FANTOM 5

2011.10.17-20
FANTOM 5 KŌYŌ MEETING
FUNCTIONAL ANNOTATION OF THE MAMMALIAN GENOME

RIKEN
Omics Science Center
RIKEN OSC

Al Forrest

1. Confidentiality
2. Reminders from February meeting
    1. Authorship
    2. Design
    3. Schedule
3. Some insights from Phase1 data
4. Goals of the meeting

Reminder from February

Authorship – main papers

Paper leaders are responsible for choosing the team they need to achieve their manuscript.

The decisions on which methods are used for each paper will be the decision of the paper leaders (this is to give them independence and the greatest chance of pushing the paper the way they envisaged).

Drafts of the manuscript including tentative author lists will be distributed by the paper leaders on the wiki prior to submission.

Collaborators will be asked to submit a contributions statement (and estimate a % number on how they contributed to the given manuscript).

If someone feels they have contributed and the main authors disagree this will be discussed prior to submission and the burden of evidence will be put on the collaborator. This is much easier to demonstrate if you are actively collaborating with someone outside of your institute.

Reminder from February

Authorship – main papers

There are no guarantees on which papers and authorship positions (this is the same for the RIKEN guys as well).

As we have a very large number of potential authors, author order will be decided based upon contributions statement. Typically on a FANTOM main paper there is a non-alphabetical front section of authors who contributed above the minimum, an alphabetical section who contributed and a senior/corresponding authors section.

Joint first authorship is possible for anyone if they contribute enough (see FANTOM2-4 mains)

Reminder from February

Authorship – satellite papers

Same rules:

- Authors submit an Authors contributions statement
- Draft manuscript made available on the wiki

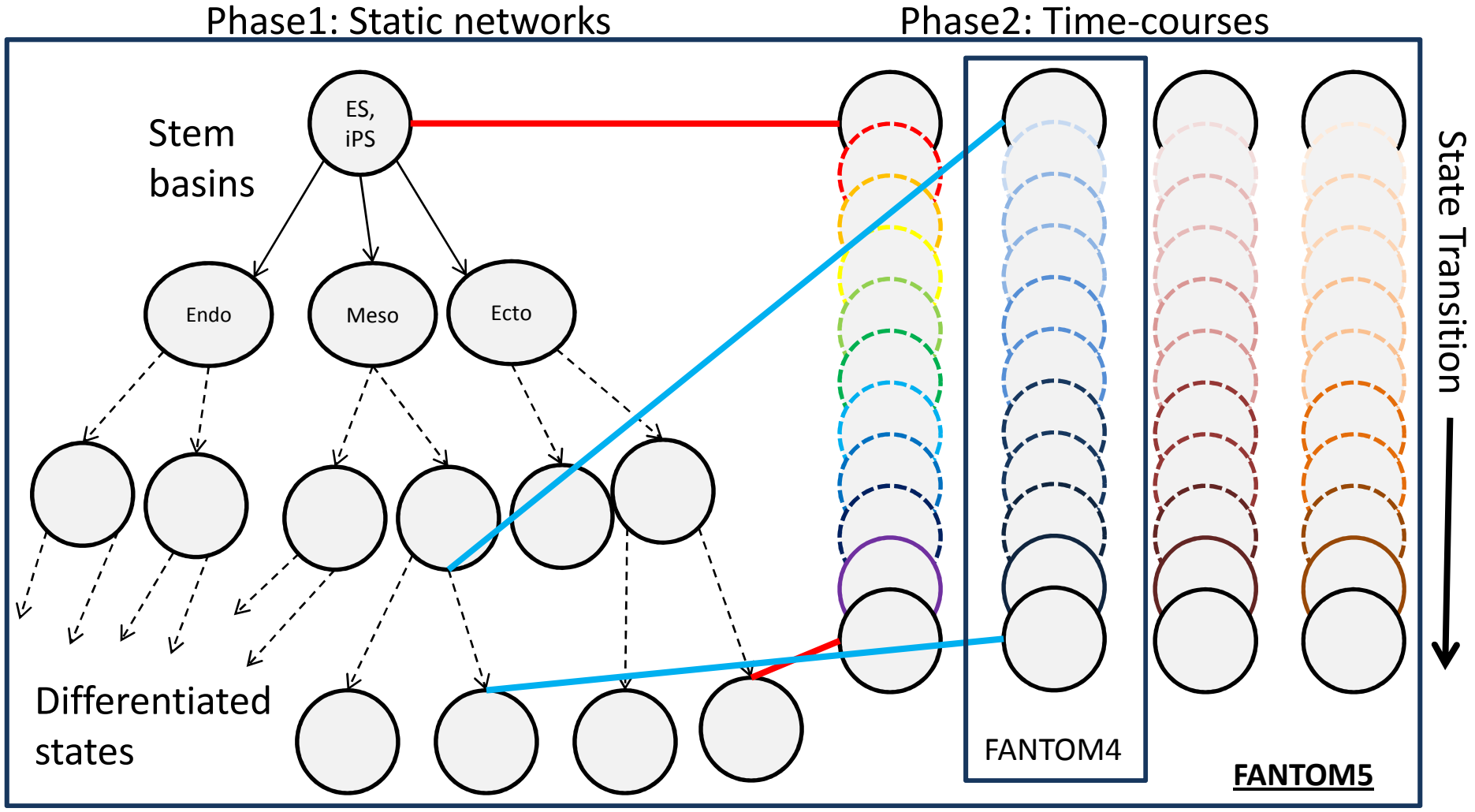
Respect and communicate with potential authors

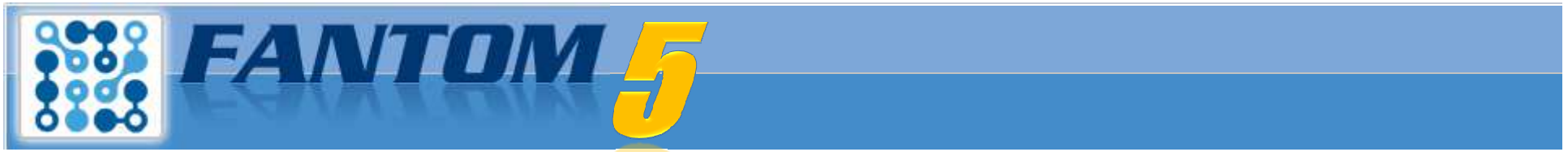
- Sample providers have invested in providing samples for the project
- Bioinformatic collaborators have invested in algorithms and analysis strategies for the project

Please collaborate.

- This is a great opportunity for wet and dry collaborations to form.
- Please during your presentations mention what help you need and what sort of collaborations you are open to.

Reminder from February





A complete mammalian promoterome

HUMAN

Focus on Human Primary cells

- Transcriptional regulatory networks 'work' at the level of a cell

Cancer cell lines

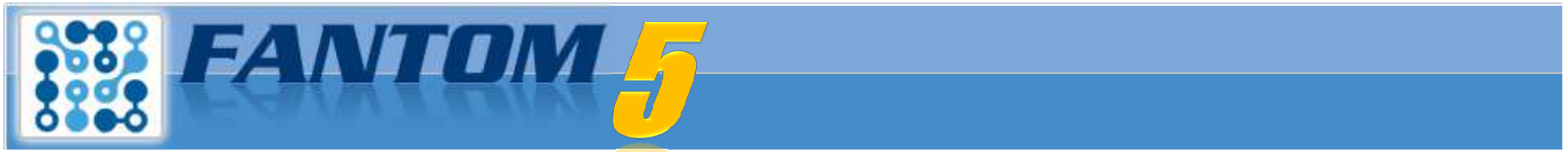
- Cancer cell line panel covering 150 different cancer subtypes (ATCC, RIKEN BRC, HSRBB)

Tissue samples (ok to capture the remaining promoters – but not as useful for TRNs)

- Capture promoters of cells missed in our primary cell collection
- Capture expression that requires the cells to be in their native tissue context

MOUSE

- Developmental time course series for all major organs during development
- Core 30 primary cell types matched to human counterparts (cross species analysis)
- Rare primary cells from mouse GFP reporter cell lines (eg. Inner ear hair cells, Intestinal stem cells, Microfold cells)



Primary cell collection ~200 types

adipose: Adipocyte - subcutaneous, Mesenchymal Stem Cells – adipose, Preadipocyte - breast, Preadipocyte - mesenteric, Preadipocyte - omental, Preadipocyte - perirenal, Preadipocyte - subcutaneous, Preadipocyte - visceral

adrenal gland: Adrenal cortical cells, Chromaffin cells

bile duct: Biliary Epithelial Cells

bladder: Smooth Muscle Cells - Bladder , Urothelial Cells

blood: Basophils, CD133+ Progenitors, CD14+ Monocytes, CD19+ B Cells, CD34+ Progenitors, CD4+ T Cells, CD8+ T Cells, Common myeloid progenitors, Dendritic Cells - monocyte immature derived, Dendritic Cells - plasmacytoid, Endothelial Progenitor Cells, Eosinophils, Erythroid Progenitors, Granulocyte-Macrophage Progenitors, Macrophage - monocyte derived, Mesenchymal Stem Cells – bone marrow, Multipotent Cord Blood Unrestricted Somatic Stem Cells, Natural Killer Cells, Neutrophils, Peripheral Blood Mononuclear Cells, promyelocytes/myelocytes

bone: Chondrocyte - de diff, Chondrocyte - re diff, Osteoblast, Osteoblast - differentiated, Synoviocyte

brain: Astrocyte - cerebellum, Astrocyte - cerebral cortex, Astrocyte - hippocampal, Cerebellar Granule Cells, Fibroblast - Choroid Plexus , Meningeal Cells, Microglia, Neural stem cells, Neuron-hippocampal, Neurons, Oligodendrocyte, Oligodendrocyte - oligospheres, Oligodendrocyte - precursors, Smooth Muscle Cells - Brain Vascular

breast: Fibroblast - Mammary , Mammary Epithelial Cell

digestive system: Intestinal epithelial cells, Lgr5+ stem cells, Paneth cells, Smooth Muscle Cells - Colonic , Smooth Muscle Cells - Intestinal

ear: Inner ear hair cells

esophagus: Esophageal Epithelial Cells, Smooth Muscle Cells - Esophageal

eye: Ciliary Epithelial Cells, Corneal Epithelial Cells, Fibroblast - Conjunctival , Iris Pigment Epithelial Cells , Keratocytes, Lens Epithelial Cells, Photoreceptor cells, Retinal Endothelial Cells, Retinal Pigment Epithelial Cells, Trabecular Meshwork Cellsheart: Cardiac Myocyte-adult, Fibroblast - Cardiac

kidney: Renal Cortical Epithelial Cells, Renal Epithelial Cells, Renal Glomerular Endothelial Cells, Renal Mesangial Cells, Renal Proximal Tubular Epithelial Cell

liver: Hepatic Sinusoidal Endothelial Cells, Hepatic Stellate Cells (lipocyte), Hepatocyte, Mesenchymal Stem Cells – hepatic

lymphatic: Endothelial Cells - Lymphatic, Fibroblast - Lymphatic

mesothelium: Mesothelial Cells

oral: Fibroblast - Gingival , Fibroblast - Periodontal Ligament, Keratinocyte - oral, Malassez-derived epithelial cells

other: Endothelial Cells - Thoracic

pancreas: Pancreatic epithelial cells, Pancreatic stromal cells

peripheral nervous system: Schwann Cells

prostate: Prostate Epithelial Cells, Prostate Stromal Cells, Smooth Muscle Cells - Prostate

reproductive system: Amnion, Amniotic Epithelial Cells, Chorion, Fibroblast - Villous Mesenchymal , Mesenchymal Stem Cells - amniotic membrane, Mesenchymal Stem Cells – umbilical, Mesenchymal Stem Cells - Wharton's Jelly , Placental Epithelial Cells, Sertoli Cells, Smooth Muscle Cells - Umbilical Vein , Smooth Muscle Cells - Uterine , Trophoblast

