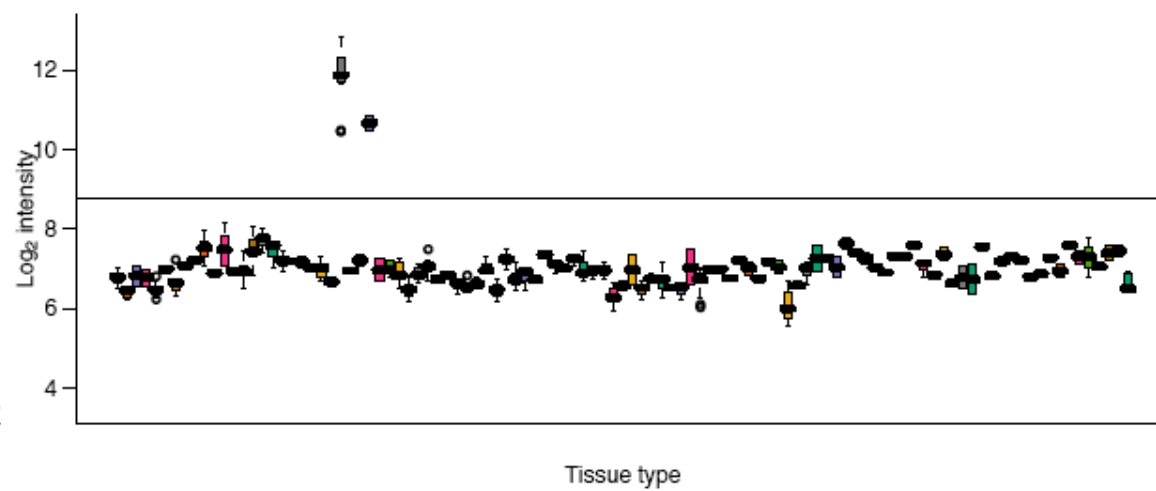
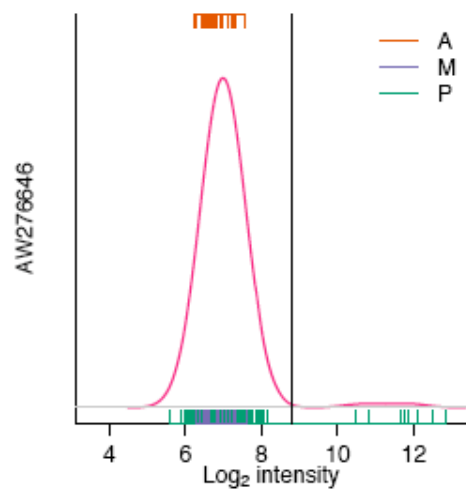
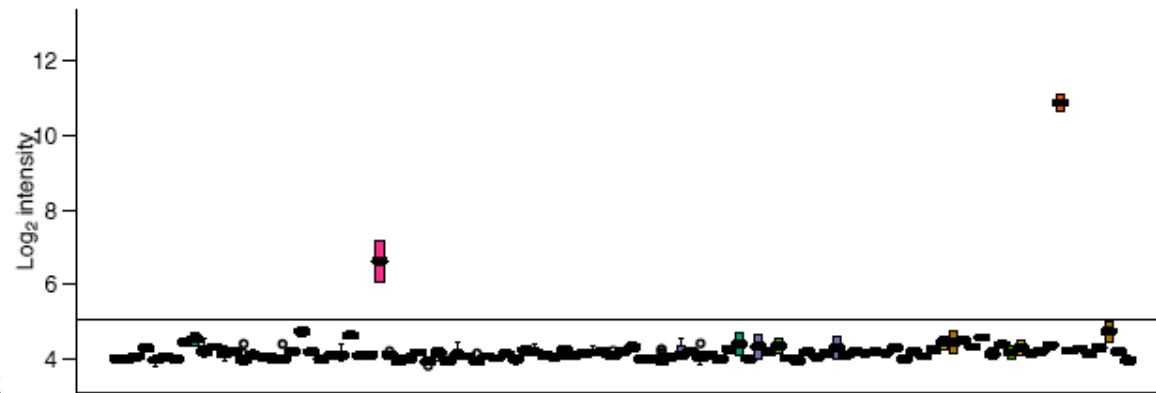
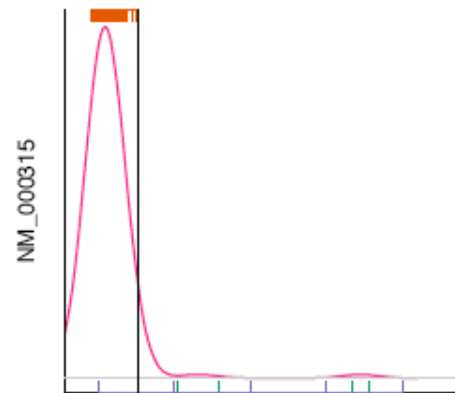


