Table S1. ENCODE RNA-seq experiments.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ENCODE ID | Cell-line | Fraction | rRNA treatment | Transcript selection |
| ENCSR000AED | GM12878 | whole\_cell | rRNA-depleted | Poly-A+ |
| ENCSR000AEC | GM12878 | whole\_cell | rRNA-depleted | Total |
| ENCSR000CQE | GM12878 | cytoplasmic | rRNA-depleted | DSN\_normalized\_Poly-A- |
| ENCSR000COR | GM12878 | cytoplasmic | rRNA-depleted | Poly-A+ |
| ENCSR000CQF | GM12878 | nuclear | rRNA-depleted | DSN\_normalized\_Poly-A- |
| ENCSR000CPO | GM12878 | nuclear | rRNA-depleted | Poly-A+ |
| ENCSR000CVT | GM12878 | nucleolar | rRNA-depleted | Total |
| ENCSR000COS | GM12878 | whole\_cell | rRNA-depleted | DSN\_normalized\_Poly-A- |
| ENCSR000COQ | GM12878 | whole\_cell | rRNA-depleted | Poly-A+ |
| ENCSR000CQT | HeLa-S3 | cytoplasmic | rRNA-depleted | DSN\_normalized\_Poly-A- |
| ENCSR000CPP | HeLa-S3 | cytoplasmic | rRNA-depleted | Poly-A+ |
| ENCSR000CQI | HeLa-S3 | nuclear | rRNA-depleted | DSN\_normalized\_Poly-A- |
| ENCSR000CPQ | HeLa-S3 | nuclear | rRNA-depleted | Poly-A+ |
| ENCSR000CQJ | HeLa-S3 | whole\_cell | rRNA-depleted | DSN\_normalized\_Poly-A- |
| ENCSR000CPR | HeLa-S3 | whole\_cell | rRNA-depleted | Poly-A+ |
| ENCSR000CQU | HepG2 | cytoplasmic | rRNA-depleted | DSN\_normalized\_Poly-A- |
| ENCSR000CPF | HepG2 | cytoplasmic | rRNA-depleted | Poly-A+ |
| ENCSR000CQK | HepG2 | nuclear | rRNA-depleted | DSN\_normalized\_Poly-A- |
| ENCSR000CPC | HepG2 | nuclear | rRNA-depleted | Poly-A+ |
| ENCSR000CPD | HepG2 | whole\_cell | rRNA-depleted | DSN\_normalized\_Poly-A- |
| ENCSR000CPE | HepG2 | whole\_cell | rRNA-depleted | Poly-A+ |
| ENCSR166QLP | HT1080 | cytoplasmic | rRNA-depleted | Poly-A+ |
| ENCSR535VTR | HT1080 | whole\_cell | rRNA-depleted | Total |
| ENCSR067UNX | HT1080 | nuclear | rRNA-depleted | Poly-A+ |
| ENCSR000AEM | K562 | whole\_cell | rRNA-depleted | Poly-A+ |
