

# Accessing FANTOM5 expression data



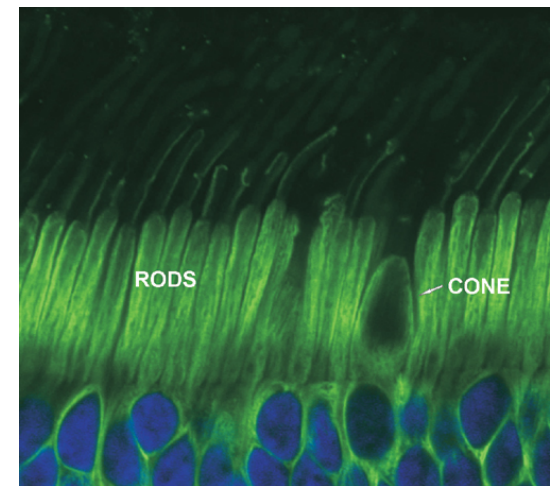
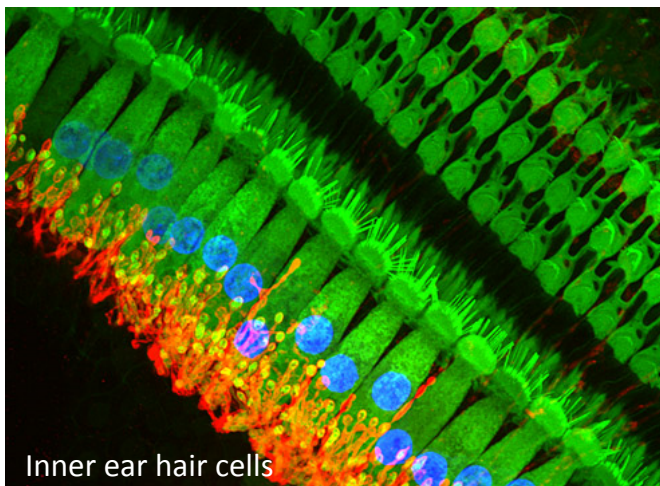
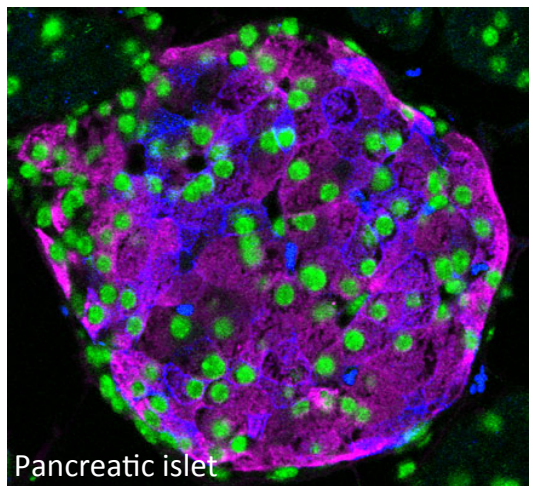
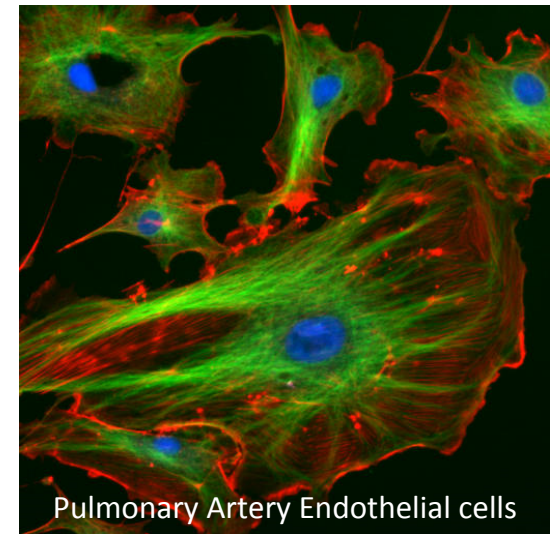
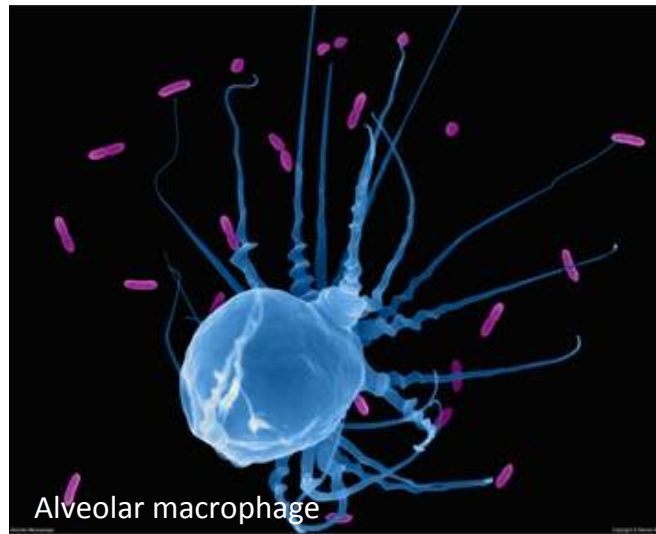
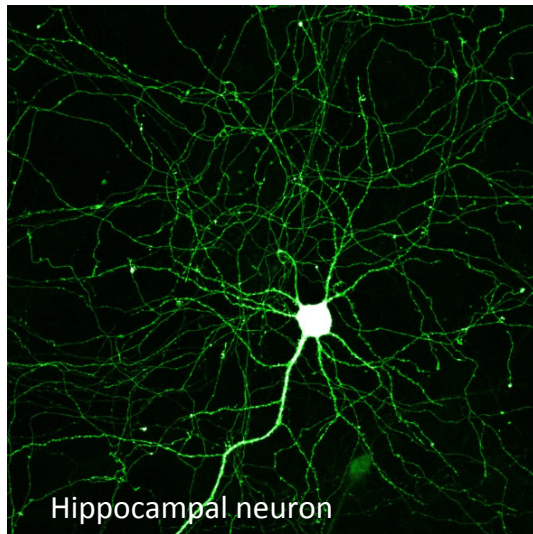
**Robin Andersson**  
robin@binf.ku.dk

