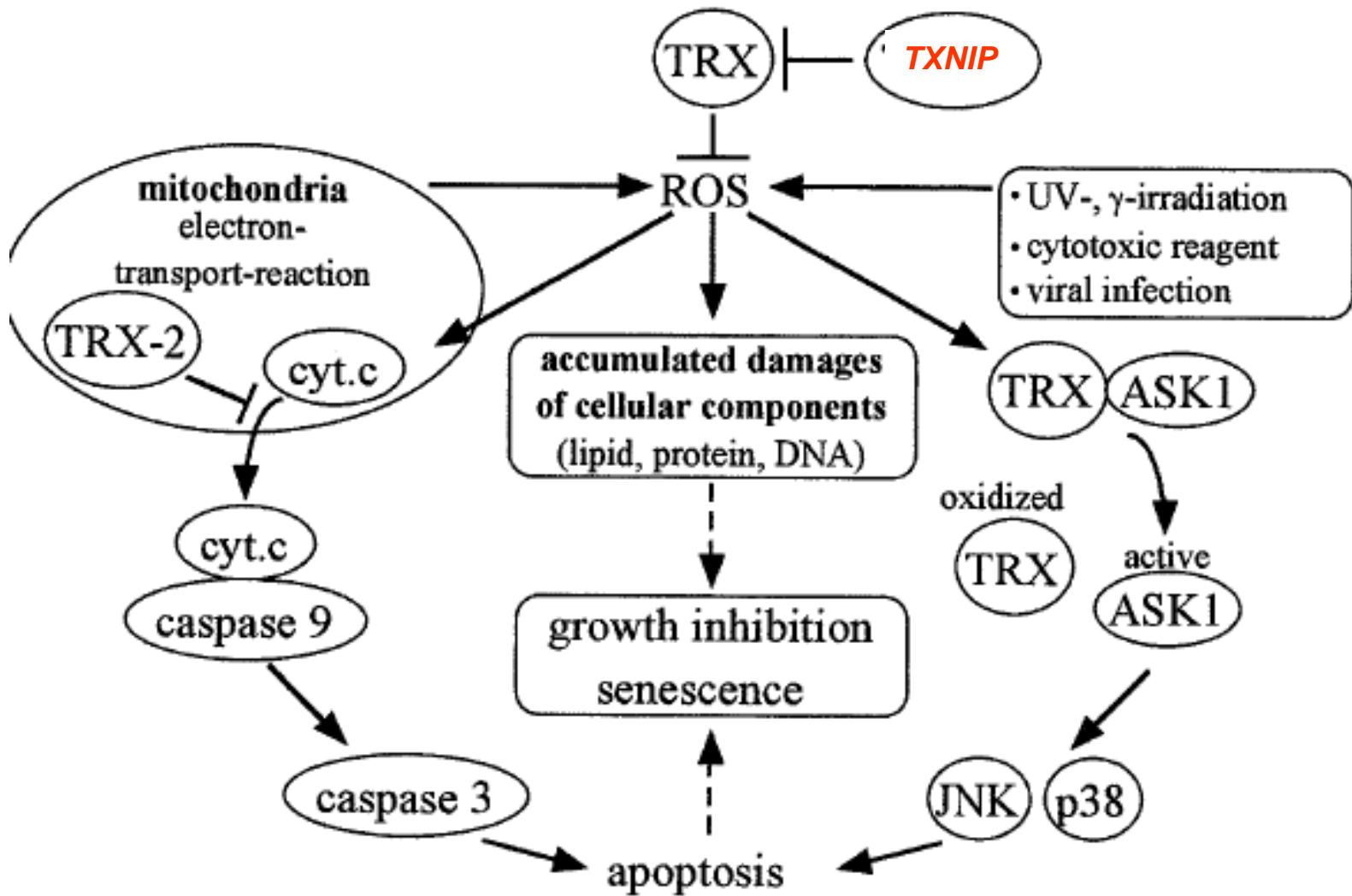
*B4GALT3*





# Direct transcriptional targets of *FOXO1A*





