# network-based analyses of transcriptomics data: new approaches to define cell signatures from tissue data

Prof Tom Freeman
Systems Immunology Group,
The Roslin Institute,
University of Edinburgh



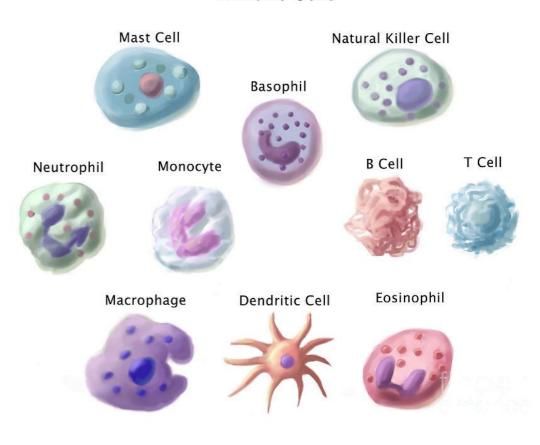






# Brain Heart Bone Lung Blood Muscle Liver **Pancreas** Kidney Egg Sperm

### Immune Cells

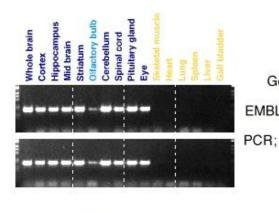


tissues

cells

- What does the expression pattern of a gene tell you about its likely function?
- What is the difference between the expression profile of tissues and cells, and how does this reflect their different functions?
- How is their expression profile changed during disease? disease classification, patient stratification, drug targeting

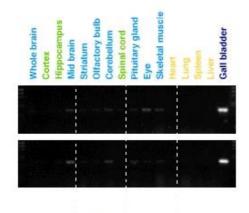
# mouse expression atlas



Gene/EST: Calbindin 28K Symbol: Calb1 EMBL Ac. No.: M23663

Primers: 95/42

Date: 8/2/96 MgCl<sub>2</sub>: 3.5 mM Temp: 55°C Cycles: x40ex

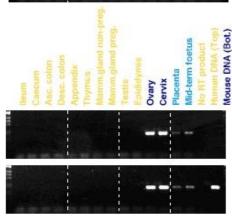


Gene/EST: Villin Symbol: Vil EMBL Ac. No.: M98454 Primers: 95/67 PCR;

Date: 11/1/96 MgCl<sub>2</sub>: 3.5 mM Temp:55°C Cycles: x40ex









### **Gene Expression Data**

Freeman TC, et al., Expression Mapping of Mouse Genes. MGI Direct Data Submission. 1998;

Filter expression by: Anatomical System 7 | Assay Type 7 | Detected? 7 | Theiler Stage 7 | Wild type? 7

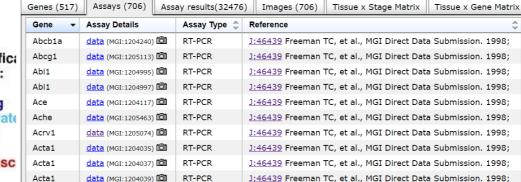
<< first < prev 1 2 3 next > last >> 100 ▼

Showing results(s) 1 - 100 of 706

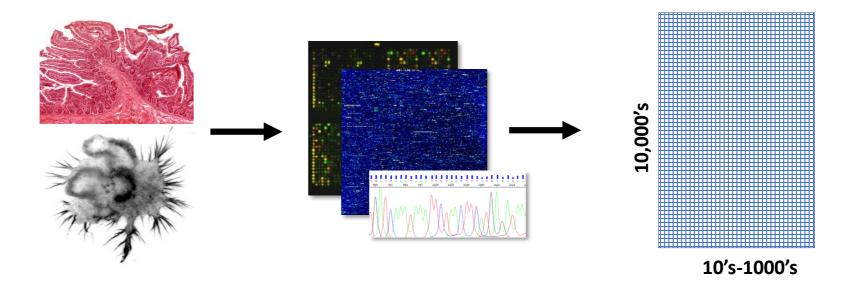
2

Amplifica signal:

Strong Moderate Trace +ve Disc



# transcriptomics analyses



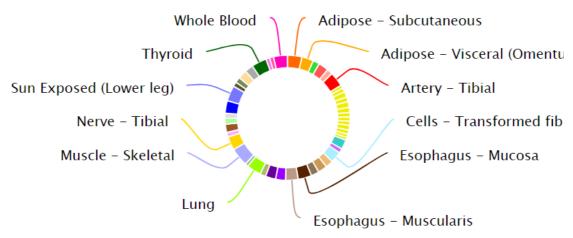




# resources of gene expression data

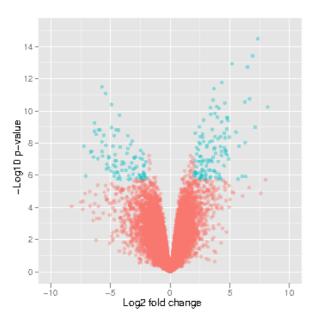


Total samples in all eQTL tissues: 10294

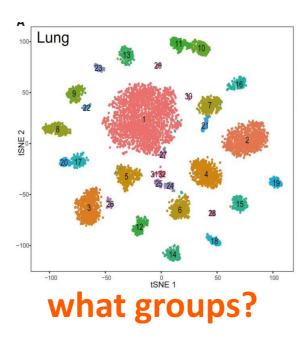


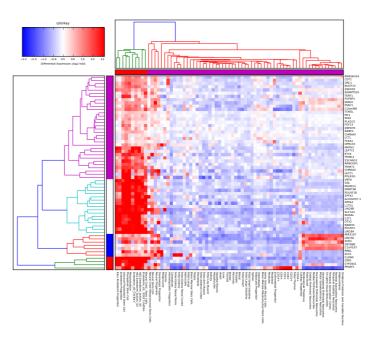






# what is different?





what is similar?



what does it mean?

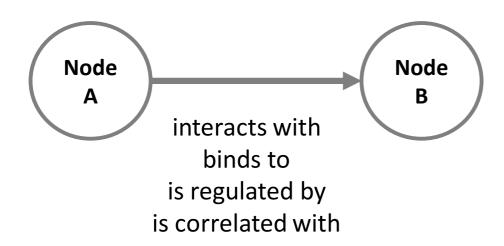




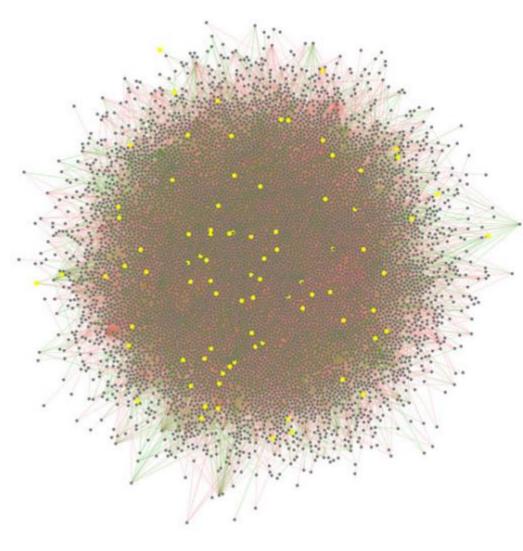
super visualisations

super computer

# biological networks





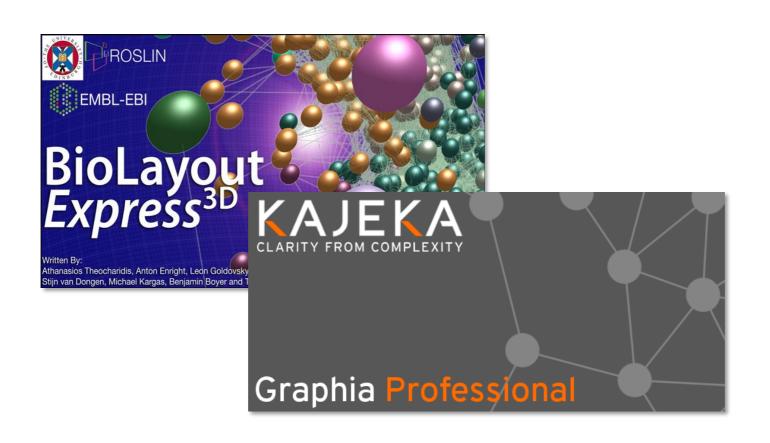


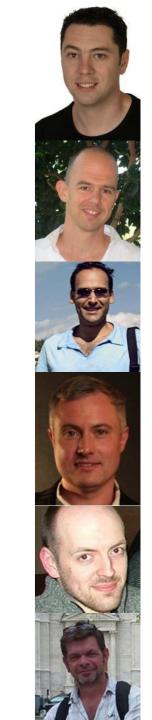
Prioritizing cancer-related genes with aberrant methylation based on a weighted protein-protein interaction network.