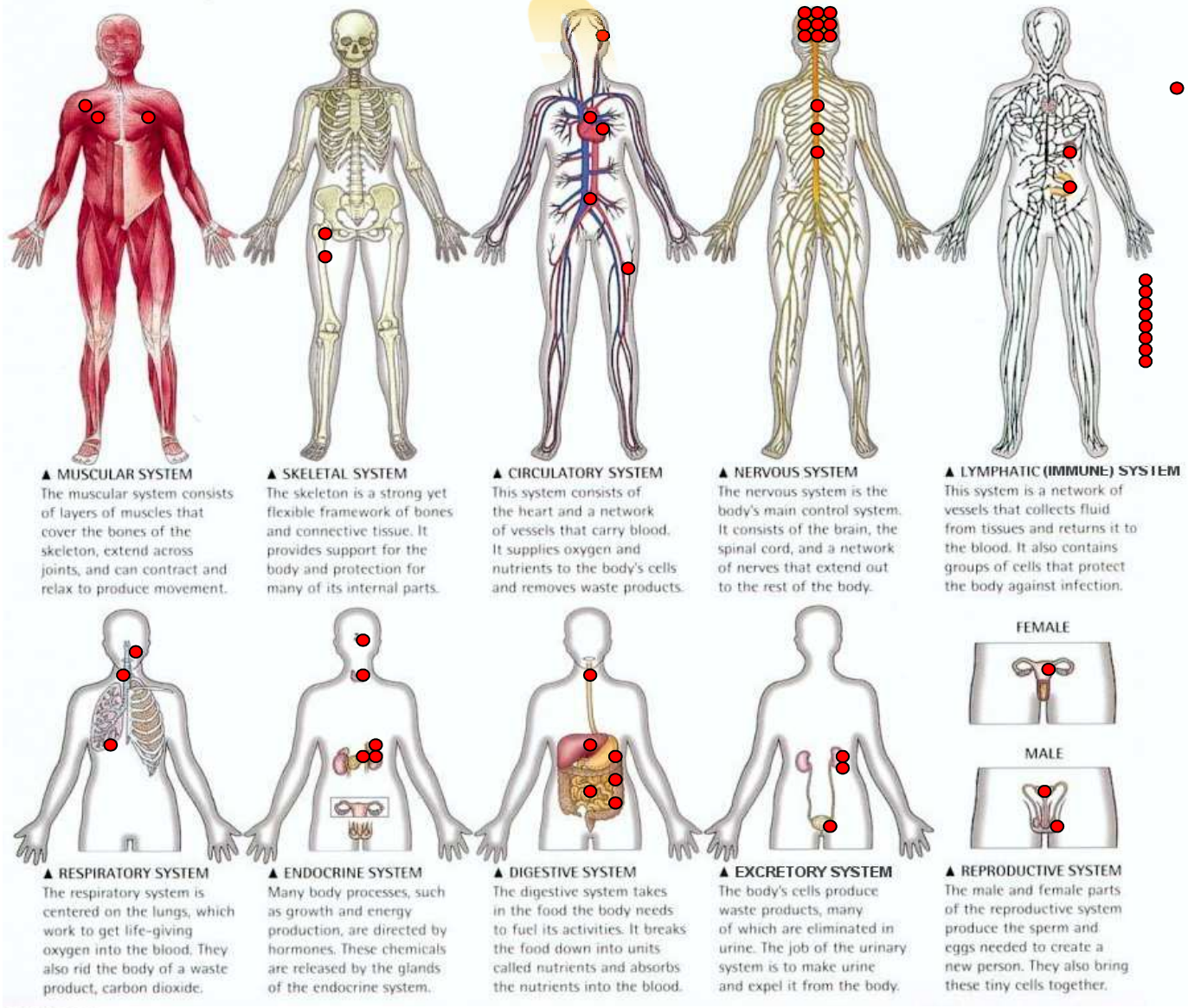
respiratory system: Alveolar Epithelial Cells, Bronchial Epithelial Cell, Bronchial Epithelial Cells, Fibroblast - Lung , Small Airway Epithelial Cells, Smooth Muscle Cells - Bronchial , Smooth Muscle Cells - Tracheal , Tracheal Epithelial Cells

skeletal muscle: Myoblast, Myocyte, Skeletal Muscle Cells, Skeletal muscle cells differentiated into Mytotubes - multinucleated, Skeletal Muscle Satellite Cells

skin: Fibroblast - Dermal , Hair Follicle Dermal Papilla Cells, Hair Follicle Inner Root Sheath Cells, Hair Follicle Outer Root Sheath Cells, Hair Germinal Matrix Cells, Keratinocyte - epidermal, Melanocyte - dark, Melanocyte - light, Sebocytes

spinal cord: Anulus Pulposus Cell, Astrocyte-spinal cord total, Mesenchymal Stem Cells - Vertebral, Nucleus Pulposus Cell, Perineurial Cells

vascular system: Endothelial Cells - Aortic, Endothelial Cells - Artery, Endothelial Cells - Microvascular , Endothelial Cells - Vein, Fibroblast - Aortic Adventitial , Fibroblast - Pulmonary Artery , Pericytes, Smooth Muscle Cells - Aortic , Smooth Muscle Cells - Brachiocephalic , Smooth Muscle Cells - Carotid , Smooth Muscle Cells - Coronary Artery , Smooth Muscle Cells - Internal Thoracic Artery , Smooth Muscle Cells - Pulmonary Artery , Smooth Muscle Cells - Subclavian Artery , Smooth Muscle Cells - Umbilical Artery

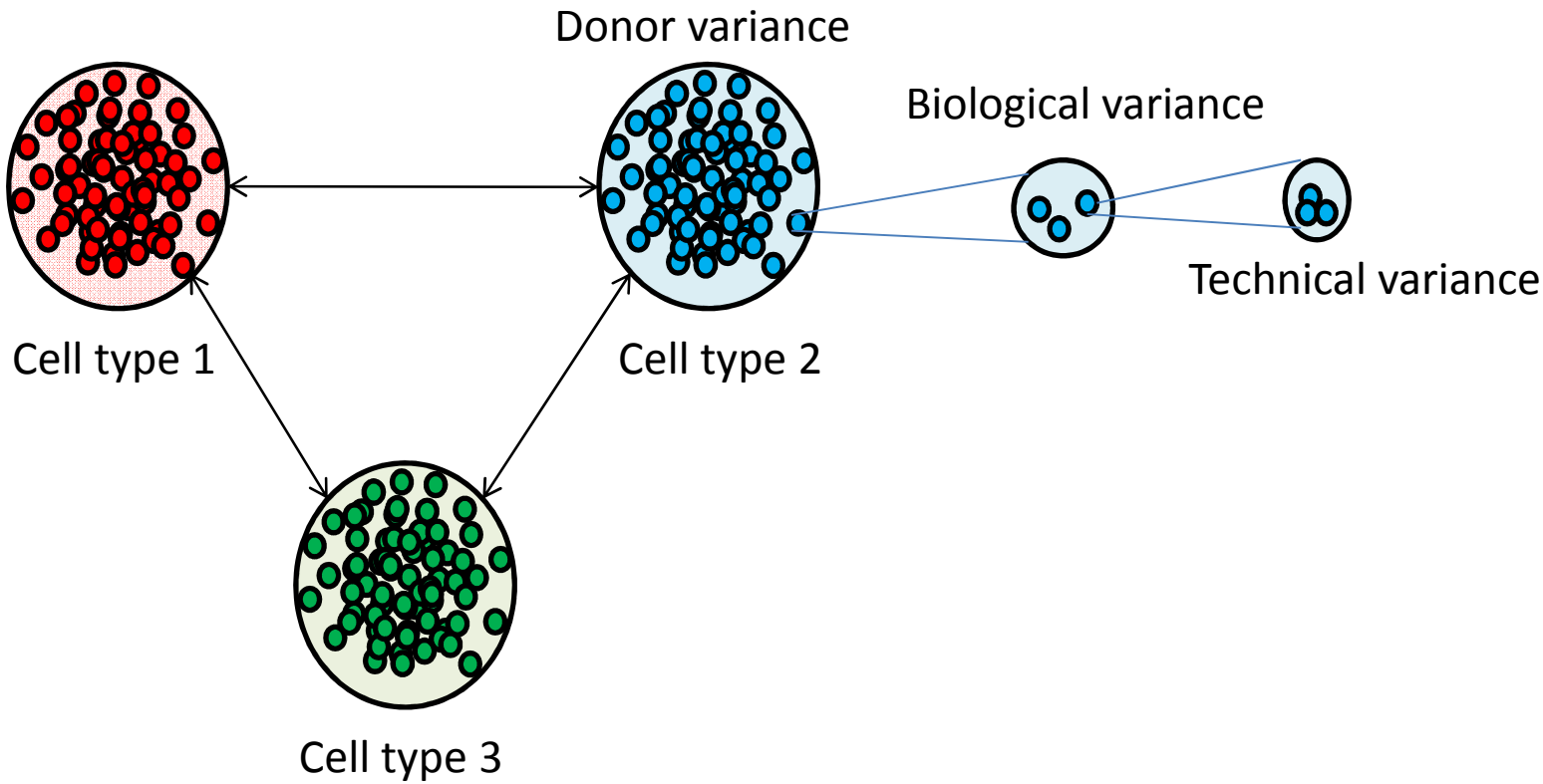


- Primary cells

RIKEN OSC
RIKEN
Omics Science Center

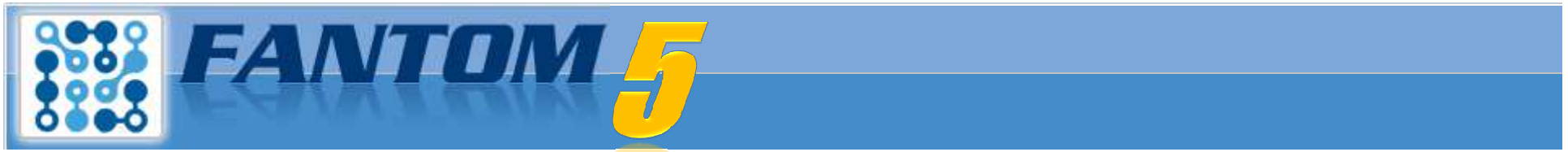
Al Forrest

Use of donor triplicates



For each **primary** cell type, 3 separate donors are used

(ie. 3 genotypes – we want networks that generalise across individuals)



Human cancer cell lines ~250

Adrenal: Neuroblastoma

Bile Duct: Carcinoma

Bladder: transitional-cell carcinoma

Blood: acute lymphoblastic leukemia (T-ALL), acute myeloid leukemia (FAB M0), acute myeloid leukemia (FAB M1), acute myeloid leukemia (FAB M3), acute myeloid leukemia (FAB M4), acute myeloid leukemia (FAB M4eo), acute myeloid leukemia (FAB M5), acute myeloid leukemia (FAB M6), acute myeloid leukemia (FAB M7), adult T-cell leukemia, Burkitt's lymphoma, chronic lymphocytic leukemia (T-CLL), Chronic myelogenous leukemia, diffuse large B-cell lymphoma, Disease: xeroderma pigentosum, Hodgkin's lymphoma, myeloma, non T non B acute lymphoblastic leukemia (ALL), Splenic lymphoma with villous lymphocytes

Bone: osteosarcoma, osteoclastoma

Brain: astrocytoma, glioma, glioblastoma, meningioma, neuroblastoma

Breast: carcinoma

Cervix: arygyrophil small cell carcinoma, epidermoid carcinoma, glassy cell carcinoma, large cell non-keratininzing squamous cell carcinoma, squamous cell carcinoma

Chorion: choriocarcinoma

Colon: carcinoma

Duodenum: papillotubular adenocarcinoma

Embryo: teratoma

Endometrium: clear cell carcinoma, carcinoma, stromal sarcoma, leiomyosarcoma, mixed mullerian tumor

Esophagus: squamous cell carcinoma

Eye: retinoblastoma

Gall bladder: Carcinoma

Hair follicle: Malignant trichilemmal cyst

Intestine: small cell gastrointestinal carcinoma

Kidney: neuroectodermal tumor, Wilm's tumor, renal cell carcinoma

Liver: cholangiocellular carcinoma, hepatoblastoma, hepatoma

Lung: giant cell carcinoma, large cell lung carcinoma, adenocarcinoma, small cell lung carcinoma, squamous cell lung carcinoma

Mesothelium: mesothelioma

Mouth: oral squamous cell carcinoma, maxillary sinue tumor

Ovary: clear cell carcinoma, Granulosa cell tumor, Krukenberg tumor, mucinous adenocarcinoma, serous adenocarcinoma, serous cystadenocarcinoma

Pancreas: ductal cell carcinoma, pancreatic carcinoma

Pelvis: neuroblastoma

Peripheral nervous system: neuroectodermal tumor, schwannoma

Prostate: cancer

Rectum: cancer

Skeletal muscle: rhabdomyosarcoma

Skin: melanoma, merkel cell carcinoma, keratocanthoma

Soft tissue: epithelioid sarcoma, extraskeletal myxoid chondrosarcoma, mesodermal tumor, myxofibrosarcoma

Stomach: gastric adenocarcinoma, cancer, gastrointestinal carcinoma, signet ring carcinoma, small cell gastrointestinal carcinoma

Synovium: sarcoma

Testis: germ cell embryonal carcinoma

Thyroid: anaplastic carcinoma, papillary adenocarcinoma, squamous carcinoma

Uterus: adenosquamous carcinoma, carcinosarcoma, endometrioid adenocarcinoma

Vulva: epidermoid carcinoma

Tissue RNA collection ~200

Immune system: Blood, Peripheral Leukocytes, Lymph Node, Spleen, Tonsil, Thymus

Brain: Whole Brain, caudate nucleus, cerebellum, cerebrum, Corpus Callosum, Dura Mater, Frontal Cortex, Frontal Lobe , gyrus temporalis, hippocampus, hypothalamus, Insula, locus coeruleus, medial frontal gyrus, medial temporal gyrus, Medulla Oblongata, Nucleus Accumbens, Occipital Lobe, Occipital Pole, Paracentral Gyrus, Parietal cortex, Parietal Lobe, Pons, Postcentral Gyrus, Putamen, substantia nigra, superior occipital gyrus, superior parietal gyrus, Temporal Lobe, amygdala, thalamus (pulvinar)

Digestive system: Cecum, Colon, Small intestine Duodenum, Rectum, Stomach, Esophagus, Large intestine

Heart: Whole Heart, Whole Diseased Heart, Whole Diseased Post-Infarction, Aorta , Left Atrium, Left Ventricle

Excretory systems: Liver, Kidney, Bladder

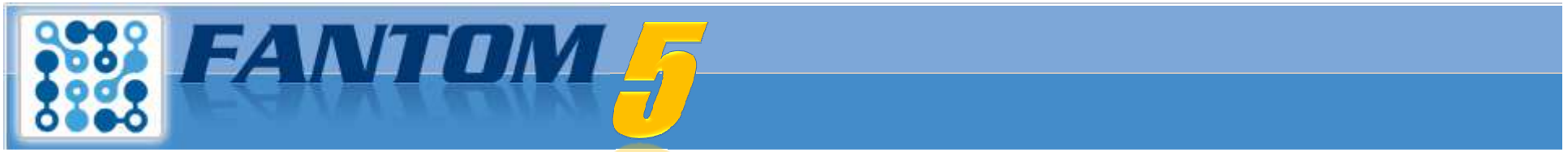
Respiratory system: Lung, Trachea, Diaphragm