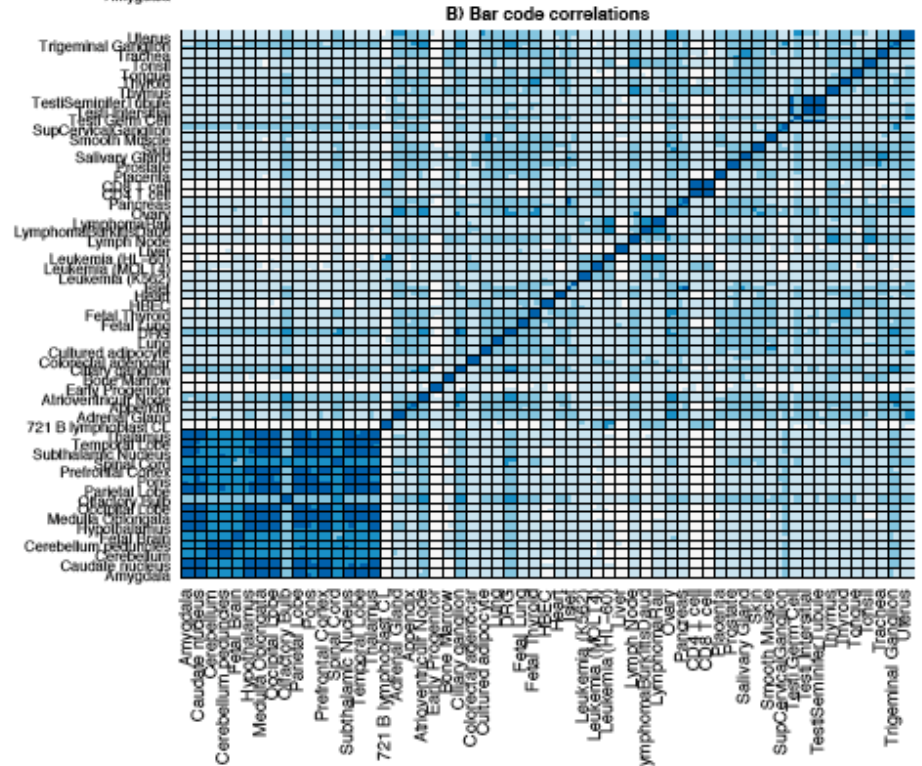
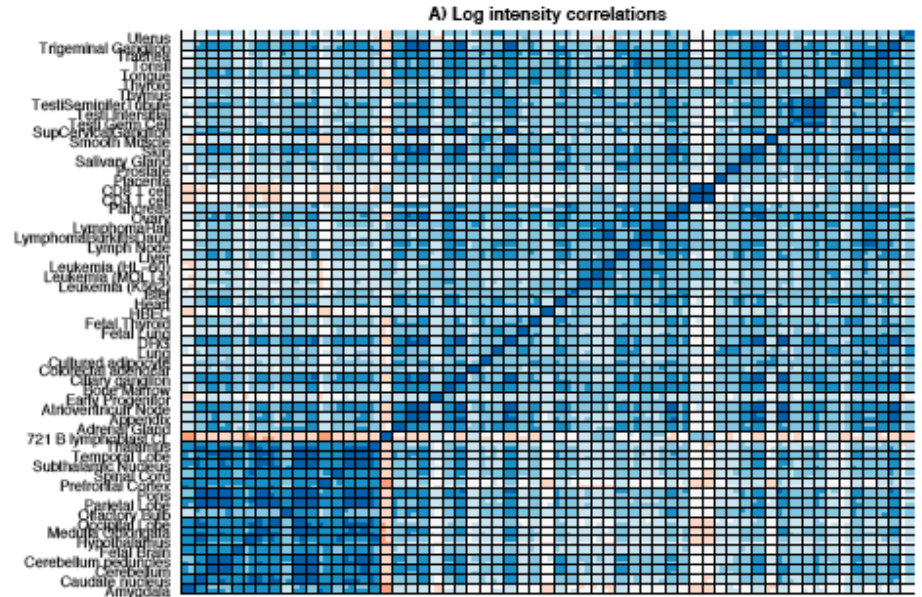
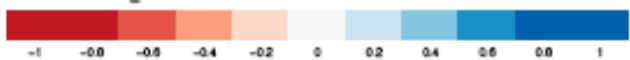
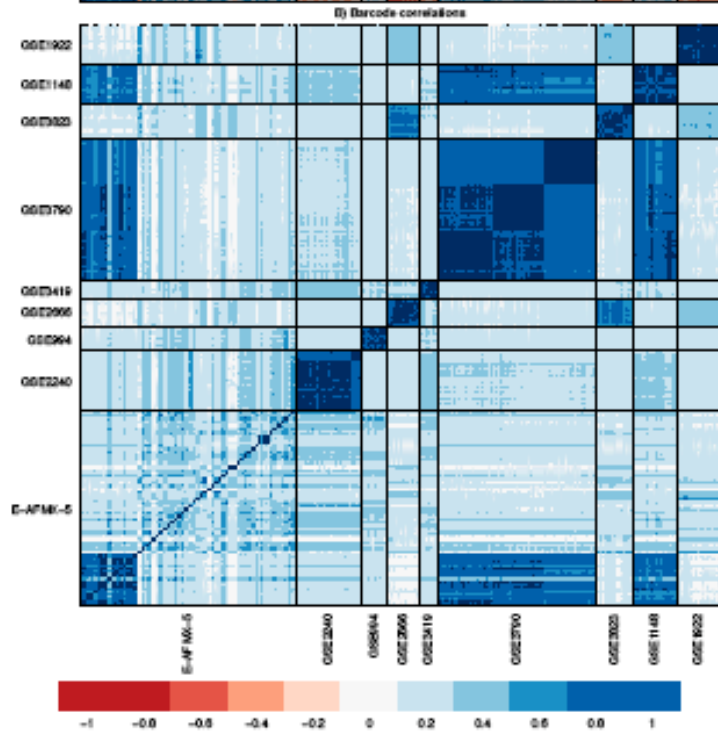
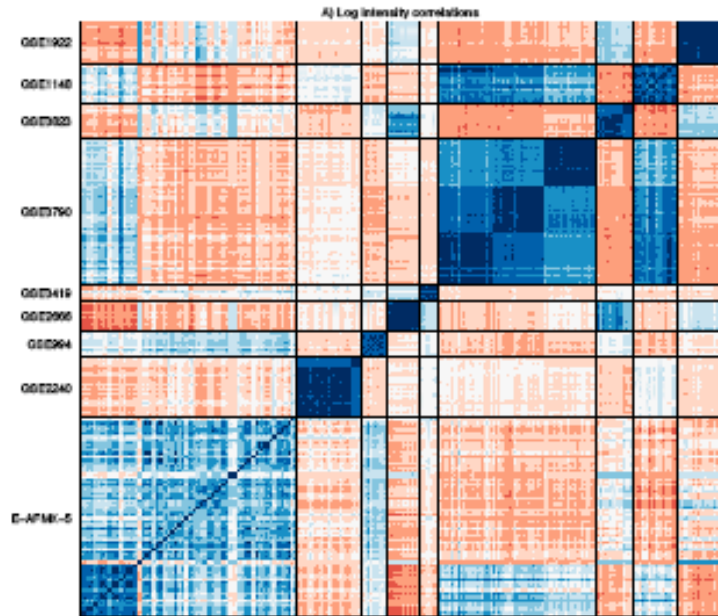
A Gene Expression Barcode for Microarray Data