| ENCSR000AEL | K562 | whole\_cell | rRNA-depleted | Total |
| ENCSR000CPY | K562 | chromatin | rRNA-depleted | Total |
| ENCSR000CPY | K562 | chromatin | rRNA-depleted | Total |
| ENCSR000CQL | K562 | cytoplasmic | rRNA-depleted | DSN\_normalized\_Poly-A- |
| ENCSR000COK | K562 | cytoplasmic | rRNA-depleted | Poly-A+ |
| ENCSR000CQM | K562 | nuclear | rRNA-depleted | DSN\_normalized\_Poly-A- |
| ENCSR000CPS | K562 | nuclear | rRNA-depleted | Poly-A+ |
| ENCSR000CPZ | K562 | nucleolar | rRNA-depleted | Total |
| ENCSR000CQA | K562 | nucleoplasm | rRNA-depleted | Total |
| ENCSR000CPG | K562 | whole\_cell | rRNA-depleted | DSN\_normalized\_Poly-A- |
| ENCSR000CPH | K562 | whole\_cell | rRNA-depleted | Poly-A+ |
| ENCSR000CTU | MCF-7 | cytoplasmic | rRNA-depleted | Poly-A+ |
| ENCSR000CTO | MCF-7 | nuclear | rRNA-depleted | Poly-A+ |
| ENCSR000CQB | MCF-7 | whole\_cell | rRNA-depleted | DSN\_normalized\_Poly-A- |
| ENCSR000CPT | MCF-7 | whole\_cell | rRNA-depleted | Poly-A+ |
| ENCSR586SEE | NCI-H460 | cytoplasmic | rRNA-depleted | Poly-A+ |
| ENCSR164OCT | NCI-H460 | whole\_cell | rRNA-depleted | Total |
| ENCSR625QJI | NCI-H460 | nuclear | rRNA-depleted | Poly-A+ |
| ENCSR291DJH | SK-MEL-5 | cytoplasmic | rRNA-depleted | Poly-A+ |
| ENCSR669KQU | SK-MEL-5 | whole\_cell | rRNA-depleted | Total |
| ENCSR201WVA | SK-MEL-5 | nuclear | rRNA-depleted | Poly-A+ |
| ENCSR569JKX | SK-N-DZ | cytoplasmic | rRNA-depleted | Poly-A+ |
| ENCSR136WGP | SK-N-DZ | whole\_cell | rRNA-depleted | Total |
| ENCSR255NYQ | SK-N-DZ | nuclear | rRNA-depleted | Poly-A+ |
| ENCSR000CTR | SK-N-SH | cytoplasmic | rRNA-depleted | Poly-A+ |
| ENCSR000CTS | SK-N-SH | nuclear | rRNA-depleted | Poly-A+ |
| ENCSR000CQP | SK-N-SH | whole\_cell | rRNA-depleted | DSN\_normalized\_Poly-A- |
| ENCSR000CPN | SK-N-SH | whole\_cell | rRNA-depleted | Poly-A+ |
| ENCSR000CTT | SK-N-SH | whole\_cell | rRNA-depleted | Poly-A+ |