Reproductive system: Testes, Ovary, Uterus, Umbilical cord, Placenta, Prostate, Epidymidis, Cervix , Penis, Vagina

Oral system: Salivary Gland, Tongue, throat

Other: Pancreas, Retina, Skeletal Muscle, Skin, Smooth Muscle, Spinal Cord, Thyroid , Eye , Adipose, Adrenal , Mammary Gland, Embryo

+~750 mouse developmental tissue samples



Reminder from February

FANTOM5 Schedule

Phase 1:

Datasets: Primary cells, tissues, cell lines (**completed!**)
Papers: Promoterome, Variome
Ume Blossom Meeting: February 20-26
Submission target: June-August (**delayed**)
New Submission target: December-March

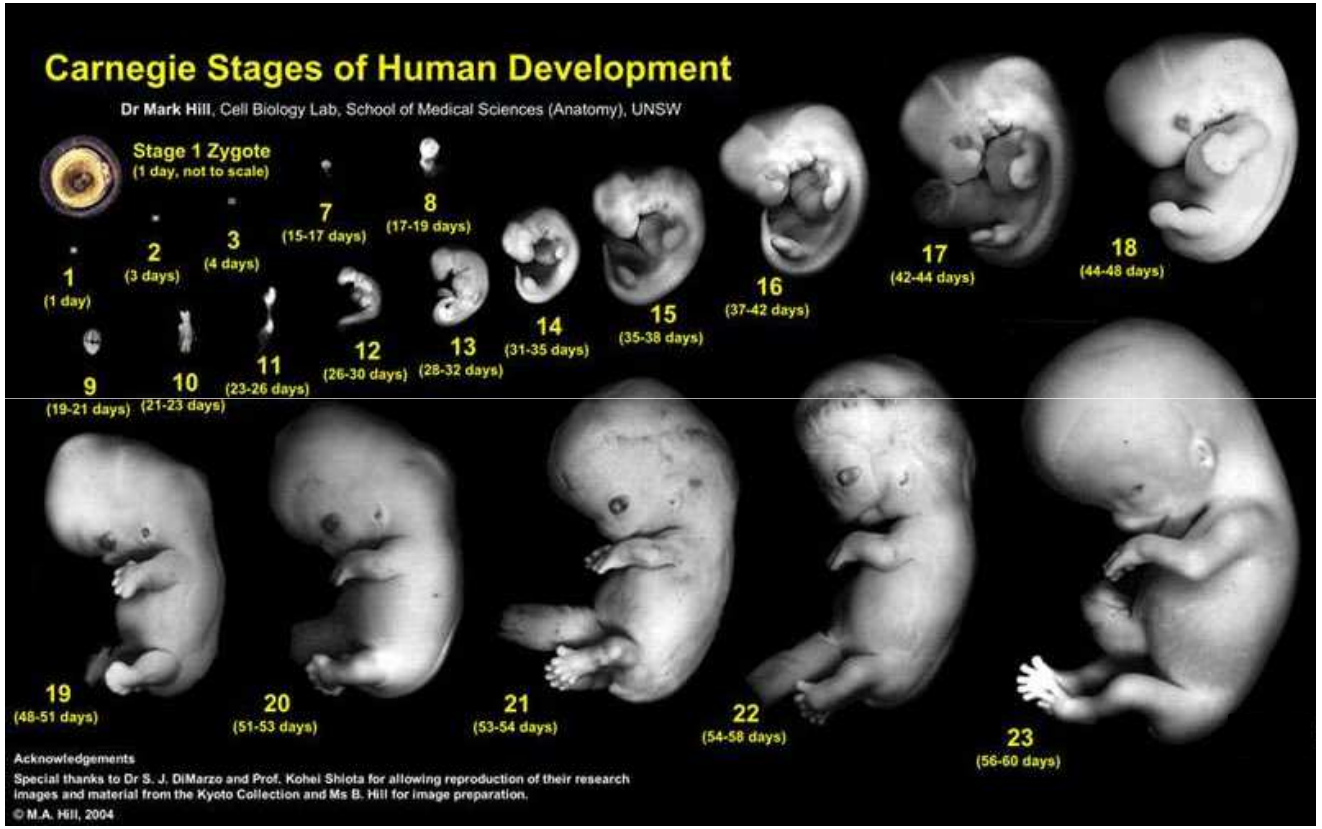
Phase 2:

Datasets: Time courses (**last libraries being started this week!**)
Papers: Cellular Basin logic, ncRNAs, cell conversion
Kōyō Meeting: October 17-20
Submission target: December-March

A complete catalog of promoters used to regulate distinct cellular states.

RIKEN Omics Science Center
Yokohama Institute, JAPAN
Unit Leader
Alistair Forrest, Ph.D.

Human development

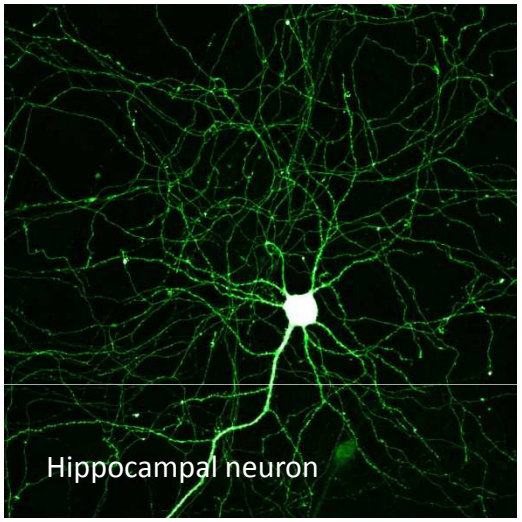


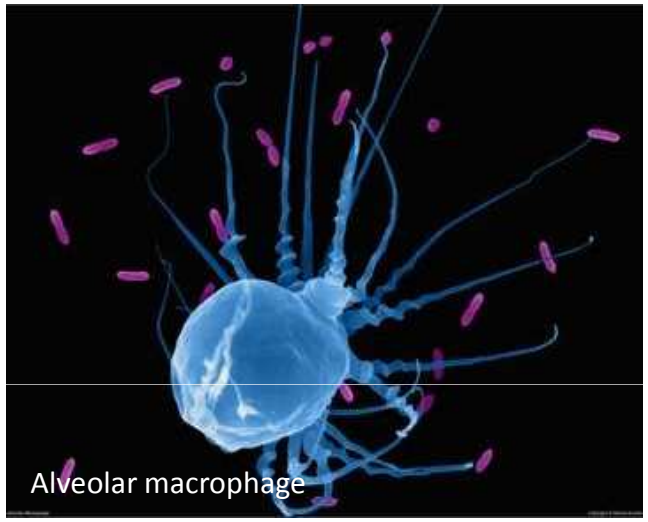
We are complex multicellular organisms composed of at least 400 distinct cell types, that have all arisen through sequential specialization from a single totipotent cell.

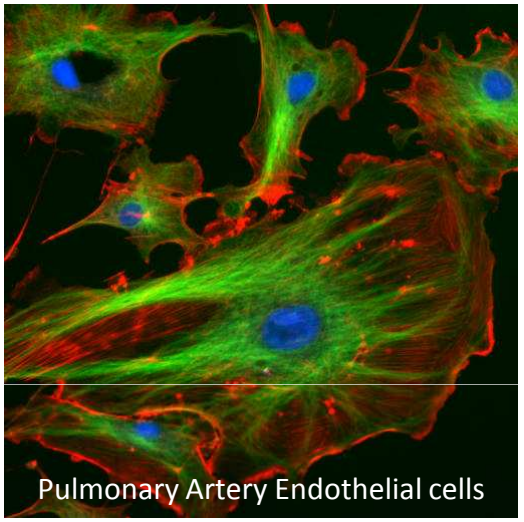
Acknowledgements: Images from Dr Mark Hill, Cell Biology Lab, UNSW

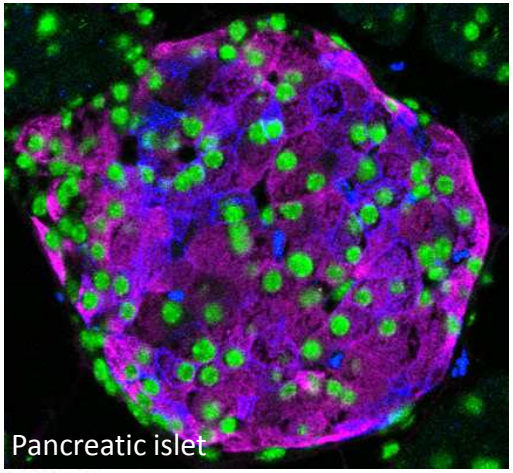
FANTOM 5

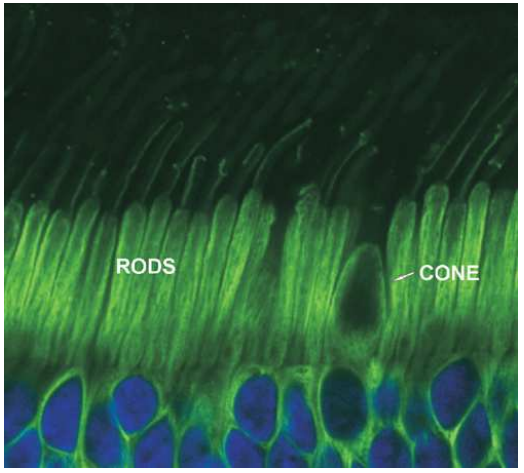
Diversity of encoded cellular states









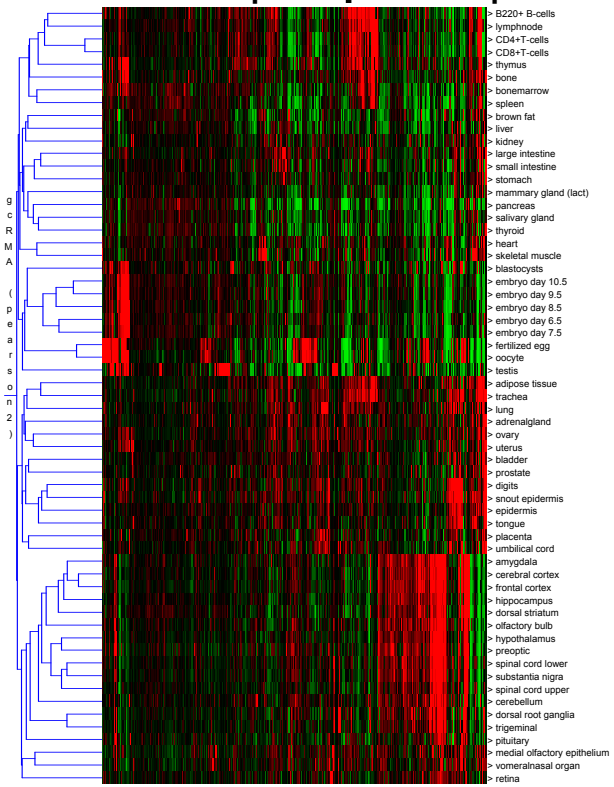


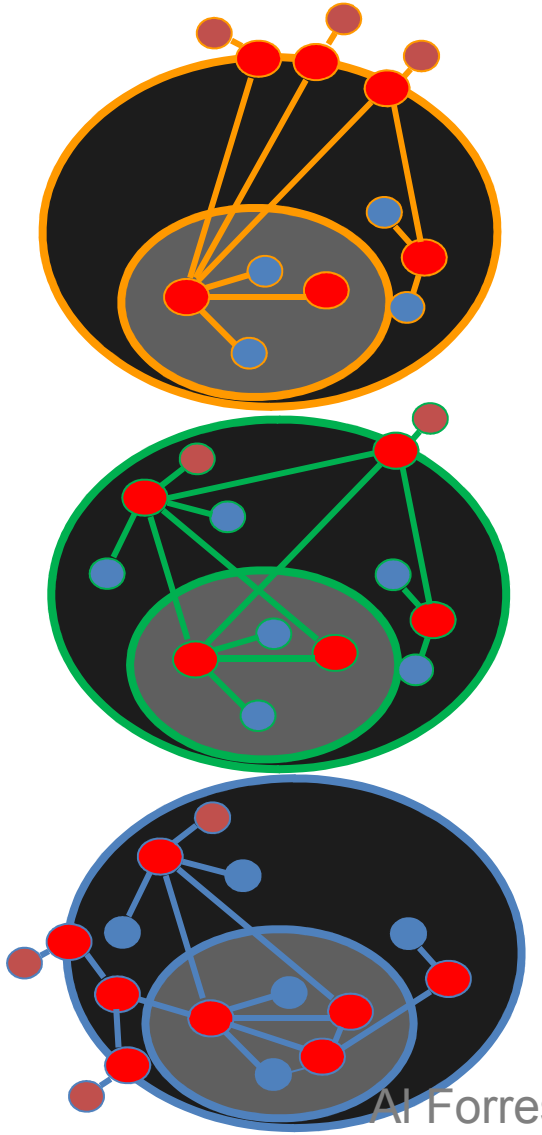
RIKEN OSC
RIKEN
Omics Science Center

Acknowledgements pending: all found on google images

Cell specific networks

- Shared core components
- Common lineage components
- Cell specific components





Forrest et al. 2006, *BMC Bioinformatics* 2006, 7:82
Mined from GNF http://symatlas.gnf.org

Building cell specific network models

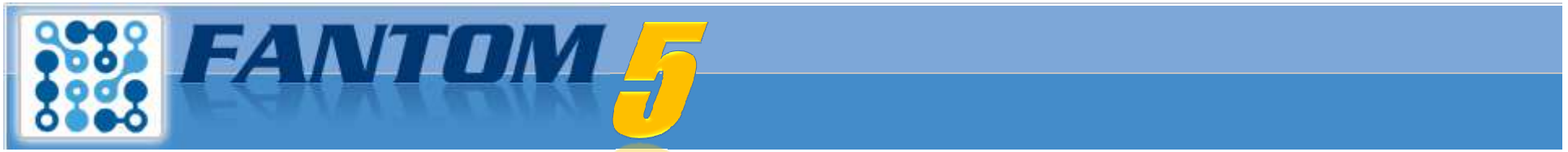
Broadly the project aims to understand how each cellular state is encoded by constructing **cell specific** transcriptional regulatory network models for the majority of cell types

Using **Cap Analysis of Gene Expression (CAGE)** across a very diverse collection of human samples (primary cells, cell lines, tissues) to measure expression and identify promoter regions.