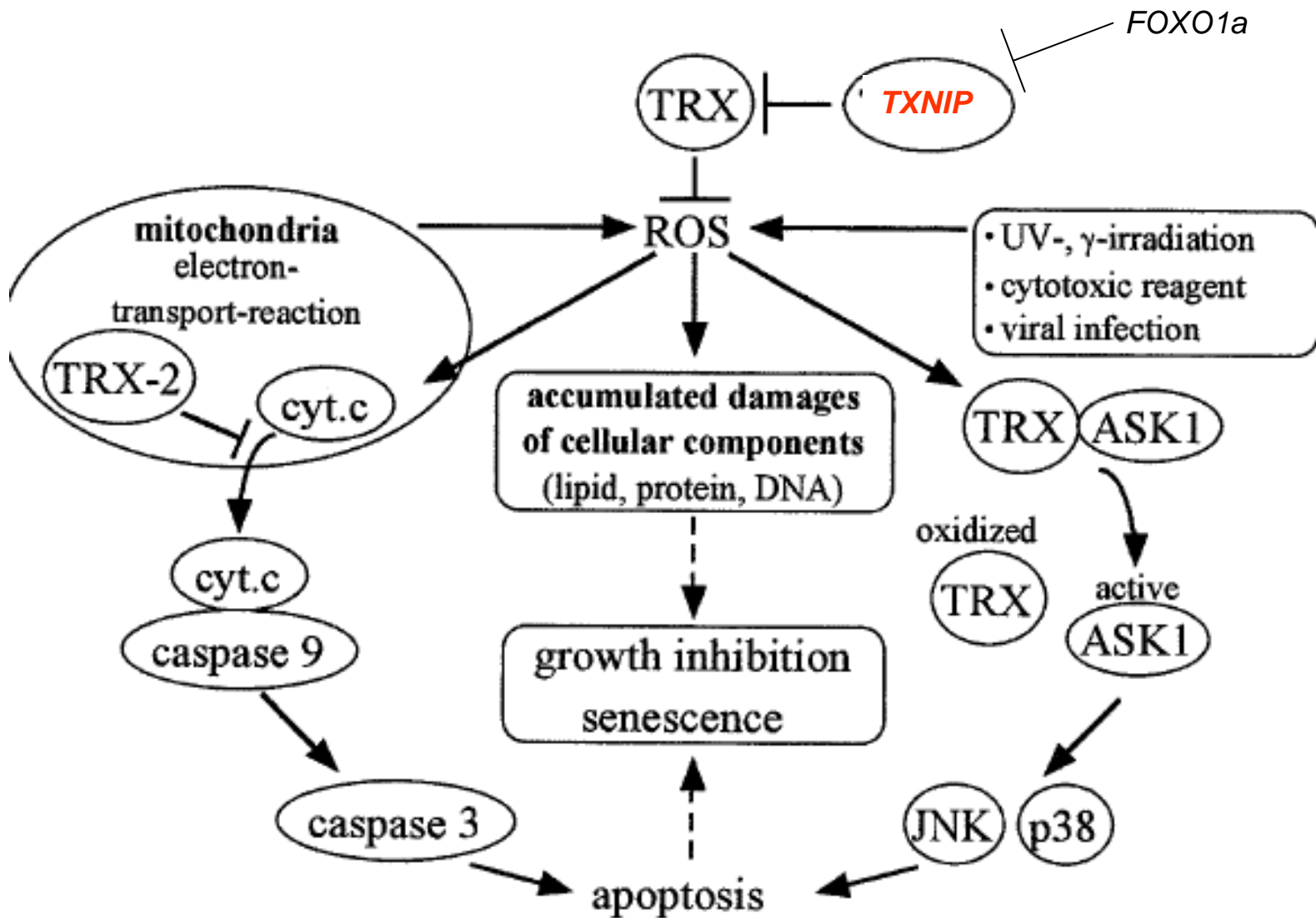
## Reactive oxygen species (ROS) detoxification and the rate of aging:

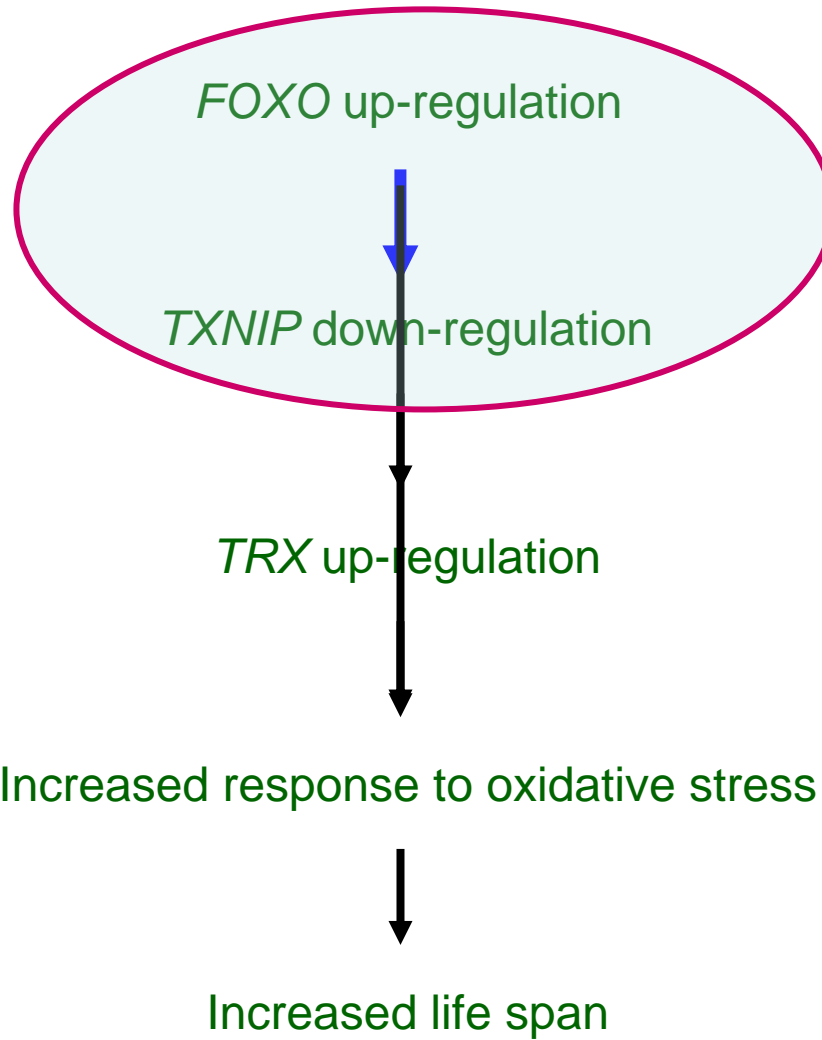
Kregel KC, Zhang HJ (2007) An integrated view of oxidative stress in aging: basic mechanisms, functional effects, and pathological considerations. *Am J Physiol Regul Integr Comp Physiol* 292(1): R18-36.

Jee C, Vanoaica L, Lee J, Park BJ, Ahnn J (2005) Thioredoxin is related to life span regulation and oxidative stress response in *Caenorhabditis elegans*. *Genes Cells* 10(12): 1203-1210.

