**Supplementary Table 1.** **Number of samples from The Cancer Genome Atlas used in model development for tissue-of-origin prediction.**

|  |  |  |
| --- | --- | --- |
| **Cancer type** | **Abbreviation** | **Number of samples** |
| Adrenocortical carcinoma | ACC | 79 |
| Bladder urothelial carcinoma | BLCA | 416 |
| Breast invasive carcinoma | BRCA | 787 |
| Cervical and endocervical carcinoma | CESC | 308 |
| Cholangiocarcinoma | CHOL | 36 |
| Colorectal adenocarcinoma (COADREAD) | CRAD | 390 |
| Diffuse large B-cell lymphoma | DLBC | 48 |
| Esophageal squamous cell carcinoma | ESCC | 91 |
| Esophageal and gastric adenocarcinoma | ESST | 486 |
| Head and neck squamous cell carcinoma | HNSC | 525 |
| Kidney renal cell carcinoma (KICH/KIRC/KIRP) | KIPAN | 652 |
| Glioma (LGG + GBM) | LGGBM | 684 |
| Liver hepatocellular carcinoma | LIHC | 377 |
| Lung adenocarcinoma | LUAD | 461 |
| Lung squamous cell carcinoma | LUSC | 364 |
| Mesothelioma | MESO | 87 |
| Ovary serous cystadenocarcinoma | OV | 10 |
| Pancreatic adenocarcinoma | PAAD | 184 |
| Pheochromocytoma and paraganglioma | PCPG | 183 |
| Prostate adenocarcinoma | PRAD | 496 |
| Skin cutaneous melanoma | SKCM | 473 |
| Thyroid carcinoma | THCA | 511 |
| Thymoma | THYM | 124 |
| Uterine corpus endometrial carcinoma | UCEC | 422 |
| Uveal melanoma | UVM | 80 |
| **Total number of cases** |  | **8,274** |