NODE expression levels are measured by Cap Analysis of Gene Expression (CAGE)

EDGE are predicted from transcription factor binding sites(TFBS) in promoter regions identified by CAGE and co-expression

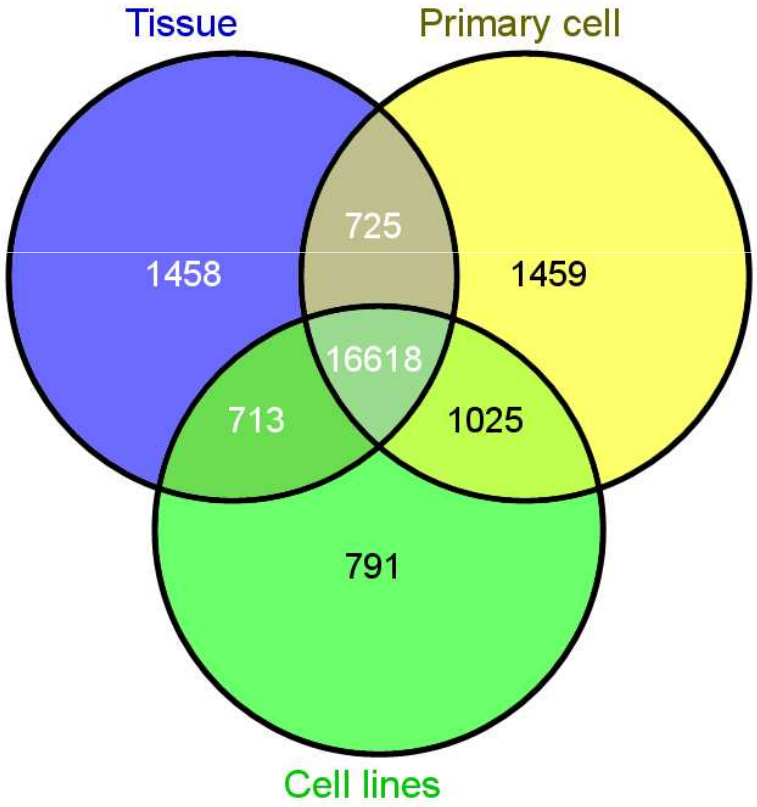
Key NODES are identified by tissue specificity of transcription factor expression and motif overrepresentation.



How does the data look? (1)

- Majority of known genes are detected
  - Primary cells, cell lines and tissues are required
- Expression atlas makes sense
  - Lineage specific markers are specific
  - “housekeeping genes” are broadly expressed
- Identifies promoter usage across 1059 samples
  - Corrects gene models
  - Identifies novel loci

All three sample types are needed to recover gene expression

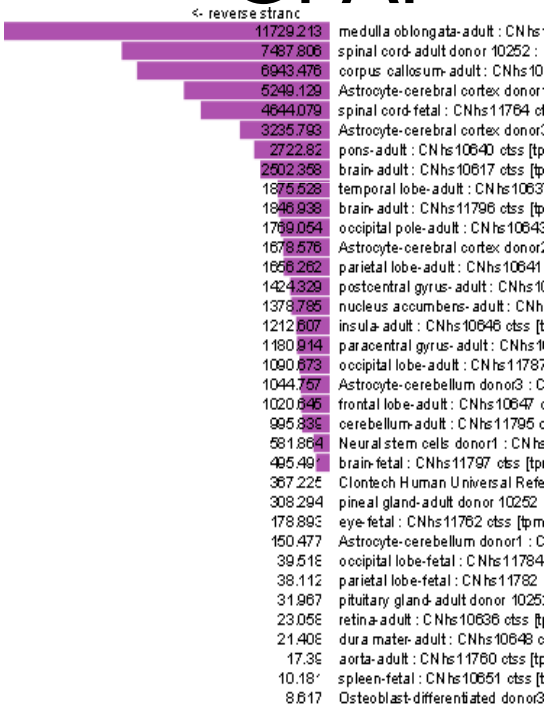


22789 Human gencode genes >=10TPM (tags per million)

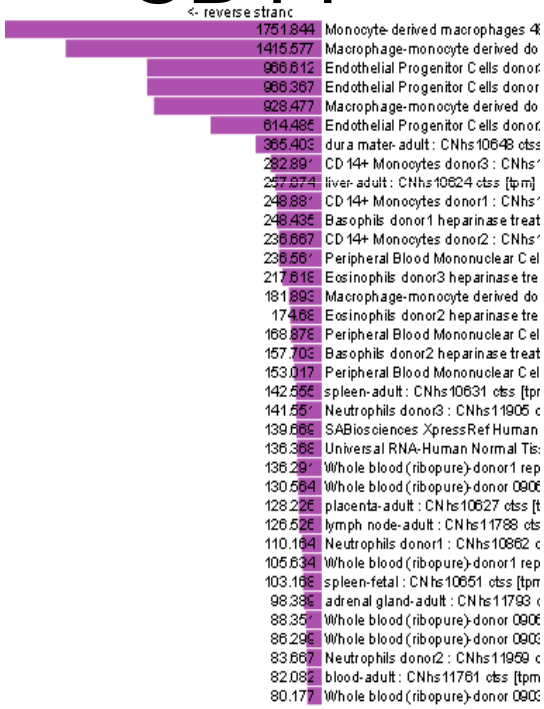
Promoter level expression atlas.

Top 30 most highly expressing samples (of 1059 in Human).

GFAP

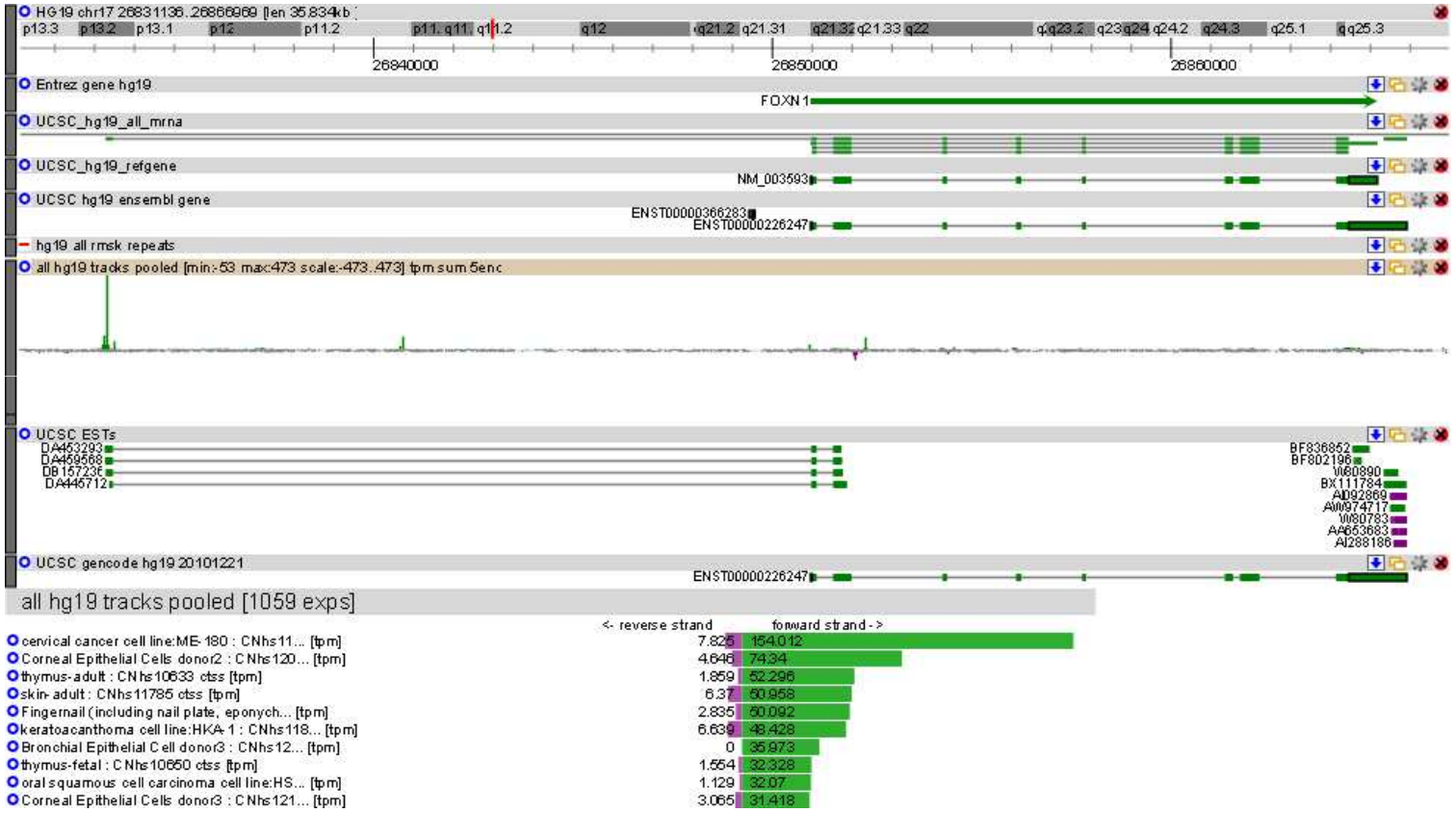


CD14



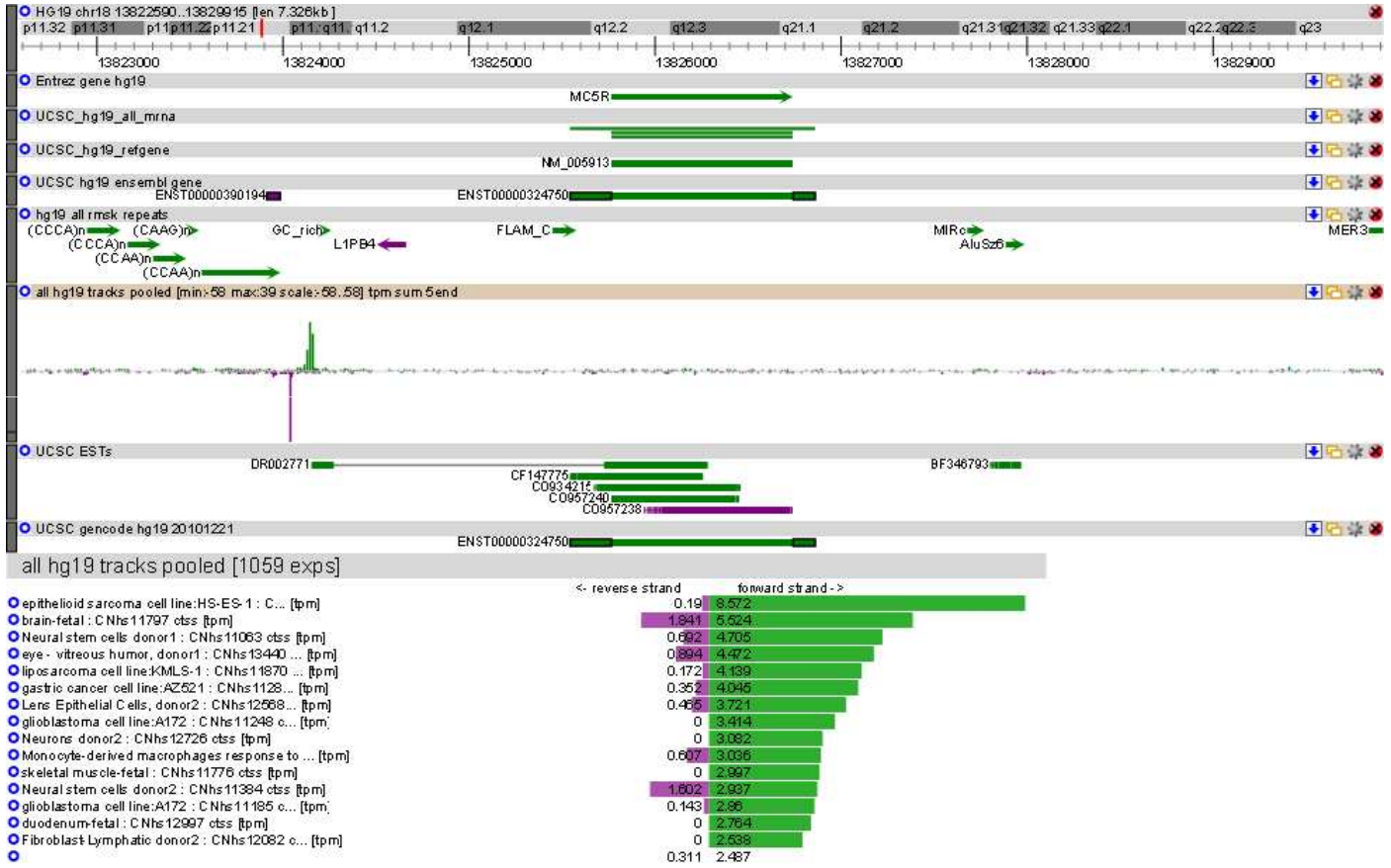
ACTB

Improved annotation



FOXN1 is transcribed from HG19::chr17:26833309-26833408
and is most highly expressed in the ME-180 cell line, corneal epithelial cells and….