Liu H, Su J, Li J, Liu H, Lv J, Li B, Qiao H, Zhang Y - BMC Syst Biol (2011)

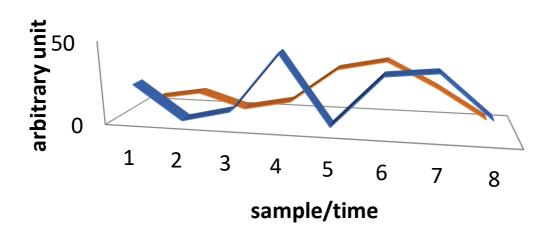
# network analysis tools

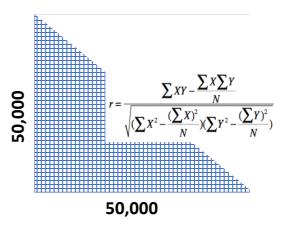
developed over 15 year period with by scientists in Cambridge and Edinburgh, now by Kajeka Ltd.



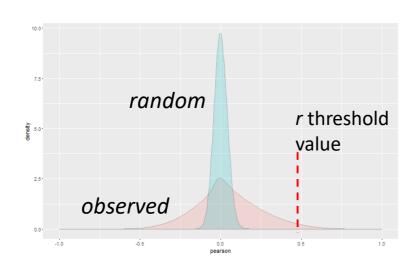


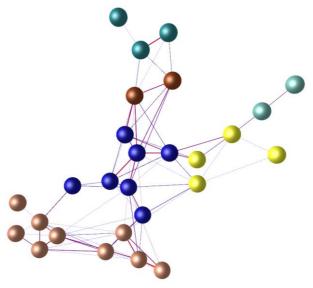
# plotting the statistically improbable: principles of correlation networks



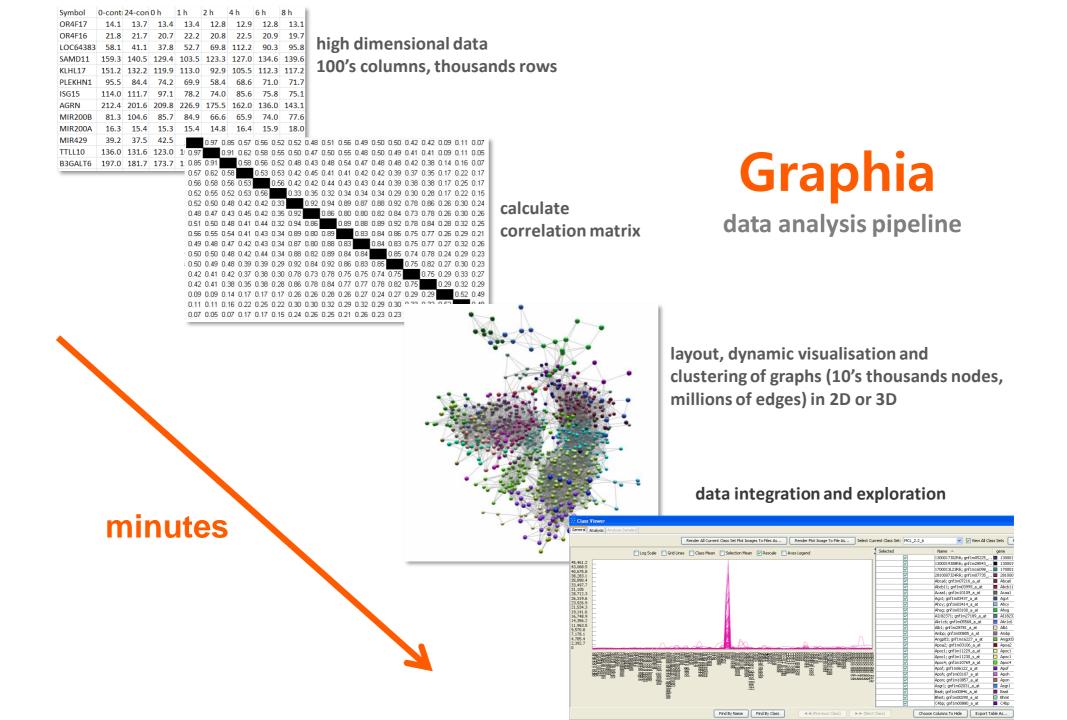


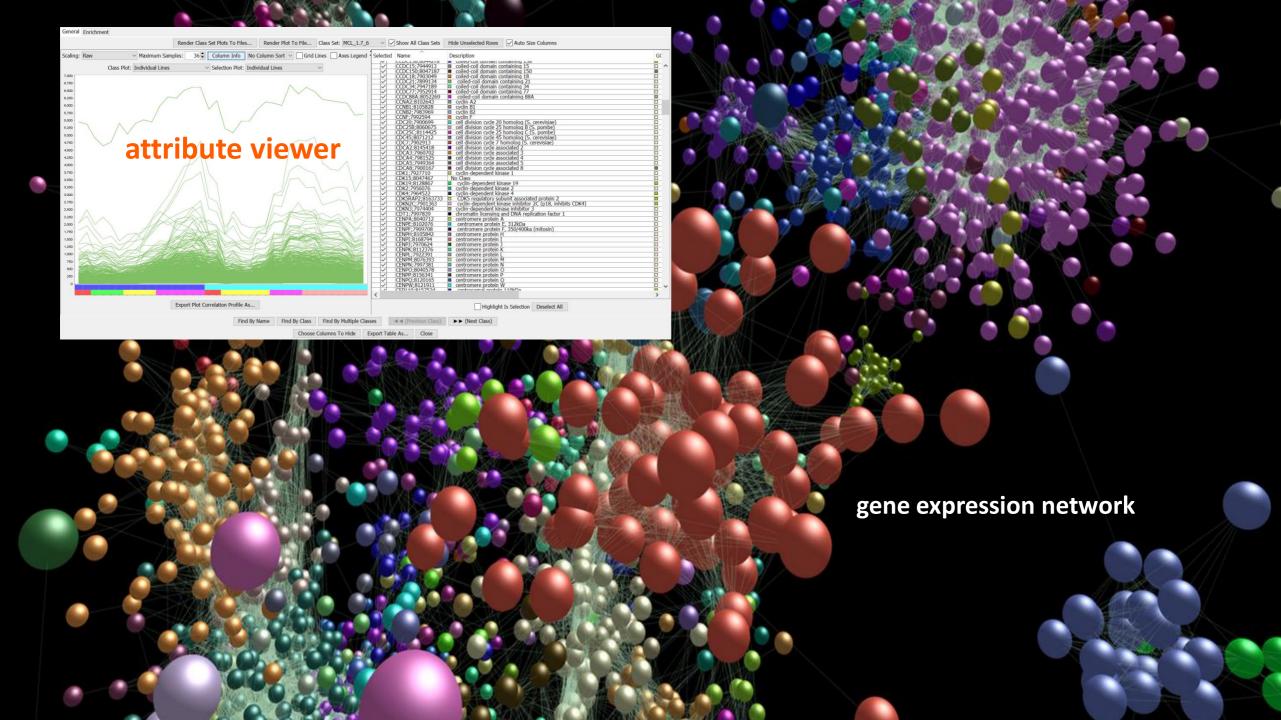
1.25 billion calculations





weighted, non-directional graph





#### Human Conjunctival Transcriptome Analysis Reveals the Prominence of Innate Defense in *Chlamydia trachomatis* Infection <sup>∀</sup>†