The Bioinformatics Centre,  
Dept. of Biology & BRIC,  
University of Copenhagen



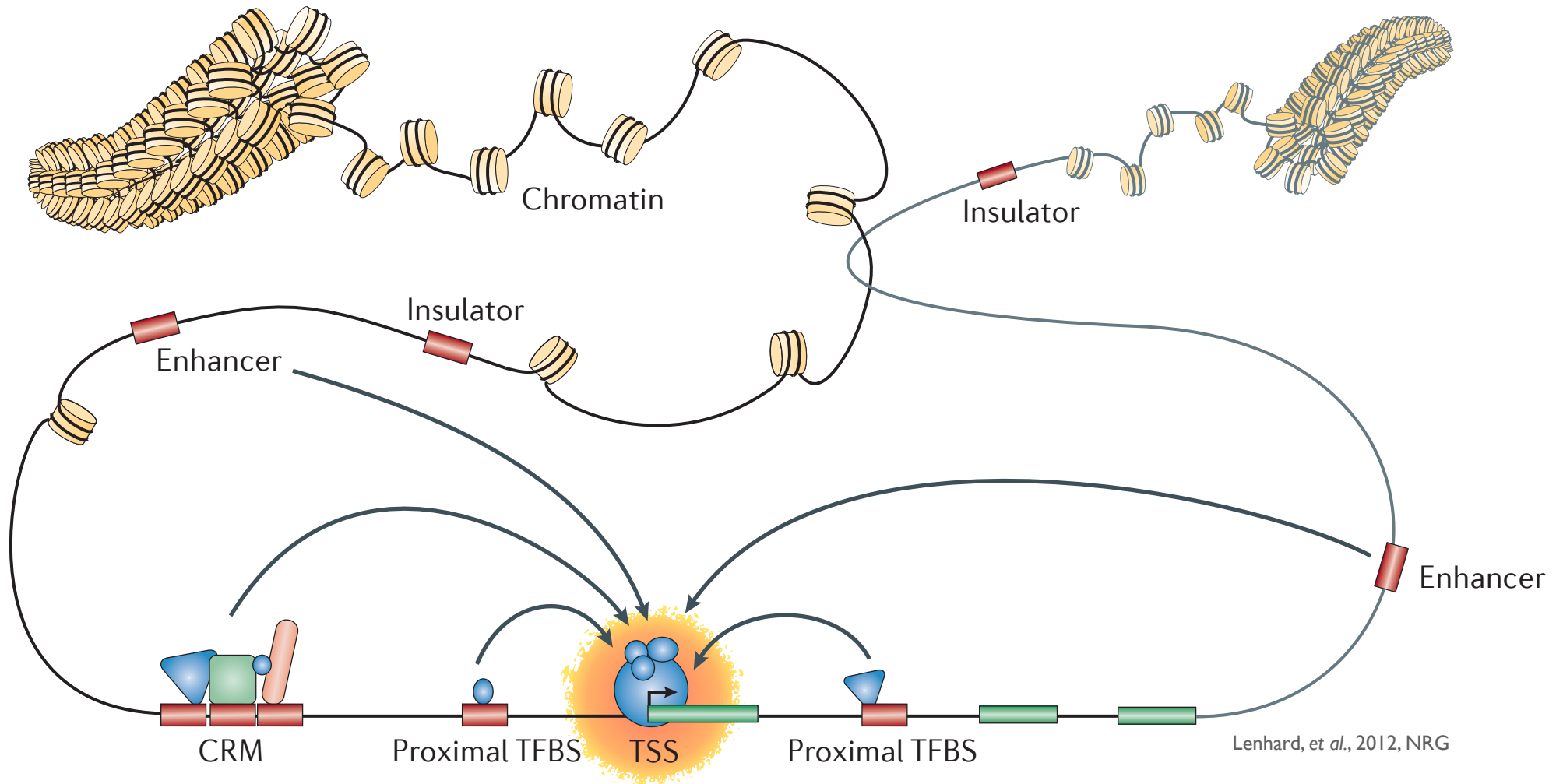
ELIXIR Innovation & SME Forum, Nov 24, 2014





- ~400 different cell types with varying function and morphology
- How are these differences encoded?
- What defines a certain cellular state?

