Fig1A - Distribution of correlations between reads mapping to L1 Subfamilies in RNA-seq experiments and # of bases of these same subfamilies in the Ref. Genome
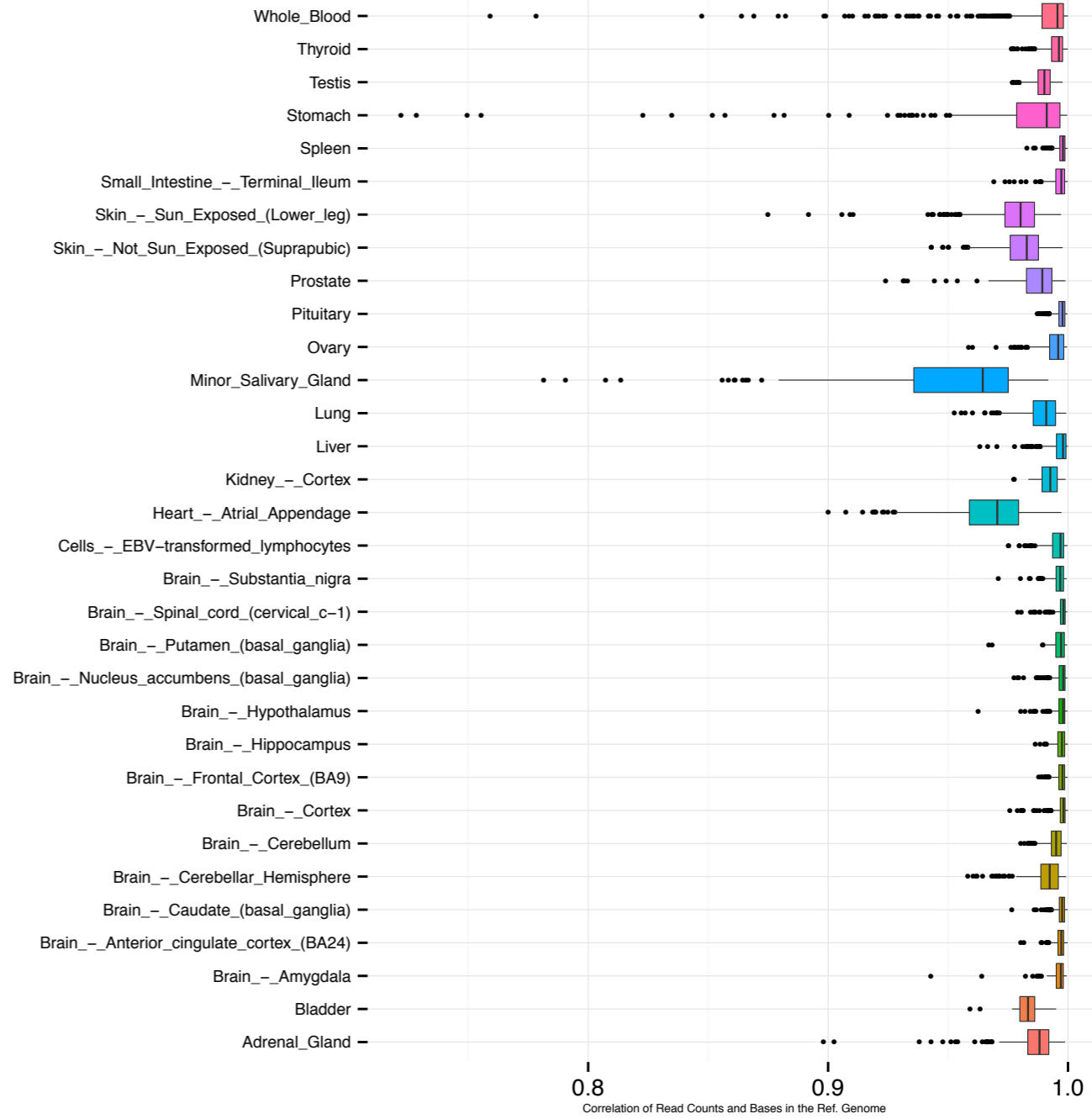