Angels Natividad, 1‡ Tom C. Freeman, 2 David Jeffries, 1,3 Matthew J. Burton, 1 David C. W. Mabey, 1 Robin L. Bailey, and Martin J. Holland 1,3\*

Department of Clinical Research, London School of Hygiene and Tropical Medicine, University of London, London WC1E 7HT, United Kingdom1; The Roslin Institute, R(D)SVS, University of Edinburgh, Roslin Biocentre, Midlothian EH25 9PS, United Kingdom2: and Medical Research Council Laboratories, Viral Diseases Programme, Fajara, Banjul, the Gambia

Received 3 August 2010/Returned for modification 17 August 2010/Accepted 24 August 2010



Contents lists available at ScienceDirect

#### Genomics



journal homepage: www.elsevier.com/locate/ygend

Functional clustering and lineage markers: Insights into cellular differentiation and gene function from large-scale microarray studies of purified primary cell populations

David A. Hume \*, Kim M. Summers, Sobia Raza, J. Kenneth Baillie, Thomas C. Freeman

The Roslin Institute, Roslin Biocentre, Roslin EH25 9PS, Midlothian, UK Royal (Dick) School of Veterinary Studies, The University of Edinburgh, Roslin Biocentre, Roslin EH25 9PS, Midlothian, UK

Derivation of marker gene signatures from human skin and their use in the interpretation of the transcriptional changes associated with dermatological disorders

Barbara B Shih O, Ajit | Nirmal, Denis | Headon, Arne N Akbar, Neil A Mabbott<sup>1\*</sup> and Tom C Freeman<sup>1\*</sup>

**Journal of Pathology** | Pathol 2017; 241: 600-613

Freeman et al. BMC Biology 2012, 10:90 http://www.biomedcentral.com/1741-7007/10/90



#### RESEARCH ARTICLE

**Open Access** 

#### A gene expression atlas of the domestic pig

Tom C Freeman<sup>1\*</sup>, Alasdair Ivens<sup>2,6</sup>, J Kenneth Baillie<sup>1</sup>, Dario Beraldi<sup>1,7</sup>, Mark W Barnett<sup>1</sup>, David Dorward<sup>1</sup>, Alison Downing<sup>1</sup>, Lynsey Fairbairn<sup>1</sup>, Ronan Kapetanovic<sup>1</sup>, Sobia Raza<sup>1</sup>, Andru Tomoiu<sup>1</sup>, Ramiro Alberio<sup>3</sup>, Chunlei Wu<sup>4</sup>, Andrew I Su<sup>4</sup>, Kim M Summers<sup>1</sup>, Christopher K Tuggle<sup>5</sup>, Alan L Archibald<sup>1\*</sup> and David A Hume<sup>1\*</sup>

#### RESEARCH ARTICLE



### A high resolution atlas of gene expression in the domestic sheep (*Ovis aries*)

Emily L. Clark \*\*, Stephen J. Bush \*\*, Mary E. B. McCulloch \*, Iseabail L. Farquhar \*\*, Rachel Young<sup>1</sup>, Lucas Lefevre<sup>1</sup>, Clare Pridans<sup>1</sup>, Hiu G. Tsang<sup>1</sup>, Chunlei Wu<sup>2</sup>, Cyrus Afrasiabi<sup>2</sup>, Mick Watson<sup>1</sup>, C. Bruce Whitelaw<sup>1</sup>, Tom C. Freeman<sup>1</sup>, Kim M. Summers<sup>1,3</sup>, Alan L. Archibald<sup>1</sup>\*, David A. Hume<sup>1,3</sup>\*





journal homepage: www.elsevier.de/imbio



Meta-analysis of lineage-specific gene expression signatures in mouse leukocyte populations

Neil A. Mabbott\*, J. Kenneth Baillie, David A. Hume\*,1, Tom C. Freeman 1

The Roslin Institute and Royal (Dick) School of Veterinary Studies, The University of Edinburgh, Roslin Biocentre, Roslin EH25 9PS, Midlothian, UK

Mabbott et al. BMC Genomics 2013, 14:632 http://www.biomedcentral.com/1471-2164/14/632



#### RESEARCH ARTICLE

**Open Access** 

An expression atlas of human primary cells: inference of gene function from coexpression networks

Neil A Mabbott\*†, J Kenneth Baillie, Helen Brown, Tom C Freeman\*† and David A Hume\*†

Doig et al. BMC Genomics 2013. 14:469 http://www.biomedcentral.com/1471-2164/14/469



#### **RESEARCH ARTICLE**

Open Access

Coexpression analysis of large cancer datasets provides insight into the cellular phenotypes of the tumour microenvironment

Tamasin N Doig<sup>1,3</sup>, David A Hume<sup>3</sup>, Thanasis Theocharidis<sup>3</sup>, John R Goodlad<sup>2</sup>, Christopher D Gregory<sup>1</sup> and Tom C Freeman<sup>3\*</sup>

doi:10.1038/nature13182

### A promoter-level mammalian expression atlas

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



over 150 publications using our software

# advantages of network approaches

- data agnostic
- fast
- scalable big data sets
- visualisation of 1,000,000s of data points
- powerful pattern finding ideal for complex data
- visually engaging
- hypothesis free explorative



#### **RESEARCH ARTICLE**

**Open Access** 

An expression atlas of human primary cells: inference of gene function from coexpression networks

Neil A Mabbott\*†, J Kenneth Baillie, Helen Brown, Tom C Freeman\*† and David A Hume\*†

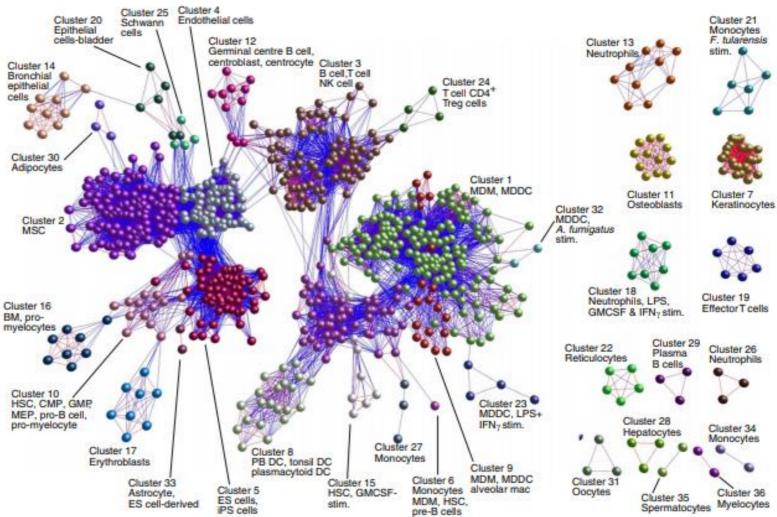


Figure 2 Clustering of samples based upon their gene global expression profiles. A Pearson correlation matrix was prepared comparing