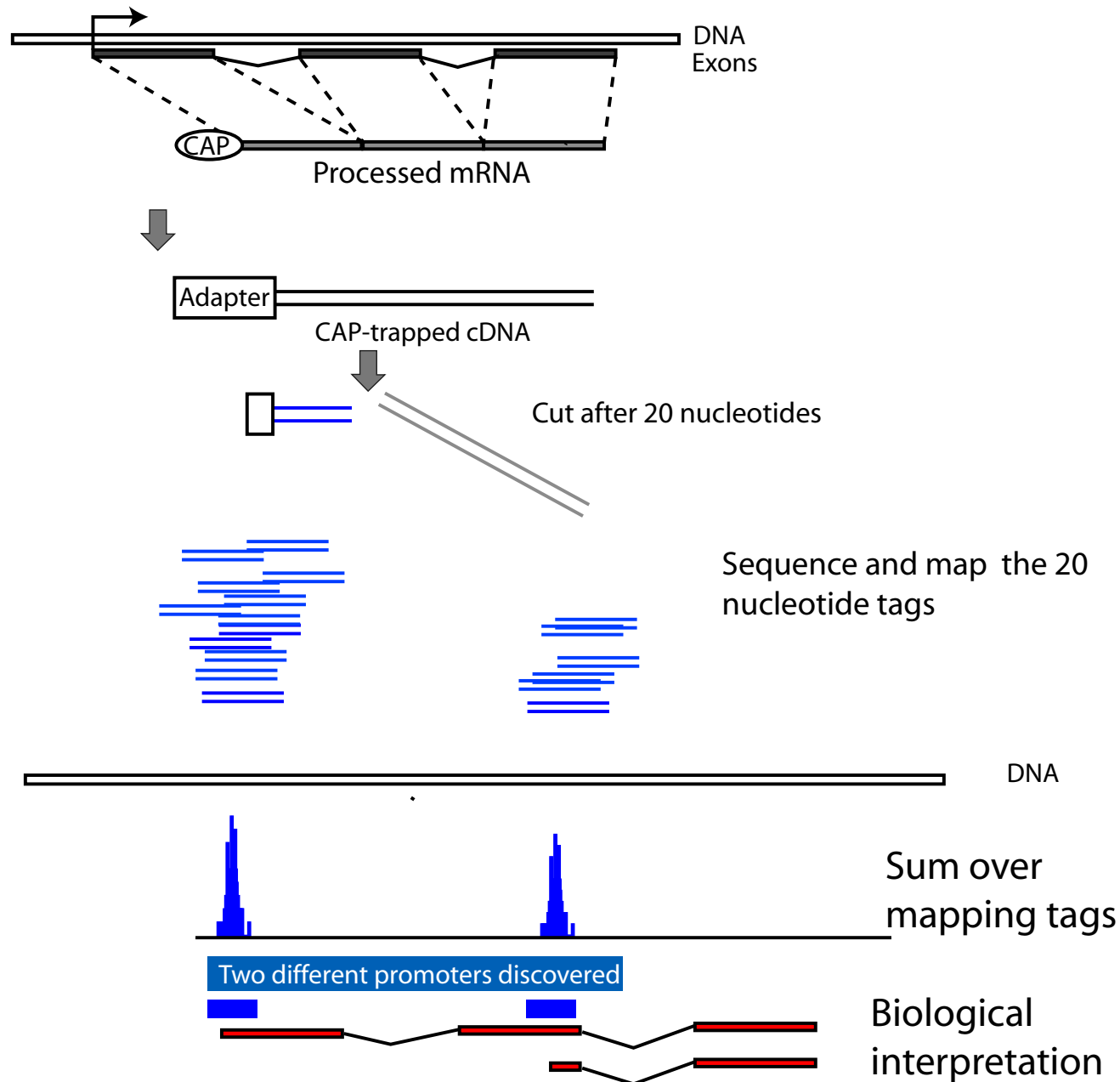




Lenhard, et al., 2012, NRG



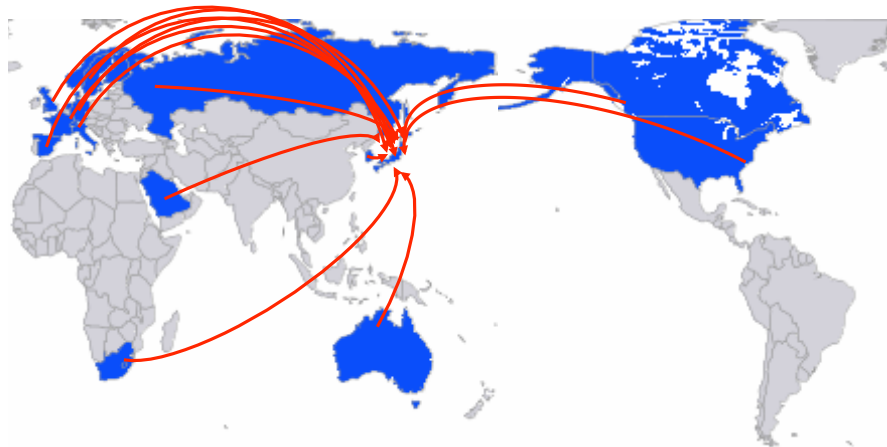
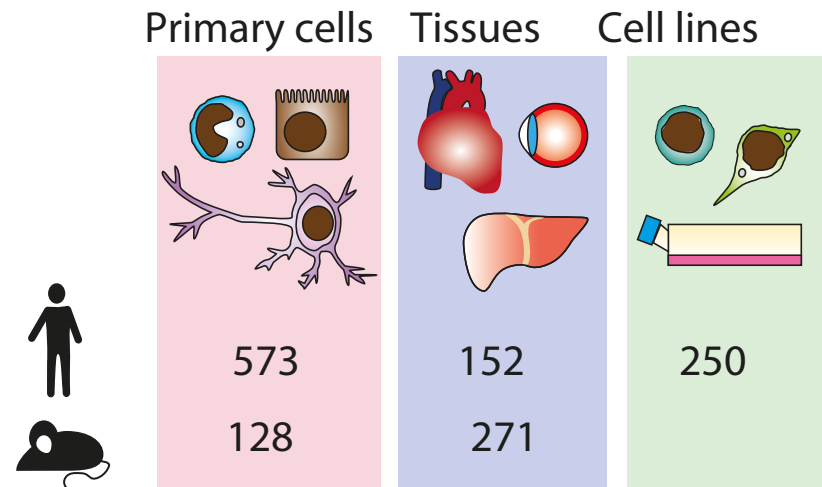
# Cap Analysis of Gene Expression (CAGE)