Improved annotation



MC5R is transcribed from HG19::chr18:13824105-13824204 and is most highly expressed in the HS-ES-1 line…

Putative novel ncRNAs

ncRNAs on Thursday!

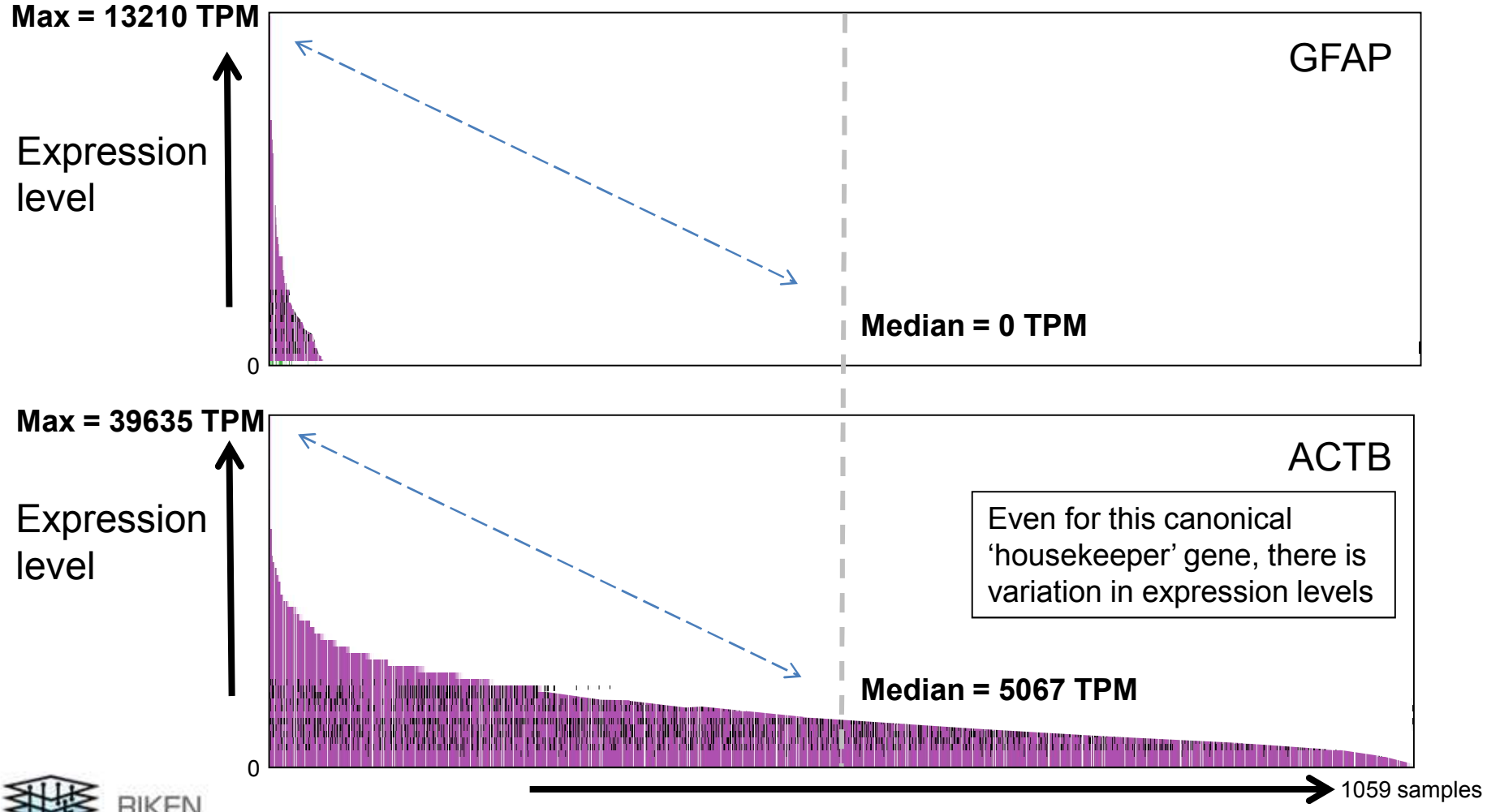
Major observations

- Two classes of gene expression modes
  - Housekeeper ‘like’ – broadly expressed at similar levels
  - Lineage specific – not detected in the majority of samples

- Composite promoter architectures
  - Similar to alternative promoters
  - Many promoters are composite and are composed of multiple closely located independently controlled regions
  - Can combine different expression modes

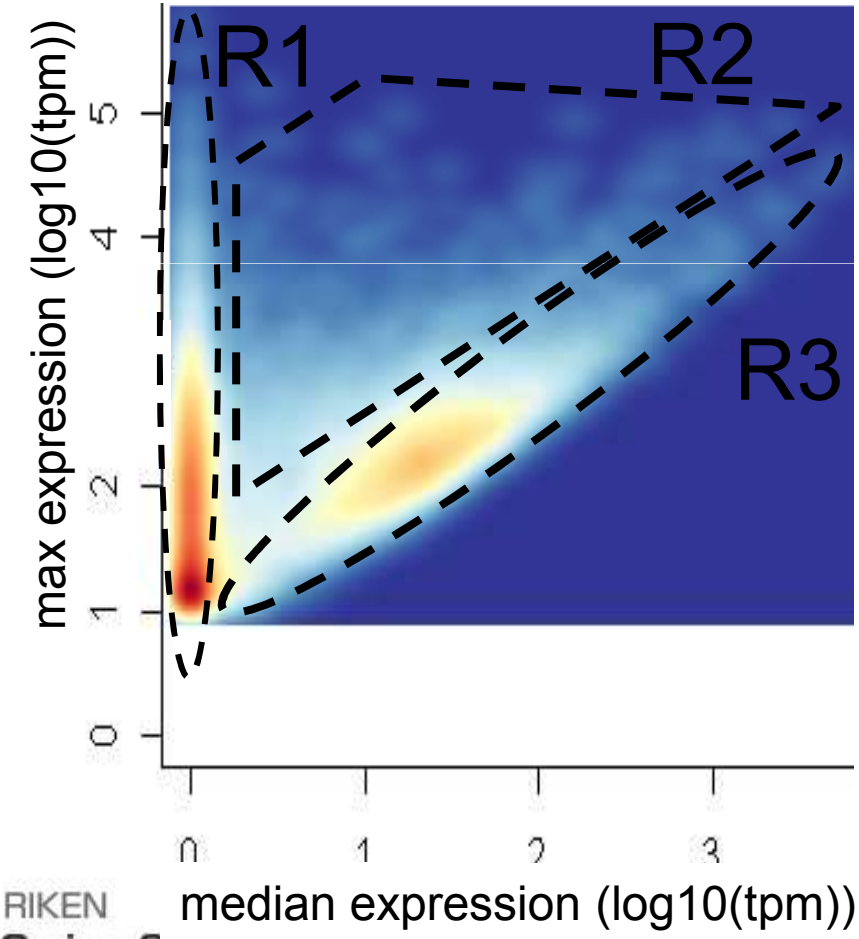
Expression Modes

Example TPM distributions for R1 and R3 across 1059 human samples



Expression Modes – update using Gencode v7

Distribution of maximum and median TPM values for Gencode genes
(based on HeliscopeCAGE data from 1059 human samples) – 22369 have max TPM >=10



Legend:

R1: Highly restricted expression
max >> median (12427 genes/53%)

R2: Partly restricted expression
max > median (5185 genes/22%)

R3: House keeping genes
max ~= median (5757 genes/25%)

Red and white indicate high areas of high and low density respectively.

*See next page for how numbers calculated for each class

Taking Gencode max >=10TPM

Counts for R1, R2 and R3 are calculated thusly:

R1 genes: *(log10(median+1) <= 0.5* = **12457**

R2 genes: *(log10(max+1) - log10(median+1) >1 AND log10(median+1) > 0.5* = **5185**

R3 genes: *(log10(max+1) - log10(median+1) <=1 AND log10(median+1) > 0.5* = **5757**

Breakdown of max source (as max is susceptible to outliers and possible gene amplifications in cell lines)

12457 genes in R1 (log10(median+1) <= 0.5

Max source from cell line	3584	(29%)
Max source from primary	4466	(36%)
Max source from tissue	4377	(35%)

5185 genes in R2 (log10(max+1) - log10(median+1) >1 AND log10(median+1) > 0.5

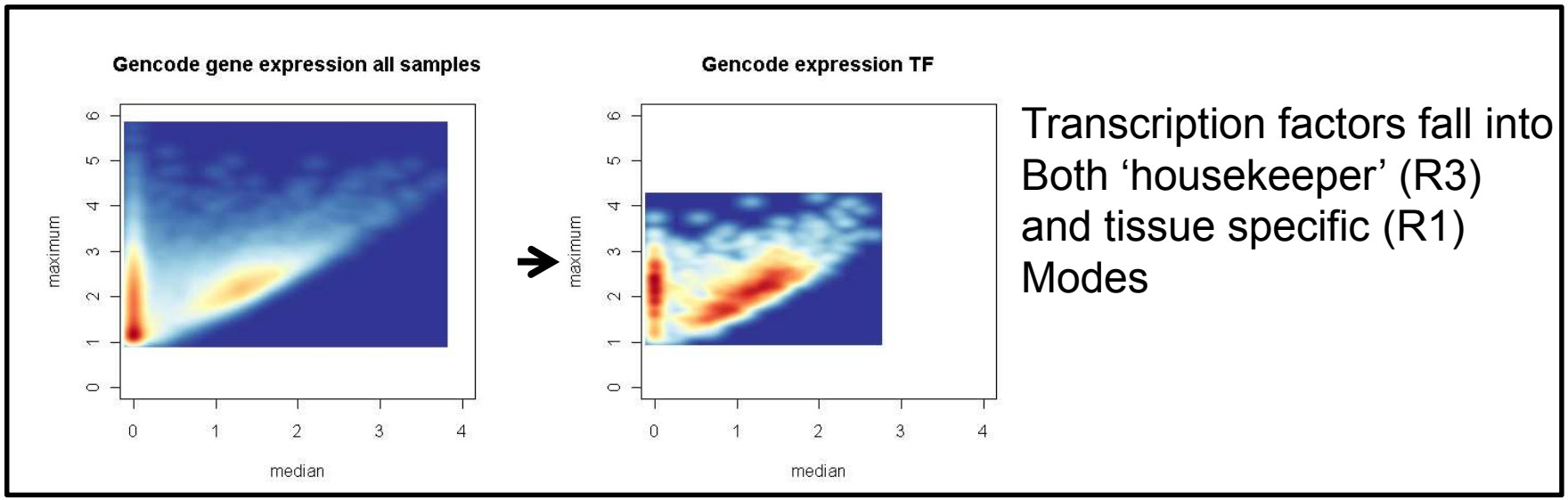
Max source from cell line	1762	(34%)
Max source from primary	2675	(52%)
Max source from tissue	748	(14%)

5757 genes in R3 (log10(max+1) - log10(median+1) <=1 AND log10(median+1) > 0.5

Max source from cell line	2711	(47%)
Max source from primary	2567	(45%)
Max source from tissue	479	(8%)

Al Forrest

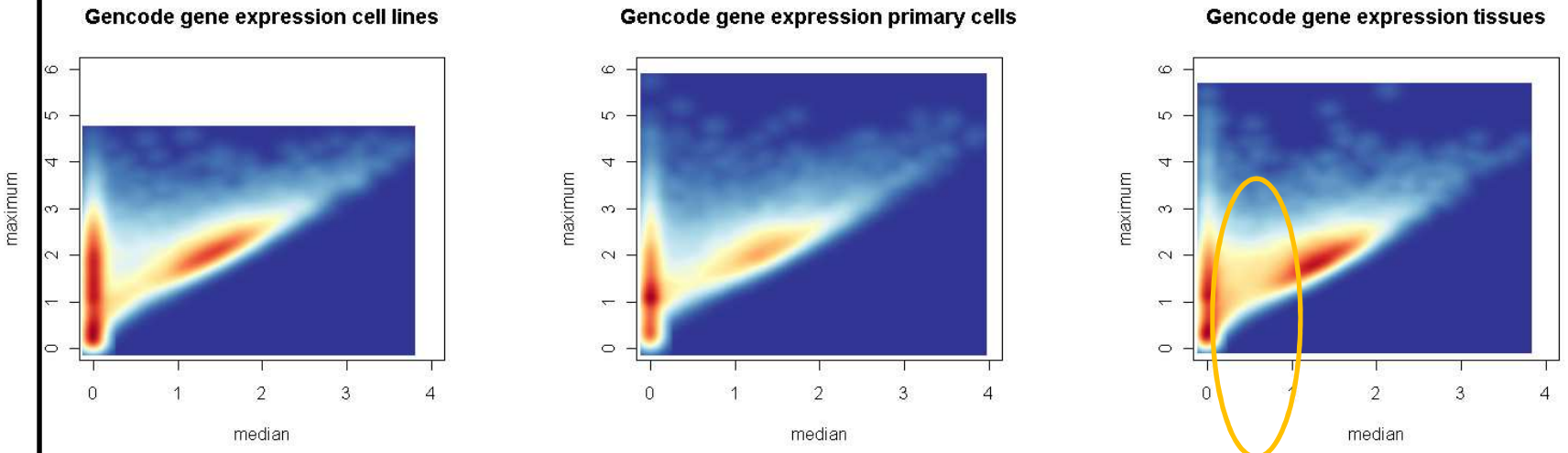
Gencode v7 analysis on Update12



Transcription factors fall into Both 'housekeeper' (R3) and tissue specific (R1) Modes

max expression (log10(tpm))