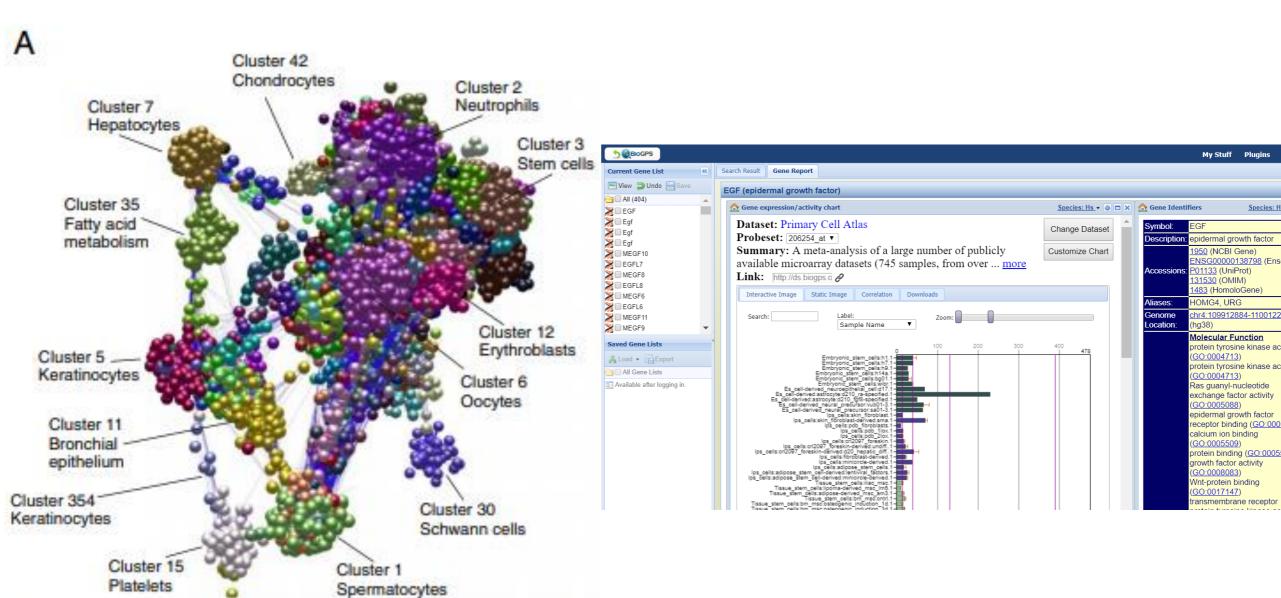


Figure 3 Network analysis of human primary cell transcriptomics data. (/

**Network visualisation and** Epithelia - rumen - tonsil clustering of the sheep gene **expression atlas** Epithelia - small intestine - liver Pituitary Gland Ribosomal Epithelia - large intestine Brain T and B Cell House Keeping (2) PLOS GENETICS House Keeping (1) \* Macrophages ardiac Muscle Cell Cycle Testes Phosphorylation Placenta Stomach - intestine Kidney Medulla Omentum

Clark EL et al. (2017) A high resolution atlas of gene expression in the domestic sheep (Ovis aries). *PLOS Genetics* **13**(9): e1006997. https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1006997

### data deconvolution

computational approaches that can identify specific signals within heterogeneous data signals

in our case identify transcriptional **signatures** of cells and pathways from mixed cell populations, i.e. tissue samples

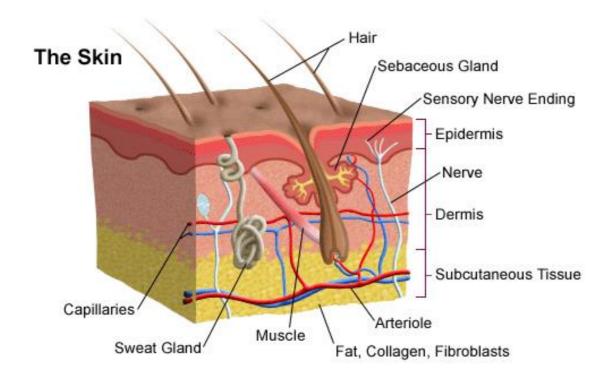
where a **signature** is a module (cluster) of genes that are robustly coexpressed because they are specifically transcribed by a given cell type or coregulated as they encode proteins that are part of the same pathway or process

variation in clinical samples



every sample is a little different

## an expression atlas of human skin









Journal of Pathology J Pathol 2017; 241: 600–613 Published online 24 February 2017 in Wiley Online Library (wileyonlinelibrary.com) 001: 10.1002/path.4864

ORIGINAL PAPER



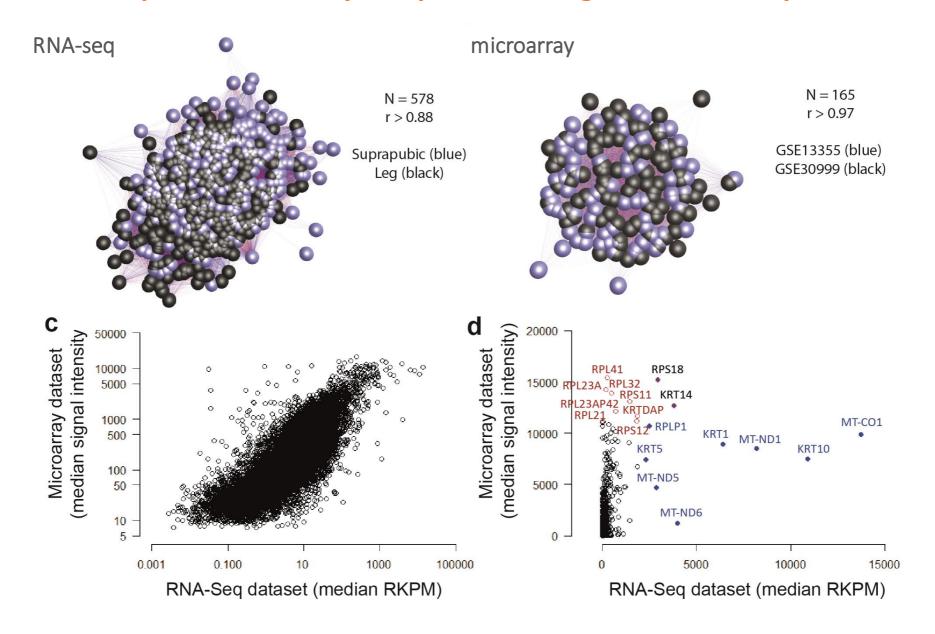
Derivation of marker gene signatures from human skin and their use in the interpretation of the transcriptional changes associated with dermatological disorders

Barbara B Shih<sup>1</sup>, Ajit J Nirmal<sup>1</sup>, Denis J Headon<sup>1</sup>, Arne N Akbar<sup>2</sup>, Neil A Mabbott<sup>1</sup> and Tom C Freeman<sup>1</sup>