# FANTOM5

- Aim: Map the vast majority of transcription initiation sites in human and mouse via single molecule CAGE



Over 120 international collaborators: Sample providers, Technology developers and Bioinformaticians from 19 countries

[fantom.gsc.riken.jp/5](http://fantom.gsc.riken.jp/5)

Contact:

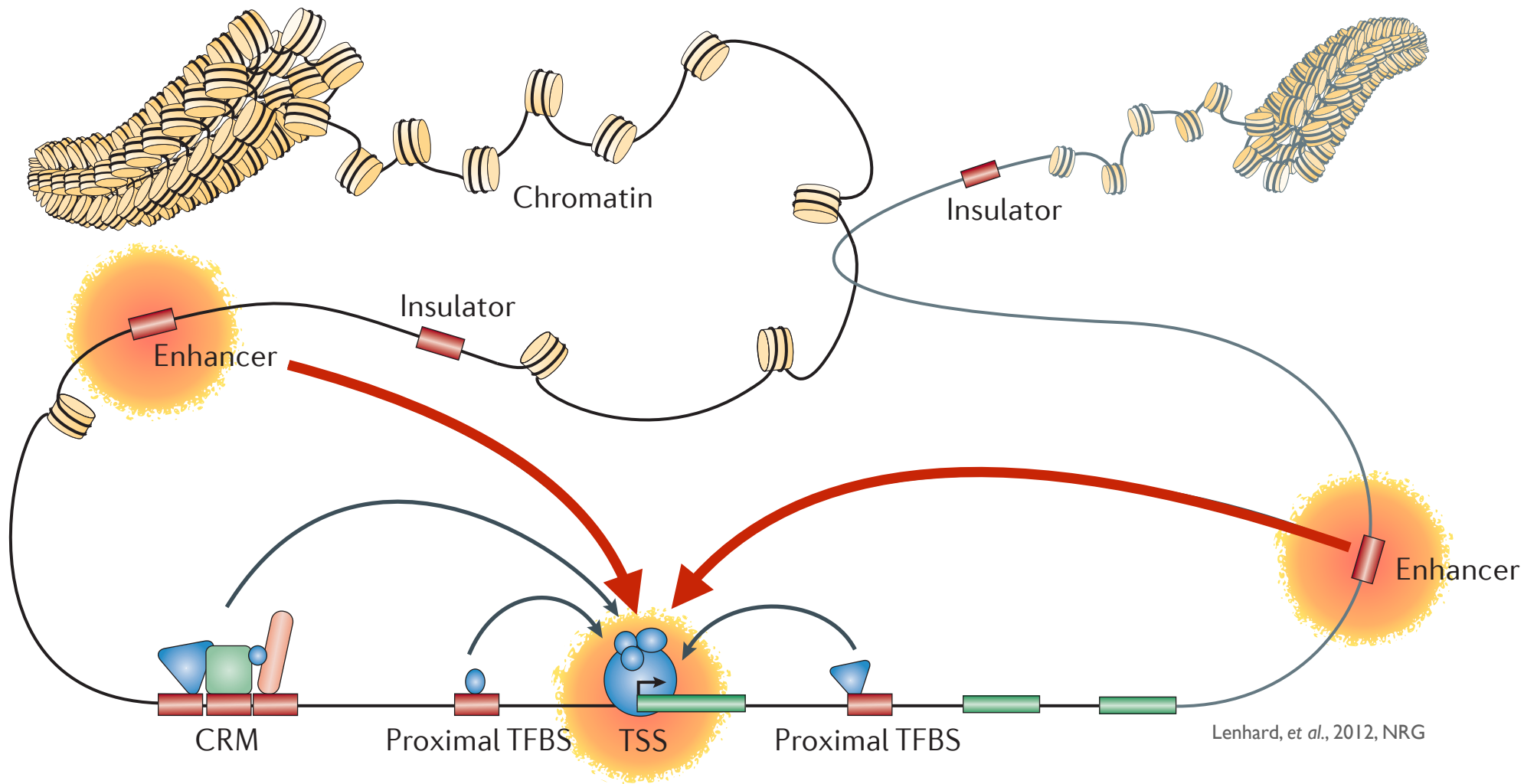
Yoshihide Hayashizaki

Piero Carninci

Alistair Forrest

Division of Genomic Technologies,  
RIKEN CLST,  
Yokohama, Japan



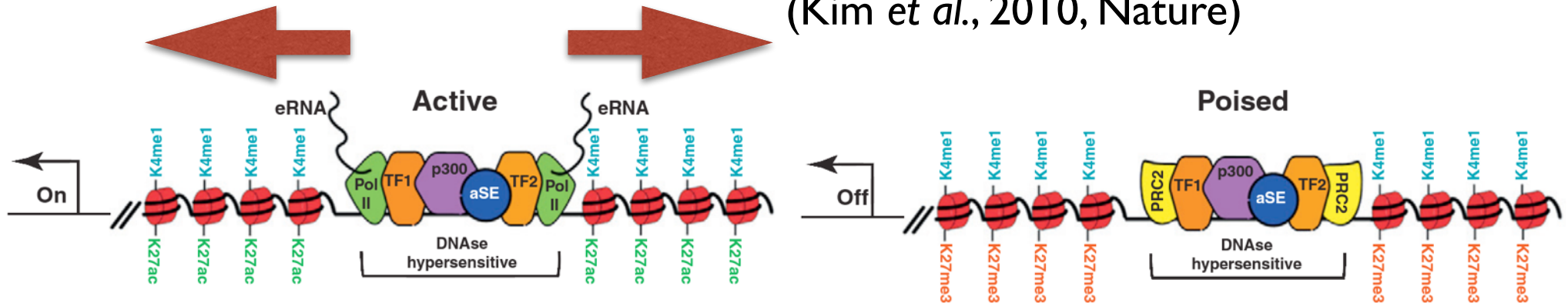


Lenhard, et al., 2012, NRG



# Localization of enhancers

Active enhancers are transcribed  
(Kim *et al.*, 2010, Nature)