Same plots but calculated using only cell line, primary or tissue data

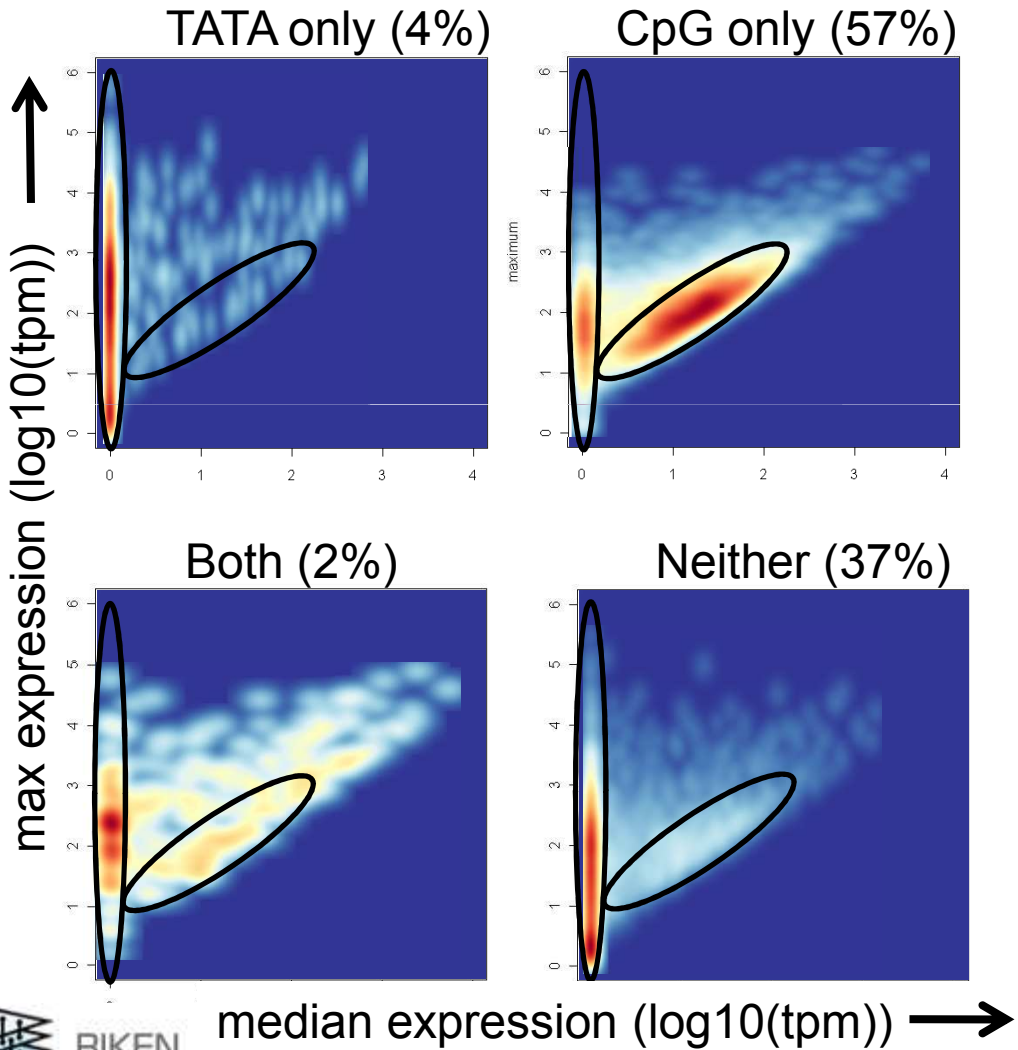


Cell mixtures in tissue
Over-estimate broad class

median expression (log10(tpm))

Al Forrest

Expression Modes – split by features

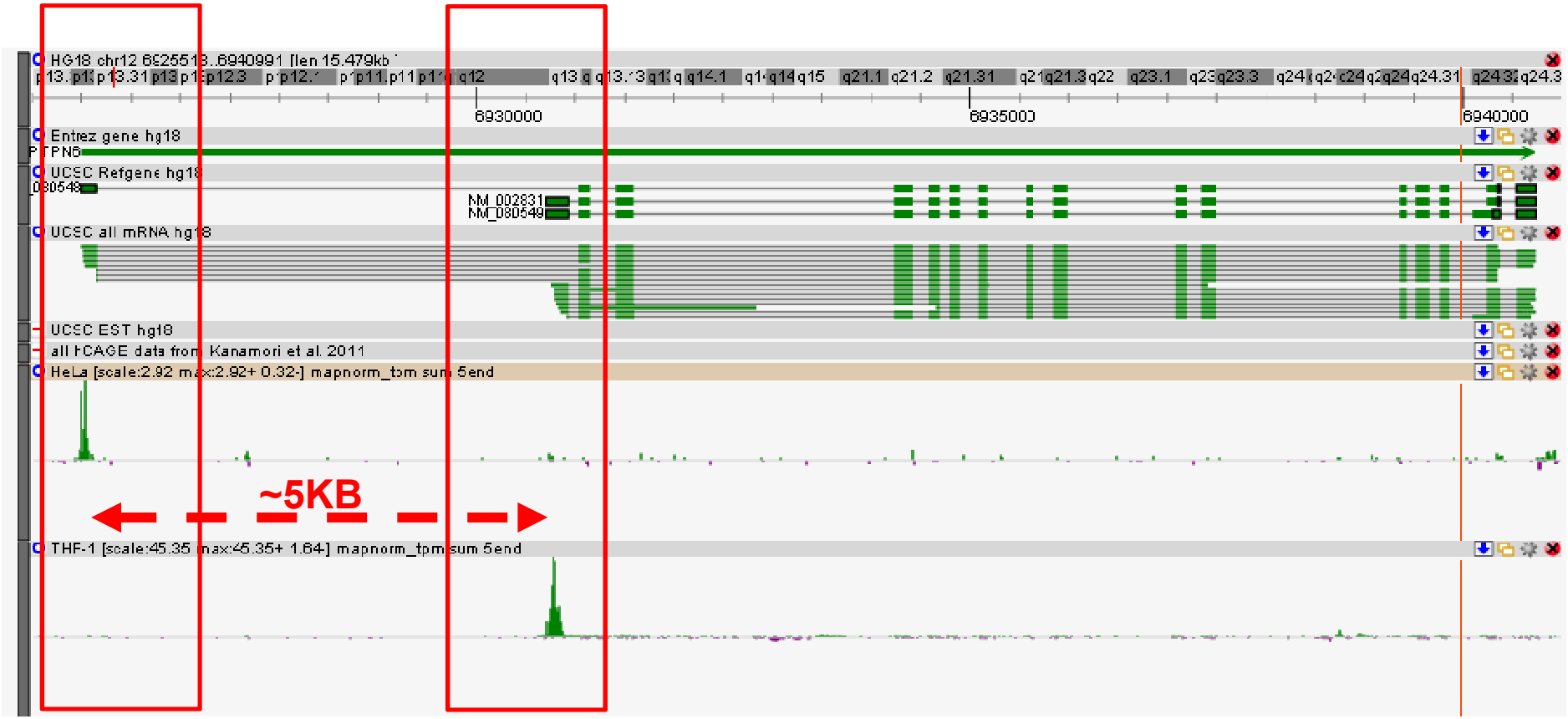


Expression modes associate with known promoter features.

Housekeepers mostly CpG

Tissue/cell specific mostly non CpG

Reminder: Alternative promoter usage in the PTPN6 locus



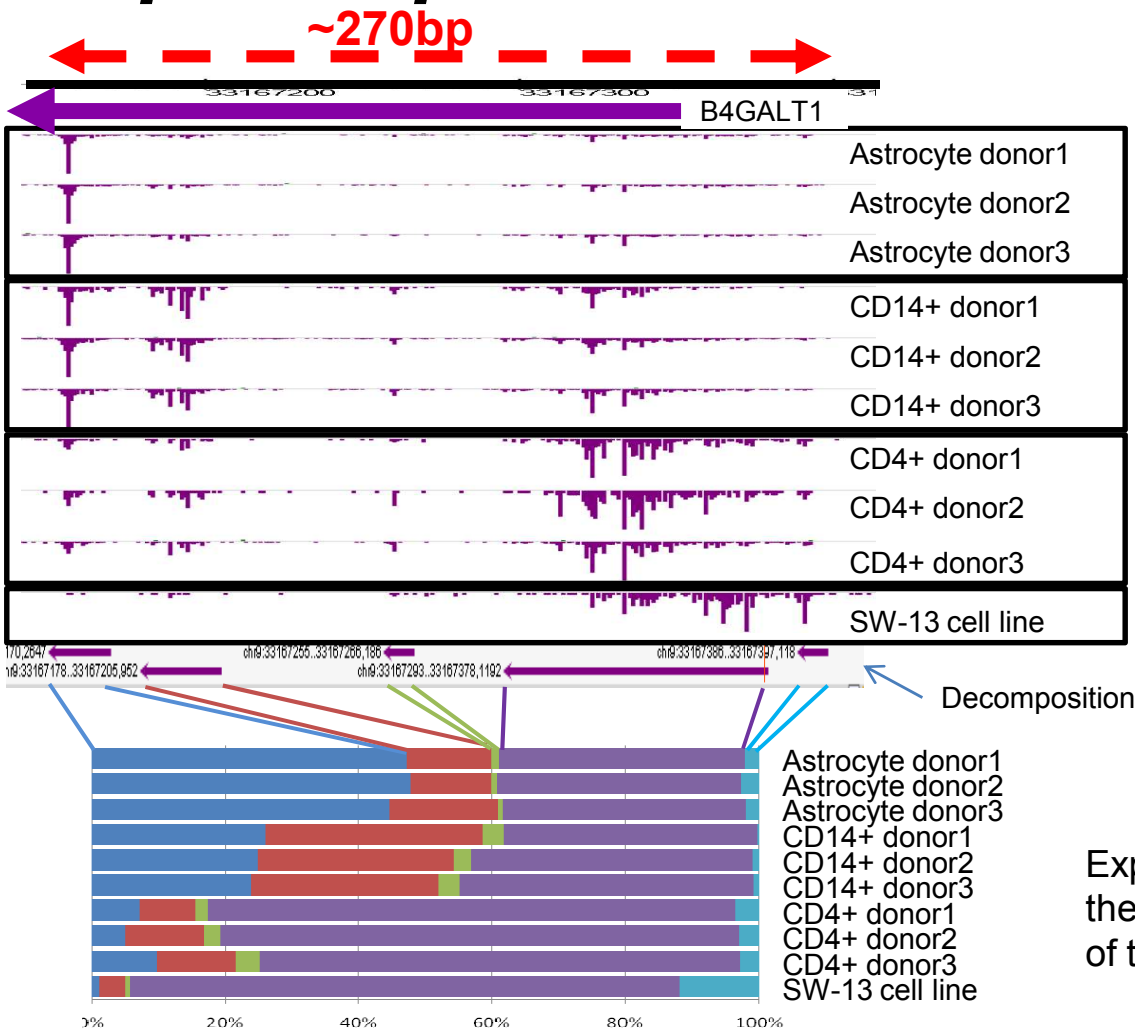
HeLa Promoter THP-1 Promoter

RIKEN Omics Science Center

Kanamori-Katayama, Itoh, Kawaji et al. 2011 Genome Research.
"Unamplified cap analysis of gene expression on a single-molecule sequencer"

Al Forrest

Composite promoter architectures



TSS preferences observed for 270bp in the B4GALT1 core promoter in primary Astrocytes, CD14+ monocytes and CD4+ T-cells

Decomposition reveals 5 independently controlled TSS regions

Expression output for the B4GALT1 locus is the sum of these TSS activities

Implications for networks

- Composite architectures
  - May improve robustness
  - Allow context specific upregulation of housekeepers
  - Allow multiple lineage dependent modes for induction

- Motif over-representation analyses
  - Separate ‘house keeper’ and ‘lineage specific’ modes
  - Weight motifs based on distance and expression for independent components in composite architectures