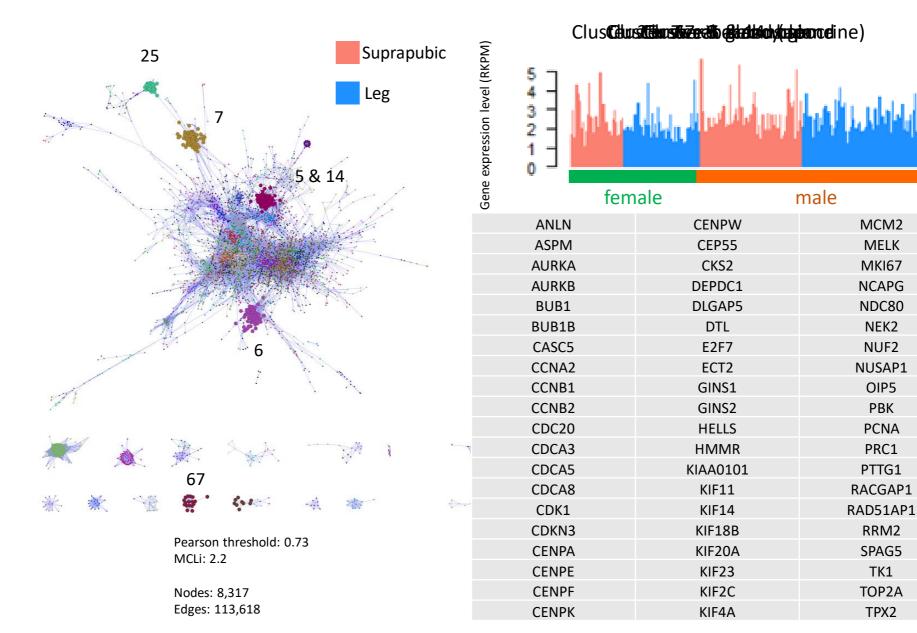
<sup>1</sup> The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian, Edinburgh, EH25 9RG, UK

Division of Infection and Immunity, University College London, 90 Gower Street, London, WCTE 6BT, UK

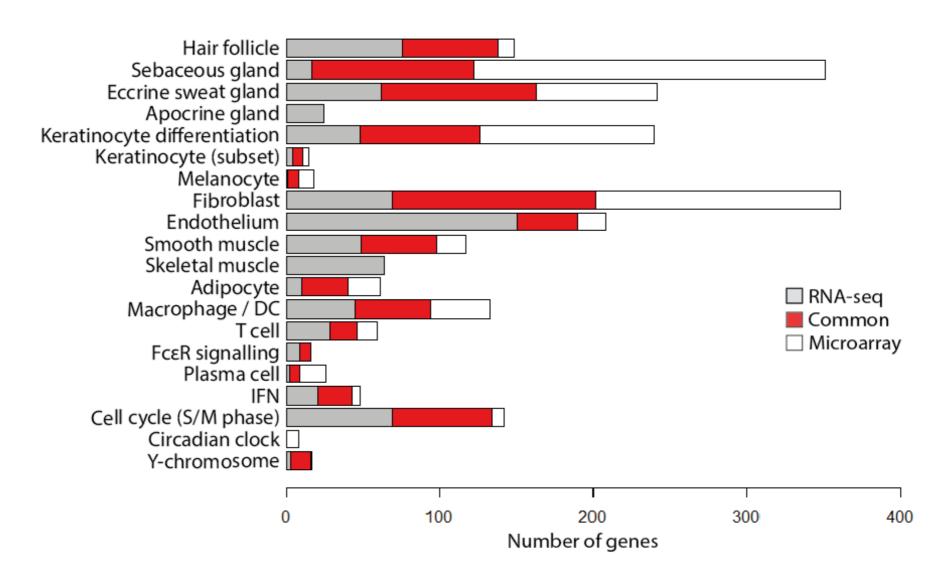
### RNA-seq and microarray sample clustering and data comparison



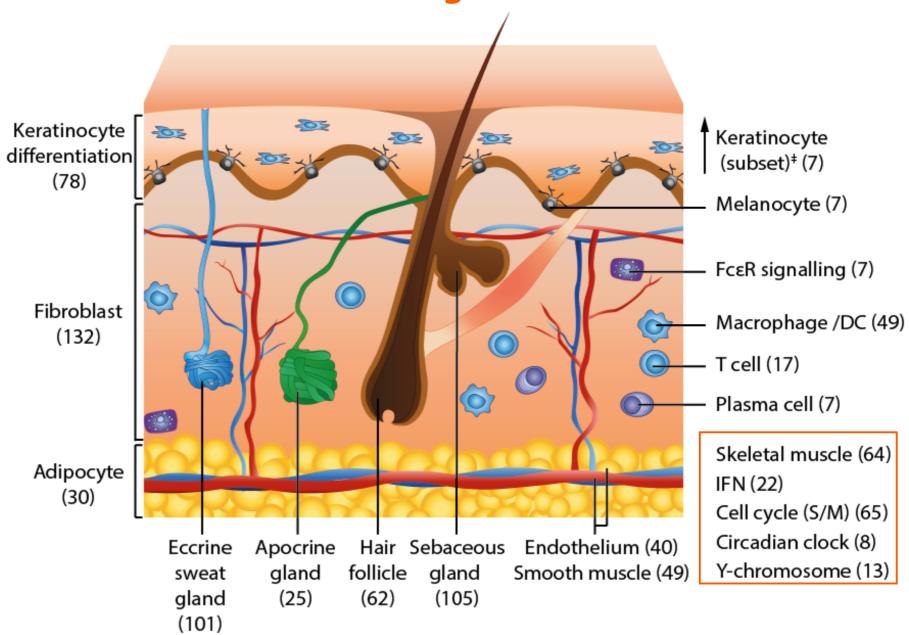
# **GTEx (RNA-seq) dataset**

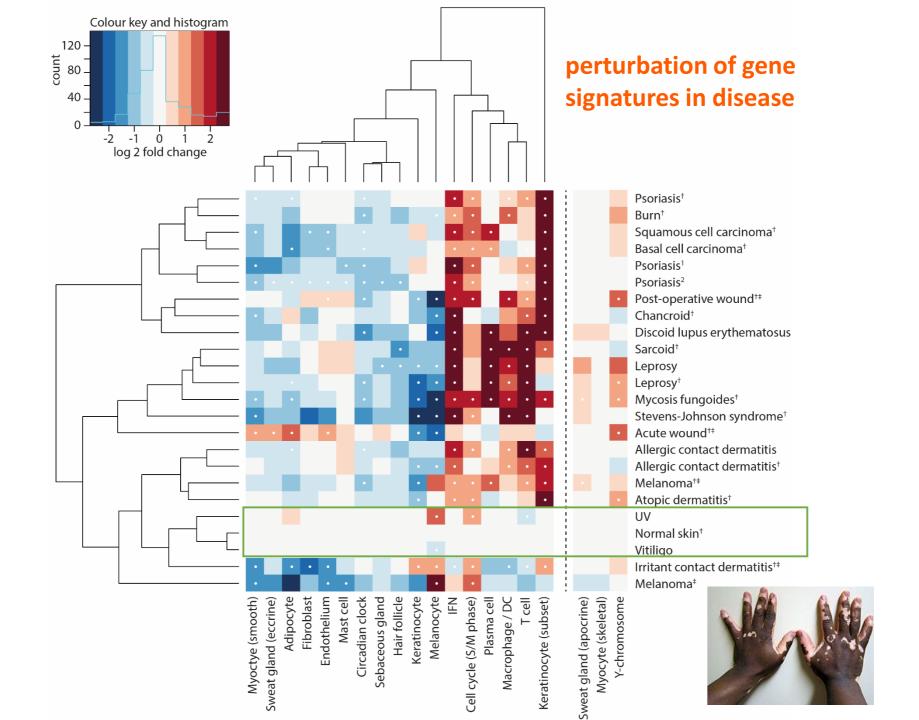


# overlap between datasets



SkinSig



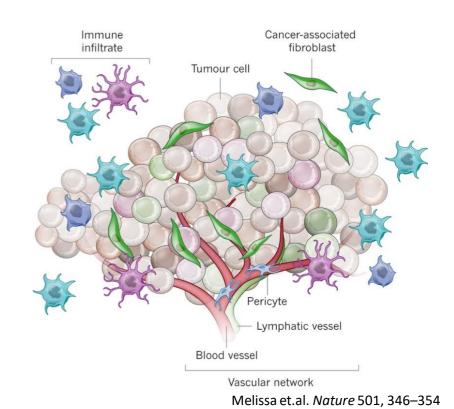


### immune microenvironment and cancer

growing interest in targeting the immune system in the treatment of cancer – the so called field of *immuno-oncology* 

numerous immune targets/strategies being explored as therapeutic targets, e.g. PD1, CTLA4, TLRs, CCRs etc

how do we analyse the immune microenvironment of tumours and relationship to outcome or treatment?