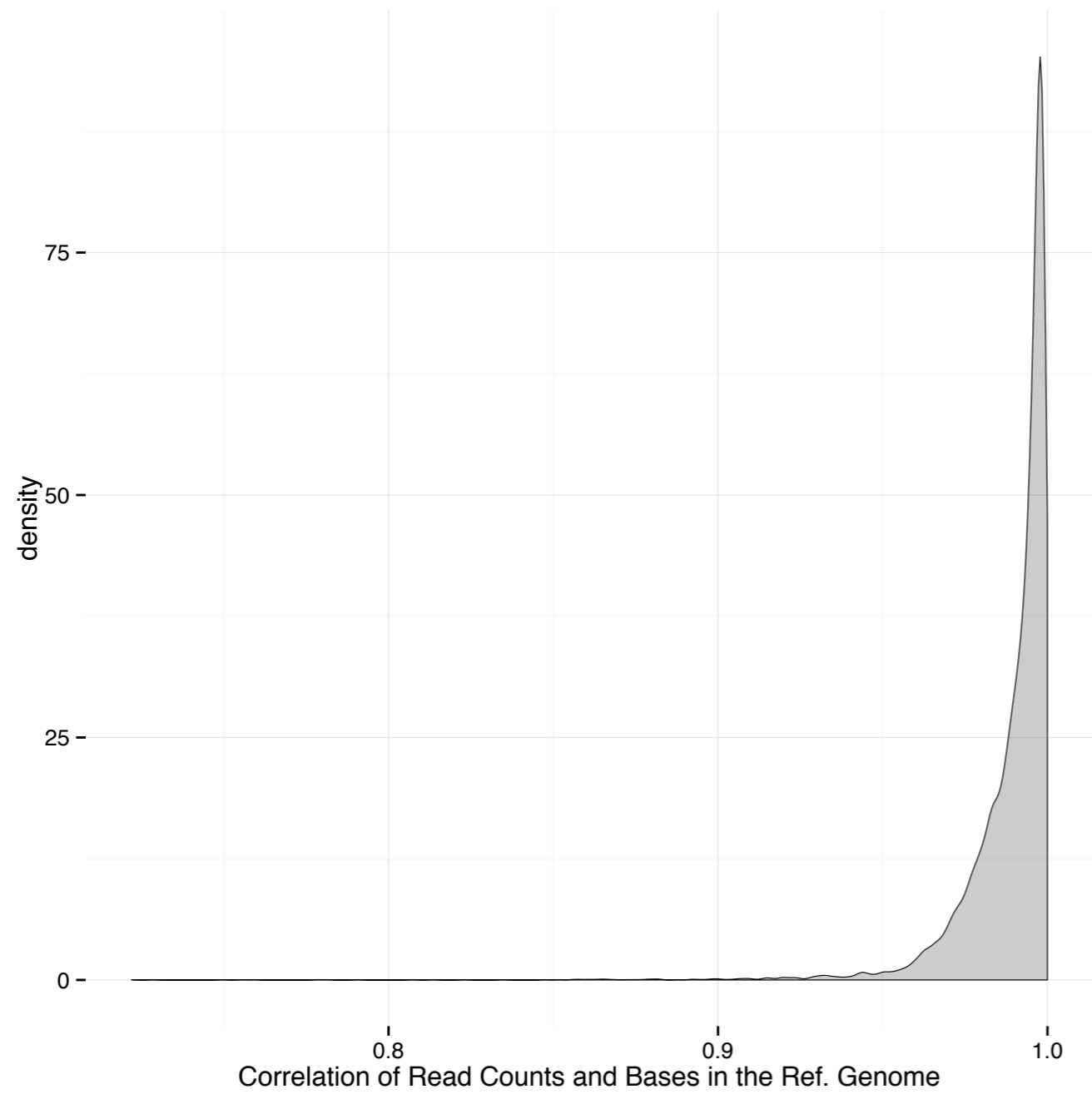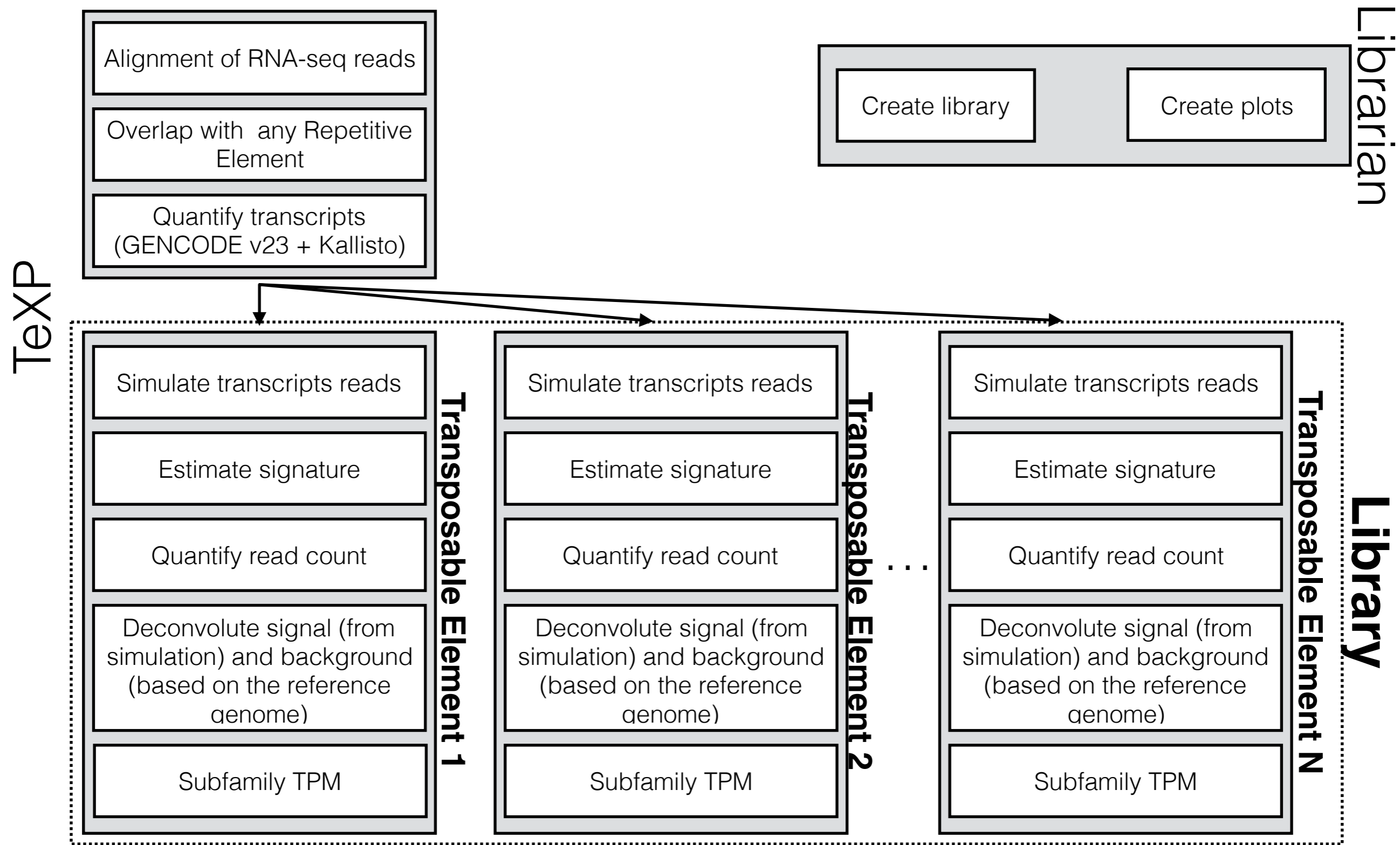
Fig1B - Pipeline