Buecker and Wysocka, 2012, Cell

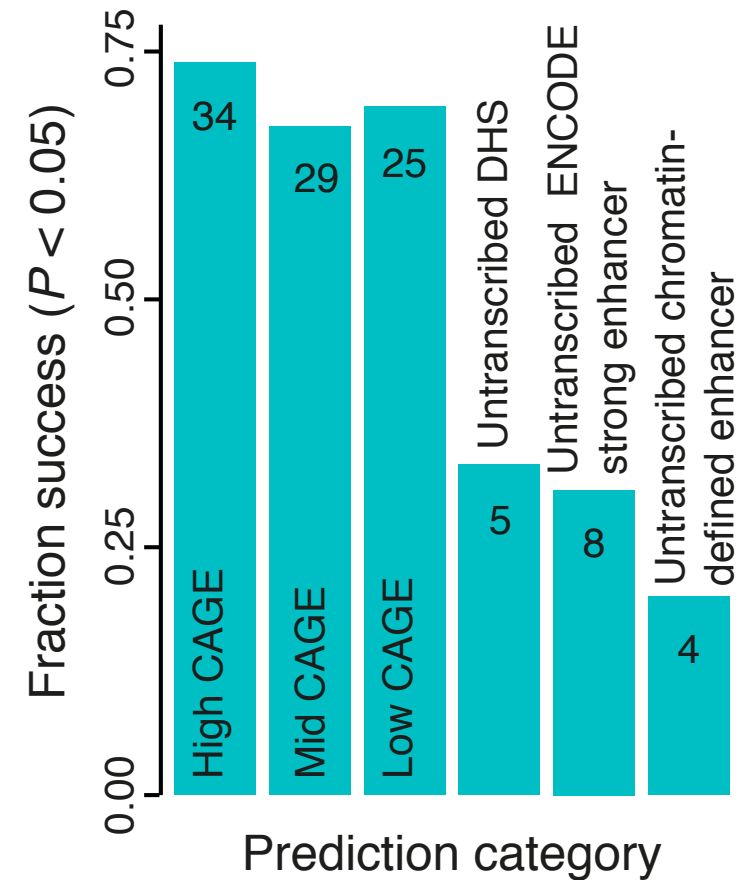
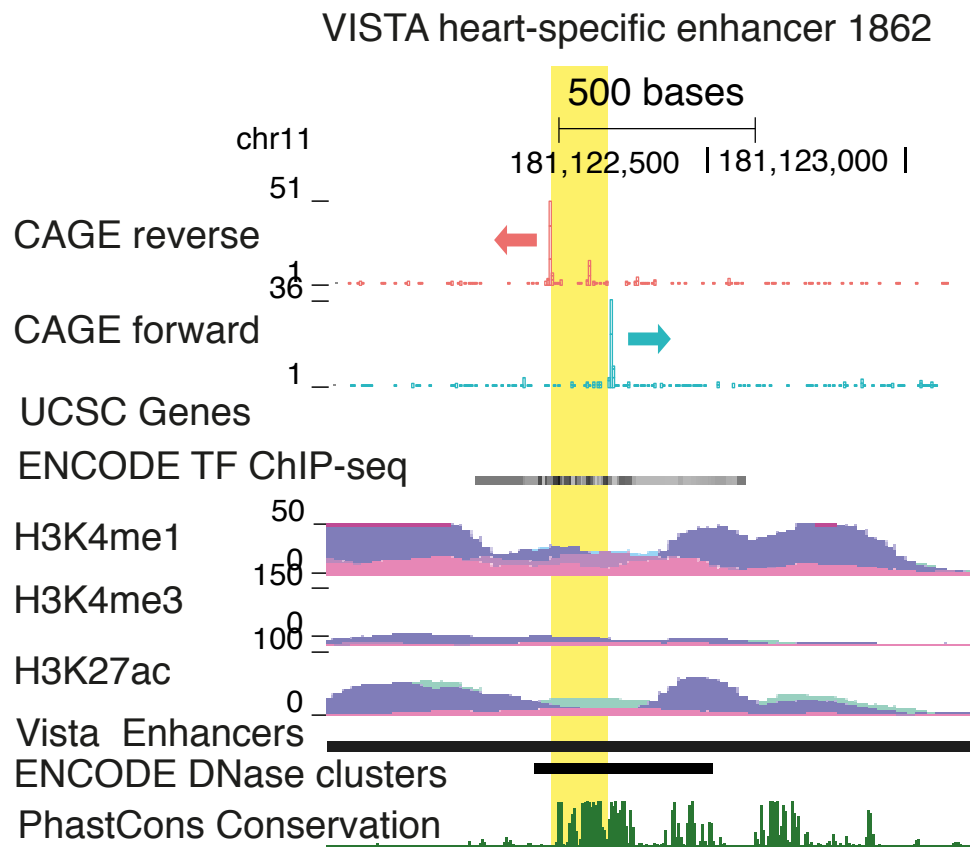
- Open chromatin (DNase hypersensitivity)
- TF binding (master regulators, co-activators)
- Chromatin context (H3K4me1, H3K27ac)
- Chromatin interactions

e.g. ENCODE





# Atlas of 44,000 human enhancers



- Divergent capped RNAs identifies active enhancers
- Transcribed enhancers are more likely to be active



