**Table S1.** TCGA FFPE biosamples used for training and validation of XGBoost model for filtering RNAseq variants.

|  |  |  |  |
| --- | --- | --- | --- |
| **TCGA tissue code** | **Cancer type** | **Sequencing site code** | **Number of samples** |
| BLCA | Bladder Urothelial Carcinoma | BL | 3 |
| BRCA | Breast cancer | A7 | 8 |
| BRCA | Breast cancer | AC | 3 |
| BRCA | Breast cancer | PL | 4 |
| CESC | Cervical squamous cell carcinoma and endocervical adenocarcinoma | PN | 1 |
| COAD | Colon adenocarcinoma | A6 | 13 |
| KIRC | Kidney renal clear cell carcinoma | B2 | 2 |
| LUAD | Lung adenocarcinoma | 44 | 12 |
| PRAD | Prostate adenocarcinoma | HC | 3 |
| UCEC | Uterine Corpus Endometrial Carcinoma | BK | 4 |