Librarian

Create library   Create plots

TeXP

Alignment of RNA-seq reads

Overlap with any Repetitive Element

Quantify transcripts (GENCODE v23 + Kallisto)

Transposable Element 1

Simulate transcripts reads

Estimate signature

Quantify read count

Deconvolute signal (from simulation) and background (based on the reference genome)

Subfamily TPM

Transposable Element 2

Simulate transcripts reads

Estimate signature

Quantify read count

Deconvolute signal (from simulation) and background (based on the reference genome)

Subfamily TPM

. . .

Transposable Element N

Simulate transcripts reads

Estimate signature

Quantify read count

Deconvolute signal (from simulation) and background (based on the reference genome)
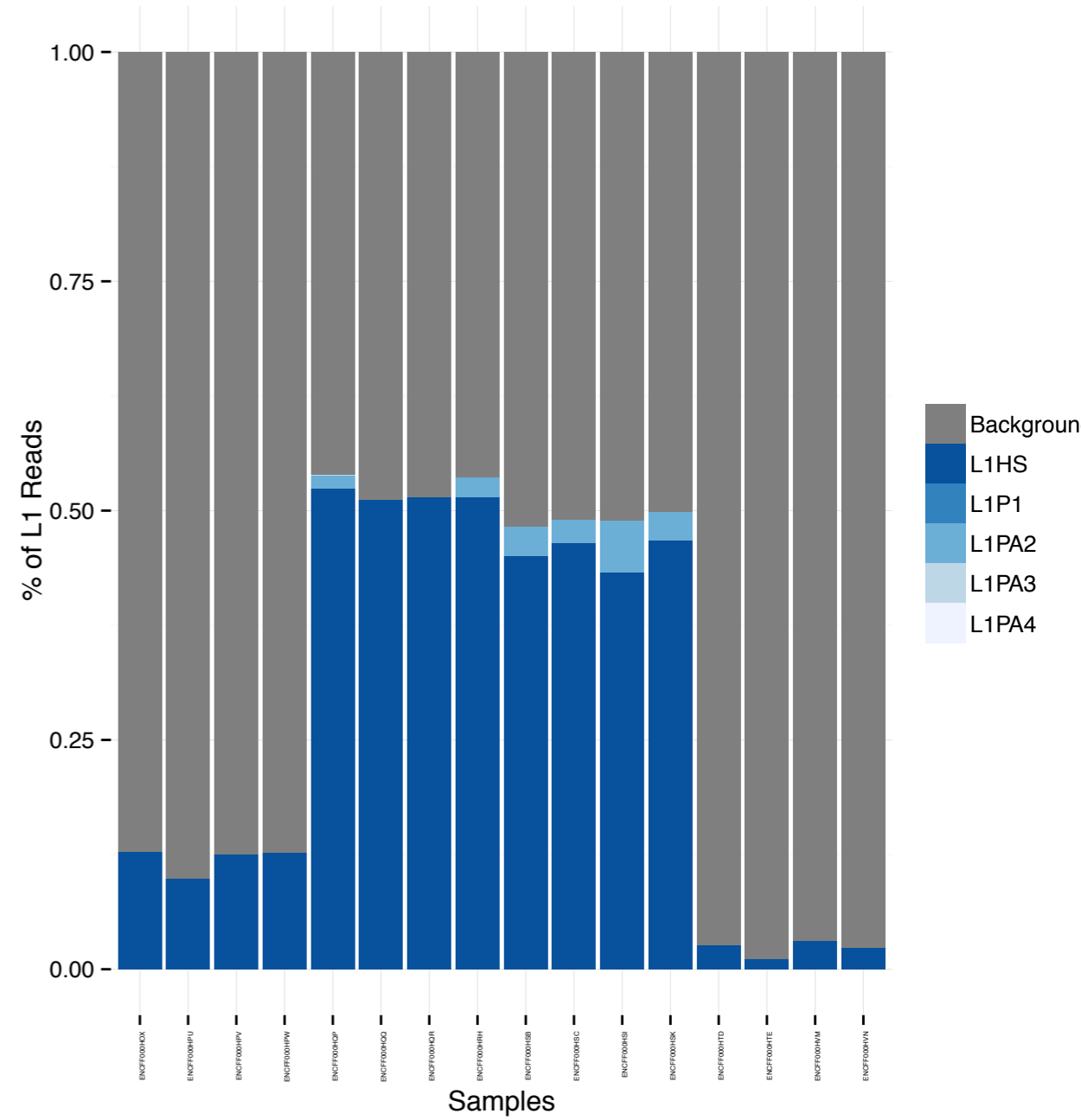
Subfamily TPM

Library

Fig1C - MCF7 after/before

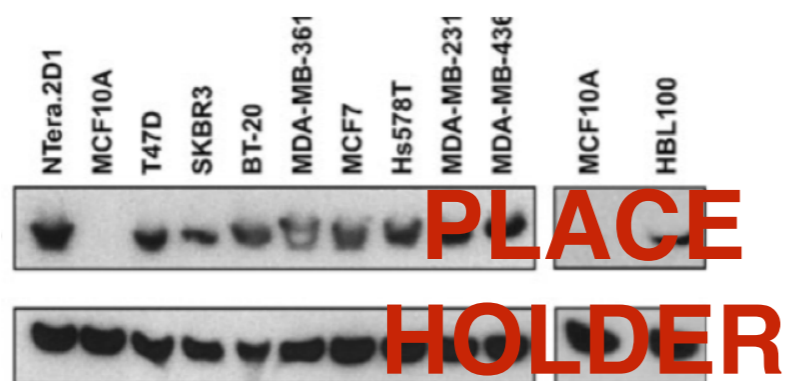# Fig1D - Controls
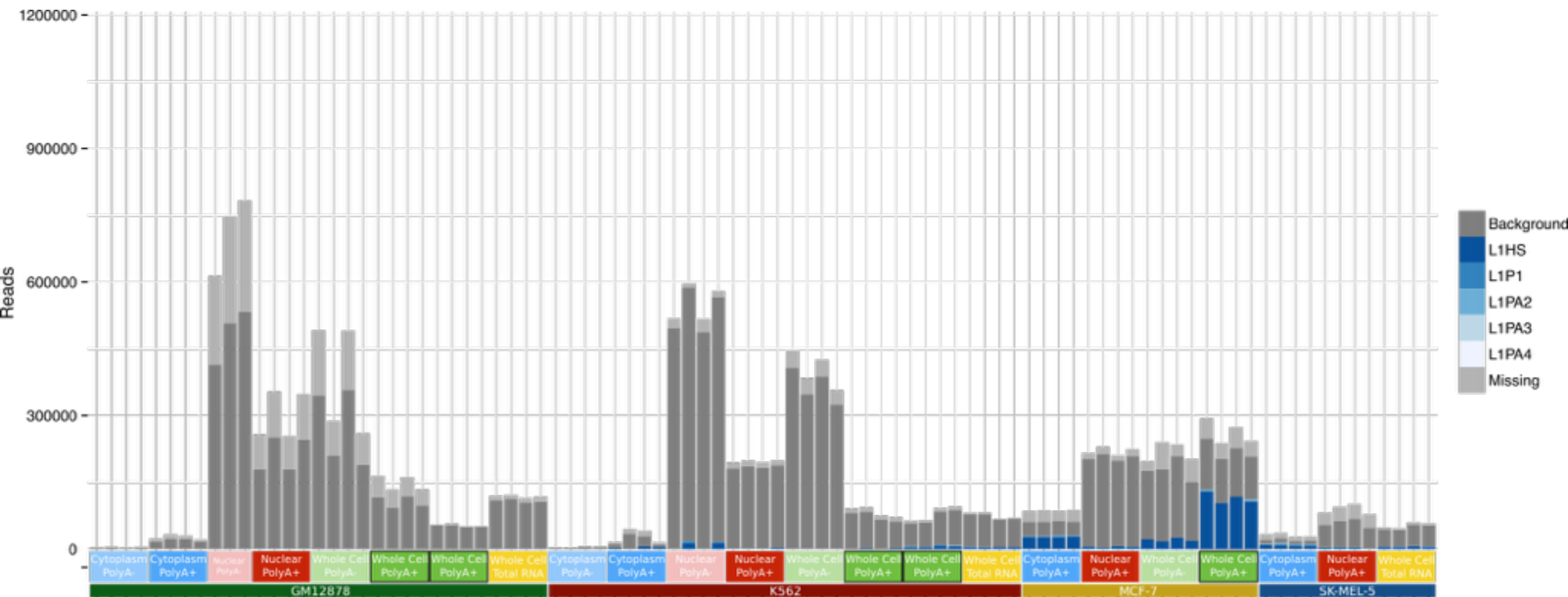
Fig1E - Validation