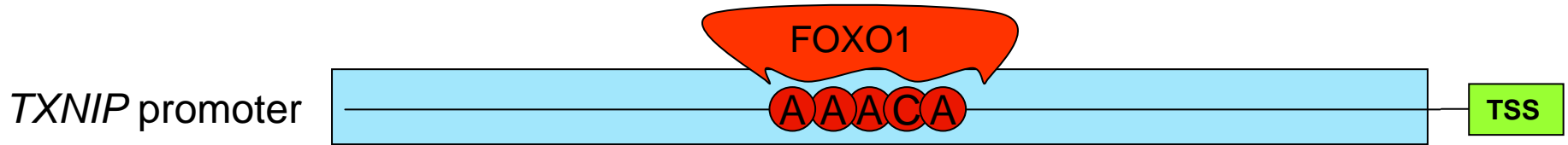
Mitsui A, Hamuro J, Nakamura H, Kondo N, Hirabayashi Y et al. (2002) Overexpression of human thioredoxin in transgenic mice controls oxidative stress and life span. *Antioxid Redox Signal* 4(4): 693-696.

Yoshida T, Oka S, Masutani H, Nakamura H, Yodoi J (2003) The role of thioredoxin in the aging process: involvement of oxidative stress. *Antioxid Redox Signal* 5(5): 563-570.