**Table S2. Number of samples per tissue tissue type.** Bladder, Kidney Cortex and minor salivary gland were eliminated from further analysis.

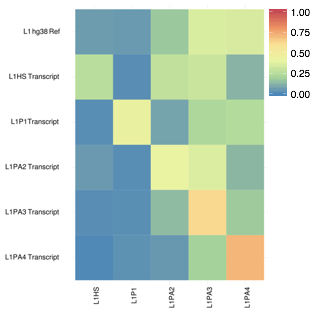
|  |  |
| --- | --- |
| 234 | Adipose\_-\_Visceral\_(Omentum) |
| 146 | Adrenal\_Gland |
| 123 | Artery\_-\_Coronary |
| 11 | Bladder |
| 81 | Brain\_-\_Amygdala |
| 99 | Brain\_-\_Anterior\_cingulate\_cortex\_(BA24) |
| 133 | Brain\_-\_Caudate\_(basal\_ganglia) |
| 115 | Brain\_-\_Cerebellar\_Hemisphere |
| 145 | Brain\_-\_Cerebellum |
| 132 | Brain\_-\_Cortex |
| 117 | Brain\_-\_Frontal\_Cortex\_(BA9) |
| 102 | Brain\_-\_Hippocampus |
| 103 | Brain\_-\_Hypothalamus |
| 123 | Brain\_-\_Nucleus\_accumbens\_(basal\_ganglia) |
| 103 | Brain\_-\_Putamen\_(basal\_ganglia) |
| 76 | Brain\_-\_Spinal\_cord\_(cervical\_c-1) |
| 71 | Brain\_-\_Substantia\_nigra |
| 200 | Breast\_-\_Mammary\_Tissue |
| 132 | Cells\_-\_EBV-transformed\_lymphocytes |
| 78 | Cells\_-\_Leukemia\_cell\_line\_(CML) |
| 300 | Cells\_-\_Transformed\_fibroblasts |
| 142 | Colon\_-\_Sigmoid |
| 178 | Colon\_-\_Transverse |
| 151 | Esophagus\_-\_Gastroesophageal\_Junction |
| 192 | Heart\_-\_Atrial\_Appendage |
| 36 | Kidney\_-\_Cortex |
| 136 | Liver |
| 280 | Lung |
| 69 | Minor\_Salivary\_Gland |
| 468 | Muscle\_-\_Skeletal |
| 335 | Nerve\_-\_Tibial |
| 108 | Ovary |
| 193 | Pancreas |
| 124 | Pituitary |
| 119 | Prostate |
| 271 | Skin\_-\_Not\_Sun\_Exposed\_(Suprapubic) |
| 395 | Skin\_-\_Sun\_Exposed\_(Lower\_leg) |
| 104 | Small\_Intestine\_-\_Terminal\_Ileum |
| 118 | Spleen |
| 205 | Stomach |
| 199 | Testis |
| 355 | Thyroid |
| 90 | Uterus |
| 88 | Vagina |
| 449 | Whole\_Blood |