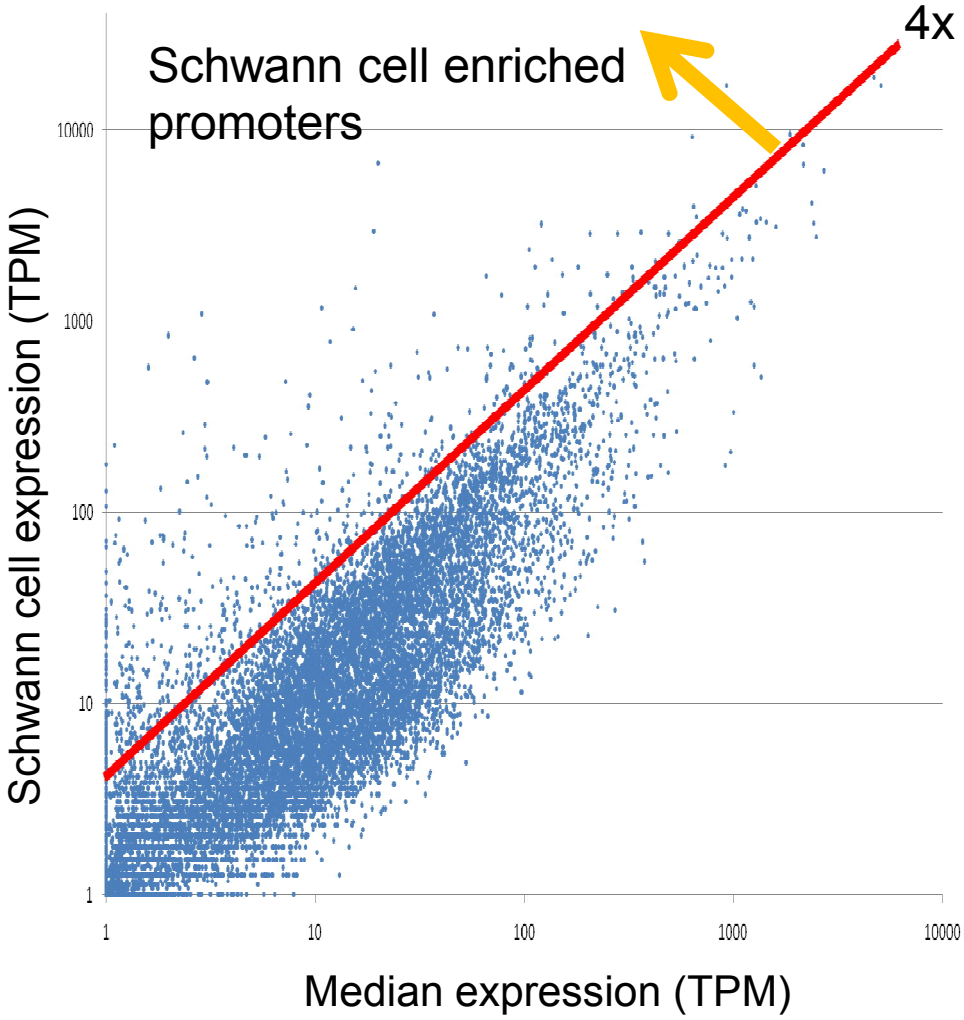
Toward cell specific networks

- Identify key transcription factors
- Identify key motifs
- Identify edges between key factors
  - MARA (motif activity response analysis) predictions
  - ChIP-seq of Key factors
  - KD-CAGE (siRNA knock-down of TF followed by CAGE)

Cell specific TFs and motifs

Schwann cell promoter level expression
-> Schwann cell enriched promoters
-> Schwann cell enriched genes
-> Schwann cell enriched TFs

	median TPM all samples	schwann cell TPM	Fold enrichment
ZIC1	0.0	91.2	92.2
SOX11	0.0	42.9	43.9
SIX1	4.8	213.9	37.3
HMGA2	4.2	193.6	37.1
ZIC2	0.1	36.8	34.5
OSR1	1.2	76.2	34.5
FOXC2	1.4	72.6	31.1
HOXA9	0.1	29.7	27.0
PRRX1	0.6	36.7	24.0
SOX10	0.0	19.7	20.7



RIKEN
Omics Science Center
RIKEN OSC

Al Forrest

Cell specific TFs and motifs

***De novo* Motif Results for Eosinophils**

Rank	Motif	P-value	log P-pvalue	% of Targets	% of Background	STD(Bg STD)	Best Match/Details	Motif File
1		1e-580	-1.336e+03	43.62%	16.68%	Cbp (0bp)	PB0058.1_Sfpi1_1 More Information | Similar Motifs Found	motif file (matrix)
2		1e		32.83%	15.14%	Cbp (0bp)	GDNF_f1_ppm_HOCOMOCOv7	motif file
3								
4								
5								
6								
7								

***De novo* Motif Results for Sertoli_Cells**

Rank	Motif	P-value	log P-pvalue	% of Targets	% of Background	STD(Bg STD)	Best Match/Details	Motif File
1		1e-46	-1.070e+02	2.56%	0.19%	0bp (0bp)	Znf263(Zf)/K562-Znf263-ChIP-Seq/Homer More Information | Similar Motifs Found	motif file (matrix)
2		1e-33		11.12%	5.16%	0bp (0bp)	MA0313.1_EAP2	motif file
3								
4								
5								
6								
7								

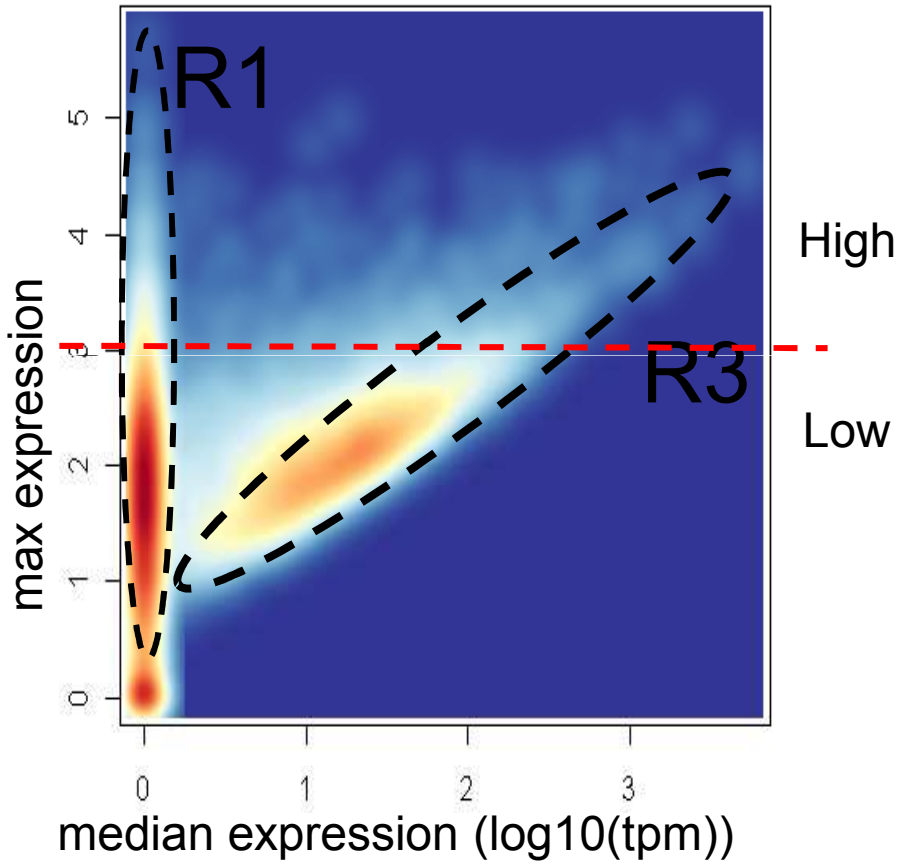
***De novo* Motif Results for Schwann_Cells**

Rank	Motif	P-value	log P-pvalue	% of Targets	% of Background	STD(Bg STD)	Best Match/Details	Motif File
1		1e-176	-4.058e+02	21.45%	6.67%	0bp (0bp)	MA0060.1_NFYA More Information | Similar Motifs Found	motif file (matrix)
2		1e-46	-1.079e+02	23.09%	11.51%	0bp (0bp)	Sp1(Zf)/Promoter/Homer More Information | Similar Motifs Found	motif file (matrix)
3		1e-35	-8.232e+01	4.32%	1.27%	0bp (0bp)	GFY(?)/Promoter/Homer More Information | Similar Motifs Found	motif file (matrix)
4		1e-30	-7.112e+01	2.40%	0.25%	0bp (0bp)	PO2F1_f1_ppm_HOCOMOCOv7 More Information | Similar Motifs Found	motif file (matrix)
5		1e-30	-7.054e+01	5.98%	1.65%	0bp (0bp)	E2F4_f1_ppm_HOCOMOCOv7 More Information | Similar Motifs Found	motif file (matrix)
6		1e-30	-7.032e+01	12.58%	5.61%	0bp (0bp)	CHR/Cell-Cycle-Exp/Homer More Information | Similar Motifs Found	motif file (matrix)
7		1e-30	-6.930e+01	12.00%	5.26%	0bp (0bp)	E2F/Hela-E2F1-ChIP-Seq/Homer More Information | Similar Motifs Found	motif file (matrix)

RIKEN
Omics Science Center
RIKEN OSC

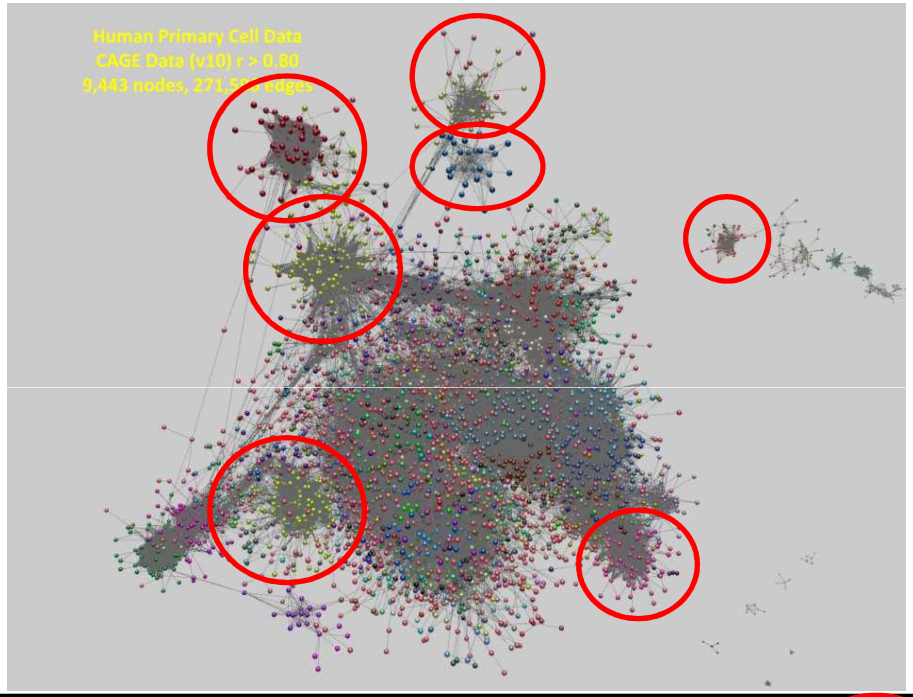
Michael Rehli + Christian Schmidl - HOMER

Other motif enrichments



- Housekeeper vs Lineage specific (R1 vs R3)
- High vs Low
  - High R3 vs low R3
  - High R1 vs low R1
- Co-expressed modules
  - Next slide

Co-expression approach example



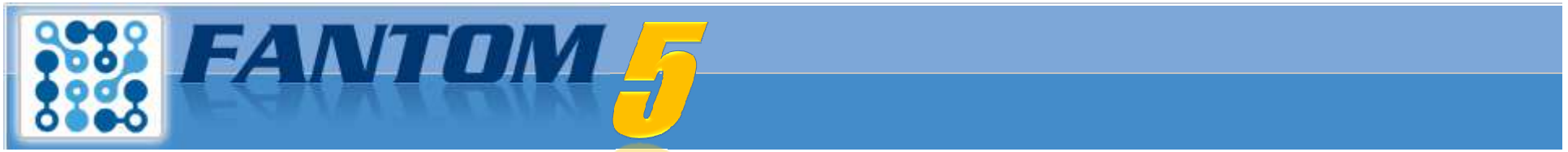
Clusters of co-regulated loci

Many different states considered

Module – motif over-representation
- Known PWMs and Ab-initio motifs
- TF co-expression

→ Motif set 1

→ Motif set 2



Phase1: tasks and questions

Gene regulation:

- How are tissue specific and broadly expressed genes encoded?
  - Tissue specific motifs? Tissue specific TFBS cluster combinations?
- How are highly expressed and weakly expressed genes encoded? (single molecule gives us this)
  - Copy number? TFBS clusters? Promoter shape/length?
- What factors are necessary to induce expression of a particular gene?
  - Insulin, Hemoglobin

Transcription factors:

- Agreed upon set of TFs – must be finalised in this meeting!!! – and curated mapping to CAGE cluster
- Which TFs can we predict the motif for?
- What is the tissue specificity of transcription factors?
  - How connected are sample specific as opposed to broadly expressed transcription factors? Are they leaves or hubs?

Motifs:

- A figure to present the motifs analyses? (Integration of multiple methods?Integrated figure with TF expression? Known motifs and ab-initio treated separately?)
- Motifs from co-expression modules?