ARTICLE

fantom.gsc.riken.jp/5/

doi:10.1038/nature13182

# A promoter-level mammalian expression atlas

The FANTOM Consortium and the RIKEN PMI and CLST (DGT)\*

ARTICLE

enhancer.binf.ku.dk

doi:10.1038/nature12787

# An atlas of active enhancers across human cell types and tissues

Robin Andersson<sup>1\*</sup>, Claudia Gebhard<sup>2,3\*</sup>, Irene Miguel-Escalada<sup>4</sup>, Ilka Hoof<sup>1</sup>, Jette Bornholdt<sup>1</sup>, Mette Boyd<sup>1</sup>, Yun Chen<sup>1</sup>, Xiaobei Zhao<sup>1,5</sup>, Christian Schmidl<sup>2</sup>, Takahiro Suzuki<sup>6,7</sup>, Evgenia Ntini<sup>8</sup>, Erik Arner<sup>6,7</sup>, Eivind Valen<sup>1,9</sup>, Kang Li<sup>1</sup>, Lucia Schwarzfischer<sup>2</sup>, Dagmar Glatz<sup>2</sup>, Johanna Raithel<sup>2</sup>, Berit Lilje<sup>1</sup>, Nicolas Rapin<sup>1,10</sup>, Frederik Otzen Bagger<sup>1,10</sup>, Mette Jørgensen<sup>1</sup>, Peter Refsing Andersen<sup>8</sup>, Nicolas Bertin<sup>6,7</sup>, Owen Rackham<sup>6,7</sup>, A. Maxwell Burroughs<sup>6,7</sup>, J. Kenneth Baillie<sup>11</sup>, Yuri Ishizu<sup>6,7</sup>, Yuri Shimizu<sup>6,7</sup>, Erina Furuhashi<sup>6,7</sup>, Shiori Maeda<sup>6,7</sup>, Yutaka Negishi<sup>6,7</sup>, Christopher J. Mungall<sup>12</sup>, Terrence F. Meehan<sup>13</sup>, Timo Lassmann<sup>6,7</sup>, Masayoshi Itoh<sup>6,7,14</sup>, Hideya Kawaji<sup>6,14</sup>, Naoto Kondo<sup>6,14</sup>, Jun Kawai<sup>6,14</sup>, Andreas Lennartsson<sup>15</sup>, Carsten O. Daub<sup>6,7,15</sup>, Peter Heutink<sup>16</sup>, David A. Hume<sup>11</sup>, Torben Heick Jensen<sup>8</sup>, Harukazu Suzuki<sup>6,7</sup>, Yoshihide Hayashizaki<sup>6,14</sup>, Ferenc Müller<sup>4</sup>, The FANTOM Consortium†, Alistair R. R. Forrest<sup>6,7</sup>, Piero Carninci<sup>6,7</sup>, Michael Rehli<sup>2,3</sup> & Albin Sandelin<sup>1</sup>

# Questions asked

- In which cell type(s) is a certain gene expressed?
- Which promoters/enhancers are “specific” to a certain cell type?
- How do promoter/enhancer usage differ across cell types?
- Where do promoters/enhancers to a certain gene locate in the genome?
- In which cell type(s) might a certain genetic variant have an effect?
- Which enhancer(s) regulate a certain gene?

**Problem:** ~1,000 CAGE libraries, ~200,000 promoters and ~44,000 enhancers - **how to best visualize and distribute data?**





# FANTOM5 SSTAR

[fantom.gsc.riken.jp/5/sstar/](http://fantom.gsc.riken.jp/5/sstar/)

- SSTAR: Semantic catalogue of Samples, Transcription initiation And Regulators
- Sample info and ontology (UBERON, Disease, Cell) term association
- CAGE promoters - expression ranks, gene association, ontology enrichment
- Transcription factor motif enrichment
- ...



# ZENBU Genome Browser



Developed at RIKEN for the FANTOM 5 project.

Next-generation Sequencing (NGS) focused genome browser.

- Work directly with sequence alignment files (BAM)
- Powerful embedded data processing/analysis
- Powerful data upload, secured sharing and collaborations
- Big-data: ability to work with 1000s of experiments