Table S3. L1 autonomous transcription level and age correlation coefficient and significance for each GTEx tissue.

|  |  |  |
| --- | --- | --- |
| Tissue | correlation | FDR |
| Lung | 0.2818789 | 7.22E-05 |
| Muscle\_-\_Skeletal | 0.1788474 | 2.20E-03 |
| Cells\_-\_Transformed\_fibroblasts | 0.2089546 | 3.60E-03 |
| Prostate | -0.3238307 | 3.60E-03 |
| Adipose\_-\_Visceral\_(Omentum) | 0.2277305 | 3.93E-03 |
| Skin\_-\_Not\_Sun\_Exposed\_(Suprapubic) | 0.1753624 | 2.77E-02 |
| Whole\_Blood | -0.1302921 | 3.58E-02 |
| Breast\_-\_Mammary\_Tissue | 0.1858914 | 4.11E-02 |
| Testis | 0.1886884 | 4.11E-02 |
| Skin\_-\_Sun\_Exposed\_(Lower\_leg) | 0.1147755 | 9.91E-02 |
| Esophagus\_-\_Gastroesophageal\_Junction | 0.1791177 | 1.11E-01 |
| Brain\_-\_Cerebellar\_Hemisphere | 0.1997148 | 1.19E-01 |
| Brain\_-\_Frontal\_Cortex\_(BA9) | 0.1880845 | 1.33E-01 |
| Pituitary | 0.182976 | 1.33E-01 |
| Brain\_-\_Cortex | 0.1588801 | 2.02E-01 |
| Brain\_-\_Substantia\_nigra | 0.2120247 | 2.09E-01 |
| Brain\_-\_Hippocampus | 0.1600052 | 2.69E-01 |
| Kidney\_-\_Cortex | 0.2710293 | 2.69E-01 |
| Heart\_-\_Atrial\_Appendage | 0.1102971 | 2.88E-01 |
| Minor\_Salivary\_Gland | -0.1837433 | 2.88E-01 |
| Uterus | 0.139628 | 3.97E-01 |
| Adrenal\_Gland | 0.102378 | 4.36E-01 |
| Nerve\_-\_Tibial | 0.06607081 | 4.36E-01 |
| Brain\_-\_Anterior\_cingulate\_cortex\_(BA24) | 0.1186109 | 4.44E-01 |
| Cells\_-\_EBV-transformed\_lymphocytes | 0.09962761 | 4.50E-01 |
| Brain\_-\_Nucleus\_accumbens\_(basal\_ganglia) | 0.08814543 | 5.62E-01 |
| Brain\_-\_Caudate\_(basal\_ganglia) | 0.07846093 | 6.02E-01 |
| Brain\_-\_Putamen\_(basal\_ganglia) | 0.08561444 | 6.13E-01 |
| Ovary | 0.07199146 | 6.96E-01 |
| Stomach | 0.05014164 | 6.97E-01 |
| Bladder | -0.2305936 | 7.03E-01 |
| Brain\_-\_Cerebellum | -0.05408436 | 7.13E-01 |
| Liver | 0.04890585 | 7.19E-01 |
| Small\_Intestine\_-\_Terminal\_Ileum | 0.05828941 | 7.19E-01 |
| Thyroid | 0.03256656 | 7.19E-01 |
| Spleen | -0.03437764 | 8.70E-01 |
| Pancreas | 0.02279996 | 8.95E-01 |
| Colon\_-\_Sigmoid | 0.02209352 | 9.19E-01 |
| Brain\_-\_Amygdala | 0.02406885 | 9.36E-01 |
| Brain\_-\_Hypothalamus | -0.01603804 | 9.36E-01 |
| Brain\_-\_Spinal\_cord\_(cervical\_c-1) | -0.02072991 | 9.36E-01 |
| Artery\_-\_Coronary | -0.008006473 | 9.74E-01 |
| Colon\_-\_Transverse | 0.001340182 | 9.86E-01 |
| Vagina | 0.003462678 | 9.86E-01 |