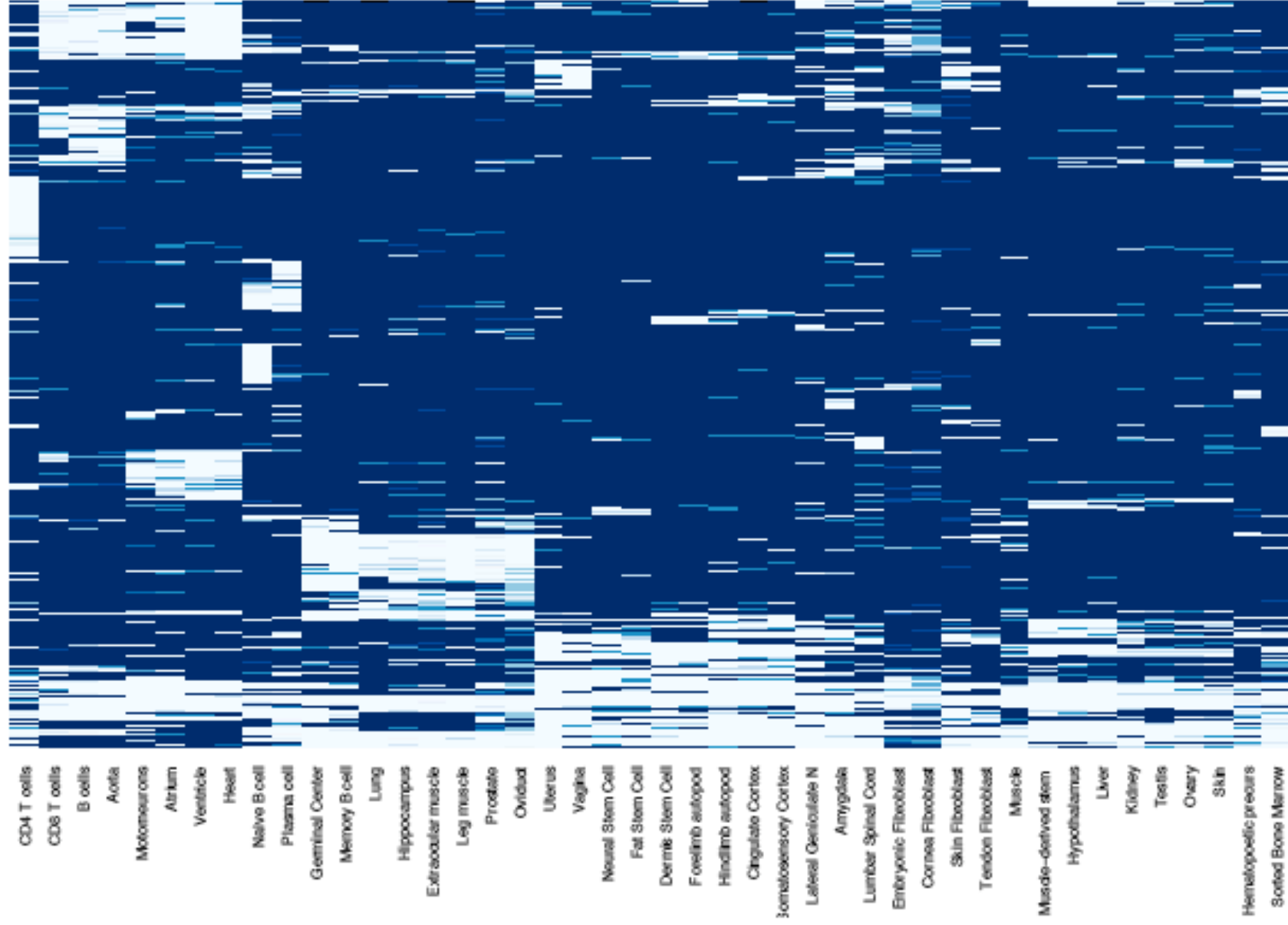
Rafael A Irizarry

- Zilliox MJ and Irizarry RA (2007) *Nature Methods* 4(11):911-913.
- <http://www.rafalab.org/barcode>

Barcode







Sample s	Comparison Type	PAM (% correct)	Bar code (% correct)
Human Normal Tissues	Different tissues	95	98
Mouse Normal Tissues	Different tissues	91	96
Alzheimer's disease	Normal versus disease	60	70
	Normal versus severe disease	83	91
Adenocarcinoma	Three different conditions	83	83
	Normal versus cancer/precursor	91	91
Bladder Cancer	Three different conditions	73	83
	Normal versus cancer	90	96
Renal Cell Carcinoma	Normal versus cancer	94	100

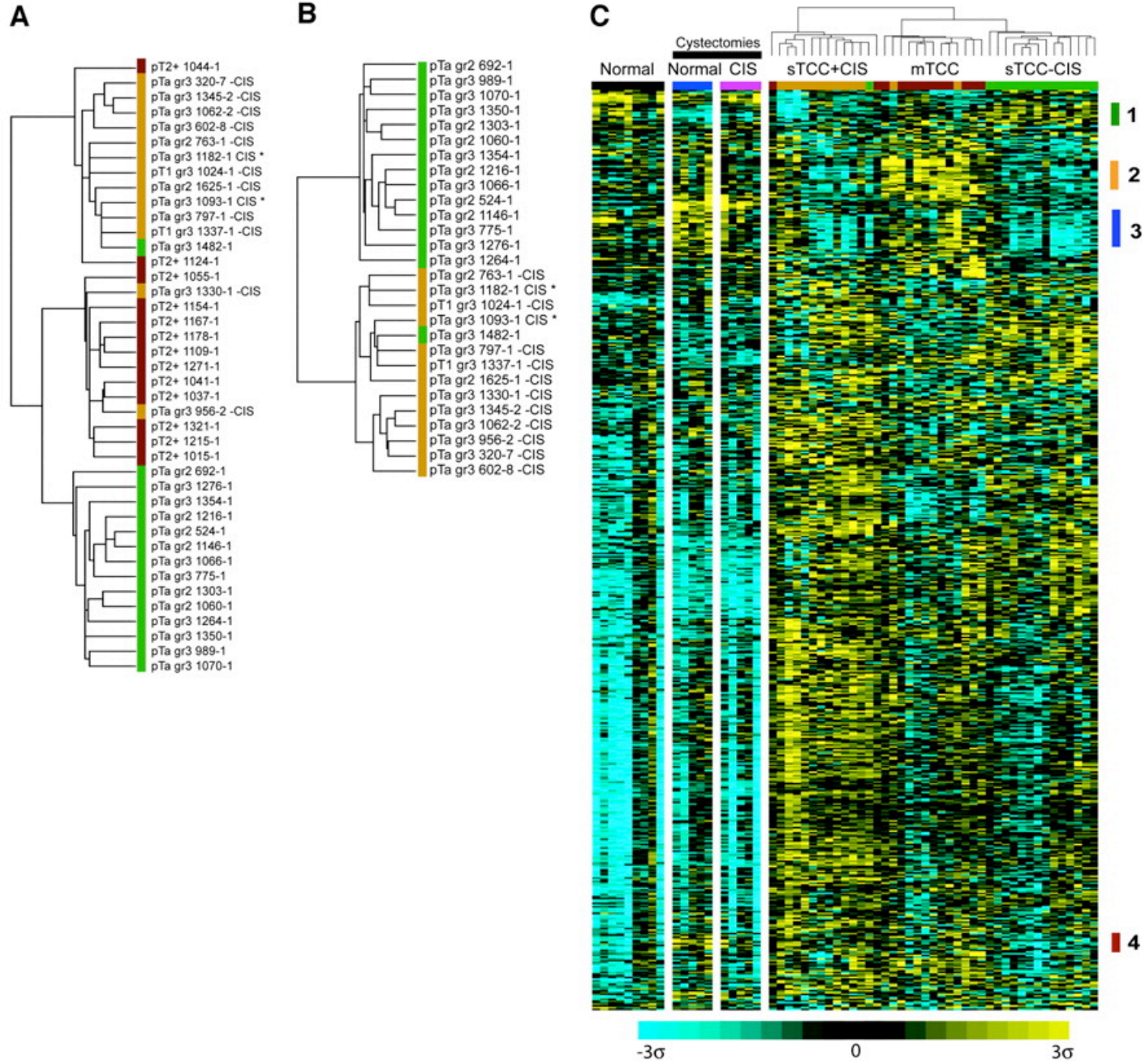
Frozen RMA (fRMA)