Table1 - Processed samples

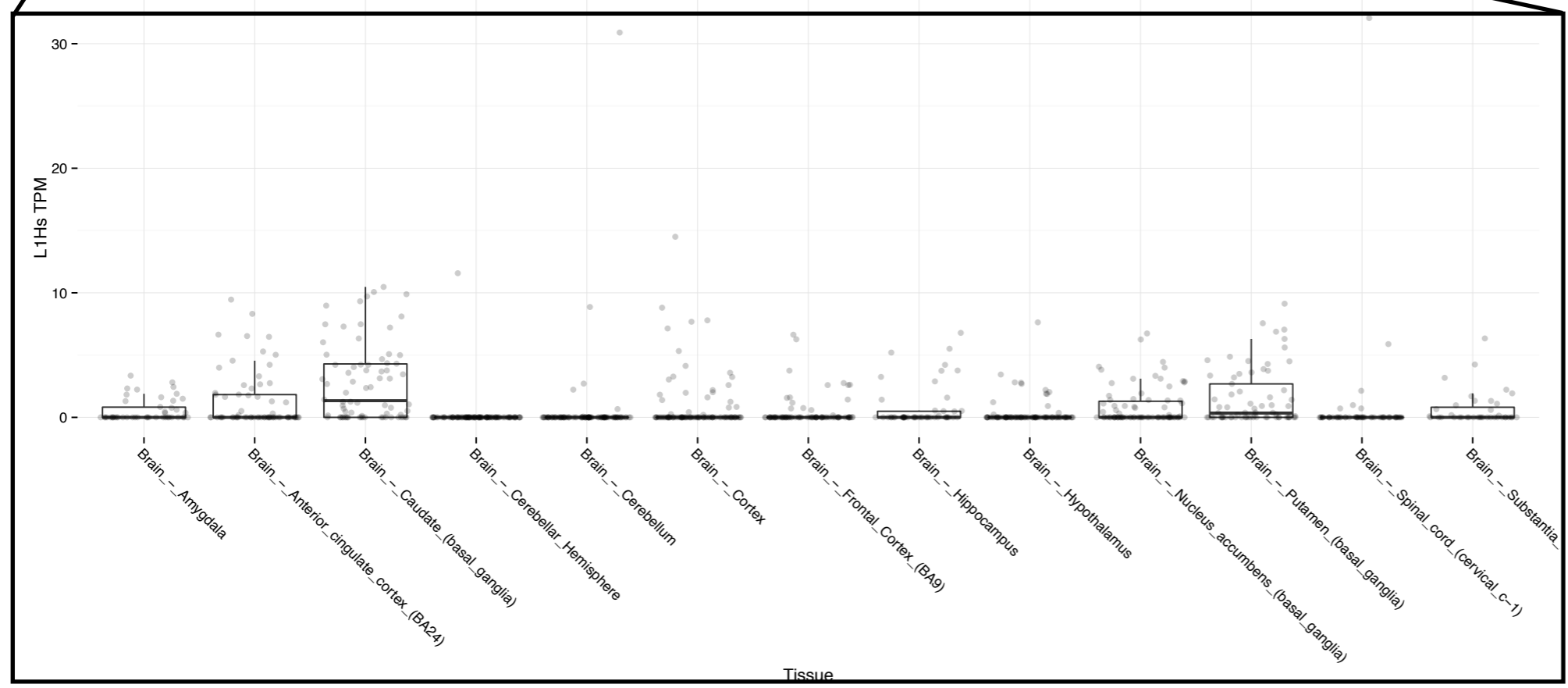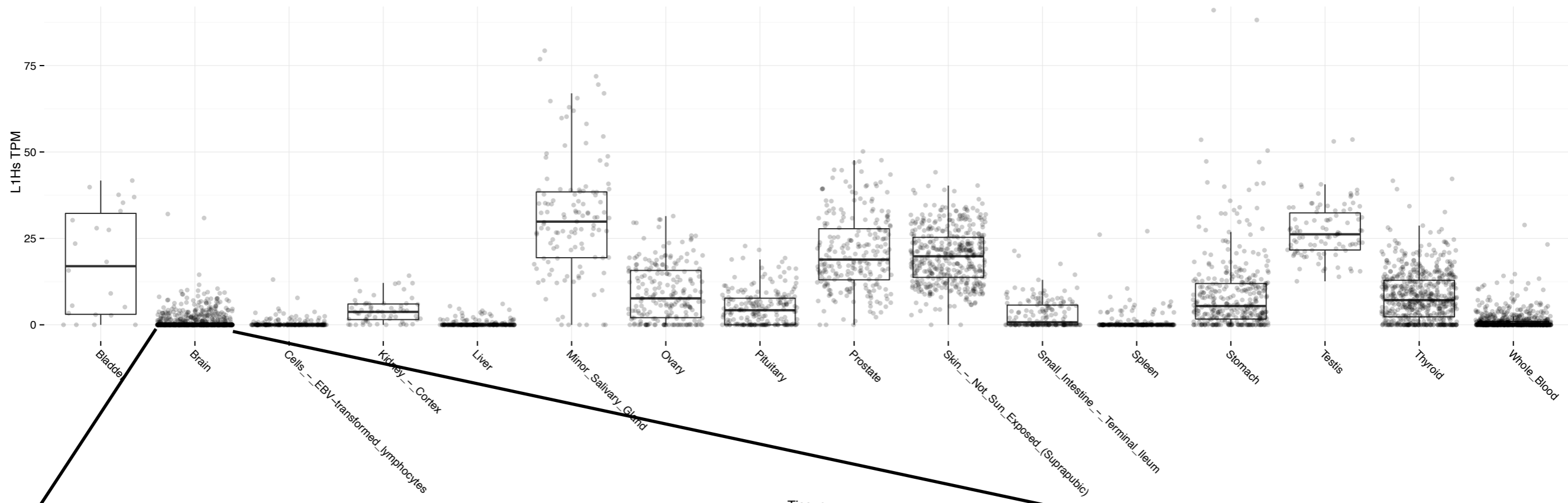| Project	RNA-seq samples	Tissues
GTex	3836	16
BrainSpan	618	1
TCGA	673	1
272 Glioblastoma	272	1
Brain single cell	466	1 (9 cell types)
Embryonic single cell	124	7 embryonic stages
PsychENCODE (Flora's iPSC)	-	-