Table S4. L1 autonomous transcription level and BMI correlation coefficient and significance for each GTEx tissue.

|  |  |  |
| --- | --- | --- |
| Tissue | Correlation | FDR |
| Breast\_-\_Mammary\_Tissue | 0.2302589 | 0.04565421 |
| Vagina | 0.278925 | 0.18697398 |
| Nerve\_-\_Tibial | 0.1051893 | 0.59869724 |
| Pituitary | -0.1785558 | 0.59869724 |
| Spleen | -0.1618882 | 0.70290553 |
| Brain\_-\_Hypothalamus | -0.1533396 | 0.76691563 |
| Liver | 0.1353808 | 0.76691563 |
| Adipose\_-\_Visceral\_(Omentum) | -0.02108249 | 0.94438298 |
| Adrenal\_Gland | 0.05035129 | 0.94438298 |
| Artery\_-\_Coronary | 0.04729189 | 0.94438298 |
| Bladder | -0.1457863 | 0.94438298 |
| Brain\_-\_Anterior\_cingulate\_cortex\_(BA24) | 0.1103182 | 0.94438298 |
| Brain\_-\_Cerebellar\_Hemisphere | -0.05415352 | 0.94438298 |
| Brain\_-\_Cerebellum | -0.1021354 | 0.94438298 |
| Brain\_-\_Cortex | 0.05534512 | 0.94438298 |
| Brain\_-\_Frontal\_Cortex\_(BA9) | 0.05059266 | 0.94438298 |
| Brain\_-\_Hippocampus | 0.04102779 | 0.94438298 |
| Brain\_-\_Nucleus\_accumbens\_(basal\_ganglia) | -0.04920859 | 0.94438298 |
| Brain\_-\_Putamen\_(basal\_ganglia) | -0.06361988 | 0.94438298 |
| Brain\_-\_Spinal\_cord\_(cervical\_c-1) | -0.03879462 | 0.94438298 |
| Brain\_-\_Substantia\_nigra | 0.0809203 | 0.94438298 |
| Cells\_-\_EBV-transformed\_lymphocytes | 0.08891989 | 0.94438298 |
| Cells\_-\_Transformed\_fibroblasts | -0.04567598 | 0.94438298 |
| Colon\_-\_Sigmoid | 0.08788036 | 0.94438298 |
| Colon\_-\_Transverse | -0.08720016 | 0.94438298 |
| Esophagus\_-\_Gastroesophageal\_Junction | -0.1032197 | 0.94438298 |
| Kidney\_-\_Cortex | 0.07034367 | 0.94438298 |
| Lung | 0.03207328 | 0.94438298 |
| Minor\_Salivary\_Gland | -0.03864037 | 0.94438298 |
| Pancreas | -0.02870926 | 0.94438298 |
| Skin\_-\_Sun\_Exposed\_(Lower\_leg) | 0.02586574 | 0.94438298 |
| Small\_Intestine\_-\_Terminal\_Ileum | 0.02866719 | 0.94438298 |
| Stomach | -0.03121609 | 0.94438298 |
| Testis | 0.02448582 | 0.94438298 |
| Thyroid | 0.04087014 | 0.94438298 |
| Uterus | 0.1448039 | 0.94438298 |
| Brain\_-\_Amygdala | -0.01585113 | 0.98823175 |
| Brain\_-\_Caudate\_(basal\_ganglia) | -0.01218631 | 0.98823175 |
| Ovary | -0.009673273 | 0.98823175 |
| Prostate | -0.01337519 | 0.98823175 |
| Skin\_-\_Not\_Sun\_Exposed\_(Suprapubic) | 0.006995845 | 0.98823175 |
| Heart\_-\_Atrial\_Appendage | 0.003893693 | 0.98833993 |
| Muscle\_-\_Skeletal | 0.000968298 | 0.98833993 |
| Whole\_Blood | 0.000691617 | 0.98833993 |