Research Article

Cancer Immunology Research

# Immune Cell Gene Signatures for Profiling the Microenvironment of Solid Tumors 19

Ajit J. Nirmal<sup>1</sup>, Tim Regan<sup>1</sup>, Barbara B. Shih<sup>1</sup>, David A. Hume<sup>1,3</sup>, Andrew H. Sims<sup>2</sup>, and Tom C. Freeman<sup>1</sup>



#### Aim

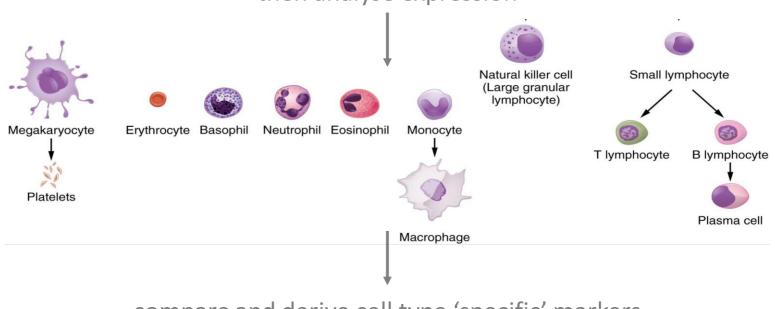
to derive robust immune gene signatures, that can be used to identify the presence and abundance of immune cells in a tissue (tumour) micro-environment from transcriptomics data



Ajit Johnson Nirmal

### derivation of markers

sort cells (usually from blood), sometimes cultured, then analyse expression

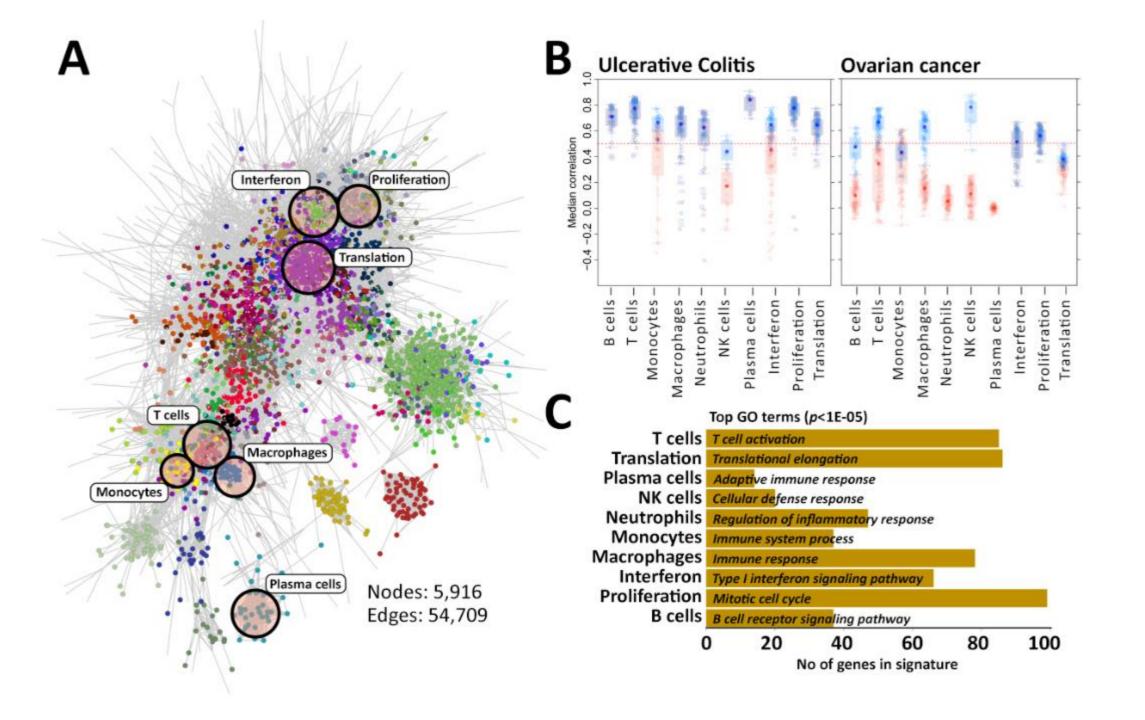


compare and derive cell type 'specific' markers

plugin into algorithm

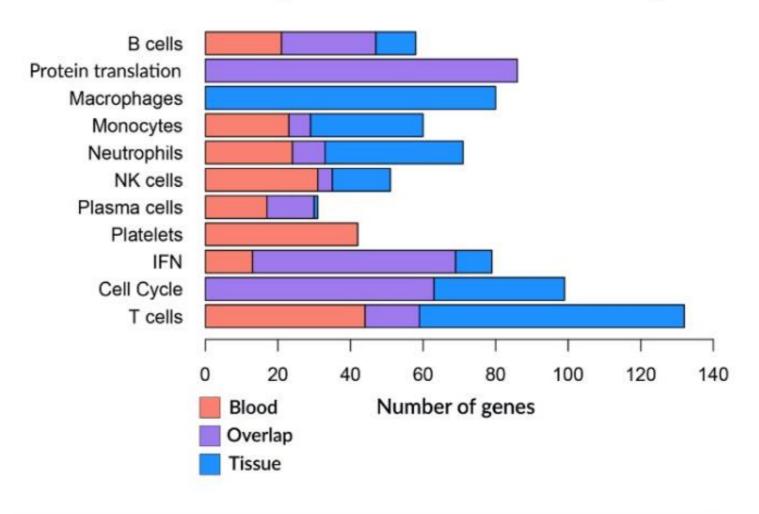
# immune cell types and subtypes defined by published gene signatures

Study	Author, year (PMID)	Cell gene signatures (defined subtypes)	Unique genes
1	Becht <i>et al.</i> (27765066)	T (3), B, NK, Mono, DC's, Neutrophils, Other (2)	321
2	Angelova <i>et al.</i> (25853550)	T (14), B (3), DC's (4), NK (4), Eosinophil, Macrophages, Mast cells, Mono, Neutrophils, MDSC	812
3	Abbas <i>et al.</i> 2005 (19568420)	T (2), B (2), Plasma cells, NK (2), Monocytes (2), DC's (2), Neutrophils	100
4	Watkins <i>et al.</i> (19228925)	T (2), Mono, B, NK, Neutrophils, Erythroblast, Megakaryocyte	1,847
5	Bindea <i>et al.</i> (24138885)	B (34), T (11), NK (3), DC's (3), Eosinophils, Macrophages, Mast cells, Neutrophils, Others (4)	582
6	Abbas <i>et al.</i> 2009 (15789058)	B, DC's, Monocyte, Neutrophil, NK, T, Other(3)	1,757
7	Newman <i>et al.</i> (25822800)	B (2), Plasma cells, T (7), NK (2), Monocytes, Macrophages (3), DC's (2), Mast cells (2), Eosinophils, Neutrophils	547



separate signatures for different applications  $ImSig_{blood}$  and  $ImSig_{tissue}$ 