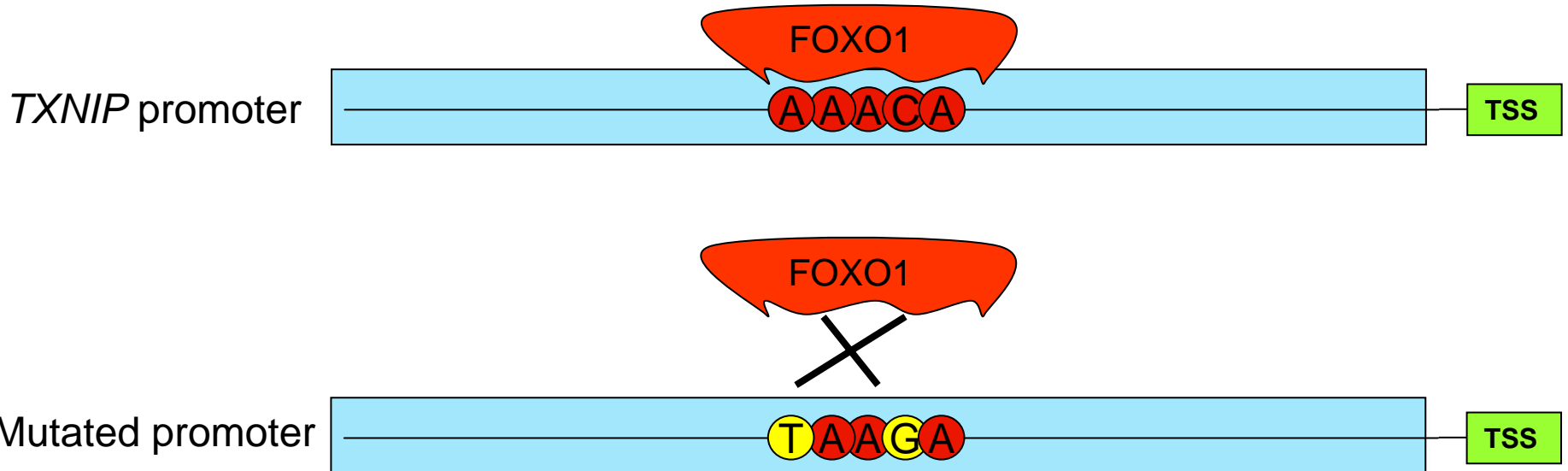




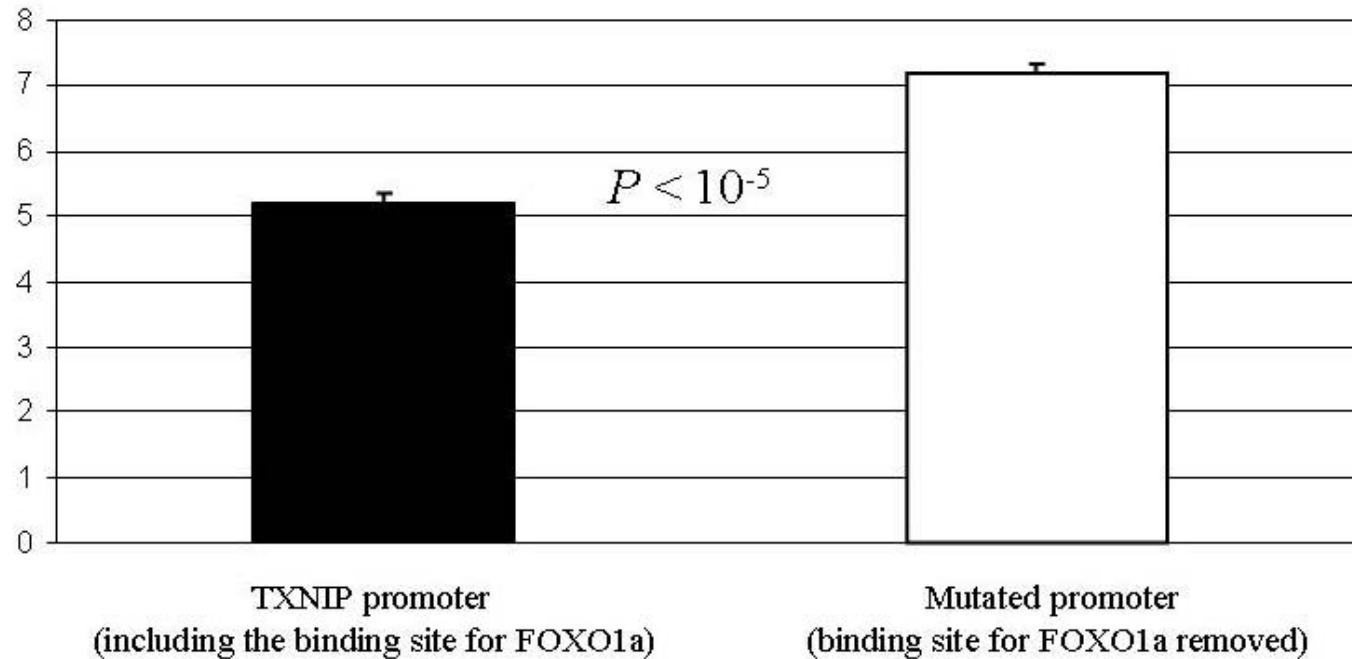
# Reporter Gene Assays



# Reporter Gene Assays



# FOXO1a directly regulates TXNIP

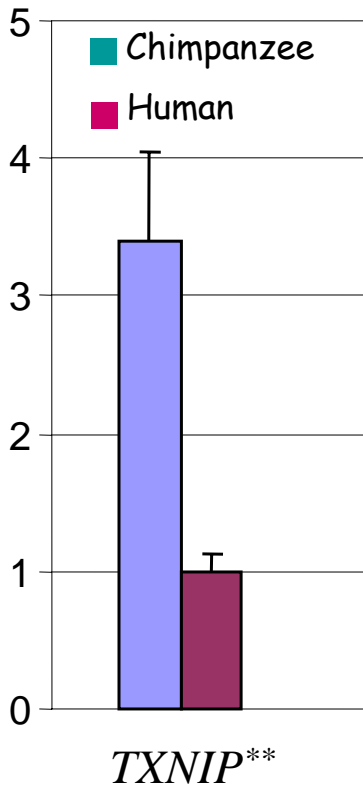


Reporter Gene Assays



## Inter-species differences in *FOXO1A* expression levels affect the direct transcriptional targets

- *FOXO1a* is a direct repressor of TXNIP
- *FOXO1a* expression is higher in humans compared to chimpanzee
- Does *TXNIP* expression level reduced in humans?



## Human

*FOXO* elevated expression



*TXNIP* reduced expression



*TRX* up-regulation



Increased response to oxidative stress

Cutler, R. G (2005)



Increased life span

Hill et al., (2001)

## Chimpanzee

*FOXO* reduced expression



*TXNIP* elevated expression



*TRX* down-regulation



Decreased response to oxidative stress



Reduced life span

## Current work: Functional study