**Figure S1. L1 Subfamily mappability fingerprint.** Simulated transcripts from putative L1Hs, L1P1, L1PA2, L1PA3 and L1PA4 were aligned to the reference genome using the same parameters as the TeXP pipeline. The proportion of reads mapped to each subfamily was calculated. L1hg38 Ref displays the proportion of bases annotated as each of these subfamilies and indicates the expected signal from pervasive transcription.

L1Subfamilies_RPKM_ENCODE.pdf

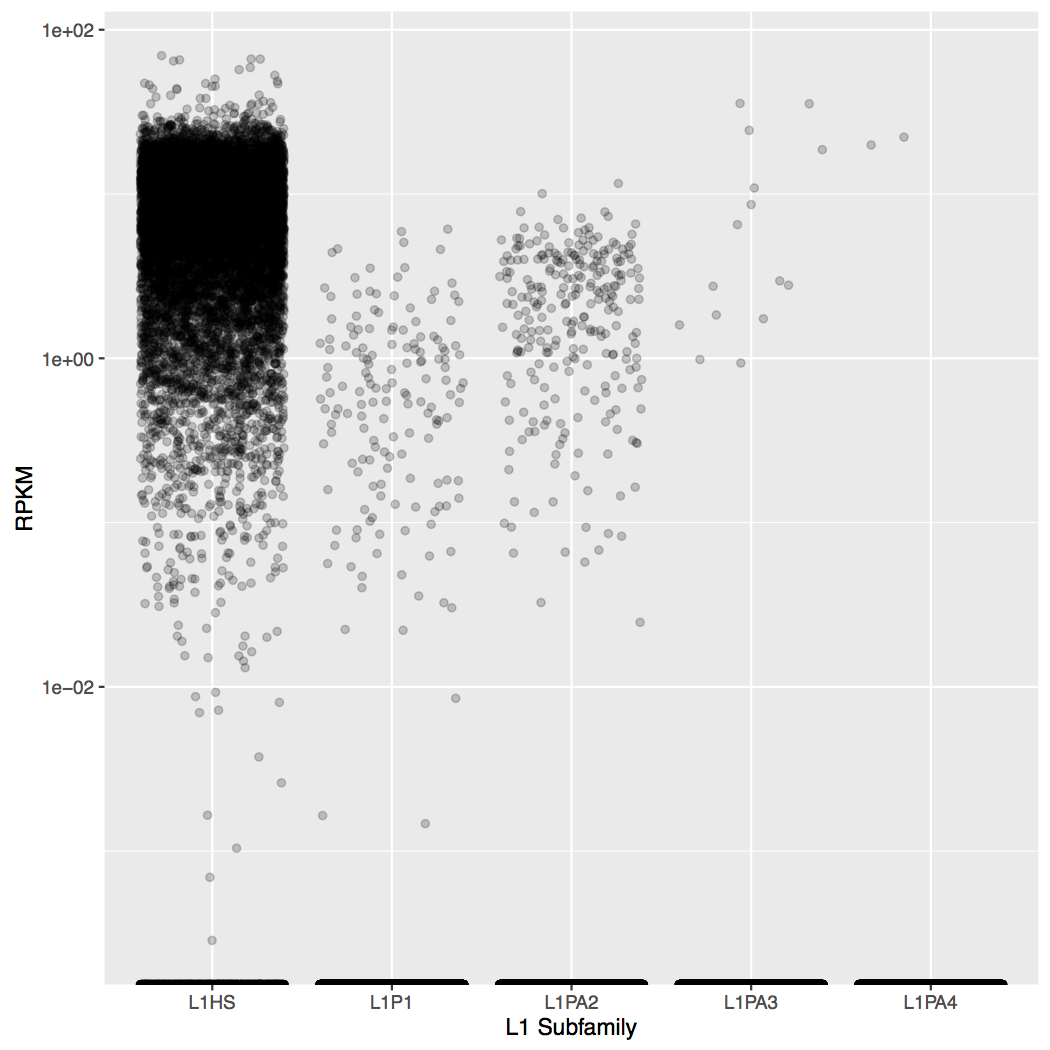
**Figure S2. L1 Subfamily RPKM in ENCODE RNA-seq samples.** Most samples have zero RPKM (more dense bar at the bottom). L1Hs has more samples with higher RPKM than any other subfamily followed by L1PA2 the second most recent L1 Subfamily.