Fig2A - L1Hs in other GTex tissues

Fig2B - L1Hs and age

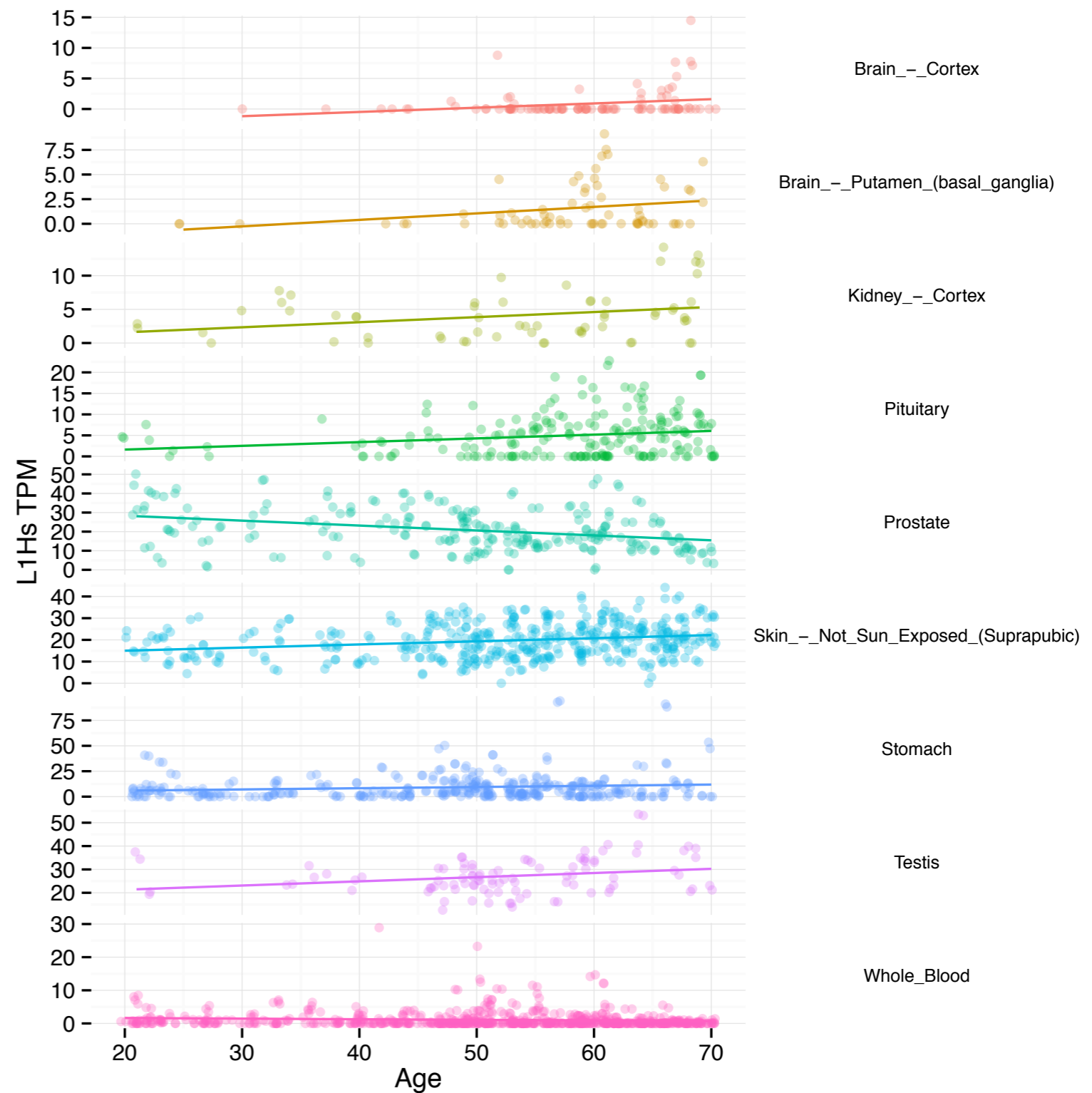| Tissue	Correlation	p-value
Brain - Cortex	0.23	1.8x10-2
Brain - Putamen	0.25	3.1x10-2
Kidney	0.27	2.9x10-2
Pituitary gland	0.17	9x10-3
Prostate	-0.31	5.4x10-8
Skin (not exposed)	0.22	1.38x10-5
Stomach	0.12	1.9x10-2
Testis	0.24	1.5x10-2
Whole blood	-0.1	4x10-3

Fig2B - Human Brain and TEs (L1HS)

+Very early development single-cell sequencing

Fig2C - Human Brain and TEs (LTR5_Hs)