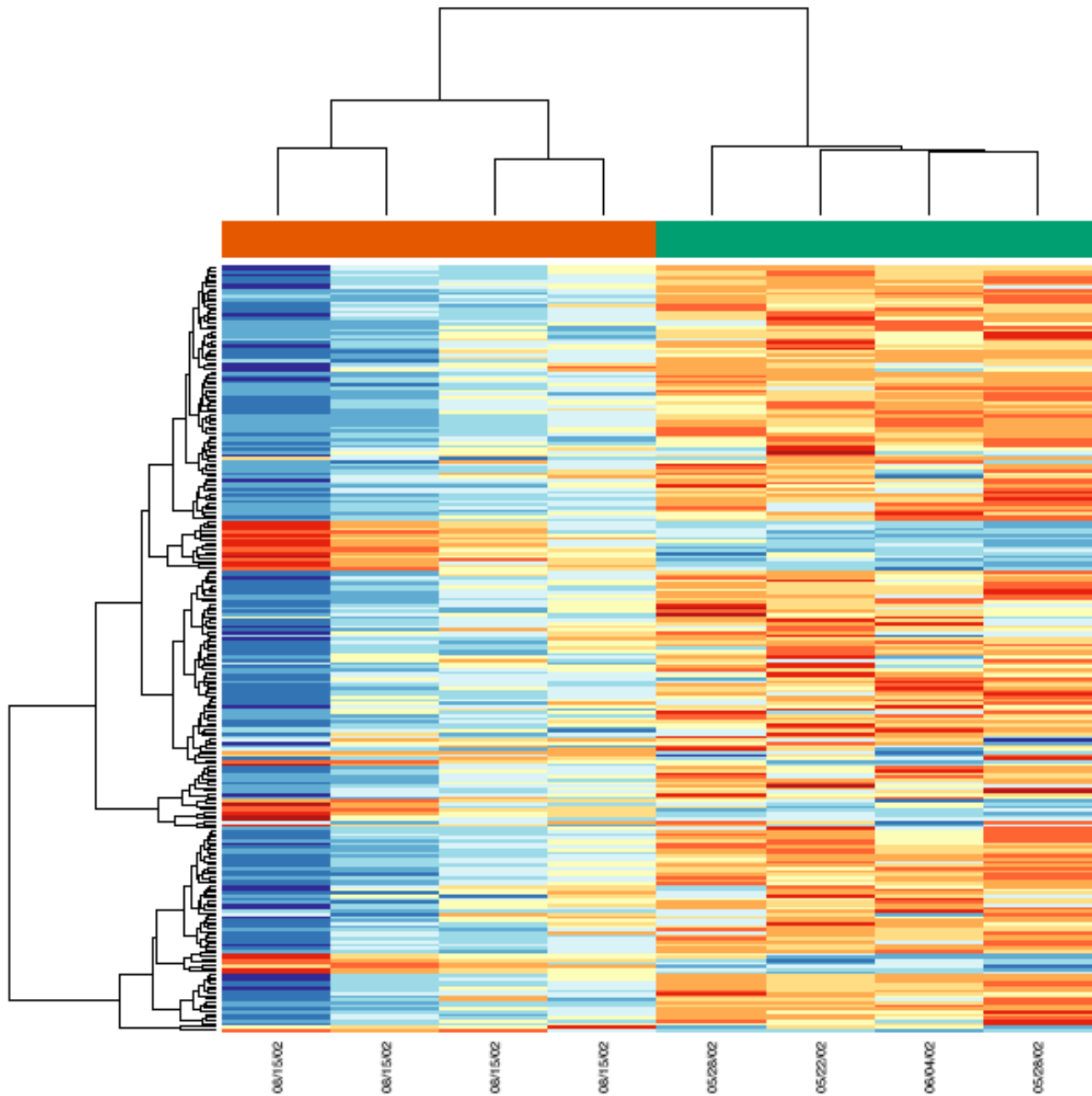
- Create a batch balancing lab and tissue.
- Background correct as usual
- Quantile normalize: keep distribution
- Run robust regression: keep probe effects
- For new sample: background correct as usual, quantile normalize to saved distribution, subtract saved probe effects, take robust average

Table 1 | Percentage accuracy comparison on independent data sets

GEO identifier	Data type	PAM (% correct)	Bar code (% correct)
GSE5388	Cortex	100	100
GSE2395	Respiratory system epithelia	0	100
GSE2665	Lymph node/tonsil	35	95
GSE1561	Breast tumor	69	100
GSE2603	Breast tumor	77	90
GSE6344	Kidney: normal versus cancer	100	100