### Overlap between blood and tissue ImSig



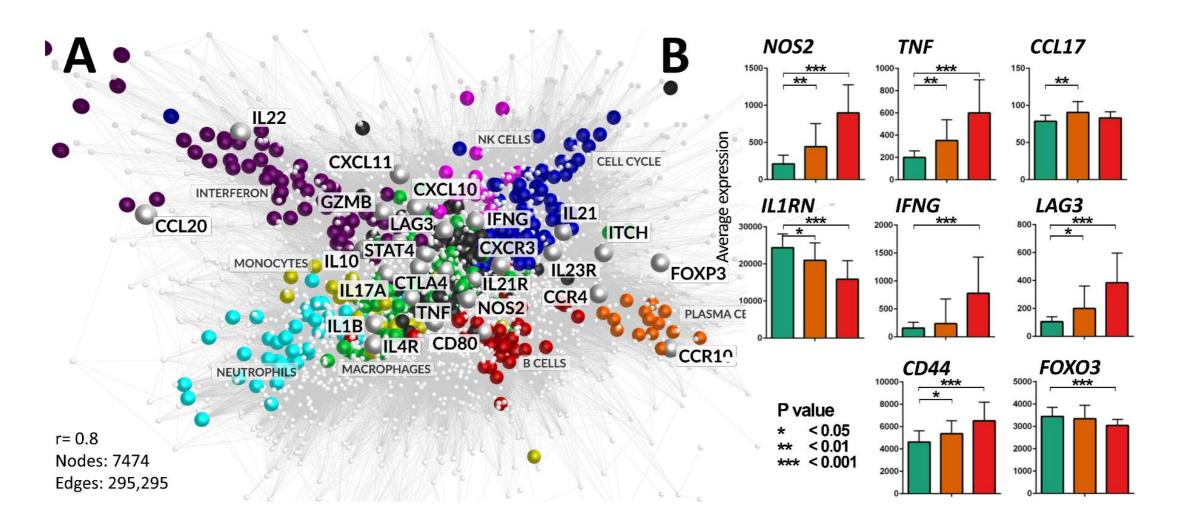
#### Monocytes Plasma cells *ImSig*<sub>tissue</sub> 22000 13500 2000 1000 nterferon Neutrophils Cell cycle Interferon B cells 4750 1500 Monocytes B cells Macrophages NK cells T cells **Platelets** Cell cycle Neutrophils 2800 1200 **Platelets** 1600 Plasma cells 2500 Nodes: 561 • (Translation 250 Edges: 32,624 NK cells r = 0.7Macrophages Translation 21000 3500 Control ■ Trachoma Patients (C. trachomatis -ve) 12000 2000 Trachoma Patients (C. trachomatis +ve) 1000 250

INFECTION AND IMMUNITY, Nov. 2010, p. 4895–4911 0019-9567/10/\$12.00 doi:10.1128/IAI.00844-10 Copyright © 2010, American Society for Microbiology. All Rights Reserved. Vol. 78, No. 11

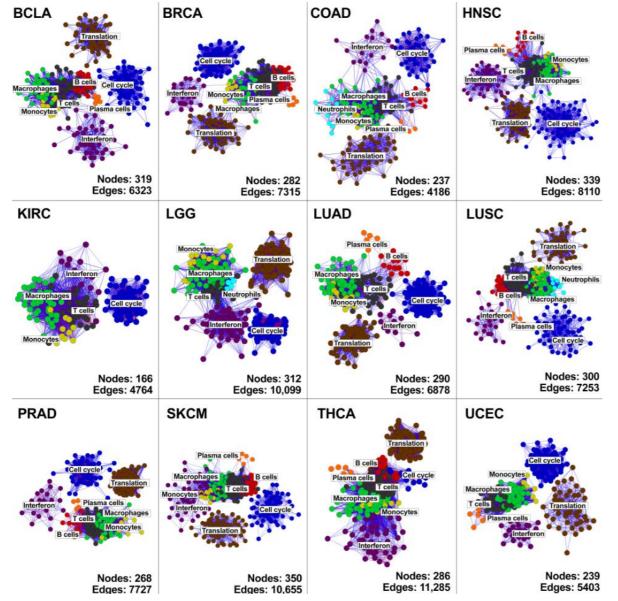
### Human Conjunctival Transcriptome Analysis Reveals the Prominence of Innate Defense in *Chlamydia trachomatis* Infection <sup>∇</sup>†

Angels Natividad, † Tom C. Freeman, David Jeffries, Matthew J. Burton, David C. W. Mabey, Robin L. Bailey, and Martin J. Holland 1,3\*

### what genes are coexpressed with signatures (subtype analysis)



### deconvolution of TCGA cancer datasets



#### **Cancer types (sample number)**

BCLA- Bladder Urothelial Carcinoma (408)

BRCA- Breast invasive carcinoma (1092)

COAD- Colon adenocarcinoma (456)

HNSC- Head and Neck squamous cell

carcinoma (501)

KIRC- Kidney renal clear cell carcinoma (530)

LGG- Brain Lower Grade Glioma (512)

LUAD- Lung adenocarcinoma (515)

LUSC- Lung squamous cell carcinoma (501)

PRAD- Prostate adenocarcinoma (496)

SKCM- Skin Cutaneous Melanoma (468)

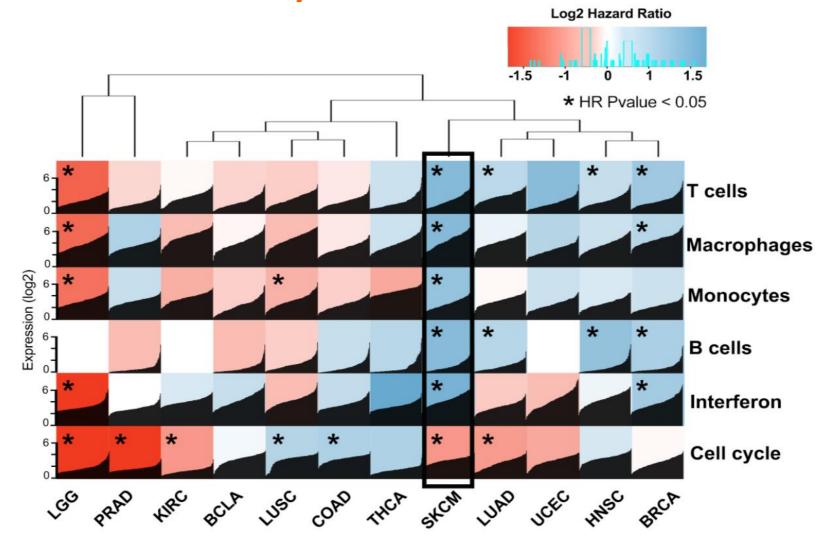
THCA- Thyroid carcinoma (502)

**UCEC- Uterine Corpus Endometrial** 

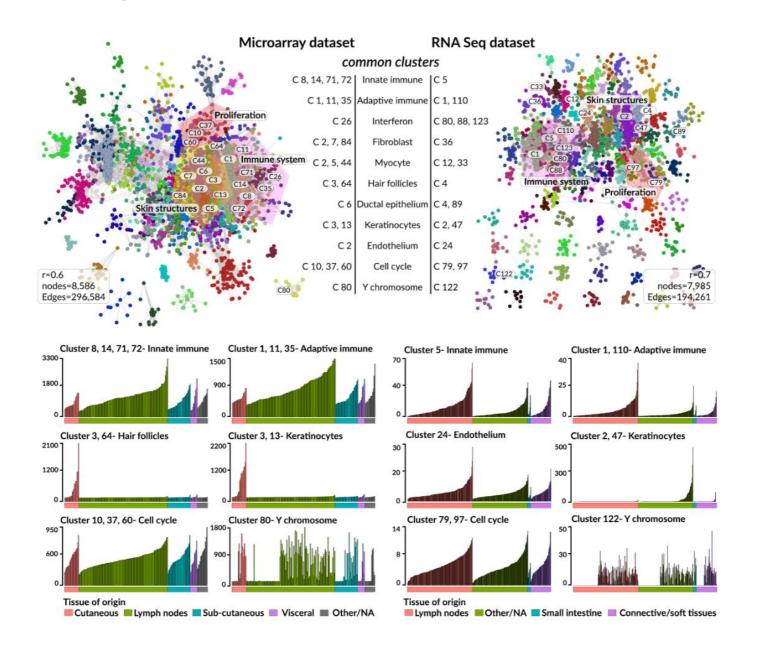
Carcinoma (555)