<http://fantom.gsc.riken.jp/zenbu>

# ZENBU genome browser



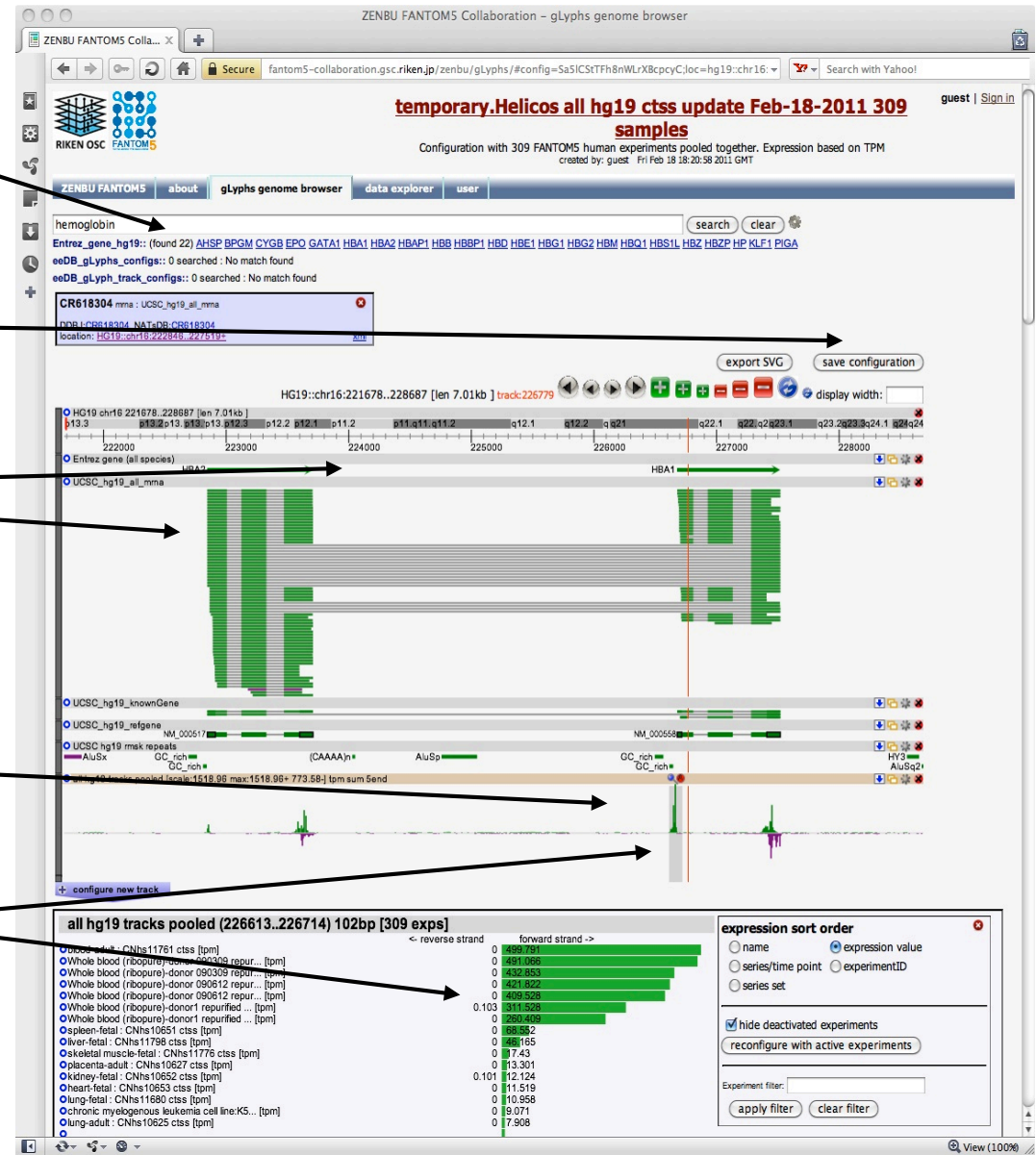
Search for genes, views, tracks with keywords or gene names

Save and share view configurations with collaborators

Highly configurable views: track rearrangement, user selectable drawing styles, scaling...

Pooled expression data: one track dynamically mixing expression from many experiments

Interactive data exploration: selection of regions in pooled expression is reflected in experiment-expression view. Allows for easy checking of expression differences of CAGE peaks in a region.



[fantom.gsc.riken.jp/zenbu](http://fantom.gsc.riken.jp/zenbu)



# Enhancer/promoter slider tools

- How do best define cell type-specificity?
- Let the end user define the rules!

enhancer.binf.ku.dk

Welcome [Search for Enhancers](#) [Motif Over-representation](#) [Pre-defined Tracks](#) [View All Enhancers](#) **FANTOM** FUNCTIONAL ANNOTATION OF TRANSCRIPTOMES OF THE HUMAN GENOME **The Bioinformatics Centre**

### FANTOM5 Enhancer Selector: Search for enhancers by expression or location [Show Details](#)

**Search** Set the searching arguments of location and cell/organ percentage information

**Location** Specify the location range [How?](#)

| Chromosome | Starting site | Ending site | Gene | Strand | Up and Down Stream | Sort by     |
|------------|---------------|-------------|------|--------|--------------------|-------------|
| ALL        | 1             | 249250621   |      | NA     | Up Down            | Do not sort |

**Primary cell expression** [How?](#) Remaining: 19.20 [Reset](#)

**Organ/tissue expression** [How?](#) Remaining: 100.00 [Reset](#)

mesothelial cell 0.00%  
monocyte 0.00%  
myoblast 0.00%  
natural killer cell 0.00%  
neuron 0.00%  
neuronal stem cell 0.00%  
neutrophil 0.00%  
osteoblast 0.00%