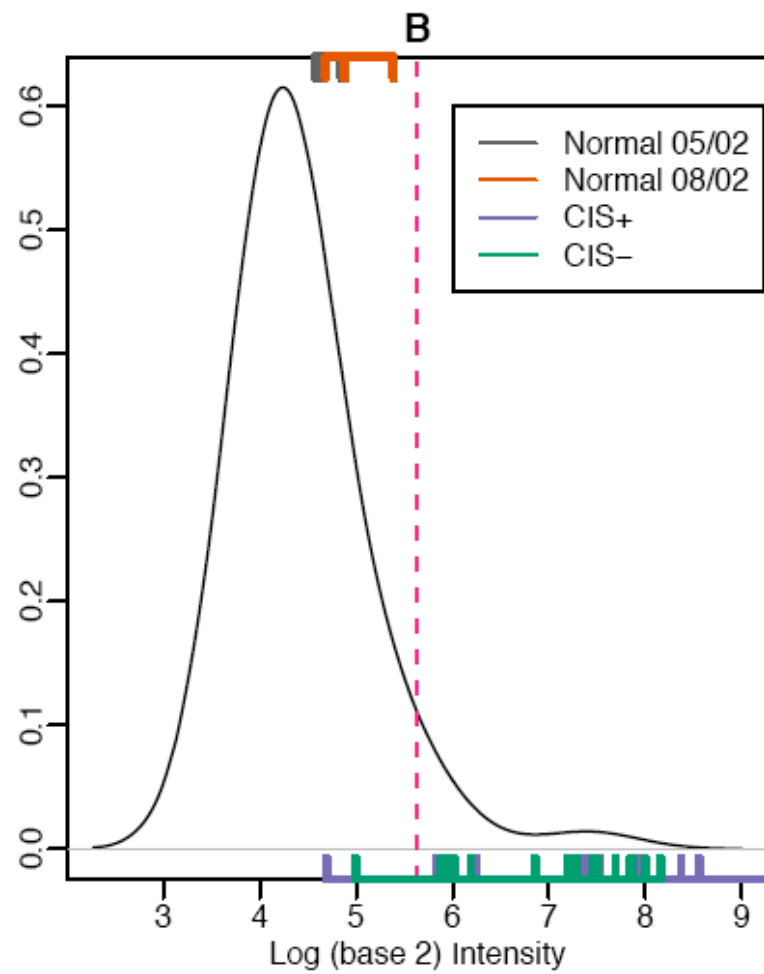
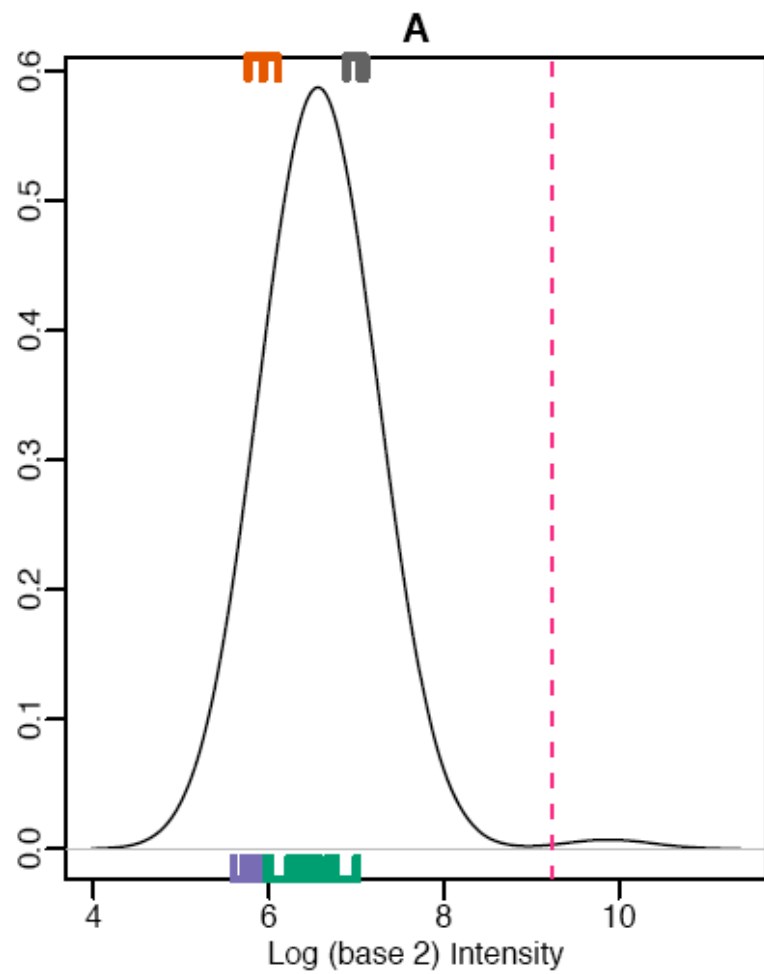
PAM versus the bar code approach in six randomly selected data sets not included in the original database. The data described in **Supplementary Table 1** were used to train the prediction algorithms. GEO, Gene Expression Omnibus.



