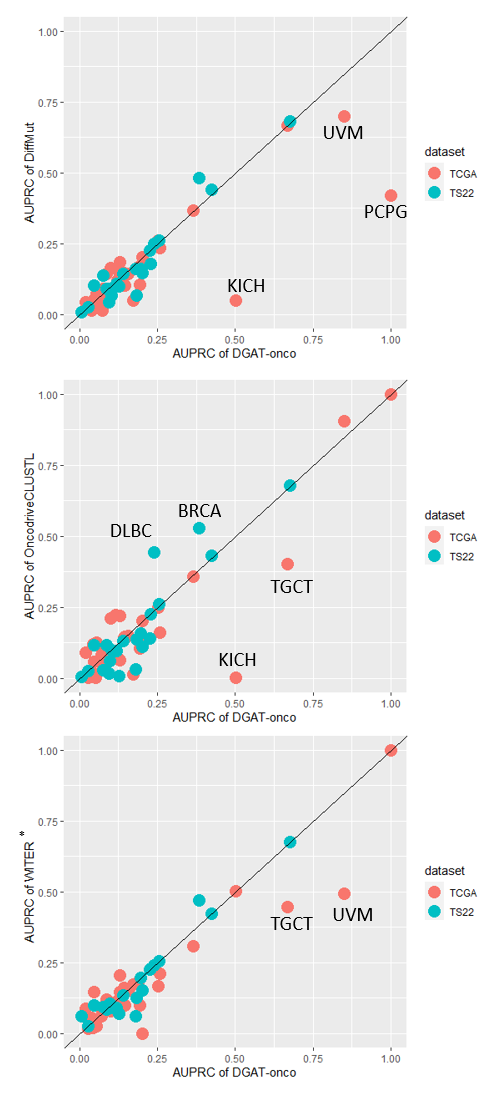
Table S1. Number of samples, genes and mutations of datasets

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Source** | **Population** | **Sample** | **Gene** | **Mutation** |
| TCGA | ACC | 92 | 4936 | 6859 |
| TCGA | BLCA | 406 | 17037 | 80874 |
| TCGA | BRCA | 965 | 16057 | 57090 |
| TCGA | CESC | 288 | 17489 | 69715 |
| TCGA | CHOL | 49 | 2567 | 2916 |
| TCGA | COAD | 398 | 18781 | 146513 |
| TCGA | DLBC | 37 | 3135 | 4038 |
| TCGA | ESCA | 182 | 11610 | 25372 |
| TCGA | GBM | 384 | 13236 | 35101 |
| TCGA | HNSC | 502 | 15469 | 57478 |
| TCGA | KICH | 66 | 1394 | 1508 |
| TCGA | KIRC | 332 | 7981 | 12887 |
| TCGA | KIRP | 280 | 9140 | 15754 |
| TCGA | LAML | 93 | 1034 | 1141 |
| TCGA | LGG | 498 | 10358 | 20765 |
| TCGA | LIHC | 361 | 13688 | 36051 |
| TCGA | LUAD | 550 | 17012 | 108220 |
| TCGA | LUSC | 484 | 17299 | 107968 |
| TCGA | MESO | 78 | 1946 | 2210 |
| TCGA | OV | 431 | 10902 | 22602 |
| TCGA | PAAD | 135 | 6804 | 10439 |
| TCGA | PCPG | 174 | 1091 | 1199 |
| TCGA | PRAD | 452 | 9102 | 16464 |
| TCGA | READ | 135 | 13850 | 40873 |
| TCGA | SARC | 235 | 9234 | 16620 |
| TCGA | SKCM | 463 | 18734 | 272761 |
| TCGA | STAD | 422 | 17635 | 105773 |
| TCGA | TGCT | 138 | 1332 | 1473 |
| TCGA | THCA | 473 | 2782 | 3652 |
| TCGA | THYM | 81 | 1684 | 1881 |
| TCGA | UCEC | 528 | 20913 | 606509 |
| TCGA | UCS | 57 | 5238 | 7220 |
| TCGA | UVM | 80 | 1000 | 1168 |
| Kim et al. Eur Urol 2015 | BLCA | 95 | 230 | 1072 |
| Pereira et al. Nat Commun 2016 | BRCA | 932 | 173 | 6721 |
| Lowery et al. Clin Cancer Res 2018 | CHOL | 53 | 96 | 187 |
| Vasaikar et al. Cell 2019 | COAD | 101 | 13638 | 40272 |
| Reddy et al. Cell 2017 | DLBC | 883 | 80 | 3604 |
| Dulak et al. Nat Genet 2013 | ESCA | 146 | 8717 | 16296 |
| Stransky et al. Science 2011 | HNSC | 39 | 3441 | 4527 |
| Sato et al. Nat Genet 2013 | KIRC | 72 | 243 | 360 |
| Tyner et al. Nature 2018 | LAML | 608 | 3335 | 9910 |
| Johnson et al. Science 2014 | LGG | 51 | 6006 | 9482 |
| Schulze et al. Nat Genet 2013 | LIHC | 240 | 7825 | 14770 |
| Imielinksi et al. Cell 2012 | LUAD | 162 | 11600 | 32864 |
| Guo et al. Cancer Res 2015 | MESO | 20 | 91 | 92 |
| Bailey et al. Nature 2016 | PAAD | 274 | 7060 | 11863 |
| Armenia et al. Nat Genet 2018 | PRAD | 657 | 13878 | 49389 |
| Tirode et al. Cancer Discov 2014 | SARC | 110 | 1004 | 1079 |
| Krauthammer et al. Nat Genet 2012 | SKCM | 91 | 5391 | 15277 |
| Wang et al. Nat Genet 2014 | STAD | 67 | 6327 | 10463 |
| Landa et al. J Clin Invest 2016 | THCA | 94 | 152 | 406 |
| Soumerai et al. Clin Cancer Res. 2018 | UCEC | 22 | 189 | 401 |
| Jones et al. Nat Commun 2014 | UCS | 15 | 5245 | 8357 |
| Johansson et al. Oncotarget 2016 | UVM | 7 | 59 | 65 |
| 1000G | Natural | 2504 | 18470 | 81271745 |

Fig. S1. Pair-wise comparison between DGAT-onco and 3 methods.



\* WITER was unable to generate oncogene predictions in some cancers with limited mutations (8 cancers in TS22 [BLCA, CHOL, DLBC, LAML, STAD, THCA, UCEC, and UVM]. We imputed these missing AUPRC with that from DGAT-onco.

The x-axis represents the AUPRC of DGAT-onco while the y-axis represents the AUPRC of 3 methods. The colour represents the dataset (i.e., TCGA or TS22).