adipose tissue 0.00%  
blood 0.00%  
0.00%  
0.00%  
0.00%  
0.00%  
0.00%

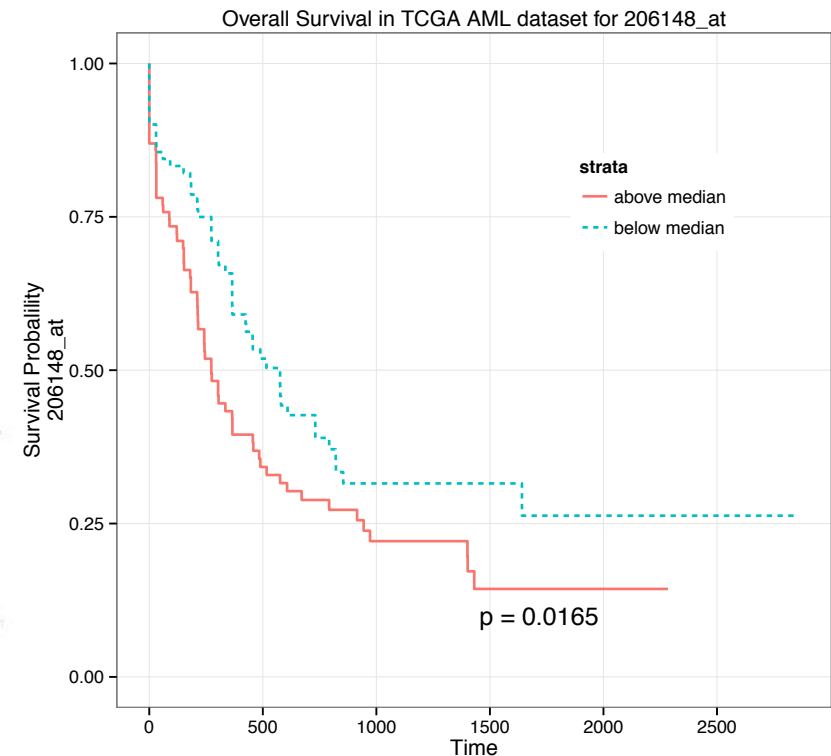
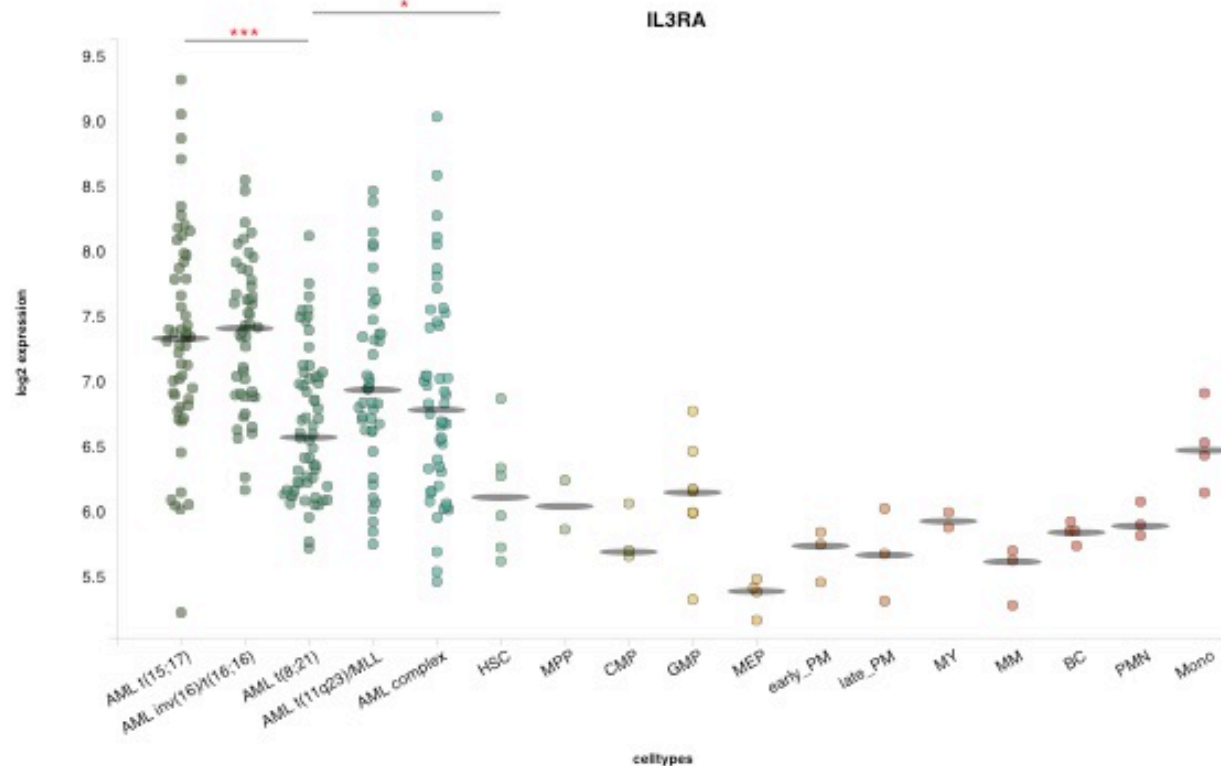
**Results**  
Show results that match [Cell AND Organ Constr](#)  
Number of hits: **36**  
[See Detailed Results >>>](#)

**Example case:**  
choose the set of enhancers whose expression is composed of >40% neurons and >40% neural stem cells





- Single gene lookup based on available microarray data from healthy and leukemic cells of the bone marrow
- Gene expression, visualization, correlation and test for influence on survival in acute myeloid leukemia
- <http://servers.binf.ku.dk/bloodspot>
- Workshop tomorrow tech track 1. N. Rapin & F. O. Bagger





# Acknowledgements

## Uni. Copenhagen

Albin Sandelin  
Ilka Hoof  
Eivind Valen  
Yun Chen  
Xiaobei Zhao  
Berit Lilje  
Kang Li  
Jette Bornholdt  
Mette Boyd  
Mette Jørgensen  
Morana Vitezic  
Hans Ienasescu  
Frederik Otzen Bagger  
Nicolas Rabin

## Uni. Hospital Regensburg

Michael Rehli  
Claudia Gebhard  
Christian Schmidl  
Lucia Schwarzfischer  
Dagmar Glatz  
Johanna Raithel

## Uni. Birmingham

Ferenc Müller  
Irene Miguel-Escalada

## Aarhus University

Torben Heick Jensen  
Peter Refsing Andersen  
Evgenia Ntini

## FANTOM 5 consortium

David Hume  
Kenneth Baillie  
Christopher Mungall  
Terrence Meehan  
Peter Heutink  
Andreas Lennartsson  
Carsten Daub

## RIKEN CLST

Yoshihide Hayashizaki  
Alistair Forrest  
Piero Carninci  
Harakazu Suzuki  
Takahiro Suzuki  
Timo Lassmann  
Masayoshi Itoh  
Hideya Kawaji  
Naoto Kondo  
Jun Kawai



## Funding

- Research Grant for RIKEN Omics Science Center from MEXT to YH.
- Grant of the Innovative Cell Biology by Innovative Technology (Cell Innovation Program) from the MEXT, Japan to YH.