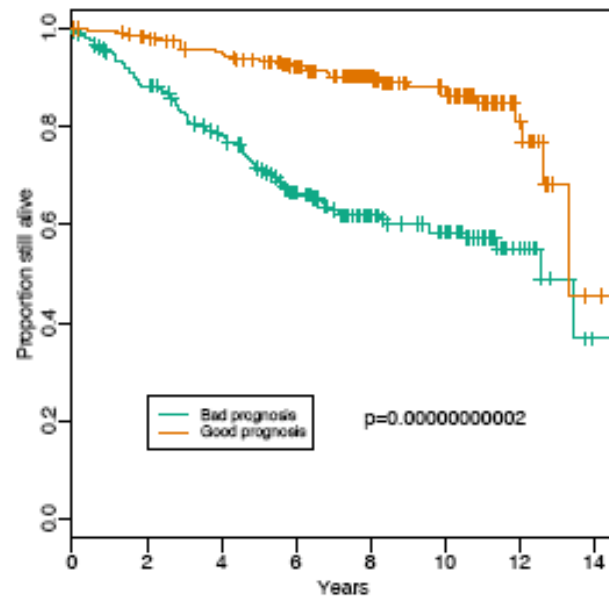


August

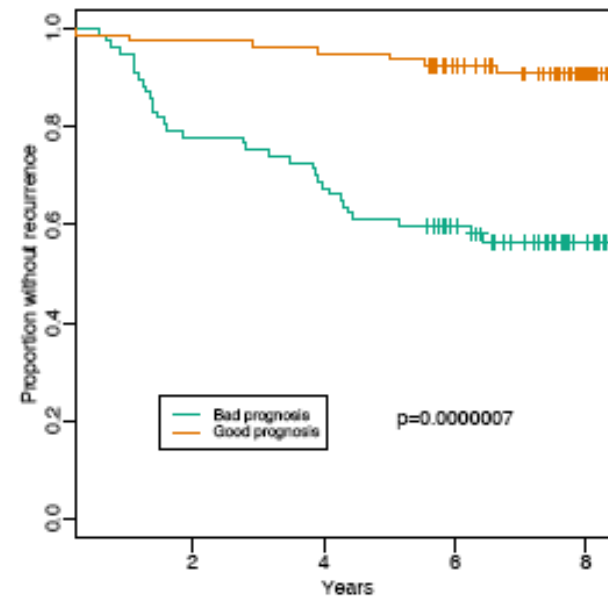
May



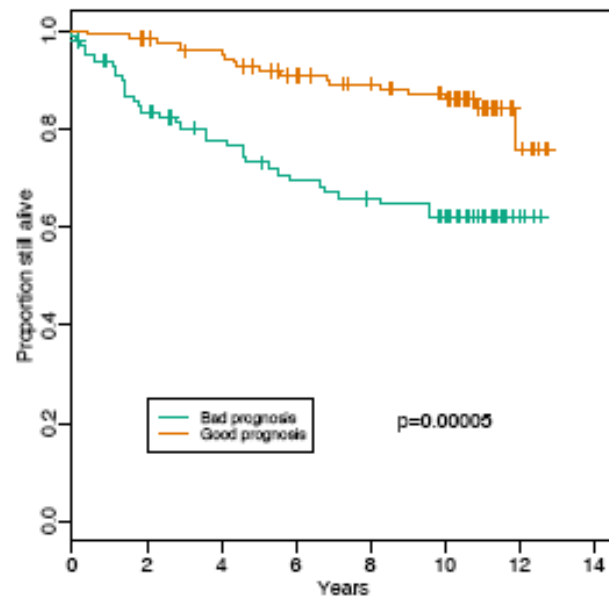