+Very early development single-cell sequencing

Fig3 - Diseased brains
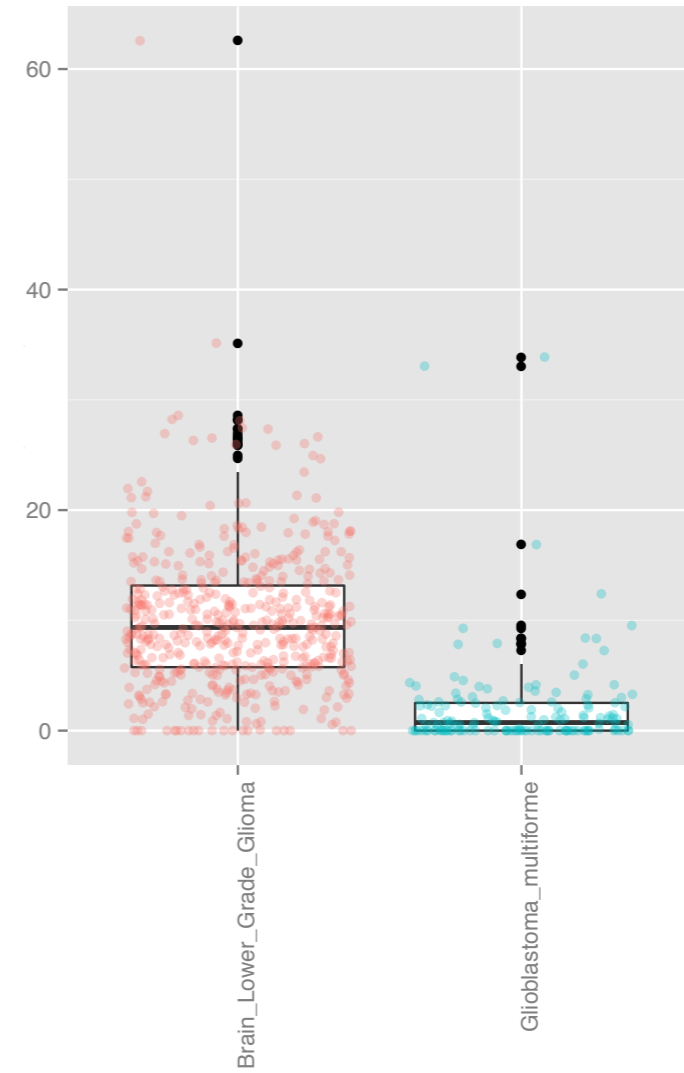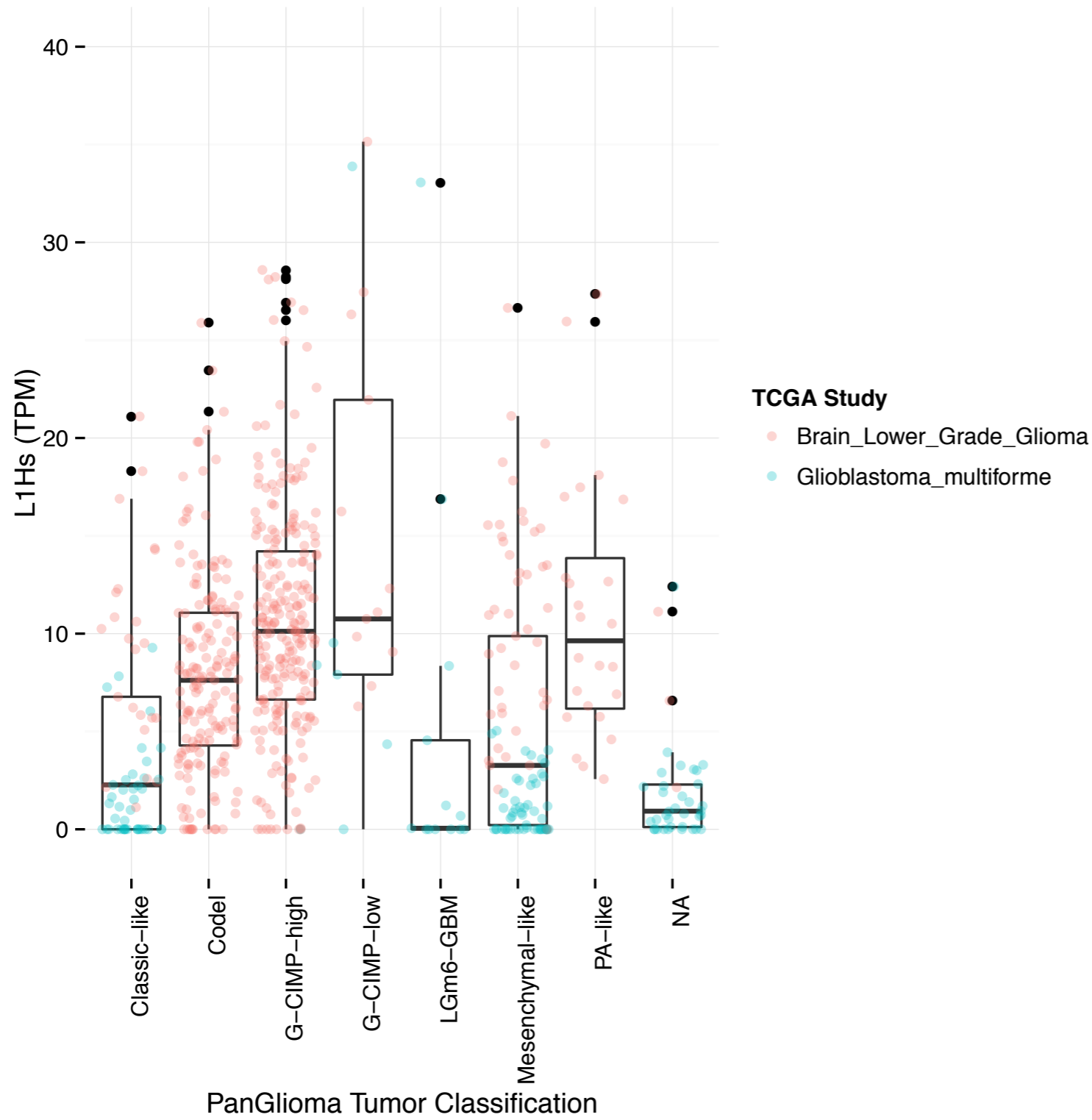
+272 Glioblastoma samples (done)
+TCGA