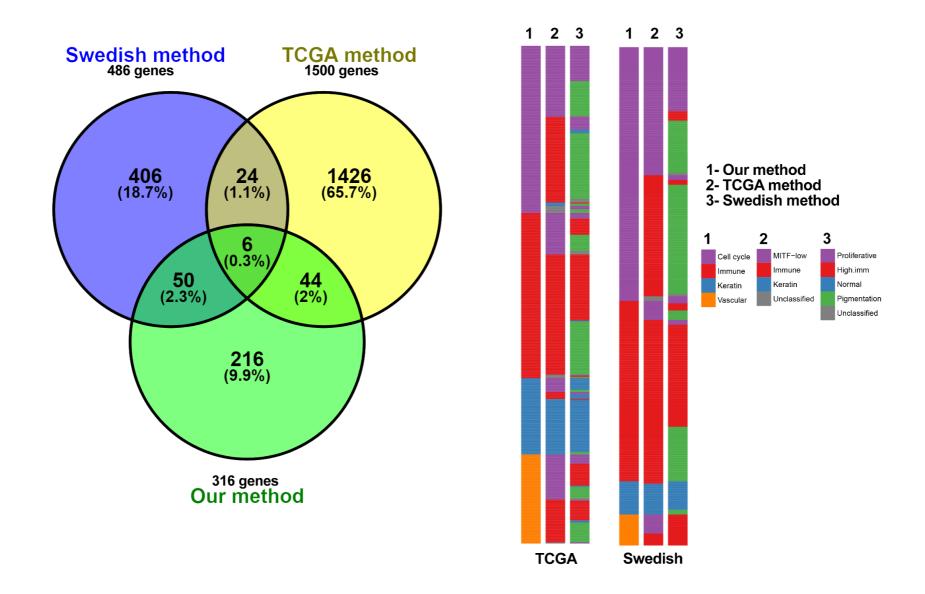
## pan-cancer hazard ratio analysis



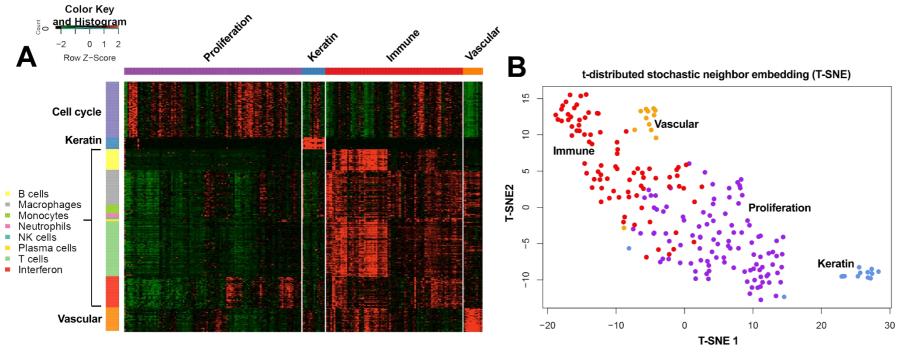
### network analysis of melanoma datasets



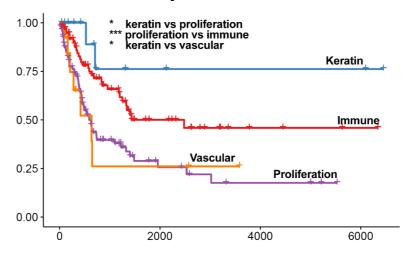
### comparison of melanoma groupings



## grouping of melanoma patients



### Survival analysis



### summary

- gene correlation network analysis approach supports exploration of tissue data allowing the derivation of cell and pathway-specific gene signatures
- knowing what constitutes normal, tells you a lot about disease
- still poor understanding of how immune microenvironment varies between cancers what does a 'good' microenvironment look like?
- what immune profile is most likely to respond to immunotherapy? Can we predict responders to a given treatment?





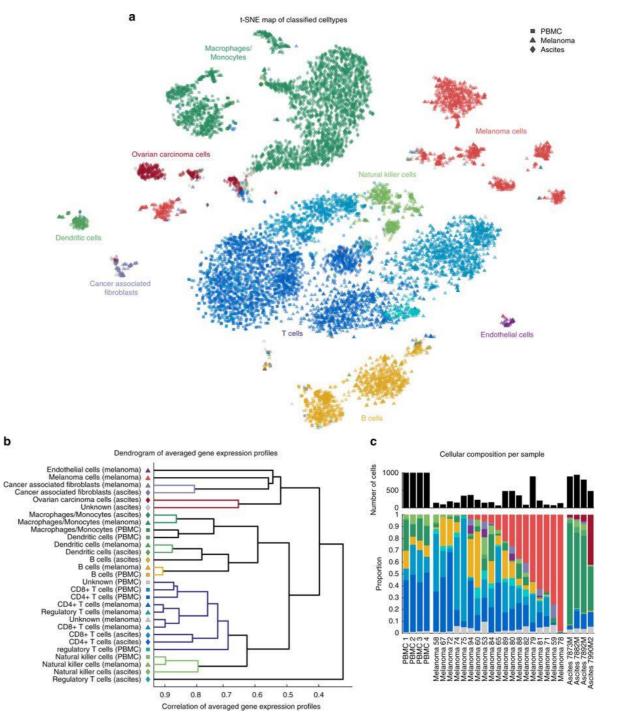
#### **ARTICLE**

DOI: 10.1038/s41467-017-02289-3

OPEN

# Estimation of immune cell content in tumour tissue using single-cell RNA-seq data

Max Schelker<sup>1,2</sup>, Sonia Feau<sup>1</sup>, Jinyan Du<sup>1</sup>, Nav Ranu<sup>1</sup>, Edda Klipp<sup>1</sup>, Gavin MacBeath<sup>1</sup>, Birgit Schoeberl<sup>1</sup> & Andreas Raue<sup>1</sup>

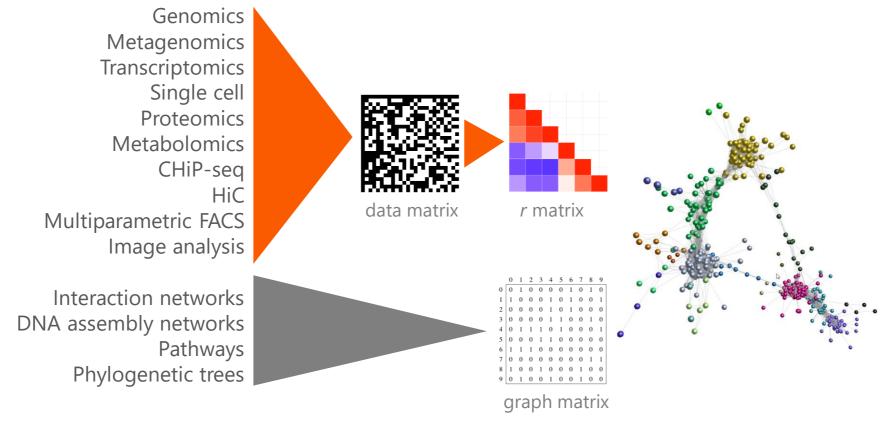




- designed and built from scratch
- user friendly
- bigger graphs, better rendering
- dynamic layout, input and filtering
- improved graph analytics, attribute handling, visualisation
- customisable for specific data types (plugin enabled)
- potential for direct connection to cloud resources







advanced network analysis solutions for the biological sciences