A



B

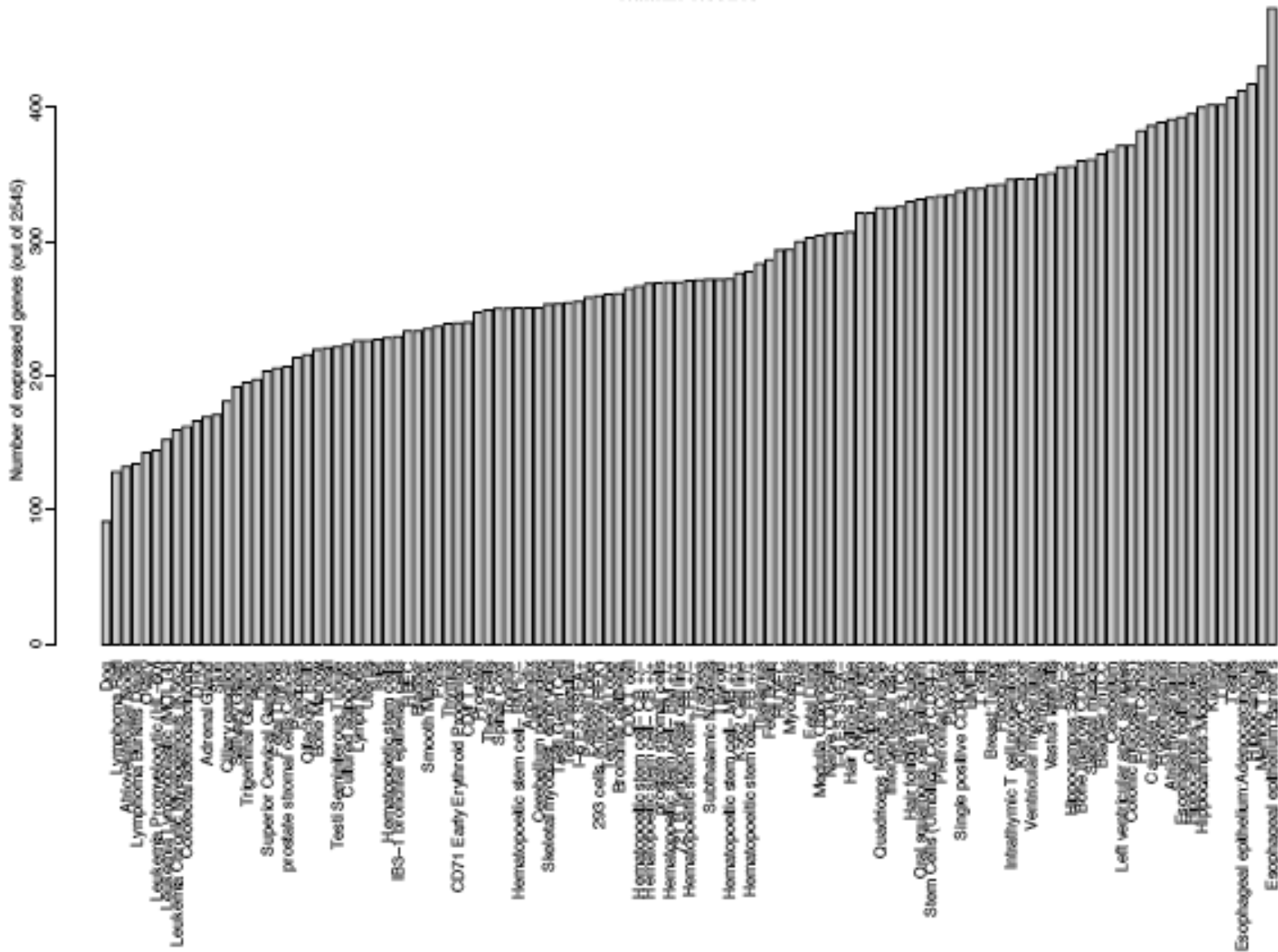


C

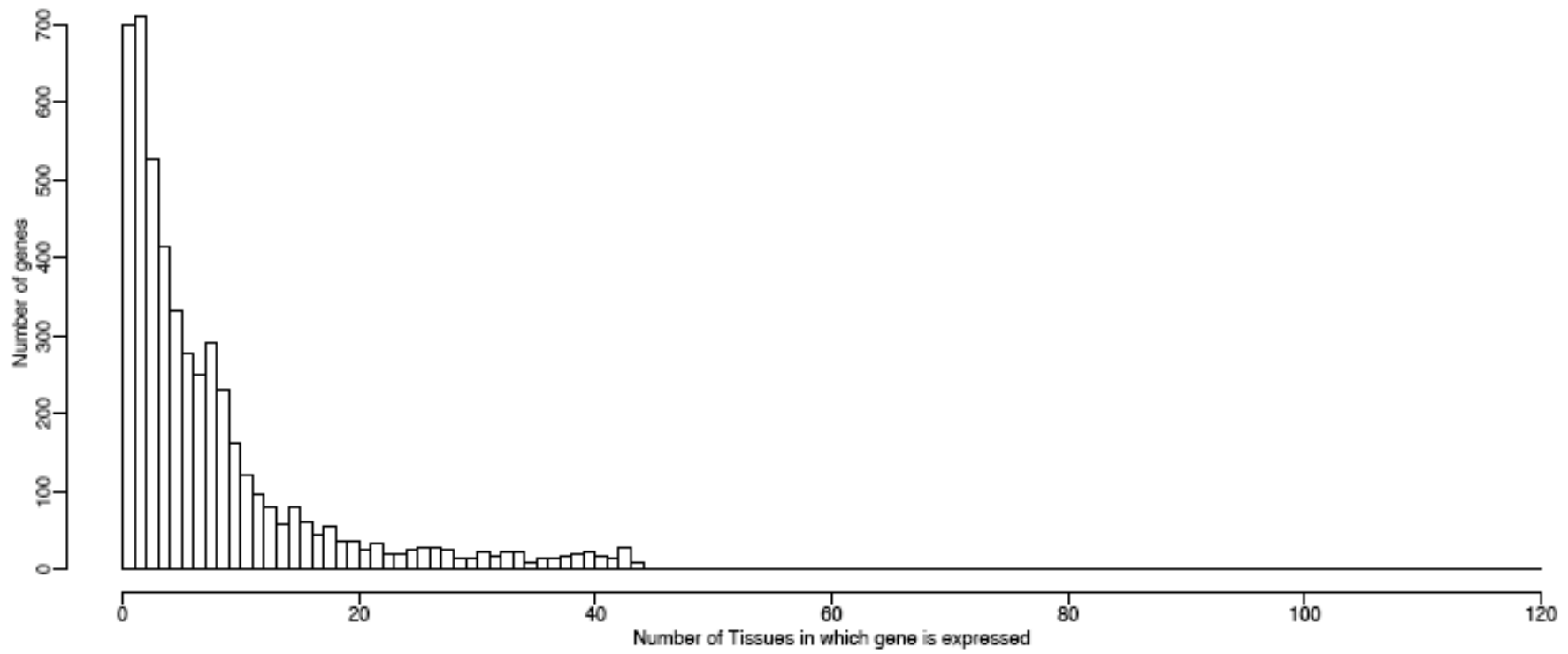


Thanks!