**Figure S3. MCF7 L1 Subfamilies read count.** Every four bars are experiments respectively from whole cell polyA-; whole cell polyA+; cytoplasm polyA+ and nuclear polyA+.



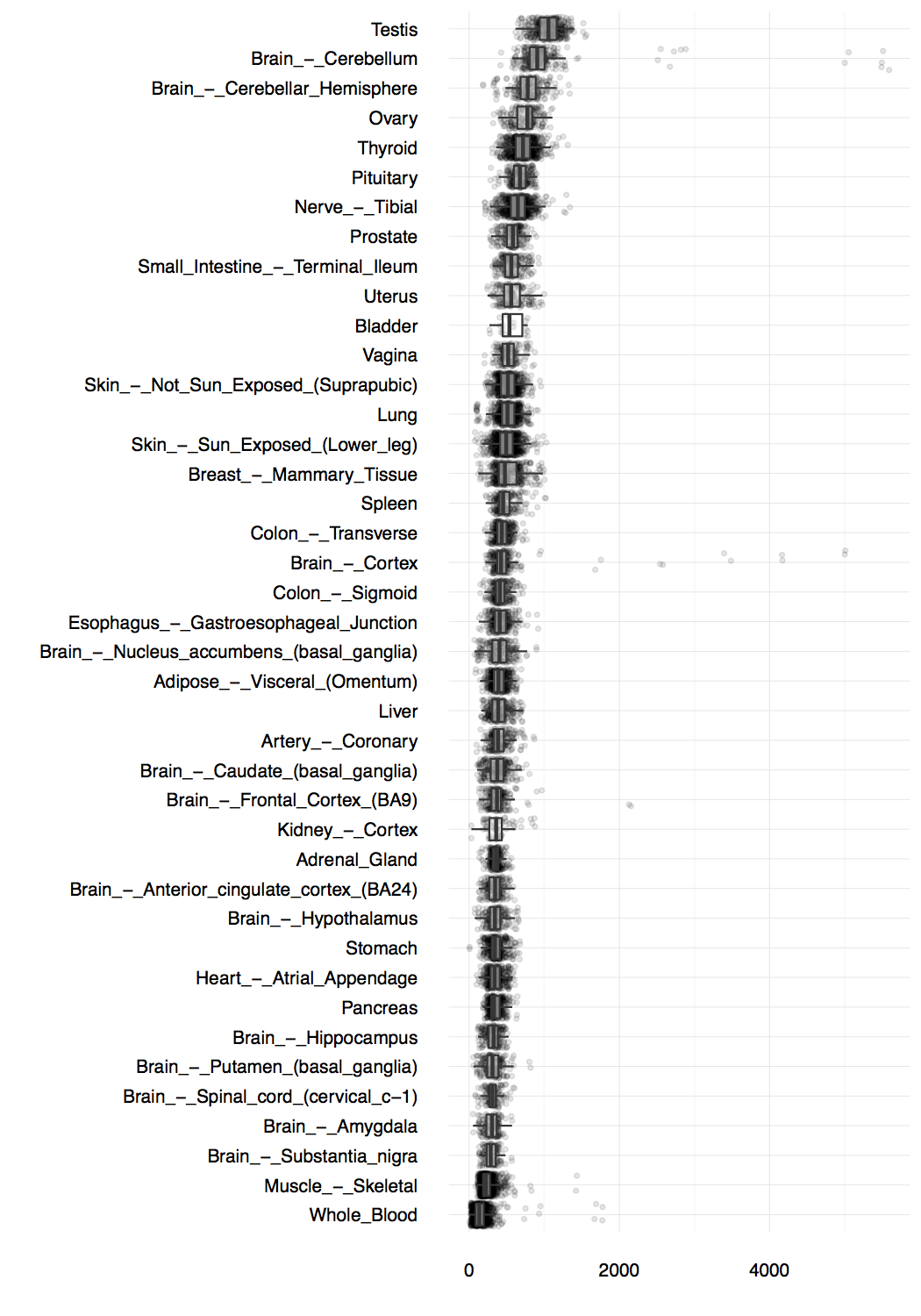
**Figure S4. MCF7 L1 Subfamilies read deconvolution.** Every four bars are experiments respectively from whole cell polyA-; whole cell polyA+; cytoplasm polyA+ and nuclear polyA+. Gray and Dark blue bars refer to pervasive transcription and L1Hs autonomous transcription signal respectively.

****

**Figure S5. L1 Subfamily RPKM in GTex samples.** Most samples have zero RPKM (more dense bar at the bottom). L1Hs has more samples with higher RPKM than any other subfamily followed by L1PA2 the second most recent L1 Subfamily.

GTEx_cell%20lines.pdf

**Figure S6. Estimation of L1Hs autonomous transcription in GTEx cell lines.** Most of EBV transformed cell lines have no autonomous transcription of L1Hs (bottom box). Transformed fibroblasts, derived from skin, have intermediate autonomous transcription of L1Hs (at lower levels than Skin samples). And K-562, derived from Leukemia tumor, has consistently high autonomous transcription of L1Hs across distinct batches.



**Figure S7. Pervasive transcription index per GTEx tissue.** Pervasive transcription index was estimated for each GTEx sample and ordered by pervasive transcription median.

../../../../../Desktop/l1hs-RPKM-cancer%20normal.pd

**Figure S8. L1Hs autonomous transcription in healthy and tumoral samples.**



**Figure S9. L1Hs transcript simulated as RNA sequencing reads.** Different aligners were used to assess the construction of mappability fingerprints in the human reference genom. One hundred independent simulations of L1Hs transcript reads were independently mapped to the reference genome using bowtie2 (red), bwa (green) and star (blue). The box plot represents the distribution of the number of reads mapped to each L1 subfamily in the reference genome.



**Figure S10. Read length effect on mappability of L1 subfamilies.** Simulating L1Hs reads with different length yield distinct proportions of reads mapped to each subfamily. As expected, the longer the read, the higher the proportion of reads correctly mapped originating subfamily.