Phase1: tasks and questions

Networks:

- How many unique combinations of transcription factors are expressed in human cells? What can we say about combinatorial regulation?
- Does promoter complexity make the network/gene expression more robust to perturbations?
- Can we say cell line networks are more unique than primary cell networks? (ie. each cell line is a unique entity)

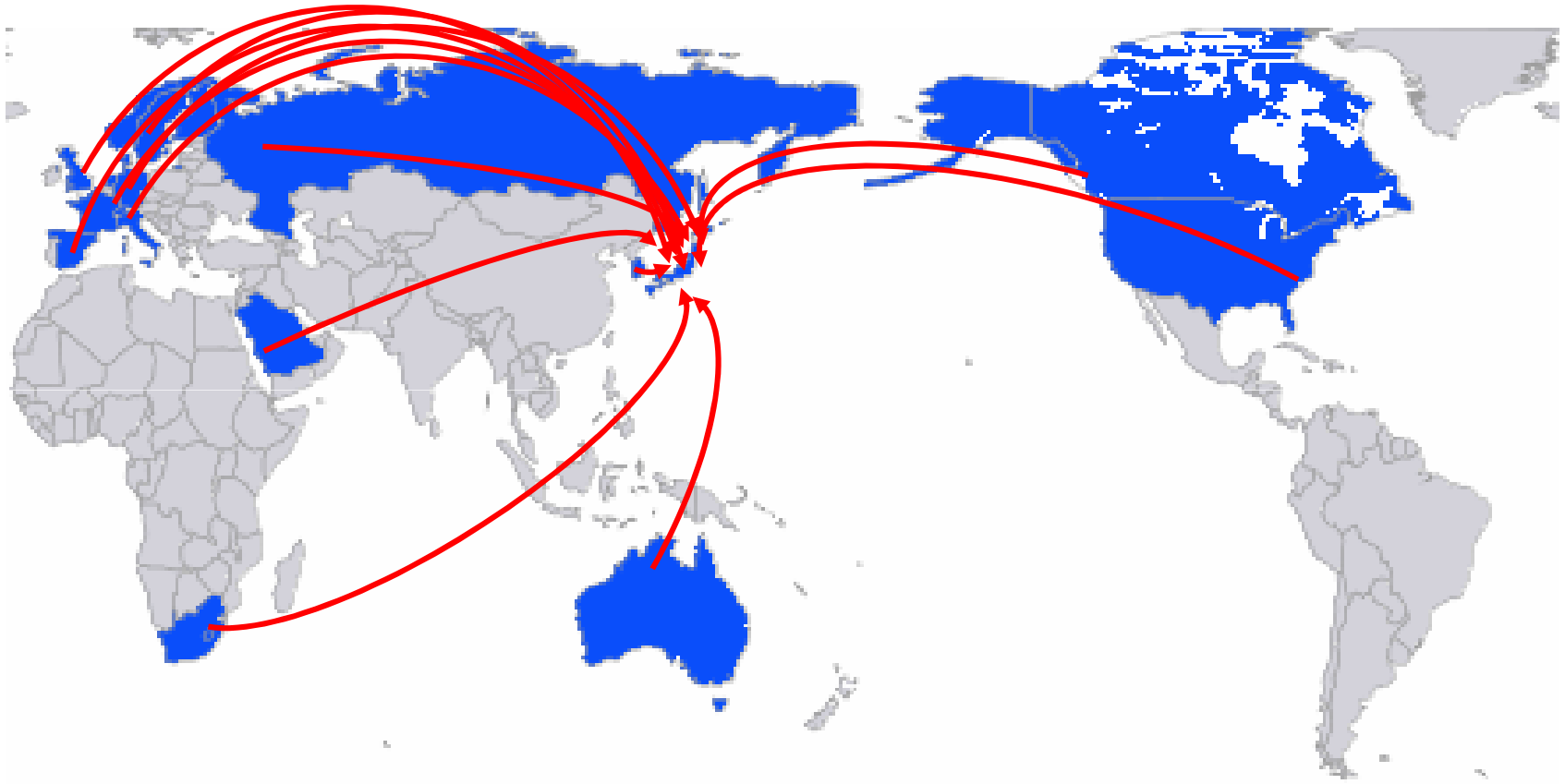
Cellular identity and evolution:

- Distributions of gene utilization in man (comparatively how many genes actually expressed in each state? Measure of transcript complexity – how diverse is the mix?)
- What can we say about multicellularity?
- Tissue specificity of genes evolved since multicellularity?
- Best figure to demonstrate tissue/cell specificity?

Other:

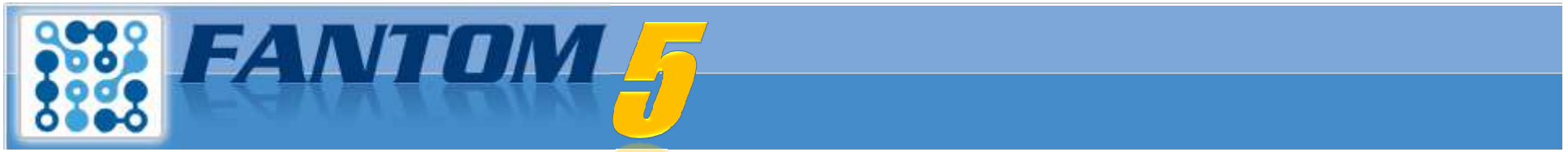
- A set of 'snippet' insights from satellites (Main paper will make a call out to the satellite)

FANTOM5 consortium



Acknowledging over 120 international collaborators: Sample providers,

Technology developers and Bioinformaticians from 19 countries



OSC acknowledgements

Informatics WP4 and 5
Timo Lassmann
Kawaji Hideya
Michiel de Hoon
Jessica Severin
Nicolas Bertin
Marina Lizio
Erik Arner
Max Burroughs

Library production
Masayoshi Itoh
Sayaka Sato
Miki Kojima
Sakai Mizuho
Mitsuyoshi Murata
Mutsumi Kanamori-Katayama

Ethics
Emi Ito
Makoto Hasegawa

Secretariat
Hiroto Atsui
Hanna Daub

Samples
Yuri Ishizu
Hiromi Nishiyori
Sachiko Ishikawa

WP9
Piero Carninci
Carsten Daub
Harukazu Suzuki
Jun Kawai
Yoshihide Hayashizaki

Funding:

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

Supporting Staff FANTOM Koyo Meeting

FANTOM5 Secretariat
Hiroto Atsui
Tsugumi Burroughs
Hanna Daub
Kayo Goto
Chinami Yuda
Kasumi Toyoda
Harumi Uruma

Infra Structure Support
Teruaki Kitakura
Takayuki Ishii
Atsushi Kondo
Nobuyuki Takeda
Akira Furukawa

Pr, Design, Photographing, Video
Emi Ito
Kazufumi Watanabe
Yoshie Hirayama

General Assitance
Mieko Adachi
Yumi Hara
Sanae Hitomi
Ayako Hosoya
Yayoi Karino
Machiko Kashiwagi
Yukiko Kurosawa
Junichi Masubuchi
Nobuko Miyajima
Hiroyuki Obata
My Persson
Tsutomu Saito
Nana Tamura
Mitsuko Tsuji
Noriko Ueda
Marina Lizio
Kazuhiro Kajiyama
Serkan Sahin
Takehiro Hashimoto

RIKEN
Omics Science Center

FANTOM 5

Invitation to final year PhD students

OSC is always looking for new postdocs.

JSPS fellowship applications have a high success rate

Japan is a very different experience

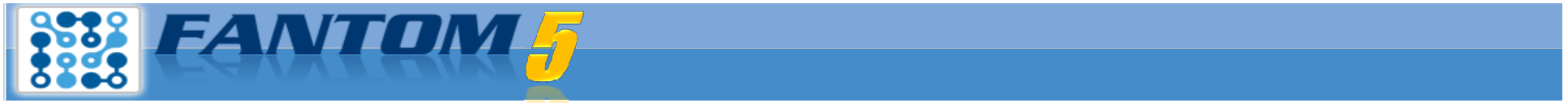
Next generation sequencing

Bioinformatics

Transcriptomics

Stem cell/cell lineage biology

Human Genetics *



Australia	*University of New South Wales*	Levon Khachigian
	University of Queensland	Christine Wells, Anthony Beckhouse
	Walter and Eliza Hall Institute	Mark Robinson, Gordon Smyth
	Western Australian Institute for Medical Research	Peter Klinken, Louise Winteringham
Canada	*McGill University*	Josée Dostie, Hisashi Miura
	University of British Columbia	Daniel Goldowitz, Thomas Jonghyun Ha
Denmark	*University of Copenhagen*	Albin Sandelin, Robin Andersson, Eivind Valen
Finland	*University of Helsinki*	Alessandro Bonetti
France	*University Pierre & Marie Curie*	Hugues Richard
Germany	*Charite Universitätsmedizin Berlin*	Magda Babina, Sven Guhl, Gundula Schulze-Tanzil
	University Hospital Regensburg	Christian Schmidl, Michael Rehli
Italy	*Dulbecco Telethon Institute*	Beatrice Bodega, Valerio Orlando
	Fondazione Bruno Kessler	Marco Chierici , Cesare Furlanello, Giuseppe Jurman, Marco Roncador
	International School for Advanced Studies (SISSA)	Stefano Gustinich, Silvia Zucchelli
	National Lab of Italian Consortium for Biotechnology	Silvano Piazza, Claudio Schneider, Roberto Verardo
Japan	*Database Center for Life Science*	Hidemasa Bono
	Keio University	Jun-ichi Furusawa, Shigeo Koyasu, Kazuyo Moro
	Kyushu University	Daisuke Sugiyama
	National Institute of Advanced Industrial Science and Technology (AIST)	Martin Frith , Junko Tsuji
	National Institute of Genetics	Takashi Gojobori, Kazuho Ikeo, Mitsuru Morimoto, Yasukazu Nakamura
	Ochanomizu Unviersity	Jun Sese, Aika Terada
	Ohu University	Mitsuhiro Ohshima
	Osaka University	Hideo Matsuda, Shigeto Seno, Yoichi Takenaka
	RIKEN, Advanced Science Institute	Mitsuko Hara, Soichi Kojima, Hideki Tatsukawa
	RIKEN, Bio Resource Center	Yukio Nakamura
	RIKEN, Bioinfomatics And Systems Engineering division	Takao Endo, Iida, Shuji Kawaguchi, Koro Nishikata, Tetsuro Toyoda
	RIKEN, Center for Developmental Biology	Hideki Enomoto, Guojun Sheng
	RIKEN, Research Center for Allergy and Immunology	Mitsuhiro Endo, Tomokatsu Ikawa, Hiroshi Kawamoto, Haruhiko Koseki, Takeshi Nagashima, Mariko Okada, Kaoru Takahashi , Noriko Yumoto
	Saitama Medical University	Masami Muramatsu, Yutaka Nakachi, Yasushi Okazaki
	The University of Tokyo	Toshio Kitamura, Kiyoshi Asai, Chieko Kai, Hiroki Sato, Takaaki Sugiyama, Misako Yoneda, Kenta Nakai, Sung-Joon Park
	Tohoku University	Hozumi Motohashi, Masayuki Yamamoto
	Tokyo Medical and Dental University	Soichi Ogishima, Hiroshi Tanaka
	Tokyo University of Pharm. and Life Sciences	Yuan Bo, Tadasuke Nozaki, Hiroo Toyoda, Yuta Yoshino
Korea	*Kyungbook National University*	Won Young Suk, Weonju Lee
Norway	*University of Bergen*	Vedran Franke, Vanja Haberle, Boris Lenhard
Russia	*Vavilov Institute of General Genetics*	Vsevolod Makeev, Yulia Medvedeva
Saubi Arabia	*King Abdullah University of Science and Technology*	Huoming Zhang, Mamoon Rashid, Ulf Schaefer, Sebastian Schmeier, Vladimir Bajic, Christian R. Voolstra, Carlo Vittorio Cannistraci, Timothy Ravasi, Arnab Pain
Spain	*Centre for Genomic Regulation (CRG)*	Sarah Djebali, Roderic Guigó
Sweden	*Karolinska Institutet*	Juha Kere, Peter Arner, Lukasz Huminiecki, Niklas Mejhert, Andreas Lennartsson
	Stockholm University	Lukas Käll
Switzerland	*Ecole Polytechnique Federale de Lausanne*	Bart Deplancke
	Eidgenössische Technische Hochschule Zürich	Michael Detmar
	University of Basel	Erik van Nimwegen, Piotr Balwierz, Mikhail Pachkov, Mihaela Zavolan
Netherlands	*Hubrecht Institute*	Hans Clevers
	Leiden University	Christine Mummery
	VU University	Peter Heutink, Patrizia Rizzu, Teunis Geijtenbeek
UK	*Imperial College London, MRC Clinical Sciences Center*	Ana Pombo
	The Roslin Institute	John Kenneth Baillie, Dario Beraldi, Geoffery Faulkner, Tom Freeman, David Hume, Christelle Robert, Kim Summers
	University of Bristol	Julian Gough
	University of Nottingham	Alan Knox
	Wellcome Trust Sanger Institute	Laurens Wilming
	Western General Hospital, MRC Human Genetics Unit	Alison Meynert, James Prendergast, Colin Semple, Martin Taylor
USA	*Cold Spring Harbor Laboratory*	Carrie Davis, Tom Gingeras
	Columbia University	Andrea Califano, Mariano Javier Alvarez, Yishai Shimoni
	Dana-Farber Cancer Institute	Jessica Mar, John Quackenbush
	Harvard School of Public Health	Gabriel Altschuler, Emmanuel Dimont, Winston Hide, Shannan Ho Sui, Oliver Marc Hofmann, Ramakrishna R Sompallae
	Massachusetts Eye and Ear Infirmary	Albert Edge, Judith Kempfle
	Massachusetts Institute of Technology	Arvind Govindarajan
	Rice University	Mary(Cindy) Farach-Carson
	The Jackson Laboratory	Richard M. Baldarelli, Terrence F. Meehan
	The Wistar Institute	Susan Zabierowski
	University of California	James Ben Brown
	University of Delaware	Swati Pradhan Bhatt
	University of Texas Health	Dianna Milewicz
	Wake Forest University	Steve Hodges
	Wayne State University	Leonard Lipovich