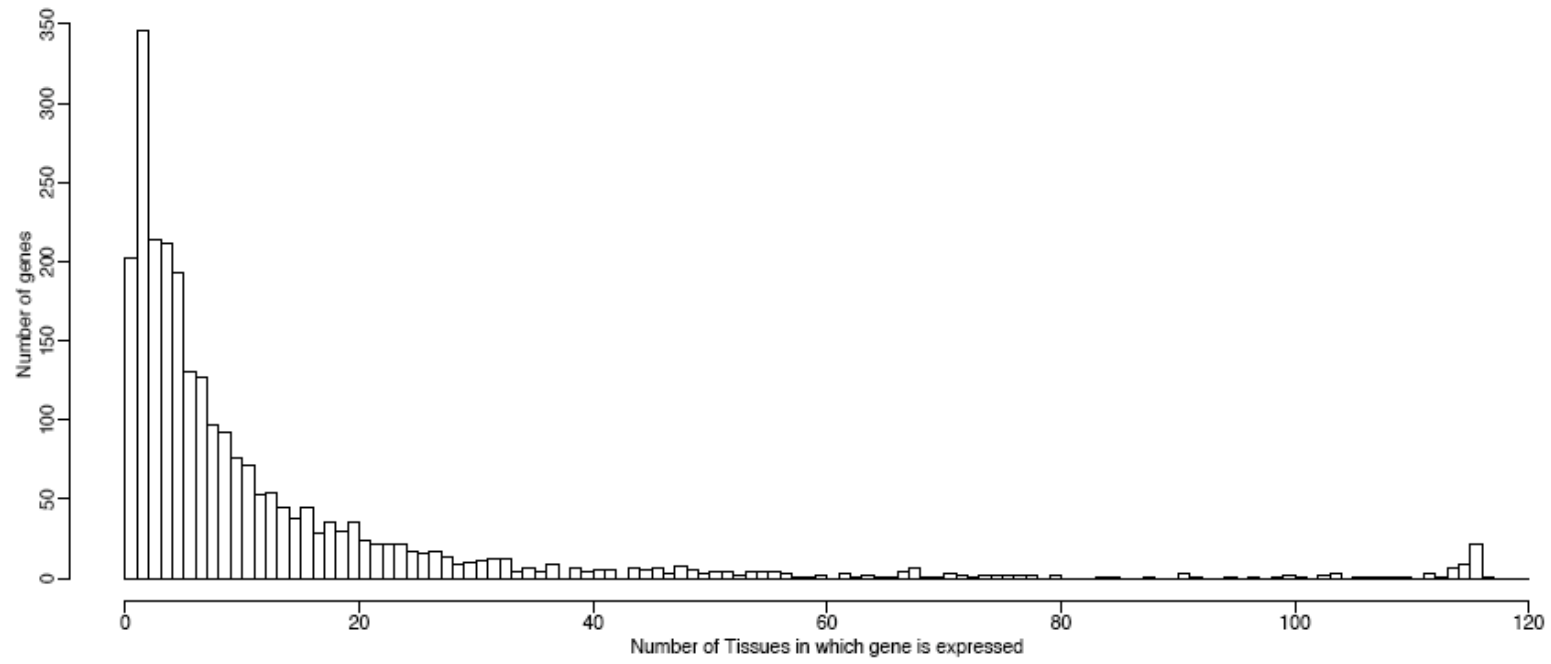
Human Tissues

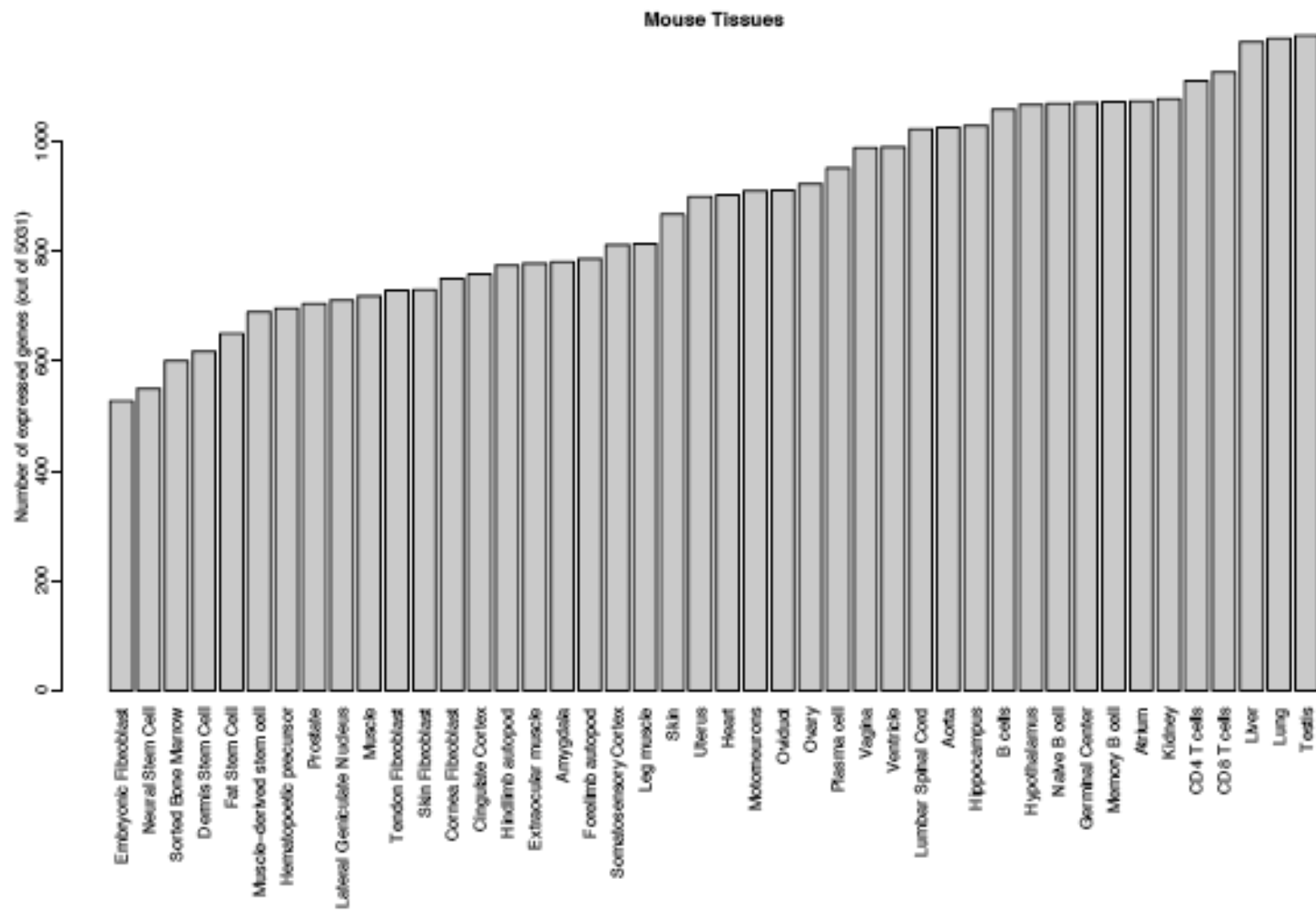


Mouse



Human





Human Tissues

Memory T cells	3	Fetal Brain	2	9	Pancreas	2	
293 cells (Kidney HEK)	5	Fetal Lung	2	HBEC	2	Parietal Lobe	2
721 B lymphoblast cell line	2	Fetal Thyroid	2	Heart	2	CD4 T cell	2
Adrenal Gland	2	Vastus lateralis	4	Hippocampus	9	CD8 T cell	2
Atrial myocardium	30	Myometrium	9	HMEC	3	Placenta	2
Amygdala	2	Bladder	4	HUVEC	3	Pons	2
Left ventricular apex samples	6	Testis	5	Hypothalamus	25	Prefrontal Cortex	2
Appendix	2	Skeletal myoblast, embryonic	3	IB3-1 bronchial epithelial cells	4	Prostate	2
Atrioventricular Node	2	Hair follicle bulge	5	Intervertebral disc	4	Quadriceps femoris muscle	3
Bronchial epithelia	15	Hair follicle sub-bulge	5	Islet	2	Salivary Gland	2
CD71 Early Erythroid Progenitor	2	Oral squamous cell epithelium	3	K562 Cell line	24	Skin	2
Bone marrow CD34+	8	Myoblasts	3	Kidney cells	5	Smooth Muscle	2
Bone Marrow	2	Hematopoietic stem cells	3	Leukemia Chronic Myelogenous (K562)	2	Spinal Cord	2
Caudate nucleus	33	Frontal Cortex	28	Leukemia Lymphoblastic (MOLT4)	2	Subthalamic Nucleus	2
Stem Cells (Umbilical Cord CD34+)	9	I-6 ES SSEA+	3	Leukemia Promyelocytic (HL-60)	2	Superior Cervical Ganglion	2
Na ⁺ 225ve CD4 cells	3	I-6 ES SSEA-	3	Liver	2	Temporal Lobe	2
Single positive CD4 cells	3	Hematopoietic stem cell- FB +/-	3	Lymph Node	2	Testi Germ Cell	2
Cerebellum	29	Hematopoietic stem cell- MB +/-	3	Lymphoma Burkitts Daudi	2	Testi Interstitial	2
Cerebellum peduncles	2	Hematopoietic stem cell- FB +/-	3	Lymphoma Raji	2	Testi Seminiferous Tubule	2
Ciliary ganglion	2	Hematopoietic stem cell- CB +/-	3	Medulla Oblongata	2	Thalamus	2
Colorectal adenocarcinoma	2	Hematopoietic stem cell- CB +/-	3	Hep-2	4	Theca cells	4
T cells	6	Hematopoietic stem cell- BM +/-	3	Colonic mucosal (IBS)	5	Thymus	2
Cultured adipocyte	2	Fibroblast	9	Skeletal muscle	3	Thyroid	2
Lung	11	HBEC	2	Occipital Lobe	2	Tongue	2
DRG	2			Olfactory Bulb	2	Tonsil	2
Esophageal Epithelium	8			Ovary	2	Intrathymic T cell progenitors	3
						Trachea	2
						Trigeminal Ganglion	2
						Uterus	2
						Ventricular myocardium	5